UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

División de Estudios de Posgrado
Facultad de Ingeniería

Condiciones y Demostraciones Equivalentes del Principio de Pontryagin

Tesis de Maestría

Que para obtener el grado de:
Maestro en Investigación de Operaciones

p r e s e n t a :

RAMON FELIPE LOPEZ GUAL

Director de Tesis: Dr. Sergio Fuentes Maya

México D. F. 1980

I N D I C E

I N D I C E (Continuación)

1. INTRODUCCION

Durante los últimos 20 años, uno de los principales objetivos a los que se han enfocado las matemáticas y la ingeniería es el análisis de problemas de toma de decisiones en sistemas físicos y organizacionales. Esta tendencia fue inspirada originalmente por los beneficios económicos resultantes del buen uso de los recursos escasos y por la demostración de que el análisis de este tipo de problemas podría hacerse en forma matemática para incrementar dichos beneficios con base en mejores decisiones.

La aparición de las computadoras digitales contribuyó, también, al desarrollo de este análisis, ya que permitió la implantación e integración de sistemas de gran escala, revolucionó las matemáticas aplicadas y facilitó la resolución de problemas complejos.

El concepto de decisión óptima aparece como el enfoque fundamental en la formulación de problemas de decisión. Con este enfoque una cantidad que resume el comportamiento de la decisión, es aislado y optimizado, mediante la selección propia entre las alternativas disponibles. La decisión óptima es tomada como la solución al problema de decisión. Es obvia la limitación que se tiene, debida a la necesidad de seleccionar un objetivo único con el cual se midan los resultados; sin embargo, la optimización ha demostrado su utilidad como una herramienta para el análisis y está firmemente cimentada en el campo de la toma de decisiones.

En el desarrollo actual de la optimización en los problemas de decisión, las técnicas clásicas han sido reexaminadas, extendidas, a veces redescubiertas y aplicadas a problemas diferentes de los que las motivó. Se ha profundizado en la teoría y nuevas técnicas han sido desarrolladas; la computación ha desechado muchas técnicas obsoletas y ha hecho a otras que anteriormente eran imprácticas, factibles y eficientes. Hoy en día, el estudio de la optimización como un tópico independiente, es visto como una rama de las matemáticas aplicadas.

Hay una gran variedad de problemas que pueden ser formulados como de optimización. Una clase importante de estos, son aquéllos en los que se involucra a un sistema dinámico. Los problemas de control pertenecen a esta clase.

Existe otra teoría, el control óptimo, desarrollada paralelamente a la optimización, que establece los fundamentos matemáticos para resolver específicamente problemas de control.

Un problema de control puede caracterizarse como sigue: Un sistema a controlar, un objetivo deseado, un conjunto de controles y una medida de costos o efectividad de las acciones de control.

Ambas teorías, la optimización y el control óptimo, tienen resultados equivalentes pero obtenidos partiendo de bases diferentes.

El propósito de esta tésis es presentar las condiciones necesarias para los problemas de control, en los cuales el control está restringido a un conjunto dado, resultado conocido como el principio del máximo de Pontryagin, y llegar a establecer equivalentemente dichas condiciones partiendo de las dos teorías mencionadas: la optimización y el control óptimo.

Para ilustrar la variedad de problemas que pueden formularse como de optimización ó control y que pueden ser resueltos aplicando el principio máximo de Pontryagin, se presentan algunos ejemplos a continuación:

1) Problema del cohete. Considere el problema de seleccionar un programa de lanzamiento u(t) para el ascenso vertical de un cohete, el cual está únicamente sujeto a las fuerzas de gravedad y de lanzamiento y el objetivo es alcanzar una altura determinada con una utilización mínima de combustible.

   Suponiendo masa unitaria, gravedad unitaria fijas y condiciones iniciales cero, la altura x(t) es gobernada por la ecuación diferencial:

$$\ddot{x}(t) = u(t) - 1 \qquad\qquad \dot{x}(0) = x(0) = 0$$

   La altura a alcanzar es $x(T) = 1$, minimizando el consumo de combustible:

$$\int_0^T |u(t)|\, dt$$

   El tiempo final T no está especificado, pero el enfoque del problema es resolverlo para cada T fija y entonces minimizar con respecto a T.

2) Problema del pozo petrolero. Una compañía petrolera tiene localizado un yacimiento petrolífero, el cual tiene una capacidad α . Se desea determinar el programa de extracción a largo plazo tal que se maximice el beneficio total descontado.

El problema consiste en encontrar la función integrable x sobre $[0,\infty)$ que representa la tasa de extracción tal que se maximice

$$\int_0^\infty F\,[x(t)]\,v(t)\,dt$$

sujeto a

$$\int_0^\infty x(t)\,dt \leq \alpha \qquad x(t) \geq 0$$

$F\,[x]$ representa la tasa de beneficio asociado con la tasa de extracción x, esta función puede considerarse como estrictamente cóncava y creciente sobre $[0,\infty)$ y $F\,[0] = 0$. La función v(t) representa el factor de descuento y puede suponerse que es contínua, positiva, estrictamente decreciente hacia cero e integrable sobre $[0,\infty)$.

3) Problema de asignación del granjero. Un granjero produce un sólo cultivo, digamos trigo. Después de la cosecha debe almacenarlo o venderlo y reinvertir comprando tierra adicional y equipo para incrementar su tasa de producción. El granjero desea maximizar la cantidad total almacenada al tiempo T.

Sea: $x_1(t)$ = la tasa de producción

$x_2(t)$ = la tasa de reinversión

$x_3(t)$ = la tasa de almacenamiento

supongamos:

$$\dot{x}_1(t) = x_2(t) \quad , \quad x_1(0) = x_0 > 0$$

$$x_2(t) + x_3(t) = x_1(t)$$

$$x_1(t) \geq 0 \, , \; x_2(t) \geq 0 \, , \; x_3(t) \geq 0$$

El granjero desea operar de tal forma de maximizar:

$$\int_0^T x_3(t)\, dt$$

4) Planeación de la producción. Una empresa que produce cierto producto desea planear su programa de producción sobre un cierto período de una manera óptima. Se supone que una función de demanda sobre el período es conocida y que esta demanda debe ser satisfecha. El exceso de inventario debe ser almacenado y el costo de almacenamiento es proporcional a la cantidad almacenada. Existe un costo de producción asociado con cada tasa de producción. Sea $x(t)$ la existencia disponible al tiempo t, $r(t)$ la tasa de producción al tiempo t y $d(t)$ la demanda al tiempo t, el sistema de producción puede ser descrito por:

$$\dot{x}(t) = r(t) - d(t) \quad , \quad \text{dado } x(0)$$

y se desea encontrar la función r tal que satisfaga:

$$\left.\begin{array}{l} r(t) \geq 0 \\ x(0) + \int_0^t \left[r(\zeta) - d(\zeta)\right] d\zeta = x(t) \geq 0 \end{array}\right\} \text{para } 0 \leq t \leq T$$

y que minimice el costo:

$$J = \int_0^T \{c\left[r(t)\right] + h.x(t)\}dt$$

donde $c\left[r\right]$ es la tasa del costo de producción para el nivel de producción r y h.x es la tasa de costo de inventario para el nivel de inventario x.

5) La implantación de un algoritmo para la optimización de problemas de control no lineales. Este algoritmo consiste en la linealización de la solución del problema no lineal de control no óptimo, mediante iteraciones sucesivas basadas en la aplicación de técnicas de optimización lineal, de manera que se obtenga un sistema óptimo no lineal. La solución lineal es obtenida mediante la aplicación del principio máximo de Pontryagin. (Tesis de Maestría. Sohler, Jerome Francis. California University. Los Angeles. Dept. of Engineering).

Este trabajo está desarrollado de la siguiete forma: en el capítulo 2, se presentan los fundamentos matemáticos y los conceptos básicos de Optimización, necesarios para el desarrollo de las demostraciones del Principio máximo de Pontryagin.

En el capítulo 3, se desarrolla la demostración del principio con las bases teóricas de la optimización. Y en el capítulo 4 se hace la demostración correspondiente a la teoría de control óptimo.

Deseo expresar mi agradecimiento al Dr. Sergio Fuentes Maya por el apoyo y enseñanzas recibidas y por la supervisión y dirección de este trabajo. Ya la Sra. Cristina Arias Gónzález por la excelente labor mecanográfica que realizó.

2. FUNDAMENTOS MATEMATICOS Y CONCEPTOS BASICOS DE OPTIMIZACION

2.1 Introducción

En el presente capítulo se establecen los fundamentos matemáticos y los conceptos básicos de optimización necesarios para el desarrollo de las demostraciones del principio máximo de Pontryagin. El desarrollo del mismo es como sigue: En las dos primeras secciones de este capítulo, 2.2 y 2.3 se introducen los conceptos de espacio vectorial y algunas de sus propiedades; se definen los subespacios, variedad lineal y convexidad. En las secciones subsiguientes, 2.4 a 2.10 se discute las propiedades básicas de espacios lineales normados. Se establece el concepto de norma y se definen propiedades analíticas y topológicas como convergencia y conjuntos abiertos y cerrados. En las secciones 2.11, 2.12 y 2.13, se introducen los conceptos de funcional lineal, espacios duales y el Teorema de Hahn-Banach. Estos conceptos son básicos para la teoría de optimización, de hecho el teorema de Hahn-Banach es considerado uno de los resultados más importantes para dicha teoría. En las secciones 2.14 y 2.15 se estudia los operadores lineales y adjuntos, los cuales son esenciales para un enfoque más profundo de problemas de optimización y permiten la mejor aplicación de los principios de ésta en situaciones complicadas. En las últimas secciones, 2.16, 2.17 y 2.18 se presentan las generalizaciones de los conceptos de diferenciales, derivadas e integrales.

2.2 Espacios vectoriales

Asociado con cada espacio vectorial, hay un conjunto de escalares que es usado para definir la multiplicación por escalares sobre el espacio. Durante el desarrollo de este trabajo tomaremos este conjunto de escalares como el conjunto de los números reales, denotado R.

Definición: Un espacio vectorial X (o bien, espacio vectorial lineal X) es un conjunto de elementos llamados vectores, junto con dos operaciones. La primera es la suma la cual asocia con cualesquiera par de vectores x, y de X un vector x+y de X, llamado la suma de "x" y "y". La segunda multiplicación por un escalar la cual asocia con cualquier vector x de X y un escalar α, un vector αx; llamado la multiplicación escalar de x por α. El conjunto X y las operaciones de adición y multiplicación escalar satisfacen además, los siguientes axiomas:

1. $x + y = y + x$ (ley conmutativa)
2. $(x + y) + z = x + (y+z)$ (ley asociativa)
3. Existe un vector nulo θ en X tal que $x+\theta=x$, para toda x en X
4. $\alpha(x+t) = \alpha x+\alpha y$ (leyes distributivas)
5. $(\alpha+\beta)x = \alpha x+\beta x$
6. $(\alpha\beta)x = \alpha(\beta x)$ (ley asociativa)
7. $0x = \theta$; $1x = x$

Algunos ejemplos de espacios vectoriales son:

Ejemplo 1. El conjunto de los números reales. Con la operación suma definida de la manera usual y la multiplicación escalar como la multiplicación ordinaria. El vector nulo es el número real cero. Las propiedades de suma ordinaria y multiplicación de números reales, satisfacen los axiomas de la definición de espacios vectoriales.

Ejemplo 2. Una extensión del ejemplo anterior es el espacio coordenado real de dimensión n. Los vectores en este espacio consisten de arreglos de n números reales, tal que un vector típico tendrá la forma ; $x = (\S_1, \S_2, \ldots, \S_n)$.

El número real $\S_k$ será referenciado como el k-ésimo componente del vector. Dos vectores serán iguales si sus correspondientes componentes son iguales. El vector nulo está definido como;

$$\theta = (0,0,\ldots,0). \text{ Si } x = (\S_1,\S_2,\ldots,\S_n)$$

y $y = (\eta_1,\eta_2,\ldots,\eta_n)$, el vector x+y estará definido como el arreglo de n números reales donde el k-ésimo elemento será igual a $\S_k + \eta_k$. El vector αx, donde α es un escalar (real), será un arreglo de n números reales donde el k-ésimo elemento será igual a $\alpha\S_k$. Los axiomas de la definición se verifican comprobando la igualdad entre componentes. Este espacio coordenado real n-dimensional es denotado por R^n.

Ejemplo 3. La colección de todas las funciones continuas definidas sobre el intervalo real $[a,b]$, denotado $c[a,b]$, forman un espacio vectorial. Diremos que $x = y$, si $x(t) = y(t)$ para t en $[a,b]$. El vector nulo será la función que es igual a cero sobre $[a,b]$, ó sea $f(t) = 0$ para t en $[a,b]$.

Definiremos la suma $(x+y)(t) = x(t) + y(t)$ y el producto escalar $(\alpha x)(t) = \alpha x(t)$.

Este espacio es conocido como el espacio vectorial de las funciones continuas sobre $[a,b]$.

Definición. Sean X y Y espacios vectoriales, definidos sobre el mismo campo de escalares. Entonces el producto cartesiano de X y Y, denotado X x Y consiste de la colección de pares ordenados (x,y) con "x" en X, "y" en Y. La adición y producto escalar estarán definidos sobre: X x Y como:

$$(x_1,y_1,) + (x_2,y_2) = (x_1+x_2,y_1+y_2) \text{ y } \alpha(x,y) = (\alpha x,\alpha y)$$

Obviamente, este espacio, producto cartesiano, es consistente con los axiomas de espacio vectorial y es en sí un espacio vectorial. Esta definición se puede generalizar fácilmente para el producto de n espacios vectoriales.

Definición. Un subconjunto no vacío M de un espacio vectorial X es llamado un subespacio de X si cada vector de la forma $\alpha x + \beta y$ está en M siempre que x,y estén en M. α,β son escalares.

Ya que al menos debe de tener un elemento, y por definición debe contener a $0x = \theta$, entonces cada subespacio contiene al vector nulo.

Un subespacio de un espacio vectorial es en sí mismo un espacio vectorial.

<u>Definición</u>; La traslación de un subespacio se le denomina - como <u>variedad lineal</u> (ó subespacio afín). Una variedad lineal V puede ser escrita como $V = x_0 + M$, donde M es un subespacio. En esta representación el subespacio M es único pero cualquier vector en V puede servir como x_0

<u>Definición</u>: Un conjunto K en un espacio vectorial, se dice que es <u>convexo</u> si dados x_1, x_2 en K, todos los puntos de la forma $\alpha x_1 + (1-\alpha)x_2$ con $0 \leq \alpha \leq 1$, están en K.

2.3 Espacios lineales normados

Los axiomas de los espacios vectoriales, solamente describen propiedades algebráicas de los elementos del espacio: suma, multiplicación escalar y combinación de ellos. Sin embargo, lo que no se toma en cuenta son conceptos topológicos como apertura, cerradura, convergencia y completez. Estos se pueden incluir mediante la introducción del concepto de medida de la distancia en un espacio vectorial.

Definición: Un espacio vectorial normado (o bien espacio vectorial lineal normado) es un espacio vectorial X sobre el cual está definida una función que mapea a cada elemento x en X a un número real $||x||$, llamado la norma de x.

La norma satisface los siguientes axiomas:

1. $||x|| \geq 0$ para toda x en X, $||x|| = 0$ sí y sólo si $x = \theta$
2. $||x+y|| \leq ||x|| + ||y||$ para toda x, y en X
3. $||\alpha x|| = |\alpha| \cdot ||x||$ para cualquier escalar α y cualquier x en X.

Como se ve, la norma es una abstracción de nuestro concepto usual de medida de distancia.

Ejemplos:

1) El espacio vectorial de los números reales con la norma definida como el valor absoluto $|.|$, que cumple los axiomas de norma, es entonces un espacio normado.

2) El espacio R^n, donde $x = (\xi_1,\xi_2,\ldots,\xi_n)$ tiene como norma: $||x|| = (\sum_{i=1}^{n} |\xi_i|^2)^{1/2}$, donde esta norma cumple los axiomas correspondientes, es por tanto, también un espacio normado.

3) Otro espacio normado es $C[a,b]$, que consiste de todas las funciones continuas sobre el intervalo real $[a,b]$ y tiene como norma de x : $||x|| = \max\{|x(t)| \; ; \; a \leq t \leq b\}$

2,4 Conjuntos abiertos y cerrados

Definición. Sea P un subconjunto de un espacio normado X. El punto $p\varepsilon P$ se dice que es un punto interior de P si existe una $\varepsilon>0$ tal que todos los vectores x, que satisfacen $||x-p||<\varepsilon$ son también elementos de P. La colección de todos los puntos interiores de P se llama Interior de P y se denota $\mathring{P}$.

Introduciremos la notación $S(x,\varepsilon)$ para la esfera (o también: vecindad o bola) abierta centrada en x y radio ε; esto es:

$$S(x,\varepsilon) = \{y: ||x-y||<\varepsilon\}$$

Definición: Un conjunto P se dice que es abierto si $P = \mathring{P}$

Definición: Un punto $x\varepsilon X$ se dice que es un punto de cerradura de un conjunto P si dado $\varepsilon>0$, existe un punto p en P que satisface $||x-p||<\varepsilon$. La colección de todos los puntos de cerradura de P es llamada la cerradura de P y se denota $\overline{P}$.

Definición: Un conjunto P se dice que es cerrado si $P = \overline{P}$.
Un conjunto es cerrado si su complemento es abierto.

Las siguientes dos proposiciones, son complementarias y sus pruebas se pueden encontrar en cualquier libro de Análisis Matemático.

Proposición 1. La intersección de un número finito de conjuntos abiertos es abierta. La unión de una colección arbitraria de conjuntos abiertos es abierta.

Proposición 2. La unión de un número finito de conjuntos cerrados es cerrada. La intersección de una colección arbitraria de conjuntos cerrados es cerrada.

2.5 Convergencia

Definición: En un espacio lineal normado una sucesión infinita de vectores $\{x_n\}$ se dice que converge al vector x, si la sucesión de números reales $\{||x-x_n||\}$ converge a cero. En este caso, se escribe $x_n \longrightarrow n$.

Otra manera de definir convergencia es en términos de esferas.

Definición: Una sucesión $\{x_n\}$ converge a x si y sólo si dado $\varepsilon>0$; la esfera $S(x,\varepsilon)$ contiene a x_n para toda n mayor que algún número N.

Dos resultados directos de esta definición son las siguientes proposiciones:

Proposición 1. Si $x_n \longrightarrow x$, se sigue que $||x_n|| \longrightarrow ||x||$.

Proposición 2. Si una sucesión converge, su límite es único.

Ahora con la noción de convergencia definida, podremos demostrar la siguiente proposición:

Proposición 3. Un conjunto F es cerrado si y sólo si toda sucesión convergente con elementos en F tiene su límite en F.

Prueba: El límite de una sucesión de F es un punto de cerradura de F y por tanto debe estar contenido en F, ya que F es cerrada.

Supongamos ahora que F no es cerrado. Entonces hay un punto de cerradura x que no está en F.

Sea x_n en F tal que $||x_n - x|| < 1/n$, esta x_n existe ya que x es un punto de cerradura de F. La sucesión $\{x_n\} \longrightarrow x$ y x no está en F, lo cual es una contradicción.

2.6 Transformaciones y Continuidad

Definición: Sean X y Y espacios vectoriales y D un subconjunto de X. Una regla que asocia a cada elemento x en D un elemento y en Y se dice que es una transformación de X a Y con dominio D. Si y corresponde a x bajo T, escribiremos y = T(x).

Si para cada y en Y existe a lo más una x en D para el cual T(x) = y, la transformación T se dice que es uno a uno.

Si para cada y en Y existe al menos una x en D para la cual T(x) = y, T se dirá que es sobre.

<u>Definición</u>: Una transformación del espacio vectorial X al espacio de los Reales se dice que es un <u>funcional</u> sobre X.

<u>Definición</u>: Una transformación T que mapea un vector del espacio X a un vector del espacio Y se dice que es lineal si para toda x_1, x_2 en X, y para cualesquiera escalares α_1, α_2 tenemos:

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$$

Al definir transformaciones, se considera que los mapeos son de un espacio abstracto a otro. Ahora si tomamos espacios en la definición, que sean normados, entonces será posible definir el concepto de continuidad.

<u>Definición</u>: Una transformación que mapea un espacio normado X a un espacio normado Y es <u>continua</u> en x_0 en X, si para cada $\varepsilon>0$, existe $\delta>0$ tal que:

$$||x-x_0||<\delta \quad \text{implica que} \quad ||T(x)-T(x_0)||<\varepsilon$$

Hay que notar que la <u>continuidad</u> depende de que ambos espacios X y Y sean normados.

Si T es continua en cada punto x_0 de X, diremos que T es continua en todo X o más simplemente que T es continua.

Una caracterización de continuidad, que será muy útil en la demostración de continuidad de transformaciones es la siguiente:

Proposición 1: Una transformación que mapea un espacio normado X a otro espacio normado Y, es continua en el punto x_0 si y solo si $x_n \longrightarrow x_0$ implica $T(x_n) \longrightarrow T(x)$.

2.7 Espacios de Banach

Definición: Una sucesión $\{x_n\}$ en un espacio normado se dice que es una sucesión de Cauchy si $||x_n - x_m|| \longrightarrow 0$ cuando $n,m \longrightarrow \infty$; i.e. dado $\varepsilon>0$, existe un entero N, tal que $||x_n - x_m||<\varepsilon$ para toda $n,m>N$

En un espacio normado, cualquier sucesión convergente es sucesión de Cauchy, ya que si $x_n \longrightarrow x$, entonces:

$$||x_n - x_m|| = ||x_n - x + x - x_m|| \leq ||x_n - x|| + ||x - x_m|| \longrightarrow 0$$

Sin embargo, en general, una sucesión de Cauchy puede no ser convergente.

Los espacios normados en los cuales, toda sucesión de Cauchy es convergente son de particular interés. En esos espacios es posible identificar a las sucesiones convergentes sin identificar explícitamente a sus límites.

Definición: Un espacio vectorial normado X es completo o de Banach si toda sucesión de Cauchy en X, tiene su límite en X.

Lema 1. Una sucesión de Cauchy es acotada.

Prueba: Sea $\{x_n\}$ una sucesión de Cauchy y sea N un entero tal que $||x_n - x_N||<1$ para $n>N$.

Para n>N, tenemos:

$$||x_n|| = ||x_n - x_N + x_N|| \leq ||x_n - x_N|| + ||x_N|| < ||x_N|| + 1$$

Teorema: En un espacio de Banach un subconjunto es completo si y solo si es cerrado.

Prueba: Si el subconjunto es completo, es obviamente cerrado ya que cada sucesión convergente tiene su límite en el subconjunto. Por otro lado, cada sucesión de Cauchy del subconjunto tiene su límite en el espacio (ya que este es Banach). Y por cerradura, debe estar en el subconjunto.

Definición: Un conjunto K en un espacio normado X se dice que es compacto si para toda sucesión arbitraria $\{x_i\}$ en K, existe una subsucesión $\{x_{i_n}\}$ que converge a un elemento x de K.

En dimensiones finitas, compacto equivale a cerrado y acotado, pero en general no es verdad cuando se habla de espacios normados. Sin embargo, todo conjunto compacto K debe ser completo ya que cualquier sucesión de Cauchy de K debe tener su límite en K.

2.8 Espacios Cociente

Definición: Sea M un subespacio de un espacio vectorial X. Dos elementos x_1, x_2 en X se dice que son equivalentes modulo M si $x_1 - x_2$ está en M. En tal caso escribiremos $x_1 \equiv x_2$.

Puede ser comprobado que esta es una relación de equivalencia y como tal, particiona el espacio X en conjuntos disjuntos ó clases. Estas clases son llamadas clases de equivalencia (cosets en inglés). Dada una x en X, esta pertenece a una única clase de equivalencia de M, que denotaremos $[x]$, y que consiste de:

$$[x] = \{y : y \equiv x\}$$

Definición: Sea M un subespacio de un espacio vectorial X. El espacio cociente X/M consiste de todas las clases de equivalencia de M; con la suma y producto por un escalar definidos por:

$$[x_1] + [x_2] = [x_1+x_2] \quad \text{y} \quad \alpha[x_1] = [\alpha x_1] \text{ , respectivamente.}$$

Definición: Definimos la norma de una clase de equivalencia en X/M como:

$$\|\ [x]\ \| = \inf\ \{\|x+m\| \ ; \quad m \quad \text{en} \quad M\}$$

i.e. $\|\ [x]\ \|$ es el ínfimo de las normas de todos los elementos en la clase de equivalencia $[x]$.

Proposición 1. Sea X un espacio de Banach, M un subespacio cerrado de X, y X/M el espacio cociente con la norma definida como la anterior. Entonces X/M es también un espacio de Banach.

2.9 Densidad y Separabilidad

Definición: Un conjunto D se dice que es denso en un espacio normado X si para cada elemento x en X y cada $\varepsilon>0$ existe d en

D tal que $||x-d||<\varepsilon$.

Si D es denso en X, existen puntos de D arbitrariamente cercanos a cada x en X. Entonces dado x, una sucesión puede ser construida de D, que sea convergente a x.

Una definición equivalente a la anterior es la siguiente:

Definición: D es denso en X si $\bar{D}$, la cerradura de D, es X.

Definición: Un conjunto E se dice que es denso en ninguna parte en un espacio normado X, si $\bar{E}$ no contiene ningún conjunto abierto.

Definición: Un espacio normado es separable si contiene un conjunto denso que sea contable.

2.10 Isomorfismo e Isometría

Definición: Dos espacios lineales normados, X y Y se dice que son isomórficos si existe un mapeo T, uno a uno, de X a Y, tal que $T(\alpha_1 x_1 + \alpha_2 \alpha_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$

Definición: Dos espacios normados, X y Y, se dice que son isometricamente isomórficos si son isomórficos y si el mapeo uno a uno, T, satisface:

$$||T(x)|| = ||x||$$

2.11 Funcionales lineales

Definición: Un funcional en un espacio vectorial X es lineal si para cualesquiera dos vectores x, y en X y cualesquiera dos escalares α y β, se sostiene que:

$$f(\alpha x+\beta y) = \alpha f(x)+\beta f(y)$$

Proposición 1. Si un funcional lineal en un espacio normado X es continuo en un punto, entonces es continuo en todo el espacio X.

Prueba. Sea f funcional lineal y continuo en x_0 en X. Sea $\{x_n\}$ una sucesión de X que converge a x en X. Entonces por la linealidad de f:

$$|f(x_n)-f(x)| = |f(x_n)+f(x_0)-f(x_0)-f(x)| = |f(x_n-x+x_0)-f(x_0)|$$

como: $x_n \longrightarrow x$

entonces: $x_n-x \longrightarrow 0$

por lo tanto: $x_n-x+x_0 \longrightarrow x_0$

y como: f es continua en x_0

tenemos: $f(x_n-x+x_0) \longrightarrow f(x_0)$

o sea: $f(x_n-x+x_0) - f(x_0) \longrightarrow 0$

lo que implica: $|f(x_n)-f(x)| = |f(x_n-x+x_0)-f(x_0)| \longrightarrow 0$

lo que demuestra que f es continua para cualquier x en X. O sea f es continua en X.

Definición: Un funcional f en un espacio normado es acotado si existe una constante M tal que $|f(x)| \leq M \; ||x||$ para toda x en X.

La constante M que es la más pequeña es llamada la norma del funcional f, y formalmente la definiremos:

Definición: La norma de el funcional f denotada $||f||$ esta dada por:

$$||f|| = \inf \{M : |f(x)| \leq M \; ||x||, \text{ para toda } x \text{ en } X\}$$

de otra manera:

$$||f|| = \sup \{|f(x)| \;/\; ||x||;\; x \varepsilon X\}$$

Proposición 2: Un funcional lineal en un espacio normado es acotado si y sólo si es continuo.

Prueba: Supongamos primero que el funcional f es acotado. Sea M tal que $|f(x)| \leq M \; ||x||$ para toda x en X. Entonces, si $x_n \quad \theta$, tenemos $|f(x_n)| \leq M \; ||x_n|| \longrightarrow 0$. Entonces, f es continuo en θ. Por lo tanto es continuo en todo el espacio.

Supongamos ahora que f es continua en θ. Entonces existe una $\delta > 0$ tal que $|f(x)| < 1$ para $||x|| \leq \delta$. Ya que para toda x en X diferente de cero, $\delta x/||x||$ tiene su norma igual a δ, tenemos:

$$|f(x)| = |f((\delta x/||x||)\;(||x||/d))| = |f(\delta x/||x||)\;|\;||x||/\delta|\;|$$

$$= |f(\delta x/||x||)| \;.\; ||x||/\delta$$

y como: $||\delta x/||x|| \; || \leq \delta$ entonces $|f(\delta x/||x||)| < 1$

por lo tanto: $|f(x)| < ||x||/\,\delta$

Haciendo $M = 1/\varepsilon$, vemos que es una cota para f. Los funcionales lineales en un espacio vectorial pueden ser vistos como elementos de un espacio vectorial; definiendo la adición y multiplicación por un escalar como sigue: Dados dos funcionales f_1 y f_2 en un espacio vectorial X, definimos su suma $f_1 + f_2$ como el funcional en X dado por $(f_1+f_2)(x) = f_1(x) + f_2(x)$ para toda x en X.

Similarmente dado un funcional lineal f, definimos αf por $\alpha f(x) = \alpha[f(x)]$. El elemento nulo de este espacio de funcionales lineales es el funcional que es idéntico a cero sobre X.

Definición: Sea X un espacio vectorial normado. El espacio de todos los <u>funcionales lineales acotados</u> (i.e. continuos) de X es llamado el <u>dual</u> de X, y es denotado por X^*.

La norma de un elemento f en X^* es:

$$\begin{aligned} \|f\| &= \sup \{|f(x)| \; ; \; \|x\| \leq 1\} \\ &= \sup \{|f(x)| \; ; \; \|x\| = 1\} \\ &= \sup \{|f(x)| / \|x\| \; ; \; x \neq \theta \} \\ &= \operatorname{Inf} \{M : |f(x)| \leq M\|x\| \; ; \; \text{para toda } x \text{ en } X\} \end{aligned}$$

El valor del funcional lineal x^* en X^* en el punto x en X es denotado por $x^*(x)$ ó por la siguiente notación: $\langle x,x^*\rangle$.

<u>Teorema</u>: X^* es un espacio de Banach.

Prueba: Hemos visto que X^* es un espacio normado, por lo que únicamente demostraremos que es completo. Sea $\{x_n^*\}$ una sucesión de Cauchy en X^*. Esto significa que $||x_n^*-x_m^*|| \longrightarrow 0$, cuando $n,m \longrightarrow \infty$. Ahora para cualquier x en X, $\{x_n^*(x)\}$ es una sucesión de Cauchy ya que : $|x_n^*(x)-x_m^*(x)| \leq ||x_n^*-x_m^*|| \cdot ||x||$.

Por tanto para cada x, existe un escalar $x^*(x)$ tal que $\{x_n^*(x)\} \longrightarrow x^*(x)$, por que la sucesión $\{x_n(x)\}$ está definida en los reales, que es completo. El funcional x^* definido sobre todo X es lineal ya que $x^*(\alpha x + \beta y) = \lim x_n^*(\alpha x + \beta y) = \lim [\alpha x_n^*(x) + \beta x_n^*(y)] = \alpha \lim x_n^*(x) + \beta \lim x_n^* (y) = \alpha x^*(x) + \beta x^*(y)$.

Ahora ya que $\{x_n^*\}$ es Cauchy, dado $\varepsilon>0$, existe una M tal que:

$|x_n^*(x)-x_m^*(x)| \leq ||x_n^*-x_m^*|| \ ||x|| < \varepsilon \ ||x||$ para toda $n,m > M$ y toda x; pero ya que $x_n^*(x) \longrightarrow x^*(x)$ tenemos que:

$$|x^*(x)-x_m^*(x)| \leq |x^*(x)-x_n^*(x)|+|x_n^*(x)-x_m^*(x)|$$

$$\leq |x^*(x)-x_n^*(x)|+||x_n^*-x_m^*|| \ ||x||$$

$$|x^*(x)-x_n^*(x)| < \varepsilon/2 \ ||x||; \text{ para toda } n > M_1$$

$$||x_n^*-x_m^*|| \ ||x|| < \varepsilon/2 \ ||x||; \text{ para toda } n,m>M_2$$

Sea $M=\max \{M_1,M_2\}$

$$\therefore |x^*(x)-x_m^*(x)| < \varepsilon/2 \ ||x||+\varepsilon/2 \ ||x||; \text{para } m > M$$

$$< \varepsilon \ ||x||; \text{ para } m > M$$

Entonces:

$$|x^*(x)| = x^*(x) - x_m^*(x) + x_m^*(x)| \leq |x^*(x) - x_m^*(x)| + |x_m^*(x)| < \varepsilon ||x|| +$$

$$||x_m^*|| \quad ||x||$$

$$\therefore \ |x^*(x)| < (\varepsilon + ||x_m^*||) \ ||x||$$

Por lo que x^* es un funcional lineal acotado.

Además, como $|x^*(x) - x_m^*(x)| < \varepsilon \ ||x||$ para $m > M$ y para toda x.

$$\text{entonces } \frac{|x^*(x) - x_m^*(x)|}{||x||} < \varepsilon \qquad \text{para } m > M \text{ y para toda } x.$$

$$\therefore \sup_{x \neq \theta} \frac{|x^*(x) - x_m^*(x)|}{||x||} < \varepsilon \qquad \text{para } m > M$$

$$\therefore \quad ||x^* - x_m^*|| < \varepsilon \qquad \text{para } m > M \ .$$

por lo tanto $x_m^* \longrightarrow x^*$. Lo que termina la prueba.

2.12 <u>Extensión de funcionales lineales</u> (Teorema de Hahn-Banach)

El teorema de Hahn-Banach, es el más importante teorema para el estudio de optimización en espacios lineales.

<u>Definición</u>: Sea f un funcional lineal definido sobre un subespacio M de un espacio vectorial X. Un funcional lineal F se dice que es una extensión de f, si F es definida sobre un subespacio N que contiene a M y si sobre M,

F es idéntico a f. En tal caso, diremos que F es una extensión de M a N.

En otros términos, el teorema de Hahn-Banach establece que un funcional lineal acotado f definido solamente sobre el sub-espacio M de un espacio normado, puede ser extendido a un funcional lineal acotado F definido sobre el espacio completo y con norma igual a la norma de f sobre M.

i.e.

$$||F|| = ||f||_M = \sup \{|f(m)| \quad ||m|| \ ; \ m \text{ en } M\}$$

Definición: Un conjunto parcialmente ordenado es un conjunto M sobre el cual está definido un ordenamiento parcial, esto es, una relación binaria la cual es escrita $\leq$ y satisface las condiciones:

i)	$a \leq a$ para toda a en M	(Reflexividad)
ii)	Si $a \leq b$ y $b \leq a$, entonces $a=b$	(Antisimetría)
iii)	Si $a \leq b$ y $b \leq c$, entonces $a \leq c$	(Transitividad)

"Parcialmente" enfatiza que M puede contener elementos a y b para los cuales no se cumple ni $a \leq b$, ni $b \leq a$. En tal caso a y b son llamados elementos incomparables. En contraste,dos elementos a y b son llamados elementos comparables si ellos satisfacen $a \leq b$ ó $b \leq a$ (ó ambos).

Un conjunto totalmente ordenado o cadena, es un conjunto parcialmente ordenado, tal que, para todo par de elementos del conjunto, son comparables. En otras palabras, una cadena es un conjunto parcialmente ordenado, que no tiene elementos incomparables.

Definición: Una cota superior de un subconjunto W de un conjunto parcialmente ordenado M, es un elemento u en M, tal que:

$$x \leq u \quad \text{para toda } x \text{ en } W$$

Dependiendo de M y W, esta u puede ó no existir.

Definición: Un elemento maximal de M es una m en M, tal que:

$$m \leq x \text{ implica que } x = m$$

Otra vez, M puede no tener elementos máximos. Note, sin embargo, que un elemento máximo no necesariamente es una cota superior.

Ejemplos:

Números Reales: Sea M el conjunto de todos los números reales, y sea $x \leq y$ en el sentido usual. M es totalmente ordenado. M no tiene elementos máximos.

Conjunto Potencia. Sea P(x) el conjunto potencia (el conjunto de todos los subconjuntos) de un conjunto X y sea $A \leq B$ si $A \subset B$, esto es, A es un subconjunto de B. Entonces, P(x) es parcialmente ordenado. El único elemento máximo de P(x) es X.

El espacio R^n. Sea M el conjunto R^n, con $x=(\xi_1,\xi_2,\ldots,\xi_n)$; $y=(\eta_1,\eta_2,\ldots,\eta_n)$ donde ξ_i, η_i son números reales y sea $x \leq y$ si $\xi_j \leq \eta_j$ para toda $j = 1,\ldots,n$, donde $\xi_j \leq \eta_j$ es en el sentido usual. Esto define un conjunto parcialmente ordenado sobre M.

Lema. Lema de Zorn. Sea $M \neq \emptyset$, un conjunto parcialmente ordenado. Suponga que toda cadena $C = M$ tiene una cota superior. Entonces M tiene al menos un elemento máximal.

En el teorema de Hahn-Banach, el objeto es extender un funcional lineal f, el cual es definido sobre un subespacio M de un espacio vectorial X y tiene ciertas propiedades de acotamiento el cual puede ser formulado en términos de un funcional sublineal.

Definición: Un funcional sublineal es un funcional p definido sobre un espacio vectorial X que satisface:

i) propiedad de subaditividad

$p(x_1+x_2) \leq p(x_1)+p(x_2)$ para toda x_1, x_2 en X.

ii) $p(\alpha x) = \alpha p(x)$ para toda $\alpha \geq 0$; x en X.

Obviamente, cualquier norma en un espacio normado es un funcional sublineal.

Teorema de Hahn-Banach (Extensión de funcionales lineales)

Sea X un espacio vectorial y p un funcional sublineal sobre X. Sea f un funcional lineal definido sobre un subespacio M de X satisfaciendo: $f(m) \leq p(m)$ para toda m en M. ...(1)
Entonces existe una extensión F de f, de M a X tal que:
$F(x) \leq p(x)$ para toda x en X ...(2)
Esto es: F es un funcional lineal sobre X, que satisface (1) sobre X y $F(x) = f(x)$ para toda x en M.

<u>Prueba</u>: Procederemos en tres pasos:

(a) El conjunto E de todas las extensiones lineales g de f que satisfacen $g(x) \leq p(x)$ sobre su dominio $D(g)$ pueden ser parcialmente ordenados y por el lema de Zorn obtener un elemento máximal. F de E.

(b) F está definido en todo el espacio X.

(c) Una relación auxiliar usada en (b).

(a) Sea E el conjunto de todas las extensiones lineales g de f, las cuales satisfacen la condición $g(x) \leq p(x)$ para toda x en $D(g)$. Claramente $E \neq \emptyset$ ya que f está en E. Sobre E ,podemos definir un ordenamiento parcial por: $g \leq h$, lo que significa que h es una extensión de g. Esto es ,por definición $D(h) \supset D(g)$ y $h(x) = g(x)$ para x en $D(g)$.

Para cualquier cadena $C \subset E$ definiremos ahora $\tilde{g}$ por $\tilde{g}(x) = g(x)$ si x está en $D(g)$; (g en C) $\tilde{g}$ es un funcional lineal cuyo dominio es

$$D(\tilde{g}) = \{\cup D(g) \ ; \ g \text{ en } C\}$$

el cual es un espacio vectorial, ya que C es una cadena. La definición de $\tilde{g}$ no es ambigua. Para una x en $D(g_1) \cap D(g_2)$, con g_1, g_2 en C, tenemos $g_1(x) = g_2(x)$ ya que C es una cadena, de modo que $g_1 \leq g_2$ ó $g_2 \leq g_1$. Claramente, $g \leq \tilde{g}$ para toda g en C. Por lo tanto es una cota superior de C. Ya que $C \subset E$ fue arbitraria, por el lema de Zorn se implica que E tiene un elemento maximal F. Por definición de E, este es una extensión lineal de f , el cual satisface

$$F(x) \leq p(x) \qquad x \text{ en } D(F) \qquad \ldots(3)$$

(b) Demostraremos ahora que D(F) es todo X. Supongamos que esto es falso. Entonces podemos escoger una y_1 en $X - D(F)$ y considere el subespacio Y_1 de X, generado por D(F) y y_1. Note que $y_1 \neq \theta$ ya que θ está en D(F). Cualquier x en Y_1 puede ser escrita:

$$x = y + \alpha y_1 \qquad y \text{ en } D(F)$$

Esta representación es única. Ya que $y + \alpha y_1 = \tilde{y}+\beta y$, con $\tilde{y}$ en D(F) implica $y-\tilde{y} = (\beta-\alpha)y_1$, donde $y-\tilde{y}$ está en D(F) siempre que y_1 no esté en D(F), lo cual unicamente es posible si $y-\tilde{y} = \theta$ y $\beta-\alpha = 0$. Esto significa unicidad. Un funcional g_1 sobre Y_1 está definido por:

$$g_1(y+\alpha y_1) = F(y) + \alpha c \qquad \ldots(4)$$

donde C es cualquier constante real. Fácilmente se puede demostrar que g_1 es lineal. Más aún, para $\alpha = 0$ tenemos $g_1(y) = F(y)$. Por lo tanto g_1 es una extensión propia de F(i.e., D(F) es un subconjunto propio de $D(g_1)$. Consecuentemente, si podemos probar que g_1 está en E, demostrando que $g_1(x) \leq p(x)$ para toda x en $D(g_1)$ ...(5) esto contradiría la maximalidad de F, por lo que $D(F) \neq X$ sería falso y D(F) = X, verdadero.

(c) Finalmente, deberemos probar que g_1 con una c convenien-
te en (4) satisface (5).

Consideremos cualquier y, z en D(F). De (1) y la propiedad subaditiva de los funcionales sublineales obtenemos:

$$F(y)-F(z) = F(y-z) \leq p(y-z) = p(y+y_1-y_1-z)$$

$$\leq p(y+y_1) + p(-y_1-z)$$

Por lo tanto:

$$-p(-y_1-z) - F(z) \leq p(y+y_1) - F(y) \qquad \ldots(6)$$

Donde y_1 es fija. Ya que y no aparece a la izquierda y z no aparece a la derecha, la desigualdad se sigue sosteniendo si tomamos el supremo sobre z en D(F) a la izquierda (llamémosle m_0)y el ínfimo sobre y en D(F) sobre la derecha, llamémosle m_1. Entonces $m_0 \leq m_1$ y para c con $m_0 \leq c \leq m_1$, tenemos de (6)

$$-p(-y_1-z) - F(z) \leq c \qquad \text{para toda z en D(F)} \qquad \ldots(7a)$$

$$c \leq p(y+y_1) - F(y) \qquad \text{para toda y en D(F)} \qquad \ldots(7b)$$

Probaremos (5) primero para negativa en (4) y después para α positiva. Para $\alpha < 0$, usamos (7a), sustituyendo z por $\alpha^{-1}y$, esto es:

$$-p(-y_1-(1/\alpha)y) - F((1/\alpha)y) \leq c$$

multiplicando por $-\alpha > 0$, obtenemos:

$$\alpha\, p(-y_1-\alpha^{-1}y) + F(y) \leq -\alpha\, c$$

De esto y(4), usando $y+\alpha y_1 = x$, obtenemos la desigualdad deseada:

$$g_1(x) = F(y)+\alpha c \leq -\alpha p(-y_1-\alpha^{-1}y) = p(\alpha y_1+y) = p(x)$$

Para $\alpha = 0$, tenemos x está en D(F) y no tenemos nada que probar. Para $\alpha > 0$, usaremos (7b) reemplazando y por $\alpha^{-1}y$ para obtener:

$$c \leq p(\alpha^{-1}y + y_1) - F(\alpha^{-1}y)$$

multiplicando por $\alpha > 0$, tenemos:

$$\alpha c \leq \alpha p(\alpha^{-1}y+y_1) - F(y) = p(x) - F(y)$$

De esto y(4)

$$g_1(x) = F(y) + \alpha c \leq p(x)$$

<u>Corolario 1</u>: Sea f un funcional lineal acotado, definido sobre un subespacio M de un espacio vectorial normado (real) X. Entonces, existe un funcional lineal acotado F, definido sobre X el cual es una extensión de f y cuya norma es igual a la norma de f, sobre M.

<u>Corolario 2</u>: Sea X_0 un elemento de un espacio normado X. Entonces existe un funcional lineal acotado F, diferente de cero, sobre X tal que:

$$F(x_0) = ||F|| \; ||x_0||$$

2.13 El dual de $C[a,b]$

Teorema. (Teorema de la Representación de Riesz) Sea f un funcional lineal acotado, sobre $X = C[a,b]$. Entonces, existe una función v de variación acotada sobre $[a,b]$ tal que para toda x en X

$$f(x) = \int_a^b x(t)\, d\, v\, (t)$$

y tal que la norma de f es la variación total de v sobre $[a,b]$. Recíprocamente, cada función de variación acotada sobre $[a,b]$ define un funcional lineal acotado en este sentido.

Prueba: Sea B el espacio de las funciones acotadas sobre $[a,b]$ con la norma de un elemento en B definida:

$$||x||_B = \sup \{x(t); a \leq t \leq b\}$$

El espacio $C[a,b]$ puede ser considerado como un subespacio de B. Ya que $[a,b]$ es un conjunto compacto y toda función continúa definida sobre un compacto es acotada. Entonces, si f es un funcional lineal sobre $X = C[a,b]$ existe por el teorema de Hahn-Banach, un funcional lineal F sobre B el cual es una extensión de f y tiene la misma norma.

Para cualquier s en $[a,b]$, definamos la función μ_s por:

$$\mu_a = 0$$

$$\mu_s(t) = \begin{cases} 1 & \text{si } a \leq t \leq s \\ 0 & \text{si } s < t \leq b \end{cases} \qquad \text{para } a < s \leq b$$

Obviamente μ_s está en B.

Definamos $v(s) = F(\mu_s)$ y demostremos que v es de variación acotada. Para este propósito, sea $a=t_0<t_1,t_2,\ldots,<t_n = b$ una partición finita de $[a,b]$. Denotemos $\varepsilon_i=\mathrm{sgn}[v(t_i)-v(t_{i-1})]$

$$\text{e.i.} \quad \varepsilon_i=\begin{cases} +1 \text{ si el signo de la diferencia } v(t_i)-v(t_{i-1}) \text{ es positivo} \\ \cdot \\ -1 \text{ si el signo de la diferencia } v(t_i)-v(t_{i-1}) \text{ es negativo} \end{cases}$$

Con estas consideraciones podemos escribir:

$$\sum_{i=1}^{n} |v(t_i)-v(t_{i-1})| = \sum_{i=1}^{n} \varepsilon_i [v(t_i)-v(t_{i-1})]$$

$$= \sum_{i=1}^{n} \varepsilon_i [F(\mu_{t_i})-F(\mu_{t_{i-1}})]$$

$$= F\left[\sum_{i=1}^{n} \varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}})\right] \leq ||F|| \; || \sum_{i=1}^{n} \varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}})||$$

Para cada t_{i-1}, t_i se puede tener los siguientes casos: (ver figura 2.13.1).

$$\mu_{t_i} - \mu_{t_{i-1}} = \begin{cases} 0 & \text{si } a \leq x \leq t_{i-1} \\ 1 & \text{si } t_i-1 < x \leq t_i \\ 0 & \text{si } t_i < x \leq b \end{cases}$$

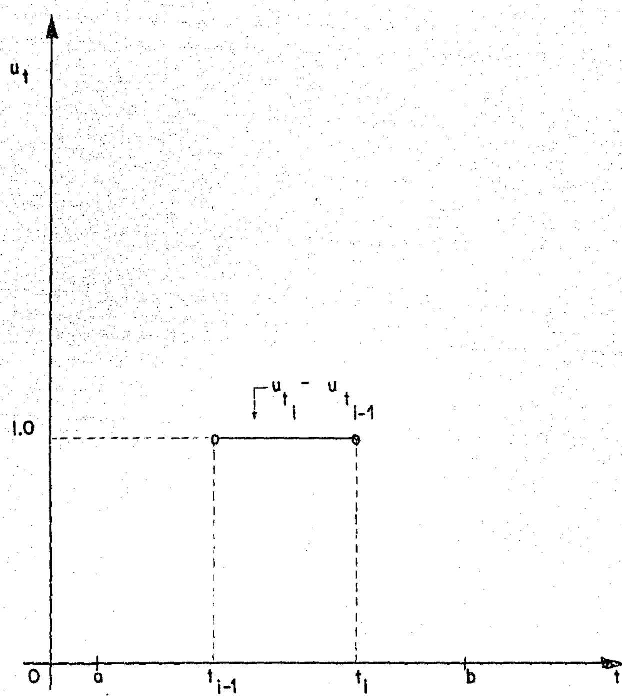

FIGURA 2.13.1

$$\varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}}) \begin{cases} 0 & \text{Si } a \le x \le t_{i-1} \\ 1 & \text{Si } t_{i-1} < x \le t_i \text{ ; } \varepsilon_i = + \\ -1 & \text{Si } t_{i-1} < x \le t_i \text{ ; } \varepsilon_i = - \\ 0 & \text{Si } t_i < x \le b \end{cases}$$

$$\sum_{i=1}^{n} \varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}}) = \pm 1 \text{ si } t_{i-1}<x\le t_i;\ \varepsilon_i=\pm;\ i=1,\ldots,n; a\le t_i\le b$$

Como $\sum_{i=1}^{n} \varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}})$ está en B

Entonces: $\left\|\sum_{i=1}^{n} \varepsilon_i(\mu_{t_i}-\mu_{t_{i-1}})\right\| = \text{Sup.}\left|\{\pm 1 \text{ si } t_{i-1}<x\le t_i;\ \varepsilon_i=\pm;\ i=1,\ldots,n; a\le t_i\le b\}\right|$

$= 1$

Por lo tanto:

$$\sum_{i=1}^{n} |v(t_i) - v(t_{i-1})| \le ||F|| = ||f||$$

Por lo que v es de variación acotada, con variación total menor o igual a $||f||$

i.e. $V.T.(v) \le ||f||$.

Derivaremos ahora una representación para f sobre X.

Sea x en X; y :

$$z(t) = \sum_{i=1}^{n} x(t_{i-1}) \left[\mu_{t_i}(t)-\mu_{t_{i-1}}(t)\right]$$

donde $\{t_i\}$ es una partición finita de $[a,b]$

Entonces:

$$||z-x||_B = \max\{z(t)-x(t); a \leq t \leq b\} = \max\{\max\{|z(t)-x(t)|;\ t_{i-1} \leq t \leq t_i\}\ i\ \}$$

$$||z-x||_B = \max\{\max\{|x(t_{i-1}) - x(t)|\ ;\ t_{i-1} \leq t \leq t_i\}\ \ i\ \ \}$$

la cual por la continuidad informe de x, tiende a ceró cuando la partición se hace cada vez más fina. Por lo tanto ya que F es contínua

$$F(z) \longrightarrow F(x) = f(x).$$

Pero : $F(z) = \sum_{i=1}^{n} x(t_{i-1}) \left[v(t_i) - v(t_{i-1})\right]$

y por definición de integral de Stieltjes

$$F(z) \longrightarrow \int_a^b x(t)dv(t)$$

Por lo tanto

$$f_{(x)} = \int_a^b x(t)dv(t)$$

la cual por propiedades de Integral de Stieltjes

$$\left| \int_a^b x(t)\ dv(t) \right| \leq ||x|| \cdot V.T.(v)$$

y en consecuencia

$$||f|| \leq V.T.\ (v)$$

Por otro lado tenemos que

$$||f|| \geq V.T.(v)$$

$$\therefore \quad ||f|| = V.T.(v)$$

2.14 Operadores lineales

Una transformación es, como discutimos anteriormente, un mapeo de un espacio vectorial a otro.

Si T mapea el espacio X al Y, escribimos $T : X \longrightarrow Y$, y si T mapea el vector x en X al vector y en Y, escribimos $y = T(x)$ y nos referiremos a y como la imagen de x bajo T.

Una transformación, puede ser definida, solamente sobre un subconjunto $D \subset X$, llamado dominio de T, aunque en la mayoría de los casos $D = X$. La colección de todos los vectores y en Y para los cuales existe una x en D, con $y = T(x)$, es llamado el rango de T.

Si $T:X \longrightarrow Y$, y S es un conjunto dado en X, denotaremos por T(S) a la imagen de S en Y definida como el subconjunto de Y que consiste de los puntos de la forma $y = T(s)$, con s en S. Similarmente, dado cualquier conjunto $P \subset Y$, denotaremos por $T^{-1}(P)$ a la imagen inversa de P, la cual es el conjunto consistente de todos los puntos x en X que satisfacen T(x) está en P.

Definición: El rango de un operador lineal $A : X \longrightarrow Y$ es denotado por R(A) ; y está dado por:

$$R(A) = \{y \text{ en } Y : Ax = y \; ; \; x \text{ en } X\}$$

Es además un subespacio de Y.

Definición: El conjunto $\{x \text{ en } X : Ax = \Theta\}$ que corresponde al operador lineal A, es llamado el espacio nulo de A y se denota por N(A). Este es un subespacio de X.

En esta sección, trataremos transformaciones lineales, a las cuales nos referiremos como operadores lineales. Y escribiremos indistintamente A(x) ó Ax.

Proposición 1. Un operador lineal sobre un espacio normado X es contínuo en todo punto de X si es contínuo en un solo punto.

Prueba: Sea A operador lineal, contínuo en x_0 en X,

X espacio normado

Sea $\{X_n\}$ una sucesión en X que converge a x en X.

Por la linealidad de A:

$$||A(x_n)-A(x)||=||A(x_n)-A(x)+A(x_0)-A(x_o)||=||A(x_n-x+x_0)-A(x_0)||$$

Como : $x_n \longrightarrow x$, entonces $x_n - x \longrightarrow \Theta$

por lo tanto: $x_n - x + x_0 \longrightarrow x_0$

Por continuidad de A en X_0

$A(x_n - x + x_0) \longrightarrow A(x_0)$

Por lo tanto $||A(x_n - x + x_0) - A(x_0)|| \longrightarrow 0$

sustituyendo $||A(x_n) - A(x)|| \longrightarrow 0$

Entonces: $A(x_n) \longrightarrow A(x)$

<u>Definición</u>: Un operador lineal A, de un espacio normado X a un espacio Y, se dice que es <u>acotado</u>, si existe una constante M tal que satisface: $||Ax|| \leq M||x||$ para toda x en X.

La más pequeña de las M, que satisface la condición anterior se denota $||A||$ y es llamada la <u>norma</u> de A.

$$||A|| = \text{Sup} \{ ||Ax|| \; ; \; ||x|| \leq 1 \}$$
$$||A|| = \text{Sup} \{ ||Ax|| / ||x|| \; ; \; ||x|| \neq \theta \}$$

<u>Proposición 2</u>: Un operador lineal es acotado si y sólo si es contínuo.

<u>Definición</u>: El espacio de todos los operadores lineales acotados del espacio normado X al espacio normado Y es denotado B(X,Y)

<u>Teorema 1</u>: Sea X y Y espacios normados, con Y completo. Entonces el espacio B(X,Y) es completo.

2.15 Operadores adjuntos

Las restricciones impuestas en muchos problemas de optimización por ecuaciones diferenciales, ecuaciones de matrices, etc, pueden ser descritos por operadores lineales. La resolución de estos problemas, casi invariablemente, llaman la consideración de un operador asociado: el adjunto. La razón para esto es que los adjuntos proporcionan un mecanismo conveniente para describir relaciones de ortogonalidad y dualidad, las cuales permiten, todo análisis de optimización.

Definición: Sean X y Y espacios normados y sea A en B(X,Y). El operador adjunto $A^* : Y^* \longrightarrow X^*$ es definido por la ecuación:

$$\langle x, A^*y^* \rangle = \langle Ax, y^* \rangle$$

Esta importante definición, requiere un poco de explicación y justificación.

Dada una y^*, fija de Y^*, la cantidad $\langle Ax, y^* \rangle$ es un escalar para cada x en X, y es por tanto un funcional sobre X. Más aún, por la linealidad de y^* y A, tenemos:

$$\begin{aligned}\langle A(\alpha x_1+\beta x_2), y^* \rangle &= \langle (\alpha Ax_1+\beta Ax_2), y^* \rangle \\ &= \langle \alpha Ax_1, y^* \rangle + \langle \beta Ax_2, y^* \rangle \\ &= \alpha \langle Ax_1, y^* \rangle + \beta \langle Ax_2, y^* \rangle\end{aligned}$$

por lo que este funcional es lineal.

Y además, ya que :

$$|\langle Ax,\ y^*\rangle| \leq ||y^*|| \cdot ||Ax|| \leq || y^* || \cdot ||A|| \cdot ||x||$$

se sigue que este funcional es acotado.

Definiremos entonces $A^* y^* = x^*$. Este operador es únicó y lineal.

Como por definición, se satisface:

$$y^* \quad (Ax) = (A^* y^*)\ x$$

para cada x en X, entonces podemos escribir

$$y^* A = A^* y^*$$

Donde el lado izquierdo denota el funcional sobre X el cual es la composición de los operadores A y y^*.

$$\begin{tikzcd} X^* & Y^* \arrow[l, "A^*"'] \\ X \arrow[r, "A"] & Y \end{tikzcd}$$

Fig. 2.15.1 Un operador y su adjunto

<u>Teorema</u>: El operador adjunto A^* de un operador A en $B(X,Y)$ es lineal y acotado, con

$$||A^*|| = ||A||$$

<u>Prueba</u>:

i) <u>Linealidad</u>

$$\begin{aligned}<x,A^*(y_1^*+y_2^*)> &= <Ax,(y_1^*+y_2^*)> \\ &= <Ax,y_1^*> + <Ax,y_2^*> \\ &= <x,A^*y_1^*> + <x,A^*y_2^*>\end{aligned}$$

Por lo que A^* es lineal.

ii) <u>Acotamiento</u>: Para toda x en X, y cualquier y^* en Y^*, se tiene:

$$|<x,A^*y^*>| = |<Ax,y^*>| \leq ||y^*|| \, ||Ax|| \leq ||y^*|| \, ||A|| \, ||x||$$

como se cumple para toda x, se cumple para el supremo.

$$\therefore \text{Sup } \{|<x,A^*y^*>|/||x||;\ x \text{ en } X\} = ||A^*y^*|| \leq ||y^*|| \, ||A||$$

entonces: $||A^*y^*|| \leq ||y^*|| \, ||A||$ ..... por lo que es acotado

iii) <u>$||A^*||=||A||$</u>; $||A^*y^*||/||y^*|| \leq ||A||$

como se cumple para cualquier y^* en Y^* se tiene:

$$\text{Sup } \{||A^*y^*||/||y^*|| \ ;\ y^* \text{ en } Y^*\} = ||A^*|| \leq ||A||$$

Sea $x_0 \neq \Theta$ un elemento cualquiera de X, por el Teorema de Hahn-Banach, existe un elemento y_0^* en Y^*, $||y_0^*|| = 1$ tal que $\langle Ax_0, y_0^* \rangle = ||Ax_0||$, entonces

$$||Ax_0|| = | \langle x_0, A^*y_0^* \rangle | \leq ||A^*y_0^*|| \, ||x_0|| \leq ||A^*|| \, ||y_0^*|| \, ||x_0||$$

$$||A^*|| \, ||y_0^*|| \, ||x_0|| = ||A^*|| \quad ||x_0||$$

$$||Ax_0|| \leq ||A^*|| \quad ||x_0||$$

$$\therefore \qquad ||Ax_0|| / \, ||x_0|| \leq ||A^*||$$

$$\therefore \quad \sup \{ ||Ax_0|| / ||x_0|| ; \; x_0 \text{ en } X \} = ||A|| \leq ||A^*||$$

$$||A|| \leq ||A^*||$$

por lo tanto

$$||A|| = ||A^*||$$

Adicionalmente al resultado anterior, el operador adjunto, cumple las siguientes relaciones algebráicas, las cuales pueden ser fácilmente demostradas, a partir de la definición básica.

Proposición 1: Los operadores adjuntos satisfacen las siguientes propiedades:

1. Si I es el operador identidad sobre un espacio normado X, entonces $I^* = I$.

2. Si A_1, A_2 están en $B(X,Y)$, entonces $(A_1+A_2)^* = A_1^* + A_2^*$.

3. Si A está en $B(X,Y)$ y α es un escalar real, entonces $(\alpha A)^* = \alpha A^*$.

4. Si A_1 está en $B(X,Y)$, A_2 en $B(Y,Z)$ entonces $(A_2A_1)^* = A_1^*A_2^*$.

5. Si A está en $B(X,Y)$ y A tiene una inversa acotada, entonces $(A^{-1})^* = (A^*)^{-1}$.

Lema: Sean X y Y espacios de Banach y sea A en $B(X,Y)$. Suponga que $R(A)$ es cerrado. Entonces, existe una constante K, tal que para cada y en $R(A)$ existe una x, que satisface:

$$Ax = y \quad ; \qquad \|x\| \leq K\|y\|$$

Prueba: Sea $N = N(A)$, y considere el espacio X/N consistente de todas las clases de equivalencia $[x]$ módulo N.

Defina $\bar{A}$: $X/N \longrightarrow R(A)$ por $\bar{A}\, x = A(x)$. Se puede verificar fácilmente que $\bar{A}$ es uno a uno, sobre [o sea $R(X/N) = R(A)$], lineal y acotado. Ya que $R(A)$ es cerrado, y es subespació de Y que es Banach, $R(A)$ es Banach. Ahora bien, N es un subespacio cerrado de X, y por lo tanto, X/N es también un espacio de Banach(ver sección 2). Ahora bien, por el Teorema Inverso de Banach, A tiene un inverso acotado. Entonces, dado y en $R(A)$ existe x tal que $\bar{A}[x] = Ax = y$; y por lo tanto: $[x] = \bar{A}^{-1}y$

$$\|[x]\| = \|\bar{A}^{-1}y\| \leq \|\bar{A}^{-1}\| \; \|y\|$$

Por definición de norma de $[x]$; existe x_0 en $[x]$

con $\|x_0\| \leq 2 \; \|[x]\|$

$\therefore \quad \|x_0\| \leq 2 \; \|\bar{A}^{-1}\| \quad \|y\|$

Haciendo $K = 2\ ||\bar{A}^{-1}||$

$||x_0|| \leq K\ ||y||$ y como x_0 en $[x]$ es de la forma $x + n$ para n en N, tenemos

$Ax_0 = Ax + An = A_x = y$ (ya que $An = \Theta$)

<u>Teorema</u>: Sean X,Y espacios de Banach y sea A en B(X,Y). Sea R(A) cerrado. Entonces:

$$R(A^*) = [N(A)]^{\perp}$$

<u>Prueba</u>:

Recordemos que:

$$R(A^*) = \{x^* \text{ en } X^*: \text{Existe } y^* \text{ en } Y^* \text{ con } A^*y^* = x^*\}$$

$[N(A)]^{\perp} = \{x^*$ en X^*: $\langle x,x^*\rangle = 0$; para toda x que cumpla $Ax=\Theta\}$

Sea ahora, x^* en $R(A^*)$. Entonces $x^*=A^*y^*$ para alguna y^* en Y^*. Para cualquier x en $N(A)$, se cumple:

$\langle x,x^*\rangle = \langle x,A^*y^*\rangle = \langle Ax,y^*\rangle = \langle \Theta,y^*\rangle = 0$

por lo que x^* está en $[N(A)]^{\perp}$ y por lo tanto $R(A^*)$ es subconjunto de $N(A)$.

Sea ahora, x^* en $[N(A)]^{\perp}$. Para y en $R(A)$ y cada x que satisfa ce $Ax = y$, el funcional $\langle x,x^*\rangle$ tiene el mismo valor. Para ello, sean x_1, x_2 satisfaciendo $Ax_1 = y$, $Ax_2 = y$. Entonces $A(x_1 - x_2) = \Theta$

Por lo que $x_1 - x_2$ esta en $N(A)$.

$$\langle x_1-x_2,x^*\rangle = \langle x_1,x^*\rangle - \langle x_2,x^*\rangle = 0$$

Estará, entonces, bien definido $f(y) = \langle x,x^* \rangle$ sobre $R(A)$.
Por el lema anterior, para cada y en $R(A)$, existe una x con $||x|| \leq K\,||y||$; $Ax = y$. De este modo:

$|f(y)| = |\langle x, x^* \rangle| \leq ||x^*||\;||x|| \leq K||x^*||\,||y||$; por lo que f es un funcional lineal acotado sobre $R(A)$.

Por el corolario al Teorema de Hahn-Banach, existe y^* en Y^*, el cual es una extensión de f. Entonces, $\langle x,A^*y^* \rangle = \langle Ax,y^* \rangle = \langle x,x^* \rangle$ se sigue que $A^*y^* = x^*$ y por tanto x^* está en $R(A^*)$ o sea $[N(A)]^{\perp}$ es subconjunto de $R(A)$. Lo que prueba que $R(A) = [N(A)]^{\perp}$.

2.16 Teoría Local

Sean:

X:espacio vertorial

Y:espacio normado

$T: D \longrightarrow Y$ donde $D \subset X$ y $R(T) \subset Y$

donde $R(T)$ es el rango de T.

Definición: Sea x en D y h arbitrario en X. Si el límite:

$$(1) \quad \delta T(x;h) = \lim_{\alpha \to 0} \frac{T(x + \alpha h) - T(x)}{\alpha}$$

existe, es llamado la diferencial de Gateaux de T en x con incremento h. Si el límite (1) existe para cada h en X, la transformación T, se dice, es diferenciable en x.

Note que:

(1) es considerado si $x + \alpha h$ en D para toda α suficientemente pequeña; para cada x en D y h considerada como variable; $\delta T(x;h)$ define una transformación de X a Y.

Si T es lineal,

$$\delta T(x;h) = \lim_{\alpha \to 0} \frac{T(x + \alpha h) - T(x)}{\alpha} = \lim_{\alpha \to 0} \frac{T(x + \alpha h - x)}{\alpha}$$

$$= \lim_{\alpha \to 0} \frac{\alpha T(h)}{\alpha} = T(h).$$

Si Y es el eje real, $T : X \longrightarrow R$; o sea T es un funcional f sobre X. Entonces: $\delta f(x;h) = \frac{d}{d\alpha} f(x+ h)\Big|_{\alpha=0} = \lim_{\alpha \to 0} \frac{f(x+\alpha h)-f(x)}{\alpha}$

Ejemplo 1: $X = E^n$ y $f(x) = f(x_1,\dots,x_n)$ un funcional teniendo derivadas parciales continuas con respecto a cada variable x_i, entonces

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

$$= \lim_{\alpha \to 0} \frac{\nabla f(x).\alpha h + \delta(||\alpha h||_X)}{\alpha}$$

$$= \nabla f(x).h$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i$$

Ejemplo 2: $X = C[0,1]$ y $f(x) = \int_0^1 g(x(t),t)\,dt$

donde g_x existe y es continua con respecto a x y t.

$$\delta f(x;h) = \lim_{\alpha \to 0} \frac{\int_0^1 g(x(t)+\alpha h(t),t)dt - \int_0^1 g(x(t),t)dt}{\alpha}$$

$$= \int_0^1 \lim_{\alpha \to 0} \frac{g(x(t)+\alpha h(t),t) - g(x(t),t)\ dt}{\alpha}$$

Para cada t fija, por el teorema del valor medio:

$$g(x(t)+\alpha h(t)-g(x(t)) = g_x(\bar{x}(t),t)\ (\alpha h(t))$$

donde $\bar{x}(t)$ está entre $x(t)$ y $x(t) + \alpha h(t)$

$$\therefore \int_0^1 \lim_{\alpha \to 0} \frac{g_x(\bar{x}(t),t)\ \alpha h(t)}{\alpha}\ dt$$

$$= \int_0^1 g_x(x(t),t)\ h(t)\ dt$$

Sean ahora:

X : espacio normado

Y : espacio normado

T : $D \subset X$, D abierto

Definición: Si para una x en D y cada h en X, existe $\delta T(x;h)$ en Y el cual es lineal y continuo con respecto a h, y

$$\lim_{||h||\to 0} \frac{||T(x+h) - T(x) - \delta T(x;h)||}{||h||} = 0$$

=> T se dice que es diferenciable en el sentido de Frechet en x y $\delta T(x;h)$ es la diferencial de Frechet de T en X con incremento h.

Proposición 1: Si la transformación T tiene una diferencial de Frechet, esta es única.

Prueba: Suponga que $\delta T(x;h)$ y $\delta' T(x;h)$ satisfacen los requerimientos de la definición:

$$\begin{aligned} ||\delta T(x;h) - \delta' T(x;h)|| &= ||-(T(x+h) - T(x)) + \delta T(x;h) + T(x+h) - \\ &\qquad T(x) - \delta' T(x;h)|| \\ &\leq ||T(x+h) - T(x) - \delta\, T(x;h)|| + \\ &\qquad ||T(x+h) - T(x) - \delta' T(x;h)|| \end{aligned}$$

entonces: $||\delta T(x;h) - \delta' T(x;h)|| = \acute{o}\,(||h||)$

Suponga que existe h tal que:

$$\frac{\|\delta T(x;h) - \delta T(x;h)\|}{\|h\|} = \lambda \neq 0$$

$\Rightarrow$ Por linealidad: $\dfrac{\|\delta T(x;\varepsilon h) - \delta T(x;\varepsilon h)\|}{\|\varepsilon h\|} = \lambda$

pero para $\varepsilon \to 0$ $\quad \delta(\|h\|) \not\to 0$ $\quad$ (contradicción)

Proposición 2: Si la diferencial de Fréchet de T existe en x, entonces la diferencial de Gateaux existe en x y son iguales.

Prueba: $\|T(x+\alpha h)-T(x)-\delta T(x;\alpha h)\| = \delta(\|\alpha h\|)$ (por hipótesis). Si $\alpha \to 0$, entonces $\|\alpha h\| \to 0$

$$\therefore \lim_{\alpha \to 0} \frac{\|T(x+\alpha h)-T(x)-\delta T(x;\alpha h)\|}{\|\alpha\|} = 0$$

Por linealidad: $\displaystyle \lim_{\alpha \to 0} \left\| \frac{T(x+\alpha h)-T(x)}{\alpha} - \delta T(x;h) \right\| = 0$

$$\Rightarrow \lim_{\alpha \to 0} \frac{T(x+\alpha h)-T(x)}{\alpha} = \delta T(x;h) .$$

Proposición 3. Si la transformación T definida sobre el conjunto abierto D en X es diferenciable Fréchet en x, entonces es continua en X.

Prueba: Dado $\varepsilon > 0$ Existe δ tal que x+h en $B_X(\delta)$, entonces $\|h\| < \delta$

$$\|T(x+h) - T(x) - \delta T(x;h)\| < \varepsilon \|h\|$$

$$\begin{aligned}\therefore \|T(x+h) - T(x)\| &\leq \|T(x+h)-T(x)-\delta T)x;h)\|+\|\delta T(x;h)\| \\ &< \varepsilon\|h\|+\|\delta T(x;h)\| \leq \varepsilon \|h\|+M_1 \|h\| \\ &= \underbrace{\varepsilon+M_1}_{M} \|h\| = M \|h\|\end{aligned}$$

2.17 Derivadas de Frechet

Suponga que la transformación T, definidas en el Dominio D $D \subset X$; (x espacio normado) es Fréchet diferenciable sobre todo D.

$\Longrightarrow$En un punto fijo $x \in D$ la diferencial de Fréchet $\delta T(x\ ;\ h)$ es por definición de la forma:

$$\delta T(x;h) = A_x h$$

Donde A_x es un operador lineal acotado de X a Y

{Esto sucede ya que $T(x;h)$ es lineal y continúa.}

Entonces como varía x sobre D; la correspondencia $x \longrightarrow A_x$ define una transformación de D a $B(X;Y)$. $\{T':D \longrightarrow B(X,Y).\}$. A esta transformación se le conoce como la derivada de Fréchet T' de T

$$\therefore \quad \delta T(x;h) = T'(x)\ h$$

Si la correspondencia: $x \longrightarrow T'(x)$ es continúa en el punto x_0 o sea: Dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$||x-x_0|| < \delta \quad \text{implica} \quad ||T'(x)-T'(x_0)||$$

$\Longrightarrow$Decimos que la derivada de Fréchet de T es contínua en x_0.

Nota: No debemos confundir esto con el enunciado que dice que $T'(x_0)$ es un mapeo contínuo de X a Y, que es una de las propiedades básicas de la derivada de Fréchet.

Si la derivada de T es contínua sobre una bola abierta S decimos que T es continuamente diferenciable según Frechet sobre S.

En el caso especial en que la transformación original es un funcional f sobre el espacio X, tenemos:

$$\delta f(x;h) = f'(x)h$$

donde $f'(x) \in X^*$ para cada x; $f'(x)$ es llamado el gradiente de f en x.

<u>NOTACION</u>: $f'(x) = \nabla f(x)$; $\langle h , f'(x)\rangle = \delta f(x;h) = f'(x)h$

Mucho de la teoría de las derivadas ordinarias puede ser generalizado a las derivadas de Fréchet. Por ejemplo: El teorema de la función implícita y la serie de Taylor tienen una extensión satisfactoria.

<u>PROPOSICION 1</u>. Si T_1 y T_2 son diferenciables Fréchet en $x\in D$, entonces $\alpha_1 T_1 + \alpha_2 T_2$ es diferenciable Fréchet en x, y:

$$(\alpha_1 T_1 + \alpha_2 T_2)'(x) = \alpha_1 T_1'(x) + \alpha_2 T_2'(x)$$

Prueba: $\alpha_1 \delta_1(x;h) + \alpha_2 \delta_2(x;h)$ es lineal y contínuo respecto a h ya que es la suma finita de lineales y contínuos.

$$\lim_{||h||\to 0} \frac{||(\alpha_1 T_1+\alpha_2 T_2)(x+h)-(\alpha_1 T_1+\alpha_2 T_2)(x)-\alpha_1 \delta T_1(x;h)-\alpha_2 \delta T_2(x;h)||}{||h||} =$$

$$\lim_{||h||\to 0} \frac{||(\alpha_1 T_1)(x+h)+\alpha_2 T_2(x+h)-\alpha_1 T_1(x)-\alpha_2 T_2(x)-\alpha_1 \delta T_1(x;h)-\alpha_2 \delta T_2(x;h)||}{||h||} \leq$$

$$\lim_{||h||\to 0} \frac{||\alpha_1 T_1(x+h)-\alpha_1 T_1(x)-\alpha_1 \delta T_2(x;h)||}{||h||} + \lim_{||h||\to 0} \frac{||\alpha_2 T_2(x+h)-\alpha_2 T_2(x)-\alpha_2 \delta T_2(x;h)||}{||h||}$$

$$= 0 + 0$$

$$\therefore \lim_{||h||\to 0} \frac{||(\alpha_1T_1+\alpha_2T_2)(x+h)-(\alpha_1T_1+\alpha_2T_2)(x)-\{\alpha_1\delta T_1(x;h)+\alpha_2\delta T_2(x;h)\}||}{||h||}=0$$

=> Por la unicidad de la diferencial de Fréchet, tenemos:

$$\delta(\alpha_1T_1+\alpha_2T_2)(x;h) = \alpha_1\delta T_1(x;h)+ \alpha_2\delta T_2(x;h)$$

$\therefore$ $\alpha_1T_1 + \alpha_2T_2$ es Fréchet diferenciable

Además:

$$\delta(\alpha_1T_1+\alpha_2T_2)(x;h)=(\alpha_1T_1+\alpha_2T_2)'(x)h=\alpha_1T'_1(x).h+\alpha_2T'_2(x)h=\alpha_1\delta T_1(x;h)+\alpha_2\delta T_2(x;h)$$

$$\therefore (\alpha_1T_1+\alpha_2T_2)'(x)\ h = \alpha_1T'_1(x)h + \alpha_2T'_2(x)\ h$$

$$(\alpha_1T_1+\alpha_2T_2)'(x)\ h = (\alpha_1T'_1(x) + \alpha_2T'_2(x))h$$

$$\Rightarrow (\alpha_1T_1+\alpha_2T_2)' = \alpha_1T'_1(x) + \alpha_2T'_2(x)$$

<u>PROPOSICION 1</u> (Regla de la cadena aplicada a derivadas de Fréchet)

Sea S una transformación mapeando un conjunto abierto $D \subset X$ a un conjunto abierto $E \subset Y$ y sea P una transformación mapeando E a un espacio normado Z. Haciendo $T = P.S$ y suponiendo que S es diferencial de Fréchet en $x \in D$ y que P es diferenciable de Fréchet en $y=s(x)$ en E. Entonces T es diferenciable de Fréchet en x, y:

$$T'(x)=P'(y)S'(x)$$

<u>Prueba</u>: Para h en x tal que x+h en D tenemos:

$$T(x+h) - T(x) = P[S(x+h)] - P[S(x)]$$

$$= P[S(x)+S(x+h)-S(x)] - P[S(x)]$$

Haciendo $g = S(x+h) - S(x)$..... {obviamente en Y}

$y = S(x)$

$\Rightarrow T(x+h) - T(x) = P\left[y+g\right] - P\left[y\right]$.......(1)

como P es diferenciable en $y = S(x)$

$$\lim_{\|g\| \to 0} \frac{\|P(y+g) - P(y) - \delta P(y;g)\|}{\|g\|} = 0$$

$\therefore$ sustituyendo de (1)

$$\lim_{\|g\| \to 0} \frac{T(x+h) - T(x) - \delta P(y;g)\|}{\|g\|} = 0$$

$\Rightarrow \|T(x+h)-T(x) - \delta P(y;g)\| = \sigma(\|g\|)$

como $\delta P(y;g) = P'(y).g$

$\Rightarrow \|T(x+h) - T(x) - P'(y)g\| = \sigma(\|g\|)$

como S es diferenciable en x

$$\lim_{\|h\| \to 0} \frac{\|S(x+h) - S(x) - \delta S(x;h)\|}{\|h\|} = 0$$

$\Rightarrow \|S(x+h) - S(x) - \delta S(x;h)\| = \sigma(\|h\|)$

$\therefore \|S(x+h) - S(x) - S'(x)h\| = \sigma(\|h\|)$

$\therefore \|g - S'(x)h\| = \sigma(\|h\|)$

Por otra parte:

$$\sigma(\|g\|)+\sigma(\|h\|)=\|T(x+h)-T(x)-P'(y)g\|+\|P'(y)\|\|g-S'(x)h\|$$
$$\geq \|T(x+h)-T(x)-P'(y)g\|+\|P'(y)\left[g-S'(x)h\right]\|$$

$$\geq \|T(x+h)-T(x)-P'(y)g+P'(y)g-P'(y)S'(x)h\|$$

$$\geq \|T(x+h)-T(x)-P'(y)S'(x)h\|$$

$$\therefore \|T(x+h)-T(x)-P'(y)S'(x)h\| \leq \delta(\|g\|) + \delta(\|h\|) \ldots\ldots (1)$$

como S es contínua en x

$$\|g\| = \|S(x+h) - S(x)\| < M\|h\|$$

o sea que cuando $\|h\| \longrightarrow 0$ ello implica que $\|g\|$ también tenderá a cero...(2)

=> de (1) sacando límite y dividiendo entre $\|h\|$

$$\lim_{\|h\| \to 0} \frac{\delta\|g\|}{\|h\|} + \lim_{\|h\| \to 0} \frac{\delta\|h\|}{\|h\|} \geq \lim_{\|h\| \to 0} \frac{\|T(x+h)-T(x)-P'(y)S'(x)h\|}{\|h\|} \geq 0$$

de (2) tenemos

$$\lim_{\|g\| \to 0} \frac{\delta\|g\|}{\|h\|} + \lim_{\|h\| \to 0} \frac{\delta\|g\|}{\|h\|} \geq \lim_{\|h\| \to 0} \frac{\|T(x+h)-T(x)-P'(y)S'(x)h\|}{\|h\|} \geq 0$$

Pero:

$$\lim_{\|g\| \to 0} \frac{\delta\|g\|}{\frac{1}{M}\|g\|} \geq \lim_{\|g\| \to 0} \frac{\delta\|g\|}{\|h\|} \geq \lim_{\|h\| \to 0} \frac{\|T(x+h)-T(x)-P'(y)S'(x)h\|}{\|h\|} \geq 0$$

entonces:

$$M \lim_{\|g\| \to 0} \frac{\delta\|g\|}{\|g\|} \geq \lim_{\|h\| \to 0} \frac{\|T(x+h)-T(x)-P'(y)S'(x)h\|}{\|h\|} \geq 0$$

$$\therefore \lim_{\|h\| \to 0} \frac{\|T(x+h)-T(x)-P'(y)S'(x)h\|}{\|h\|} = 0$$

la pregunta ahora es: ¿es lineal?

Sea $\alpha h_1 + \beta h_2$

$P'(y)S'(x)(\alpha h_1+\beta h_2)$ => por la linealidad de $S(x;h)$ respecto a h en x.

$S'(x)(\alpha h_1+\beta h_2)=S'(x)\alpha h_1+S'(x)\beta h_2 = \alpha S'(x)h + \beta S'(x)h_2$

donde $(\alpha S'(x)h_1 + \beta S'(x)h_2)$ en Y

ahora por la linealidad de $\delta P(y;h')$

$P'(y)\ \{S'(x)\ \alpha\ h_1 + S'(x)\beta h_2\} = \alpha P'(y)S'(x)h_1+\beta P'(y)S'(x)h$

$\therefore$ es lineal

ahora: ¿es acotada?

$S'(x)h$ en Y

=> $||P'(y)S'(x)h|| \leq M\ ||S'(x)h||$ ya que $P'(y)$ es acotada

Pero $||S'(x)h|| \leq N\ ||h||$ ya que $S'(x)$ es acotada

$\therefore\ ||P'(y)S'(x)h|| \leq MN\ ||h|| = K\ ||h||$

$\therefore$ es acotada

<u>PROPOSICION 2</u>: Sea T diferenciable de Fréchet en un dominio abierto D. Sea $x \varepsilon D$ y considere que $x+\alpha h$ en D para toda α; $0 \leq \alpha \leq 1$

Entonces: $||T(x+h)-T(x)|| \leq ||h|| \sup\limits_{0<\alpha<1} ||T'(x+\alpha h)||$

<u>Prueba</u>

Sea $y^* \neq 0$ en Y^* alineado con el elemento $\{T(x+h)-T(x)\} \varepsilon Y$

- Ver [Ref 11] pag. 112

Sea la función: $Q(\alpha) = y^* T(x)+ h)$ definida sobre el intervalo $[0,1]$

Entonces:

$$Q'(\alpha) = \lim_{\Delta\alpha \to 0} \frac{y^*[T(x)+\alpha h+\Delta\alpha h)] - y^*[T(x+\alpha h)]}{\Delta\alpha}$$

$$= \lim_{\Delta\alpha \to 0} \frac{y^*[T(x+\alpha h+\Delta\alpha h) - T(x+\alpha h)]}{\Delta\alpha} \quad \{\text{ por la linealidad de } y^*\}$$

$$= y^*\left[\lim_{\Delta\alpha \to 0} \frac{T(x+\alpha h-\Delta\alpha h) - T(x+\alpha h)}{\Delta\alpha}\right] \quad \{\text{ por la continuidad de } y^*\}$$

Pero como suponemos por hipótesis que para T existe la diferencial de Fréchet en D; entonces la diferencial de Gateaux obtenida es igual a la de Fréchet.

$$\therefore \quad Q'(\alpha) = y^*\left[T'(x+\alpha h)\ h\right]$$

Para funciones de variable real, el teorema del valor medio nos dice:

$$Q(1) - Q(0) = Q'(\alpha_0) \qquad 0 < \alpha_0 < 1$$

$$Q(1)-Q(0)=y^*\left[T(x+h)\right] - y^*\left[T(x)\right]=y^*\left[T(x+h)-T(x)\right]=y^*\left[T'(x+\alpha_0)h\right] = Q'(\alpha_0)$$

$$||y^*\left[T(x+h)-T(x)\right]||=||y^*\left[T'(x+\alpha_0 h)h\right]||\leq||y^*|| \ ||T'(x+\alpha_0 h)h||$$

$$\leq||y^*|| \sup_{0<\alpha<1} ||T'(x+\alpha h)h||$$

$$\leq||y^*|| \sup_{0<\alpha<1} ||T'(x+\alpha h)|| \ ||h||$$

$$||y^*||\,.\,||T(x+h)-T(x)||=|y^*\left[T(x+h)-T(x)\right]| \leq ||y^*||\ ||h||\sup_{0<\alpha<1}||T'(x+\alpha h)|| \quad \begin{cases}\text{Por hipó-}\\ \text{tesis de}\\ \text{alineación}\end{cases}$$

$$\therefore\quad ||T(x+h)-T(x)|| \leq ||h|| \sup_{0<\alpha<1} ||T'(x+\alpha h)||$$

Si $T:X \longrightarrow Y$ es diferenciable de Fréchet sobre un dominio abierto $D\subset X$ la derivada T' mapea D a $B(x,y)$ -{ ya que $\delta T(x;h) = T'(x)h$}- y puede a su vez ser Fréchet diferenciable sobre un subconjunto $D_1 \subset D$. En este caso la derivada de Fréchet de T' es llamada la segunda derivada de Fréchet de T; y se denota T''.

PROPOSICION 3: Sea T dos veces diferenciable de Fréchet sobre un dominio abierto D. Sea x en D y suponga que $x+\alpha h$ en D para toda α; $0 \leq \alpha \leq 1$.

Entonces:

$$||T(x+h)-T(x)-T'(x)h|| \leq \frac{1}{2}\ ||h||^2 \sup_{0<\alpha<1} ||T''(x+\alpha h)||$$

Prueba: Sea y^* en Y^* alineado con $\{T(x+h)-T(x)-T'(x)h\}$

Sea $Q(\alpha) = y^*\left[T(x+\alpha h)\right]$ definida sobre el intervalo $\left[0,1\right]$

con: $Q'(\alpha) = y^*\left[T'(x+\alpha h)h\right]$ ... (ver proposición anterior)

$$\Rightarrow Q''(\alpha) = \lim_{\Delta\alpha\to 0}\frac{y^*\left[T'(x+\ h+\Delta\alpha h)h\right]-y^*\left[T'(x+\alpha h)h\right]}{\Delta\alpha} = \lim_{\Delta\alpha\to 0}\frac{y^*\left[T'(x+\alpha h+\Delta\alpha h)h-T'(x+\alpha h)h\right]}{\Delta\alpha}$$

$$= \lim_{\Delta\alpha\to 0}\frac{y^*\left[\{T'(x+\alpha h+\Delta\alpha h)-T'(x+\alpha h)\}\,h\right]}{\Delta h} \text{ ... por linealidad de } T'(x) \text{ respecto a } h$$

$$= y^* \lim_{\Delta\alpha\to 0}\frac{\{T'(x+\alpha h+\Delta\alpha h)-T'(x+\alpha h)\}h}{\Delta\alpha} \text{ ... por continuidad de } y^*$$

como suponemos dos veces diferencial de Fréchet y si es de Fréchet =>es igual a la de Gateaux

$$\Rightarrow Q''(\alpha) = y^* \{T''(x+\alpha h)h\}h$$

Para funciones de variable real el teorema de Taylor nos dice:

$$Q(1) = Q(0) + Q'(0) + \frac{1}{2} Q''(\alpha_0) \quad \ldots \quad 0 < \alpha_0 < 1$$

$$y^* T(x+h) = y^*T(x) + y^*[T'(x)h] + \frac{1}{2} y^* \{T''(x+ h)h\}h$$

$$y^*[T(x+h) - T(x) - T'(x)h] = \frac{1}{2} y^* \{T''(x+\alpha_0 h)h\}h$$

$$||y^*||\,||T(x+h)-T(x)-T'(x)h|| = ||y^*[T(x+h)-T(x)-T'(x)h]|| = ||\frac{1}{2} y^*\{T''(x+\alpha_0 h)h\}h||$$

$$\leq \frac{1}{2}||y^*||\;||T'(x+\alpha_0 h)h||\;||h||$$

$$\leq \frac{1}{2}||y^*||\,||T''(x+\alpha_0 h)||\,||h||\,||h|| =$$

$$\frac{1}{2}||h||^2||y^*||\,||T''(x+\alpha_0 h)||$$

$$\leq \frac{1}{2}||y^*||\,||h||^2 \sup_{0<\alpha<1} ||T''(x+\alpha h)||$$

$$\therefore ||T(x+h) - T(x) - T'(x)h|| \leq \frac{1}{2}\,||h||^2 \sup_{0<\alpha<1} ||T''(x+\alpha h)||$$

2.18 La integral de Riemann-Stieltjes

En esta sección, se tratará el proceso de integración un poco más general que en el concepto de integral de Riemann, y estudiaremos la integral de Riemann Stieltjes, la cual envuelve dos funciones f y α. La representaremos simbólicamente por:

$$\int_a^b f(x)\ d\alpha(x)$$

La integral usual (de Riemann) ocurre como un caso especial cuando $\alpha(x)=x$.

Cuando α tiene derivada contínua, la definición es tal que la integral de Stieltjes

$\int_a^b f(x)\ d\alpha(x)$ se convierte en una integral de Riemann de la forma $\int_a^b f(x)\ \alpha'(x)dx$. Sin embargo, la integral de Stieltjes sigue teniendo sentido cuando α no es diferenciable y aún cuando α es discontínua.

Se trabajará en un intervalo compacto $[a,b]$ y se considerará, a menos que se establezca otra cosa, que las funciones f,g,α, $\beta,\ldots$ son de valor real y acotadas sobre $[a,b]$.

<u>Definición</u>: Sea $P = \{x_0,x_1,\ldots,x_n\}$ una partición de $[a,b]$ y sea t_k un punto en el subintervalo $[x_{k-1}, x_k]$. Una suma de la forma:

$$S(P,f,\alpha) = \sum_{k=1}^{n} f(t_k)\ \Delta\alpha_k$$

donde $\Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1})$

es llamada suma de Riemann-Stieltjes de f con respecto a α. Decimos que f es integrable según Riemann-Stieltjes con respecto a α sobre $[a,b]$ y lo escribiremos como "f en $R(\alpha)$ sobre $[a,b]$" si existe un número A que tenga la siguiente propiedad:

Para toda $\varepsilon>0$, existe una partición P_ε de $[a,b]$ tal que para toda partición P mas fina que P_ε y para cualquier punto t_k en $[x_{k-1}, x_k]$, tenemos que $|S(P,f,\alpha)-A| < \varepsilon$.

Cuando este número A existe, este es determinado en forma única y es denotado por $\int_a^b f\,d\alpha$ ó por $\int_a^b f(x)\,d\alpha(x)$. Y decimos también que la integral de Riemann-Stieltjes $\int_a^b f d\alpha$ existe.

En el caso especial en que $\alpha(x)=x$, escribiremos $S(P,f)$ en lugar de $S(P,f,\alpha)$ y f en R en lugar de f en $R(\alpha)$, la integral es llamada de Riemann y denotada por $\int_a^b f dx$ ó $\int_a^b f(x)dx$.

<u>Definición</u>: Sea P una partición de $[a,b]$ y sea

$$M_k = \sup\{f(x): x \text{ en } [x_{k-1},x_k]\}$$

$$m_k = \inf\{f(x): x \text{ en } [x_{k-1},x_k]\}$$

Los números:

$$U(P,f,\alpha) = \sum_{k=1}^{n} M_k(f)\Delta\alpha_k$$

y

$$L(P,f,\alpha) = \sum_{k=1}^{n} m_k(f)\Delta\alpha_k$$

son llamados respectivamente, las sumas de Stieltjes superior e inferior de f con respecto a α para la partición P.

Definición: Se dice que f satisface la condición de Riemann con respecto a α sobre $[a,b]$ si para toda $\varepsilon>0$, existe una partición P_ε tal que para P más fina que P_ε se implica

$$0 \leq U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$$

Teorema: Si f es continúa sobre $[a,b]$ y si α es de variación acotada sobre $[a,b]$, entonces f está en $R(\alpha)$ sobre $[a,b]$.

Prueba: Es suficiente probar el teorema cuando α es creciente con $\alpha(a) < \alpha(b)$. Ya que $[a,b]$ es compacto y f es contínua sobre este intervalo, esto implica, continuidad uniforme, tal que si $\varepsilon > 0$ es dada, podemos encontrar $\delta>0$ (que dependa únicamente de ε) tal que $|x-y|<\delta$ implica $|f(x)-f(y)|<\varepsilon/A$, donde $A = 2[\alpha(b)-\alpha(a)]$. Si P_ε es una partición con norma $||P_\varepsilon||<\delta$, (La norma de una partición P es la longitud del subintervalo más largo de P y es denotada $||P||$), entonces para P más fina que P_ε debemos tener:

$$M_k(f) - m_k(f) \leq \varepsilon/A$$

ya que $M_k(f) - m_k(f) = \sup\{f(x) - f(y) : x, y \text{ en } [x_{k-1},x_k]\}$. Multiplicando la desigualdad por $\Delta\alpha_k$ y sumando tenemos:

$$U(P,f,\alpha)-L(P,f,\alpha) \leq \frac{\varepsilon}{A} \sum_{k=1}^{n} \Delta\alpha_k = \frac{\varepsilon}{A} \cdot \frac{A}{2} = \frac{\varepsilon}{2} < \varepsilon$$

por lo que se cumple la condición de Riemann. Por lo tanto f está en $R(\alpha)$ sobre $[a,b]$.

Para el caso especial en el que $\alpha(x)=x$, el teorema anterior y teorema de la integración por partes, generan el siguiente corolario:

<u>Teorema</u>: Cada una de las siguientes condiciones es suficiente para la existencia de la integral de Riemann

$$\int_a^b f(x)dx$$

a) f es contínua sobre $[a,b]$

b) f es de variación acotada sobre $[a,b]$

<u>Teorema</u>: Sea α de variación acotada sobre $[a,b]$ y suponga que f está en $R(\alpha)$ sobre $[a,b]$. Se define F por la ecuación

$$F(x) = \int_a^x f d\alpha \qquad \text{si } x \text{ está en } [a,b]$$

entonces tenemos:

i) F es de variación acotada sobre $[a,b]$

ii) Todo punto de continuidad de α es también punto de continuidad de F.

iii) Si α es creciente sobre $[a,b]$, la derivada $F'(x)$ existe en cada punto x en (a,b) donde $\alpha'(x)$ existe y donde f es contínua. Para cada x, tenemos:

$$F'(x) = f(x)\alpha'(x)$$

CRITERIO DE LEBESGUE PARA LA EXISTENCIA DE LA INTEGRAL DE RIEMANN.

Definición: Una colección F de conjuntos se dice que es una cubierta de un conjunto dado S, si $S \subset \bigcup_{A \text{ en } F} A$. La colección F se dice también que cubre S.

Si F es una colección de conjuntos abiertos, entonces F es llamada una cubierta abierta de S.

Definición: Un conjunto S de números reales, se dice que tiene medida cero si, para toda $\varepsilon>0$, existe una cubierta contable de S, consistente de conjuntos abiertos, y la suma de las longitudes de estos es menor que ε.

Si los intervalos son denotados por (a_k,b_k), la definición requiere que: $S \subset \bigcup_k (a_k,b_k)$

y $\sum_k (b_k-a_k) < \varepsilon$

Teorema: Sea F una colección contable de conjuntos en R.

$$F = \{F_1, F_2, \ldots \}$$

cada uno de los cuales tiene medida cero. Entonces su unión

$$S = \bigcup_{k=1}^{\infty} F_k$$

también tiene medida cero.

Prueba: Dado $\varepsilon>0$, existe una cubierta contable de F_k, de intervalos abiertos cuya suma de longitudes es menor que $\varepsilon/2^k$. La unión de todas estas cubiertas es en si misma una cubierta contable de S, de conjuntos abiertos y la suma de todas las longitudes es menor que

$$\sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

Ejemplo: Ya que el conjunto consistente de un solo punto tiene medida cero, se sigue que todo subconjunto contable de R, tiene medida cero.

Teorema (Criterio de Lebesgue para la integral de Riemann)

Sea f definida y acotada sobre $[a,b]$ y sea D el conjunto de discontinuidades de f sobre $[a,b]$. Entonces f tiene integral de Riemann sobre $[a,b]$ si y sólo si D tiene medida cero.

Definición: Una propiedad se dice que se sostiene casi en todas partes (almost everywhere) en un subconjunto S de R^1, si se sostiene en todo S, excepto en un conjunto de medida cero.

3. EL PRINCIPIO MAXIMO DE PONTRYAGIN (Prueba de Luenberger).

3.1 Fundamentos

El principio máximo de Pontryagin proporciona un conjunto de condiciones necesarias para problemas de control, en los cuales el control u(t) está restringido a un conjunto dado.
En este capítulo, desarrollaremos una forma del principio máximo, desde un punto de vista abstracto, explotando la definición básica de un operador adjunto.

Sean X, U espacios lineales normados y $g[x,u]$ un funcional sobre X x U ; y consideremos una restricción de la forma:

$$A\,[x,\ u] = \Theta \qquad (1)$$

Donde A es un mapeo de X x U a X. La transformación A, describe un sistema de ecuaciones, que pueden ser un conjunto de ecuaciones diferenciales, ecuaciones integrales, ecuaciones diferenciales parciales, etc.

Suponemos que A define una única función implícita x(u). También suponemos que A y g son diferenciables Fréchet con respecto a x, y que las derivadas $A_x[x,u]$, $g_x[x,u]$ son contínuas sobre X x U.

Finalmente suponemos que la función implícita x(u) satisface una condición de Lipschitz de la forma:

$$||x(u) - x(v)|| \leq K||u-v|| \qquad (2)$$

El problema de control, es encontrar (x,u) que minimice $J = g\,[x,u]$, que satisfaga $A[x,u]=\Theta$ y u en Ω, donde Ω es un subconjunto de U.

Ya que x es determinada en forma única de u, el funcional objetivo J puede ser considerado como dependiente únicamente de u, ya que $J(u) = g[x(u),u]$.

Introducimos ahora el funcional de Lagrange.

Para x en X, u en U, λ^* en X^*, definimos:

$$L[x,u,\lambda^*] = \lambda^* A[x,u] + g[x,u] \tag{3}$$

La siguiente proposición puede ser vista como la base para el establecimiento de un número de condiciones necesarias para problemas de control conectados con varios tipos de sistemas.

<u>Proposición 1</u>. Para cualquier u en Ω, donde Ω es cualquier subconjunto previamente establecido de U, sea λ^* una solución de la ecuación

$$\lambda^* A_x[x(u),\ u] + g_x[x(u),u] = \Theta \tag{4}$$

entonces para v en Ω, $J(u) = g[x(u),\ u]$

$$J(u)-J(v) = L[x(u),u,\lambda^*] - L[x(u),v,\lambda^*] + \delta(||u-v||)$$

Donde $\delta(||u-v||)$ representa una función tal que

$$\lim_{||u-v||\to 0} \frac{\delta(||u-v||)}{||u-v||} = 0$$

Prueba:

Por definición:

$$J(u) - J(v) = g[x(u),u] - g[x(v),v]$$
$$= g[x(u),u] - g[x(u),v] + g[x(u)v] - g[x(v),v]$$

Por definición de Diferencial de Gateaux

$$\lim_{\alpha \to 0^+} \frac{T(x+\alpha h)-T(x)}{\alpha} = \delta T(x;h)$$

y si es lineal

$$\lim_{\alpha \to 0^+} \frac{T(x+\alpha h)-T(x)}{\alpha||h||} = \delta T(x;h/||h||) \quad ; \quad \text{o sea}$$

$$\lim_{\alpha \to 0^+} \frac{T(x+\alpha h)-T(x)- ||\alpha h|| \quad \delta T(x;h/||h||)}{\alpha||h||} = 0$$

y por la misma linealidad, tenemos

$$\lim_{\alpha \to 0^+} \frac{T(x+\alpha h)-T(x)- \delta T(x;\alpha h)}{\alpha||h||} = 0$$

de la definición de "σ"

$$T(x+\alpha h) - T(x) - \delta T(x;\alpha h) = \sigma(||\alpha h||)$$

o sea:

$$T(x+\alpha h) - T(x) = \delta T(x;\alpha h) + \sigma(||\alpha h||)$$

recordemos que $\delta T(x;\alpha h) = T'(x).\alpha h$, donde $T'(x)$ es la derivada de Fréchet de la transformación T.

Por lo tanto, para nuestro caso tenemos:

$$g[x(u),\bar{v}] - g[x(v),\bar{v}] = g_x[x(v),\bar{v}]\ [x(u)-x(v)] + \sigma(||x(u)-x(v)||)$$

Por lo tanto:

$$\begin{aligned} J(u)-J(v) = {} & g[x(u),\bar{u}] - g[x(u),\bar{v}] + g_x[x(v),\bar{v}]\,[x(u)-x(v)] + \\ & \sigma(||x(u) - x(v)||) \end{aligned}$$

$$\begin{aligned} J(u)-J(v) = {} & g[x(u),\bar{u}] - g[x(u),\bar{v}] + g_x[x(v),\bar{v}]\,[x(u)-x(v)] + \\ & g_x[x(u),\bar{u}]\,[x(u)-x(v)] - g_x[x(u),\bar{u}]\,[x(u)-x(v)] + \\ & \sigma(||x(u) - x(v)||) \end{aligned}$$

$$\begin{aligned} J(u)-J(v) = {} & g[x(u),\bar{u}] - g[x(u),\bar{v}] + g[x(u),\bar{u}]\,[x(u)-x(v)] + \\ & (g_x[x(v),\bar{v}] - g_x[x(u),\bar{u}]\,[x(u)-x(v)] + \sigma(||x(u)-x(v)||) \end{aligned}$$

Demostraremos ahora que:

$$(g_x[x(v),\bar{v}] - g_x[x(u),\bar{u}])\,[x(u)-x(v)] + \sigma(||x(u)-x(v)||) = \sigma(||u-v||)$$

Para ello tenemos que:

$$||(g_x[x(v),\bar{v}] - g_x[x(u),\bar{u}])\ \frac{[x(u)-x(v)]}{||u-v||}|| \leq$$

$$||\ g_x[x(v),\bar{v}] - g_x[x(u),\bar{u}]\,||\ \frac{||x(u)-x(v)||}{||u-v||}$$

$$\leq K\ ||g_x[x(v),\bar{v}] - g_x[x(u),\bar{u}]||\ \text{(por Lipschitz)}$$

Por lo tanto:

$$||(g_x[x(v),\bar{v}] - [g_x\ x(u),\bar{u}])\frac{[x(u)-x(v)]}{||u-v||}\ || < \varepsilon \qquad \ldots.(a)$$

Siempre que:

$$\|g_x[x(v), \bar{v}] - g_x[x(u), \bar{u}]\| \leq \varepsilon/K \qquad \ldots(b)$$

Dada la continuidad de g_x, existe $\delta>0$ tal que (b) se sostiene, siempre que:

$$\|[x(v), \bar{v}] - [x(u), \bar{u}]\| < \delta$$

donde

$$\|[x(v), \bar{v}] - [x(u), \bar{u}]\| = \text{Máx}\{\|x(v)-x(u)\|, \|u-v\|\}$$

o sea, que dada cualquier $\varepsilon>0$, existe $\delta>0$ tal que (a) se cumple siempre que:

$$\|u - v\| < \delta$$

Por lo tanto:

$$(g_x[x(v), \bar{v}] - g_x[x(u), \bar{u}])[x(u)-x(v)] = \sigma(\|u-v\|)$$

Por otro lado, sea

$$\sigma(\|x(u)-x(v)\|) = f[x(u), x(v)]$$

Aplicando la definición de límite:

$$\frac{\|f[x(u), x(v)]\|}{\|x(u)-x(v)\|} < \varepsilon/K, \text{ siempre que } \|x(u)-x(v)\| < \delta$$

o sea

$$\|f[x(u), x(v)]\| < \varepsilon/K \, \|x(u)-x(v)\| < \varepsilon \, \|u-v\| \qquad \text{(por Lipschitz)}$$

Pero

$$||x(u)-x(v)|| \leq K||u-v|| < \delta, \text{siempre que } ||u-v|| < \delta/K$$

Hemos demostrado que dada cualquier $\varepsilon > 0$, existe $\delta' = \delta/K > 0$ tal que:

$$\frac{||f[x(u),x(v)]||}{||u - v||} < \varepsilon, \text{ siempre que } ||u-v|| < \delta'$$

o sea $\quad f[x(u),\ x(v)] \ = \ \sigma(||u-v||)$

De lo que se sigue que:

$$J(u)-J(v) = g[x(u),\bar{u}] - g[x(v),\bar{v}] + g_x[x(u),\bar{u}]\,[x(u)-x(v)] + \sigma(||u-v||) \qquad \ldots(c)$$

De manera semejante

$$\Theta = A[x(u),\bar{u}] \ - \ A[x(v),\bar{v}]$$

$$= A[x(u),\bar{u}] \ - \ A[x(u),\bar{v}] + [A_x\ x(u),\bar{u}]\,[x(u)-x(v)]$$

$$+ \sigma(||u-v||)$$

o sea

$$0 = \lambda^* \ A[x(u),\bar{u}] - \lambda^* A[x(u),\bar{v}] + \lambda^* A_x[x(u),\bar{u}]\,[x(u)-x(v)]$$

$$+ \sigma(||u-v||) \qquad \ldots(d)$$

Sumando (c) y (d) tenemos:

$$J(u) - J(v) = \lambda^* A[x(u),\bar{u}] - \lambda^* A[x(u),\bar{v}] +$$

$$+ \lambda^* A_x[x(u),\bar{u}][x(u) - x(v)]$$

$$+ g[x(u),\bar{u}] - g[x(u),\bar{v}] + g_x[x(u),\bar{u}][x(u)-x(v)]$$

$$+ \delta(||u-v||)$$

como por hipótesis:

$$\lambda^* A_x[x(u),\bar{u}] + g_x[x(u),\bar{u}] = \theta$$

Entonces:

$$J(u) - J(v) = L[x(u),u,\lambda^*] - L[x(u),v,\lambda^*] + \delta(||u-v||)$$

La importancia del resultado anterior, es que proporciona una manera, de primer orden, para determinar el cambio en J, debido a un cambio en u, sin necesidad de volver a calcular la función implícita x.

Lema 1. (Lema de Bellman-Gronwall).

Sea la constante $c_1 \geq 0$. Sean u(t) y k(t) funciones contínuas por partes (o sea, funciones cuyo conjunto de puntos de discontinuidad, es finito y además existe el límite por la derecha y por la izquierda en esos puntos) sobre $[t_0,t_1]$.
Sea $k(t) \geq 0$ para t en $[t_0, t_1]$.

$$\text{Si } u(t) \leq c_1 + \int_{t_0}^{t} K(t')\, u(t')dt' \qquad \ldots(1)$$

para toda t en $[t_0, t_1]$, entonces:

$$u(t) \leq c_1 \exp \{\int_{t_0}^{t} K(t')dt'\} \text{ para toda t en } [t_0, t_1] \qquad \ldots(2)$$

<u>Prueba</u>: Sea U(t) el lado derecho de (1), entonces $u(t) \leq U(t)$.

Multiplicando ambos lados de (1) por la función no negativa:

$$K(t) \exp\left\{ - \int_{t_0}^{t} K(t')dt' \right\}$$

tenemos:

$$u(t)K(t) \exp \{- \int_{t_0}^{t} K(t')dt'\} \leq U(t)K(t) \exp \{- \int_{t_0}^{t} K(t')dt'\}$$

entonces:

$$u(t)K(t) \exp \{- \int_{t_0}^{t} K(t')dt'\} - U(t)K(t) \exp \{- \int_{t_0}^{t} K(t')dt'\} \leq 0$$

Analicemos como es $\dot{U}(t)$

Sea $\alpha(t) = t$ y $f(t) = K(t)u(t)$

ya que α es creciente y por tanto de variación acotada sobre $[t_0, t_1]$, y f es integrable pues está definida y acotada sobre $[t_0, t_1]$ y además el número de discontinuidades de f sobre $[t_0, t_1]$ es finito o sea de medida cero. Entonces si

$$F(t) = \int_{t_0}^{t} f(\zeta)\, d\alpha(\zeta)$$

su derivada existe y es igual a:

$$F'(t) = f(t)\ \alpha'(t) = K(t)u(t)$$

Para las t donde $K(t)$, $u(t)$ son contínuas.

Por lo tanto la derivada de $U = C_1 + \int_{t_0}^{t} u(t')k(t')dt'$ existe y es igual a:

$\dot{U}(t)=u(t)K(t)$ para las t donde $u(t)$ y $K(t)$ son contínuas.

por tanto, debido a que

$$\frac{d}{dt}\{U(t)\exp\{-\int_{t_0}^{t} K(t')dt'\}\}=\dot{U}(t)\exp\{-\int_{t_0}^{t} K(t')dt'\}-U(t)k(t)\exp\{\int_{t_0}^{t} K(t')dt'\}$$

Por lo tanto:

$$\frac{d}{dt}\{U(t)\ \exp\ \{-\int_{t_0}^{t} K(t')dt'\}\} \leq 0 \qquad \text{a.e.}$$

Ya que esta última expresión se cumple para casi toda t en $[t_0, t_1]$, (o sea es a.e. sobre $[t_0, t_1]$). Entonces, integrándola entre t_0 y t, y notando que en $t = t_0$, $U(t_0) = C_1$ y la exponencial toma valor de cero, obtenemos para toda t en $[t_0, t_1]$

$$\int_{t_0}^{t} \frac{d}{d\zeta}\{U(\zeta)\exp\ \{-\int_{t_0}^{\zeta} K(t')\}\}\ d\ \zeta \leq 0$$

o sea

$$U(\zeta)\exp\ \{-\int_{t_0}^{\zeta} K(t')dt'\}\Big]_{t_0}^{t} \leq 0 \qquad \text{para toda } t \text{ en } [t_0,t_1]$$

Por lo tanto:

$$U(t) \exp\{-\int_{t_0}^{t} K(t')dt'\} - C_1 \leq 0$$

$$U(t) \exp\{-\int_{t_0}^{t} K(t')dt'\} \leq C_1$$

$$U(t) \leq C_1 \exp\{\int_{t_0}^{t} K(t')dt'\}$$

Por (1) $u(t) \leq U(t)$, por lo tanto

$$u(t) \leq C_1 \exp\{\int_{t_0}^{t} K(t')dt'\} \qquad \text{para toda } t \text{ en } [t_0, t_1]$$

<u>Lema 2</u>: La función λ en NBV^n $[t_0, t_1]$ que satisface:

$$\dot{\lambda}(t) = -f_x^T[x_0(t), u_0(t), t]\ \lambda(t) - l_x^T\ [x_0(t), u_0(t), t]$$

$$\lambda(t_1) = 0$$

Satisface la ecuación:

$$\lambda A_x[x_0, u_0] + g_x[x_0, u_0] = \Theta \qquad \text{en } NBV^n\ [a, b]$$

donde

$$A[x, u] = x(t) - x(t_0) - \int_{t_0}^{t} f(x, u, t')dt'$$

$$g[x, u] = \int_{t_0}^{t_1} l(x, u, t')dt'$$

y suponemos que las funciones f,l son continuamente diferenciables con respecto a x. Además (x,u) está en C^n $[t_0, t_1]$ x U, donde U es el espacio de las funciones contínuas por partes de la derecha. (O sea que el valor de la función u en U en un punto ζ de discontinuidad, es igual al límite de u(t) cuando t tiende a ζ por la derecha).

$$\text{i.e.} \quad u(\zeta) = \lim_{t \to \zeta^+} u(t)$$

Prueba: Se necesita demostrar que para cada h en $C^n[t_0, t_1]$

$$\langle \lambda A_x(x_0,u_0), h \rangle + \langle g_x(x_0,u_0), h \rangle = 0 \tag{1}$$

o sea

$$\langle \lambda, A_x(x_0,u_0)h \rangle + \langle g_x(x_0,u_0)h \rangle = 0 \tag{2}$$

Ya que

$$A_x(x,u) = 1 - \int_{t_0}^{t} f_x(x,u,\zeta)d\zeta$$

$$g_x(x,u) = \int_{t_0}^{t_1} l_x(x,u,\zeta)d\zeta$$

y por la representación de Riez, tenemos que para toda h en $C^n[t_0,t_1]$, (2) puede expresarse como:

$$\int_{t_0}^{t_1} \dot{\lambda}(h) \left[h(t) - \int_{t_0}^{t} f_x(x_0,u_0,\zeta)h(\zeta)d\zeta\right] dt \;+$$

$$\int_{t_0}^{t_1} l_x(x_0,u_0,t)h(t)dt = 0$$

el lado derecho de la expresión se puede desarrollar como:

$$\int_{t_0}^{t_1} \dot{\lambda}(t)h(t)dt - \int_{t_0}^{t_1} d\lambda(t) \int_{t_0}^{t} f_x(x_0,u_0,\zeta)h(\zeta)d\zeta \;+$$

$$\int_{t_0}^{t_1} l_x(x_0,u_0,t)h(t)dt$$

Integrando el segundo término por partes y recordando que $\lambda(t_1) = 0$, llegamos a:

$$\int_{t_0}^{t_1} \dot{\lambda}(t)h(t)dt \;+\; \int_{t_0}^{t_1} \lambda(t)\; f_x(x_0,u_0,t)h(t)dt \;+$$

$$\int_{t_0}^{t_1} l_x(x_0,u_0,t)h(t)dt$$

o sea

$$\int_{t_0}^{t_1} \left[\dot{\lambda}(x) + \lambda(t)\; f_x(x_0,u_0,t) + l_x(x_0,u_0,t)\right] h(t)dt$$

pero

$$\dot{\lambda}(t) = -f_x(x_0,u_0,t)\lambda(t) - l_x(x_0,u_0,t)$$

$$\therefore \quad \int_{t_0}^{t_1} \left[\dot{\lambda}(t)+\lambda(t)f_x(x_0,u_0,t) + l_x(x_0,u_0,t)\right] h(t)dt = 0$$

3.2 Principio máximo de Pontryagin

Aplicaremos ahora estos resultados al problema:

$$\text{Min} \quad g(x,u) = \int_{t_0}^{t_1} l(x(t),u(t),t)dt$$

$$\dot{x}(t) = f(x(t),u(t),t)$$

$$x(t_0) = a$$

$u(t)$ en Ω subconjunto conocido de R^m donde f y l tienen derivadas parciales contínuas con respecto a x. Además f satisface la condición de Lipschitz con respecto a x y u como sigue:

$$||f(x,u) - f(y,v)|| \leq M\left[||x-y|| + ||u-v||\right] \qquad \text{para } M > 0$$

y $||.||$ la norma euclidiana.

El espacio de controles admisibles U es el espacio de funciones contínuas por partes por la derecha en el intervalo $[t_0, t_1]$ y con la norma de L_1 o sea:

$$||u|| = \int_{t_0}^{t_1} \{|u_1(t)| + \ldots + |u_m(t)|\}\, dt$$

Teorema

Sean x_0, u_0 óptimos para el problema de minimizar

$$J(u) = \int_{t_0}^{t_1} l(x,u)dt$$

sujeto a: $\dot{x}(t) = f(x,u)$; $x(t_0) = a$; $u(t)$ en Ω

Sea λ la solución a la ecuación:

$$-\dot{\lambda}(t) = f_x \lambda(t) + l_x \quad ; \quad \lambda(t_1) = 0 \qquad \ldots(1)$$

donde las derivadas parciales son evaluadas a lo largo de la trayectoria óptima y defina la función de Hamilton por:

$$H(x,u,\lambda,t) = \lambda(t) f(x,u) + l(x,u) \qquad \ldots(2)$$

Entonces para toda t en $[t_0, t_1]$

$$H[x_0(t), u_0(t), \lambda(t)] \leq H[x_0(t), u, \lambda(t)] \qquad \ldots(3)$$

para toda u en Ω.

<u>Prueba</u>: Por simplicidad sea $m = 1$; i.e. los controles son funciones escalares. Tomemos $X = C^n [t_0, t_1]$ y para U tomemos el espacio de las funciones contínuas por partes por la izquierda con la norma L_1 y definamos:

$$A[x,u] = x(t) - x(t_0) - \int_{t_0}^{t} f(x(\zeta), u(\zeta), \zeta) d\zeta$$

$$g[x,u] = \int_{t_0}^{t_1} l(x(t), u(t), t) d$$

Si x, $x+\delta x$ corresponden a u, $u + \delta u$ respectivamente en $\{(x,u) \; ; \; A[x,u] = \Theta\}$ tenemos:

$$||\delta x(t)||_{R^n} = ||\int_{t_0}^{t} \{f\left[x(\zeta)+\delta x(\zeta), u(\zeta)+ u(\zeta), \zeta\right] -$$

$$f\left[x(\zeta), u(\zeta), \zeta\right]\} \, d\zeta||$$

$$\leq \int_{t_0}^{t} ||f\left[x + \delta x, u+\delta u, \zeta\right] - f\left[x, u, \zeta\right]|| \, d\zeta$$

$$\leq \int_{t_0}^{t} M\{||\delta x(\zeta)||_{R^n} + |\delta u(\zeta)|\} \, d\zeta \quad \left\{\begin{array}{l}\text{por la con-}\\ \text{dición de}\\ \text{Lipschitz}\end{array}\right.$$

$$\leq \int_{t_0}^{t} M \, ||\delta x(\zeta)||_{R^n} d\zeta + \int_{t_0}^{t_1} M \, |\delta u(\zeta)| \, d\zeta \, ,$$

Aplicando el lema de Bellman-Gronwall

$$||\delta x(t)||_{R^n} \leq M \exp \{M(t-t_0)\} \int_{t_0}^{t_1} |\delta u(\zeta)| \, d\zeta$$

$$\leq M \exp \{M(t_1-t_0)\} \int_{t_0}^{t_1} |\delta u(\zeta)| \, d\zeta$$

por lo tanto:

$$||\delta x|| \leq K \, ||\delta u||$$

y la transformación $A[x, u]$ satisface la condición de Lipschitz requerida para la proposición 1 de este capítulo.

ya que $\quad x(u) = x \quad$ y $\quad x(u+\delta u) = x + \delta x$

por lo tanto:

$$||x(u+ \delta u) - x(u)|| = ||(x+\delta x) - x|| = ||\delta x|| \leq K||\delta u|| = K||(u+\delta u)-u||$$

Además por el lema 2 demostrado anteriormente, λ satisface la ecuación adjunta:

$$\lambda A_x(x,u) + g_x(x,u) = \Theta$$

El funcional:

$$\int_{t_0}^{t_1} H(x,u,\lambda)dt$$

es idéntico al funcional de Lagrange:

$$L[x,u,\lambda] = \lambda A[x,u] + g[x,u]$$

excepto por el término $\int_{t_0}^{t_1} \lambda x(t) \quad \dot{\lambda}(t)dt$

lo que puede comprobarse ya que:

$$H(x,u,\lambda,t) = \lambda(t)\ f(x,u) + l(x,u)$$

Entonces:

$$\int_{t_0}^{t_1} \left[\lambda(t)f(x,u)+l(x,u)\right] dt = \int_{t_0}^{t_1} \lambda(t)f(x,u)dt + \int_{t_0}^{t_1} l(x,u)dt =$$

$$\int_{t_0}^{t_1} \lambda(t)f(x,u)dt + g(x,u)$$

Integrando por partes el primer término:

$$u = \lambda(t) \quad ; \quad dv = f(x,u)\ dt$$

$$du = \dot{\lambda}(t)dt \qquad v = \int_{t_0}^{t} f(x,u,\zeta)d\zeta$$

$$uv = \lambda(t)\int_{t_0}^{t} f(x,u,\zeta)d\zeta\Big]_{t_0}^{t_1} = 0 \quad \text{ya que } \lambda(t_1) = 0$$

$$-\int_{t_0}^{t_1}\left[\dot{\lambda}(t)\int_{t_0}^{t} f(x,u,\zeta)d\zeta\right] dt$$

$$\therefore \int_{t_0}^{t_1} H(x,u,\lambda) = -\int_{t_0}^{t_1}\left[\dot{\lambda}(t)\int_{t_0}^{t} f(x,u,\zeta)d\zeta\right]dt+g(x,u)$$

Y por otro lado:

$$L[x,u,\lambda] = \lambda A[x,u] + g[x,u]$$

Por la representación de Riez:

$$\lambda A[x,u] = \int_{t_0}^{t_1}\dot{\lambda}(t)\left[x(t) - x(t_0) - \int_{t_0}^{t} f(x,u,\zeta)d\zeta\right] dt =$$

$$= \int_{t_0}^{t}\dot{\lambda}(t)x(t)dt - \lambda(t)x(t_0)\Big]_{t_0}^{t_1}$$

$$= \int_{t_0}^{t_1}\dot{\lambda}(t)x(t)dt - \lambda(t)x(t_0)\Big]_{t_0}^{t_1} - \int_{t_0}^{t_1}\left[\dot{\lambda}(t)\int_{t_2}^{t_1}f(x,u,\zeta)d\zeta\right]dt$$

en el segundo término $\lambda(t_1) = 0$ por hipótesis y $\lambda(t_0) = 0$ por estar λ en NBV

$$\therefore \lambda A[x,u] = \int_{t_0}^{t_1}\dot{\lambda}(t)x(t)dt - \int_{t_0}^{t_1}\left[\dot{\lambda}(t)\int_{t_0}^{t} f(x,u,\zeta)d\zeta\right]dt$$

Por lo tanto

$$L[x,u,\lambda]= \int_{t_0}^{t_1} \dot{\lambda}(t)x(t)dt - \int_{t_0}^{t_1} [\lambda(t) \int_{t_0}^{t} f(x,u,\zeta)d\zeta]dt +g(x,u)$$

que es lo que quería comprobarse. Este término para el caso no es importante ya que no depende de u. Por lo que la proposición 1, demostrada anteriormente nos da:

$$J(u_0)-J(u) = \int_{t_2}^{t_1} [H(x_0,u_0,\lambda) - H(x_0,u,\lambda)]dt +\delta||u-u_0||....(4)$$

Mostraremos ahora que (4) implica (3). Supongamos lo contrario, que existe $\bar{t}$ en $[t_0, t_1]$ y $\bar{u}$ en Ω tal que

$$H(x_0(\bar{t}),u_0(\bar{t}),\lambda(\bar{t})) > H(x_0(\bar{t}), \bar{u}, \lambda(\bar{t}))$$

En vista de la continuidad por partes por la derecha de u y la continuidad por la derecha de f, l,λ y la continuidad de x, se sigue que existe un intervalo $[t',t'']$ conteniendo a $\bar{t}$ y una $\varepsilon>0$ tal que:

$$H[x_0(t),u_0(t),\lambda(t)] = H[x_0(t),\bar{u},\lambda(t)] > \varepsilon$$

para toda t en $[t', t'']$.

Sea ahora u(t) la función contínua por partes por la derecha igual a $u_0(t)$ si t no está en $[t', t'']$ e igual a $\bar{u}$ en $[t',t'']$ de (4) tenemos:

$$J(u_0) - J(u) = \int_{t'}^{t''} [H(x_0,u_0,\lambda)-H(x_0,\bar{u},\lambda)]dt + \delta(||u-u_0||) >$$

$$> \varepsilon (t''-t') + \delta (||u-u_0||) \qquad ...(5)$$

pero $||u-u_0|| = \int_{t'}^{t''} \left|\bar{u}(t) - u_0(t)\right| dt$

Sea $K = \sup \{ |\bar{u}(t)-u_0(t)| \quad ; \quad t' \leq t \leq t'' \}$

entonces $||u - u_0|| \leq K(t''-t')$

Sea $g = \delta(||u-u_0||)$; por definición de "δ" pequeña, esto implica que:

$$\frac{||g||}{||u - u_0||} < \varepsilon/K \quad \text{siempre que } ||u-u_0|| < \delta$$

O sea:

$||g|| < \varepsilon/K \quad ||u-u_0|| < \varepsilon(t''-t')$

siempre que: $||u-u_0|| < K(t''-t') < \delta$

Con esto hemos demostrado que dada cualquier $\varepsilon>0$, existe $\delta' = \delta/K$ tal que:

$$\frac{||g||}{t'' - t'} < \varepsilon \quad \text{siempre que} \quad t''-t' < \delta'$$

o sea que $g = \delta(t'' = t')$, lo que implica que

$$J(u_0)-J(u) > \varepsilon \; (t''-t') + \delta(t''-t')$$

seleccionando $[t' t'']$ suficiente pequeño tendremos:

$$J(u_0) - J(u) > 0$$

lo que contradice la optimalidad de u_o.

En forma breve, este resultado nos dice que si una función de control minimiza el funcional objetivo, sus valores en cada instante deben minimizar también la función de Hamilton. La ecuación adjunta planteada por Pontryagin es un poco diferente que la nuestra, dando por resultado que la función de Hamilton deba ser maximizada, de ahí el nombre de Principio Máximo de Pontryagin en lugar de Principio Mínimo.

4. PRINCIPIO MAXIMO DE PONTRYAGIN. (Prueba de Halkin).

4.1 Polisistemas dinámicos. Evolución de los sistemas.

Consideremos un sistema cuyo estado es descrito por un vector X del espacio R^n y cuya evolución está dada por la ecuación diferencial

$$\dot{x} = f(x,u,t) \tag{1}$$

donde u es un vector de control del espacio R^m y $f(x,u,t)$ es una función vectorial n-dimensional de x,u, y t.

Supondremos que la función $f(x,u,t)$ está definida para toda x en R^n, toda u en R^m y toda t en $[0,1]$ que es dos veces continuamente diferenciable con respecto a X, continua con respecto a u y continúa por partes con respecto a t.

4.2 Funciones de control.

Sea Ω un subconjunto acotado de R^m. Este conjunto Ω es llamado el conjunto de los vectores de control admisible. Sea F la clase de todas las funciones continuas por partes de $[0,1]$ a Ω. Esto implica en particular que toda función en F es acotada. La clase F es llamada la clase de las funciones admisibles de control. Si μ y υ están en F, ello significa que el valor de ellas en el punto t será u(t) y v(t), respectivamente.

Si μ pertenece a F entonces el conjunto $\Xi(\mu)$ será definido como el conjunto de puntos en (0,1) en los cuales la función μ y la función $f(x,u,t)$ son contínuas con respecto a t.

Ya que por definición, toda función μ en la clase F y la función f(x,u,t) son contínuas con respecto a t, se sigue que existe únicamente un número finito de puntos en el intervalo $[0,1]$ que no pertenecen al conjunto $\Xi(\mu)$ y además que el conjunto $\Xi(\mu)$ es abierto.

4.3 Trayectorias

Si μ es una función de control en la clase F, denotaremos por $X(t;\mu)$ a una función contínua en t, diferenciable sobre $\Xi(\mu)$ y tal que:

(i) $\dot{x}(t,\mu) = f(x(t;\mu), u(t), t)$ para toda t en $\Xi(\mu)$.

(ii) $x(0;\mu) = \theta$

De la teoría de las ecuaciones diferenciales ordinarias [Ref. 4,5] se sabe que para toda μ en F tenemos una y sólo una de las siguientes posibilidades.

(i) La función $x(t;\mu)$ existe, es única y es acotada sobre el intervalo $[0,1]$.

(ii) Existe una ζ en $[0,1]$ tal que, la función $x(t;\mu)$ existe, es única y es acotada para toda t en $[\sigma,\zeta^*]$, con $\zeta^*<\zeta$ y es no acotada cuando t tiende a ζ. Cuando este es el caso, ζ es llamada un tipo de escape finito de la variable de estado X para la función de control μ.

Muchos desarrollos en la teoría de control óptimo son simplificados grandemente si suponemos a priori que la posibilidad de un tiempo de escape finito de la variable de estado X es rechazado para cualquier función de control μ en la clase F. Se asumirá que existe una constante $M < \infty$ tal que:

$$|f(x,u,t)| \leq M(||x||+1)$$

Para toda X en R^n, toda u en Ω y toda t en $[0,1]$.

Con la ayuda de esta suposición, es fácil probar que para toda función de control μ en la clase F, la función $x(t;\mu)$ existe, es única y acotada sobre el intervalo $[0,1]$ (*).

4.4 <u>Problema de optimización</u>. Restricciones terminales y criterio de representación.

Hemos establecido que el vector de estado X es cero al tiempo t=0. Sea r un entero tal que $0 \leq r \leq n-1$. Sea $\{S_1, S_2, ..., S_r\}$ un conjunto de números reales. Estableceremos que este conjunto representa el valor de los primeros r componentes del vector de estado X al tiempo final (t=1).

El problema de optimización se establece formalmente como sigue:

Se da el conjunto:

$$S = \{x : x \varepsilon R^n ; X_i = S_i \text{ para } i=1,2,...,r\}$$

y deseamos encontrar la función de control υ en la clase F tal que:

(i) $x(1;\upsilon)$ en S.

(ii) para toda μ en F tal que

(*) ver [ref. 10] págs. 247-256

$x(1;\mu)$ en S

se cumple la relación

$$x_n(1;\mu) \leq x_n(1;\upsilon)$$

La función de control que satisface las condiciones anteriores es llamada una función de control óptima, y la trayectoria correspondiente $x(t;\upsilon)$ es llamada la trayectoria óptima.

Podemos entonces resumir el problema de control planteado anteriormente como:

$$\max x_n(1)$$

$$\dot{x}(t) = f[x(t),\ u(t),t] \qquad t\varepsilon[0,1]$$

$$x(0) = 0$$

$$x_k(1) = s_k \qquad k=1,\ldots,r(r<n)$$

$$u\varepsilon F$$

donde $f:R^n x R^m x[0,1]\rightarrow R^n$ f continuamente diferenciable de segundo orden con respecto a x, continua con respecto a u y continua a tramos con respecto a t.

F la clase de controles admisibles consistente de la función $u(\cdot)$ en $[0,1]$ que son continuas a tramos y satisfacen $u(t)\varepsilon\Omega$, $t\varepsilon[0,1]$ donde $\Omega c R^m$ es acotado.

$x(t) = [x_1(t), \ldots, x_n(t)]$ vector de estados

$s_1, \ldots, s_r$ valor final de los primeros r $(r<n)$ componentes del vector x.

4.5 Observaciones sobre la estructura de la función $f(x,u,t)$

Se deben mencionar dos diferencias fundamentales entre la formulación del problema de optimización dada anteriormente y el problema tratado por Pontryagin.

En la formulación dada, se permite a la función $f(x,u,t)$ ser dependiente de la variable x_n, la cual es maximizada al tiempo $t = 1$. Esta dependencia es permitida por las siguientes razones: el hacer la suposición que $f(x,u,t)$ es independiente de x_n conduce a simplificaciones muy pequeñas en los desarrollos posteriores pero, rompe la simetría entre las variables de estado. Además en muchos problemas prácticos aparece la dependencia de las ecuaciones diferenciales sobre la variable a ser maximizada.

También, en contradicción con la formulación de Pontryagin no se requerirá la independencia, respecto del tiempo, del sistema diferencial. Ya que en el caso de un sistema diferencial de tiempo variable el cual es continuamente diferenciable respecto a t, podemos agregar una variable de estado artificial y transformar el problema a uno como el tratado por Pontryagin. Sin embargo, en el caso general de un sistema de tiempo variable, el cual es solamente continuo por partes con respecto a t, la transformación anterior no puede realizarse y el resultado de Pontryagin no es aplicable y se requerirán los resultados contenidos en este capítulo. Suponemos que $t = 0$ es el tiempo inicial y que $t=1$ es el final.

4.6 <u>Principio de la evolución óptima</u>. Conjuntos alcanzables.

En esta sección introducimos algunos conceptos geométricos, los cuales son de fundamental importancia a la teoría de control óptimo.

Si t está en $[0,1]$ y si X es un estado tal que existe una función de control μ en la clase F con $x(t;\mu) = x$, diremos que el estado X es alcanzable al tiempo t. Para toda t en $[0,1]$ definiremos $W(t)$ como el conjunto de todos los estados que son alcanzables al tiempo t.

i.e. $$W(t) = \{x(t;\mu): \mu \text{ está en } F\}$$

Definición: La función $H(x,u,t,\lambda)$ está definida por la relación:

$$H(x,u,t,\lambda) = \langle f(x,u,t),\lambda\rangle \qquad \lambda \text{ en } R^n$$

y es denominada Hamiltoneano.

Teorema: Principio máximo de Pontryagin: Si υ es una función de control óptimo, entonces, existe una función $\lambda(t)$, definida y contínua sobre $[0,1]$, diferenciable en los puntos de - continuidad de t sobre $(0,1)$ de las funciones υ y $f(x,\upsilon,t)$ y diferente de cero tal que:

(i) $H(x(t;\upsilon),\upsilon(t),t,\lambda(t)) \geq H(x(t,\upsilon),u,t,\lambda(t))$

para toda t en $\Xi(\upsilon)$ y toda u en Ω

(ii) $\dot{\lambda}(t) = - \left. \frac{\partial H(x,\upsilon(t),t,\lambda(t))}{\partial x} \right|_{x=x(t,\upsilon)}$ para toda t en Ξ (υ)

(iii) $\lambda_i(1) = 0 \qquad i = r+1,\ldots, n-1$

(iv) $\lambda_n(1) \geq 0$

4.7 Prueba del principio del máximo para un polisistema dinámico elemental.

Suposiciones: La función f(x,u,t) es independiente del vector x y toma la forma Q(u,t). O sea, que la evolución del polisistema dinámico bajo consideración está descrito por la ecuación:

$$\dot{x} = Q(u,t)$$

De acuerdo a las suposiciones hechas en la sección 4.1 para la función f(x,u,t), supondremos además en esta parte que la función Q(u,t) es contínua con respecto a u, y contínua por partes con respecto a t y que es uniformemente acotada sobre $[0,1]$ para toda u en Ω.

Bajo estas suposiciones, la trayectoria x(t,u) toma la forma:

$$x(t,u) = \int_0^t Q(u(\zeta),\zeta)d\zeta)$$

y tenemos inmediatamente que

$|x(t,u)| \leq M$ para toda u en F y toda t en $[0,1]$

Consideremos el conjunto:

$S^+ = \{x:\ x \text{ en } S,\ x_n > x_n^*(1)\} = \{x:x \text{ en } S,\ x_n > x_n(1;\upsilon)\}$

Es inmediato que este conjunto es convexo.

Los conjuntos S^+ y W no pueden tener ningún punto común, ya que la existencia de algún punto contradiría la optimalidad de la función υ.

Proposición: El conjunto de los estados alcanzables al tiempo t = 1.

$W = \{x(1;\mu)\ :\ \mu \text{ en } F\} = \{\int_0^1 Q(u(t),t)dt : \mu \text{ en } F\}$ es un conjunto convexo.

Prueba: Ver [Ref. 10] págs. 213-214 y 235-238.

Ya que los conjuntos S^+ y W son convexos y no tienen puntos comunes, entonces existe un hiperplano que los separa (de separación), no trivial. Esto significa que existe un vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ diferente de cero, y un número real c (la distancia del origen a este hiperplano de soporte) tal que:

$\langle x, \lambda\rangle \leq c$ si x en W

$\langle x, \lambda\rangle \geq c$ si x en S^+

y puede extenderse a

$<x, \lambda> \geq c$ si x en $\bar{S}^+$, donde $\bar{S}^+$ es la cerradura de S^+. O sea: $\bar{S}^+ = \{x : x \text{ en } S,\ x_n \geq x_n(1,\upsilon)\}$

Como por construcción $x(1,\upsilon)$ pertenece a W y S^+ entonces:

$$< x(1;\upsilon), \lambda > = c$$

esto implica que

$$< x, \lambda > \leq <x(1;\upsilon), \lambda> \quad \text{si x está en W}$$

$$< x, \lambda > \geq <x(1;\upsilon), \lambda> \quad \text{si x está en } \bar{S}^+$$

<u>Proposición</u>: $<Q(v(t),t),\lambda> \geq <Q(u,t),\lambda)>$ para toda t en $\Xi(\upsilon)$ y toda u en Ω.

<u>Prueba</u>: Supongamos que no se cumple la proposición para algún punto t_0 en $\Xi(\upsilon)$, entonces existe una u en Ω, y $\eta>0$, tal que:

$$<Q(v(t_0),t_0),\lambda> < <Q(u,t_0),\lambda> - 2\alpha$$

Ya que $\Xi(\upsilon)$ es un conjunto abierto, existe un intervalo $[t_0-\varepsilon, t_0+\varepsilon]$ con $\varepsilon>0$ en que las funciones $Q(v(t),t)$ y $Q(u,t)$ son contínuas y se satisface:

$$<Q(v(t),t),\lambda> < <Q(u,t),\lambda> - \alpha \quad \text{si t está en } [t_0-\varepsilon, t_0+\varepsilon]$$

Si definimos la función:

$$v^*(t) = \begin{cases} v(t) & \text{si t no está en } [t_0-\varepsilon, t_0+\varepsilon] \\ u(t) & \text{si t está en } [t_0-\varepsilon, t_0+\varepsilon] \end{cases}$$

esto implica que:

$$< \int_0^1 Q(v^*(t),t)dt,\lambda> \geq < \int_0^1 Q(v(t),t)dt,\lambda> + 2\varepsilon\alpha$$

o sea : $<x(1,v^*),\lambda> \geq <x(1;v),\lambda> +2\varepsilon\alpha$

como por construcción $x(1,v^*)$ está en W, tenemos una contradicción a la relación:

$<x,\lambda> \leq <x(1,\upsilon),\lambda>$ si x está en W, esto termina la prueba.

Esta proposición, es equivalente a la condición (i) del principio de pontryagin. Ya que λ es constante, la condición (ii) del principio, resulta trivial.

Para demostrar las condiciones (iii), (iv) considere los vectores unitarios e_i, donde $i=r+1,\ldots,n$. Puede observarse que los vectores $x(1,\upsilon)+e_i$ y $x(1;\upsilon)-e_i$ pertenecen a $\bar{S}^+$ para toda $i = r+1,\ldots,n-1$ (porque $x(1,\upsilon)$ está en S).

Ya que $<x,\lambda> \geq c$ si x en $\bar{S}^+$

tenemos:

$$<x(1,\upsilon) - e_i,\lambda> \geq < x(1,\upsilon),\lambda >$$

$$<x(1,\upsilon) + e_i,\lambda> \geq < x(1,\upsilon),\lambda >$$

de donde se concluye que

$$< e_i,\lambda > = 0$$

o bien

$$\lambda_i = 0 \quad i = r + 1,\ldots,n - 1$$

que satisface la condición (iii).

Además, puesto que $x(1,\upsilon) + e_n$ está en $\bar{S}^+$ se cumple que:

$$<x(1,\upsilon) + e_n,\lambda> \;\geq\; < x(1,\upsilon)\;,\;\lambda >$$

o sea

$$< e_n,\lambda > \;\geq 0$$

es decir

$$\lambda_n \;\geq 0 \qquad\qquad (\text{condición (iv)})$$

4.8 Prueba del principio del máximo para un polisistema dinámico lineal.

En esta sección supondremos que la función $f(x,u,t)$, tiene la forma particular:

$$f(x,u,t) = A(t)x + Q(u,t)$$

De acuerdo a las suposiciones iniciales, en esta sección supondremos que la función matricial $A(t)$ es contínua por partes con respecto a t y la función vectorial $Q(u,t)$ es continua con respecto a u y contínua por partes con respecto a t.

Asociado al sistema lineal

$$\dot{x} = A(t)\; x$$

existe una matriz $G(t)$ llamada matriz de transición definida y contínua para toda t en $[0,1]$, para las cuales la matriz $A(t)$ es contínua, es única y tal que tiene las siguientes propiedades (*).

(*) Ver [Ref. 4] págs. 62-66

a) $\dot{G}(t) = -G(t)A(t)$ para toda t en $[0,1]$ donde $A(t)$ y $Q(u,t)$ son contínuas.

b) $G(1) = I$

c) $G(t)$ es no singular y su inversa satisface:

$$\dot{G}^{-1}(t) = A(t)G^{-1}(t) \; ; \; G^{-1}(1) = I$$

Proposición: Para toda μ en F la función vectorial $x(t,\mu)$ toma la forma:

$$x(t;\mu) = G^{-1}(t) \int_0^t G(\zeta)\, Q(u(\zeta),\zeta)d\zeta$$

para toda t en $[0,1]$

Prueba: Sabemos que $x(t,\mu)$ existe y es única tenemos solamente que verificar que el postulado de la proposición satisface la ecuación diferencial:

$$\frac{d}{dt}x(t;\mu) = A(t)x(t,\mu) + Q(u(t),t)$$

para toda t en $\Xi(\mu)$ y la condición de frontera $x(0,\mu) = 0$.

Derivando $x(t,\mu)$ tenemos para toda t en $\Xi(\mu)$

$$\frac{dx(t,\mu)}{dt} = \frac{dG^{-1}(t)}{dt} \int_0^t G(\zeta)Q(u(\zeta),\zeta)d\zeta$$

$$+ G^{-1}(t)\frac{d}{dt}\int_0^t G(\zeta)\, Q(u(\zeta),\zeta)d\zeta$$

$$= A(t)G^{-1}(t) \int_0^t G(\zeta)\, Q(u(\zeta),\zeta)d\zeta$$

$$+ G^{-1}(t)G(t)Q(u(t),t)$$

$$A(t)x(t,\mu) + Q(u(t),t)$$

Y tenemos para la condición final:

$$x(1,\mu) = \int_0^1 G(t)Q(u(t),t)dt.$$

Definiendo el conjunto W como el de todos los estados alcanza bles al tiempo t=1 desde el estado inicial x=0 al tiempo t=0 con alguna función de control en la clase F, tenemos

$$W = \{x(1,\mu)\ ;\ \mu \text{ en } F\}$$

o sea

$$W = \{\int_0^1 G(t)Q(u(t),t)dt\ :\ \mu \text{ en } F\}$$

Siguiendo paso a paso el procedimiento usado en la sección anterior, se llega a probar que para cada función de control óptimo υ existe un vector constante λ, distinto de cero, tal que:

a) $< G(t)Q(v(t),t),\lambda > \leq < G(t)Q(u,t),\lambda >$
para toda t en $\Xi(\upsilon)$ y toda u en Ω.

b) $\lambda_i = 0$ para $i = r+1,\ldots,n-1$

c) $\lambda_n \geq 0$

Como puede observarse, los resultados (b) y (c) coinciden con las condiciones (iii) y (iv) del principio del máximo de Pontryagin.

Definiendo la función vectorial $\lambda(t)$ con:

$$\lambda(t) = G^T(t)\lambda \quad \text{para toda } t \text{ en } [0,1]$$

donde $G^T(t)$ es la matriz transpuesta de G(t). Y se tiene que $\lambda(t)$ es contínua, diferenciable por partes y diferente de cero sobre el intervalo $[0,1]$ por las propiedades de G(t) y λ.

Recordando la definición de operadores adjuntos y que el operador adjunto de una matriz G(t) es su transpuesta, podemos escribir:

$$\langle G(t)\ Q(u,t),\lambda\rangle \quad = \quad \langle Q(u,t),\ G^T(t)\lambda\rangle$$

por lo tanto, el resultado (a) puede escribirse:

$$\langle Q(v(t),t),\ G^T(t)\lambda\rangle \geq \langle Q(u,t),G^T(t)\lambda\rangle$$

o sea

$$\langle Q(v(t),t),\ \lambda(t)\rangle \geq \langle Q(u,t),\lambda(t)\rangle$$

lo que implica

$$\langle (A(t)x(t,\upsilon) + Q(v(t),t)),\lambda(t)\rangle \geq \langle (A(t)x(t,\upsilon) + Q(u,t)),\lambda(t)\rangle$$

para toda t en $\Xi(\upsilon)$ y toda u en Ω que es la condición (i) del principio del máximo de Pontryagin.

De la definición de $\lambda(t)$ tenemos:

$$\frac{d\lambda(t)}{dt} = \left(\frac{d}{dt}\ G(t)\lambda\right)^T \quad = \left(-G(t)A(t)\right)^T\ \lambda$$

$$= -A^T(t)\ G^T(t)\lambda \quad = \quad -A^T(t)\lambda(t)$$

Por lo que se cumple la condición (ii) del principio del máximo.

Con esto queda demostrado el principio del máximo para un polisistema dinámico lineal.

4.9 Prueba del principio del máximo para un polisistema dinámico general.

Sea υ una función de control óptimo para el problema de optimización establecido. Sea A(t) una función matricial de nxn cuyo elemento (i,j) es

$$\left.\frac{\partial\ f_i(x,v(t),t)}{\partial\ x_j}\right|_{x=x(t;\upsilon)}$$

De las suposiciones hechas, sabemos que la función matricial A(t) es contínua por partes sobre $[0,1]$, y sea G(t) la matriz de transición asociada al sistema cuando la función de control es υ, esta función es contínua en t sobre $[0,1]$ y diferenciable respecto a t sobre $\Xi(\upsilon)$.

Debe observarse que las matrices A(t) y G(t) han sido definidas particularmente, para la función de control υ. Matrices similares pero diferentes, deben ser definidas para las otras funciones de control en la clase F.

La propiedad característica del sistema lineal considerado en la sección anterior es que las matrices A(t) y G(t) son las mismas para toda función de control en la clase F.

Definamos para cada función de control en la clase F, a la función vectorial $v(t,u)$ por la relación:

$$y(t,\mu) = x(t,\mu)-x(t,\upsilon) \text{ para toda } t \text{ en } [0,1]$$

A esta función se le conoce como la trayectoria variacional para el control μ, respecto al control υ.

Sea $\Xi^*(\mu) = \Xi(\mu) \cap \Xi(\upsilon)$

La función $y(t,\mu)$ es contínua respecto a t en $[0,1]$, diferenciable respecto a t en $\Xi^*(\mu)$ y se tiene:

$$\dot{y}(t,\mu) = f(x(t,\mu), u(t)\ t) - f(x(t,\upsilon), v(t), t)$$

para toda t en $\Xi^*(\mu)$

Definamos a las funciones vectoriales $Q(u,t)$ y $k(t,\mu)$ como:

$$Q(u,t)=f(x(t,\upsilon),u,t) - f)x(t,\upsilon),v(t),t)$$

$$k(t,\mu)=f(x(t,\mu),u(t),t)-f(x(t,\upsilon),v(t),t)$$

$$-Q(u(t),t)-A(t)(x(t,\mu)-x(t,\upsilon))$$

La función $Q(u,t)$ es contínua respecto a u y continúa por partes respecto a t. La función $k(t,\mu)$ es contínua por partes respecto a t, para toda μ en F. Sustituyendo tenemos:

$\dot{y}(t,\mu) = A(t)\ \ y(t,\mu) + Q(u(t),t) + k(t,\mu)$ para toda t en $\Xi^*(\mu)$

Por lo tanto:

$y(t,\mu) = G^{-1}(t) \displaystyle\int_0^t G(\zeta)(Q(u(\zeta),\zeta)+ k(\zeta,\mu))d\zeta$ para toda t en $[0,1]$.

La función $k(t,\mu)$ es igual a cero para el problema lineal considerado en la sección anterior, sin embargo, no lo es para el problema no lineal considerado en esta sección.

Para cada μ en F sea $z(t,\mu)$ una función vectorial de t definida y contínua sobre $[0,1]$, y diferenciable sobre $\Xi^*(\mu)$ tal que:

i) $\dot{z}(t,\mu) = A(t)z(t,\mu) + Q(u(t),t)$

ii) $z(0,\mu) = 0$

es decir:

$$z(t,\mu) = G^{-1}(t) \int_0^t G(\zeta)Q(u(\zeta),\zeta)d\zeta \qquad \text{para toda } t \text{ en } [0,1]$$

Por lo tanto:

$$y(t,\mu) - z(t,\mu) = G^{-1}(t) \int_0^t G(\zeta)k(\zeta,\mu)d\zeta \qquad \text{para toda } t \text{ en } [0,1]$$

Sean los conjuntos W y $\tilde{W}$ definidos de la siguiente manera:

$$W = \{x(1,\upsilon) + y(1,\mu) : \mu \text{ en } F\}$$

$$\tilde{W} = \{x(1,\upsilon) + z(1,\mu) : \mu \text{ en } F\}$$

Equivalentemente, pueden escribirse:

$$W = \{x(1,\mu) : \mu \text{ en } F\}$$

$$\tilde{W} = \{x(1,\upsilon) + z : z \text{ en } Z\}$$

donde

$$Z = \{z(1,\mu) : \mu \text{ en } F\}$$

De los resultados de la sección anterior, se sabe que el conjunto Z es convexo y por lo tanto $\tilde{W}$ es convexo. El conjunto S^+, definido anteriormente, es también convexo.

<u>Lema</u>: Si no existe ningún hiperplano que separe los conjuntos convexos S^+ y $\tilde{W}$, entonces los conjuntos S^+ y W tienen al menos un punto en común.

Prueba: Ver [Ref. 10] págs. 242-247.

<u>Corolario</u>: Si los conjuntos S^+ y W no tienen ningún punto en común entonces existe un hiperplano que separa los conjuntos convexos S^+ y $\tilde{W}$.

Los conjuntos S^+ y W no pueden tener ningún punto en común pues de ser así, se contradiría la optimalidad de la función de control υ. Por lo tanto, aplicando el corolario anterior, se concluye la existencia de un hiperplano que separa los conjuntos convexos S^+ y $\tilde{W}$. Siguiendo los mismos pasos utilizados en la prueba del polisistema elemental, se prueba que para la función de control óptimo, existe una función $\lambda(t)$, definida, contínua, diferenciable por partes y diferente de cero sobre $[0,1]$ tal que:

a) $\langle Q(v(t),t),\ \lambda(t)\rangle \geq \langle Q(u,t),\lambda(t)\rangle$ para toda t en $\Xi(\upsilon)$ y toda y en .

b) $\frac{d}{dt}\lambda(t) = -A^T(t)\lambda(t)$ para toda t en $\Xi(\upsilon)$

c) $\lambda_i(1) = 0$ para $i = 1+1,\ldots, n-1$

d) $\lambda_n(1) \geq 0$

Sustituyendo la definición dada de $Q(u,t)$ en la parte (a) de los resultados, llegamos:

$$<Q(v(t),t),\lambda(t)> \;\geq\; <Q(u,t),\lambda(t)> =$$

$$<(f(x(t,\upsilon),v(t),t) - f(x(t,\upsilon),v(t),t)),\ \lambda(t)> \;\geq$$

$$<(f(x(t,\upsilon),u,t) - f(x(t,\upsilon),v(t),t)),\ \lambda(t)>$$

$$\therefore\ <f(x(t,\upsilon),v(t),t),\lambda(t)> \;\geq\; <f(x(t,\upsilon),u,t),\lambda(t)>$$

que es precisamente la condición (i) del principio del máximo.

Sustituyendo en el resultado (b) la definición dada en esta sección de la función matricial A(t) tenemos:

$$\frac{d}{dt}\lambda(t) = -A^T(t)\lambda(t)$$

$$= - \begin{bmatrix} \left.\dfrac{\partial f_1(x,v(t),t)}{\partial x_1}\right|_{x=x(t,\upsilon)} & \cdots & \left.\dfrac{\partial f_n(x,v(t),t)}{\partial x_1}\right|_{x=x(t,\upsilon)} \\ \vdots & \ddots & \vdots \\ \left.\dfrac{\partial f_1(x,v(t),t)}{\partial x_m}\right|_{x=x(t,\upsilon)} & \cdots & \left.\dfrac{\partial f_n(x,v(t),t)}{\partial x_n}\right|_{x=x(t,\upsilon)} \end{bmatrix} .\lambda(t)$$

$$= - \left.\frac{\partial f(x,v(t),t)}{\partial x}\right|_{x=x(t,\upsilon)} \qquad \lambda(t) = - \left.\frac{\partial H(x,v(t),t,\lambda(t))}{\partial x}\right|_{x=x(t,\upsilon)}$$

que es la condición (ii) del principio del máximo. Los resultados (c) y (d) son exactamente las condiciones (iii) y (iv) del principio del máximo. Esto termina la prueba general del principio del máximo de Pontryagin.

REFERENCIAS BIBLIOGRAFICAS

1. Apostol, Tom M., MATHEMATICAL ANALYSIS, Second Edition, Adison-Wesley, 1974.

2. Arrow, K.J., Hurwicz, L. y Uzawa H., STUDIES IN LINEAR AND NON-LINEAR PROGRAMMING, Stanford Univ. Press. 1964.

3. Athans, M. y Falb, P.L, OPTIMAL CONTROL: An introduction to the theory and its applications, Mc-Graw-Hill, 1966.

4. Desoer, C.A., NOTES FOR A SECOND COURSE ON LINEAR SYSTEMS, D. Van-Nostrand, 1970.

5. Elsgoltz, L., ECUACIONES DIFERENCIALES Y CALCULO VARIACIONAL, Ediciones de Cultura Popular, 1975.

6. Fleming, W.H. y Rishel, R.W., DETERMINISTIC AND STOCHASTIC OPTIMAL CONTROL, Springer-Verlag, 1975.

7. Hestenes, Magnus R., CALCULUS OF VARIATIONS AND OPTIMAL CONTROL THEORY, John Wiley, 1967.

8. Kolmogorov, A.N. y Fomin S.V., ELEMENTOS DE LA TEORIA DE FUNCIONES Y EL ANALISIS FUNCIONAL, Editorial Mir. Moscú., 1975.

9. Kreyszig, E., INTRODUCTORY FUNCTIONAL ANALYSIS WITH APPLICATIONS, John Wiley, 1978.

REFERENCIAS BIBLIOGRAFICAS (Cont.)

10. Leitmann, G., TOPICS IN OPTIMIZATION, Academic Press, 1967.

11. Luenberger, David G., OPTIMIZATION BY VECTOR SPACE METHODS, John Wiley, 1969.

12. Luenberger, David G., INTRODUCTION TO DYNAMIC SYSTEMS: Theory, models and applications, John Wiley, 1979.

13. Pontryagin, L.S., Boltianski, V., Gamkrelidze R.V. y Mishechenko, E., THEORY MATHEMATIQUE DES PROCESSUS OPTIMAUX, Editions de Moscou, 1974.

14. Royden, H.L, REAL ANALYSIS, Mac-Millan, 1963.

15. Rudin, W., FUNCTIONAL ANALYSIS, Mac Graw-Hill, 1974.

16. Smith, D.R., VARIATIONAL METHODS IN OPTIMIZATION, Prentice Hall, 1974.

17. Taylor, A.E., INTRODUCTION TO FUNCTIONAL ANALYSIS, John Wiley, 1958.