UNIVERSIDAD DE PANAMA
VICERRECTORIA DE INVESTIGACIÓN Y POSTGRADO
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y TECNOLOGÍA
MAESTRIA EN MATEMÁTICA

EXTREMOS CONDICIONADOS EN ESPACIOS DE BANACH

METODOS DE LOS MULTIPLICADORES DE LAGRANGE

EFRAIN CASIS KAWANO

Tesis presentada como uno de los requisitos para optar al grado de Maestro en Ciencias con Especialización en Matemática Pura.

Panamá, República de Panamá

2005

APROBADO POR:

Dr. Rogelio Rosas
PRESIDENTE

M.Sc. Josué Ortiz
MIEMBRO

MSc. Maria Dixiana Espinoza
MIEMBRO

REPRESENTANTE DE LA VICERRECTORÍA DE INVESTIGACIÓN Y POSTGRADO

FECHA: 21 de diciembre de 2005

DEDICATORIA

Con la convicción de que este trabajo representa un esfuerzo para quien ha llegado a la conclusión de que un hombre es verdaderamente libre cuando se abre camino a través de la superación intelectual y considerando además, que éste representa entre muchas otras cosas, un mensaje para mis queridos hijos, en cuanto a que es necesario que nos empeñemos en alcanzar los más altos niveles de conocimientos, en el campo en que nos destacamos como profesional, porque sólo así, podremos estar seguro que nos merecemos un respeto por lo que sabemos Por esta razón deseamos con este trabajo rendirle homenaje póstumo a mis queridos padres quienes me sirvieron de inspiración para realizarlo, y como tributo por sus desvelos y preocupaciones a mi querida esposa y a todos mis cinco hijos

EFRAÍN

AGRADECIMIENTO

En un mundo tan complejo como el nuestro, se hace necesario que una persona esté permanentemente estableciéndose metas, que lo mantengan haciendo suficiente esfuerzos para acondicionar el futuro, no solamente de uno mismo, sino de todos aquellos seres que nos rodean y que se les anteponen a nuestros consejos verbales, las fuerzas que desatan los modernos adelantos tecnológicos Y es por esto, que el presente trabajo para mí representó el producto que me imprimieron dos enormes fuerzas motivadoras, al punto que no medimos el nivel de fatiga que la misma me causó de una forma o de otra

Dichas fuerzas fueron, por una parte el cumplir con el requisito que ofrece nuestra primera casa de estudios para optar por el título de Maestría en Matemática Pura y que en este caso se trata de ésta Tesis de Graduación, y por el otro, y tal vez para mí el más importante, el de darle un ejemplo a mis hijos en el sentido de que observen que el camino que nos queda para saltar de la miseria a la abundancia se inicia en el cerebro humano Este es el secreto Y el estudio es el camino para despejar todas las incertidumbres que se nos presenten

Por nuestra parte, dejamos constancia de nuestro agradecimiento a todas aquellas personas que de una u otra forma nos ayudaron en esta encomiable tarea después de **Dios**, tales como al **Doctor Rogelio Rosas**, quien sabiamente me dirigió este trabajo, a todos los brillantes profesores de la Maestría que nos ilustraron y sirvieron de inspiración en los contenidos de este trabajo, a la adorable joven **Greta Salazar** que se encargó con su gran experiencia de darle el toque de estética que este trabajo posee y para finalizar, quiero agradecerle a mi esposa Ivana y a mis queridas hijas **Tiziana y Ornella**, quienes me apoyaron con su paciencia y ecuanimidad para que llegara hasta el final con optimismo

"MUCHAS GRACIAS A TODOS"

EFRAÍN

RESUMEN

En este trabajo se considera el problema de la determinación de los extremos de una función numérica definida en un abierto de un espacio de Banach real, bajo la suposición que la variable independiente esta sujeta a restricciones funcionales, o sea el problema de los extremos condicionados en espacios de Banach El resultado mas importante que se presenta en el primer capítulo, consiste en aclarar que los problemas de extremo en espacios de Banach complejos con condiciones de diferenciabilidad son triviales, lo que justifica la consideración de espacios reales esta aclaración es pertinente y la misma justifica el carácter limitado (a espacios reales) de todo el primer capítulo En el segundo capítulo se presentan los resultados esperados El Teorema de Ljusternik sobre condiciones necesarias de extremo condicionado (De los Multiplicadores de Lagrange) y un Teorema sobre condiciones suficientes de extremo En ambos casos se dan algunos ejemplos y se hacen comentarios acerca de las hipótesis, dándole formas mas manejables a las mismas (como por ejemplo las condiciones de independencia de las restricciones) Se incluye una interpretación del Multiplicador de Lagrange como diferencial de las restricciones con respecto al cambio de dichas condiciones, representación esta de utilidad en las Teorías Económicas

Summary

In this work, the problem of determine extrema for numerical functions defined on a real Banach space is considered, asuming that the independent variables is subject to functional restriccions This is the problem of conditional extreme Outstanding result from Chapter 1 is that this kind of problems, under differentiability conditions for complex spaces are trivial since differentiability implies the character of locally constants for such functions So the problem must be analize in real space Chapter 2 presents the classical results on Lagrange Multipliers (necesary conditions) and sufficent conditions for such kind of extreme, In both cases examples and coments are giving in order to explain the hipothesis of the main propositions An interpretation of the

Lagrange Multipliers as differentials of the restrictions, a very usefull point of view in Economics is included in this second chapter

INDICE

INTRODUCCIÓN

Uno de los problemas más importante del análisis matemático es el de la determinación de los extremos de una función, sean estos máximos o mínimos locales, o extremos condicionados, es decir cuando las variables independientes que intervienen están ligadas entre si por relaciones funcionales Para enfrentar el problema de los extremos se ha desarrollado el cálculo diferencial, una herramienta que permite sustituir localmente las funciones dadas por funciones más sencillas, como las afines, las cuadráticas, etc, y pasar el problema a éste tipo de aproximaciones

El programa anterior se desarrolla sin mucha dificultad en los espacios de dimensión finita y por su sencillez es deseable extender estos procedimientos a espacios mas generales Hay que notar que las operaciones involucradas en el cálculo de las derivadas implican la existencia en el dominio de las funciones de operaciones de tipo vectorial y también una estructura topológica que permita calcular límites, por esto el ambiente adecuado para una generalización del cálculo diferencial es el de los espacios normados y con más precisión el de los espacios de Banach reales

El cálculo diferencial provee las condiciones necesarias y suficientes para la existencia de estructura locales, en este sentido vale el siguiente resultado

"Para que una función doblemente diferenciable definida en un abierto tenga un extremo en un punto interior de su dominio es necesario que en dicho punto el diferencial sea nulo

Este punto es de extremo si la segunda diferencial de la función es una forma cuadrática definida positiva "

Para los extremos condicionados es decir la determinación de máximos o mínimos de una función $f \quad A \subseteq \mathbf{R}^n \to \mathbf{R}$, pero con condiciones sobre la variable independiente del tipo

$g(x) = 0$

existe un método de solución, llamado método de los Multiplicadores de Lagrange, pero este método se puede aplicar solo si se dan ciertas condiciones sobre la función y sobre los vínculos, como se establece en el siguiente teorema

Teorema (De Lagrange)

Sean f y g definidas y diferenciables en una región D Para que f alcance un valor extremo en un punto $\mathbf{P}_0$ en D, donde $\nabla g \neq 0$, es necesario que exista un número λ tal que

$\nabla f(\mathbf{P}_0) + \lambda \nabla g(\mathbf{P}_0) = \mathbf{0}$

y que

$g(\mathbf{P}_0) = \mathbf{0}$

Un análisis atento de este teorema permite señalar que el mismo puede aplicarse en un contexto geométrico más amplio, específicamente se puede utilizar para resolver el siguiente problema

" Determinar los extremos de f $U \rightarrow \mathbf{R}$ si $x \in \mathbf{V}$, donde **V** se puede representar localmente por ecuaciones del tipo $g(x) = 0$, con funciones g "suaves"

Por ejemplo, determinar el máximo o mínimo de

$F(x,y) = x + y$

en la circunferencia unitaria

Aquí la dependencia de las variables independientes esta dada por

$x^2 + y^2 = 1$

Este tipo de subconjunto (como la circunferencia) es lo que se llama una variedad diferenciable, y en los espacios $\mathbf{R}^n$ éstas se pueden definir, con apoyo de los teoremas de funciones implícitas e inversas

Este trabajo se articula en dos capítulos

En el primer capitulo de esta tesis se desarrolla el marco teórico necesario para resolver el problema de los extremos condicionados Se trata de elementos básicos del Análisis Funcional, como el principio de contracción de Banach-Cacciopoli, el Teorema de Hahn-Banach (para el caso real) y el Teorema de la Aplicación Abierta, y el Teorema de Representación de Riez-Fisher, se introducen los conceptos propios del cálculo diferencial en espacios de Banach Diferenciabilidad según Fréchet, diferenciales de orden superior y polinomios de Taylor y se establecen las condiciones necesarias y suficiente de extremos locales, también en este capítulo se consideran los teoremas de funciones implícitas e inversas En este primer capítulo se aclara también porque la aplicación de las técnicas del Cálculo Diferencial a los problemas de extremo tiene sentido sólo en los espacios vectoriales reales, justificando así la supuesta limitación de considerar los resultados presentados solo en espacios de Banach reales y no en su forma mas general en los espacios complejos

El segundo capítulo se presenta de concepto de Variedad Diferenciable regular, presentada de cierta manera como “conjunto solución “ de ecuaciones del tipo g(x) = 0, para ciertas clases de funciones g (con cierto grado de diferenciabilidad) y se dan los elementos geométricos necesarios para un adecuada visualización de los resultados que se presentan en este capítulo También se formulan y demuestran los teoremas de Ljustermik (generalización del teorema de Lagrange) -dando una interpretación del significado de los Multiplicadores de Lagrange- y el Teorema que da condiciones suficientes (de segundo grado) para la determinación de los extremos condicionados y se presentan aplicaciones y aclaraciones sobre estos resultados

CAPÍTULO I

Generalidades

En este primer Capítulo se presentan algunos conceptos generales acerca de los Espacios de Banach, que sirven de base para una extensión de los conceptos relacionados con los extremos de funciones a valores reales Se tratan específicamente los Teoremas de contracción, los de Hahn-Banach (caso real) el Teorema de la aplicación abierta y los elementos del Cálculo Diferencial en espacios de Banach que son de utilidad en los problemas de máximo y mínimo Como referencias básicas se indican las siguientes obras [Dieudonné, J Éléments d'Analyse 1 CapítuloVIII Gauthier-Villars, Paris, 1960] , [Nachbin, Leopoldo Introducáo A Analise Funcional Espacos de Banach e Cálculo diferencial (Segunda parte) Monografías Matemáticas de La O E A Washington, 1976] y [Rudin, Walter Functional Analysis Tata Mc Graw Hill, 1973]

En todo este trabajo, con X se indicará un espacio normado real, cuya norma de ser pertinente se indicará $\| \|$ A la par de los espacios normados se consideran también sus duales, es decir los espacios vectoriales reales de las aplicaciones lineales $f : X \to R$ que son continuas, normado con

$$\|f\| = \sup_{\|x\|=1} |f(x)|$$

El dual de X , resulta ser un espacio de Banach y se indicará X^*

Definición de Aplicación Lipschitziana y de Contracción

Sea X un espacio de Banach $D \subset X$ y $T : D \to X$ una aplicación Se dice que T es una **Lipschitziana** si existe $c > 0$ tal que Para todo $x, y \in D$

$$\|T(x) - T(y)\| \leq c\|x - y\| \qquad (1)$$

Si $T . D \to D$ es Lipschitziana y la constante c es menor que 1 se dice que T es una **contracción**

Observación: La condición establecida en la definición por la desigualdad (1) asegura que las aplicaciones Lipschitzianas son uniformemente continuas en su dominio

El siguiente resultado es tal vez el mas importante a propósito de la contracciones

Teorema de Banach-Caccioppoli

(Principio de contracción) Sea X un espacio de Banach $D \subset X$ cerrado y $T \quad D \to D$ una contracción Entonces T tiene un punto fijo único, o sea que existe un y solo un punto $x \in D$ tal que $T(x) = x$

Mas que una demostración este teorema amerita algunos comentarios El procedimiento para demostrar que efectivamente existe al menos un punto fijo se basa en el método de las aproximaciones sucesivas En este método se parte de un punto cualquiera $x_0 \in D$, y se construye una sucesión aplicando sucesivamente T, así

$x_1 = T(x_0), x_2 = T(x_1), \ldots$ etc Se construye así una sucesión en D, luego se demuestra que es de Cauchy y como el espacio es completo ésta converge a un punto de D porque este conjunto es cerrado En cuanto a la unicidad esta es consecuencia del carácter contractivo de la aplicación pues la condición (1) cuando los puntos son fijos equivale a

$$\|x-y\| \le c\|x-y\|$$

Condición ésta última que se puede satisfacer si y solo si los dos miembros de la desigualdad son iguales a 0, lo que lleva a la igualdad de los puntos

El principio de contracción se utiliza para determinar si una dada ecuación tiene solución única en ciertas condiciones, por ejemplo en el caso de los Problemas de Cauchy de las ecuaciones diferenciales

Un ejemplo del uso del Teorema de Banach-Caccioppoli en el sentido de este trabajo lo proporciona el siguiente resultado

Teorema de Homeomorfismo

Sea X un espacio de Banach y $T . X \to X$ una contracción, entonces la aplicación $I - T : X \to X$ es un homeomorfismo

Demostración

Sea $y \in X$ y $T_y(x) = T(x) + y$ Obviamente T_y es contractiva, por lo que existe un único x tal que

$T_y(x) = x$

o lo que es lo mismo

$x - T(x) = y$

Esto asegura que $I - T$ es suryectiva e inyectiva

$I - T$ es continua y en cuanto a la continuidad de su inversa $(I - T)^{-1}$ obsérvese que

$\|(I - T)(x) - (I - T)(x')\| = \|x - x' - (T(x) - T(x')\| \geq (1 - \alpha)\|x - x'\|$

si α es la constante de contracción de T Teniendo presente que $x = (I - T)^{-1}(I - T)(x)$,

$\|(I - T)^{-1}(x) - (I - T)^{-1}(x')\| \leq \frac{1}{1 - \alpha}\|x - x'\|$

o sea que $(I - T)^{-1}$ es continua■

Observación 1 El teorema anterior es una extensión no lineal del teorema de Neuman.

“Si $A \in L(X, X)$ y $\|A\| < 1$ entonces $I - A$ es un homeomorfismo y

$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n$ "

Observación 2: Si $A, A^{-1} \in L(X,X)$ y $T : X \to X$ es Lipschitziana con constante α, para que $A-T$ sea un homeomorfismo basta que

$$\alpha < \frac{1}{\|A^{-1}\|}$$

En efecto $A-T = A(I-A^{-1}T)$ Ahora bien

$$\|A^{-1}T(x) - A^{-1}T(x')\| \le \|A^{-1}\| \|T(x)-T(x')\| \le \|A^{-1}\| \alpha \|x-x'\|$$

Por tanto si $\alpha < \dfrac{1}{\|A^{-1}\|}$ entonces $\alpha\|A^{-1}\| < 1$ Por la Observación 1, $I - A^{-1}T$ es un homeomorfismo y $A-T$ también por ser compuesta de homeomorfismos

Observación 3: Sea Sea X un espacio de Banach $D \subset X$ abierto T $D \to X$ es un **homeomorfismo local en** $x \in D$ si existen dos vecindades abiertas de x y $T(x)$, U,V respectivamente tales que $T/_U$ es un homeomorfismo de U en V Pues bien, si $T : D \to X$ y existen $S(x_0, r) \subset D$ y $A \in L(X,X)$ con $A^{-1} \in L(X,X)$ tales que

$\|T(x) - T(x') - A(x-x')\| \le \alpha \|x-x'\|$ para todo $x, x' \in S(x_0, r)$ y $\alpha < \dfrac{1}{\|A\|}$ entonces T es un homeomorfismo local en x_0

El Teorema de Hahn-Banach

De las numerosas versiones del Teorema de Hahn-Banach interesa aquí la que establece la posibilidad de extender aplicaciones lineales continuas, definidas en un subespacio a todo el espacio de Banach, conservando la norma Aunque este resultado es válido en el caso general de los espacios complejos, para los fines de este trabajo-y como se aclarará después- es suficiente la forma real del Teorema En la vía de demostrar el Teorema se necesita el siguiente resultado

Lema. Sea M un subespacio de un espacio normado X y $f : M \to R$ un funcional lineal tal que para una oportuna constante $K > 0$ resulte $f(x) \leq K\|x\|$

Sea x_0 un elemento fijo de X Entonces, para todo número real c , las siguientes desigualdades son equivalentes

$f(x) + \mu c \leq K\|x + \mu x_0\| \qquad \forall x \in M, \forall \mu \in R \qquad (1)$

$-K\|x + x_0\| - f(x) \leq c \leq K\|x + x_0\| - f(x) \quad \forall x \in M \quad (2)$

Además, existe al menos un número real c que satisface la (2) y por tanto la (1).

Demostración:

$(1) \Rightarrow (2)$ Si $\mu = 1$ de (1) resulta $f(x) + c \leq K\|x + x_0\|$, mientras que si $\mu = -1$ y en (1) se considera $-x \in M$ entonces $f(-x) - c \leq K\|x + x_0\|$ Es obvio que de estas dos desigualdades se deduce la (2)

$(2) \Rightarrow (1)$ Considérese primero $\mu > 0$ Substituyendo en la (2) x por $\frac{x}{\mu}$ $(\in M)$ se tiene

$c \leq K\left\|\frac{x}{\mu} + x_0\right\| - f(\frac{x}{\mu}) = \frac{1}{\mu}(K\|x + x_0\| - f(x)$, luego $\mu c \leq K\|x + \mu x_0\| - f(x)$ y de esta por una simple transposición de términos se obtiene la (1)

Para valores negativos de μ, como $-\mu > 0$ se puede proceder como antes pero utilizando la parte izquierda de la (2) llegando nuevamente a la (1)

El caso $\mu = 0$ es trivial, por esto de (2) se deduce la (1)

En cuanto a la última afirmación del lema, si $x, y \in M$, como f es lineal resulta

$f(x) - f(y) = f(x - y) \leq K\|x - y\| \leq K\|x + x_0\| + K\|y + x_0\|$, transponiendo oportunamente términos en esta desigualdad se tiene lo siguiente

$-K\|y + x_0\| - f(y) \leq K\|x + x_0\| + f(x)$, y esto para cualesquiera $x, y \in M$ Esto asegura la existencia del supremo del miembro izquierdo y del ínfimo del derecho y además que $\sup_{y \in M}(-K\|y + x_0\| - f(y)) \leq \sup_{x \in M}(K\|x + x_0\| + f(x))$

Cualquier número real c comprendido entre estos extremos satisface la (2), con lo que queda completamente demostrado el lema ■

Observación: Como se trata de aplicaciones lineales definidas en un espacio real a valores en R, la condición $f(x) \leq K\|x\|$ equivale a la continuidad de la aplicación

Teorema de Hahn-Banach

Sea M un subespacio de un espacio normado(real) X y $f: M \rightarrow R$ un funcional lineal y continuo Entonces existe continuo $F \in X^{*}$ tal que $F/_{M} = f$ y $\|F\| = \|f\|$

A continuación se presenta un esbozo de la demostración de este Teorema, indicando como se utiliza el **Lema** en la demostración.

Por medio del Lema de Zorn se prueba que existen extensiones maximales de f, es decir aplicaciones lineales continuas $\hat{f}: \hat{M} \rightarrow R$, donde $M \subset \hat{M}$ y $\hat{f}/_{\hat{M}} = f$ que conservan la norma Pues bien, una extensión maximal de f con las características dadas tiene como dominio X En efecto, si este no fuera el caso y si $\hat{f}: \hat{M} \rightarrow R$ es una extensión maximal , para $x_0 \in X - M$, la función $\hat{f}$ puede extenderse al subespacio generado por $\hat{M} \cup \{x_0\}$- que denotaremos $\hat{M}'$-tomando un $c \in R$ que satisfaga cualquiera de las desigualdades del **Lema**, y poniendo

$$\hat{f}'(x + \mu x_0) = \hat{f}(x) + \mu c, \quad \forall x \in \hat{M}, \forall \mu \in R$$

Esta función es continua, lo cual está garantizado por la (1), y por esta misma razón su norma coincide con la de la función original Lo anterior contradice el carácter maximal de la extensión por lo que el dominio de ésta última tiene que coincidir con todo el espacio

Hay varios corolarios del Teorema de Hahn-Banach que juegan un papel muy importante en el Análisis Funcional, por esta razón se incluyen aquí

Corolario 1 Sea x_0 un elemento no nulo de X Entonces existe $f_0 \in X^*$ tal que $f_0(x_0) = \|x_0\|$ y $\|f_0\| = 1$

Demostración: Sea $M = \{\lambda x_0 | \lambda \in R\}$ el subespacio generado por x_0, normado con la norma inducida dela de X La aplicación $f : M \to R$ definida por.

$f(x) = \lambda\|x_0\|$ si $x = \lambda x_0$ es lineal, y continua, ya que $|f(x)| = |\lambda|\|x_0\| = \|x\|$

Además, como $|f(x)| = \|x\|, \ \forall x \in M$ resulta $\|f\| = 1$ Aplicando el Teorema de Hahn-Banach se obtiene el resultado■

Corolario 2 X^{*} separa los puntos de X

Demostración En efecto, si $x \in X$ y para todo $f \in X^{*}$ resulta $f(x) = 0$ entonces $x = 0$, por esto si $x, y \in X, x \neq y$, entonces existe $f \in X^{*}$ tal que $f(x) \neq f(y)$ ■

Estos hechos se utilizan con mucha frecuencia para probar la igualdad de puntos de un espacio normado, probando que sus valores coinciden para todo funcional lineal

El Teorema de la Aplicación Abierta

Siguiendo las ideas expuestas en la introducción se considera ahora otro importante resultado del Análisis Funcional, el denominado "Teorema de la Aplicación abierta" Aquí es necesario precisar el lenguaje

Definición de Aplicación Abierta

Una aplicación entre espacios topológicos se dice **abierta** en un punto de su dominio si la imagen de una vecindad del punto por la aplicación contiene una vecindad de la imagen del punto en el codominio Si la aplicación es abierta en cada punto del dominio entonces transforma abiertos del dominio en abiertos del codominio

En general no toda aplicación continua es abierta, es fácil convencerse de esto considerando aplicaciones constantes a valores en espacios topológicos en el que los conjuntos unitarios no son abiertos (tal es el caso de

R), otro ejemplo en este sentido lo proporcionan, en el caso finito las aplicaciones lineales de un espacio en otro de dimensión mayor

En el caso de las transformaciones lineales entre espacios normados, como quiera que la topología está completamente determinada por las vecindades del origen, es suficiente, para probar el carácter abierto de una tal aplicación comprobar que la misma es abierta en el origen En efecto

Considérese una aplicación lineal f entre dos espacios normados abierta en el origen Si ahora se considera una vecindad de un punto cualquiera del dominio x, V_x, que se puede suponer simétrica, entonces $x-V_x$ es vecindad de 0 y por la hipótesis que V es abierta en el origen existe una vecindad de 0 en el codominio W tal que $W \subset f(x-V_x)=f(x)-f(V_x)$ De aquí se deduce que $f(x)+W \subset f(V_x)$ ya que la V_x es simétrica

Como en el caso del Teorema de Hahn-Banach, hay muchos Teoremas que tienen este nombre, la versión que se presenta aquí es la mas adecuada al ambiente de este trabajo, o sea una versión en espacios de Banach

Teorema de la Aplicación Abierta.

Si f es una aplicación lineal continua, de un espacio normado X en un espacio de Banach Y, suryectiva, entonces la aplicación es abierta

Para una demostración de esta propiedad puede consultarse [Rudin, W Op Cit Pags 46-49]

Teorema de Riesz Fisher

El Teorema de Riesz-Fisher afirma que en los espacios de Hilbert, los funcionales lineales se caracterizan como simples productos internos Lo anterior es una generalización de lo que ocurre en los espacios de dimensión finita

Teorema de Representación de Riesz-Fisher

Sea H un espacio de Hilbert Si H^* indica su dual entonces

$$f \in H^* \Leftrightarrow \exists! h \in H : \forall x \in H, \quad f(x) = \langle x,h \rangle$$

Además $\|f\| = \|h\|$

Aplicaciones Diferenciables

Los conceptos de cálculo diferencial en espacios de Banach son similares a los usuales en los espacios euclidianos Exposiciones detalladas al respecto pueden consultarse en las obras ya mencionadas de Nachbin y Dieudonné y en la mas reciente de Ambrosetti y Prodi [Ambrosetti, A, Prodi, G A Primer of Nonlinear Análisis, Cambridge University Press,Cambridge, 1993] En este apartado se consideran aquellos aspectos del Cálculo Diferencial que son indispensables en el tratamiento de los problemas de extremo

Definición de Aplicación Diferenciable

Sean X, Y espacios de Banach sobre el mismo cuerpo de escalares, A un subconjunto abierto de X, $x_0 \in A$ y $f : A \to Y$ una aplicación Se dice que f es diferenciable en x_0 si existe $L \in L(X,Y)$ tal que

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)-L(h)}{\|h\|}=0 \qquad (1)$$

Si f es diferenciable en x_0, la aplicación L es única, se indica $df(x_0)$ y se llama diferencial (de Frechet) de f en x_0

Si f es diferenciable en x, para todo $x \in A$, y $df : A \to Y$ es continua se dice que f es de clase C^1

Ejemplos básicos de aplicaciones diferenciables lo constituyen las mismas funciones lineales continuas $L \in L(X,Y)$ En efecto en este caso L es diferenciable en todo punto de X y su diferencial coincide con la misma aplicación ya qúe

$$L(x+h)-L(x)=L(h)$$

Otros ejemplos sencillo lo proporcionan las aplicaciones afines continuas Estas tienen la forma $b+L$, con $L \in L(X,Y)$ y $b \in Y$

Observación: La diferenciabilidad de una aplicación en un punto x_0 equivale a lo siguiente Existen, una vecindad $V \subset A$ de x_0, y una aplicación $\omega(x_0,\cdot): V \to Y$ tales que·

$$f(x_0 + h) = f(x_0) + L(h) + \|h\|\omega(x_0, h)$$

y

$$\lim_{h \to 0} \omega(x_0, h) = 0$$

Subsiste el siguiente resultado:

Proposición . Si f es diferenciable en x_0, entonces f es continua en x_0.

Obsérvese que de la definición o de su forma equivalente se deduce que

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

O sea la continuidad de la función en el punto

Véase [Nachbin, Op Cit Prop 29, pag 72]

El conjunto de funciones que son diferenciables en un punto es- con las operaciones usuales-un espacio vectorial y la aplicación de este espacio en $L(X,Y)$ que a cada función f su diferencial $df(x_0)$ es lineal, además estas aplicaciones son estables para la composición de funciones ya que vale la siguiente proposición

Proposición. (Regla de la Cadena) Si X, Y, Z son espacios de Banach, f, g son funciones definidas en subconjuntos abiertos de X y Y a valores en Y y en Z respectivamente y en $x_0 \in X$, f diferenciable , y g es diferenciable en $f(x_0)$, entonces $g \circ f$ es diferenciable en x_0 y $d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$

[Nachbin, Op Cit Prop 36, pags 80-81]

La Regla de la Cadena permite obtener condiciones necesarias para que una función diferenciable tenga un extremo en un punto interior de su dominio En efecto subsiste el siguiente resultado

Proposición (Condición necesaria de extremo)

Sea X un espacio de Banach, A un subconjunto abierto de X, $a \in A$ y f una aplicación diferenciable de A en R tal que $f(a)$ es un extremo relativo de la función Entonces $df(a) = 0$

Demostración:

Sin pérdida de generalidad supóngase que $f(a)$ es un máximo relativo y sea $\delta > 0$ tal que la bola abierta de centro a y radio δ - $B_\delta(a)$ -esté contenida en A y resulte

$f(x) \leq f(a)$

para todo $x \in B_\delta(a)$ Sea $h \in X$ tal que $\|h\| < \delta$, entonces para todo $t \in (-1,1)$ $a + th \in B_\delta(a)$. La aplicación $\Phi : (-1,1) \to R$ definida $\Phi(t) = f(a+th)$ es compuesta de funciones diferenciables, por lo tanto es diferenciable, además tiene un máximo en $t = 0$ Por esto

$$\frac{d}{dt}\Phi(t)\Big|_{t=0} = 0$$

Pero, por el la Regla de la Cadena

$$\frac{d}{dt}\Phi(t)\Big|_{t=0} = df(a)(h)$$

Por tanto $df(a)(h) = 0$ para todo $h \in X$, $\|h\| < \delta$ y esto asegura que $df(a) = 0$.■

Derivadas de Orden Superior-Fórmula de Taylor

Supóngase que $f \in C^1(A,Y)$ y considérese la aplicación $df : A \to L(X,Y)$. Sea $a \in A$, si df es diferenciable en a, se dice que f es dos veces diferenciable en a y se llama segunda diferencial de f en a, a la aplicación $d(df)(a)$, que se indica $d^2 f(a)$

Como $d^2 f(a) : X \to L(X,Y)$ y es lineal y continua, $d^2 f(a) \in L(X, L(X,Y))$ y por tanto es una aplicación bilineal continua (véase [Ambrosetti, A, Prodi, G Op Cit Pags 23-26])

De manera similar se definen las derivadas de orden superior como aplicaciones multilineales continuas de X en Y

Lo anterior permite considerar aplicaciones de clase $C^n(A,Y)$, siempre que los diferenciales del orden indicado sean continuos

Para $f \in C^n(A,Y)$, si $(h,h,...,h) \in X^n$, se pone $d^n f(a)(h,h,...,h) = d^n f(a)(h)^n$

Un hecho importante es que las diferenciales superiores son simétricas, esto no es más que una generalización del Teorema de Schwarz que afirma, bajo hipótesis de continuidad que

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Un importante resultado con respecto a las funciones de clase $C^n(A,Y)$ es el siguiente Teorema de Taylor

Teorema de Taylor.

Sea $f \in C^n(A,Y)$, $a, a+h \in A$ y tales que el intervalo $[a, a+h]$ esté contenido en A, entonces

$$f(a+h) = f(a) + df(a)(h) + ... + \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1} d^n f(a+th)dt(h)^{n-1}$$ [1]

[1] Los integrantes son de Riemann y como es bien conocido las funciones continuas son integrables en este sentido, aun si un rango es un espacio Banach, y esta integral tiene las propiedades usuales de las integrales escalares de Riemann, en particular es válido el Teorema del Calculo

Demostración:

Como que el intervalo $[a, a+h]$ esté contenido en A, si $t \in [0,1]$, entonces $a+th \in A$ Por esto se puede considerar una aplicación de $[0,1]$ en A, poniendo $\varphi(t) = a+th$ Considerando ahora la aplicación $f \circ \varphi$, ésta resulta de clase $C^n(A,Y)$ y

$$(f \circ \varphi)^{(k)}(t) = d^k f(a+th)(h)^k, \qquad k = 1,2,\ldots,n$$

Como se deduce de aplicar repetidamente la regla de la cadena y del hecho que $\varphi'(t)(h) = h$ para todo $t \in [0,1]$

La fórmula se obtiene entonces aplicando integración por partes a partir de la identidad

$$\int_0^1 df(a+th)dt(h) = f(a+h) - f(a)$$

pues $\frac{d}{dt}(f \circ \varphi)(t) = df(a+th)(h)$ ■

Observación: El último término de la formula tiende a 0 cuando h tiende a cero, ya que en las hipótesis del Teorema el mismo puede acotarse con una cantidad del tipo

$$K\|h\|^n$$

con $K > 0$ una oportuna constante

Así como la Regla de la Cadena proporciona condiciones necesarias de extremo para funciones a valores reales, la fórmula de Taylor produce condiciones suficientes de extremo, esto queda de manifiesto en la siguiente proposición

Proposición (Condiciones suficientes de extremo)

Sea X un espacio de Banach, A un subconjunto abierto de X, $a \in A$ y f una aplicación de clase C^2 de A en R tal que $df(a) = 0$ y la forma cuadrática $d^2 f(a)$ sea definida (positiva o negativa) Entonces $f(a)$ es un valor extremo local de la función

Demostración:

Supóngase que $d^2 f(a)$ es definida positiva, o sea que para todo $h \in X - \{0\}$, $d^2 f(a)(h)^2 > 0$ Si h se escoge convenientemente en una vecindad de 0, resultará

$d^2 f(a+th)(h)^2 > 0$ para todo $t \in [0,1]$ Aplicando la fórmula de Taylor resulta

$$f(a+h) = f(a) + \frac{1}{0!}\int_0^1 d^2 f(a+th)(h)^2 dt > f(a)$$

Esta desigualdad asegura que el valor $f(a)$ es un mínimo relativo de la función ■

Derivadas Parciales

Sea $X = X_1 \times X_2$ un producto de espacios de Banach, Y un espacio de Banach y $f : A \to Y$, $A \subset X$ un abierto El hecho que el dominio de la función sea parte de un producto cartesiano permite considerar aplicaciones parciales en las que una de las dos coordenadas del punto del dominio está fija

Es conveniente precisar lo anterior

Sea $a = (a_1, a_2) \in A$ y $A_2 = \{x_2 \in X_2 | (a_1, x_2) \in A\}$ este conjunto es abierto en X_2, pues la aplicación $x_2 \xrightarrow{\varphi} (a_1, x_2)$ es continua y $A_2 = \varphi^{-1}(A)$ De manera análoga se define un abierto $A_1 \subset X_1$ Quedan definidas entonces dos aplicaciones $f_i : A_i \to Y$, $i = 1,2$ así

$f_1(x_1) = f(x_1, a_2) \qquad f_2(x_2) = f(a_1, x_2)$

Si f_i es diferenciable en $x_i = a_i$, su diferencial se llama derivada parcial de f con respecto a x_i indica $D_i f(a_1, a_2)$

Estas derivadas parciales juegan un rol similar a sus homónimas en R^n Esta afirmación es confirmada por los siguientes resultados

Proposición.

Sea $X = X_1 \times X_2$ un producto de espacios de Banach, Y un espacio de Banach, $f : A \to Y$, $A \subset X$ un abierto y f es diferenciable en $(a_1, a_2) \in A$

entonces las aplicaciones $f_i : A_i \to Y, i = 1,2$ son diferenciables en $x_i = a_i, i = 1,2$ y

$$df(a_1,a_2)(h_1,h_2) = D_1 f(a_1,a_2)(h_1) + D_2 f(a_1,a_2)(h_2)$$

Proposición.

Sea $X = X_1 \times X_2$ un producto de espacios de Banach, Y un espacio de Banach, $f : A \to Y$, $A \subset X$ un abierto

Entonces si las las aplicaciones $f_i : A_i \to Y, i = 1,2$ son de clase $C^1(A_i)$, f es diferenciable en $(a_1,a_2) \in A$ y

$$df(a_1,a_2)(h_1,h_2) = D_1 f(a_1,a_2)(h_1) + D_2 f(a_1,a_2)(h_2)$$

(Véase [Nachbin, Op. Cit Pags 106-108])

Funciones Inversas y Funciones Implícitas.

La idea original del Cálculo Diferencial es que las aplicaciones se comportan, al menos localmente como sus diferenciales En este sentido, si el diferencial es una transformación invertible, con inversa continua, lo mismo puede esperarse de la función original Esta percepción es confirmada por el siguiente Teorema, cuya demostración puede consultarse en las obras ya mencionadas, indicando que la misma se apoya fuertemente en las propiedades de contracción indicadas en la primera sección del capítulo, en especial en las observaciones 1, 2 y 3 del Teorema de Homeomorfismo

El Teorema de la función Inversa (local) se enuncia en los siguientes términos

Teorema de la Función Inversa

Sea $f \in C^1(A,B)$, donde A, B son subconjuntos abiertos de dos espacios de Banach X y Y Supóngase que en un punto $x_0 \in A$, $df(x_0)$ posee inversa continua Entonces f es localmente invertible en $y_0 = f(x_0)$, con inversa de clase C^1 Con mayor precisión Existe una vecindad $U \subset A$ de x_0 y una vecindad $V \subset B$ de y_0 tales que

i f es un homeomorfismo de U en V,

ii $f^{-1} \in C^1(V,U)$

iii $df^{-1}(y) = (df(x))^{-1}, \quad x = f^{-1}(y)$

Por último se considera un Teorema sobre funciones implícitas

Teorema (Hildebrandt-Graves)

Sean X, Y, Z son espacios de Banach, A un abierto de $X \times Y$ y $f: A \to Z$ de clase C^1, sea $(a,b) \in A$ y $f(a,b) = 0$. Si $k \to D_Y f(a,b)(k)$ es un isomorfismo de Y en Z, entonces existe una bola $B_X(a,\delta)$ de X y una y solo una función $\phi: B_X(a,\delta) \to Y$ tales que

i $\phi(a) = b$

ii Para todo $x \in B_X(a,\delta)$, $(x,\phi(x)) \in A$ y $f(x,\phi(x)) = 0$

iii ϕ es diferenciable y $d\phi(x) = -(D_Y f(x,\phi(x)))^{-1} \circ D_X f(x,\phi(x))$

Véase por ejemplo [Ambrosetti, A, Prodi, G Op Cit., pags 32-34, para el teorema de función inversa y pags 36-38]

Funciones Diferenciables Reales en Espacios Complejos

Se han considerado solo espacios de Banach reales, esto merece una explicación En esta sección se aclara que por la naturaleza de los problemas considerados, el caso de espacios complejos es en cierto sentido trivial, al menos desde el punto de vista de la teoría de los extremos

El primer resultado en el sentido antes indicado lo proporciona la siguiente proposición

Proposición.

Sea X un espacio de Banach complejo, A un abierto de X y $f : A \to R$ una aplicación diferenciable en A Entonces f es localmente constante, es decir que si $a \in A$ y $B(a,\delta) \subset A$, entonces para todo $x \in B(a,\delta)$, $f(x) = f(a)$

Demostración:

Sea $a \in A$ y $B(a,\delta) \subset A$, y sea $B_\delta = \{\lambda \in C \mid |\lambda| < \delta\}$ Para cada $u \in X$ con $\|u\| = 1$ la aplicación $\lambda \to a + \lambda u$, de B_δ en X es diferenciable, por la regla de la cadena, también es diferenciable la compuesta de ésta con la f, o sea que $\lambda \to f(a + \lambda u)$ es una aplicación holomorfa de un abierto de C en R y como su dominio es conexo, pues es una bola, resulta constante e igual a $f(a)$ para todo $\lambda \in B_\delta$

Ahora bien, si $x \in B(a,\delta)$, existe $u \in X$ con $\|u\| = 1$ y $\lambda \in B_\delta$ tal que $x = a + \lambda u$ y por tanto $f(x) = f(a)$ Como esto es cierto para cada $x \in B(a,\delta)$ se obtiene la tesis de la proposición ■

Por ultimo se considera un Teorema sobre funciones implícita

Observaciones:

1 En realidad, en las condiciones de la proposición anterior f es constante en cada componente conexa de A ya que si B es una componente conexa de A y $x, y \in B$ es posible determinar un número finito de bolas abiertas, $B(u_1), \ldots, B(u_n)$ contenidas en B tales que $x \in B(u_1), y \in B(u_n)$ y $B(u_i) \cap B(u_{i+1}) \neq \phi$ para $i = 1,2,\ldots,n-1$[1] Por la proposición anterior resulta

$f(x) = f(u_1) = \ldots\ldots = f(u_n) = f(y)$

Esto asegura que f es constante en B

2 Como f es localmente constante, entonces $df(x) = 0$ para todo $x \in A$, por esto los criterios sobre derivadas para la búsqueda de extremos son inaplicables en estos casos (X complejo), ya que los mismos no discriminan entre los puntos del dominio por lo que son totalmente inoperantes

3 Una consecuencia, curiosa, de la observación 1 es que para funciones del tipo indicado en la proposición, si $K \subset A$ es compacto, entonces la

---

[1] Para esta y otras afirmaciones sobre conjuntos conexos y componentes conexas de conjuntos abiertos véase Pini, B. (1973) Primo Corso di Analisi Matemática, CLUEB, Bologna

restricción de la función al compacto tiene codominio finito En efecto, como las componentes conexas cubren todo A, también cubren el compacto, por esto existe un número finito de estas tales que $K \subset \bigcup_{k=1}^{n} B_k$ y si $c_1,.....,c_n$ son los valores (constantes) de la función en estas componentes entonces $f(K) \subset \{c_1,....,c_n\}$

CAPITULO II

Extremos Condicionados

Así como en el capítulo 1 se consideraron las condiciones que aseguran que una función dos veces diferenciable posee en un punto interior de su dominio un extremo, en este capítulo se analizarán las condiciones que pueden asegurar esta misma ocurrencia cuando las variables independientes del problema están relacionadas entre sí por medio de vínculos funcionales Este es el llamado problema de los extremos condicionados

Variedades Regulares Y Espacio Tangente.

Sean X, Y dos espacios de Banach reales y $\phi : X \to Y$ una aplicación de clase $C^{(1)}$ y V el espacio nulo (o núcleo) de ϕ

$$V = \{x \in X | \phi(x) = 0\}$$

Se dice que V es una **variedad regular** si para todo $x \in V$ la aplicación $df(x): X \to Y$ es suryectiva (El Teorema de la aplicación abierta asegura que $df(x)$ es una aplicación abierta)

Si V es una variedad regular y $x_0 \in V$, el espacio nulo de $df(x_0)$, $T_V(x_0)$ se llama **espacio tangente** de V en x_0

$$T_V(x_0) = \{h \in X | df(x_0) = 0\}$$

Observaciones:

1. Como ϕ es clase $C^{(1)}$ y $T_V(x_0) = d\phi(x_0)^{-1}(0)$ el espacio tangente de V en x_0 es un subespacio cerrado de X, ya que $df(x_0)$ es lineal y continua

2. Si $X = R^n, Y = R^m$ y $\phi = (\phi_1, \dots, \phi_m)$, la condición de suryectividad de $df(x_0)$ equivale a que para todo $x_0 \in V$, la matriz jacobiana de ϕ tenga rango igual a m

$$rango\begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial \phi_m}{\partial x_1} & \dots & \frac{\partial \phi_m}{\partial x_n} \end{pmatrix}(x_0) = m$$

El espacio tangente determina una variedad afín $x_0 + T_V(x_0)$, llamada variedad (afín) tangente a V en x_0, que se corresponde a los conceptos de recta y plano tangentes a curvas y superficies en la geometría ordinaria y como en estas situaciones, los puntos de la variedad tangente están próximos a los de la variedad regular, al menos en una vecindad oportuna del punto de contacto x_0, estas afirmaciones se precisan en la siguiente proposición

Proposición. Sean X, Y dos espacios de Banach reales y $\phi : X \to Y$ una aplicación de clase $C^{(1)}$ y sea $V = \{x \in X | \phi(x) = 0\}$ una variedad regular Entonces, para todo $h \in T_V(x_0)$ existe $x_h \in V$ tal que $\|x_h - x_0 - h\| = o(\|h\|)$ y para cada $x \in V$ existe $h_x \in T_V(x_0)$ tal que $\|x_0 + h_x - x\| = o(\|x - x_0\|)$

Demostración Sea $X/T_V(x_0)$ el espacio cociente de X con respecto a $T_V(x_0)$ y $A: X/T_V(x_0) \to Y$ la aplicación definida poniendo

$$A(T) = df(x_0)(h), \qquad h \in T, \quad T \in X/T_V(x_0)$$

Antes que nada es conveniente señalar que A es efectivamente una aplicación pues sus valores dependen solo de T y no del elemento $h \in T$ En efecto si $h, h' \in T$ entonces $df(x_0)(h) = df(x_0)(h')$

A es un isomorfismo de $X/T_V(x_0)$ sobre Y, ya que por hipótesis $df(x_0)$ es suryectiva y por otra parte si $A(T) = A(T')$, entonces si $h \in T, h' \in T'$ resulta $df(x_0)(h) = df(x_0)(h')$ y por tanto las clases T y T' coinciden, o sea que la aplicación también es inyectiva También es continua ya que

$$\|A(T)\| = \|df(x_0)(h)\| \leq \|df(x_0)\| \cdot \|h\| \quad \forall h \in T$$

por lo tanto

$$\|A(T)\| \leq \|df(x_0)\| \cdot \|T\|$$

ya que $\|T\| = \inf\limits_{h \in T} \|h\|$

Sea $h \in T_V(x_0)$ y $(u_n),(T_n)$ con $T_n \in X/T_V(x_0), u_n \in T_n$ indiquen las sucesiones definidas como sigue

$$u_0 = 0, \quad T_0 = T_V(x_0)$$

y si se han definido $u_1, \ldots, u_{n-1}; T_1, \ldots, T_{n-1}$ entonces se pone

$$T_n = T_{n-1} - A^{-1}(\phi(x_0 + h + u_{n-1}))$$

y $u_n \in T_n$ tal que

$$\|u_n - u_{n-1}\| \le 2\|T_n - T_{n-1}\|$$

Obsérvese que esto último siempre es posible porque $\|T_n - T_{n-1}\| = \inf_{u \in T_{n-1}} \|u - u_{n-1}\|$ Como $u_{n-1} \in T_{n-1}$ se tiene que $A(T_{n-1}) = df(x_0)(u_{n-1})$ y por tanto $T_{n-1} = A^{-1}(df(x_0)(u_{n-1}))$ y

$$T_n = -A^{-1}(\phi(x_0 + h + u_{n-1}) - df(x_0)(u_{n-1}))$$

análogamente

$$T_{n-1} = -A^{-1}(\phi(x_0 + h + u_{n-2}) - df(x_0)(u_{n-2}))$$

Por lo anterior

$$T_n - T_{n-1} = -A^{-1}(\phi(x_0 + h + u_{n-1}) - \phi(x_0 + h + u_{n-2}) - df(x_0)(u_{n-1} - u_{n-2}))$$

Ahora bien, $(d\phi(x_0 + h + u_{n-2} + t(u_{n-1} - u_{n-2}) - df(x_0))(u_{n-1} - u_{n-2}))$ es una función continua en el intervalo $[0,1]$ por lo que

$$T_n - T_{n-1} = A^{-1}\int_0^1 (d\phi(x_0 + h + u_{n-2} + t(u_{n-1} - u_{n-2}) - df(x_0))(u_{n-1} - u_{n-2}))dt$$

Fijado $\varepsilon > 0$ sea $\delta_\varepsilon > 0$ tal que $\|x - x_0\| < \delta_\varepsilon \Rightarrow \|df(x) - df(x_0)\| < \varepsilon$, lo que puede hacerse por las hipótesis de continuidad de las derivadas Si $r > 0$ con $2r < \delta_\varepsilon$, $\|h\| \le r$ y $\max\{\|u_{n-1}\|, \|u_{n-2}\|\} \le r$ entonces

$$\|u_{n-2}+t(u_{n-1}-u_{n-2})\| = \|tu_{n-1}+(1-t)u_{n-2}\| \leq r$$

Por lo anterior, para cada $t \in [0,1]$ será $\|h+u_{n-2}+t(u_{n-1}-u_{n-2})\| \leq 2r < \delta_\varepsilon$ y por tanto

$$\|d\phi(h+u_{n-2}+t(u_{n-1}-u_{n-2})-d\phi(x_0))\| < \delta_\varepsilon$$

Teniendo en cuenta la representación de $T_n - T_{n-1}$ como integral resulta

$$\|T_n - T_{n-1}\| \leq \|A^{-1}\| \ \varepsilon \ \|u_{n-1}-u_{n-2}\|$$

y

$$\|u_n - u_{n-1}\| \leq 2\|T_n - T_{n-1}\| \leq 2\|A^{-1}\| \cdot \varepsilon \ \|u_{n-1}-u_{n-2}\|$$

Si se escoge ε de manera que

$$2\varepsilon\|A^{-1}\| \leq \frac{1}{2}$$

Se tiene entonces que

$$\|u_n - u_{n-1}\| \leq \frac{1}{2}\|u_{n-1}-u_{n-2}\|$$

De lo anterior sigue que si $\|h\| \leq r$ y $\|u_j\| \leq r$, $j = 1,\ldots.,n-1$ entonces

$\|u_n - u_{n-1}\| \leq \frac{1}{2^{n-1}}\|u_1 - u_0\| = \frac{1}{2^{n-1}}\|u_1\|$ (porque $u_0 = 0$) [1]

y como $u_n = u_1 + (u_2 - u_1) + \ldots\ldots + (u_n - u_{n-1})$ aplicando la desigualdad triangular y la última estimación será

[1] Esta desigualdad asegura que la sucesión (u_n) es de Cauchy, si se puede asegurar que la norma del segundo término se mantiene acotada

$\|u_n\| \leq 2\|u_1\|$

$u_0 = 0$ y $T_1 = T_0 - A^{-1}(\phi(x_0 + h))$ y $\|u_1\| = \|u_1 - u_0\| \leq 2\|T_1 - T_0\| \leq 2\|A^{-1}\|\|\phi(x_0 + h)\|$

como ϕ es diferenciable resulta

$\phi(x_0 + h) = \phi(x_0) + d\phi(x_0)(h) + \omega(h) = \omega(h)$

ya que $x_0 \in V$ y $h \in T_V(x_0)$, por esto

$\|u_1\| \leq 2\|A^{-1}\| \|\omega(h)\|$

Si r se escoge adecuadamente entonces $\|\omega(h)\| \leq \dfrac{1}{4\|A^{-1}\|}\|h\|$ para $\|h\| \leq r$,

por lo que

$\|u_1\| \leq \dfrac{1}{2}\|h\| \leq \dfrac{1}{2}r$

Como ya se señaló, esto asegura que la sucesión (u_n) es de Cauchy y si u es su límite resulta

$\|u\| = \lim\|u_n\| \leq 2\|u_1\| \leq 4\|A^{-1}\| \|\omega(h)\| = o(\|h\|)$ pues $\|\omega(h)\| = o(\|h\|)$

También la sucesión (T_n) en el espacio cociente es de Cauchy porque $\|T_{n+k} - T_n\| \leq \|u_{n+k} - u_n\|$, y si T es su límite (El cociente es completo), resulta

$T = \lim T_n = \lim(T_{n-1} - A^{-1}(\phi(x_0 + h + u_{n-1}))) = T - A^{-1}(\phi(x_0 + h + u))$

Por tanto $A^{-1}(\phi(x_0 + h + u)) = 0$ y como A^{-1} es un isomorfismo resulta

$\phi(x_0 + h + u) = 0$ y $x_0 + h + u \in V$

Si se pone $x_h = x_0 + h + u$ entonces $\|x_h - x_0 - h\| = \|u\| = o(\|h\|)$ De esta manera queda demostrada la primera afirmación

Para la segunda afirmación supóngase que $x_0 + k \in V$ Como

$$\phi(x_0 + k) - \phi(x_0) = d\phi(x_0)(k) + \omega(k)$$

donde $\|\omega(k)\| = o(\|k\|)$ y $\phi(x_0 + k) = \phi(x_0) = 0$ entonces

$$d\phi(x_0)(k) = -\omega(k)$$

Sea $T \in X / T_V(x_0)$ tal que $k \in T$, entonces $A(T) = df(x_0)(k) = -\omega(k)$, por tanto

$$T = -A^{-1}(\omega(k)), \qquad \|T\| \leq \|A^{-1}\| \cdot \|\omega(k)\|$$

Sea $u \in T$ tal que $\|u\| \leq 2\|T\| \leq 2\|A^{-1}\| \; \|\omega(h)\|$ Como $u, k \in T$, $k - u \in T_V(x_0)$ por lo que si se pone $x = x_0 + k,\ k - u = h$ entonces

$$\|x_0 + h - x\| = \|u\| = o(\|k\|) = o(\|x - x_0\|) \ \blacksquare$$

Extremos Condicionados

Planteamiento de los problemas de extremo condicionado.

Existe una amplia clase de problemas de extremo que se puede enfrentar con el método de los multiplicadores de Lagrange [Vease por ejemplo "Lagrange Functions" en la enciclopedia de Matemáticas de Kluwert], pero en este trabajo se considera uno en particular y que se puede formular en los siguientes términos

Sean X,Y dos espacios de Banach reales y $A \subset X$ un abierto Sean $f : A \to R$ y $\phi : A \to Y$ funciones El problema consiste en determinar los extremos de f si la variable independiente $x \in A$ está sujeta a la condición $\phi(x) = 0$ Por ejemplo en el caso de máximo el problema es de la forma

Determinar

$$\sup f(x) \qquad x \in A,$$

si

$$\phi(x) = 0$$

Como se verá este problema puede resolverse, al menos teóricamente si se imponen a las funciones involucradas condiciones oportunas Dado que el enfoque es el del cálculo diferencial, estas condiciones se refieren a las clases de diferencialbilidad, y precisamente, para lograr condiciones necesarias se supondrá que

1 $f \in C^{(1)}$

2 El espacio nulo de ϕ, $V = \{x \in A | \phi(x) = 0\}$ es una variedad regular, por lo que $\phi \in C^{(1)}$ y para todo $x \in V$ $df(x): X \to Y$ es suryectiva

El siguiente resultado proporciona las condiciones necesarias para la existencia de extremos condicionados

Teorema (De Ljusternik O De Los Multiplicadores De Lagrange)

Sean X, Y dos espacios de Banach reales y $A \subset X$ un abierto y $f : A \to R$, $\phi : A \to Y$ funciones que satisfacen las condiciones 1 y 2 Si en $a \in V$ la restricción de f a V tiene un extremo relativo, entonces existe $\lambda \in Y^*$ tal que poniendo $F(x) = f(x) - \lambda(\phi(x))$, $x \in A$ resulta

$$dF(a) = df(a) - \lambda(d\phi(a)) = 0$$

Demostración:

Sea $h \in T_V(a)$ y $df(a)(h) = c \in R$ Supóngase que $c \neq 0$ Por la proposición anterior, si $t \in R$ con $|t|$ suficientemente pequeño, al punto $a + th$ se le puede asociar un punto $a + th + u \in V$ con $\|u\| = o(\|th\|)$ Como f es diferenciable resulta

$$f(a+th+u) - f(a) = df(a)(th+u) + \omega(th+u) = ct + df(a)(u) + \omega(th+u)$$

Pero

$$|df(a)(u)+\omega(th+u)| \leq \|df(a)\|\cdot\|u\|+|\omega(th+u)|$$

y como $\|u\| = o(\|th\|)$

$$|\omega(th+u)| = o(\|th+u\|) = o(\|th\|)$$

por esto si t es suficientemente pequeño, los valores de la diferencia $f(a+th+u)-f(a)$, tendrán el mismo signo del producto ct, lo que quiere decir que tendrán tanto valores positivos como valores negativos, lo que contradice la hipótesis que a sea un punto de mínimo de la restricción de f a V

Así pues, para todo $h \in T_V(a)$ resulta

$$df(a)(h) = 0$$

Hecha esta observación, si $h_1, h_2 \in T \in X/T_V(a)$ entonces $h_1 - h_2 \in T_V(a)$ y esto quiere decir que $df(a)(h_1-h_2)=0$, o lo que es lo mismo que $df(a)(h_1) = df(a)(h_2)$ Esto permite definir una aplicación $\alpha : X/T_V(a) \to R$ por medio de la fórmula

$$\alpha(T) = df(a)(h), \qquad T \in X/T_V(a), \quad h \in T$$

Obviamente esta aplicación es lineal y además

$$\|\alpha(T)\| = \|f(a)(h)\| \leq \|df(a)\| \; \|h\| \qquad \forall h \in T \qquad \text{m}$$

luego

$$\|\alpha(T)\| \leq \|df(a)\| \; \|T\|$$

así pues, α es continua Pero, con la notación utilizada en la proposición anterior, $A(T) = d\phi(a)(h)$, por lo que $T = A^{-1}(d\phi(a)(h))$, por lo tanto

$$df(a)(h) = \alpha(T) = \alpha(A^{-1}(d\phi(a)(h)) \qquad \forall h \in T$$

y como las clases de equivalencia constituyen una partición de X, la última igualdad vale en todo el espacio, es decir que

$$df(a) = \alpha \circ A^{-1} \circ d\phi(a)$$

Si se pone $\lambda = \alpha \circ A^{-1}$ y $F(x) = f(x) - \lambda(\phi(x)),\ x \in A$ resulta

$$dF(a) = df(a) - \alpha \circ A^{-1} \circ d\phi(a) = 0 \quad \blacksquare$$

Ejemplos y aplicaciones del Teorema de Lujsternik

1 Sean $f_0, f_1, \ldots, f_n$ funciones reales definidas en un espacio de Banach X de clase $C^{(1)}$

Poniendo $f = (f_1, \ldots, f_n)$, sea a un punto de extremo de f_0 en el conjunto

$$V = \{x \in V \mid f(x) = 0\}$$

que se supone es una variedad regular (en a) En este caso, $Y = R^n$ y $(R^n)^* = R^n$ Como los funcionales lineales de R^n son de la forma $\lambda(\xi_1, \ldots, \xi_n) = \lambda_1\xi_1 + \cdots + \lambda_n\xi_n$ para oportunos valores de $\lambda_1, \ldots, \lambda_n \in R$, el Teorema de Lujsternik asegura que existen tales valores, llamados **multiplicadores de Lagrange**, para los cuales

$$df(a) = \sum_{j=1}^{n} \lambda_j df_j(a)$$

Esto no es otra cosa que el Teorema clásico de Lagrange sobre multiplicadores en los casos en que se da un número finito de restricciones (a valores reales)

La condición que la variedad $V = \{x \in X \mid f(x) = 0\}$ sea regular en a quiere decir que los funcionales $df_1(a), \ldots, df_n(a)$ son linealmente independientes, como estas funciones son de clase $C^{(1)}$ existe entonces una vecindad de a tal que en todos sus puntos se cumple la condición de regularidad

En efecto sea $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$, entonces $\lambda_1 df_1(a) + \cdots + \lambda_n df_n(a) \neq 0$

Como la aplicación $x \to \sum_{j=1}^{n} \lambda_j df_j(x)$ es continua, existe una vecindad $B(a,\delta)$ tal que para todo $x \in B(a,\delta)$ resulta $\lambda_1 df_1(x) + \cdots + \lambda_n df_n(x) \neq 0$ y como esto es cierto para todos los $(\lambda_1,\ldots,\lambda_n) \neq (0,\ldots,0)$, los vectores $df_1(x),\ldots,df_n(x)$ son linealmente independientes y la variedad es regular en todos los puntos de $B(a,\delta)$

Si $X = R^m$, $m > n$, la condición de regularidad en un punto equivale a que el rango de la matriz jacobiana de f tenga rango m, y si esto se da en un punto obviamente también se da en una vecindad del punto en cuestión

2 Sea $X = \{x \in (C^{(1)}([a,b];R))^n \,|\, x(a) = x(b)\}$ Sean $F_0, F_1, \ldots, F_k$ funciones de $[a,b] \times R^n \times R^n$ en R de clase $C^{(1)}$ y sea

$$f_j(x) = \int_a^b F_j(t, x(t), x'(t))dt; \qquad j = 0,1,\ldots,k; \quad x \in X$$

Supóngase que $df_1(x),\ldots,df_k(x)$ sean linealmente independientes para cada $x \in X$

Sean $c_1,\ldots,c_k \in R$ y $\phi : X \to R^k$ la función definida poniendo

$$\phi(x) = (f_1(x) - c_1, \ldots, f_k(x) - c_k); \qquad x \in X$$

y

$$V = \{x \in X \,|\, \phi(x) = 0\}$$

Pues bien si en $x^0 \in X$ la función f_0/V tiene un mínimo local, entonces existen $\lambda_1,\ldots,\lambda_k \in R$ tales que x^0 es un extremo local de

$$f_0 - \sum_{j=1}^{k} \lambda_j f_j$$

(El problema de minimizar f_0/V se llama Problema isoperimétrico de Lagrange)

Las condiciones que se dan aquí sobre los diferenciales de las restricciones son distintas a las planteadas en el Teorema de Ljusternik, en cuanto se postula la independencia lineal de los diferenciales de las restricciones, pero como se muestra a continuación, estas condiciones son equivalentes

Proposición. Sea X un espacio de Banach y $\Lambda_1,\ldots,\Lambda_n \in X^*$ linealmente independientes, entonces la aplicación $\Lambda : X \to R^n$ definida

$$\Lambda(x) = (\Lambda_1(x),\ldots,\Lambda_n(x)) \qquad x \in X$$

es suryectiva

Demostración: Sea N la intersección de los núcleos de las aplicaciones $\Lambda_i, i = 1,\ldots,n$, es decir

$$N = \{x \in X \mid \Lambda_i(x) = 0, i = 1,\ldots,n\}$$

N es un subespacio cerrado de X El espacio cociente X/N es isomorfo a $\Lambda(X)$, y por tanto es de dimensión finita $m \leq n$

Supóngase que $m < n$ y sea $T_1 = e_1 + N, \ldots, T_m = e_m + N$ una base del cociente, entonces los vectores $\Lambda(T_1), \ldots, \Lambda(T_m)$ generan $\Lambda(X)$ Por lo anterior la matriz de la transformación Λ con respecto a la base escogida en el cociente y la usual en R^n, o sea

$$\begin{pmatrix} \Lambda_1(e_1) & \cdots & \Lambda_1(e_m) \\ \vdots & \ddots & \vdots \\ \Lambda_n(e_1) & \cdots & \Lambda_n(e_m) \end{pmatrix}$$ [1]

tiene rango m y sin pérdida de generalidad podemos suponer que las primeras m filas son linealmente independientes y las restantes son combinaciones lineales de éstas, así

$$\Lambda_i(e_j) = \sum_{h=1}^{m} \lambda_h \Lambda_h(e_j), \quad i = m+1, \ldots, n; \quad j = 1, \ldots, m$$

De estas igualdades se deduce que para todo $x \in X$ resulta

$$\Lambda_i(x) = \sum_{h=1}^{m} \lambda_h \Lambda_h(x) = (\sum_{h=1}^{m} \lambda_h \Lambda_h)(x)$$

O sea que por ejemplo, Λ_{m+1} es combinación lineal de $\Lambda_i, i = 1, \ldots, m$

$\rightarrow|\leftarrow$

Por tanto $m = n$ y Λ es suryectiva ■

---

[1] Si $T \in X/N$ entonces $\Lambda(x) = \Lambda(y)$ $\forall x, y \in T$, por esto $\Lambda_i(T_j) = \Lambda_i(e_j)$ y la matriz es la que se exhibe

Observación: Una consecuencia de la proposición anterior es el siguiente criterio "numérico" de independencia lineal de funcionales lineales

Para que n funcionales lineales (continuos o no), $\Lambda_1,\ldots,\Lambda_n : X \to R$ sean linealmente independientes es necesario y suficiente que existan n puntos de X, $e_1,\ldots,e_n$ tales que

$$\det\left(\begin{pmatrix} \Lambda_1(e_1) & \cdots & \Lambda_1(e_n) \\ \vdots & \ddots & \vdots \\ \Lambda_1(e_1) & \cdots & \Lambda_n(e_n) \end{pmatrix}\right) \neq 0 \ .$$

Por supuesto, los $e_1,\ldots,e_n$ deben escogerse linealmente independientes

Interpretación De Los Multiplicadores De Lagrange

Si en un punto $a \in V$, la función f/V tiene un extremo relativo y se cumplen todas las condiciones del Teorema de Ljusternik, por estas mismas condiciones los valores extremos en una vecindad de a dependen del valor $\phi(x) = b \in Y$, en una vecindad de 0 en Y, en el sentido que, por ejemplo si

$$\min_{x \in V_b} f(x) = z^* \quad V_b = \left\{ x \in X \,\middle|\, \phi(x) = b \right\}$$

entonces z^* es una función diferenciable de b y

$$d_b z^* = \lambda^*$$

donde λ^* es el correspondiente valor del multiplicador de Lagrange

Así en muchos problemas aplicados, las z tienen un significado de ganancias o pérdidas, mientras que ϕ señala una distribución de recursos asignados para generar estas pérdidas o ganancias El multiplicador es entonces un factor de la variación unitaria de estas asignaciones Para este comentario véase la "Enciclopaedia" de Kluwert ya citada

Condiciones Suficientes De Extremo Condicionado

Las condiciones planteadas en el Teorema de Ljusternik son condiciones necesarias para la existencia de un extremo condicionado y son de primer orden, es decir involucran las primeras derivadas de las funciones involucradas, tanto la función objetivo como la función que proporciona las restricciones de la variable independiente Como en el caso de las funciones numéricas, pueden obtenerse condiciones de segundo orden que aseguren que en un punto en el que se den estas condiciones necesarias, se presente efectivamente el extremo, para esto debe recurrirse a los diferenciales segundos En los casos de dimensión finita, estas condiciones se reducen a que el diferencial segundo de la función de Lagrange sea una forma cuadrática definida, cuando se restringe al espacio tangente a la restricción en el punto de extremo [Véase por ejemplo, el primer volumen de la obra Fundamentals of Mathematical Análisis, de V I Ilyin y E G Poznyak] Algo semejante ocurre en la situación más general que se analiza en este trabajo

Considérese como antes dos espacios de Banach reales, X,Y y en cada uno de ellos dos normas, $\| \|'_X, \| \|'_Y$ dominadas por las normas originales de los espacios, es decir, que para un $C>0$ resulta

$$\|x\|'_X \leq C\|x\|_X, \|y\|'_Y \leq C\|y\|_Y \quad \forall x \in X, y \in Y$$

En lo sucesivo se omitirán los subíndices X,Y en la medida que esta omisión no lleve a confusiones

Supóngase que las siguientes hipótesis se verifican

I $f: X \to R, \phi: X \to Y$ son de clase C^2 en una vecindad U de $a \in X$ y

(a) $\left|d^2 f(a)(h,k)\right| \leq c_0 \|h\|' \|k\|'$, $\left|d^2\phi(a)(h,k)\right| \leq c_1 \|h\|' \|k\|'$ para todo $h,k \in X$ con c_0, c_1 constantes positivas,

(b) $\left|d^2 f(x)(h,k) - d^2 f(a)(h,k)\right| \leq \eta_0(\|x-a\|)\|h\|' \|k\|'$ y $\left|d^2\phi(x)(h,k) - d^2\phi(a)(h,k)\right| \leq \eta_1(\|x-a\|)\|h\|' \|k\|'$ para todo $x \in U$ y para todo $h,k \in X$, donde η_0, η_1 $R^+ \to R^+$ son funciones localmente acotadas con límite 0 a la derecha en el origen, o sea $\lim_{t \to 0^+} \eta_i(t) = 0, i = 0,1$

II (a) Es $X = \ker(d\phi(a)) \oplus X_1$ con X_1 un espacio de Banach y $d\phi(a): X_1 \underset{sury}{\overset{1-1}{\to}} Y$, (b) $d\phi(a)^{-1}$ $Y \to X_1$ es continua con respecto a la topología inducida por la norma $\| \|'$

(Si $\| \|' = \|\cdot\|$, las condiciones (a) y (b) de I y (b) de II se deducen de las otras hipótesis)

Subsiste entonces el siguiente Teorema

Teorema.

Sea $\phi(a) = 0$ y se verifiquen las condiciones I y II Sean $\delta_0 > 0$ y $\lambda \in Y^*$, continua con respecto a la norma $\| \|'$, tales que

(d1) $df(a) - \lambda(d\phi(a)) = 0$

(d2) $d^2 f(a)(h,h) - \lambda(d^2\phi(a)(h,h)) \geq \delta_0 \|h\|'^2 \quad \forall h \in \ker(d\phi(a))$

Entonces $f(a)$ es un mínimo local de f/V; $V = \{x \in X | \phi(x) = 0\}$

Demostración: Considérese la aplicación $\Phi \cdot \ker(d\phi(a)) \oplus X_1 \to Y$ definida de la siguiente manera

$$h \in \ker(d\phi(a)), x_1 \in X_1, \quad \Phi(h, x_1) = \phi(a + h + x_1)$$

Como $\Phi(0,0) = 0$ y la derivada parcial $d_{x_1}\Phi(0,0) = d\phi(a)$ es un isomorfismo de X_1 en Y por el Teorema de la Función Implícita, existe una y solo una $h \to x_1(h)$ definida y continua en una vecindad del 0 de $\ker(d\phi(a))$ tal que $\Phi(h, x_1(h)) = 0$ en dicha vecindad Pues bien

$$\lim_{\|h\| \to 0} \frac{\|x_1(h)\|'}{\|h\|'} = 0$$

En efecto Por la fórmula de Taylor se tiene que

$$0=\phi(a+h+x_1(h))-\phi(a)=d\phi(a)(h+x_1(h))+\int_0^1(1-t)d^2\phi(a+t(h+x_1(h)))(h+x_1(h)^2dt\ {}^2$$

Dado que $d\phi(a)(h)=0$, teniendo en cuenta la (b) de II y las condiciones (a) y (b) de I resulta

$$\|x_1(h)\|' \le K_0\,\|d\phi(a)(h+x_1(h))\|' \le K_0\int_0^1\|d^2\phi(a+t(h+x_1(h))(h+x_1(h))^2\|dt$$

$\le K_1(\|h+x_1(h)\|')^2$, luego

$$\|x_1(h)\|' \le K_1(\|h\|'+\|x_1(h)\|')^2 = K_1(\|h\|'^2+2\|x_1(h)\|'\,\|h\|'+\|x_1(h)\|'^2)$$

dividiendo esta última desigualdad por $\|h\|'$ se tiene lo siguiente

$$\frac{\|x_1(h)\|'}{\|h\|'}\le K_1(\|h\|'+2\|x_1(h)\|'+\|x_1(h)\|'\frac{\|x_1(h)\|'}{\|h\|'})$$

despejando el cociente resulta

$$\frac{\|x_1(h)\|'}{\|h\|}(1-k_1\|x_1(h)\|')\le k_1(\|h\|'+2\|x_1(h))\|$$

hora bien cuando $\|h\|\to 0$, las condiciones sobre las normas $\|\ \|'$, aseguran que los dos términos dentro del paréntesis también tienen límite 0 y por tanto

---

[2] Se conviene que en el caso de una aplicación bilineal, β, se denotará $\beta(h,h)=\beta(h)^2$

$$\lim_{\|h\|\to 0} \frac{\|x_1(h)\|'}{\|h\|'} = 0$$

Considérese ahora la función de Lagrange $F(x) = f(x) - \lambda(\phi(x))$ Por las condiciones dadas esta es una función de clase C^2, tal que $dF(a) = 0$ Por la fórmula de Taylor resulta

$$F(x) - F(a) = \int_0^1 (1-t) d^2F(a + t(x-a))(x-a)^2 dt$$

$$= \frac{1}{2} d^2F(a)(x-a)^2 + \int_0^1 (1-t)[d^2F(a+t(x-a))(x-a)^2 - d^2F(a)(x-a)^2]dt$$

Teniendo presente que $d^2F(x) = d^2f(x) - \lambda(d^2\phi(x))$, que λ es continua con respecto a ambas normas y la condición (b) de I se tiene que para una función $\eta \cdot R^+ \to R^+$, convergente a cero a la derecha en el origen resulta

$$[d^2F(a+t(x-a) - d^2F(a)](x-a)^2 \geq -\eta(\|x-a)\|\|x-a\|^2$$

Multiplicando esta desigualdad por $(1-t) \geq 0$ e integrando se obtiene

$$F(x) - F(a) \geq \frac{1}{2} d^2F(a)(x-a)^2 - \frac{\eta}{2}(\|x-a\|)\|x-a\|^2$$

Obsérvese que $F(x) - F(a) = f(x) - f(a) - \lambda(\phi(x) - \phi(a))$

Si $h \in \ker(d\phi(a))$ y es suficientemente pequeño resulta

$$f(a+h+x_1(h)) - f(a) = F(a+h+x_1(h)) - F(a)$$

porque $\phi(a+h+x_1(h)) = \phi(a) = 0$, por tanto

$$f(a+h+x_1(h))-f(a)\geq\frac{1}{2}d^2F(a)(h+x_1(h))^2-\frac{\eta}{2}(\|h+x_1(h)\|\ \|h+x_1(h)\|'^2$$
$$=\frac{1}{2}[d^2F(a)(h)^2+2d^2F(a)(h,x_1(h))+d^2F(a)(x_1(h))^2]-\eta(\|h+x_1(h)\|\ \|h+x_1(h)\|'^2$$

Por la hipótesis (d_2) $d^2F(a)(h)^2\geq\delta_0\|h\|'^2$, por otra parte

$$\left|d^2F(a)(h,x_1(h))\right|\leq\left|d^2f(a)(h,x_1(h))\right|+\left|\lambda(d^2\phi(a)(h,x_1(h)))\right|$$

Teniendo en cuenta las con$\{x_n\}$diciones de continuidad expresadas por la (a) de I lo mismo que la continuidad de λ se puede concluir que para una oportuna constante $\tilde{C}$ resulta

$$\left|d^2F(a)(h,x_1(h))\right|\leq\tilde{C}\|h\|'\ \|x_1(h)\|'$$

En definitiva

$$f(a+h+x_1(h))-f(a)\geq\frac{1}{2}\delta_0\|h\|'^2+\frac{1}{2}\delta_0\|x_1(h)\|'^2-\tilde{C}\|h\|'\ \|x_1(h)\|'-\eta(\|h+x_1(h)\|)\|h+x_1(h)\|'$$
$$=\frac{\|h\|'^2}{2}[\frac{\delta_0}{2}+\frac{\delta_0}{2}\frac{\|x_1(h)\|'^2}{\|h\|'^2}-\tilde{C}\frac{\|x_1(h)\|'}{\|h\|'^2}-\eta(\|h+x_1(h)\|)(1+\frac{\|x_1(h)\|'}{\|h\|'^2})]$$

Como $\lim\limits_{\|h\|\to 0}\dfrac{\|x_1(h)\|'}{\|h\|'}=0$, para h suficientemente pequeño resulta

$$f(a+h+x_1(h))-f(a)\geq\frac{1}{4}\delta_0\|h\|'^2$$

Lo anterior implica que $f(a)$ es un mínimo relativo de f/V En efecto, en caso contrario existiría una sucesión en X, $\{x_n\}$ tal que $x_n\to a$ y para todo

$n, \phi(x_n)=0, \quad f(x_n)-f(a)<0$, pero $x_n = a+h_n+x_1(h_n)$ a partir de un índice oportuno, contradiciendo el último acotamiento Así el Teorema queda demostrado ■

Observación: Si $X=R^n$ y $Y=R^m$, como todas las normas en estos espacios son equivalentes las condiciones I se deben cumplir para la norma escogida, pero estas condiciones no son otra cosa que la continuidad de la segunda derivada como forma bilineal y la continuidad de la transformación que a los puntos del dominio asigna la segunda diferencial En cuanto a las condiciones II, basta tomar $X_1 = X/\ker(d\phi(a))$

Una aplicación del Teorema sobre condiciones
Suficientes de extremo condicionado

Las condiciones introducidas con anterioridad al Teorema que establece las condiciones suficientes de extremo condicionadas parecen un tanto artificiales, en el siguiente ejemplo se puede apreciar que en muchos problemas éstas tienen un carácter natural

Sea $[a,b]$ un intervalo compacto de R y sean

- $X = C([a,b];R^n)\times C([a,b];R^r)$,
- $Y = C([a,b];R^n)$

Con $\|\cdot\|$ se indica la norma usual en $C([a,b];R^m)$ y $\|\,\|'$ denote la norma inducida en este espacio cuando se le considera como subespacio de $L^2([a,b];R^m)$, o sea

- $\|x\|' = (\int_a^b \langle x(t), x(t)\rangle)^{\frac{1}{2}}$

Para todo $x \in C([a,b];R^m)$ resulta $\|x\|' \leq \sqrt{b-a}\|x\|$, o sea que se cumple con la condición que exista $C > 0$ tal que $\|x\|' \leq C\|x\|$

Sean $F: R \times R^n \times R^r \to R$ y $g: R \times R^n \times R^r \to R^n$ dos aplicaciones de clase C^2

Se llama control un elemento $C([a,b];R^r)$ y se indicará con u Los elementos de $C([a,b];R^n)$ se denotarán con la letra x

El problema en consideración es el siguiente

Minimizar el funcional

$$f(x,u) = \int_a^b F(t, x(t), u(t))dt$$

Con $u \in C([a,b];R^r)$ *y* $x \in C([a,b];R^n)$ *tal que*

$$x(t) = x_0 + \int_a^t g(\tau, x(\tau), u(\tau))d\tau \quad \textit{(Sistema de control)}$$

donde x_0 *es un elemento determinado de* R^n

Supóngase que (x^0, u^0) satisface el sistema de control Para simplificar las notaciones $\frac{\partial F}{\partial x_j}(t)$ indique a $\frac{\partial F}{\partial x_j}(t, x^0(t), u^0(t)))$ y análogamente para $\frac{\partial F}{\partial u_k}(t), \frac{\partial g}{\partial x_j}(t), \frac{\partial g}{\partial u_k}(t)$

Sea $\psi = (\psi_1, \ldots, \psi_n)$ una solución del problema de Cauchy

$$\begin{cases} -\psi_i'(t) = \sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(t)\psi_j(t) + \frac{\partial F}{\partial x_i}(t) \\ \psi_i(a) = 0 \qquad i = 1,2,\ldots,n \end{cases}$$ [3]

[3] Este es un problema de Cauchy para un sistema de ecuaciones lineales no homogéneo con coeficientes continuos y por lo tanto admite siempre soluciones, véase por ejemplo [Coddington and Levison Theory of Ordinary Differential Equations, Tata Mcgraw-Hill, New Delhi, págs 74-75]

Póngase también

$$H(t,x,u)=\sum_{j=1}^{n}\psi_j(t)g_j(t)+F(t,x,u)$$

Las siguientes hipótesis garantizan el cumplimiento de las condiciones (d_1) y (d_2) de la proposición anterior

Hipótesis 1 $\frac{\partial H}{\partial u_k}(t,x^0(t),u^0(t))=0 \qquad \forall t\in[a,b];\, k=1,2\ldots,r$

Si en lugar de $\frac{\partial^2 H}{\partial x_i \partial x_j}(t,x^0(t),u^0(t))$ se pone $\frac{\partial^2 H}{\partial x_i \partial x_j}(t)$ y similares entonces

Hipótesis 2 Existe $\delta_0>0$ tal que

$$\int_a^b\Big(\sum_{i,j=1}^{n}\frac{\partial^2 H}{\partial x_i\partial x_j}(t)x_i(t)x_j(t)+2\sum_{j=1}^{n}\sum_{k=1}^{r}\frac{\partial^2 H}{\partial x_j\partial u_k}(t)x_j(t)u_k(t)+$$
$$\sum_{h,k=1}^{r}\frac{\partial^2 H}{\partial u_h\partial u_k}(t)u_h(t)u_k(t)\quad\Big)dt\geq\delta_o\|u\|^r$$

Para todo $(x,u)\in X$ tal que

$$\begin{cases} x_i'(t)=\sum_{j=1}^{n}\frac{\partial g_i}{\partial x_j}(t)x_j(t)+\sum_{k=1}^{r}\frac{\partial g_i}{\partial u_k}(t)u_k(t) \\ x_i(a)=0 \qquad i=1,2,\ldots,n\end{cases}\qquad (*)$$

Pues bien, si se cumplen estas dos hipótesis, el punto (x^0,u^0) es localmente optimal, o sea que existe $\varepsilon_0>0$ tal que para toda solución (x,u) del sistema de control tal que

$0<\|x-x^0\|+\|u-u^0\|<\varepsilon_0$ resulta

$$\int_a^b F(t,x(t),u(t))\geq\int_a^b F(t,x^0(t),u^0(t))$$

El sistema ($*$) puede escribirse sintéticamente en la forma vectorial siguiente

$$\begin{cases} x'(t)=\sum_{j=1}^{n}\frac{\partial g}{\partial x}(t)x(t)+\sum_{k=1}^{r}\frac{\partial g}{\partial u}(t)u(t) \\ x(a)=0\end{cases}\qquad \begin{pmatrix}*\\ *\end{pmatrix}$$

Como se ha señalado, se trata de ver que estas condiciones aseguran el cumplimiento de las del teorema anterior

Ahora bien

$$df(x^0,u^0)(x,u)=\int_a^b\left(\sum_{j=1}^{n}\frac{\partial F}{\partial x_j}(t)x_j(t)+\sum_{k=1}^{r}\frac{\partial F}{\partial u_k}(t)u_k(t)\right)dt$$

Si

$$\phi(x,u)(t)=x(t)-x_0-\int_a^t g(s,x(s),u(s))ds$$

El diferencial de esta función es

$$d\phi(x^0,u^0)(x,u)(t)=x(t)-\int_a^t\left(\sum_{j=1}^{n}\frac{\partial g}{\partial x_j}(s)x_j(s)+\sum_{k=1}^{r}\frac{\partial g}{\partial x_k}(s)u_k(s)\right)ds$$

El espacio nulo de esta transformación, $\ker\{d\phi(x^0,u^0)\}$, es precisamente el conjunto de los pares (x,u) que son soluciones del sistema $(*)$

Póngase $X_1=Y\times\{0\}\subset X$ Todo $(x,u)\in X$ se puede escribir de la forma

$$(x,u)=(\tilde{x},u)+(x-\tilde{x},0)$$

donde $\tilde{x}$ es una solución de la $(*)$ Por otra parte si $(x,u)\in\ker\{d\phi(x^0,u^0)\}\cap X_1$ tiene que ser $u=0$ y como x es una solución de la $(*)$, también $x(a)=0$ y $x'=\sum_{j=1}^{n}\frac{\partial g}{\partial x_j}(t)x_j$ por lo que $x=0$ y $X=\ker\{d\phi(x^0,u^0)\}\oplus X_1$ [4]

Considérese ahora la restricción $d\phi(x^0,u^0)/X_1$ Si $y\in Y$ la ecuación de Volterra

$$y(t)=x(t)-\int_a^t\sum_{j=1}^{n}\frac{\partial g}{\partial x_j}(s)x_j(s)ds$$

tienen una y solo una solución para cada $y\in Y$ y por tanto $d\phi(x^0,u^0)/X_1:X_1\to Y$ es un isomorfismo

[4] La única solución del sistema $x'=\sum_{j=1}^{n}\frac{\partial g}{\partial x_j}(t)x_j$ tal que $x(a)=0$ es $x=0$

Por otra parte si $\tilde{G}$ es el núcleo resolvente de la ecuación de Volterra considerada resulta

$$x(t) = y(t) + \int_a^t \tilde{G}(t,\tau)y(\tau)d\tau$$

De esta ecuación se deduce que $\|x\|' \leq C\|y\|'$

Todo lo anterior indica que se cumplen las condiciones preliminares del teorema anterior

Condición (d_1) Sea $\lambda \in Y^*$ el funcional definido por

$$\lambda(x) = -\int_a^b \sum_{j=1}^n \psi_j'(t)x_j(t)dt$$

Donde ψ es la solución de $(*)$ Entonces

$$df(x^0,u^0)(x,u) - \lambda(d\phi(x^0,u^0))(x,u) = \int_a^b (\sum_{j=1}^n \frac{\partial F}{\partial x_j}(t)x_j(t) + \sum_{k=1}^r \frac{\partial F}{\partial u_k}(t)u_k(t))dt +$$
$$+\int_a^b \sum_{j=1}^n \psi_j'(t)[x_j(t) - \int_a^t (\sum_{k=1}^n \frac{\partial g_j}{\partial x_k}(s)x_k(s) + \sum_{h=1}^r \frac{\partial g_r}{\partial u_h}(s)u_h(s))ds]dt$$

Invirtiendo el orden de integración, teniendo presente que $\psi(a)=0$ y considerando $(*)$ y la Hipótesis 1 se obtiene

$$df(x^0,u^0)(x,u) - d\phi(x^0,u^0)(x,u) = \int_a^b [\sum_{j=1}^n (\psi_j'(t) + \sum_{k=1}^n \frac{\partial g_k}{\partial x_j}(t)\psi_k(t))x_j(t) +$$
$$\sum_{j=1}^r (\sum_{k=1}^n \frac{\partial g_k}{\partial u_j}(t)\psi_k(t) + \frac{\partial F}{\partial u_j}(t)u_j(t))]dt = 0$$

Esto no es otra cosa que la condición (d_1) del Teorema

En cuanto a la condición (d_2) resulta

$$d^2 f(x^0,u^0)(x,u)^2 - \lambda(d^2\phi(x^0,u^0)(x,u)^2) =$$
$$= \int_a^b [\sum_{i,j=1}^n (\sum_{l=1}^n \frac{\partial^2 g_l}{\partial x_i \partial x_j}(t)\psi_l(t) + \frac{\partial^2 F}{\partial x_i \partial x_j}(t))x_i(t)x_j(t) +$$
$$2\sum_{j=1}^n \sum_{k=1}^r (\sum_{l=1}^n \frac{\partial^2 g_l}{\partial x_j \partial u_k}(t)\psi_l(t) + \frac{\partial^2 F}{\partial x_j \partial u_k}(t))x_j(t)u_k(t) +$$
$$\sum_{h,k=1}^r (\sum_{l=1}^n \frac{\partial^2 g_l}{\partial u_h \partial u_k}(t)\psi_l(t) + \frac{\partial^2 F}{\partial u_h \partial u_k}(t))u_h(t)u_k(t)]dt$$

Por la Hipótesis 2, esta variación resulta $\geq \delta_0 \int_a^b \sum_{k=1}^{r} [u_k(t)]^2 dt$ para todos los pares (x,u) del núcleo $\ker\{d\phi(x^0,u^0)\}$

Por otra parte si $(x,u) \in \ker\{d\phi(x^0,u^0)\}$ entonces

$$x(t) = \int_a^t [\sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(s) x_j(s) + \sum_{k=1}^{r} \frac{\partial g}{\partial u_k}(s) u_k(s)] ds$$

y por tanto, para una oportuna constante $k_1 > 0$

$$\sum_{j=1}^{n} [x_j(t)]^2 \leq k_1 \int_a^t (\sum_{j=1}^{n} [x_j(s)]^2 + \sum_{k=1}^{r} [u_k(s)]^2) ds$$

Del Lema de Gronwall [5] se deduce que $\|x\|' \leq k_2 \|u\|'$, para una constante $k_2 > 0$

De esto se deduce finalmente que para todo $(x,u) \in \ker\{d\phi(x^0,u^0)\}$

$$d^2 f(x^0,u^0)(x,u)^2 - d^2\phi(x^0,u^0)(x,u)^2 \geq \delta_0'(\|x\|' + \|u\|')$$

Cumpliéndose así la segunda condición del Teorema , lo que asegura que efectivamente se tiene el mínimo local en (x^0,u^0).

---

[5] La versión del Lema de Gronwall invocada es la siguiente

"*Si* $f(t) \leq \beta(t) + \int_\tau^t \alpha(s) f(s) ds, \quad t \in [t, t+a] \quad a > 0$ entonces

$f(t) \leq \beta(\tau) e^{\int_\tau^t \alpha(s) ds} + \int_\tau^t \beta'(s) e^{\int_s^t \alpha(u) du}$ "

BIBLIOGRAFÍA

Alpha C. Chiang:(1998) "Métodos Fundamentales de Economía Matemátıca" Tercera Edıcıón Mc Graw-Hıll .

Ambrosetti, A; Prodi, G (1993) A Pnmer of Nonlınear Analysıs, Cambndge Unıversıty Press, Cambndge

Coddington and Levison (1955) Theory of Ordınary Dıfferentıal Equatıons, Mc Graw-Hıll, New York

Dieyfus, S. (1965) : Dınamyc Propıety and the Calculus of Vanatıon, Academıc Press, New York

Dieudonné, J (1960) Éléments d'Analyse 1 Capítulo VIII Gauthıer-Vıllars, Parıs

Edwin Krayszing : Untroductory Funcıonal Analysıs Wıth Applıcatıons

Intriligator, Michael D. (1971) : "MathematıcaL Optımızatıon and Economıc

Ilyin, V. I. , E. G Poznyak, E. G (1982), Fundamentals of Mathematıcal Analysıs, IR, Moscou

Nachbin; Leopoldo (1976) Introducáo A Analıse Funcıonal Espacos de Banach e Cálculo dıferencıal (Segunda parte) Monografías Matemátıcas de La O E A Washıngton

Nashed, M. Z (1971) Dıfferentıabılıty and Related propnetıes of Nonlınear Operators Some Aspects of the role of Dıfferentıals ın Nonlınear Analysıs, Academıc press

Pini, Bruno (1973): Prımo Corso dı Analısıs Matematıca, CLUEB, Bologna

Pini, Bruno:(1982) Lezıonı dı Analısıs Matematıca Dı secondo Lıvello Parte Prıma, edıtıce CUEB, Bologna

Pini, Bruno (1972) Secondo Corso dı Analısıs Matematıca/Parte I Cooperatıva Lıbrarıa Unıversıtarıa – Bologna

Supper, Patric (1960): Axıomatıc set theory, D Van Nostrad Company, Inc Prınceton

Rudin, Walter (1973) Functıonal Analysıs Tata Mc Graw Hıll, 1973

Varsan, C. : Nonlınıear Functıonal Analysıs and Applıcatıons to Optımal Control Theory, IAEA-SMR, 17/30-