Análisis Funcional No Lineal
curso 2006/07

JESÚS GARCIA FALSET
Departament d'Anàlisi Matemàtica

23 de junio de 2007

Índice general

Introducción

El Análisis funcional no lineal puede ser entendido como un tratamiento abstracto unificado, con métodos propios, de ciertos problemas de Topología, Análisis funcional y Matemática aplicada. Simplificando mucho, se trata del estudio de ecuaciones no lineales en espacios normados (tanto desde el punto de vista de la existencia de solución como del cálculo efectivo de la misma, o de una aproximación) así como del tratamiento de problemas de optimización, en los que la función a optimizar y las restricciones, si las hay, son de carácter no lineal. Aquí se introducirá al lector en dos tópicos paradigmáticos dentro de esta rama de las matemáticas:

(a). Métodos variacionales en optimización.

(b). Teoría del punto fijo.

Métodos variacionales.

Los fundamentos del Cálculo de Variaciones fueron dados en los siglos XVII y XVIII por Bernouilli, Euler, Legendre y Jacobi. Esta rama de las matemáticas intenta solucionar ciertas clases de problemas de máximos y mínimos, las cuales tienen en comun el hecho de que cada uno de ellas está asociada con determinadas expresiones integrales. El ejemplo más sencillo de uno de estos problemas es el siguiente:

Sea $f : \mathbb{R}^3 \to \mathbb{R}$ una función de clase $\mathcal{C}^2$, y denotamos, como es usual $\mathcal{C}^1([a,b])$ la funciones con derivada continua en $[a,b]$. Dada una función $\phi \in \mathcal{C}^1([a,b])$, definimos el funcional:

$$F(\phi) := \int_a^b f(t, \phi(t), \phi'(t))dt.$$

Entonces F es una función real sobre $\mathcal{C}^1([a,b])$. El problema consiste en encontrar una función $\psi \in \mathcal{C}^1([a,b])$ de forma que $F(\psi)$ sea el valor máximo o mínimo de F sujeto a la restricción de que el valor de ψ en los puntos inicial y final este predeterminado.

Ejemplo 0.0.1 *Consideremos todas las funciones de clase* $\mathcal{C}^1[a,b]$ *tal que tienen valores fijos en los extremos del intervalo* $[a,b]$. *Entonces sus gráficas se pueden considerar como trayectorias que unen dos puntos del plano. La pregunta es: ¿ Cual es la función cuya trayectoria tiene longitud mínima?*

Está claro que la longitud de la gráfica de una función en estas circunstancias viene dada por:

$$l(G(\phi)) := \int_a^b \sqrt{1+\phi'^2(t)}dt.$$

Con lo cual podemos definir la función $f(x,y,z) = \sqrt{1+z^2}$ *y entonces el problema se reduce a encontrar el mínimo de la función*

$$F(\phi) := \int_a^b f(t,\phi(t),\phi'(t))dt,$$

con las restricciones de que ϕ *tenga determinados valores prefijados en los extremos del intervalo* $[a,b]$.

Desde un punto de vista histórico, quizás el más famoso de estos problemas consiste en encontrar el mínimo tiempo en que puede descender, por la acción de la gravedad, una pequeña cuenta de collar ensartada en un alambre que une un punto con otro cercano y más bajo. El tiempo que la cuenta tarda en completar el descenso depende de la forma del alambre a través del cual se desliza. (Naturalmente, no se considera la influencia del rozamiento, etc). Este problema clásico se conoce como *problema de la braquistócrona.*

Ejemplo 0.0.2 (Braquistócrona) *Se trata de calcular el tiempo de descenso de una cuenta que se desliza sin rozamiento a través de un alambre que une dos puntos prefijados.*

Podemos representar el alambre como una curva diferenciable $y = \varphi(x)$ *en el plano* (x,y) *que une los puntos*

$$P_0 = (x_0, y_0), \quad P_1 = (x_1, y_1).$$

Por simplicidad supondremos $y_0 > y_1 > 0$ $x_0 < x_1$, *y llamaremos* g *a la aceleración de la gravedad.*

Sea $T(x)$ *el tiempo que tarda la cuenta en alcanzar el punto de la curva* $(x,\varphi(x))$. *Naturalmente* $T : [x_0, x_1] \to \mathbb{R}$, *y* $T(x_0) = 0$.

Sea $v(x)$ *la celeridad de la cuenta en el punto* $(x,\varphi(x))$. *Si la cuenta parte del reposo,* $v(x_0) = 0$.

Si aproximamos, para h pequeño, el arco de la curva entre los puntos

$$(x, \varphi(x)),\ (x+h, \varphi(x+h))$$

por el segmento rectilíneo que los une, y suponemos recorrido este segmento a celeridad constante $v(x)$, tendremos:

$$v(x) \sim \frac{\sqrt{h^2 + (\varphi(x+h) - \varphi(x))^2}}{T(x+h) - T(x)}$$

es decir

$$\frac{T(x+h) - T(x)}{h} \sim \frac{\sqrt{1 + (\frac{\varphi(x+h)-\varphi(x)}{h})^2}}{v(x)}$$

lo que nos dice que T es derivable en x con derivada:

$$T'(x) = \frac{\sqrt{1 + (\varphi'(x))^2}}{v(x)}.$$

Por otra parte, la cuenta en la posición $(x, \varphi(x))$ tendrá energía cinética

$$\frac{1}{2} m v(x)^2$$

y energía potencial

$$mg\varphi(x),$$

donde m es la masa del cuerpo que desciende.

La ley de conservación de la energía impone que la suma de ambas energías permanezca constante en cada punto del recorrido. En particular, si la cuenta está en reposo en P_0, tendrá en dicho punto energía cinética inicial nula y energía potencial mgy_0 por lo que para cada x en $[x_0, x_1]$,

$$\frac{1}{2} m v(x)^2 + mg\varphi(x) = mgy_0$$

y resultará

$$v(x) = \sqrt{2g(y_0 - \varphi(x))}.$$

Con lo cual,

$$T'(x) = \frac{\sqrt{1 + \varphi'(x)^2}}{\sqrt{2g(y_0 - \varphi(x))}}$$

Por tanto, suponiendo que la función anterior T sea absolutamente continua, es decir que cumpla la igualdad:

$$T(x_1) - T(x_0) = \int_{x_0}^{x_1} T'(x) dx,$$

el tiempo total de descenso T será:

$$T = T(x_1) - T(x_0) = \int_{x_0}^{x_1} \frac{\sqrt{1+\varphi'(x)^2}}{\sqrt{2g(y_0 - \varphi(x))}} dx.$$

Este tiempo depende de φ y puede ser considerado como un funcional $T(\varphi)$ con dominio $D(T)$ dado por el conjunto de funciones continuamente diferenciables sobre $[x_0, x_1]$ verificando las condiciones de frontera $\varphi(x_0) = y_0$, $\varphi(x_1) = y_1$, tales que exista y sea finita la integral

$$T(\varphi) := \int_{x_0}^{x_1} \frac{\sqrt{1+\varphi'(x)^2}}{\sqrt{2g(y_0 - \varphi(x))}} dx.$$

El planteamiento inicial del problema obliga a que la curva $y = \varphi(x)$ que minimice el tiempo de descenso, pase por los puntos P_0 y P_1, lo cual podemos traducirlo diciendo que φ ha de pertenecer a $D(T)$ y verificar las restricciones

$$\varphi(x_0) = y_0, \ \ \varphi(x_1) = y_1.$$

Esta claro que estamos ante un problema de minimizar el funcional

$$F(\varphi) = \int_{x_0}^{x_1} f(t, \varphi(t), \varphi'(t)) dt$$

donde $f(x, y, z) = \frac{\sqrt{1+z^2}}{\sqrt{2g(y_0 - y)}}$.

Esta función f está definida en el abierto de $\mathbb{R}^3$

$$G := \mathbb{R} \times (-\infty, y_0) \times \mathbb{R}$$

Con lo cual si queremos minimizar el funcional F lo tendremos que hacer sobre el conjunto

$$D_G := \{\varphi \in \mathcal{C}^1([x_0, x_1]) \ : \ \forall t \in [x_0, x_1], \quad (t, \varphi(t), \varphi'(t) \in G\}.$$

Para simplificar los cálculos podemos suponer, sin pérdida de generalidad, que $P = (0, 0)$ y el punto $Q = (x_1, y_1)$ tiene coordenadas $x_1 > 0$ e $y_1 < 0$. En estas circunstancias el problema de la Braquistocrona se reduce a estudiar la existencia de mínimos del funcional:

$$F(\varphi) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1+\varphi'(x)^2}}{\sqrt{\varphi(x)}} dx$$

sobre el conjunto

$$D_G := \{\varphi \in \mathcal{C}^1([x_0, x_1]) \ : \ \forall t \in [x_0, x_1], \quad (t, \varphi(t), \varphi'(t) \in G\}.$$

Siendo $G = \mathbb{R} \times (0, +\infty) \times \mathbb{R}$.

En este curso, trataremos problemas del tipo de los ejemplos anteriores, dando resultados que nos permitan saber cuando dichos problemas tienen solución y cómo obtener dicha solución.

Teoremas de punto fijo

La ecuación
$$x^3 + 3x^2 + 5x + 1 = 0$$
no es más que un problema, (el de encontrar los números reales o complejos) x de forma que la ecuación anterior se transforme en una identidad.

De igual modo, la ecuación diferencial
$$y'' - 3y' + y = 0$$
también es un problema, quizá planteado con menos precisión, que trata de encontrar las funciones y, de una variable real t, de forma que sean dos veces derivables en su dominio D y cumpliendo que $y''(t) - 3y'(t) + y(t) = 0$ para todo $t \in D$. La imprecisión está en no concretar el conjunto de funciones donde se busca la solución.

Estas ecuaciones (numérica la primera y funcional la segunda) i.e., estos problemas, pueden plantearse de manera unificada como casos particulares de una situación abstracta muy general.

Consideremos la función $T : \mathbb{R} \to \mathbb{R}$ definida por
$$T(x) = x^3 + 3x^2 + 6x + 1.$$
Un número real x_0 que verifica
$$x_0 = T(x_0) = x_0^3 + 3x_0^2 + 6x_0 + 1$$
debe cumplir que
$$x_0^3 + 3x_0^2 + 5x_0 + 1 = 0,$$
es decir, es una solución de la primera ecuación que hemos introducido.

De igual modo, una función $y(t)$ verificará
$$y(t) = -y''(t) + 3y'(t)$$
para todo t de su dominio, si y sólamente si, es solución de la ecuación diferencial anteriormente introducida. Luego, si definimos la aplicación T que a cada función y de un cierto conjunto de funciones $\mathcal{M}$ asocia la función
$$T(y) = -y'' + 3y'$$

entonces $y \in \mathcal{M}$ es solución de la ecuación diferencial anterior sii $T(y) = y$.

Vemos pues que buscar la solución de una ecuación en un conjunto puede ser lo mismo que buscar elementos de este conjunto que se transformen en sí mismos mediante una aplicación determinada.

Definición 0.0.3 *Si M es un conjunto no vacío y $T : M \to M$ es una aplicación, un punto $x \in M$ que se aplica en sí mismo, es decir que $T(x) = x$ se llama punto fijo o punto invariante de T en M.*

Los ejemplos anteriores sugieren que la existencia de puntos fijos de algunas aplicaciones está intimamente relacionada con la existencia de soluciones de ciertas ecuaciones, numéricas y funcionales.

Si $T : M \to M$ es una aplicación, para tener algún resultado general que garantice (en M) la existencia de puntos fijos de T, hay que imponer algunas condiciones tanto a la aplicación T como al conjunto M. Estos resultados se llaman resultados de punto fijo.

Unos teoremas de punto fijo usan propiedades topológicas de M y de T. Así surge la Teoría topológica del punto fijo.

Otros teoremas de punto fijo usan propiedades métricas de T o/y de M. De este modo surge la Teoría métrica del punto fijo.

En este curso, trataremos el teorema del punto fijo de Banach como ejemplo más relevante de la teoría métrica y los teoremas de Brouwer y Schauder como ejemplos de la teoría topológica.

Capítulo 1

La diferencial de Gâteaux y de Fréchet.

1.1. Introdución.

La Teoría de la diferenciación en espacios de dimensión infinita tiene sus comienzos en 1887 con V. Volterra, al considerar éste las derivadas variacionales de funciones de $\mathcal{C}([a,b])$ en $\mathbb{R}$.

Una de las más importantes motivaciones para el desarrollo del Cálculo Diferencial en espacios de dimensión infinita proviene de la teoría clásica del Cálculo de Variaciones.

La necesidad de desarrollar una teoría de la diferenciación en espacios de dimensión infinita, parece ser que, la puso de manifiesto J. Hadamard. Esta tarea la iniciaron sus discipulos M. Fréchet y M.R. Gâteaux.

La propuesta de este capítulo consiste en dar una exposición del Cálculo diferencial en dimensión infinita que posteriormente será de utilidad para el estudio del Cálculo de Variaciones.

1.2. Diferencial de Gâteaux.

Sea X un espacio normado, $\mathcal{U}$ un abierto no vacío de X.

Sea $F : \mathcal{U} \subseteq X \to \mathbb{R}$ una función. Fijado un vector no nulo $h \in X$ y $u_0 \in \mathcal{U}$, como $\mathcal{U}$ es un conjunto abierto existirá $\delta > 0$ tal que si $|t| < \delta$, entonces $u_0 + th \in \mathcal{U}$. Por lo tanto, si además $t \neq 0$ tiene sentido escribir:

$$\frac{F(u_0 + th) - F(u_0)}{t}.$$

Definición 1.2.1 *Si existe y es finito el siguiente límite:*

$$\lim_{t\to 0} \frac{F(u_0 + th) - F(u_0)}{t},$$

se llamará derivada de Gâteaux de F *en la dirección* $h \in X$ *en el punto* $u_0 \in \mathcal{U}$.

Como el límite de una función si existe es único, entonces una función puede tener a lo sumo una derivada de Gâteaux en el punto u_0 y en la dirección $h \in X$. Usaremos la notación $\delta F(u_0, h)$ para referirnos al límite anterior. Es natural convenir que $\delta F(u_0, 0) = 0$.

Observación 1.2.2 *El valor de* $\delta F(u_0, h)$ *es la derivada direccional de* F *en* u_0 *en la dirección de* h *y por lo tanto, la derivada de Gâteaux anterior se puede considerar una generalización de las derivadas direccionales del Cálculo diferencial de varias variables.*

La variación de Gâteaux de F *en un punto depende sólamente del comportamiento local de* F *cerca del punto. Además, la variación de Gâteaux es una operación lineal, en el sentido de que si* F *y* G *tienen variación de Gâteaux en un punto* u_0 *en la dirección* h *entonces,* $\delta(\alpha F + \beta G)(u_0, h) = \alpha\delta F(u_0, h) + \beta\delta G(u_0, h)$.

Puede suceder que exista $\delta F(u_0, h)$ para todo $h \in X$, o en otras palabras que la aplicación $h \to \delta F(u_0, h)$ tenga por dominio todo X. En este caso, si la aplicación

$$h \to \delta F(u_0, h)$$

es lineal y continua, (es decir si pertenece al dual topológico de X) diremos que F es *diferenciable en el sentido de Gâteaux en* u_0 y para referirnos a esta aplicación usaremos la notación $F'(u_0) : X \to \mathbb{R}$ tal que $F'(u_0)(h) := \delta F(u_0, h)$.

Observación 1.2.3 *El concepto de función diferenciable Gâteaux no es completamente general. Así hay quien distingue el caso en que el dominio de la función anterior sea todo el espacio, diciendo entonces que la función es diferenciable Gâteaux y quien, como nosotros, solo en el caso en que además sea continua y lineal diremos que* F *es diferenciable Gâteaux.*

Ejemplo 1.2.4 *Sea* X *un espacio normado y definimos la función* $F : X \to \mathbb{R}$ *definida por* $F(x) = \|x\|^2$.

Si $h \in X$ *es no nulo, y tomamos* $u_0 = 0$ *se tiene que*

$$\delta F(0,h) = \lim_{t\to 0} \frac{t^2\|h\|^2}{t} = 0$$

con lo cual F es diferenciable en el sentido de Gâteaux en $u_0 = 0$ y además su diferencial es la función identicamente nula.

Ejemplo 1.2.5 *Sea X un espacio normado y definimos la función $F : X \to \mathbb{R}$ definida por $F(x) = \|x\|$.*

Si $h \in X$ es no nulo, y tomamos $u_0 = 0$ se tiene que

$$\delta F(0,h) = \lim_{t\to 0} \frac{|t|\|h\|}{t}.$$

Es claro que el límite anterior no existe, por lo tanto el funcional F no tiene derivada de Gâteaux en ninguna dirección en el punto $u_0 = 0$.

Ejemplo 1.2.6 *Consideremos el espacio de Banach $X := \mathcal{C}([a,b])$ con su norma usual y consideremos el funcional $J : X \to \mathbb{R}$ definido por*

$$J(y) = \int_a^b (\sin^3(x) + y^2(x))dx,$$

el cual está definido en todo X. Veamos ahora la variación de Gâteaux en una dirección v y en un punto y. Tenemos que evaluar:

$$\frac{J(y+tv) - J(y)}{t} = \frac{1}{t}\int_a^b [(y+tv)^2(x) - y^2(x)]dx =$$

$$\frac{1}{t}\int_a^b [y^2(x) + 2ty(x)v(x) + t^2v^2(x) - y^2(x)]dx =$$

$$2\int_a^b y(x)v(x)dx + t\int_a^b v^2(x)dx.$$

Con lo cual se tiene que

$$\delta J(y,v) = 2\int_a^b y(x)v(x)dx.$$

Esto significa que existe la derivada de Gâteaux en toda dirección y además es una aplicación lineal y continua por lo tanto es Gâteaux diferenciable.

1.3. Diferencial de Fréchet.

Definición 1.3.1 *Se dice que F es diferenciable en u_0 en el sentido de Frechet si existe una aplicación lineal y continua $L : X \to \mathbb{R}$ tal que*

$$0 = \lim_{h\to 0} \frac{F(u_0 + h) - F(u_0) - L(h)}{\|h\|}.$$

Proposición 1.3.2 *La aplicación lineal y continua L que verifica la definición anterior, en caso de existir es única.*

Prueba. Supongamos que L y S sean dos aplicaciones en las condiciones de la definición anterior y consideremos $w \in S_X$. En este caso, se tendrá que para todo $\epsilon > 0$ existe $\delta > 0$ tal que si, $t \in \mathbb{R}$ y $0 < |t| < \delta$, entonces $u_0 + tw \in \mathcal{U}$ y

$$|F(u_0 + tw) - F(u_0) - L(tw)| < \epsilon\|tw\| = \epsilon|t|$$

y existe $\delta' > 0$ tal que si $0 < |t| < \delta'$, entonces $u_0 + tw \in \mathcal{U}$ y

$$|F(u_0 + tw) - F(u_0) - S(tw)| < \epsilon\|tw\| = \epsilon|t|$$

Con lo cual, para $|t| < \min\{\delta, \delta'\}$

$$|L(tw) - S(tw)| < 2\epsilon|t|,$$

consecuentemente

$$|L(w) - S(w)| < 2\epsilon, .$$

Como $\epsilon > 0$ es arbitrario, se debe cumplir que $L(w) = S(w)$. Pero dos aplicaciones lineales y continuas que coinciden sobre la esfera unidad de X han de ser iguales.

Si F es diferenciable en u_0, la unicidad que nos proporciona la proposición anterior nos permite.

Definición 1.3.3 *LLamaremos diferencial de Fréchet de la función F en el punto u_0 a la única aplicación lineal y continua L que verifica la definición anterior. Se denotará por $DF(u_0)$.*

Ejemplo 1.3.4 *Sea $(X, \|.\|)$ un espacio normado, para la función real $F : X \to \mathbb{R}$ definida por $F(x) = \|x\|^2$, se tiene que F es diferenciable en el* *origen, con diferencial $DF(0) = \ominus$, (donde $\ominus$ representa la aplicación lineal nula sobre X).*

En efecto,

$$\lim_{h\to 0} \frac{F(0+h) - F(0) - \ominus(h)}{\|h\|} = \lim_{h\to 0} \frac{\|h\|^2}{\|h\|} = 0.$$

Naturalmente si X es de dimensión finita la diferencial de Fréchet de una función es la diferencial ordinaria y viceversa ya que, toda aplicación lineal entre espacios de dimensión finita es continua.

Teorema 1.3.5 *Sea $F : \mathcal{U} \subseteq X \to \mathbb{R}$ una función diferenciable Fréchet en $u_0 \in \mathcal{U}$, entonces F es continua en dicho punto.*

Prueba.

Como $\mathcal{U}$ es un abierto existirá $\delta > 0$ tal que si $\|h\| < \delta$, entonces $u_0 + h \in \mathcal{U}$. Para h en las condiciones anteriores se tiene:

$$F(u_0+h) - F(u_0) = \|h\| \frac{F(u_0+h) - F(u_0) - DF(u_0)(h)}{\|h\|} + D(F(u_0)(h)$$

Como $DF(u_0)$ es lineal y continua se cumplirá que

$$\lim_{h\to 0} DF(u_0)(h) = 0.$$

Por lo tanto,

$$\lim_{h\to 0} F(u_0+h) = F(u_0).$$

La propiedad que se lista seguidamente se demuestra de la misma forma que en dimensión finita, adaptando ligeramente la prueba.

Si $F_1, F_2 : U \to \mathbb{R}$ son Fréchet diferenciables en u_0 y α, β son números reales, entonces $\alpha F_1 + \beta F_2$ es diferenciable en u_0 con

$$D(\alpha F_1 + \beta F_2)(u_0) = \alpha DF_1(u_0) + \beta DF_2(u_0).$$

1.4. Relación entre las dos diferenciales.

Proposición 1.4.1 *Sea $F : \mathcal{U} \subseteq X \to \mathbb{R}$ una función diferenciable en el sentido de Fréchet en u_0 entonces, F es diferenciable en el sentido de Gâteaux en ese punto, y además $DF(u_0) = F'(u_0)$.*

Prueba. Sea $h \in X$ un vector no nulo. Queremos ver que existe $\delta F(u_0, h)$ y que es igual a $DF(u_0)(h)$.

En efecto, como F es diferenciable Fréchet sabemos que $DF(u_0) \in X^\star$ con lo cual fijado h se tendrá que

$$\lim_{t \to 0} \frac{F(u_0 + th) - F(u_0) - DF(u_0)(th)}{\|th\|} = 0,$$

lo cual implica que

$$0 = \lim_{t \to 0} \frac{|t|}{t} \frac{F(u_0 + th) - F(u_0) - tDF(u_0)(h)}{|t|\|h\|} = 0$$

y por lo tanto

$$\lim_{t \to 0} \left(\frac{F(u_0 + th) - F(u_0)}{t} - DF(u_0)(h)\right) = 0.$$

Esto significa que existe $\delta F(u_0, h) = DF(u_0)(h)$. Como además, $DF(u_0)$ es lineal y continua ya tenemos el resultado.

Ejemplo 1.4.2 *$F(x, y) = \frac{xy^3}{x^2+y^6}$ si $(x, y) \neq (0, 0)$ y $F(0, 0) = 0$. Es un ejercicio sencillo ver que esta función no es continua en el punto $(0, 0)$ y por lo tanto no puede ser diferenciable Fréchet en dicho punto. Sin embargo, las derivadas direccionales son todas nulas, y por lo tanto la función es diferenciable Gâteaux en el origen.*

Como pone de manifiesto el ejemplo anterior el recíproco a la proposición anterior no se verifica. No obstante, si F es Gâteaux diferenciable en u_0 y si la igualdad

$$\lim_{t \to 0} \frac{F(u_0 + tv) - F(u_0)}{t} = F'(u_0)(v)$$

se verifica uniformemente en $v \in S_X$, entonces F es Fréchet diferenciable en u_0 con $DF(u_0) = F'(u_0)$.

En efecto, dado cualquier $h \in X$ no nulo, se tiene que

$$\lim_{t \to 0} \frac{F(u_0 + t\frac{h}{\|h\|}) - F(u_0)}{t} = F'(u_0)\left(\frac{h}{\|h\|}\right)$$

se verifica uniformemente en h. Esto significa que dado $\epsilon > 0$ existe $\delta > 0$ de forma que para cualquier vector no nulo $h \in X$, si $0 < |t| < \delta$

$$\left|\frac{F(u_0 + t\frac{h}{\|h\|}) - F(u_0)}{t} - F'(u_0)\left(\frac{h}{\|h\|}\right)\right| < \epsilon$$

Luego si $w \in X$ es tal que $0 < \|w\| < \delta$ entonces

$$|\frac{F(u_0 + \|w\|\frac{w}{\|w\|}) - F(u_0)}{\|w\|} - F'(u_0)(\frac{w}{\|w\|})| < \epsilon$$

o lo que es lo mismo,

$$|F(u_0 + w) - F(u_0) - F'(u_0)(w)| < \epsilon\|w\|.$$

Luego F es Fréchet diferenciable en u_0, y por unicidad, $DF(u_0) = F'(u_0)$.

El siguiente resultado es la extensión de la condición suficiente de diferenciabilidad, i.e.; que toda función de clase $\mathcal{C}^1$ es diferenciable.

Teorema 1.4.3 *Sea $F : \mathcal{U} \to \mathbb{R}$. Si La diferencial de Gâteaux $F'(v)$ existe en algún entorno de $x \in \mathcal{U}$ y es continua en x, entonces F es diferenciable Fréchet en x y además, $DF(x) = F'(x)$.*

Prueba.

Escribimos para $h \in X$

$$w(x, h) = F(x + h) - F(x) - F'(x)(h).$$

Como F está definida en algún entorno de x, existirá $\delta > 0$ tal que $w(x, h)$ tiene sentido siempre que $\|h\| < \delta$.

Consideremos la función real de variable real

$$H(t) = F(x + th) - tF'(x)(h).$$

$H(t)$ tiene sentido para todo $t \in [0, 1]$. Además,

$$H(1) - H(0) = F(x + h) - F'(x)(h) - F(x) = w(x, h)$$

Observar que, si $\tau \in]0, 1[$ y $|s|$ es suficientemente pequeño,

$$H(\tau + s) - H(\tau) = F(x + (\tau + s)h) - F(x + \tau h) - sF'(x)(h).$$

Luego

$$H'(\tau) = \lim_{s\to 0} \frac{1}{s}[H(\tau + s) - H(\tau)] = F'(x + \tau h)(h) - F'(x)(h)$$

existirá para $\tau \in]0, 1[$, (siempre que $\|h\| < \delta$).

Desde luego, H es continua en 0, en efecto,

$$|H(s)-H(0)| = |F(x+sh)-sF'(x)(h)-F(x)| = |s|\left|\frac{F(x+sh)-F(x)}{s}-F'(x)(h)\right|,$$

y por la definición de derivada de Gâteaux se puede concluir que

$$\lim_{s\to 0^+} H(s) = H(0).$$

De igual modo podemos ver que H es continua en 1, ya que

$$|H(s)-H(1)| = |s-1|\left|\frac{F(x+h+(s-1)h)-F(x+h)}{s-1}-F'(x)(h)\right|$$

Luego de la definición de Gâteaux diferenciabele se obtiene que

$$\lim_{s\to 1^-} H(s) = H(1).$$

Con lo cual, podemos aplicar el teorema del valor medio a la función H en el intervalo $[0,1]$. Existe $\tau = \tau(h) \in]0,1[$ con

$$w(x,h) = H(1) - H(0) = H'(\tau) = F'(x+\tau h)(h) - F'(x)(h)$$

Entonces,

$$|w(x,h)| = |F'(x+\tau h)(h) - F'(x)(h)| \leq \|F'(x+\tau h) - F'(x)\|\|h\|.$$

por lo que, si $0 < \|h\| < \delta$

$$\frac{|w(x,h)|}{\|h\|} \leq \|F'(x+\tau h) - F'(x)\|.$$

Siendo continua F' existirá δ_1 con $0 < \delta_1 < \delta$ de modo que, dado cualquier $\epsilon > 0$, si $0 < \|h\| < \delta_1$

$$\frac{w(x,h)|}{\|h\|} \leq \|F'(x+\tau h) - F'(x)\| < \epsilon.$$

Lo que prueba que F es Fréchet diferenciable en x.

1.5. Extremos locales.

1.5.1. Extremos locales sobre conjuntos abiertos.

Definición 1.5.1 *Sea X un espacio normado y sea $F : \Omega \subseteq X \to \mathbb{R}$ una función real con dominio el abierto Ω. Se dice que F tiene un mínimo local (máximo local) en $x_0 \in \Omega$, si existe $\epsilon > 0$ tal que para cada $x \in B(x_0, \epsilon) \subseteq \Omega$ $F(x) \geq F(x_0)$ ($F(x) \leq F(x_0)$).*

Para el estudio de extremos locales de funciones de varias variables reales, sabemos que el cálculo diferencial ordinario nos proporciona mucha información. Veremos ahora cómo podemos utilizar los conceptos que hemos introducido en los apartados anteriores.

Proposición 1.5.2 *Sea X un espacio normado y D un subconjunto abierto y no vacío de X. Si un funcional $F : D \to \mathbb{R}$ tiene un extremo local en $x \in D$, y además F tiene varición de Gâteaux en x, entonces $\delta F(x_0, h) = 0$, para todo $h \in X$.*

Prueba.

Supongamos, sin perdidda de generalidad, que x es un mínimo local de F en D, entonces existirá $\delta > 0$ tal que $x + th \in D$ siempre que $|t| < \delta$ y además $F(x + th) \geq F(x)$. Por tanto,

$$\frac{F(x+th) - F(x)}{t} \geq 0 \quad \text{para } t > 0$$
$$\frac{F(x+th) - F(x)}{t} \leq 0 \quad \text{para } t < 0$$

y como F tiene variación de Gâteaux en el extremo local, sabemos que existe $\delta F(x, h)$ y esto significa que los límites laterales existen y son iguales, con lo cual $\delta F(x, h) = 0$.

La proposición que acabamos de presentar nos dá una condición necesaria de extremo local. Sin embargo, queda muy lejos, como pone de manifiesto el caso de funciones reales de varias variables, que ésta sea una condición suficiente.

Seguidamente veremos como para funcionales convexos si que tenemos una condición suficiente.

Recordemos que un funcional convexo $f : D \subset X \to \mathbb{R}$, donde D es un subconjunto convexo de un espacio normado X es una función que cumple:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y), \quad \forall \alpha \in [0,1], \quad \forall x, y \in D.$$

Proposición 1.5.3 *Sea X un espacio normado y D un subconjunto abierto y no vacío de X. Si un funcional convexo (esto significará que es convexo sobre los subconjuntos convexos de D) $F : D \to \mathbb{R}$ si existe $x \in D$ tal que $\delta F(x_0, h) = 0$, para todo $h \in X$. Entonces x es un mínimo local de F en D.*

Prueba Como D es abierto entonces existirá una bola centrada en x y contenida en D, llamemosla B, como toda bola es un conjunto convexo, podemos afirmar que F es convexo sobre la bola B. (Los puntos de B se pueden escribir como $x + h$ para $\|h\|$ suficientemente pequeña.

Como

$$0 = \delta F(x, h) = \lim_{t \to 0^+} \frac{F(x + th) - F(x)}{t} =$$

$$\lim_{t \to 0^+} \frac{F((1 - t)x + t(x + h)) - F(x)}{t} \leq$$

$$\lim_{t \to 0^+} \frac{(1 - t)F(x) + tF(x + h) - F(x)}{t} =$$

$$F(t + h) - F(x)$$

Luego, x es un mínimo local.

Observación 1.5.4 *Si en la proposición anterior suponemos que D es convexo, entonces obtenemos que x es un mínimo absoluto de F sobre D.*

Si miramos con atención la prueba de la última proposición, se desprende que si F es un funcional convexo sobre D, entonces $F(x + v) - F(x) \geq \delta F(x, v)$, siempre que $y + v \in D$.

Ejemplo 1.5.5 *Vamos a calcular los mínimos absolutos del funcional $F : \mathcal{C}([a, b]) \to \mathbb{R}$ definido por*

$$F(y) := \int_a^b [\sin^3(x) + y^2(x)]dx.$$

Claramente F es un funcional convexo definido en todo el espacio $\mathcal{C}([a, b])$, con lo cual su dominio es un conjunto abierto y convexo.

Por otra parte, según muestra el ejemplo 1.2.6 éste funcional es Gâteaux diferenciable en todo punto y además su diferencial de Gâteaux viene dada por:

$$\delta F(y, v) = 2 \int_a^b y(x)v(x)dx, \quad \forall y, v \in \mathcal{C}([a, b]).$$

De este modo, por la proposición 1.5.3 y la observación 1.5.4, los mínimos absolutos de este funcional vienen dados por los puntos $y \in \mathcal{C}([a,b])$ *tal que* $\delta F(y,v) = 0$ *para todo* $v \in \mathcal{C}([a,b])$.

Consecuentemente nos tenemos que plantear la ecuación

$$2\int_a^b y(x)v(x)dx = 0, \quad \forall v \in \mathcal{C}([a,b]). \tag{1.1}$$

Es evidente que la función identicamente igual a cero ($y(x) = 0 \quad \forall x \in [a,b]$) *verifica la ecuación (1.1). Además es fácil deducir que es la única función que la verifica (esto lo veremos con más detalle en el siguiente capítulo, en los lemas variacionales).*

Definición 1.5.6 *Sea* X *un espacio normado,* $\mathcal{U}$ *un abierto no vacío de* X. *Sea* $F : \mathcal{U} \to \mathbb{R}$ *un funcional de forma que tiene variación de Gâteaux en cualquier punto de* $\mathcal{U}$. *Dado* $x_0 \in \mathcal{U}$ *y dado* $h \in X \setminus \{0\}$ *se llama variación de Gâteaux segunda en el punto* x_0 *y en la dirección* h *al valor del siguiente límite siempre que exista:*

$$\delta^2 F(x_0,h) = \lim_{s\to 0} \frac{\delta F(x_0 + sh, h) - \delta F(x_0,h)}{s}.$$

Teorema 1.5.7 (Teorema de Taylor) *Sea* X *un espacio normado y* $\mathcal{U}$ *un subconjunto abierto de* X. *Supongamos que* $u_0 \in \mathcal{U}$ *y que el segmento de extremos* u_0, u_0+h *está contenido en* $\mathcal{U}$. *Supongamos además que* $J : \mathcal{U} \to \mathbb{R}$ *es dos veces Gâteaux-derivable en todo punto del segmento anterior.*

Entonces existe c *en el segmento de extremos* u_0 *y* $u_0 + h$ *tal que*

$$J(u_0 + h) = J(u_0) + \delta J(u_0,h) + \frac{1}{2}\delta^2 J(c,h).$$

Prueba. Sea $f :]-r, 1+r[\to \mathbb{R}$ la función dada por

$$f(t) := J(u_0 + th)$$

Sabemos que, por suponer que J es dos veces G-derivable en los puntos del segmento, estamos afirmando que f es derivable en $[0,1]$ ya que;

$$\lim_{s\to 0} \frac{f(t+s) - f(t)}{s} = \lim_{s\to 0} \frac{J(u_0 + th + sh) - J(u_0 + th)}{s} = \delta J(u_0 + th, h).$$

Además,

$$\frac{f'(t+s) - f'(t)}{s} = \frac{\delta J(u_0 + th + sh, h) - \delta J(u_0 + th, h)}{s}$$

por lo que

$$f''(t) = \delta^2 J(u_0 + th, h)$$

Aplicando la fórmula de Mc Laurin a la función f tendremos que existirá $\theta \in (0, 1)$ con

$$f(1) = f(0) + f'(0) + \frac{1}{2} f''(\theta).$$

Con lo cual,

$$J(u_0 + h) = J(u_0) + \delta J(u_0, h) + \frac{1}{2}\delta^2 J(u_0 + \theta h, h).$$

Teorema 1.5.8 *Sea $F : \mathcal{U} \to \mathbb{R}$ un funcional. El funcional F tiene en $u_0 \in \mathcal{U}$ un mínimo local siempre que:*

(i) $\delta F(u_0, h) = 0$ para cada $h \in X$.

(ii) $\forall h \in X, \exists \delta^2 F(u, h)$ en un entorno de u_0 y además existe una constante positiva c tal que

$$\delta^2 F(u_0, h) \geq c\|h\|^2 \quad \forall h \in X.$$

(iii) Dado $\epsilon > 0$existe $\mu > 0$ (dependiendo de ϵ) tal que

$$|\delta^2 F(u, h) - \delta^2 F(u_0, h)| \leq \epsilon \|h\|^2$$

para cualquier $u, h \in X$ con $\|u - u_0\| \leq \mu$.

Prueba.

Definimos la función $f(t) = F(u_0 + th)$. Patra $|t|$ lo suficientemente pequeño sabemos, ver la demostración del teorema de Taylor, que

$$f'(t) = \delta F(u_0 + th, h), \quad f''(t) = \delta^2 F(u_0 + th, h).$$

Luego las hipótesis que estamos haciendo sobre la existencia de la variación segunda de F nos permiten aplicar el teorema de Taylor anterior para obtener que existe $\theta \in]0, 1[$ tal que

$$F(u_0 + h) - F(u_0) = \frac{1}{2}\delta^2 F(u_0 + \theta h, h) =$$

$$\frac{1}{2}\delta^2 F(u_0, h) + \frac{1}{2}[\delta^2 F(u_0 + \theta h, h) - \delta^2 F(u_0, h)] \geq$$

$$\frac{1}{2} c\|h\|^2 + \frac{1}{2}[\delta^2 F(u_0 + \theta h, h) - \delta^2 F(u_0, h)]$$

Si damos $\epsilon = \frac{c}{4}$ en la condición (iii) del enunciado, existirá $\mu > 0$ tal que

$$|\delta^2 F(u,h) - \delta^2 F(u_0,h)| \leq \frac{c}{2}\|h\|^2$$

siempre que $\|u - u_0\| < \mu$.

Si en particular, tomamos h tal que $\|h\| < \mu$, entonces $\|u_0 + \theta h - u_0\| \leq \|h\| < \mu$, luego

$$|\delta^2 F(u_0 + \theta h, h) - \delta^2 F(u_0,h)| \leq \frac{c}{2}\|h\|^2$$

o lo que es lo mismo,

$$-\frac{c}{2}\|h\|^2 < \delta^2 F(u_0 + \theta h, h) - \delta^2 F(u_0,h) < \frac{c}{2}\|h\|^2.$$

Sustituyendo obtenemos

$$F(u_0 + h) - F(u_0) \geq \frac{c}{2}\|h\|^2 - \frac{c}{4}\|h\|^2 \geq 0$$

lo que prueba que F tiene un mínimo local en u_0.

El siguiente ejemplo pone de manifiesto que las condiciones que se imponen a la segunda derivada no son supérfugas, en el sentido de que su positividad sólo no asegura que el punto crítico sea un mínimo.

Ejemplo 1.5.9 *Consideremos el funcional $J : l_2 \to \mathbb{R}$ definido por*

$$J((x_n)) = \sum_{n=1}^{\infty} \frac{x_n^2}{n^3} - \sum_{n=1}^{\infty} x_n^4.$$

Veamos que el origen no es un mínimo local para J, aunque $J'(0) = 0$ y que no obstante $J''(0)(h,h) \geq 0$ para todo $h \in l_2$.

En efecto, El origen no es un mínimo local ya que

$$J((0, ..., 0, \frac{1}{n}, 0, ...)) = \frac{1}{n^5} - \frac{1}{n^4} < 0 = J(0).$$

Dados $u = (u_n), h = (h_n) \in l_2$, calculemos

$$J(u + th) = \sum_{n=1}^{\infty} \frac{(u_n + th_n)^2}{n^3} - \sum_{n=1}^{\infty} (u_n + th_n)^4 =$$

$$\sum_{n=1}^{\infty} \frac{u_n^2 + 2tu_nh_n + t^2h_n^2}{n^3} - \sum_{n=1}^{\infty} (u_n^4 + 4u_n^3 th_n + 6u_n^2 t^2 h_n^2 + 4u_n t^3 h_n^3 + t^4 h_n^4) =$$

$$J(u) + t\sum_{n=1}^{\infty}[\frac{2u_nh_n}{n^3} - 4u_n^3h_n] + t^2\sum_{n=1}^{\infty}[\frac{h_n^2}{n^3} - 6u_n^2h_n^2 - 4u_nth_n^3 - t^2h_n^4]$$

De donde,

$$\lim_{t\to 0}\frac{J(u+th)-J(u)}{t} = \sum_{n=1}^{\infty}[\frac{2u_nh_n}{n^3} - 4u_n^3h_n]$$

Luego existe $\delta J(u,h)$ y además la función $h \to \delta(u,h)$ es lineal y continua en h. Puede escribirse

$$\delta J(u,h) = \langle J'(u), h\rangle$$

donde $J'(u)$ es el vector de l_2,

$$(\frac{2u_n}{n^3} - 4u_n^3)$$

En particular $J'(0) = 0$. A partir de aquí podemos calcular la variación segunda en el origen:

Como

$$\frac{\delta J(0+sh,h)-\delta J(0,h)}{s} = \frac{1}{s}\delta J(sh,h) =$$

$$= \frac{1}{s}\sum_{n=1}^{\infty}[\frac{2sh_nh_n}{n^3} - 4(sh_n)^3h_n]$$

entonces exisrirá

$$\lim_{s\to 0}\frac{\delta J(0+sh,h)-\delta J(0,h)}{s} = \sum_{n=1}^{\infty}\frac{2h_n^2}{n^3}$$

Así pues, existe la diferencial segunda de Gâteaux en el origen y

$$\delta^2 J(0)(h,h) = \sum_{n=1}^{\infty}\frac{2h_n^2}{n^3} \geq 0.$$

1.6. Problemas

Problema 1.6.1 *Sea $(X, \|.\|_2)$ el espacio $\mathcal{C}([0,1])$ con la norma*

$$\|f\|_2 := \left(\int_0^1 f^2(t)dt\right)^{\frac{1}{2}}.$$

Probar que el funcional $J : X \to \mathbb{R}$ definido por $J(f) = f(0)$ es lineal pero no es continuo en cada punto de X.

Problema 1.6.2 *Sea* $J : \mathcal{C}([0,1]) \to \mathbb{R}$ *definido, para cada* $f \in \mathcal{C}([0,1])$ *por*

$$J(f) = \int_0^1 \sin(x) f^3(x) dx.$$

Probar que J *es continuo.*

Problema 1.6.3 *Sea* $J : \mathcal{C}^1([a,b]) \to \mathbb{R}$ *definido, para cada* $f \in \mathcal{C}^1([a,b])$*, por*

$$J(f) = \int_a^b [x^2 f'(x)^2] dx.$$

Hallar $\delta J(f,h)$ *para cualquiera* $f, h \in \mathcal{C}^1([a,b])$*.*

Problema 1.6.4 *Sea* $J : \mathcal{C}([0,1]) \to \mathbb{R}$ *un funcional definido por*

$$J(f) = \int_0^1 [2\tan(x)(1 - \cos(x)) f(x) - f^2(x)] dx.$$

Calcular la derivada de Gâteaux de J *en todo punto y en toda dirección. Encontrar un máximo local de* J*.*

Problema 1.6.5 *Sea* $J : \mathcal{C}([0,2]) \to \mathbb{R}$ *un funcional definido por*

$$J(f) = \int_0^2 [3x^5 - 12x^2 f(x) + 3x f(x)^2 + 2 f(x)^3] dx.$$

Obtener todos los posibles extremos locales.

Capítulo 2

Métodos variacionales en optimización.

Aunque la solución de Jakob Bernoulli de 1696 al problema de la braquistócrona que le planteó su hermano Johann marcó la introducción de las consideraciones variacionales, no fue hasta los trabajos de Euler (1742) y de Lagrange (1755) en que la ahora conocida teoría sistematica del Cálculo de variaciones emergió. Inicialmente, fue restringido a encontrar condiciones necesarias en orden a que una función integral

$$F(y) = \int_a^b f(x, y(x), y'(x))dx$$

pudiese tener un extremo local sobre el conjunto

$$D \subseteq \{y \in \mathcal{C}^1([a,b]) \ : \ y(a) = a_1; \ y(b) = b_1\}$$

El estudio de encontrar condiciones necesarias para minimizar el problema planteado arriba, será el objetivo principal de este capítulo.

2.1. Funcionales de Lagrange

Como hemos visto en el caso de la **Braquistócrona** los funcionales que aparecen deben ser minimizados en conjuntos de la forma:

$$D_G := \{\varphi \in \mathcal{C}^1[a,b] \ : \ \forall t \in [a,b], \ (t, \varphi(t), \varphi'(t)) \in G\}$$

donde G es un abierto de $\mathbb{R}^3$.

Seguidamente probaremos que D_G es un abierto del espacio de Banach $(\mathcal{C}^1[a,b], \|.\|)$, (donde $[a,b]$ es un intervalo compacto de la recta real y donde la norma $\|.\|$ se define por

$$\|\varphi\| = \|\varphi\|_\infty + \|\varphi'\|_\infty)$$

En efecto, si D_G es vacío evidentemente es abierto.

Si $\varphi_0 \in D_G$, sea

$$M = \{(t, \varphi_0(t), \varphi_0'(t)) \ : \ t \in [a,b]\}$$

Por la continuidad de las funciones φ_0 y φ_0' en el intervalo compacto $[a,b]$, el conjunto M será un compacto de $\mathbb{R}^3$ contenido en el abierto G. Por lo tanto, existe $r > 0$ tal que

$$B_1(M,r) := \{(t,x,y) \ : \ d_1((t,x,y),M) < r\} \subset G.$$

donde d_1 es la distancia dada por la norma $\|(a,b,c)\|_1 = |a| + |b| + |c|$.

Si tomamos $\varphi \in \mathcal{C}^1[a,b]$ con $\|\varphi - \varphi_0\| < r$, se tendrá, para cada $t \in [a,b]$ que $(t, \varphi_0(t), \varphi_0'(t)) \in M$ y

$$\|(t, \varphi_0(t), \varphi_0'(t)) - (t, \varphi(t), \varphi'(t))\|_1 \leq \|\varphi_0 - \varphi\| < r$$

Luego

$$d_1((t,\varphi(t),\varphi'(t)),M) \leq d_1((t,\varphi(t),\varphi'(t)),(t,\varphi_0(t),\varphi_0'(t)) < r$$

Por lo tanto para cada $t \in [a,b]$

$$(t, \varphi(t), \varphi'(t)) \in G$$

es decir $\varphi \in D_G$.

Definición 2.1.1 *Dado un abierto G de $\mathbb{R}^3$ y una función F real de clase $\mathcal{C}^1(G)$ se llamará funcional de Lagrange a:*

$$J : D_G \to \mathbb{R},$$
$$J(\varphi) := \int_a^b F(t, \varphi(t), \varphi'(t))dt.$$

Como hemos visto en la introducción este tipo de funcionales son los que dan origen al cálculo de variaciones.

Nuestro objetivo, ahora, es probar que los funcionales de Lagrange admiten variación de Gâteuax en todas direciones y en todo punto de su dominio.

Para probar que existe $\delta J(\varphi, h)$ para todo $\varphi \in D_G$ y toda $h \in \mathcal{C}^1[a,b]$ primeramente necesitamos el siguiente lema de integración.

Lema 2.1.2 *Sea $\mathcal{U}$ un abierto de $\mathbb{R}$ y sea $f : [a,b] \times \mathcal{U} \to \mathbb{R}$ una función que cumple:*

(i) Para cada $x \in \mathcal{U}$ fijo, la función $t \to f(t,x)$ es integrable en $[a,b]$(de esta forma nos aseguramos que $\varphi(x) = \int_a^b f(t,x)dt$ existe).

(ii) Para cada $t \in [a,b]$ fijo, la función $x \to f(t,x)$ es de clase $\mathcal{C}^1$.

(iii) La función $t \to D_2 f(t,x)$ es integrable en $[a,b]$ para cada $x \in \mathcal{U}$.

(iv) Además, existe $M > 0$ tal que $|D_2 f(t,x)| \leq M$ para todo $(t,x) \in [a,b] \times \mathcal{U}$.

Entonces la función

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

es derivable y

$$\varphi'(x) = \int_a^b D_2 f(t,x)dt.$$

Prueba.

Dado $x \in \mathcal{U}$ tenemos que estudiar la siguiente expresión:

$$\lim_{h \to 0} \frac{\varphi(x+h) - \varphi(x)}{h}$$

Para probar la existencia de este límite utilizaremos la caracterización sucesional. Supongamos que (t_k) es una sucesión tal que $\lim_{k\to\infty} t_k = 0$ y además $t_k \neq 0$ para todo $k \in \mathbb{N}$. Si desarrollamos el cociente anterior nos queda :

$$\frac{\varphi(x+t_k) - \varphi(x)}{t_k} = \int_a^b \frac{f(t,x+t_k) - f(t,x)}{t_k} dt$$

Ahora, como por (ii) a la función $f(t,.)$ le podemos aplicar el teorema del valor medio, nos quedara que existe $\theta_k \in]0,1[$ tal que:

$$\int_a^b \frac{f(t,x+t_k) - f(t,x)}{t_k} = \int_a^b D_2 f(t, x + \theta_k t_k)dt.$$

LLamemos

$$h_k(t) := D_2 f(t, x + \theta_k t_k)$$

por (iii) estas funciones son integrables. Además $h_k \to D_2 f(.,x)$ puntualmente. Entonces, por (iv), podemos aplicar el teorema de la convergencia dominada y obtenemos:

$$\int_a^b D_2 f(t,x)dt = \lim_{k\to\infty} \int_a^b h_k(t)dt =$$

$$\lim_{k\to\infty} \int_a^b D_2 f(t, x+\theta_k t_k)dt = \lim_{k\to\infty} \frac{\varphi(x+t_k)-\varphi(x)}{t_k}.$$

Lo cual prueba el resultado.

Una vez visto este lema pasemos a demostrar la existencia de la variación de Gâteaux del funcional de Lagrange. Sea $\varphi \in D_G$ y $h \in \mathcal{C}^1[a,b]$ dos funciones fijas.

Consideremos ahora la función de dos variables

$$f(t,\epsilon) := F(t, \varphi(t)+\epsilon h(t), \varphi'(t)+\epsilon h'(t))$$

Claramente esta función es la composición de la función F la cual es diferenciable y la función $g(t,\epsilon) = (t, \varphi(t)+\epsilon h(t), \varphi'(t)+\epsilon h'(t))$, que tiene derivada parcial continua respecto a la segunda variable y vale: $D_2 g(t,\epsilon) = (0, h(t), h'(t))$. Veamos que en estas condiciones existe $D_2 f(t,\epsilon)$. En efecto,

$$D_2 f(t,\epsilon) = \lim_{s\to 0} \frac{f(t,\epsilon+s)-f(t,\epsilon)}{s} =$$

$$\lim_{s\to 0} \frac{F(g(t,\epsilon)+[g(t,\epsilon+s)-g(t,\epsilon)])-F(g(t,\epsilon))}{s}$$

Ahora como F es diferenciable sabemos que el límite anterior se puede expresar:

$$\lim_{s\to 0} DF_{g(t,\epsilon)}(\frac{g(t,\epsilon+s)-g(t,\epsilon)}{s}) + \frac{|s|}{s}\|\frac{g(t,\epsilon+s)-g(t,\epsilon)}{s}\|\xi(s)$$

Teniendo en cuenta las consideraciones habituales sobre la diferenciabilidad se tiene que

$$D_2 f(t,\epsilon) = DF_{g(t,\epsilon)}(D_2 g(t,\epsilon)) = D_2 F(g(t,\epsilon))h(t) + D_3 F(g(t,\epsilon))h'(t).$$

Una vez visto este resultado, como la función f está en las condiciones del lema anterior. Si definimos

$$\phi(\epsilon) = \int_a^b f(t,\epsilon)dt$$

se tendrá que ϕ es derivable en $\epsilon = 0$ y además su derivada será:

$$\delta J(\varphi,h) = \phi'(0) = \int_a^b [D_2 F(t,\varphi(t),\varphi'(t))h(t) + D_3 F(t,\varphi(t),\varphi'(t))h'(t)]dt. \tag{2.1}$$

2.2. Lemas variacionales

Como se ha venido poniendo de manifiesto en el apartado anterior, deseamos estudiar los mínimos locales de los funcionales de Lagrange. Como hemos visto que dichos funcionales tienen derivada de Gâteaux en todo punto y en toda dirección, tenemos que estudiar: para qué funciones $\varphi \in D_G$ se cumple que $0 = \delta J(\varphi, h) \ \ \forall h \in \mathcal{C}^1[a,b]$.

Los siguientes resultados de Cálculo elemental serán de utilidad para nuestro proposito:

Lema 2.2.1 *Sea $f : [a,b] \to \mathbb{R}$ una función continua.*

(a) Si $\int_a^b f(t)h(t)dt = 0$ para toda función $h \in \mathcal{C}[a,b]$. Entonces f es identicamente cero en $[a,b]$.

(b) Si $\int_a^b f(t)h(t)dt = 0$ para toda función $h \in \mathcal{C}[a,b]$ verificando que $\int_a^b h(t)dt = 0$. Entonces f es constante en $[a,b]$.

Prueba.

(a) Basta tomar $h = f$. entonces

$$\int_a^b f^2(t)dt = 0$$

y como f^2 es continua y no negativa, se obtiene que $f \equiv 0$.

(b). Llamemos $c := \frac{\int_a^b f(t)dt}{b-a}$ y consideremos la función $\varphi := f - c$. Es claro que

$$\int_a^b \varphi(t)dt = \int_a^b (f(t) - c)dt = 0,$$

y entonces, por hipótesis se tiene que

$$\int_a^b f(t)\varphi(t)dt = 0,$$

con lo cual

$$\int_a^b (f(t) - c)^2 dt = \int_a^b (f(t) - c)\varphi(t)dt =$$
$$\int_a^b f(t)\varphi(t)dt - c\int_a^b \varphi(t)dt = 0$$

Por la continuidad de φ y de lo anterior se deduce:

$$(f(t) - c)^2 = 0 \ \ \forall t \in [a,b],$$

y por lo tanto $f \equiv c$.

Lema 2.2.2 (du Bois-Reymond) *Sea $f : [a,b] \to \mathbb{R}$ una función continua. Si*

$$\int_a^b f(t)h'(t)dt = 0 \quad \forall h \in \mathcal{D}_0([a,b]),$$

donde $\mathcal{D}_0([a,b]) := \{h \in \mathcal{C}^1[a,b] \ : \ h(a) = h(b) = 0\}$. Entonces f es constante en $[a,b]$.

Prueba.
Si c es cualquier función constante, la función

$$h(x) := \int_a^x (f(t) - c)dt$$

es derivable y además,

$$h'(t) = f(t) - c$$

que es continua en $[a,b]$. Además $h(a) = 0$.

Por lo tanto, si elegimos

$$c = \frac{1}{b-a}\int_a^b f(t)dt,$$

entonces

$$h(b) = \int_a^b (f(t) - c)dt = \int_a^b f(t)dt - c(b-a) = 0.$$

Con lo cual, podemos asegurar que $h \in \mathcal{D}_0([a,b])$.

Consecuentemente,

$$\int_a^b (f(t) - c)^2 dt = \int_a^b (f(t) - c)h'(t)dt =$$

$$= \int_a^b f(t)h'(t)dt - c(h(b) - h(a)) = 0.$$

Luego, por la continuidad de f, concluimos que $f \equiv c$.

Proposición 2.2.3 *Sean $f, g : [a,b] \to \mathbb{R}$ dos funciones continuas. Si*

$$\int_a^b [g(t)h(t) + f(t)h'(t)]dt = 0$$

para toda $h \in \mathcal{D}_0([a,b])$. Entonces $f \in \mathcal{C}^1[a,b]$ y $f' = g$.

Prueba.
Sea $G : [a.b] \to \mathbb{R}$ la función definida por

$$G(x) := \int_a^x g(t)dt.$$

Por el teorema fundamental del Cálculo, G es derivable con $G' = g$. Además, G es absolutamente continua en $[a, b]$. Entonces integrando por partes,

$$0 = \int_a^b [g(t)h(t) + f(t)h'(t)]dt = \int_a^b g(t)h(t)dt + \int_a^b f(t)h'(t) =$$

$$= [h(t)G(t)]_a^b - \int_a^b G(t)h'(t)dt + \int_a^b f(t)h'(t)dt =$$

$$= \int_a^b (f(t) - G(t))h'(t)dt.$$

En definitiva, tenemos que para toda $h \in \mathcal{D}_0[a, b]$,

$$\int_a^b (f(t) - G(t))h'(t)dt = 0,$$

lo que nos dá, segun el lema de du Bois Reymond, que existe una constante c tal que

$$f(t) - G(t) = c \quad \forall t \in [a, b].$$

Esto implica que $f = G + c \in \mathcal{C}^1[a, b]$ y que $f' = G' = g$, como se pretendía demostrar.

Corolario 2.2.4 *Sea $g : [a, b] \to \mathbb{R}$ una función continua. Si*

$$\int_a^b g(t)h(t)dt = 0$$

para toda $h \in \mathcal{D}_0[a, b]$. Entonces g es la función identicamente nula.

Prueba.
Es suficiente considerar en la proposición anterior la función f identicamente nula.

2.3. Ecuación de Euler-Lagrange.

Como aplicación directa de los lemas variacionales, estudiamos el problema de minimizar $J : D_G \to \mathbb{R}$, donde G es un abierto de $\mathbb{R}^3$, F es una función real de clase $\mathcal{C}^1(G)$.

Como hemos visto en las secciones anteriores D_G es un abierto de $(\mathcal{C}^1[a,b], \|.\|)$ y el funcional

$$J(\varphi) = \int_a^b F(t, \varphi(t), \varphi'(t))dt$$

tiene varición de Gâteaux:

$$\delta J(\varphi, h) = \int_a^b [h(t)D_2F(t, \varphi(t), \varphi'(t)) + h'(t)D_3F(t, \varphi(t), \varphi'(t))]dt$$

Si $\varphi \in D_G$ es un extremo local de J, entonces sabemos que para toda $h \in \mathcal{C}^1[a,b]$,

$$0 = \delta J(\varphi, h) = \int_a^b [h(t)D_2F(t, \varphi(t), \varphi'(t)) + h'(t)D_3F(t, \varphi(t), \varphi'(t))dt$$

Si ahora definimos las funciones:

$$g(t) = D_2F(t, \varphi(t), \varphi'(t)), \quad f(t) = D_3F(t, \varphi(t), \varphi'(t))$$

Está claro que ambas funciones son continuas en $[a,b]$, y podemos aplicar la proposición 2.2.3, obteniendo:

$$f \in \mathcal{C}^1[a,b] \text{ y } f' = g,$$

o sea que φ necesariamente debe cumplir

$$\frac{d}{dt}D_3F(t, \varphi(t), \varphi'(t)) = D_2F(t, \varphi(t), \varphi'(t)) \tag{2.2}$$

que se llama *Ecuación de Euler-Lagrange* asociada al funcional J.

Las soluciones de la ecuación de Euler-Lagrange se llaman **extremales** de J, y no tienen por qué ser extremos locales del funcional J pero, si φ es un extremo local de J, entonces debe ser una solución de la ecuación.

Vamos a precisar un poco más:

Si $F \in \mathcal{C}^1(G)$. Siguiendo el razonamiento anterior, si φ es un extremal, tanto h como

$$t \to D_3F(t, \varphi(t), \varphi'(t))$$

tienen derivadas continuas en $[a, b]$ podemos integrar por partes en la segunda integral de la siguiente igualdad

$$0 = \int_a^b D_2F(t, \varphi(t), \varphi'(t))h(t)dt + \int_a^b h'(t)D_3F(t, \varphi(t), \varphi'(t))dt$$

y queda

$$\int_a^b h'(t)D_3F(t, \varphi(t), \varphi'(t))dt =$$

$$[h(t)D_3F(t, \varphi(t), \varphi'(t))]_a^b - \int_a^b h(t)\frac{d}{dt}D_3F(t, \varphi(t), \varphi'(t))dt$$

Sustituyendo, tenemos que para cada h nos queda:

$$0 = \delta J(\varphi, h) =$$

$$\int_a^b h(t)[D_2F(t, \varphi(t), \varphi'(t)) - \frac{d}{dt}D_3F(t, \varphi(t), \varphi'(t))]dt + [h(t)D_3F(t, \varphi(t), \varphi'(t))]_a^b$$

Si φ cumple la ecuación de Euler-Lagrange el integrando de la última integral es nulo, lo que nos dice que

$$0 = [h(t)D_3F(t, \varphi(t), \varphi'(t))]_a^b$$

Escogiendo, en particular, una función h tal que $h(b) = 0$, y $h(a) \neq 0$ quedará la ecuación

$$D_3F(a, \varphi(a), \varphi'(a)) = 0$$

Escogiendo, en particular, una función h tal que $h(a) = 0$, y $h(b) \neq 0$ queda:

$$D_3F(b, \varphi(b), \varphi'(b)) = 0$$

Luego cualquier extremo local φ de J debe ser solución del problema de contorno:

$$\left\{ \begin{array}{l} \frac{d}{dt}D_3F(t, \varphi(t), \varphi'(t)) - D_2F(t, \varphi(t), \varphi'(t)) = 0 \\ D_3F(a, \varphi(a), \varphi'(a)) = 0 \\ D_3F(b, \varphi(b), \varphi'(b)) = 0. \end{array} \right. \tag{2.3}$$

2.3.1. Extremos locales en conjuntos generales.

Sea $(X, \|.\|)$ un espacio normado, D un subconjunto no vacío de X y $J : D \to \mathbb{R}$ un funcional real.

Definición 2.3.1 *Un punto $y_0 \in D$ se dice que es un mínimo local de J sobre D si existe $r > 0$ tal que y_0 es un mínimo absoluto de J sobre el conjunto $D_r(y_0) := \{y \in D \;:\; \|y - y_0\| < r\}$.*

Si queremos minimizar el funcional J sobre D, teniendo en cuenta lo obtenido para conjuntos abiertos, es natural considerar para cada $y \in D$ aquellas direciones $v \in X$ en las cuales la restrición de J a D admite variación en y, i.e.; deseamos distinguir aquellas direcciones $v \in X$ para las cuales:

(i) $y + \epsilon v \in D$ para todo ϵ suficientemente pequeño; y

(ii) Que exista $\delta J(y, v)$.

A dichas direcciones las llamaremos *admisibles* para y en D, o $D-$admisibles en y. Observemos que si v es $D-$admisibles en y, entonces cada múltiplo cv para $c \in \mathbb{R}$ es siempre admisible.

No es difícil ver que.

Proposición 2.3.2 *Si y_0 es un extremo local para J en D, entonces*

$$\delta J(y_0, v) = 0 \;\; \forall \text{ direccion } v \; D - \text{admisible en } y_0$$

En el caso de querer minimizar $J : D \subset D_G \cap \{\varphi(a) = a_1, \; \varphi(b) = b_1\} \to \mathbb{R}$, gracias a la proposicion 2.3.2 obtenemos la siguiente ecuación de Euler-Lagrange:

Si existe $y \in D$ tal que $\delta J(y : v) = 0$ para toda v admisible entonces, se verifica la ecuación 2.2, i.e.,

$$\frac{d}{dt} D_3F(t, \varphi(t), \varphi'(t)) = D_2F(t, \varphi(t), \varphi'(t))$$

Ahora , no podemos seguir tal como haciamos en el caso anterior puesto que en nuestro caso la funciones v admisibles verifican $v(a) = v(b) = 0$.

2.4. Casos especiales.

En general, es difícil encontrar cualquier solución para la ecuación de Euler-Lagrange. Sin embargo, cuando una o más de las variables de F no aparecen explicitamente, entonces podemos al menos simplificar la ecuación diferencial. En esta sección analizaremos tres casos.

(a) Cuando $F(x, y, z) = F(z)$.

En este caso como $D_2F(x, y, z) \equiv 0$ la ecuación de Euler-Lagrange se reduce a:

$$\frac{d}{dt}D_3F(\varphi'(t)) = 0$$

lo cual nos dice que $D_3F(\varphi'(t))$ es constante. Por lo tanto, si φ es un extremal del funcional J correspondiente, entonces φ' debe pertenecer a un conjunto de nivel de D_3F, esto siempre ocurre cuando φ es una función afín, ya que en este caso su derivada es constante.

Como ejemplo de este caso vamos a estudiar el primer ejemplo que aparece en la introducción:

Ejemplo 2.4.1 *Sean $(a, c), (b, d)$ dos puntos del plano. Veamos que la función $\varphi_0(t) = (b-t)\frac{c}{b-a}+(t-a)\frac{d}{b-a}$ es la función de $\mathcal{C}^1[a, b]$ de menor longitud entre todas las que verifican que $\varphi(a) = c$ y $\varphi(b) = d$.*

Como vimos en la introducción se trata de minimizar el funcional

$$J(\varphi) = \int_a^b \sqrt{1 + (\varphi'(t))^2}dt$$

sobre el conjunto $D_G = \mathcal{C}^1[a, b] \cap \{\varphi\ : \varphi(a) = c,\ \varphi(b) = d\}$.

En este caso, la función $F(x, y, z) = \sqrt{1 + z^2}$, que evidentemente es de clase $\mathcal{C}^1$, es una función del tipo (a). Por lo tanto, las funciones afines son extremales del funcional J. Entre las funciones afines la única que verifica la condición $\varphi(a) = c,\ \varphi(b) = d$, es

$$\varphi_0(t) = (b - t)\frac{c}{b - a} + (t - a)\frac{d}{b - a}.$$

(b) Cuando $F(x, y, z) = F(x, z)$.

En este caso de nuevo se tiene que $D_2F(x, y, z) = 0$ con lo cual la Ecuación de Euler-Lagrange nos dice que

$$D_3F(t, \varphi'(t)) = \text{constante}$$

Para ilustrar este caso veamos el siguiente ejemplo.

Ejemplo 2.4.2 *Geodésicas sobre una esfera.*

Para las compañías aereas, es esencial saber la ruta más corta entre dos ciudades. De hecho, como la Tierra es redonda puede considerarse como una esfera, es claro que esto requiere el conocimiento de las geodésicas sobre una superficie esférica.

Cada punto $Y = (x, y, z)$ localizado sobre la superficie de la esfera de radio R (excepto los polos) se pueden expresar mediante coordenadas esféricas de la forma siguiente:

$$Y = (R\cos(\theta)\sin(\varphi), R\sin(\theta)\sin(\varphi), R\cos(\varphi))$$

para un único $\varphi \in]0, \pi[$ y un único $\theta \in [0, 2\pi[$. Además, dados dos puntos distintos A y B sobre la superfície, supondremos que los ejes están elegidos de forma que A está en el polo norte ($\varphi = 0$), mientras que B tiene coordenadas esféricas (R, θ_0, φ_1) para $\varphi_1 > 0$ Entonces una curva que une A con B sobre la superficie de la esfera estará determinada por la curva $v : [0, 1] \to \mathbb{R}^2$ dada por

$$v(t) = (\varphi(t), \theta(t))$$

y de forma que $\varphi(0) = 0$, $\theta(1) = \theta_0$, $\varphi(1) = \varphi_1$ (consideremos además que ambas funciones son continuamente diferenciables). Entonces las curvas que une los dos puntos vendrán dadas por

$$Y(t) = R(\cos(\theta(t))\sin(\varphi(t)), \sin(\theta(t))\sin(\varphi(t)), \cos(\varphi(t)))$$

Calculando su derivada y su longitud nos queda

$$L(Y) = \int_0^1 \|Y'(t)\|dt = R\int_0^1 \sqrt{\sin^2(\varphi(t))\theta'(t)^2 + \varphi'(t)^2}dt$$

El problema que queremos resolver es el de encontrar la v que nos dé la geodésica de longitud más corta.

El problema que abarcamos aquí, será el anterior pero restringido a aquellas curvas v que són gráficas de funciones. Así en este caso sabemos que existe $y : [0, \varphi_1] \to \mathbb{R}$ tal que $G(y) = v([0, 1])$ y de forma que $\theta(t) = y(\varphi(t))$.

Con lo cual utilizando el teorema de cambio de variable el funcional que debemos minimizar será

$$J(y) = R\int_0^{\varphi_1} \sqrt{1 + (y'(\varphi)\sin(\varphi))^2}d\varphi$$

sobre el conjunto

$$D = \{y \in \mathcal{C}^1[0, \varphi_1] \ : \ y(\varphi_1) = \theta_0\}$$

Luego el funcional J está asociado a la función $F(x,y,z) = F(x,z) = R\sqrt{1+(z\sin(x))^2}$ Claramente esta función es de clase $\mathcal{C}^1$, y es del tipo (b). Como

$$D_3F(\varphi,z) = \frac{Rz\sin^2(\varphi)}{\sqrt{1+z^2\sin^2(\varphi)}} = k$$

Como cuando $\varphi = 0$ la expresión anterior es nula. Los extremales verifican:

$$\frac{Ry'(\varphi)\sin^2(\varphi)}{\sqrt{1+(y'(\varphi))^2\sin^2(\varphi)}} = 0$$

Lo cual nos dice que $y'(\varphi) = 0$ y por tanto $y(\varphi)$ es constante. Y esto quiere decir que $\theta(t)$ es constante. Como por hipótesis $\theta(1) = \theta_0$, se tendrá que $\theta(t) = \theta_0$. Lo que corresponde al arco de un circunferencia de radio R que une los dos puntos i.e.,

$$Y(t) = (R\cos(\theta_0)\sin(\varphi(t)), R\sin(\theta_0)\sin(\varphi(t)), R\cos(\varphi(t))).$$

(c) Caso $F(x,y,z) = F(y,z)$.

Si suponemos que φ es de clase $\mathcal{C}^2$, entonces podemos aplicar la regla de la cadena, como $D_1F = 0$, para obtener que

$$\frac{d}{dt}F(t,\varphi(t),\varphi'(t)) = \varphi'(t)D_2F(t,\varphi(t),\varphi'(t)) + D_3F(t,\varphi(t),\varphi'(t))\varphi''(t)$$

Con lo cual

$$\frac{d}{dt}[F(t,\varphi(t),\varphi'(t) - \varphi'(t)D_3F(t,\varphi(t),\varphi'(t))] =$$

$$\frac{d}{dt}F(t,\varphi(t),\varphi'(t)) - \varphi''(t)D_3F(t,\varphi(t),\varphi'(t)) - \varphi'(t)\frac{d}{dt}D_3F(t,\varphi(t),\varphi'(t))$$

Ahora sustituyendo y cancelando nos queda:

$$\frac{d}{dt}[F(t,\varphi(t),\varphi'(t) - \varphi'(t)D_3F(t,\varphi(t),\varphi'(t))] =$$

$$-\varphi'(t)[\frac{d}{dt}D_3F(t,\varphi(t),\varphi'(t)) - D_2F(t,\varphi(t),\varphi'(t)),$$

ahora teniendo presente que φ verifica la ecuación de Euler-Lagrange, podemos concluir:

$$F(t,\varphi(t),\varphi'(t)) - \varphi'(t)D_3F(t,\varphi(t),\varphi'(t)) = \text{constante}$$

Aquí hay que resaltar que si la última igualdad es cierta en un intervalo donde φ no se anula, entonces φ verifica la ecuación de Euler-Lagrange.

Ejemplo 2.4.3 (La braquistocrona) *Según hemos visto en la introducción y eligiendo adecuadamente los puntos. Este problema se reduce a obtener los mínimos del funcional*

$$J(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1+y'(t)^2}}{\sqrt{y(t)}} dt$$

sobre el conjunto

$$D = \{y \in \mathcal{C}^1[0,x_1] \ : y \geq 0, \ y(0)=0, \ y(x_1)=y_1 \text{ y } \int_0^{x_1} \frac{1}{\sqrt{y(t)}} dt < \infty\}.$$

El conjunto D donde tenemos que buscar los mínimos de J es un conjunto sensiblemente diferente a los que hemos visto en los ejemplos anteriores, ya que aquí aparte de tener las condiciones sobre los extremos del intervalo $[0,x_1]$ tenemos además una condición frontera, es decir, sólo podemos aceptar funciones y de forma que $\int_0^{x_1} \frac{1}{\sqrt{y(t)}} dt < \infty$. La condición frontera nos lleva a que las direcciones admisibles para las que se puede calcular la variación de Gâteaux no sea muy amplia. En efecto,

Dado $y \in D$, para que $J(y+\epsilon v)$ esté bien definido debemos tener que $y+\epsilon v \geq 0$ siempre que $|\epsilon|$ sea suficientemente pequeño. Además, como hemos visto en ejemplos precedentes $v(0) = v(x_1) = 0$. Finalmente, si consideramos $|v(t)| \leq y(t)$ se tiene que v es una posible dirección admisible en y. En efecto, para una tal v tomamos $|\epsilon| \leq \frac{1}{2}$:

$$y(x) + \epsilon v(x) \geq y(x) - |\epsilon| v(x) \geq y(x) - \frac{1}{2} y(x) = \frac{1}{2} y(x) \geq 0.$$

Además,

$$\int_0^{x_1} \frac{dx}{\sqrt{y(x)+\epsilon v(x)}} \leq \sqrt{2} \int_0^{x_1} \frac{dx}{\sqrt{y(x)}} < +\infty.$$

Llamando $F(x,y,z) = \frac{1}{\sqrt{2g}} \frac{\sqrt{1+z^2}}{\sqrt{y}}$, las derivadas parciales de F son:

$D_2F(x,y,z) = \frac{-1}{2\sqrt{2g}}\frac{\sqrt{1+z^2}}{\sqrt{y^3}}$,
$D_3F(x,y,z) = \frac{z}{\sqrt{2g}\sqrt{1+z^2}\sqrt{y}}$.

Con lo cual, haciendo uso de la derivación bajo la integral (ver el ranonamiento hecho despues del lema 2.1.2), se obtiene:

$$\delta J(y,v) = \frac{1}{\sqrt{2g}}\int_0^{x_1}[-\frac{1}{2}\frac{\sqrt{1+y'(t)^2}}{\sqrt{y(t)^3}}v(t) + \frac{y'(t)}{\sqrt{y(t)}\sqrt{1+y'(t)^2}}v'(t)]dt$$

La cual es finita siempre que

$$\int_0^{x_1}\frac{1}{\sqrt{y(t)^3}}dt$$

sea finita. Por lo tanto, llegamos a que el problema de la Braquistocrona se puede plantear sobre el conjunto

$$D_1 := \{y \in \mathcal{C}^1[0,x_1]\ : y \geq 0,\ y(0)=0,\ y(x_1)=y_1 \text{ y } \int_0^{x_1}\frac{1}{\sqrt{y(t)^3}}dt < \infty\}.$$

En este caso $F(t,y,z) = F(y,z) = \frac{\sqrt{1+z^2}}{\sqrt{y}}$. Luego estamos en las condiciones del caso (c) y por lo tanto la ecuación de Euler-Lagrange nos dice que si buscamos un extremal de clase $\mathcal{C}^2$ deberá verificar:

$$\frac{\sqrt{1+y'(t)^2}}{\sqrt{y(t)}} - y'(t)(\frac{y'(t)}{\sqrt{y(t)}\sqrt{1+y'(t)^2}}) = \text{constante}.$$

Lo que es equivalente a escribir:

$$\frac{1}{\sqrt{y(t)}\sqrt{1+y'(t)^2}} = \frac{1}{c},$$

y por lo tanto

$$y(t)(1+y'(t)^2) = c^2.$$

Para resolver la ecuación anterior, sustituimos y' por $\frac{dy}{dx}$ y separamos variables, entonces tenemos:

$$dx = (\frac{y}{c^2-y})^{\frac{1}{2}}dy.$$

Introduzcamos una nueva variable ϕ haciendo

$$(\frac{y}{c^2 - y})^{\frac{1}{2}} = \tan(\phi)$$

de modo que $y = c^2 \sin^2(\phi)$, $dy = 2c^2 \sin(\phi)\cos(\phi)d\phi$, y

$$dx = \tan(\phi)dy = 2c^2 \sin^2(\phi)d\phi = c^2(1 - \cos(2\phi))d\phi$$

Integrando queda

$$x = \frac{c^2}{2}(2\phi - \sin(2\phi)) + c_1$$

La solución que buscamos debe pasar por el origen, luego se debe cumplir que $x = y = 0$ cuando $\phi = 0$, en consecuencia, $c_1 = 0$. Así pues,

$$x = \frac{c^2}{2}(2\phi - \sin(2\phi)), \quad y = c^2 \sin^2(\phi) = \frac{c^2}{2}(1 - \cos(2\phi)).$$

Si ahora hacemos $a = \frac{c^2}{2}$ y $\theta = 2\phi$, entonces las ecuaciones anteriores se convierten en

$$\begin{cases} x = a(\theta - \sin(\theta)) \\ \\ y = a(1 - \cos(\theta)), \end{cases}$$

Estas son las ecuaciones paramétricas de la cicloide, generada por un punto sobre una circunferencia de radio a que rueda por el eje de abcisas. De este modo, vemos que las únicas funciones de clase $\mathcal{C}^2$ que verifican la ecuación de Euler-Lagrange son las cicloides.

2.5. Condiciones suficientes para un mínimo.

Las ecuaciones de Euler-Lagrange son condiciones necesarias pero no suficientes para caracterizar un mínimo local del funcional:

$$J(y) := \int_a^b F(t, y(t), y'(t))dt$$

sobre el conjunto de la forma

$$\mathcal{D} := \{y \in \mathcal{C}^1([a,b]) \ : \ y(a) = a_1, \ y(b) = b_1\}.$$

Sin embargo, en presencia de convexidad (fuerte) de $F(\underline{x}, y, z)$ caracterizan de forma única al mínimo.

2.5.1. Funciones convexas.

Si $F \in \mathcal{C}^1(\mathbb{R}^3)$, en este caso, por la interpretación de la derivada de Gâteaux, se tiene que dado $y \in \mathbb{R}^3$, $v \in \mathbb{R}^3$

$$\delta F(y, v) = \langle \nabla F(y), v \rangle,$$

además si F es una función convexa se verifica:

$$\langle \nabla F(y), v \rangle = D_v F(y) = \lim_{t \to 0} \frac{F(y + tv) - F(y)}{t} =$$
$$\lim_{t \to 0} \frac{F(t(y + v) + (1 - t)y) - F(y)}{t} \leq$$
$$\lim_{t \to 0} \frac{tF(y + v) + (1 - t)F(y) - F(y)}{t} = F(y + v) - F(y).$$

De donde se deduce que:

$$F(y + v) - F(y) \geq \langle \nabla F(y), v \rangle = \delta F(y, v),$$

y es estrictamente convexa cuando la igualdad, en la desigualdad anterior, sólamente se dá cuando $v = 0$.

También, podemos observar que minimizar una función convexa es obtener que $\nabla F(y) = 0$ con lo cual claramente y es un mínimo de F. Este hecho sugiere la siguiente definición

Definición 2.5.1 *Sea X un espacio normado y $D \subset X$ un subconjunto no vacío de X. Un funcional $J : D \to \mathbb{R}$ diremos que es [estrictamente] convexo sobre D si cuando $y, y + v \in D$ entonces $\delta J(y, v)$ está definido y además*

$$J(y + v) - J(y) \geq \delta J(y, v)$$

[con igualdad sii $v = 0$]

Observación 2.5.2 *En esta definición no se supone que el dominio del funcional sea convexo. Pero en el caso en que $\delta J(y, v) = 0$ la definición anterior nos confirma que y es un mínimo de J en la dirección de v.*

Proposición 2.5.3 *Sea X un espacio normado y D un subconjunto no vacío de X. Si $J : D \to \mathbb{R}$ es un funcional [estrictamente] convexo, entonces para cada $y_0 \in D$ tal que $\delta J(y_0, v) = 0$ $\forall y_0 + v \in D$ se tiene que y_0 es un mínimo de J sobre D [único].*

Prueba.

Como queremos probar que y_0 es un mínimo, lo que tenemos que ver es que $\forall y \in D$ se verifica que $J(y_0) \leq J(y)$.

Dado un punto $y \in D$, se cumple que $y = y_0 + (y - y_0)$ por lo tanto llamando $v = y - y_0$ se tiene que $y_0, y_0 + v \in D$. Ahora, aplicando la definición de funcional convexo se tiene que existe $\delta J(y_0, v)$ y además

$$J(y_0 + v) - J(y_0) \geq \delta J(y_0, v).$$

Finalmente, como por hipótesis $\delta J(y_0, v) = 0$ se deduce que

$$J(y_0) \leq J(y_0 + v) = J(y).$$

2.5.2. Funcionales de Lagrange convexos

Para simplificar los cálculos supondremos que $F \in \mathcal{C}^1([a,b] \times \mathbb{R}^2)$ y definimos

$$J(y) = \int_a^b F(t, y(t), y'(t))dt,$$

como hemos visto en (2.1) se tiene que para qualesquiera $y, v \in \mathcal{C}^1([a,b])$

$$\delta J(y,v) = \phi'(0) = \int_a^b [D_2F(t,y(t),y'(t))v(t) + D_3F(t,y(t),y'(t))v'(t)]dt.$$

Si queremos ver la convexidad de J hemos de estudiar si se verifica la desigualdad

$$J(y+v) - J(y) \geq \delta J(y,v)$$

o lo que es lo mismo

$$\int_a^b F(t, y+v, y'+v')dt - \int_a^b F(t,y,y')dt \geq \delta J(y,v). \tag{2.4}$$

Ahora teniendo en cuenta que la integral es monótona, si probáramos que para cada $t \in [a,b]$ se verifica:

$$F(t,y+v,y'+v') - F(t,y,y') \geq D_2F(t,y(t),y'(t))v(t) + D_3F(t,y(t),y'(t))v'(t)$$

ya tendríamos (2.4).

Con lo cual si llamamos $y = y(t),\ z = y'(t),\ v = v(t),\ w = v'(t)$ nos queda:

$$F(t, y+v, z+w) - F(t, y, z) \geq$$

$$D_2F(t, y, z)v + D_3F(t, y, z)w \quad \forall (t, y, z), (t, y+v, z+w) \in]a, b[\times \mathbb{R}^2.$$

Lo cual significa que si fijamos $t \in]a, b[$ la función $G : \mathbb{R}^2 \to \mathbb{R}$ definida como:

$$G(y, z) = F(t, y, z)$$

verifica que

$$G(y+v, z+w) - G(y, z) \geq \langle \nabla G(y, z), (v, w) \rangle,$$

es decir, G es una función convexa en $\mathbb{R}^2$.

Esta restricción sobre la convexidad parcial de F cuando fijamos t será esencial para nuestro proposito, además nos sugiere la siguiente definición.

Definición 2.5.4 *$F(\underline{t}, y, z)$ se dice que es [estrictamente] convexa sobre $S \subseteq \mathbb{R}^3$ si $F \in \mathcal{C}^1(S)$ y cumple:*

$$F(t, y+v, z+w) - F(t, y, z) \geq D_2F(t, y, z)v + D_3F(t, y, z)w \qquad (2.5)$$

para cualesquiera $(t, y, z), (t, y+v, z+w) \in S$. [igualdad si, y sólo si, $v = 0$ y $w = 0$].

Teorema 2.5.5 *Sea D un dominio de $\mathbb{R}^2$ y para $a_1, b_1 \in \mathbb{R}$ consideramos el conjunto*

$$\mathcal{D} := \{y \in \mathcal{C}^1([a, b]) \;:\; y(a) = a_1,\; y(b) = b_1,\; (y(t), y'(t)) \in D\}.$$

Si $F(\underline{t}, y, z)$ es [estrictamente] convexa en $[a, b] \times D$, entonces el funcional

$$J(y) = \int_a^b F(t, y(t), y'(t))dt$$

es [estrictamente] convexo sobre $\mathcal{D}$. Además, cada $y \in \mathcal{D}$ que verifique

$$\frac{d}{dt} D_3F(t, y(t), y'(t)) = D_2F(t, y(t), y'(t))$$

es un mínimo de J sobre $\mathcal{D}$ [único].

Prueba

Cuando $y, y+v \in \mathcal{D}$ la desigualdad (2.5) nos dice que para cada $t \in]a,b[$ fijo se verifica:

$$J(y+v) - J(y) \geq \delta J(y,v)$$

con lo cual J es convexo.

Finalmente, si $y \in \mathcal{D}$ verifica que

$$\frac{d}{dt} D_3 F(t, y(t), y'(t)) = D_2 F(t, y(t), y'(t))$$

como además la variación de Gâteaux de J viene dada por

$$\delta J(y,v) = \int_a^b [D_2 F(t,y,y')v + D_3 F(t,y,y')v']dt$$

de ambas igualdades resulta que

$$\delta J(y,v) = \int_a^b \frac{d}{dt}(D_3 F(t,y,y')v)dt =$$

$$v(b) D_3 F(b, y(b), y'(b)) - v(a) D_3 F(a, y(a), y'(a))$$

Ahora bien, como las direcciones admisibles del problema que estamos estudiando son aquellas que cumplen que $v(a) = v(b) = 0$.

Se debe verificar:

$$J(y,v) = \int_a^b \frac{d}{dt}(D_3 F(t,y,y')v)dt = 0.$$

Con estos hechos podemos aplicar la proposición 2.5.3 para concluir que y es un mínimo.

Observando la prueba del resultado anterior no es difícil ver que se verifican los siguientes resultados:

Teorema 2.5.6 *Sea D un dominio de $\mathbb{R}^2$ y consideramos el conjunto*

$$\mathcal{D} := \{y \in \mathcal{C}^1([a,b]) \ : \ (y(t), y'(t)) \in D\}.$$

Si $F(\underline{t}, y, z)$ es [estrictamente] convexa en $[a,b] \times D$, entonces el funcional

$$J(y) = \int_a^b F(t, y(t), y'(t))dt$$

es [estrictamente] convexo sobre $\mathcal{D}$. Además, cada $y \in \mathcal{D}$ que verifique

$$\frac{d}{dt} D_3F(t, y(t), y'(t)) = D_2F(t, y(t), y'(t))$$

$$D_3F(a, y(a), y'(a)) = D_3F(b, y(b), y'(b)) = 0$$

es un mínimo de J sobre $\mathcal{D}$ [único].

Teorema 2.5.7 *Sea D un dominio de $\mathbb{R}^2$ y para $b_1 \in \mathbb{R}$ consideramos el conjunto*

$$\mathcal{D} := \{y \in \mathcal{C}^1([a, b]) \ : \ y(b) = b_1, \ (y(t), y'(t)) \in D\}.$$

Si $F(\underline{t}, y, z)$ es [estrictamente] convexa en $[a, b] \times D$, entonces el funcional

$$J(y) = \int_a^b F(t, y(t), y'(t))dt$$

es [estrictamente] convexo sobre $\mathcal{D}$. Además, cada $y \in \mathcal{D}$ que verifique

$$\frac{d}{dt} D_3F(t, y(t), y'(t)) = D_2F(t, y(t), y'(t)),$$

$$D_3F(a, y(a), y'(a)) = 0$$

es un mínimo de J sobre $\mathcal{D}$ [único].

Ahora veremos cómo aplicar estos resultados a los ejemplos de la sección anterior.

Ejemplo 2.5.8 *El funcional de Lagrange asociado a la función $F(x, y, z) = \sqrt{1 + z^2}$ estudia, como hemos visto en el Ejemplo 2.4.1 la distancia entre dos puntos del plano. Ahora veremos que el extremal obtenido en tal ejemplo es un mínimo absoluto.*

En efecto, según el resultado anterior será suficiente con demostrar que $F(\underline{t}, y, z)$ es estrictamente convexa. Para ello, dados dos números reales $z, w \in \mathbb{R}$, si $w > 0$ definimos la función $g : [z, z + w] \to \mathbb{R}$ por $g(t) = \sqrt{1 + t^2}$ teniendo presente que las derivadas de dicha función son:

$$g'(t) = \frac{t}{\sqrt{1 + t^2}} \text{ y } g''(t) = \frac{1}{(1 + t^2)\sqrt{1 + t^2}}$$

se cumplirá, aplicando el teorema del valor medio, y teniendo presente que g' es creciente:

$$\begin{aligned} F(t,y+v,z+w)-F(t,y,z) = & \ \sqrt{1+(z+w)^2}-\sqrt{1+z^2} \\ = & \ g(z+w)-g(z) \\ \geq & \ \tfrac{zw}{\sqrt{1+z^2}} \end{aligned}$$

Por otra parte tenemos:

$$D_2F(t,y,z)=0 \ \text{ y } \ D_3F(t,y,z)=\frac{z}{\sqrt{1+z^2}}$$

Con lo cual

$$F(t,y+v,z+w)-F(t,y,z) \geq vD_2F(t,y,z)+wD_3F(t,y,z).$$

En el caso en que $w<0$ razonamos de la misma forma tomando la función: $h:[z+w,z]\to\mathbb{R}$ definda como $h(t)=\sqrt{1+t^2}$.

Ejemplo 2.5.9 *El funcional de Lagrange asociado a la función* $F(x,y,z)=\sqrt{1+(z\sin(x))^2}$ *estudia, como hemos visto en el Ejemplo 2.4.2 las geodésicas sobre un esfera. Ahora veremos que el extremal obtenido en tal ejemplo es un mínimo absoluto.*

En efecto, según el Teorema 2.5.7 será suficiente con demostrar que $F(\underline{t},y,z)$ es estrictamente convexa ya que el extremal obtenido en el ejemplo (2.4.2) verifica: $D_3F(0,y(0),y'(0))=0$.

Razonando como en el ejemplo anterior, se llega a que

$$\begin{aligned} F(t,y+v,z+w)-F(t,y,z) = & \ \sqrt{1+((z+w)\sin(t))^2}-\sqrt{1+(z\sin(t))^2} \\ \geq & \ \sin(t)^2\tfrac{zw}{\sqrt{1+(z\sin(t))^2}} \end{aligned}$$

Por otra parte tenemos:

$$D_2F(t,y,z)=0 \ \text{ y } \ D_3F(t,y,z)=\frac{z\sin^2(t)}{\sqrt{1+(z\sin(t))^2}}$$

Con lo cual

$$F(t,y+v,z+w)-F(t,y,z) \geq vD_2F(t,y,z)+wD_3F(t,y,z).$$

2.6. Extremos condicionados.

Sea X un espacio normado. Sea D un subconjunto abierto de X, y sean J, K dos funcionales los cuales están definidos y tienen variación de Gâteaux sobre D.

Consideremos el problema de encontrar extremos para J sobre aquellos vectores de D que satisfacen la restrición

$$K(x) = k$$

Para simplificar la exposición, usaremos el símbolo $D(K = k)$ para representar el conjunto $D \cap \{x \in X \; : \; K(x) = k\}$. Además, siempre supondremos que $D(K = k) \neq \emptyset$.

El problema de extremos que consideraremos consiste en encontrar un extremo local de J de entre los vectores de $D(K = k)$.

Observación 2.6.1 *La variación de J, ahora, no necesariamente se anula sobre el extremo local de J en $D(K = k)$, ya que en general este conjunto no es abierto.*

Ejemplo 2.6.2 *Consideremos el problema de encontrar los extremos del funcional $J(x) = x^2$ sujetos a la restricción $K(x) = x^2 + 2x + \frac{3}{4} = 0$. Es claro que los extremos corresponden a los puntos $x = -\frac{1}{2}$, $x = -\frac{3}{2}$, y en estos puntos $J'(x) \neq 0$.*

El método de Euler-Lagrange nos permite resolver muchos problemas de extremos condicionados sin ninguna consideración directa sobre los conjuntos $D(K = k)$. Para establecer el método de los multiplicadores, necesitamos la noción de *variación débilmente continua.*

Definición 2.6.3 *Si J es un funcional el cual tiene variación sobre un conjunto abierto D contenido en un espacio normado X y para algún vector $x \in D$ se tiene que*

$$\lim_{y \to x} \delta J(y, h) = \delta J(x, h)$$

es cierto para cualquier vector $h \in X$, entonces diremos que la variación de J es débilmente continua en x.

En la práctica sólamente necesitaremos comprobar que para cada $h \in X$ la diferencia $\delta J(y, h) - \delta J(x, h)$ puede hacerse arbitrariamente pequeña para todos aquellos vectores los cuales están lo suficientemente próximos a x.

Ejemplo 2.6.4 *Sea $K : \mathcal{C}^1([a,b]) \to \mathbb{R}$ definido por $K(y) = y(x_0)$, donde $x_0 \in [a,b]$ fijo. Entonces la variación de Gâteaux de K es débilmente continua en cualquier punto del espacio.*

En efecto, es claro que

$$\frac{K(y+\epsilon h) - K(y)}{\epsilon} = h(x_0),$$

por lo tanto $\delta K(y,h) = h(x_0)$. A partir de aquí se ve fácilmente que δK es débilmente continua.

Observación 2.6.5 *La noción de variación débilmente continua es simplemente una generalización a funcionales de la noción de la continuidad de las derivadas parciales de primer orden.*

Teorema 2.6.6 (Teorema de Euler-Lagrange) *Sean $J, K_1, K_2, ..., K_m$ funcionales reales definidos sobre un abierto D de un espacio normado X, tales que admiten variación de Gâteaux débilmente continua en D. Supongamos que x^* es un extremo local de J sobre el conjunto M definido por*

$$D \cap \{x \in X \ : \ K_i(x) = k_i, \ i = 1, ..., m\}.$$

Entonces se cumple al menos una de las siguientes condiciones:
(i) Para cualquiera $h_1, h_2, ..., h_m \in X$, se tiene que

$$det(\delta K_i(x^*, h_j)) = 0.$$

(ii) Existen $\lambda_1, \lambda_2, ..., \lambda_m$ escalares tales que

$$J'(x^*) = \sum_{i=1}^{m} \lambda_i K_i'(x^*)$$

Observación 2.6.7 *El teorema simplemente garantiza que todos los posibles extremos locales deben estar comprendidos entre la colección de puntos que verifican las condiciones (i) o (ii) del teorema.*

Prueba.
Será suficiente probar que la condición (ii) debe ser cierta en el caso en que no se verifique la condición (i).

Supondremos, entonces que se pueden encontrar $\overline{h}_1, ..., \overline{h}_m \in X$ de forma que

$$det(\delta K_i(x^*, \overline{h}_j)) \neq 0 \tag{2.6}$$

En este caso debemos demostrar que (ii) es cierto.

Esto lo demostraremos utilizando (2.6) y el siguiente hecho:

$$\begin{vmatrix} \delta J(x^*, h) & \delta J(x^*, h_1) & .. & . & \delta J(x^*, h_m) \\ \delta K_1(x^*, h) & \delta K_1(x^*, h_1) & .. & . & \delta K_1(x^*, h_m) \\ . & . & .. & . & \\ . & . & .. & . & \\ \delta K_m(x^*, h) & \delta K_m(x^*, h_1) & .. & . & \delta K_m(x^*, h_m) \end{vmatrix} = 0 \tag{2.7}$$

Veremos que el determinante anterior es nulo, desarrollandolo por los adjuntos de la primera columma se tiene:

$$\delta J(x^*, h) det(\delta K_i(x^*, h_j)) + \sum_{i=1}^{m} \delta K_i(x^*, h) \mu_i(h_1, ..., h_m) = 0$$

donde los escalares $\mu_i(h_1, ..., h_m)$ dependen de $h_1, ..., h_m$. Ahora tomando $h_j = \overline{h}_j$ para $j = 1, 2, ..., m$ y teniendo en cuenta (2.6) se tiene que:

$$\delta J(x^*, h) = \sum_{i=1}^{m} \lambda_i \delta K_i(x^*, h)$$

para todo $h \in X$. Siendo $\lambda_i = -\frac{\mu_i(\overline{h}_1, ..., \overline{h}_m)}{det(\delta K_i(x^*, \overline{h}_j))}$.

Así, para demostrar el teorema es suficiente con obtener (2.7). Lo cual lo realizaremos por reducción al absurdo. Supongamos que existe $h, h_1, ..., h_m \in X$ de forma que el determinante no sea cero.

Como $x^* \in M$ es un extremo local de J en M, entonces existirá una bola $B(x^*)$ contenida en D de forma que x^* será un extremo absoluto de J en $B(x^*) \cap M$, lo cual nos permite afirmar que podemos encontrar:

$$U := \{(\alpha, \beta_1, ..., \beta_m) \in \mathbb{R}^{m+1} \ : \ \alpha^2 + \beta_1^2 + ... + \beta_m^2 < \rho^2\}$$

de manara que

$$x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m \in B(x^*) \subseteq D.$$

Ahora definimos la función $F : U \subseteq \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ de la siguiente manera:

$$F_1(\alpha, \beta_1, ..., \beta_m) = J(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m)$$

$$F_2(\alpha, \beta_1, ..., \beta_m) = K_1(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m)$$

$$.$$

$$.$$

$$.$$

$$F_{m+1}(\alpha, \beta_1, ..., \beta_m) = K_m(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m).$$

Si ahora calculamos las derivadas parciales de esta función obtenemos que:

$$D_1 F_1(\alpha, \beta_1, ..., \beta_m) = \lim_{\epsilon \to 0} \frac{F_1(\alpha + \epsilon, \beta_1, ..., \beta_m) - F_1(\alpha, \beta_1, ..., \beta_m)}{\epsilon} =$$

$$\lim_{\epsilon \to 0} \frac{J(x^* + (\alpha + \epsilon)h + \beta_1 h_1 + ... + \beta_m h_m) - J(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m)}{\epsilon} =$$

$$\delta J(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m, h).$$

Del mismo modo se obtiene que

$$D_2 F_1(\alpha, \beta_1, ..., \beta_m) = \delta J(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m, h_1)$$

$$...$$

$$D_{m+1} F_1(\alpha, \beta_1, ..., \beta_m) = \delta J(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m, h_m)$$

Así sucesivamente se tendrá:

$$D_1 F_{m+1}(\alpha, \beta_1, ..., \beta_m) = \delta K_m(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m, h)$$

$$...$$

$$D_{m+1} F_{m+1}(\alpha, \beta_1, ..., \beta_m) = \delta K_m(x^* + \alpha h + \beta_1 h_1 + ... + \beta_m h_m, h_m)$$

Por lo tanto, dado que $J', K_1', ..., K_m'$ són débilmente continuas en D, podemos considerar que F es una función de clase $\mathcal{C}^1(U; \mathbb{R}^{m+1})$ cuyo determinante Jacobiano en el punto $(0, 0, ..., 0)$ es no nulo, i.e.,

$$\det(JF(0,0,...,0)) = \begin{vmatrix} \delta J(x^*, h) & \delta J(x^*, h_1) & .. \;. & \delta J(x^*, h_m) \\ \delta K_1(x^*, h) & \delta K_1(x^*, h_1) & .. \;. & \delta K_1(x^*, h_m) \\ . & . & .. \;. & \\ . & . & .. \;. & \\ \delta K_m(x^*, h) & \delta K_m(x^*, h_1) & .. \;. & \delta K_m(x^*, h_m) \end{vmatrix} \neq 0$$

Aplicando ahora el teorema de la función inversa a F en el origen tendremos; que existirá un abierto V de $\mathbb{R}^{m+1}$ tal que $F(0, ..., 0) = (J(x^*), k_1, ..., k_m) \in$

V. Además, existirá una función $G : V \to U$ de clase $\mathcal{C}^1$ de forma que para todo $(j, y_1, ..., y_m) \in V$ se tiene que $(j, y_1, ..., y_m) = F(G(j, y_1, ..., y_m))$. Es decir;

$$j = F_1(G(j, y_1, ..., y_m)) = J(x^* + G_1 h + ... + G_{m+1} h_m)$$

$$y_1 = F_2(G(j, y_1, ..., y_m)) = K_1(x^* + G_1 h + ... + G_{m+1} h_m)$$

$$....$$

$$y_m = F_{m+1}(G(j, y_1, ..., y_m)) = J(x^* + G_1 h + ... + G_{m+1} h_m).$$

En particular, $G(J(x^*), k_1, ..., k_m) = (0, ..., 0)$, como $(J(x^*), k_1, ..., k_m) \in V$, y V es un subconjunto abierto, entonces podremos encontrar $(j_1, k_1, ..., k_m)$, $(j_2, k_1, ..., k_m) \in V$ de forma que $j_1 < J(x^*) < j_2$. Si ahora calculamos las imágenes de estos dos puntos tendremos:

$$(\alpha^*, \beta_1^*, ..., \beta_m^*) = G(j_1, k_1, ..., k_m)$$

$$(\alpha^{**}, \beta_1^{**}, ..., \beta_m^{**}) = G(j_2, k_1, ..., k_m).$$

Y los vectores

$$x^* + \alpha^* h + \beta_1^* h_1 + ... + \beta_m^* h_m = v^*$$

$$x^* + \alpha^{**} h + \beta_1^{**} h_1 + ... + \beta_m^{**} h_m = v^{**}$$

están en la bola $B(x^*)$, además,

$$F(\alpha^*, \beta_1^*, ..., \beta_m^*) = (j_1, k_1, ..., k_m)$$

$$F(\alpha^{**}, \beta_1^{**}, ..., \beta_m^{**}) = (j_2, k_1, ..., k_m)$$

esto significa que

$$J(v^*) = j_1, \ J(v^{**}) = j_2$$

y además,

$$K_1(v^*) = k_1 = K_1(v^{**}), ..., K_m(v^*) = k_m = K_m(v^{**})$$

con lo cual los vectores v^*, v^{**} cumplen las restricciones, pero en cambio se tiene que $J(v^*) < J(x^*) < J(v^{**})$ lo cual es imposible puesto que x^* es un extremo absoluto de J en $M \cap B(x^*)$.

2.6.1. Problema de Lagrange de extremos fijos.

Sea G un subconjunto abierto no vacío de $\mathbb{R}^3$, sea $F \in \mathcal{C}^2(G, \mathbb{R})$. Como es habitual consideremos

$$D_G := \{y \in \mathcal{C}^1[a,b] : \ \forall t \in [a,b], \ (t, y(t), y'(t)) \in G\}.$$

Buscamos los extremos del funcional $J : D_G \to \mathbb{R}$,

$$J(y) = \int_a^b F(t, y(t), y'(t))dt,$$

sobre el conjunto

$$M = \{y \in D_G : \ y(a) = y_0, \ y(b) = y_1\}.$$

Éste es un problema de extremos condicionados, ya que si llamamos:

$$K_1 : D_G \to \mathbb{R} \text{ tal que } K_1(\varphi) = \varphi(a),$$

$$K_2 : D_G \to \mathbb{R} \text{ tal que } K_2(\varphi) = \varphi(b).$$

Entonces se trata de obtener los estremos del funcional J sobre el conjunto $D_G \cap \{K_1 = y_0\} \cap \{K_2 = y_1\}$.

Como hemos visto en el ejemplo 2.6.4 los funcionales K_1, K_2 tienen variación de Gâteaux débilmente continua en todo punto de su dominio.

Por otra parte, si suponemos que F está bajo las hipótesis de los problemas (2.7.7), o ,(2.7.8) se tiene que J admite variación de Gâteaux débilmente continua en todo punto de D_G.

Los hechos anteriores nos permiten aplicar el teorema de Euler-Lagrange para estudiar los estremos del problema. De este modo, si φ es solución del problema debe cumplirse alguna de las siguientes alternativas:

(a) Para cualesquiera $h_1, h_2 \in \mathcal{C}^1[a,b]$, se debe cumplir:

$$\det(\delta K_i(\varphi, h_j)) = 0.$$

(b) Existen dos números reales λ_1, λ_2 de forma que para cada $h \in \mathcal{C}^1[a,b]$

$$\delta J(\varphi, h) = \lambda_1 \delta K_1(\varphi, h) + \lambda_2 \delta K_2(\varphi, h).$$

La alternativa (a) supone que para todo $h_1, h_2 \in \mathcal{C}^1[a,b]$,

$$\begin{vmatrix} \delta K_1(\varphi, h_1) & \delta K_1(\varphi, h_2) \\ \delta K_2(\varphi, h_1) & \delta K_2(\varphi, h_2) \end{vmatrix} = \begin{vmatrix} h_1(a) & h_2(a) \\ h_1(b) & h_2(b) \end{vmatrix} = 0$$

lo cual evidentemente no puede darse.

Por lo tanto, se debe cumplir la alternativa (b), es decir que existen dos números reales de forma que

$$\delta J(\varphi, h) = \lambda_1 \delta K_1(\varphi, h) + \lambda_2 \delta K_2(\varphi, h),$$

luego

$$\delta J(\varphi, h) = \lambda_1 h(a) + \lambda_2 h(b).$$

Teniendo en cuenta el valor de la variación de Gâteaux del funcional de Lagrange J se debe cumplir:

$$\int_a^b [D_2 F(t, \varphi(t), \varphi'(t))h(t) + D_3 F(t, \varphi(t), \varphi'(t))h'(t)]dt = \lambda_1 h(a) + \lambda_2 h(b).$$

En particular, si tomamos $h \in \mathcal{C}^1[a, b]$ de forma que $h(a) = h(b) = 0$, se debe verificar:

$$\int_a^b [D_2 F(t, \varphi(t), \varphi'(t))h(t) + D_3 F(t, \varphi(t), \varphi'(t))h'(t)]dt = 0.$$

Como ya sabemos, esta ecuación nos permite afirmar que si φ es una solución, entonces debe cumplir la ecuación de Euler-Lagrange

$$D_2 F(t, \varphi(t), \varphi'(t)) = \frac{d}{dt} D_3 F(t, \varphi(t), \varphi'(t)),$$

junto con las condiciones de frontera $\varphi(a) = y_0,\ \varphi(b) = y_1$.

2.7. Problemas

Problema 2.7.1 *Probar que el espacio $\mathcal{C}^1([a, b])$ con la norma $\|f\| = \|f\|_\infty + \|f'\|_\infty$ es un espacio de Banach.*

Problema 2.7.2 *Minimizar, si es posible, el funcional $J(f) = \int_0^1 (f'(t))^2 dt$, en $\mathcal{C}^1([0, 1])$.*

Problema 2.7.3 *Minimizar, si es posible, el funcional*

$$J(y) = \int_0^1 x^2 (y'(x))^2 dx,$$

sobre los siguientes conjuntos

(a). $M = \{y \in \mathcal{C}^1([0, 1]) : \ y(0) = 0,\ y(1) = 1\}$.

(b). $M = \{y \in \mathcal{C}^1([0, 1]) : \ y(1) = 1\}$.

Problema 2.7.4 *Minimizar, si es posible, el funcional*

$$J(\varphi) = \int_0^1 [(\varphi'(x))^2 - 2x\varphi(x)]dx,$$

sobre el conjunto

$$M := \{\varphi \in \mathcal{C}^1([1,2]) : \ \varphi(1) = 0, \ \varphi(2) = -1\}$$

Problema 2.7.5 *Una persona camina con velocidad constante v_0 desde el origen de coordenadas y a lo largo de una curva dada paramétricamente en función del tiempo t por las ecuaciones*

$$\gamma(t) = (x(t), y(t)).$$

Sea $\alpha(t)$ el ángulo formado por el eje OX y la recta tangente descrita por la persona, en el punto $\gamma(t)$.

Se tiene entonces que

$$x'(t) = v_0 \cos(\alpha(t)), \ y'(t) = v_0 \sin(\alpha(t)).$$

Con lo cual,

$$x(t) = v_0 \int_0^t \cos(\alpha(s))ds, \ y(t) = v_0 \int_0^t \sin(\alpha(s))ds.$$

Si termina su recorrido en algún punto del eje OX en el tiempo $t = T$, entonces el área rodeada por el camino recorrido y el eje OX es

$$A = \frac{1}{2} \int_0^T [x(t)y'(t) - y(t)x'(t)]dt,$$

por lo que puede considerarse como dependiente de α.

Maximizar esta área con la restricción de que el espacio recorrido v_0T ha de ser igual a la longitud del camino recorrido que se supone previamente dado.

Problema 2.7.6 *Sea $y = \gamma(x)$ una curva simple que pasa por los puntos $(0,0)$ y (l, R), con $l, R > 0$. Supongamos que $\gamma(x) \geq 0$ para todo $x \in [0, l]$. Si el arco $\{(x, \gamma(x)) \ x \in [0, l]\}$ gira alrededor del eje $= X$ genera un cuerpo de revolución.*

Si un gas de baja densidad se desplaza con velocidad constante en la dirección del vector $(1, 0)$ e incide sobre dicho cuerpo, realiza sobre él una *fuerza (o bien podemos decir que este cuerpo presenta una resistencia) que puede calcularse (siguiendo un razonamiento de Físisca elemental que omitimos) y resulta depender de γ (es decir, de la forma del cuerpo) según la función*

$$F(\gamma) = k \int_0^l \gamma(x)(\gamma'(x))^3 dx.$$

Se trata de encontrar la función $\gamma \in \mathcal{C}^2([0,l])$ que genera la menor resistencia posible.

Comprobar que la ecuación de Euler-Lagrange que genera este problema de extremos fijos es

$$(y')^3 - 3\frac{d}{dx}(yy'^2) = 0,$$

y estudiar sus soluciones.

Sean $G = (a-\epsilon, b+\epsilon) \times \mathbb{R} \times \mathbb{R}$, F una función de clase $\mathcal{C}^2(G,\mathbb{R})$ y $D_G = \{y \in \mathcal{C}^1([a,b]) :\ \forall t \in [a,b],\ (t, y(t), y'(t)) \in G\}$.

Sabemos que el conjunto D_G es abierto en el espacio $\mathcal{C}^1([a,b])$. A continuación vamos a dar condiciones para asegurar que la variación de Gâteaux del funcional de Lagrange

$$J(y) = \int_a^b F(t, y(t), y'(t))dt$$

sea débilmente continua en cada punto de D_G.

Problema 2.7.7 *Si suponemos que las derivadas parciales de segundo orden de F están acotadas sobre el abierto G. Probar que la variación de Gâteaux de J es débilmente continua.*

Problema 2.7.8 *Si suponemos que las derivadas de primer orden de F son uniformemente continuas en el abierto G. Probar que la variación de Gâteaux de J es débilmente continua.*

2.7.1. Problemas isoperimétricos.

Uno de los más antiguos problemas de optimización es el de encontrar el área máxima que se puede encerrar por una curva de perímetro fijo.

De acuerdo con Virgilio, este problema se remonta a la época de la fundación de la ciudad de Cartago (año 850 a.c.) y se conoce como el problema de la reina Dido de Cartago.

En este apartado daremos la siguiente formulación simplificada del problema de Dido:

Entre las gráficas de funciones $\varphi \in \mathcal{C}^1([-1,1])$ que unen los puntos $P = (-1,0)$ y $Q = (1,0)$ determinar si existe alguna función φ de manera que la longitud de su gráfica sea un valor fijo $l > 2$ y de forma que el área encerrada por dicha gráfica y el segmento PQ sea máxima.

El funcional a maximizar es el área, que viene dada por:

$$A(\varphi) = \int_{-1}^{1} \varphi(x)dx.$$

Las restricciones son de dos clases:

Primera, la longitud de la gráfica determinada por φ debe ser l. Con lo cual

$$\int_{-1}^{1} \sqrt{1 + (\varphi'(x))^2}dx = l.$$

Segunda, $\varphi(-1) = 0 = \varphi(1)$.

Por lo tanto se busca un máximo de A sobre el conjunto

$$\mathcal{M} := \{\varphi \in \mathcal{C}^1([-1,1]) : \ K_1(\varphi) = l, \ K_2(\varphi) = 0, \ K_3(\varphi) = 0\}$$

donde,

$$K_1(\varphi) = \int_{-1}^{1} \sqrt{1 + (\varphi'(x))^2}dx.$$

$$K_2(\varphi) = \varphi(-1), \ K_3(\varphi) = \varphi(1).$$

Observación 2.7.9 *Los antiguos griegos llamaron problema isoperimétrico al de hallar la curva cerrada plana de longitud dada que encierra área máxima.*

Problema 2.7.10 *Probar que A, K_1, K_2, K_3 tienen variación de Gâteaux en cualquier punto $\varphi \in \mathcal{C}^1([-1,1])$ y en cualquier dirección $h \in \mathcal{C}^1([-1,1])$.*

Problema 2.7.11 *Probar además que dichas variaciones son débilmente continuas en $\mathcal{C}^1([-1,1])$.*

Según los dos problemas anteriores, estamos bajo las hipótesis del Teorema de Euler-Lagrange. Luego si y es un punto de $\mathcal{M}$ en el que A tiene un extremo local se debe cumplir una de las siguientes alternativas:

(i) Para cualquier $h_1, h_2, h_3 \in \mathcal{C}^1([-1,1])$ se tiene que

$$det(\delta K_i(y, h_j)) = 0.$$

(ii) Existen $\lambda_1, \lambda_2, \lambda_3$ de forma que

$$\delta A(y,h) = \sum_{i=1}^{3} \lambda_i \delta K_i(y,h) \quad \forall h \in \mathcal{C}^1([-1,1])$$

Caso de cumplirse la alternativa (ii) tendríamos:

Problema 2.7.12 *(a) Escribir en nuestro problema concreto la expresión (ii).*

(b) Probar que

$$\int_{-1}^{1} [h(t) - \frac{\lambda_1 y'((t)}{\sqrt{1+(y'(t))^2}} h'(t)]dt = 0$$

siempre que $h \in \{\varphi \in \mathcal{C}^1([-1,1]) : \ \varphi(-1) = \varphi(1) = 0\}$

Problema 2.7.13 *Probar que*

$$-\frac{\lambda_1 y'((t)}{\sqrt{1+(y'(t))^2}} \in \mathcal{C}^1([-1,1])$$

y además que

$$[-\frac{\lambda_1 y'((t)}{\sqrt{1+(y'(t))^2}}]' = 1$$

(en particular, el problema anterior nos dice que $\lambda_1 \neq 0$).

El problema anterior nos permite decir que

$$-\frac{\lambda_1 y'((t)}{\sqrt{1+(y'(t))^2}} = t + C$$

para una determinada constante C. Si llamamos $\lambda = -\lambda_1$ y la expresión anterior la elevamos al cuadrado nos queda:

$$\frac{\lambda^2 (y'((t))^2}{1+(y'(t))^2} = (t+C)^2$$

Problema 2.7.14 *Obtener las soluciones de la ecuación diferencial anterior. Es decir, obtener las soluciones de las ecuaciones diferenciales:*

$$y'(t) = \frac{t+C}{\sqrt{\lambda^2 - (t+C)^2}}$$

y

$$y'(t) = -\frac{t+C}{\sqrt{\lambda^2 - (t+C)^2}}$$

Problema 2.7.15 *(a) Si exigimos que las soluciones de las ecuaciones anteriores estén en $\mathcal{M}$. Probar que $C = 0$.*

(b) Ver que los puntos de las gráficas de las soluciones de las dos ecuaciones anteriores están sobre una circunferencia de radio $|\lambda|$.

Teniendo en cuenta que las soluciones de las ecuaciones diferenciales debemos encontrarlas en $\mathcal{M}$ y que $l > 2$

Problema 2.7.16 *Probar que $D = -\sqrt{\lambda^2 - 1}$ donde $\lambda > 0$ es la única solución de la ecuación $\sin(\frac{l}{2\lambda}) = \frac{1}{\lambda}$.*

En definitiva,

Problema 2.7.17 *Probar que los posibles extremos locales de A sobre $\mathcal{M}$ son:*

$$y_1(t) = \sqrt{\lambda^2 - t^2} - \sqrt{\lambda^2 - 1}.$$
$$y_2(t) = -\sqrt{\lambda^2 - t^2} + \sqrt{\lambda^2 - 1}$$

Finalmente, nos queda por estudiar los puntos de $\mathcal{M}$ que verifican la condición (i) del teorema de Euler-Lagrange.

Problema 2.7.18 *Probar que si $y \in M$ verifica la condición (i) del teorema de Euler-Lagrange debe cumplir:*

(a)

$$\int_{-1}^{1} \frac{y'(t)h_1'(t)}{\sqrt{1+(y'(t))^2}} dt = 0.$$

(b) De donde se debe cumplir

$$\frac{y'(t)}{\sqrt{1+(y'(t))^2}} = k$$

Esto nos permite afirmar que los posibles extremos que obtenemos de esta condición ya están considerados en el caso anterior.

Capítulo 3

Principio de Contracción de Banach.

3.1. Introducción

El origen de la Teoría métrica del punto fijo, descansa en el método de aproximaciones sucesivas para probar la existencia y unicidad de soluciones de ecuaciones diferenciales. Este método está ligado a celebres matemáticos del siglo XIX como Picard. Sin embargo, es el matemático polaco Stefan Banach quien obtuvo un planteamiento general de las ideas que subyacían. El teorema del punto fijo, conocido generalmente como *principio de contracción de Banach*, aparece en forma explícita en la tesis de Banach de 1922 donde fue usado para establecer la existencia de una solución de una ecuación integral. Desde entonces, por su simplicidad y utilidad, ha llegado a ser uno de los resultados más utilizados en la resolución de problemas de existencia en muchas ramas del Análisis.

3.2. Teorema del punto fijo de Banach.

Sea $\mathcal{M}$ un espacio métrico con función distancia ρ. Una aplicación $T : M \to M$ se llama *lipschitziana* si existe una constante $k \geq 0$ tal que para todo $x, y \in M$,

$$\rho(Tx, Ty) \leq k\rho(x, y). \qquad (1)$$

La constante k más pequeña que verifica la desigualdad (1) se llama *constante de Lipschitz* de la aplicación T, y la denotaremos por $k(T)$.

Una aplicación $T : M \to M$ se llama *contracción* si $k(T) < 1$.

Teorema 3.2.1 *(Principio de contracción de Banach). Sea (M, ρ) un espacio métrico completo y sea $T : M \to M$ una contracción. Entonces T tiene un único punto fijo en M, y para cada $x_0 \in M$ la sucesión de iteradas $(T^n x_0)$ converge a este punto fijo.*

Demostración.- Sea $x_0 \in M$ y definimos la sucesión de iteradas (x_n) por $x_{n+1} = Tx_n$ (equivalentemente, $x_n = T^n x_0$, $n = 1, 2, 3, ...$). Observese que para cualesquiera índices $n, p \in \mathbb{N}$,

$$\rho(x_n, x_{n+p}) = \rho(T^n x_0, T^{n+p} x_0) \leq k(T)^n \rho(x_0, T^p x_0).$$

Para simplificar la notación, llamaremos k a la constante $k(T)$, entonces nos queda:

$$\begin{gathered}\rho(x_n, x_{n+p}) \leq k^n(\rho(x_0, Tx_0) + ... + \rho(T^{p-1} x_0, T^p x_0) \leq \\ k^n(1 + k + ... + k^{p-1})\rho(x_0, Tx_0) \leq k^n(\tfrac{1-k^p}{1-k})\rho(x_0, Tx_0).\end{gathered} \tag{3.1}$$

Esto prueba que (x_n) es una sucesión de Cauchy, y como M es un espacio métrico completo existirá $x \in M$ tal que $\text{lím}_{n\to\infty} x_n = x$. Para ver que x es el único punto fijo de T, observemos que,

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Tx_n = Tx.$$

y, además, si $x = Tx$ e $y = Ty$ se tiene

$$\rho(x, y) = \rho(Tx, Ty) \leq k\rho(x, y),$$

obteniendose que $\rho(x, y) = 0$.

Por otra parte, se observa que $\rho(x_n, x) \to 0$.

Observación 3.2.2 *De la desigualdad (3.1) se obtine, teniendo en cuenta que $k < 1$, que $\rho(x_n, x_{n+p}) \leq \frac{k^n}{1-k}\rho(x_0, T(x_0))$. Con lo cual, si x es el punto fijo de T se debe cumplir:*

$$\rho(x_n, x) = \lim_{p\to\infty} \rho(x_n, x_{n+p}) \leq \frac{k^n}{1-k}\rho(x_0, T(x_0)),$$

que es lo que se denomina cota de error a priori.

Como para cualesquiera enteros positivos n, m,

$$\rho(x_{n+1}, x_{n+m+1}) \leq \sum_{i=1}^{m} \rho(x_{n+i}, x_{n+1+i}) \leq \sum_{i=1}^{m} k^i \rho(x_n, x_{n+1}) \leq \\ \leq \sum_{i=1}^{\infty} k^i \rho(x_n, x_{n+1}) \leq \frac{k}{1-k} \rho(x_n, x_{n+1})$$

teniendo en cuenta la continuidad de la función distancia se obtiene:

$$\rho(x_{n+1}, x) = \lim_{m\to\infty} \rho(x_{n+1}, x_{n+m+1}) \leq \frac{k}{1-k} \rho(x_n, x_{n+1}),$$

que se denomina cota del error a posteriori.

3.3. Aplicaciones del principio de contracción.

3.3.1. El método de Picard-Lindelöf.

Sea $f : [t_0 - c, t_0 + c] \times \mathbb{R} \to \mathbb{R}$ una función continua. El problema de valores iniciales de Cauchy es el problema de encontrar una función $x : [t_0 - c, t_0 + c] \to \mathbb{R}$ continuamente derivable que satisfaga la siguiente ecuación diferencial:

$$\begin{cases} x'(t) = f(t, x(t)) & t \in [t_0 - c, t_0 + c] \\ x(t_0) = p_0 \end{cases}$$

El resultado clásico de Picard-Lindelöf establece que si f es lipschitziana con respecto a la segunda variable, i.e., si existe $L > 0$ de forma que si para todo $x, y \in \mathbb{R}$,

$$|f(t, x) - f(t, y)| \leq L|x - y|, \quad \forall t \in [t_0 - c, t_0 + c],$$

entonces la ecuación anterior tiene una única solución.

Prueba.

Consideremos el espacio de Banach $(C([t_0 - c, t_0 + c]), \|.\|_\infty)$ Sobre dicho espacio definimos la aplicación $F : C([t_0 - c, t_0 + c]) \to C([t_0 - c, t_0 + c])$ donde a cada función $x \in C([t_0 - c, t_0 + c])$ le hacemos corresponder la función $F(x)$ definida de la forma siguiente:

$$F(x)(t) = p_0 + \int_{t_0}^{t} f(s, x(s))ds$$

Es claro que si $x \in C([t_0 - c, t_0 + c]$, por el teorema fundamental del cálculo, se tiene que $F(x) \in C([t_0 - c, t_0 + c])$. Además, claramente una solución del problema de Cauchy será un punto fijo de dicha aplicación.

Observar que si $, y \in C([t_0 - c, t_0 + c])$, entonces

$$|F(x)(t) - F(y)(t)| = |\int_{t_0}^{t} f(s, x(s))ds - \int_{t_0}^{t} f(s, y(s))ds| \leq$$

$$|\int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))|ds| \leq |\int_{t_0}^{t} L|x(s) - y(s)|ds| \leq L|t - t_0| \|x - y\|_\infty.$$

Ahora, teniendo presente la definición de la norma $\|.\|_\infty$ se concluye:

$$\|F(x) - F(y)\|_\infty \leq L2c\|x - y\|_\infty.$$

Esta última expresión nos dice que la aplicación F es lipschitziana con constante de Lipschitz $k(F) \leq L2c$.

Consideremos ahora la aplicación $F^2 := F \circ F : C([t_0 - c, t_0 + c]) \to C([t_0 - c, t_0 + c])$ y haciendo el mismo razonamiento que antes, se tiene:

$$|F^2(x)(t) - F^2(y)(t)| \leq |\int_{t_0}^{t} |f(s, F(x)(s)) - f(s, F(y)(s))|ds| \leq$$

$$|\int_{t_0}^{t} L|F(x)(s) - F(y)(s)|ds| \leq |\int_{t_0}^{t} L| \int_{t_0}^{s} L|x(u) - y(u)|du|ds| \leq$$

$$L^2\|x - y\|_\infty| \int_{t_0}^{t} |s - t_0|ds = L^2 \frac{|t - t_0|^2}{2} \|x - y\|_\infty$$

Por lo tanto,

$$\|F^2(x) - F^2(y)\|_\infty \leq \frac{(2cL)^2}{2} \|x - y\|_\infty.$$

Repitiendo este proceso inductivamente se obtiene que

$$\|F^n(x) - F^n(y)\|_\infty \leq \frac{(2cL)^n}{n!} \|x - y\|_\infty.$$

Por otra parte, como $\lim_{n\to\infty} \frac{(2cL)^n}{n!} = 0$. Entonces existirá una iterada de F la cual será contractiva, digamos F^{n_0}. Aplicando el principio de Banach a dicha iterada se tendrá: existe una única $x_0 \in C([t_0 - c, t_0 + c])$ de forma que $F^{n_0}x_0 = x_0$. Veamos ahora que x_0 también es un punto fijo de F. En caso contrario, tenemos:

$$\|F(x_0) - x_0\|_\infty = \|F^{n_0+1}(x_0) - F^{n_0}(x_0)\|_\infty \leq k(F^{n_0})\|F(x_0) - x_0\|_\infty.$$

Lo cual es una contradicción con el hecho de que F^{n_0} es contractiva y por lo tanto $k(F^{n_0}) < 1$.

La unicidad es consecuencia del teorema de Banach:

En efecto, supongamos que y_1 es otra solución del problema. Entonces se debe cumplir que $F(y_1) = y_1$. Por lo tanto, y_1 es un punto fijo de F^{n_0}, que es una aplicación contractiva. El teorema de Banach nos dice F^{n_0} sólamente tiene un punto fijo, con lo cual $y_1 = x_0$.

Ejemplo 3.3.1 *Veamos que el siguiente problema de valores iniciales:*

$$\begin{cases} y'(x) = x + y(x) & x \in \mathbb{R} \\ y(0) = 1 \end{cases}$$

tiene solución única y mediante el teorema anterior podemos calcularla.

Existencia y unicidad.

Dado $n \in \mathbb{N}$, consideramos el intervalo compacto $[-n, n]$. Sobre dicho compacto definimos la función:

$$f : [-n, n] \times \mathbb{R} \to \mathbb{R} \;:\; f(x, y) = x + y$$

Claramente f está bajo las hipótesis del teorema de Picard-Lindelöf, ya que $|f(x, y) - f(x, z)| \leq |y - z|$. Luego el problema anterior en el intervalo $[-n, n]$ es:

$$\begin{cases} y'(x) = f(x, y(x)) & x \in [-n, n] \\ y(0) = 1 \end{cases}$$

Por el teorema Picard-Lindelöf existe una única función $y_n : [-n, n] \to \mathbb{R}$ que es solución de dicho problema. Como ésto lo podemos hacer para cada número entero positivo y además la unicidad de la solución nos permite decir que si y_n e y_m son las soluciones del problema en los intervalos respectivos $[-n, n]$ y $[-m, m]$, entonces $y_n(x) = y_m(x)$ siempre que $x \in [-n, n] \cap [-m, m]$.

La función $y : \mathbb{R} \to \mathbb{R}$ definida de la siguiente forma:

Dado $x \in \mathbb{R}$, existirá $n \in \mathbb{N}$ tal que $x \in [-n, n]$, entonces $y(x) := y_n(x)$. Es la única solución del problema de valores iniciales.

Cálculo de la solución

Como es habitual en el tratamiento de las ecuaciones diferenciales, consideremos su ecuación integral equivalente:

$$y(x) = 1 + \int_0^x [t + y(t)]dt,$$

En este caso, como estamos bajo las hipótesis del teorema de Picard-Lindelöf, las iteradas de una función, que vendrán dadas por la fórmula

$$y_n(x) = 1 + \int_0^x [t + y_{n-1}(t)]dt,$$

convergen a la única solución.

Si tomamos $y_0(t) = 1$, el método anterior nos da

$$y_1(x) = 1 + \int_0^x (t+1)dt = 1 + x + \frac{x^2}{2!},$$

$$y_2(x) = 1 + \int_0^x \left(1 + 2t + \frac{t^2}{2!}\right) dt = 1 + x + x^2 + \frac{x^3}{3!},$$

$$y_3(x) = 1 + \int_0^x \left(1 + 2t + t^2 + \frac{t^3}{3!}\right) dt = 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{4!},$$

$$y_4(x) = 1 + \int_0^x \left(1 + 2t + t^2 + \frac{t^3}{3} + \frac{t^4}{4!}\right) dt = 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{4(3)} + \frac{x^5}{5!}.$$

Razonando por inducción se llega a que

$$y_n(x) = 1 + x + 2\left(\frac{x^2}{2!} + \frac{x^3}{3!} + ... + \frac{x^n}{n!}\right) + \frac{x^{n+1}}{(n+1)!}.$$

Luego tomando límites, nos queda:

$$y(x) = \lim_{n\to\infty} y_n(x) = 1 + x + 2(e^x - x - 1) + 0 = 2e^x - x - 1.$$

3.4. Aplicaciones no expansivas.

En esta sección intentaremos generalizar el Teorema de Banach para una clase de aplicaciones más general que las contractivas.

Definición 3.4.1 *Una aplicación $T : M \to M$ se dice que es contráctil si*

$$d(Tx, Ty) < d(x, y)$$

para cada $x, y \in M$ con $x \neq y$.

Ahora, veamos que el Principio de contracción de Banach no se verifica para esta clase de aplicaciones. En efecto,

Ejemplo 3.4.2 *Consideremos la aplicación*

$$\begin{array}{rccl} T: & [1,+\infty[& \longrightarrow & [1,+\infty[\\ & x & \mapsto & T(x) = \frac{x^2+1}{x} \end{array}$$

Obviamente $([1,+\infty[,|.|)$ *es un espacio métrico completo y* T *es una aplicación contráctil sin puntos fijos en* $[1,+\infty[$.

La clase de aplicaciones contráctiles está incluida el la clase de las aplicaciones denominadas no expansivas, es decir:

Definición 3.4.3 *Una aplicación* $T : M \to M$ *se dice que es no expansiva si*

$$d(Tx,Ty) \leq d(x,y)$$

para cada $x, y \in M$.

El mayor interés de las aplicaciones no expansivas se encuentra cuando se trabaja en espacios de Banach. Por lo tanto, concentraremos nuestra atención en las aplicaciones no expansivas definidas sobre espacios de Banach. En este sentido se tiene:

Definición 3.4.4 *Sea* $(X, \|.\|)$ *un espacio de Banach y sea* C *un subconjunto no vacío de* X. *Una aplicación* $T : C \to C$ *se dice que es no expansiva si*

$$\|Tx - Ty\| \leq \|x - y\|$$

para cada $x, y \in C$.

Es fácil observar que las aplicaciones no expansivas no siempre tienen puntos fijos:

Ejemplo 3.4.5 *Sea* X *un espacio de Banach, dado* $a \in X \setminus \{0\}$ *definimos la aplicación* $T : X \to X$ *por* $T(x) = x + a$. *Claramente* T *es no expansiva y no tiene puntos fijos en* X.

Por otra parte, cuando las aplicaciones no expansivas tienen punto fijo éste no necesariamente es único (pensar en la aplicación identidad sobre un espacio de Banach). También hay otra diferencia fundamental respecto al principio de contracción ya que incluso cuando una aplicación no expansiva tiene un punto fijo la sucesión de iteradas de un punto en general no converge.

Ejemplo 3.4.6 *En el espacio de Banach* $(\mathbb{R}^2, \|.\|_2)$. *Consideremos su bola unidad cerrada* B *y sea* $T : B \to B$ *la siguiente aplicación:* $T(x, y) = (y, x)$ *para cada* $(x, y) \in B$. *Es claro que* T *es no expansiva y además los puntos de la bola que de la forma* (x, x) *son puntos fijo de* T.

Sin embargo, la sucesión de iteradas $\{T^n(1, 0)\}$ *no converge.*

Si pretendemos obtener resultados de punto fijo para este tipo de aplicaciones es de interés la siguiente observación:

Proposición 3.4.7 *Sea* X *un espacio de Banach y sea* K *un subconjunto no vacío, cerrado acotado y convexo de* X. *Si* $T : K \to K$ *es una aplicación no expansiva, entonces* $\inf\{\|x - Tx\| : \ x \in K\} = 0$.

Prueba.

Fijemos $z \in K$ y $\epsilon \in]0, 1[$, ahora consideremos la aplicación $T_\epsilon : K \to K$ definida por

$$T_\epsilon x = \epsilon z + (1 - \epsilon)Tx$$

Claramente T_ϵ es contractiva, ya que

$$\|T_\epsilon x - T_\epsilon y\| \leq (1 - \epsilon)\|Tx - Ty\| \leq (1 - \epsilon)\|x - y\|$$

Entonces, por el principio de contracción existirá $x_\epsilon \in K$ de forma que $x_\epsilon = T_\epsilon x_\epsilon$. De donde se desprende que

$$\|x_\epsilon - Tx_\epsilon\| \leq \epsilon \mathrm{diam}(K).$$

El resultado se obtiene haciendo tender ϵ hacia 0.

Como una consecuencia inmediata de la proposición anterior obtenemos un primer resultado de punto fijo para estas aplicaciones:

Corolario 3.4.8 *Sea* X *un espacio de Banach y sea* K *un subconjunto no vacío, compacto y convexo de* X. *Si* $T : K \to K$ *es una aplicación no expansiva, entonces* T *tiene al menos un punto fijo en* K.

En virtud del ejemplo 3.4.6 está claro que no se puede obtener un resultado de convergencia parecido al teorema 3.5.6. No obstante, hay métodos de convergencia a un punto fijo que funcionan en situaciones bastante generales. El resultado que veremos a continuación se llama *iteración de Krasnoselskij*.

Teorema 3.4.9 *Sea K un subconjunto compacto convexo de un espacio de Hilbert H. Si $T : K \to K$ es una aplicación no expansiva. Entonces para cada $x_0 \in K$, y para cualquier número fijo $\lambda \in]0,1[$ la sucesión de iteradas de Krasnoselskij (x_n) dada por*

$$x_{n+1} = (1-\lambda)x_n + \lambda T x_n \tag{3.2}$$

converge a un punto fijo de T.

Prueba.

Como hemos visto antes, es claro que T tiene algún punto fijo en K. Sea pues p un punto fijo de T en K.

Por otra parte, dado $x_0 \in K$; como K es convexo está claro que la sucesión (x_n) definida como en (3.2) está formada por elementos de K.

Lo primero que veremos ahora es que la sucesión $(x_n - Tx_n)$ converge a cero. En efecto,

$$x_{n+1} - p = (1-\lambda)(x_n - p) + (1-\lambda)(Tx_n - p).$$

Por otra parte, para cada constante a se tiene:

$$a(x_n - Tx_n) = a(x_n - p) + a(p - Tx_n).$$

Entonces, teniendo en cuenta que la norma es la inducida por el producto escalar nos queda:

$$\|x_{n+1} - p\|^2 = (1-\lambda)^2\|x_n - p\|^2 + \lambda^2\|Tx_n - p\|^2 + 2\lambda(1-\lambda)\langle Tx_n - p, x_n - p\rangle.$$

Además,

$$a^2\|x_n - Tx_n\|^2 = a^2\|x_n - p\|^2 + a^2\|Tx_n - p\|^2 - 2a^2\langle Tx_n - p, x_n - p\rangle.$$

Ahora sumando las dos expresiones anteriores y teniendo presente que T es no expansiva y además que $Tp = p$, nos queda:

$$\|x_{n+1} - p\|^2 + a^2\|x_n - Tx_n\|^2 \leq$$
$$[2a^2 + (1-\lambda)^2 + \lambda^2]\|x_n - p\|^2 + 2[\lambda(1-\lambda) - a^2]\langle Tx_n - p, x_n - p\rangle.$$

Si elegimos ahora a tal que $a^2 \leq \lambda(1-\lambda)$, entonces aplicando la desigualdad de Cauchy-Schwartz, en la última desigualdad se obtiene:

$$\|x_{n+1} - p\|^2 + a^2\|x_n - Tx_n\|^2 \leq$$

$(2a^2+\lambda^2+(1-\lambda)^2+2\lambda(1-\lambda)-2a^2)\|x_n-p\|^2 = [\lambda+(1-\lambda)]^2\|x_n-p\|^2 = \|x_n-p\|^2.$

Tomando $a = \lambda(1-\lambda)$ y despejando en la desigualdad anterior se tiene:

$$a^2\|x_n - Tx_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2.$$

Como lo anterior es cierto para todo número natural, podemos escribir que dado $N \in \mathbb{N}$ se cumple:

$$\lambda(1-\lambda)\sum_{n=0}^{N}\|x_n - Tx_n\|^2 \leq \sum_{n=0}^{N}\left[\|x_n - p\|^2 - \|x_{n+1} - p\|^2\right] =$$
$$\|x_0 - p\|^2 - \|x_{N+1} - p\|^2 \leq \|x_0 - p\|^2$$

Es decir la serie de números positivos

$$\sum_{n=0}^{\infty}\|x_n - Tx_n\|^2$$

es convergente, entonces por el criterio del resto se debe verificar:

$$\lim_{n\to\infty}\|x_n - Tx_n\| = 0.$$

Veamos ahora que (x_n) converge a un punto fijo q de T. En efecto,

Como (x_n) está acotada y K es un compacto, entonces existirá (x_{n_k}) una subsucesión convergente a un determinado $q \in K$. Por la continuidad de T se debe verificar que q es un punto fijo de T ya que:

Por una parte

$$x_{n_k} \to q,$$

por otra parte, la continuidad de T nos dice que

$$Tx_{n_k} \to Tq$$

y finalmente, como

$$x_n - Tx_n \to 0$$

se debe cumplir que $q = Tq$.

Únicamente nos queda por ver que $x_n \to q$. En efecto, la sucesión $(\|x_n - q\|)$ es decreciente, ya que:

$$\|x_{n+1} - q\| = \|(1-\lambda)(x_n - q) + \lambda(Tx_n - Tq)\| \leq$$
$$(1-\lambda)\|x_n - q\| + \lambda\|Tx_n - Tq\| \leq \|x_n - q\|.$$

Como la sucesión $(\|x_n - q\|)$ es decreciente, entonces será convergente y como tiene una subsucesión que tiende a cero se obtiene el resultado.

3.5. Problemas

Problema 3.5.1 *Dada la ecuación en* $\mathbb{R}$,

$$t^5 + t + 1 = 0,$$

demostrar que tiene una única raiz real y reducirla a un modo que pueda resolverse mediante un procedimiento iterativo.

Problema 3.5.2 *Demostrar que existe una única función continua* $f : [-1,1] \to \mathbb{R}$ *tal que*

$$f(t) = t + \frac{1}{2}\sin(f(t)), \ \forall t \in [-1,1].$$

Problema 3.5.3 *Sean* $P, Q : \mathbb{R} \to \mathbb{R}$ *funciones continuas. Probar la existencia de una única función* $y : \mathbb{R} \to \mathbb{R}$ *continuamente diferenciable que es solución de la ecuación*

$$\begin{cases} y'(t) + P(t)y(t) = Q(t) & t \in \mathbb{R} \\ y(t_0) = p_0 \end{cases}$$

Problema 3.5.4 *¿Podemos afirmar que el problema de valores iniciales*

$$\begin{cases} y'(t) = y^2(t) & t \in]-1,1[\\ y(0) = 1 \end{cases}$$

tiene una única solución?.

Problema 3.5.5 *Estudiar si el problema de valores iniciales:*

$$\begin{cases} y'(t) = 2t(1 + y(t)) & t \in \mathbb{R} \\ y(0) = 0 \end{cases}$$

tiene una única solución. En caso afirmativo, obtener la solución exacta por el método de Picard.

Problema 3.5.6 *Probar que si* (M,d) *es un espacio métrico compacto y si* $T : M \to M$ *es una aplicación contráctil. Entonces* T *tiene un único punto fijo* x_0, *y además, para cada* $x \in M$, $lim_{n\to\infty} T^n x = x_0$.

Problema 3.5.7 *En el espacio de Banach* $X := \mathcal{C}([0,1])$ *con la norma habitual, probar que el conjunto*

$$K := \{f \in X : \ 0 = f(0) \leq f(t) \leq f(1) = 1, \ \forall t \in [0,1]\},$$

es cerrado y acotado.

Se define $T : X \to X$ *por* $(Tf)(t) := tf(t)$*. Probar que*

(a) $T(K) \subseteq K$.

(b) $\|Tf - Tg\| < \|f - g\|$.

(c) T *no tiene puntos fijos en* K.

Probar que la aplicación $J : K \to \mathbb{R}$ *definida por* $J(f) = \|f - Tf\|$*, verifica que* $\inf\{J(f) : f \in K\} = 0$.

3.5.1. Ecuaciones integrales.

Una ecuación integral es un problema que consiste en obtener una función incógnita, cuando ésta cumple cierta identidad mediante operadores integrales.

Por ejemplo,

$$x(t) = \mu \int_a^b F(t, s, x(s))ds + y(t)$$

donde F e y son funciones datos y x es la función incognita, se llama Ecuación integral de Urysohn.

Veamos cómo el teorema de Banach puede utilizarse para probar la existencia de soluciones de algunas ecuaciones integrales tipo Urysohn.

Consideremos el intervalo de la recta real $[a, b]$ y supongamos que la función $F : [a, b] \times [a, b] \times \mathbb{R} \to \mathbb{R}$ es tal que para todo $t \in [a, b]$ y para toda función continua $x \in C([a, b])$, la aplicación

$$t \to \int_a^b F(t, s, x(s))ds$$

es continua. Además verifica la siguiente condición

$$|F(t, s, z_1) - F(t, s, z_2)| \leq G(t, s)|z_1 - z_2|.$$

Supongamos, además que

$$|\mu|(b - a) \max\{G(t, s) : (t, s) \in [a, b] \times [a, b]\} < 1.$$

Problema 3.5.8 *Probar que la ecuación de Urysohn*

$$x(t) = \mu \int_a^b F(t, s, x(s))ds + y(t),$$

tiene solución siempre que F *esté en las condiciones anteriores.*

Problema 3.5.9 *Probar que existe una función continua* $x : [0, \frac{1}{2}] \to \mathbb{R}$ *que satisface la ecuación integral*

$$x(s) - \int_0^{\frac{1}{2}} \cos(st)\cos(x(t))dt = e^s.$$

Capítulo 4

Los Teoremas de Brouwer y Schauder

4.1. El teorema de Brouwer.

Definición 4.1.1 *Un espacio topológico (X, τ) se dice que tiene la propiedad del punto fijo topológica (fppt para abreviar) si toda función continua $T : X \to X$ tiene algún punto fijo.*

No es sorprendente que la propiedad anterior sea un invariante topológico. En efecto,

Lema 4.1.2 *Si X, Y son dos espacios topológicos homeomorfos y si X tiene la fppt, entonces Y también la tiene.*

Prueba. Sea $h : X \to Y$ un homeomorfismo con $h(X) = Y$, y supongamos que $f : Y \to Y$ es continua. Construimos la función

$$g = h^{-1} \circ f \circ h : X \to X.$$

Como f es continua, h es un homeomorfismo, entonces g es continua; luego existirá $x \in X$ tal que $g(x) = x$, lo cual implica que $y = h(x)$ es un punto fijo de f.

Otra observación muy útil acerca de la propiedad del punto fijo topológica es la siguiente:

Lema 4.1.3 *Si X es un espacio topológico que tiene la fppt e Y es un retracto de X, entonces Y tiene la fppt.*

($Y \subseteq X$ es un retracto de X si existe una función continua $r : X \to Y$ de forma que $r(y) = y$ $\forall y \in Y$).

Prueba. Sea $r : X \to Y$ una retracción y sea $f : Y \to Y$ una aplicación continua. Entonces podemos extender de forma continua la función f a todo X definiendo:

$$g := f \circ r : X \to X.$$

Ahora, sólamente hay que tener en cuenta que los puntos fijos de g deben ser puntos fijos de f.

Probablemente la observación más elemental respecto a la propiedad del punto fijo topológica sea el teorema de Bolzano es decir, el hecho que cualquier función continua $f : [a, b] \to [a, b]$ tiene un punto fijo (i.e., $[a, b]$ tiene la fppt). Como consecuencia del Lema 4.1.2 todos los intervalos compactos y arcos simples tienen la fppt. Sin embargo, probar que conjuntos tan simples como los triángulos de $\mathbb{R}^2$ también tienen la fppt es un argumento mucho más sofisticado. Este hecho es un caso especial del siguiente teorema probado por L.E.J. Brouwer en 1912.

Teorema 4.1.4 (Brouwer) *La bola cerrada unidad de* $(\mathbb{R}^n, \|.\|_2)$ *tiene la propiedad del punto fijo topológica.*

Existen muchas pruebas del teorema de Brouwer. Algunas son muy cortas, aunque usan resultados de Topología Algebraica, o de Integración en Variedades. Otras en cambio, son muy laboriosas, pero elementales. Aquí incluiremos una prueba, elemental, de dicho teorema que fue publicada en 1978 por J. Milnor.

En todo la demostración denotaremos por B^n y S^{n-1} la bola y la esfera unidad de $(\mathbb{R}^n, \|.\|_2)$.

Sea $A \subseteq \mathbb{R}^n$. Una aplicación continua $f : A \to \mathbb{R}^n$ se llamará de clase $\mathcal{C}^1$ si f admite una extensión continua en un entorno abierto de A la cual es continuamente diferenciable.

Para simplificar la intuición, nos referiremos a una aplicación $f : A \to \mathbb{R}^n$ como: *un campo vectorial*. Al hacer esto, estamos pensando en f como una aplicación que a cada vector $x \in A$ le hace corresponder el vector que tiene como punto inicial el x y punto final $x + f(x)$.

Una aplicación (campo vectorial) $f : A \to \mathbb{R}^n$ se dice que no se anula si $f(x) \neq 0$ para todo $x \in A$. Se llama normalizada si $\|f(x)\|_2 = 1$ para todo $x \in A$, (i.e., $f : A \to S^{n-1}$). El campo $f : S^{n-1} \to \mathbb{R}^n$ se llama tangente a S^{n-1} si $\langle x, f(x)\rangle = 0$ para todo $x \in S^{n-1}$.

Lema 4.1.5 *Si A es un subconjunto compacto de $\mathbb{R}^n$ y $f : A \to \mathbb{R}^n$ es de clase $\mathcal{C}^1$ en A. Entonces existe $L > 0$ tal que*

$$\|f(x) - f(y)\|_2 \leq L\|x - y\|_2 \qquad \forall x, y \in A$$

Prueba. Para cada $x \in A$, llamamos $\mathcal{U}_x$ a una bola abierta que contiene a x y sobre cuya frontera f es continuamente diferenciable.

Como A es compacto, A puede recubrirse con un número finito, $\mathcal{U}_1, ..., \mathcal{U}_p$, de tales bolas.

Llamemos

$$c_{i,j}^k = \text{máx}\{|\frac{\partial f_i(x)}{\partial x_j}| \ : \ x \in \overline{\mathcal{U}_k}\} \quad i, j = 1, ..., n, \quad k = 1, ..., p.$$

Como el conjunto $\mathcal{U}_k$ es convexo, el teorema del valor medio y el hecho de que $\|.\|_2 \leq \|.\|_1$ nos permite obtener para cada $x, y \in \mathcal{U}_k$:

$$\|f(x)-f(y)\|_2 \leq \sum_{i=1}^{n} |f_i(x)-f_i(y)| \leq \sum_{i=1}^{n}(\sum_{j=1}^{n} c_{i,j}^k\|x-y\|_2) = \sum_{i,j=1}^{n} c_{i,j}^k\|x-y\|_2.$$

Por otra parte, por compacidad se puede comprobar que existe $\varepsilon > 0$ de forma que si $x, y \in A$ y no pertenecen al mismo $\mathcal{U}_k$, entonces $\|x - y\|_2 \geq \varepsilon$. En efecto, dado $x \in \mathcal{U}_k \cap A$ está claro que $\text{dist}(x, A \cap (\mathbb{R}^n \setminus \mathcal{U}_k) = \epsilon_x > 0$.

Por lo tanto, $A \subseteq \bigcup_{x \in A} B(x, \frac{\epsilon_x}{2})$. Como A es compacto, se tendrá:

$$A \subseteq \bigcup_{i=1}^{m} B(x_i, \frac{\epsilon_i}{2}).$$

Llamemos $2\varepsilon := \text{mín}\{\epsilon_1, ..., \epsilon_m\} > 0$. Supongamos que $x, y \in A$ y que $\|x-y\|_2 < \varepsilon$. Entonces, como $x \in B(x_i, \frac{\epsilon_i}{2})$ tenemos que $\|x_i - y\| < \epsilon_i$, luego por el razonamiento anterior si $x_i \in \mathcal{U}_k$ se deduce que $x, y \in \mathcal{U}_k$.

Como A es compacto y f es continua se tiene que $f(A)$ es compacto, con lo cual se cumplirá:

$$\|f(x) - f(y)\|_2 \leq \frac{1}{\varepsilon}\text{diam}(f(A))\|x - y\|_2.$$

Si llamamos $L := \text{máx}\{L_1, ..., L_p, \frac{1}{\varepsilon}\text{diam}(f(A))\}$, donde $L_k = \sum_{i,j=1}^n c_{i,j}^k$, $k = 1, ...p$ se obtiene el resultado.

Sea $F : A \to \mathbb{R}^n$ un campo vectorial y para cada $t \in \mathbb{R}$ definimos $f_t : A \to \mathbb{R}^n$ por: $f_t(x) = x + tF(x)$, $\forall x \in A$.

Lema 4.1.6 *Si A es un dominio cerrado y acotado (i.e., la clausura de un subconjunto abierto conexo) y si $F : A \to \mathbb{R}^n$ es de clase $\mathcal{C}^1$ en A, entonces existe $\varepsilon > 0$ de forma que*

$$\begin{array}{rrcl} \varphi : &]-\varepsilon, \varepsilon[& \longrightarrow & \mathbb{R} \\ & t & \mapsto & \varphi(t) = vol(f_t(A)) \end{array}$$

es un polinomio en t de grado a lo sumo n.

Prueba. Sea L la constante de lipschitz de F que podemos encontrar gracias al lema 4.1.5. Ahora consideremos ε cualquier número del intervalo abierto $]0, \frac{1}{L}[$. Esto lo hacemos así puesto que de este modo nos aseguramos que f_t es una aplicación inyectiva sobre A siempre que $|t| < \varepsilon$.

En efecto, si $f_t(x) = f_t(y)$, se debe cumplir que $x - y = t(F(y) - F(x))$, por lo tanto

$$\|x - y\|_2 = |t|\|F(x) - F(y)\|_2 \leq |t|L\|x - y\|_2$$

lo cual sólamente es cierto si $x = y$.

Por otra parte si consideramos el determinante

$$det(I + t(\frac{\partial F_i(x)}{\partial x_j})) = 1 + ta_1(x) + t^2a_2(x) + ... + t^n a_n(x)$$

donde cada a_i $i = 1, 2, ..., n$ es una función continua en x, está claro que podemos encontrar un ε menor que el anterior de forma que f_t siga siendo inyectiva en A y además

$$det(I + t(\frac{\partial F_i(x)}{\partial x_j})) > 0$$

para todo $x \in A$ y $|t| < \varepsilon$.

Para un tal t, el teorema de la función inversa nos dice que cada f_t es una aplicación inyectiva de clase $\mathcal{C}^1$ con inversa de clase $\mathcal{C}^1$. Es decir es un cambio de variable o homeomorfismo. De este modo, podemos calcular el $vol(f_t(A))$ mediante la fórmula del cambio de variable:

$$vol(f_t(A)) = \int_{f_t(A))} 1dx = \int_A det(I + t(\frac{\partial F_i(x)}{\partial x_j}))dx = \alpha_0 + t\alpha_1 + ... + t^n\alpha_n.$$

donde $\alpha_i = \int_A a_i(x)dx$ y $\alpha_0 = vol(A)$.

Lema 4.1.7 *Supongamos que $F : S^{n-1} \to \mathbb{R}^n$ es un campo vectorial normalizado, de clase $\mathcal{C}^1$ y tangente a S^{n-1}. Entonces existe $t > 0$ suficientemente pequeño de forma que $f_t(S^{n-1}) = \sqrt{1+t^2}S^{n-1}$.*

Prueba. Primero, veamos que $f_t(S^{n-1}) \subseteq \sqrt{1+t^2}S^{n-1}$. En efecto, dado $y \in S^{n-1}$ se tiene que

$$\|f_t(y)\|_2^2 = \langle y + tF(y), y + tF(y)\rangle = \|y\|_2^2 + t^2\|F(y)\|_2^2 + 2t\langle y, F(y)\rangle.$$

Como F es tangente a S^{n-1} se tiene que $\langle y, F(y)\rangle = 0$. Además, como F está normalizado se cumple que $\|F(y)\|_2 = 1$. Por lo tanto,

$$\|f_t(y)\|_2^2 = 1 + t^2,$$

es decir $f_t(y) \in \sqrt{1+t^2}S^{n-1}$.

Segundo, extendemos el campo F de la forma siguiente:

$$\begin{array}{rccl} \widetilde{F}: & \mathbb{R}^n \setminus \{0\} & \longrightarrow & \mathbb{R}^n \\ & x & \mapsto & \widetilde{F}(x) = \|x\|_2 F(\frac{x}{\|x\|_2}). \end{array}$$

Consideremos el conjunto $A := \{x \in \mathbb{R}^n \ : \ \frac{1}{2} \leq \|x\|_2 \leq \frac{3}{2}\}$, y sea $|t| < \text{mín}\{\frac{1}{3}, \frac{1}{L}\}$, donde L es la constante de lipschitz de $\widetilde{F}$ sobre A que se obtiene por el lema 4.1.5.

Lo que se trata de probar ahora es que para un t en las condiciones anteriores se cumple que dado $y \in \sqrt{1+t^2}S^{n-1}$ se tiene que $y \in f_t(S^{n-1})$.

Para verlo, fijemos un $z \in S^{n-1}$ y definimos:

$$\begin{array}{rccl} G: & A & \longrightarrow & \mathbb{R}^n \\ & x & \mapsto & G(x) = z - t\widetilde{F}(x). \end{array}$$

Como $|t| < \frac{1}{3}$, entonces $G(x) \in A$, con lo cual $G : A \to A$. Además, como $|t| < \frac{1}{L}$, G es una contracción, entonces aplicando el principio de contracción de Banach, existirá un único $x_0 \in A$ tal que $G(x_0) = x_0$, es decir, $x_0 = z - t\widetilde{F}(x_0)$.

De aquí se obtiene:

$$x_0 + t\|x_0\|_2 F(\frac{x_0}{\|x_0\|_2}) = z.$$

Por lo tanto, como F está normalizado, se tiene

$$t\|x_0\|_2 = \|z - x_0\|_2,$$

ahora teniendo en cuenta que estamos trabajando con la norma procedente del producto escalar:

$$t^2\|x_0\|_2^2 = \|z - x_0\|_2^2 = \|z\|_2^2 + \|x_0\|_2^2 - 2\langle z, x_0\rangle.$$

Con lo cual,

$$t^2\|x_0\|_2^2 = 1-\|x_0\|_2^2-2\langle t\|x_0\|_2 F(\frac{x_0}{\|x_0\|_2}), x_0\rangle = 1-\|x_0\|_2^2-2t\|x_0\|_2\langle F(\frac{x_0}{\|x_0\|_2}), x_0\rangle.$$

Finalmente, como F es tangente a S^{n-1} se cumple que $\langle F(\frac{x_0}{\|x_0\|_2}), x_0\rangle = 0$. Lo que nos permite concluir:

$$\|x_0\| = \frac{1}{\sqrt{1+t^2}}.$$

De donde se tiene que si $y = \sqrt{1+t^2}z$ y llamamos $v = \sqrt{1+t^2}x_0 \in S^{n-1}$. Se tiene que, $v + tF(v) = \sqrt{1+t^2}z$, es decir, existe $v \in S^{n-1}$ tal que $f_t(v) = y$.

El lema previo usa campos vectoriales tangentes a S^{n-1}. Si n es par, es decir si $n = 2k$ para algún $k \in \mathbb{N}$, entonces es bastante fácil encontrar ejemplos concretos de éste tipo de campos vectoriales. Quizá uno de los ejemplos más simples sea:

$$F(x_1, x_2, ..., x_{2k-1}, x_{2k}) = (x_2, -x_1, x_4, -x_3, ..., x_{2k}, -x_{2k-1}).$$

Sin embargo, cuando n es impar la situación es totalmente diferente.

Teorema 4.1.8 *No existen campos vectoriales normalizados de clase $\mathcal{C}^1$ tangentes a S^{2k}.*

Prueba. Lo haremos por reducción al absurdo. Supongamos que F es un campo vectorial de esta forma.

Consideremos $0 < a < 1 < b$ y extendemos F, como en el lema anterior, a $\widetilde{F}$ sobre el dominio $A = \{x \in \mathbb{R}^n \ : \ a \leq \|x\|_2 \leq b\}$.

Observar que $\widetilde{F}$ es tangente a cualquier esfera concéntrica con S^{2k} la cual está contenida en A.

Sea $f_t(x) = x + t\widetilde{F}(x)$. Por el lema 4.1.7, $f_t(A) = \sqrt{1+t^2}A$ para $t > 0$ lo suficientemente pequeño (sea $x \in A$, entonces $\frac{x}{\|x\|_2} \in S^{2k}$, con lo cual por el lema anterior $\frac{x}{\|x\|_2} + tF(\frac{x}{\|x\|_2}) \in \sqrt{1+t^2}S^{2k}$ lo que significa que

$x + t\widetilde{F}(x) \in \sqrt{1+t^2}\|x\|S^{2k} \subseteq \sqrt{1+t^2}A$. Acabamos de ver que se cumple que $f_t(A) \subseteq \sqrt{1+t^2}A$.

Para obtener la desigualdad contraria razonamos de la forma siguiente: sea $y = \sqrt{1+t^2}z$, donde $z \in A$. Entonces $z = \|z\|\frac{z}{\|z\|}$. Por lo tanto, $y = \sqrt{1+t^2}\|z\|\frac{z}{\|z\|} = \|z\|\sqrt{1+t^2}\frac{z}{\|z\|}$, ahora por el lema anterior existirá $v \in S^{n-1}$ tal que $y = \|z\|(v + tF(v))$. Si llamamos $w = \|z\|v \in A$ obtenemos que

$$y = w + t\widetilde{F}(w) = f_t(w),$$

i.e., $\sqrt{1+t^2}A \subseteq f_t(A)$.

Ahora teniendo en cuenta las propiedades de la integral de Lebesgue se tendrá

$$vol(f_t(A)) = (\sqrt{1+t^2})^{2k+1} vol(A).$$

Sin embargo, $(\sqrt{1+t^2})^{2k+1}$ no es un polinomio en t en un entorno de cero, lo cual está en contradicción con el lema 4.1.6.

Teorema 4.1.9 (Teorema de la bola de Hairy) *No existen campos vectoriales continuos y tangentes a S^{2k} que no se anulen.*

Prueba. Supongamos que F es un campo vectorial en las condiciones de las hipótesis. llamemos $m = \text{mín}\{\|F(x)\|_2 \ : \ x \in S^{2k}\}$, obviamente $m > 0$. Por el teorema de aproximación de Weierstrass aplicado a cada componente de F se cumplirá:

Existe $P : S^{2k} \to \mathbb{R}^{2k+1}$ tal que $\|P(x) - F(x)\|_2 < \frac{m}{2}$ para todo $x \in S^{2k}$. Donde cada componente de P es un polinomio. De este modo, P es de clase $\mathcal{C}^\infty$ y además P no se anula sobre S^{2k}. En efecto,

$$\|P(x)\|_2 \geq \|F(x)\|_2 - \|P(x) - F(x)\|_2 > \frac{m}{2}.$$

Ahora modificamos P de la forma siguiente:

$$\widetilde{P}(x) = P(x) - \langle P(x), x\rangle x.$$

Claramente $\widetilde{P}$ es de clase $\mathcal{C}^\infty$ y tangente a S^{2k}. Además,

$$\|\widetilde{P}(x)\|_2 \geq \|P(x)\|_2 - \|\widetilde{P}(x) - P(x)\|_2 > \frac{m}{2} - |\langle P(x), x\rangle| =$$

$$= \frac{m}{2} - |\langle P(x) - F(x), x\rangle| \geq \frac{m}{2} - \|P(x) - F(x)\| > 0.$$

Finalmente normalizando $\widetilde{P}$, i.e., definiendo $P_1(x) = \frac{\widetilde{P}(x)}{\|\widetilde{P}(x)\|_2}$ llegamos a una contradicción con el teorema anterior.

Antes de empezar con la demostración del teorema de Brouwer necesitamos alguna notación adicional.

El espacio vectorial $\mathbb{R}^n$ puede verse como un subespacio vectorial de $\mathbb{R}^{n+1}$ identificando cada punto $x = (x_1, ..., x_n) \in \mathbb{R}^n$ con el punto $(x_1, ..., x_n, 0) \in \mathbb{R}^{n+1}$. De esta forma, cualquier punto de $\mathbb{R}^{n+1}$ se puede representar como (x, x_{n+1}) siendo $x \in \mathbb{R}^n$ y $x_{n+`1} \in \mathbb{R}$.

La esfera unidad $S^n \subset \mathbb{R}^{n+1}$ se puede dividir en dos semi-esferas- la superior o norte:

$$S^n_+ := \{(x, x_{n+1}) \in S^n \ : \ x_{n+1} \geq 0\}$$

y la semiesfera inferior o sur:

$$S^n_- := \{(x, x_{n+1}) \in S^n \ : \ x_{n+1} \leq 0\}$$

de esta forma la esfera $S^{n-1} \subset \mathbb{R}^n$ puede identificarse con el ecuador de S^n ya que $S^{n-1} = S^n_+ \cap S^n_-$.

Los puntos $e_{n+1} = (0, ..., 0, 1) \in \mathbb{R}^{n+1}$ y $-e_{n+1} = (0, ..., 0, -1) \in \mathbb{R}^{n+1}$ se pueden identificar respectivamente con el polo norte y polo sur de la esfera.

La proyección estereográfica desde e_{n+1} a S^n es la aplicación $S_+ : \mathbb{R}^n \to S^n$ que asigna a cada punto de $\mathbb{R}^n$ el punto de la intersección de S^n con la semirecta abierta que empieza en e_{n+1} y contiene al punto x. En concreto,

$$S_+(x) = (\frac{2x}{1 + \|x\|_2^2}, \frac{\|x\|_2^2 - 1}{1 + \|x\|_2^2}).$$

Esta aplicación es de clase $\mathcal{C}^\infty$ y transforma la bola B^n en S^n_-. También, para $x \in S^{n-1}$, $S_+(x) = x$. De forma análoga, la proyección estereogáfica S_- desde $-e_{n+1}$ a S^n se define:

$$S_-(x) = (\frac{2x}{1 + \|x\|_2^2}, \frac{1 - \|x\|_2^2}{1 + \|x\|_2^2}).$$

Ahora estamos en condiciones de probar el teorema de Brouwer.

Prueba del Teorema 4.1.4.

Empezaremos con el caso $n = 2k$. Supongamos que existe una aplicación continua $f : B^{2k} \to B^{2k}$ sin punto fijo. Entonces el campo vectorial $F_1(x) = x - f(x)$ no se anula sobre B^{2k}, y en cualquier punto de la esfera $x \in S^{2k-1}$ el campo está dirigido hacia fuera, i.e.,

$$\langle F_1(x), x\rangle = \langle x - f(x), x\rangle = \|x\|_2 - \langle x, f(x)\rangle = 1 - \langle f(x), x\rangle > 0.$$

(la desigualdad anterior es consecuencia de que $\langle f(x), x\rangle = \|x\|\|f(x)\|\cos(\theta)$ de donde se deduce que si $\langle x, f(x)\rangle = 1$, entonces $f(x) = x$ lo cual es una contradicción.)

Ahora definimos el campo vectorial:

$$F(x) = x - \frac{1 - \|x\|_2^2}{1 - \langle x, f(x)\rangle} f(x).$$

El campo F tampoco se anula sobre B^n. En efecto,

Si $F(x_0) = 0$, entonces

$$x_0 = \frac{1 - \|x_0\|_2^2}{1 - \langle x_0, f(x_0)\rangle} f(x_0),$$

por lo tanto

$$x_0 - x_0\langle x_0, f(x_0)\rangle = (1 - \|x_0\|_2^2) f(x_0),$$

con lo cual

$$\langle x_0 - x_0\langle x_0, f(x_0)\rangle, x_0\rangle = (1 - \|x_0\|_2^2)\langle f(x_0), x_0\rangle.$$

De donde se deduce que

$$\|x_0\|_2^2 - \langle x_0, f(x_0)\rangle \|x_0\|_2^2 = \langle x_0, f(x_0)\rangle - \langle x_0, f(x_0)\rangle \|x_0\|_2^2,$$

es decir

$$\|x_0\|_2^2 = \langle x_0, f(x_0)\rangle$$

luego $1 - \|x_0\|_2^2 = 1 - \langle x_0, f(x_0)\rangle$, con lo cual

$$0 = F(x_0) = x_0 - f(x_0).$$

Lo cual es una contradicción, ya que estamos suponiendo que f no tiene puntos fijos.

Por otra parte, es fácil observar que si $x \in S^{2k-1}$, entonces $F(x) = x$.

Ahora, para cada $x \in B^n$ consideremos el segmento que une el punto x con el punto $x + F(x)$, i.e., el conjunto

$$\{x + tF(x) \;:\; t \in [0, 1]\}.$$

La imagen de este segmento bajo la proyección estereográfica S_+ es un arco diferenciable con punto inicial sobre la semiesfera inferior:

$$t \longmapsto S_+(x + tF(x)).$$

Ahora asignamos a cada punto $y = S_+(x) \in S_-^{2k}$ el vector el tangente a este arco en el punto y. Más preciso, consideramos el campo vectorial T_- sobre S_-^{2k} por:

$$T_-(y) = \frac{d}{dt}S_+(x + tF(x))|_{t=0} = \lim_{t\to 0} \frac{S_+(x + tF(x)) - S_+(x)}{t} =$$

$$= \frac{2}{(1 + \|x\|_2^2)^2}((1 + \|x\|_2^2)F(x) - 2\langle x, F(x)\rangle x, 2\langle x, F(x)\rangle).$$

Entonces T_- es continuo, no se anula y es tangente a S_-^{2k}. Además, sobre el ecuador de S^{2k} T_- se dirige hacia arriba, i.e., $T_-(y) = e_{n+1}$ para cada $y \in S^{2k-1}$.

De manera totalmente análoga a la anterior, si tomamos el campo vectorial $-F$ en lugar del campo F y proyectamos sobre los arcos desde $-e_{n+1}$ por S_- para obtener el campo vectorial T_+ definido sobre S_+^{2k} por

$$T_+(y) = \frac{d}{dt}S_-(x - tF(x))|_{t=0} =$$

$$\frac{2}{(1 + \|x\|_2^2)^2}(-(1 + \|x\|_2^2)F(x) + 2\langle x, F(x)\rangle x, 2\langle x, F(x)\rangle).$$

Este campo vectorial sobre el ecuador se dirige hacia arriba y $T_-(y) = T_+(y) = e_{n+1}$ para cada $y \in S^{2k-1}$.

Finalmente, definimos para cada $y \in S^{2k}$ el siguiente campo vectorial:

$$T(y) := \begin{cases} T_-(y) & y \in S_-^{2k} \\ T_+(y) & y \in S_+^{2k} \end{cases}$$

Con lo cual T es un campo vectorial continuo, no se anula y es tangente a S^{2k} lo cual es imposible por el Teorema de la bola de Hairy. Luego f debe tener un punto fijo sobre B^{2k}.

Para terminar la prueba nos hace falta demostrarlo para el caso $n = 2k - 1$. En este caso, es suficiente observar que si $f : B^n \to B^n$ es una función continua que no tiene puntos fijos, entonces la función

$$\begin{array}{rccl} g : & B^{2k} & \longrightarrow & B^{2k} \\ & (x, x_{2k}) & \mapsto & g(x, x_{2k}) = (f(x), 0). \end{array}$$

La función g es continua en B^{2k} y no tiene puntos fijos, lo cual contradice el caso anterior.

4.1.1. Consecuencias del teorema de Brouwer.

Corolario 4.1.10 *Toda bola cerrada de* $(\mathbb{R}^n, \|.\|_2)$ *tiene la fppt.*

Prueba. Consideremos la bola cerrada $U(x_0, R)$ es claro que dicha bola la podemos escribir como $U(x_0, R) = x_0 + RU(0, 1)$.

Consideremos ahora una aplicación continua $T : U(x_0, R) \to U(x_0, R)$. Para ver que dicha aplicación tiene un punto fijo es suficiente considerar la aplicación $G : U(0, 1) \to U(0, 1)$ definida por

$$G(x) = \frac{1}{R}(T(x_0 + Rx) - x_0)$$

Por el teorema de Brouwer existe $x_1 \in U(0, 1)$ tal que $G(x_1) = x_1$. Con lo cual, el punto $y = x_0 + Rx_1$ es un punto fijo de T.

Lema 4.1.11 *Cada subconjunto* C *no vacío, cerrado y convexo de* $(\mathbb{R}^n, \|.\|_2)$ *es un retracto de* $\mathbb{R}^n$.

Prueba. Sea $x \in \mathbb{R}^n$ veamos que existe un único punto $R(x) \in C$ de forma que

$$\|R(x) - x\| = \inf\{\|x - y\| : y \in C\}.$$

En efecto llamemos $d = \inf\{\|x - y\| : y \in C\}$, entonces dado $n \in \mathbb{N}$ podemos encontrar $x_n \in C$ de forma que $\|x_n - x\| \leq d + \frac{1}{n}$. Como (x_n) es acotada y C es cerrado, podemos suponer que $x_{n_k} \to c$. Claramente se observa que

$$\|c - x\| = \inf\{\|x - y\| : y \in C\}.$$

Veamos ahora que c es único. Lo haremos por reducción al absurdo. Si p es otro punto verificando lo mismo que c por la convexidad de C se tendrá que $\frac{c+p}{2} \in C$. Con lo cual $d \leq \|\frac{c+p}{2} - x\|$. Sin embargo, si consideramos que $c \neq p$ llegamos a la siguiente contradicción:

$$d^2 \leq \|\frac{c+p}{2} - x\|^2 = \|\frac{c - x + p - x}{2}\|^2 = \frac{1}{4}(2d^2 + 2\langle c - x, p - x\rangle).$$

Pero además sabemos que

$$0 < \|c - x - (p - x)\|^2 = 2d^2 - 2\langle c - x, p - x\rangle,$$

es decir,

$$\langle c - x, p - x\rangle < d^2.$$

Luego

$$d^2 \leq \|\frac{c+d}{2} - x\|^2 = \frac{1}{4}(2d^2 + 2\langle c - x, p - x\rangle) < d^2.$$

Lo cual es una contradicción y por tanto podemos concluir que $c = p$.

Ahora es suficiente definir $R(x) = c$. Sólamente nos queda por ver que R es una aplicación continua.

Para ello será suficiente ver que para cualquier $x \in \mathbb{R}^n$ y para cualquier $y \in C$ se cumple:

$$\langle y - R(x), x - R(x)\rangle \leq 0 \tag{4.1}$$

En efecto, si asumimos (4.1) se tiene

$$\langle R(y) - R(x), x - R(x)\rangle \leq 0 \text{ y } \langle R(x) - R(y), y - R(y)\rangle \leq 0$$

Con lo cual

$$\langle R(y) - R(x), R(y) - R(x) + x - y\rangle \leq 0$$

De donde se obtiene que

$$\|R(x) - R(y)\|^2 \leq \|R(x) - R(y)\|\|x - y\|,$$

es decir,

$$\|R(x) - R(y)\| \leq \|x - y\|.$$

Lo único que nos queda por justificar es la expresión (4.1). Veamos que es cierta,

dado $n \in \mathbb{N}$ definimos $v_n = (1 - \frac{1}{n})R(x) + \frac{1}{n}y \in C$

$$\|x - R(x)\| \leq \|x - v_n\| = \|x - R(x) - \frac{1}{n}(y - R(x)\|,$$

por lo tanto

$$\|x - R(x)\|^2 \leq \|x - R(x)\|^2 - 2\frac{1}{n}\langle x - R(x), y - R(x)\rangle + \frac{1}{n^2}\|y - R(x)\|^2$$

Es decir $2\langle x - R(x), y - R(x)\rangle \leq \frac{1}{n}\|y - R(x)\| \to 0$ cuando $n \to \infty$.

Corolario 4.1.12 *Cada subconjunto no vacío, cerrado, acotado y convexo C de $(\mathbb{R}^n, \|.\|_2)$ tiene la fppt.*

Prueba.
Como C está acotado lo podemos incluir en alguna bola cerrada de $\mathbb{R}^n$ y entonces será un retracto de dicha bola. Ahora teniendo en cuenta que las bolas tienen la fppt y el Lema 4.1.3 obtenemos el resultado.

Corolario 4.1.13 *(Brouwer) Sea C un subconjunto no vacío, cerrado, acotado y convexo de un espacio de Banach real de dimensión finita $(X, \|.\|)$. Entonces C tiene la fppt.*

Prueba. Supongamos que $\dim(X) = n$, es bien conocido que $(X, \|.\|)$ es homeomorfo a $(\mathbb{R}^n, \|.\|_2)$. Con lo cual aplicando el lema 4.1.2 se obtiene el resultado.

Es muy sencillo construir ejemplos que muestren que ninguna de las hipótesis del teorema de Brouwer puede ser suprimida sin perder la tesis.

Ejemplo 4.1.14 *Consideremos $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(x) = x + a$ donde a es un vector no nulo de $\mathbb{R}^n$. Claramente T es continua, no tiene punto fijo y $\mathbb{R}^n$ es cerrado y convexo (pero no acotado).*

Ejemplo 4.1.15 *Consideremos $T : [0, 1] \cup [2, 3] \to [0, 1] \cup [2, 3]$ definida por $T(x) = x + 2$ si $x \in [0, 1]$ y $T(x) = x - 2$ si $x \in [2, 3]$. Claramente T es continua, sin puntos fijos y el dominio de T es compacto (pero no convexo).*

4.2. El Teorema de Schauder.

La extensión del Teorema de Brouwer a espacios normados de dimensión infinita fue dada por J. Schauder en 1930, partiendo de una linea de razonamiento expuesta en ciertos trabajos anteriores de G.D. Birkhoff y O.D. Kellogg que demostraron que los compactos de $L^2[0, 1]$ tienen la propiedad del punto fijo para auto-aplicaciones continuas.

Veremos una demostración del teorema de Schauder apoyándola en el teorema de Brouwer.

Teorema 4.2.1 (Schauder) *Cualquier subconjunto no vacío, convexo y compacto de un espacio de Banach tiene la fppt.*

Prueba. Sea K un compacto convexo de un espacio de Banach X. Consideremos $T : K \to K$ una aplicación continua.

Lo único que tenemos que ver es:

$$\inf\{\|x - Tx\| \ : \ x \in K\} = 0, \tag{4.2}$$

ya que en este caso podemos encontrar una sucesión (x_n) de elementos de K de forma que

$$\lim_{n\to\infty} \|x_n - Tx_n\| = 0.$$

Ahora bien, como K es compacto, podremos encontrar una subsucesión (x_{n_k}) que será convergente, llamemos $x_0 \in K$ al valor de su límite. Es claro ahora por la continuidad de T que x_0 es un punto fijo.

Veamos ahora que se verifica (4.2). Para ver (4.2) es suficiente probar que dado un $\varepsilon > 0$ fijo, existe $x_\varepsilon \in K$ tal que $\|x_\varepsilon - Tx_\varepsilon\| \leq \varepsilon$.

Consideremos el conjunto siguiente:

$$\mathcal{U} := \{B(x, \varepsilon) \ : \ x \in K\}.$$

El conjunto $\mathcal{U}$ es un recubrimiento por conjuntos abiertos de K. Como K es compacto, podremos extraer un subrecubrimiento finito,i.e., existiran $a_1, ..., a_p$ elementos de K de forma que

$$K \subseteq \bigcup_{i=1}^{p} B(a_i, \varepsilon).$$

Esto nos permite construir las siguientes funciones reales:

$$\begin{array}{rccl} F_i : & K & \longrightarrow & \mathbb{R} \\ & x & \mapsto & F_i(x) = \begin{cases} 0, & \|x - a_i\| \geq \varepsilon \\ \varepsilon - \|x - a_i\|, & \|x - a_i\| \leq \varepsilon \end{cases} \end{array}$$

Dado $x \in K$ está claro que existirá un $i_0 \in \{1, ..., p\}$ de forma que $\|x - a_{i_0}\| < \varepsilon$, con lo cual podemos afirmar que

$$\sum_{i=1}^{p} F_i(x) > 0.$$

Este hecho nos permite construir la aplicación

$$\begin{array}{rccl} \varphi : & K & \longrightarrow & K_0 := K \cap Lin(a_1, ..., a_p) \\ & x & \mapsto & \varphi(x) = \frac{\sum_{i=1}^{p} F_i(x) a_i}{\sum_{i=1}^{p} F_i(x)} \end{array}$$

Obviamente φ es una función continua, y para cada $x \in K$ se verifica:

$$\|x - \varphi(x)\| = \|\frac{\sum_{i=1}^{p} F_i(x)(a_i - x)}{\sum_{i=1}^{p} F_i(x)}\|$$

Teniendo en cuenta que si $\|x - a_i\| \geq \varepsilon$, entonces $F_i(x) = 0$. Podemos concluir que

$$\|x - \varphi(x)\| \leq \frac{\sum_{\{i\ : \|x-a_i\|<\varepsilon\}} F_i(x)\|a_i - x\|}{\sum_{\{i\ : \|x-a_i\|<\varepsilon\}} F_i(x)} \leq \varepsilon.$$

Ahora construimos la aplicación $\tilde{T} : K \to K$ definida por $\tilde{T} = \varphi \circ T$. De este modo hemos construido una aplicación $\tilde{T} : K_0 \to K_0$ continua y K_0 compacto convexo de un espacio de Banach de dimensión finita, entonces le podemos aplicar el teorema de Brouwer y obtener $x_\varepsilon \in K_0$ tal que $\tilde{T}(x_\varepsilon) = x_\varepsilon$. Por lo tanto,

$$\|x_\varepsilon - Tx_\varepsilon\| \leq \|x_\varepsilon - \tilde{T}x_\varepsilon\| + \|\tilde{T}x_\varepsilon - Tx_\varepsilon\| = \|\varphi(T(x_\varepsilon)) - T(x_\varepsilon)\| \leq \varepsilon.$$

4.3. Compacidad en espacios normados.

En dimensión finita es relativamente sencillo identificar los conjuntos compactos ya que son los conjuntos cerrado y acotados. En dimensión infinita esto cambia de forma importante y como veremos en los siguientes ejemplos la hipótesis de compacidad no puede ser reemplazada.

Ejemplo 4.3.1 *En el espacio de las sucesiones convergentes a cero* $(c_0, \|.\|_\infty)$. *Podemos considerar sobre su bola cerrada unidad la siguiente aplicación:*

$$T(x_1, x_2, x_3, ...) := (1, x_1, x_2, x_3, ...).$$

Claramente esta aplicación deja invariante a la bola unidad cerrada, es continua (de hecho es una isometría) y no puede tener puntos fijos en $(c_0, \|.\|_\infty)$.

Como hemos visto, el teorema de Schauder actua sobre conjuntos compactos convexos, el problema que se nos plantea ahora es reconocer los compactos en espacios de dimensión infinita. Para este fin introducimos los siguientes conceptos.

Definición 4.3.2 *Sea D un subconjunto de un espacio de Banach $(X, \|.\|)$ y sea $\epsilon > 0$. Se dice que $A \subset X$ es una $\epsilon-$red de D, cuando para cada punto $x \in D$ existe al menos un punto $a \in A$ tal que $\|x - a\| \leq \epsilon$.*

Un conjunto D se llama totalmente acotado cuando dado cualquier $\epsilon > 0$ existe una ϵ-red finita de D.

Es una consecuencia obvia de la definición que todo subconjunto compacto es totalmente acotado. Nuestro objetivo ahora es ver la implicación contraria.

Teorema 4.3.3 *Sea D un subconjunto cerrado y totalmente acotado de un espacio de Banach, entonces D es compacto.*

Prueba. Tenemos que ver que toda sucesión (x_n) de D, tiene al menos un punto de acumulación.

Consideremos la familia $\mathcal{U} = \{B(x, 1) \ : \ x \in A\}$ donde A es una 1-red finita de D. Puesto que estas bolas recubren todo D y además hay sólamente un número finito, al menos una de estas bolas, llamémosla B_1, contiene una subsucesión infinita $x_1^1, ..., x_n^1, ...$ de la sucesión (x_n).

Consideremos ahora en B_1 una $\frac{1}{2}$-red finita de (x_n^1). Haciendo el mismo razonamiento que antes, podemos encontrar una bola, llamémosla B_2, que contiene una subsucesión (x_n^2) de (x_n^1). Razonando de la misma forma se llega:

Dado $m \in \mathbb{N}$ existe una bola $B_m \subset B_{m-1}$ de radio $\frac{1}{m}$ tal que contiene una subsucesión (x_n^m) de la sucesión inicial (x_n). Por el principio de las bolas encajadas se tiene que

$$\bigcap_{i=1}^{\infty} B_i \neq \emptyset$$

lo cual implica que (x_n) tiene un punto de acumulación.

Corolario 4.3.4 *Sea C un subconjunto totalmente acotado de un espacio de Banach, entonces C es relativamente compacto.*

En Análisis no lineal, uno de los espacios de Banach más interesantes es el espacio $(C([a, b]), \|.\|_\infty)$. Para identificar los subconjuntos relativamente compactos de este espacio se utiliza el llamado teorema de Ascoli-Arzelá.

Para enunciarlo, necesitaremos los conceptos siguientes.

Definición 4.3.5 *Una familia $\mathcal{F} \subseteq C([a, b])$ se llama equiacotada, cuando existe un número $K > 0$ tal que*

$$|\varphi(x)| \leq K$$

para todas las funciones $\varphi \in \mathcal{F}$ *y para todo* $x \in [a,b]$.

La família $\mathcal{F}$ *se llama equicontinua, cuando para cada* $\epsilon > 0$ *existe un* $\delta > 0$ *tal que*

$$|\varphi(x) - \varphi(y)| \leq \epsilon$$

para toda función $\varphi \in \mathcal{F}$ *siempre que* $|x - y| \leq \delta$

Teorema 4.3.6 (Ascoli-Arzelá) *Una familia* $\mathcal{F}$ *de funciones de* $C([a,b])$ *es relativamente compacta si,y sólo si, es una familia equiacotada y equicontinua.*

4.3.1. La medida de no compacidad de Kuratowski.

Una versión modificada del teorema de Schauder muy útil es la siguiente:

Teorema 4.3.7 *Sea* K *un conjunto no vacío, cerrado y convexo de un espacio de Banach* X *y sea* $T : K \to K$ *una aplicación continua de forma que* $\overline{T(K)}$ *es compacto. Entonces* T *tiene un punto fijo en* K.

Recordemos que las aplicaciones que transforman conjuntos acotados en conjuntos relativamente compactos se llaman *aplicaciones compactas*. En vista de esta definición y del teorema anterior se tiene:

Corolario 4.3.8 *Sea* K *un conjunto no vacío, cerrado, acotado y convexo de un espacio de Banach* X *y sea* $T : K \to K$ *una aplicación continua y compacta. Entonces* T *tiene un punto fijo en* K.

Prueba. Como K es convexo, cerrado y $T(K) \subseteq K$ es claro que $K_0 = \overline{co}(T(K))$ es un subconjunto de K, T-invariante. Si K_0 fuese compacto, entonces podríamos aplicar el teorema anterior y obtener el resultado. Luego el objetico a partir de ahora es probar este hecho.

Para resolver la cuestión que se nos plantea en la demostración anterior introducimos la medida de no compacidad de Kuratowski.

Definición 4.3.9 *Sea* D *un subconjunto acotado de un espacio de Banach* $(X, \|.\|)$. *A una família finita de subconjuntos acotados de* X, $P = \{P_i, ..., P_n\}$ *verificando que*

$$D \subseteq \bigcup_{i=1}^{n} P_i,$$

se le llama partición finita de D, *y se define la norma de dicha partición como:*

$$\|P\| := \sup\{\text{diam}(P_i) \ : \ i = 1, ...n\}.$$

La medida de no compacidad de Kuratowski $\phi(D)$ de D es:

$$\phi(D) = \inf\{\|P\| : \ P \text{ es una particion finita de } D\}.$$

Es claro, a partir de la definición, que si A es un conjunto acotado, entonces $\phi(A) \leq \operatorname{diam}(A)$.

Proposición 4.3.10 *La medida de no compacidad de Kuratowski satisface las siguientes propiedades:*
(a). $\phi(D) = \phi(\overline{D})$.
(b). $\phi(D) = 0$ si, y sólo si, $\overline{D}$ es compacto.
(c). Si $A \subseteq B$, entonces $\phi(A) \leq \phi(B)$.
(d). $\phi(A \cup B) = \max\{\phi(A), \phi(B)\}$.
(e). $\phi(A + B) \leq \phi(A) + \phi(B)$.
(f). $\phi(\lambda A) = |\lambda|\phi(A)$.

Todas las propiedades anteriores son consecuencia inmediata de la definición de medida de no compacidad. Sólamente remarcaremos que (b) sale como consecuencia de que un conjunto es compacto si, y sólo si, es cerrado y completamente acotado.

Veamos ahora la propiedad que nos interesa para obtener la segunda forma del teorema de Schauder.

Teorema 4.3.11 $\phi(\overline{co}(A)) = \phi(A)$.

Prueba.

Para la demostración seguiremos varios pasos:

Paso 1. Para cada conjunto acotado A, $\phi(N_\epsilon(A)) \leq \phi(A) + 2\epsilon$. Donde $N_\epsilon(A) = \bigcup_{x \in A} B(x, \epsilon)$.

Para ver este paso, únicamente hay que percatarse de que si S es un conjunto acotado, entonces $\operatorname{diam}(N_\epsilon(S)) \leq \operatorname{diam}(S) + 2\epsilon$.

Paso 2. Supongamos que $C = A \cup B$ donde A, B son dos conjuntos acotados y convexos. Entonces

$$\phi(co(C)) \leq \max\{\phi(A), \phi(B)\}.$$

En efecto, sea $x \in co(C)$ y supongamos que $x \notin A \cup B$. Entonces existiran $x_i \in C$ y $\lambda_i \geq 0$, $i = 1, ..., n$, con $\sum_{i=1}^n \lambda_i = 1$, tal que $x = \sum_{i=1}^n \lambda_i x_i$. Reagrupando, podemos suponer que $x_i \in A$ para $1 \leq i \leq m$ y $x_i \in B$ para $m + 1 \leq i \leq n$. Por lo tanto,

$$x = (\sum_{i=1}^{m} \lambda_i) \frac{\sum_{i=1}^{m} \lambda_i x_i}{\sum_{i=1}^{m} \lambda_i} + (\sum_{i=m+1}^{n} \lambda_i) \frac{\sum_{i=m+1}^{n} \lambda_i x_i}{\sum_{i=m+1}^{n} \lambda_i},$$

de esta forma se obtiene que cada $x \in co(C) \setminus (A \cup B)$ puede escribirse en la forma

$$x = \alpha u + (1 - \alpha)v$$

donde $u \in A$, $v \in B$ y $0 \leq \alpha \leq 1$.

Como A y B son conjuntos acotados podemos asegurar que existe $M > 0$ de forma que $\|a\| + \|b\| \leq M$ para todo $a \in A$ y para todo $b \in B$. Con lo cual, dado $\epsilon > 0$ podemos elegir $n \in \mathbb{N}$ tal que $\frac{M}{n} < \epsilon$. Para cada $i = 0, 1, ..., n$, definimos

$$C_i = \{x \in co(C) \ : \ x = \frac{i}{n} u + (1 - \frac{i}{n})v \text{ para algun } u \in A, \ v \in B\}.$$

Notar que dado $\alpha \in]0, 1[$, entonces $\alpha \in [\frac{i}{n}, \frac{i+1}{n}]$ para algún $0 \leq i \leq n-1$, con lo cual si $x \in co(C) \setminus (A \cup B)$ se cumple que $x \in N_\epsilon(C_i)$ para algún $i = 0, 1, ..., n$. Consecuentemente,

$$co(C) \subseteq (\bigcup_{i=1}^{n} N_\epsilon(C_i)) \cup A \cup B,$$

lo que implica:

$$\phi(co(C)) \leq \text{máx}\{\phi(A), \phi(B), \text{máx}\{\phi(N_\epsilon(C_i)) \ : \ i = 1, ..., n\}\} \leq$$

$$\leq \text{máx}\{\phi(A), \phi(B), \text{máx}\{\phi(C_i) + 2\epsilon \ : \ i = 1, ..., n\}\}.$$

Ahora por (e) y (f), para cada i se tiene,

$$\phi(C_i) = \phi(\frac{i}{n} A + (1 - \frac{i}{n})B) \leq \frac{i}{n}\phi(A) + (1 - \frac{i}{n})\phi(B),$$

lo que implica

$$\phi(C_i) \leq \text{máx}\{\phi(A), \phi(B)\}.$$

Por lo tanto,

$$\phi(co(C)) \leq \text{máx}\{\phi(A), \phi(B)\} + 2\epsilon.$$

El paso 2 se obtiene ahora del hecho de que $\epsilon > 0$ es arbitrario.

Paso 3. Supongamos que $C = \bigcup_{i=1}^n A_i$ donde cada A_i es un conjunto acotado y convexo. Entonces

$$\phi(co(C)) \leq \max\{\phi(A_i) \; : \; i = 1, ..., n\}.$$

Esto es una consecuencia inmediata del paso 2.

Una vez vistos estos tres pasos, veamos el resultado:

Supongamos que $\phi(A) < a$. Entonces A puede escribirse como una unión finita $A \subseteq \bigcup_{i=1}^n A_i$, donde $\operatorname{diam}(A_i) \leq a$ y como $\operatorname{diam}(co(A_i)) = \operatorname{diam}(A_i)$ podemos asumir que cada A_i es convexo. Con lo cual tenemos que

$$co(A) \subseteq co\left(\cup_{i=1}^n A_i\right).$$

Entonces, por el paso 3 se tiene:

$$\phi(co(A)) \leq \max\{\phi(A_i) \; : \; i = 1, ..., n\} \leq \max\{\operatorname{diam}(A_i) \; : \; i = 1, ..., n\} \leq a.$$

Como lo anterior es cierto para todo a tal que $\phi(A) < a$ la prueba del teorema está completa.

4.3.2. Principio de Leray-Schauder.

Sea X un espacio de Banach. Como ya se ha visto, muchos problemas analíticos de encontrar solución se pueden reducir a encontrar un punto fijo de una determinada función $f : X \to X$. Sin embargo, el teorema de Schauder en muchos de estos problemas no se puede usar puesto que para ello necesitamos que la aplicación f sea compacta y además deje invariante un determinado conjunto cerrado, acotado y convexo de X.

La hipótesis que se suele usar para obtener una versión más conveniente del teorema de Schauder se llama *condición de frontera de Leray-Schauder*. Dada una aplicación $f : X \to X$, la aplicación f satisface esta condición si existe $r > 0$ tal que si $\|x\| = r$, entonces $f(x) \neq \lambda x$ para todo $\lambda > 1$. En particular, como $\|\lambda x\| = \lambda\|x\|$, es claro que f satisface la condición de frontera de Leray-Schauder si $\|x\| = r$ implica que $\|f(x)\| \leq r$, esta es la forma más habitual de expresar la condición de frontera de Leray-Schauder.

Observar que esta condición no impone ninguna restricción sobre cómo debe ser r.

Teorema 4.3.12 (Principio de Leray-Schauder) *Sea X un espacio de Banach y $f : X \to X$ una aplicación continua y compacta. Si f verifica la condición de frontera de Leray-Schauder, entonces f tiene un punto fijo.*

Prueba.

La condición de Leray-Schauder nos asegura la existencia de $r > 0$ tal que si $\|x\| = r$, entonces $f(x) \neq \lambda x$ para todo $\lambda > 1$. Consideremos $C := B_r = \{x \in X \ : \ \|x\| \leq r\}$, i.e., la bola centrada en el cero y de radio el r de la condición.

Si restringimos la función f a dicha bola nos queda $f : B_r \to X$. No hay ninguna razón para esperar que f deje invariante dicha bola. Así que introducimos la función $\rho : X \to B_r$ (la retracción de X sobre B_r la cual viene dada por:

$$\rho(x) = \begin{cases} x, & \|x\| \leq r \\ \frac{r}{\|x\|}x, & \|x\| \geq r \end{cases}$$

Ahora definimos la aplicación $f^* = \rho \circ f : B_r \to B_r$.

Como f es compacta sabemos que $f(B_r)$ está continido en un compacto, llamemosle K, entonces $f^*(B_r) \subseteq \rho(K)$, el cual es compacto ya que ρ es una aplicación continua, así f^* es también una aplicación compacta. Esto significa que f^* está bajo las hipótesis del corolario 4.3.8 y por lo tanto existe $x_0 \in B_r$ tal que $f^*(x_0) = x_0$.

Finalmente vamos a probar que $f(x_0) \in B_r$, con lo cual tendremos que $x_0 = f^*(x_0) = \rho(f(x_0)) = f(x_0)$. Lo haremos por reducción al absurdo.

Supongamos que $f(x_0) \notin B_r$ lo que quiere decir que $\|f(x_0)\| > r$, por lo tanto

$$x_0 = f^*(x_0) = \rho(f(x_0)) = \frac{rf(x_0)}{\|f(x_0)\|}.$$

Esto significa,

$$\|x_0\| = \|\frac{rf(x_0)}{\|f(x_0)\|}\| = r.$$

Reescribiendo la expresión anterior se tiene

$$f(x_0) = \frac{\|f(x_0)\|}{r}x_0 = \lambda x_0$$

con $\lambda > 1$ Lo cual es una contradicción por la condición de frontera de Leray-Schauder.

Ejemplo 4.3.13 *La ecuación integral*

$$u(x) = 2\int_a^b \sin(u(s))ds + e^x$$

tiene una solución $y \in \mathcal{C}([a,b])$.

En efecto, consideremos la aplicación $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ definida por

$$T(u)(x) = 2\int_a^b \sin(u(t))dt + e^x.$$

Esta aplicación evidentemente está bien definida y es continua. Además si calculamos

$$\left|2\int_a^b \sin(u(t))dt + e^x\right| \leq 2(b-a) + e^b,$$

luego si llamamos $r = 2(b-a)+e^b$ se tiene que $\|T(u)\| \leq r$ sea cual sea u, de donde se deduce que si $\|u\| = r$, entonces $\|T(u)\| \leq r$. (Condición frontera de Leray-Schauder).

Luego lo único que nos queda por ver es que T es compacta.

Para ello primero veremos que si B es un acotado de $\mathcal{C}([a,b])$, entonces $T(B)$ es equiacotado.

Para cualquier función $T(u) \in T(B)$ se tiene, por el razonamiento anterior, que $\|Tu\| \leq r$. Luego $T(B)$ es equiacotado.

Por último, para probar que $T(B)$ es equicontinuo razonamos del siguiente modo.

Como la exponencial es uniformemente continua en $[a,b]$ se cumple que dado $\epsilon > 0$ existe $\delta > 0$ tal que si $|x - x'| \leq \delta$ entonces $|e^x - e^{x'}| < \epsilon$. Con lo cual

$|T(u)(x) - T(u)(x')| = |e^x - e^{x'}| < \epsilon$. Siendo δ independiente de la función u que tomemos. Luego $T(B)$ es equicontinuo y, una vez más, por el teorema de Arzela-Ascoli concluimos que T es una aplicación compacta.

4.3.3. Aplicaciones condensantes.

Como hemos visto hasta ahora los teoremas de punto fijo en espacios de Banach se enmarcan en dos categorias de naturaleza distinta:

Teoremas topológicos del punto fijo (Teorema de Brouwer, Schauder).

Teoremas métricos del punto fijo (Teorema de Banach).

Hay sin embargo una larga clase de aplicaciones las cuales sirven de puente entre la teoría métrica y topológica del punto fijo. Esta clase de aplicaciones fue introducida por Darbo en 1955.

Definición 4.3.14 *Sea K un subconjunto de un espacio de Banach $(X, \|.\|)$. Una aplicación $T : K \to X$ se llama condensante, si T es acotada, continua* *y para todo subconjunto acotado D de K se cumple que*

$$\phi(T(D)) < \phi(D)$$

siempre que $\phi(D) > 0$.

El siguiente teorema fue probado por Sadovskii en 1967. En 1955 Darbo probó el mismo teorema pero para una clase más restrictiva de aplicaciones.

Teorema 4.3.15 *Supongamos que K es un conjunto no vacío, acotado, cerrado y convexo de un espacio de Banach X y supongamos que $T : K \to K$ es una aplicación condensante. Entonces T tiene un punto fijo.*

Prueba. Fijemos $x \in K$ y sea Σ la colección de todos los subconjuntos cerrados y convexos D de K de forma que $x \in D$ y $T : D \to D$. LLamemos

$$B = \bigcap_{D \in \Sigma} D,$$

y sea

$$C = \overline{co}\{T(B) \cup \{x\}\}.$$

Como $x \in B$ y $T : B \to B$ se debe cumplir que $C \subseteq B$. Esto, a su vez, implica que $T(C) \subseteq T(B) \subseteq C$. Como $x \in C$ se sigue que $C \in \Sigma$. Por lo tanto, $B \subseteq C$, de donde concluimos que $C = B$.

Ahora tenemos que $T(C) = T(B) \subseteq C$ y

$$\phi(C) = \phi(\overline{co}\{T(B) \cup \{x\}\}) = \phi(\{T(B) \cup \{x\}\}) = \phi(T(B)) = \phi(T(C)).$$

Como T es condensante la única solución es que $\phi(C) = 0$, es decir que C sea compacto. Ahora bien como $T : C \to C$ es una aplicación continua y C es un compacto convexo, podemos aplicar el teorema de Schauder y de esta forma T debe tener un punto fijo.

Como hemos visto en el apartado anterior, el teorema de Schauder se puede suavizar usando las condiciones frontera de Leray-Schauder. El mismo resultado es válido para aplicaciones condensantes y éste resultado fue obtenido por Petryshyn en 1971.

Teorema 4.3.16 (Principio de Leray-Schauder para condensantes) *Sea X un espacio de Banach y $T : X \to X$ una aplicación continua y condensante. Si T verifica la condición de frontera de Leray-Schauder, entonces T tiene un punto fijo.*

Prueba.

La condición de Leray-Schauder nos asegura la existencia de $r > 0$ tal que si $\|x\| = r$, entonces $T(x) \neq \lambda x$ para todo $\lambda > 1$. Consideremos $C := B_r = \{x \in X \; : \; \|x\| \leq r\}$, i.e., la bola centrada en el cero y de radio el r de la condición.

Si restringimos la función T a dicha bola nos queda $T : B_r \to X$. No hay ninguna razón para esperar que T deje invariante dicha bola. Así que introducimos la función $\rho : X \to B_r$ (la retracción de X sobre B_r la cual viene dada por:

$$\rho(x) = \begin{cases} x, & \|x\| \leq r \\ \frac{r}{\|x\|}x, & \|x\| \geq r \end{cases}$$

Veamos ahora que la aplicación ρ no aumenta la medida de no compacidad (cuando una aplicación verifica esta condición se dice que es una 1-set contracción). En efecto, sea M un subconjunto acotado y no vacío de X, está claro que si tomamos $x \in M$ pueden ocurrir dos casos:

$$x \in M \Rightarrow \begin{cases} \|x\| \leq r \\ \|x\| > r \end{cases}$$

Si $\|x\| \leq r$, entonces $\rho(x) = x \in M \subseteq \mathrm{co}(M \cup \{0\})$.

Si $\|x\| > r$, entonces $\rho(x) = \frac{r}{\|x\|}x$. Llamando $\lambda = \frac{r}{\|x\|}$ queda que $\rho(x) = \lambda x + (1 - \lambda)0 \in \mathrm{co}(M \cup \{0\})$.

Por lo tanto, acabamos de ver que $\rho(M) \subset \mathrm{co}(M \cup \{0\})$.

Teniendo en cuenta las propiedades de la medida de compacidad se concluye que $\phi(\rho(M)) \leq \phi(M)$.

Ahora definimos la aplicación $T^* = \rho \circ T : B_r \to B_r$.

Veamos ahora que T^* es condensante. Si A es un subconjunto acotado y no compacto de la bola, se tendrá:

$$\phi(T^*(A)) = \phi(\rho(T(A))) \leq \phi(T(A)) < \phi(A).$$

Esto significa que T^* está bajo las hipótesis del teorema 4.3.15 y por lo tanto existe $x_0 \in B_r$ tal que $T^*(x_0) = x_0$.

Finalmente vamos a probar que $T(x_0) = x_0$.

Si $T(x_0) \in B_r$, entonces $x_0 = T^*(x_0) = \rho(T(x_0)) = T(x_0)$.

Si $T(x_0) \notin B_r$, se tendrá que $\|T(x_0)\| > r$, por lo tanto

$$x_0 = T^*(x_0) = \rho(T(x_0)) = \frac{rT(x_0)}{\|T(x_0)\|}.$$

Esto significa,

$$\|x_0\| = \|\frac{rT(x_0)}{\|T(x_0)\|}\| = r.$$

Reescribiendo la expresión anterior se tiene

$$T(x_0) = \frac{\|T(x_0)\|}{r} x_0 = \lambda x_0$$

con $\lambda > 1$ Lo cual es una contradicción por la condición de frontera de Leray-Schauder.

4.4. Aplicaciones.

4.4.1. Teorema de Peano.

Consideremos el problema consistente en encontrar alguna función derivable y cuya gráfica pase por el punto del plano (x_0, y_0) y que para cierto número positivo r se cumpla que $y'(x) = f(x, y(x))$ siempre que $x \in]x_0 - r, x_0 + r[$, donde f es una función continua en un abierto D del plano que contiene a (x_0, y_0). Este problema se escribe habitualmente de la siguiente forma:

$$\begin{cases} y'(x) = f(x, y(x)) \\ \\ y(x_0) = y_0. \end{cases} \tag{4.3}$$

Teorema 4.4.1 (Peano) *El problema anterior tiene solución.*

Prueba. Sea $r > 0$ tal que $S = \{(x, y) : \ \|(x, y) - (x_0, y_0)\|_\infty \leq r\} \subseteq D$. Podemos escoger $h > 0$ de forma que $h \leq r$ y $hK \leq r$, donde

$$K = \text{máx}\{|f(x, y)| : \ (x, y) \in S\}.$$

Consideremos el espacio de Banach $(\mathcal{C}([x_0 - h, x_0 + h]), \|.\|_\infty)$. Consideremos la bola cerrada $U(y_0, r)$, donde y_0 es la función que toma el valor constante y_0.

Definimos el operador $T : U(y_0, r) \to U(y_0, r)$ donde a cada función $u \in U(y_0, r)$ le hacemos corresponder la función $T(u)$ definida de la siguiente forma:

$$T(u)(x) := y_0 + \int_{x_0}^{x} f(t, u(t))dt,$$

observar que dicho operador está bien definido ya que si $u \in U(y_0, r)$ entonces para cada $t \in [x_0 - h, x_0 + h]$,

$$|u(t) - y_0| \leq \|u - y_0\|_\infty \leq r,$$

por lo que $y_0 - r \leq u(t) \leq y_0 + r$, y

$$(t, u(t)) \in [x_0 - h, x_0 + h] \times [y_0 - r, y_0 + r] \subseteq S,$$

y tiene sentido $f(t, u(t))$.

Por el teorema fundamental del cálculo es claro que $Tu \in \mathcal{C}([x_0 - h, x_0 + h])$, además

$$\|Tu - y_0\|_\infty = \text{máx}\{|\int_{x_0}^{x} f(t, u(t))dt| \ : \ x \in [x_0 - h, x_0 + h]\} \leq hK \leq r.$$

Una vez visto que el operador T está bien definido y que deja invariante a la bola $U(y_0, r)$. Veamos que $T(U(y_0, r))$ es un conjunto equicontinuo.

$$|Tu(x) - Tu(y)| = |\int_{x_0}^{x} f(t, u(t))dt - \int_{x_0}^{y} f(t, u(t))dt| \leq K|x - y|$$

Como este razonamiento es independiente de la función u escogida, se tiene que efectivamente $T(U(y_0, r))$ es un conjunto equicontinuo y equiacotado (por ser acotado para la norma del espacio).

Veamos ahora que el operador T es continuo.

En efecto, como f es continua en S y S es un compacto de $\mathbb{R}^2$, es claro que f será uniformemente continua sobre S (Teorema de Heine-Cantor). De este modo, dado $\epsilon > 0$ existirá $\delta > 0$ de forma que si $(x, y_1), (x, y_2) \in S$ con $|y_1 - y_2| = \|(x, y_1) - (x, y_2)\|_\infty \leq \delta$ entonces

$$|f(x, y_1) - f(x, y_2)| < \frac{\epsilon}{h}.$$

Si $u, v \in U(y_0, r)$ con $\|u - v\|_\infty < \delta$, entonces para todo $t \in [x_0 - h, x_0 + h]$, se cumplirá

$$|u(t) - v(t)| \leq \|u - v\|_\infty < \delta.$$

Entonces

$$\|Tu - Tv\|_\infty = \text{máx}\{|\int_{x_0}^{x} (f(t, u(t)) - f(t, v(t)))dt| \ : x \in [x_0 - h, x_0 + h]\} \leq$$

$$\leq \max\{|x_0 - x|\frac{\epsilon}{h} \ : \ x \in [x_0 - h, x_0 + h]\} \leq h\frac{\epsilon}{h} = \epsilon.$$

En definitiva $T : U(y_0, r) \to U(y_0, r)$ es un operador compacto y por lo tanto le podemos aplicar el teorema de Schauder, y así existe una función $y = Ty$ en $U(y_0, r)$ es decir, para todo $x \in [x_0 - h, x_0 + h]$ se verifica que

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt,$$

desde donde es inmediato ver que y es solución del problema (4.3).

4.5. Problemas

Problema 4.5.1 *Sea A un conjunto acotado y* $\mathrm{co}(A)$ *su envoltura convexa.*
(a). Probar que dado $x \in X$, se tiene que

$$\sup\{\|x - b\| \ : \ b \in \mathrm{co}(A)\} = \sup\{\|x - a\| \ : \ a \in A\}.$$

(b) Probar que $\mathrm{diam}(A) = \mathrm{diam}(\mathrm{co}(A))$.

Problema 4.5.2 *Sea K un subconjunto de un espacio de Banach X. Supongamos que $T_1 : K \to X$ es una aplicación contractiva y supongamos que $T_2 : K \to X$ es una aplicación compacta. Entonces $T = T_1 + T_2$ es condensante.*

Problema 4.5.3 *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función continua. Probar que el sistema*

$$\begin{cases} x &= 10^{27} + \sin(f(x, y)) \\ y &= \cos(f(x, y)) \end{cases}$$

tiene una solución en $\mathbb{R}^2$.

Problema 4.5.4 *Sea α un número real y sea $f \in \mathcal{C}([a, b])$ una función fija. Probar que la ecuación*

$$u(x) = \alpha \int_a^b \sin(u(t))dt + f(x)$$

tiene solución.

Problema 4.5.5 *Probar que la siguiente ecuación integral tiene solución:*

$$\cos(s) = f(s) + \int_0^1 \frac{s + t}{1 + s + t + f^2(t)}dt.$$

Problema 4.5.6 *¿Podemos afirmar que el problema de valores iniciales*

$$\begin{cases} y'(t) = y^2(t) & t \in]-1, 2[\\ y(0) = 1 \end{cases}$$

tiene al menos una solución?.

Problema 4.5.7 *Sea C un subconjunto cerrado y convexo de un espacio de Banach $(X, \|.\|)$ tal que $0 \in C$. Probar que toda aplicación $T : C \to C$ continua y condensante verificando la condicin de Leray-Schauder sobre C tiene un punto fijo en C.*

4.5.1. Soluciones de ecuaciones integrales.

Sea $F : [0,1] \times [0,1] \times \mathbb{R} \to \mathbb{R}$ una función continua en su dominio y acotada.

Consideremos la ecuación integral

$$g(s) = f(s) + \int_0^1 F(s,t,f(t))dt,$$

donde la función g, es continua en [0,1], es un dato y donde f es la función incógnita, que buscamos en $\mathcal{C}([0,1])$.

Problema 4.5.8 *Aplicar el Teorema de Schauder para obtener una solución de la ecuación integral.*

Para obtener la existencia de una solución de la ecuación integral podríamos actuar del modo siguiente:

Consideremos el espacio de Banach $(\mathcal{C}([0,1]), \|.\|_\infty)$ que es donde buscamos la función f. Como F está acotada, sea $K > 0$ de forma que

$$|F(u,v,w)| \leq K \quad \forall (u,v,w) \in [0,1] \times [0,1] \times \mathbb{R}.$$

Dada $f \in \mathcal{C}([0,1]$ definimos $T(f) : [0,1] \to \mathbb{R}$ de la forma siguiente:

$$T(f)(s) := g(s) - \int_0^1 F(s,t,f(t))dt.$$

Problema 4.5.9 *Ver que un punto fijo de T sería una solución de nuestro problema.*

Como consecuencia del problema anterior, nos planteamos si el operador T está bien definido.

Problema 4.5.10 *Probar que para cada* $f \in \mathcal{C}([0,1])$ *se tiene que* $T(f) \in \mathcal{C}([0,1])$.

Problema 4.5.11 *Ver que para cada* $f \in \mathcal{C}([0,1])$ *se tiene que* $\|T(f)\|_\infty \leq \|g\|_\infty + K$.

Si llamamos $K' := \|g\|_\infty + K$, el problema anterior nos asegura que $T(U(0,K')) \subseteq U(0,K')$.

Problema 4.5.12 *Probar que* $T(U(0,K'))$ *es un conjunto equicontinuo. Para hacer esto tener en cuenta que* F *es uniformemente continua en* $[0,1]\times[0,1]\times[-K',K']$.

Problema 4.5.13 *Probar que* T *es un operador continuo.*

4.5.2. Aplicaciones admitiendo un centro.

Definición 4.5.14 *Sea* $(X,\|.\|)$ *un espacio de Banach y* C *un subconjunto no vacío de* X. *Una aplicación* $T : C \to X$ *diremos que tiene un centro en* $y_0 \in X$ *si se cumple:*

$$\|Tx - y_0\| \leq \|x - y_0\|, \quad \forall x \in C.$$

Problema 4.5.15 *Comprobar que si* y_0 *es un centro de* $T : C \to X$ *y además* $y_0 \in C$, *entonces* y_0 *es un punto fijo de* T.

Problema 4.5.16 *Probar que si* C *es un conjunto cerrado de un espacio de Banach* X, *y* $T : C \to C$ *es una aplicación contractiva (i.e. existe* $k \in (0,1)$ *tal que* $\|Tx - Ty\| \leq k\|x - y\|$*) entonces* T *tiene un centro en* C.

En vista de los problemas anteriores será interesante, desde el punto de vista de la existencia de puntos fijos, estudiar las aplicaciones que admiten un centro fuera de su dominio.

Problema 4.5.17 *Probar que la aplicación de Beals (ver ejemplo (4.3.1), admite como centro al* $(2,0,0,...)$.

Sea $(E,\|.\|)$ un espacio de Banach real de dimensión finita. Sea C un subconjunto cerrado no vacío de E. Consideremos ahora $y_0 \in E \setminus C$.

Problema 4.5.18 *Probar que existe* $c \in C$ *tal que* $\|c - y_0\| = \min\{\|x - y_0\| \ : \ x \in C\}$.

Problema 4.5.19 *Probar que si además* C *es convexo el conjunto* $P(y_0) := \{c \in C \ : \ \|c - y_0\| = \text{mín}\{\|x - y_0\| \ : \ x \in C\}$ *es compacto convexo.*

Problema 4.5.20 *Sea* $(E, \|.\|)$ *un espacio de Banach real de dimensión finita. Sea* C *un subconjunto cerrado y convexo no vacío de* E. *Si* $T : C \to C$ *es una aplicación continua que admite un centro* $y_0 \in E$. *Probar que* T *tiene un punto fijo.*

Capítulo 5

Aplicaciones Multivaluadas.

5.1. Introducción.

Las aplicaciones multivaluadas aparecen de modo natural cuando se consideran imágenes inversas de aplicaciones univaluadas no inyectivas. También para garantizar la existencia de soluciones de determinadas ecuaciones no lineales en derivadas parciales es necesario introducir aplicaciones multivaluadas.

La Teoría de Control es otro marco donde es frecuente el uso de aplicaciones multivaluadas. Por último añadiremos que la optimización también conduce al uso de aplicaciones multivaluadas.

En este capítulo mostraremos dos teoremas de punto fijo para aplicaciones multivaluadas que pueden considerarse como las extensiones naturales del teorema de Banach y del Teorema de Schauder

5.2. Consideraciones generales

Una aplicación entre dos conjuntos X, Y es toda regla por la cual a cada elemento $x \in X$ le hacemos corresponder uno y sólo uno del conjunto Y,(llamemosle $f(x) \in Y$).

El conjunto formado por todos los subconjuntos de un conjunto dado Y se denota usualmente por 2^Y.

Definición 5.2.1 *Sean X, Y dos conjuntos no vacios, toda aplicación entre X y 2^Y se llama aplicación multivaluada entre X e Y.*

Una aplicación multivaluada $F : X \to 2^Y$ asocia a cada elemento $x \in X$ un elemento $F(x) \in 2^Y$, es decir un subconjuto y sólo uno de Y.

Se entiende que si $A \neq \emptyset \subseteq X$, entonces

$$F(A) = \cup_{a \in A} F(a).$$

Si consideramos una aplicación multivaluada $F : X \to 2^Y$, como $\emptyset \in 2^Y$ nada impide formalmente que para algún $x \in X$ pueda ocurrir que $F(x) = \emptyset$. Algunas veces se excluye al conjunto vacío de los posibles valores que pueda tomar F. Sin embargo, hay casos en que esto no es lo adecuado:

Ejemplo 5.2.2 (operador inverso generalizado) *Sean X, Y conjuntos no vacíos $S : Y \to X$ una aplicación. Se llama operador solución T de la ecuación $S(y) = x$ al operador que asocia a cada $x \in X$ el conjunto $T(x) \in 2^Y$ de las soluciones de la ecuación anterior. Si no existe solución para algún x se escribe $T(x) = \emptyset$. Esta aplicación T se llama también aplicación inversa generalizada de S.*

Definición 5.2.3 *Sea la aplicación multivaluada $F : X \to 2^Y$. Diremos $x \in X$ es un punto fijo de F si se cumple que $x \in F(x)$.*

Dada la aplicación multivaluada $F : X \to 2^Y$, se llama gráfica de dicha aplicación al siguiente conjunto:

$$Gr(F) = \{(x, y) \in X \times Y \ : \ y \in F(x)\},$$

evidentemente x es un punto fijo de F si se cumple que $(x, x) \in Gr(F)$.

5.2.1. La métrica de Hausdorff.

Sea (M, d) un espacio métrico y denotamos por $\mathcal{M}$ a la familia de todos los subconjuntos cerrados acotados y no vacíos de M. Para cada $A \in \mathcal{M}$ y $\epsilon > 0$ se define el ϵ-entorno de A como el conjunto:

$$N_\epsilon(A) = \{x \in M \ : \ \mathrm{dist}(x, A) < \epsilon\},$$

donde como es habitual $\mathrm{dist}(x, A) = \inf\{d(x, y) \ : \ y \in A\}$. Ahora para $A, B \in \mathcal{M}$ se define :

$$H(A, B) := \inf\{\epsilon > 0 : A \subseteq N_\epsilon(B) \text{ y } B \subseteq N_\epsilon(A)\}.$$

Teorema 5.2.4 *$(\mathcal{M}, H)$ es un espacio métrico. (La métrica H se llama métrica de Hausdorff)*

Prueba.

Para ver que H es una métrica, primero observemos:

Si $x \in A$, entonces $\inf\{d(x,y) \ : y \in A\} = 0$.

Con lo cual $x \in N_\epsilon(A)$ para cada $\epsilon > 0$.

Por lo tanto $A \subseteq N_\epsilon(A)$ para cada $\epsilon > 0$, lo que nos asegura que $H(A,A) = 0$.

Por otra parte, si $x \in N_\epsilon(A)$ para cada $\epsilon > 0$, entonces para cada $n \in \mathbb{N}$ existirá $y_n \in A$ tal que $d(x, y_n) < \frac{1}{n}$. Lo que nos dice que $x \in \overline{A}$. Como por hipótesis A es cerrado se obtiene que $x \in A$. El hecho anterior nos permite afirmar que si $H(A,B) = 0$ entonces $A = B$. Puesto que es obvio desde la definición que para cada $A, B \in \mathcal{M}$, $H(A,B) = H(B,A)$ y $H(A,B) \geq 0$.

Luego, para ver que H es una métrica sólo tenemos que probar la desigualdad triangular.

Sean $A, B, C \in \mathcal{M}$, $\sigma = H(A,C)$, $\mu = H(C,B)$, y sea $\rho > 0$.

Como $A \subseteq N_{\sigma+\frac{\rho}{4}}(C)$, si $a \in A$ existirá $c \in C$ de forma que $d(a,c) \leq \sigma + \frac{\rho}{2}$.

Por otra parte, como $C \subseteq N_{\mu+\frac{\rho}{4}}(B)$, existirá $b \in B$ tal que $d(c,b) \leq \mu + \frac{\rho}{2}$.

Por lo tanto, para cada $a \in A$ existirá $b \in B$ de forma que

$$d(a,b) \leq d(a,c) + d(c,b) \leq \sigma + \mu + \rho.$$

La desigualdad anterior nos dice que $A \subseteq N_{\sigma+\mu+\rho}(B)$ para cada $\rho > 0$. Luego

$$\inf\{\epsilon > 0 : \ A \subseteq N_\epsilon(B)\} \leq \sigma + \mu.$$

Intercambiando los papeles de A y B en la argumentación anterior se obtiene:

$$\inf\{\epsilon > 0 : \ B \subseteq N_\epsilon(A)\} \leq \sigma + \mu.$$

Entonces,

$$H(A,B) \leq \sigma + \mu = H(A,C) + H(C,B).$$

Ahora veremos como es la convergencia sucesional mediante la métrica de Hausdorff.

Proposición 5.2.5 *Sea (M,d) un espacio métrico. Sea (A_n) una sucesion de elementos de $\mathcal{M}$ y $A \in \mathcal{M}$ de forma que $\lim_{n\to\infty} H(A_n, A) = 0$. Entonces*

$$A = \bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m} = \bigcap_{\epsilon>0} \bigcup_{n\geq 1} \bigcap_{m\geq n} N_\epsilon(A_m).$$

Prueba

Sea $\epsilon > 0$. Como por hipótesis $\lim_{n\to\infty} H(A_n, A) = 0$, existirá $n_0(\epsilon) \in \mathbb{N}$ de forma que si $m \geq n_0(\epsilon)$ se tendrá que

$$A \subseteq N_\epsilon(A_m) \text{ y } A_m \subseteq N_\epsilon(A).$$

De las inclusiones anteriores se obtine que para cada $\epsilon > 0$:

$$A \subseteq \bigcap_{m \geq n_0(\epsilon)} N_\epsilon(A_m) \subseteq \bigcup_{n\geq 1} \bigcap_{m\geq n} N_\epsilon(A_m)$$

luego se da la inclusión

$$A \subseteq \bigcap_{\epsilon>0} \bigcup_{n\geq 1} \bigcap_{m\geq n} N_\epsilon(A_m)$$

Veamos que se cumple la inclusión

$$\bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m} \subseteq A.$$

En efecto, como para cada $m \geq n_0(\epsilon)$ se verifica que $A_m \subseteq N_\epsilon(A)$ entonces

$$\overline{\bigcup_{m\geq n_0(\epsilon)} A_m} \subseteq N_\epsilon(A).$$

Por lo tanto, para cada $\epsilon > 0$ se cumple que

$$\bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m} \subseteq \overline{\bigcup_{m\geq n_0(\epsilon)} A_m} \subseteq N_\epsilon(A)$$

consecuentemente,

$$\bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m} \subseteq \bigcap_{\epsilon>0} N_\epsilon(A) = \overline{A} = A.$$

Hasta ahora hemos visto que

$$\bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m} \subseteq A \subseteq \bigcap_{\epsilon>0} \bigcup_{n\geq 1} \bigcap_{m\geq n} N_\epsilon(A_m)$$

Para terminar la demostración consideremos un

$$x \in \bigcap_{\epsilon>0} \bigcup_{n\geq 1} \bigcap_{m\geq n} N_\epsilon(A_m).$$

Entonces, dado $\epsilon > 0$, existirá $n_0(\epsilon) \in \mathbb{N}$ de forma que si $m \geq n_0(\epsilon)$ se tendrá $x \in N_\epsilon(A_m)$.

Consideremos $n \in \mathbb{N}$ fijo, y tomemos $m \geq \max\{n, n_0(\epsilon)\}$ está claro que $x \in N_\epsilon(A_m) \subseteq N_\epsilon(\bigcup_{m\geq n} A_m)$ con lo cual,

$$x \in \overline{\bigcup_{m\geq n} A_m}$$

y por lo tanto

$$x \in \bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m}$$

lo que acaba la prueba de la proposición.

Teorema 5.2.6 *Si* (M, d) *es un espacio métrico completo. Entonces* $(\mathcal{M}, H)$ *también lo es.*

Prueba.

Sea (A_n) una sucesión de Cauchy de elementos de $\mathcal{M}$. La proposición anterior identifica al único candidato posible para ser límite de la sucesión (A_n).

Llamemosle

$$A := \bigcap_{n\geq 1} \overline{\bigcup_{m\geq n} A_m}.$$

Ahora veamos que $A \in \mathcal{M}$ y que además $\lim_{n\to\infty} H(A_n, A) = 0$.

Claramente A es cerrado por ser intersección de conjuntos cerrados.

Para ver que A es acotado será suficiente ver que $\bigcup_{m\geq n} A_m$ es acotado para algun $n \in \mathbb{N}$.

Dado el 1, como la sucesión (A_n) es de Cauchy, se cumplirá que existirá $n_0 \in \mathbb{N}$ tal que $H(A_m, A_{n_0}) < 1$ siempre que $m \geq n_0$.

Esto significa que para cada $m \geq n_0$, $A_m \subseteq N_1(A_{n_0})$, luego

$$\bigcup_{m\geq n_0} A_m \subseteq N_1(A_{n_0})$$

con lo cual, dados $x, y \in \bigcup_{m\geq n_0} A_m$ se tiene que existen $x_0, x_1 \in A_{n_0}$ de forma que

$$d(x, y) \leq d(x, x_0) + d(x_0, x_1) + d(x_1, y_0) \leq 4 + diam(A_{n_0}).$$

Luego A es acotado.

Hasta aquí hemos visto que $A \in \mathcal{M}$.

Sea $\epsilon > 0$, como la sucesión es de Cauchy, para cada $k \in \mathbb{N}$ podremos encontrar $N_k \geq 1$ tal que

$$H(A_n, A_m) < \frac{\epsilon}{2^{k+1}} \quad \forall n, m \geq N_k.$$

Tomemos $n_0 \geq N_0$ y $x_0 \in A_{n_0}$. Entonces elegimos $n_1 > \max\{n_0, N_1\}$ y $x_1 \in A_{n_1}$ de tal modo que $d(x_0, x_1) < \frac{\epsilon}{2}$ (esto es consecuencia de que $d(x_0, A_{n_1}) \leq H(A_{n_0}, A_{n_1}) < \frac{\epsilon}{2}$). Entonces si (n_k) es una sucesión estrictamente creciente de naturales con $n_k \geq N_k$ podemos construir inductivamente una sucesión (x_k) de forma que $x_k \in A_{n_k}$ y cumpliendo que $d(x_k, x_{k+1}) \leq \frac{\epsilon}{2^{k+1}}$. Esto nos dice que la sucesión (x_k) es de Cauchy en M y como (M, d) es completo, entonces existirá $x_0 \in M$ tal que $x_k \to y_0$.

Por otra parte, como (n_k) es estrictamente creciente, dado $n \geq 1$ existirá $k_n \geq 1$ tal que $n_{k_n} \geq n$. Entonces

$$x_k \in \bigcup_{m \geq n} A_m \quad \forall k \geq k_n$$

y así

$$y_0 \in \overline{\bigcup_{m \geq n} A_m} \quad \forall n \geq 1.$$

Luego hemos visto que $y_0 \in A$, lo que demuestra que $A \neq \emptyset$. Además,

$$d(y_0, x_0) = \lim_{n \to \infty} d(x_n, x_0) \leq \lim_{n \to \infty} \sum_{k=0}^{n-1} d(x_k, x_{k+1}) \leq \epsilon.$$

Para todo $n_0 \geq N_0$ y para todo $x_0 \in A_{n_0}$ se tiene que existe $x \in A$ tal que $d(x, x_0) \leq \epsilon$. Entonces $A_{n_0} \subseteq N_\epsilon(A)$. Lo único que nos queda por ver es que $A \subseteq N_\epsilon(A_n)$ siempre que $n \geq n_0$.

Sea $x \in A$, entonces $x \in \overline{\bigcup_{m \geq N_0} A_m}$, luego existe $m \geq N_0$ y existe $y \in A_m$ de forma que $d(x, y) < \frac{\epsilon}{2}$. También, si $n \geq N_0$ tenemos que $d(x, A_n) \leq d(x, A_m) + H(A_m, A_n) < \epsilon$. Lo que implica que $A \subseteq N_\epsilon(A_n)$ como queríamos demostrar.

5.2.2. Conjuntos autosimilares

Sea (M, d) un espacio métrico completo.

Problema 5.2.7 *Sea $\mathcal{MC}$ la familia de subconjuntos compactos no vacíos de M. Probar que $(\mathcal{MC}, H)$ es un subespacio métrico cerrado de $(\mathcal{M}, H)$.*

Como es habitual, entenderemos por aplicación k-contractiva en (M, d) a toda aplicación $T : M \to M$ que cumple

$$d(Tx, Ty) \leq kd(x, y) \quad \forall x, y \in M, \ k \in [0, 1[.$$

Problema 5.2.8 *Dado $i \in \{1, 2, ..., n\}$ sea $T_i : M \to M$ una aplicción k_i-contractiva. Probar que si $C \in \mathcal{MC}$, entonces el conjunto*

$$\bigcup_{i=1}^{n} T_i(C)$$

es también un elemento de $\mathcal{MC}$.

El problema anterior nos dice que aplicación $F : \mathcal{MC} \to \mathcal{MC}$ dada por

$$F(C) = \bigcup_{i=1}^{n} T_i(C)$$

está bien definida.

Problema 5.2.9 *Probar que la aplicación F anterior es una k-contracción en $(\mathcal{MC}, H)$, con $k := \text{máx}\{k_i \ : \ i = 1, 2, ...n\}$*

Problema 5.2.10 *Sean $T_1, ..., T_n$ una familia finita de k_i-contracciones en un espacio métrico completo (M, d). Probar que existe un único compacto no vacío K de M de forma que*

$$K = \bigcup_{i=1}^{n} T_i(K).$$

En particular, si (M, d) es el espacio euclídeo $\mathbb{R}^p$ y las aplicaciones T_i se definen por

$$T_i(x) = k_i x + b_i$$

donde b_i es un vector prefijado de $\mathbb{R}^p$ y $0 < k_i < 1$, $(i = 1, 2, ..., n)$, entonces el único conjunto compacto de $\mathbb{R}^p$ que cumple la tesis del problema anterior se llama *conjunto autosimilar* respecto a las aplicaciones $T_1, ..., T_n$. Estos conjuntos están íntimamente relacionados con la teoría de fractales.

Problema 5.2.11 *Consideremos la recta real* $\mathbb{R}$ *como espacio métrico completo, si consideramos las aplicaciones*

$$T_1(x) = \frac{1}{3}x, \quad T_2(x) = \frac{1}{3}x + \frac{2}{3}$$

Comprobar que el conjunto autosimilar respecto a T_1, T_2 *es el conjunto ternario de Cantor.*

5.3. Contracciones multivaluadas

La definición de continuidad para aplicaciones multivaluadas $F : X \to 2^Y$ requiere disponer de topologías, tanto X como en 2^Y.

Cuando, en particular (X, d), (Y, d') son espacios métricos y $F : X \to \mathcal{M}(Y)$, donde $\mathcal{M}(Y)$ es la colección de los subconjuntos cerrados y acotados de Y, en dicho conjunto podemos considerar la métrica de Hausdorff. En este caso, decimos que F es continua si lo es como aplicación del espacio métrico (X, d) al espacio métrico $(\mathcal{M}(Y), H)$.

En 1969 S. Nadler observó que el principio de contracción de Banach se puede extender al caso multivaluado. La idea clave para lograrlo es la siguiente: Si A y B son dos conjuntos no vacíos cerrados y acotados de un espacio métrico y si $x \in A$, entonces dado $\epsilon > 0$ debe existir $y \in B$ tal que

$$d(x, y) \leq H(A, B) + \epsilon,$$

esto se da porque la definición de distancia de Hausdorff asegura que para cada $\mu > 0$

$$A \subseteq N_{\rho+\mu}(B),$$

donde $\rho = H(A, B)$.

Teorema 5.3.1 *Sea* (M, d) *un espacio métrico completo y sea* $\mathcal{M}$ *la colección de todos los subconjuntos de* M *no vacíos acotados y cerrados. Supongamos que* $T : M \to \mathcal{M}$ *es una contracción en el sentido de que existe* $k \in (0, 1)$ *tal que*

$$H(T(x), T(y)) \leq kd(x, y), \quad \forall x, y \in M.$$

Entonces T *tiene un punto fijo en* M.

Prueba.

Tomemos $x_0 \in M$ y $x_1 \in T(x_0)$. Por la obsevación que hemos hecho al principio de la sección debe existir $x_2 \in T(x_1)$ tal que

$$d(x_1, x_2) \leq H(T(x_0), T(x_1)) + k.$$

Similarmente, exisistirá $x_3 \in T(x_2)$ de forma que

$$d(x_2, x_3) \leq H(T(x_1), T(x_2)) + k^2.$$

Procediendo por inducción obtenemos una sucesión (x_n) de elementos de M con la propiedad de que si $i \in \mathbb{N}$, $x_{i+1} \in T(x_i)$ y además,

$$d(x_i, x_{i+1}) \leq H(T(x_{i-1}), T(x_i)) + k^i \leq kd(x_{i-1}, x_i) + k^i \leq$$

$$k[H(T(x_{i-2}), T(x_{i-1})) + k^{i-1}] + k^i \leq$$

$$k^2 d(x_{i-2}, x_{i-1}) + 2k^i \leq ... \leq k^i d(x_0, x_1) + ik^i.$$

Por lo tanto,

$$\sum_{i=0}^{\infty} d(x_i, x_{i+1}) \leq d(x_0, x_1) \left(\sum_{i=0}^{\infty} k^i \right) + \sum_{i=0}^{\infty} ik^i < \infty.$$

Lo que demuestra que la sucesión (x_n) es de Cauchy, como M es completo existirá $x \in M$ de forma que $x_n \to x$. También, como T es continua $H(T(x_n), T(x)) \to 0$. Como $x_n \in T(x_{n-1})$,

$$\lim_{n \to \infty} \operatorname{dist}(x_n, T(x)) = \lim_{n \to \infty} \inf\{d(x_n, y) : \ y \in T(x)\} = 0.$$

Esto implica que

$$\operatorname{dist}(x, T(x)) = \inf\{d(x, y) : \ y \in T(x)\} = 0,$$

y como $T(x)$ es un conjunto cerrado se debe cumplir que $x \in T(x)$.

Observación 5.3.2 *Si comparamos el resultado que hemos obtenido con el principio de contracción podemos destacar: Primero, en contraste con el caso univaluado aquí no obtenemos la unicidad del punto fijo. Segundo, como M es completo, $(\mathcal{M}, H)$ también es completo, pero este hecho no se utiliza en la prueba.*

5.4. El teorema de Kakutani.

Definición 5.4.1 *Si X, Y son espacios métricos, la aplicación $F : X \to 2^Y \setminus \{\emptyset\}$ es* **cerrada** *en $x \in X$ si se cumple que para cada sucesión (x_n) con $x_n \to x$ y para cada sucesión (y_n) con $y_n \in F(x_n)$ cumpliendo que $y_n \to y$ se tiene que $y \in F(x)$.*

Se dice que F es cerrada en X si lo es en cada punto de X.

Proposición 5.4.2 *Si la aplicación $F : X \to 2^Y \setminus \{\emptyset\}$ es cerrada en $x \in X$ entonces $F(x)$ es un conjunto cerrado de Y.*

Prueba

Si (y_n) es una sucesión de elementos de $F(x)$ de forma que $y_n \to y$, entonces como $y_n \in F(x)$ para cada $n \in \mathbb{N}$, consideremos la sucesión constante $x_n = x$ para todo $n \in \mathbb{N}$, la cual evidentemente converge a x, ahora la definición de aplicación cerrada nos dice que $y \in F(x)$.

Definición 5.4.3 *La aplicación $F : X \to 2^Y \setminus \{\emptyset\}$se llama* **semicontinua superiormente** *en $x \in X$ si para todo conjunto abierto U de Y tal que $F(x) \subset U$ se tiene que existe un abierto V de X con $x \in V$ y cumpliendo que $F(V) \subseteq U$.*

La aplicación F se llama semicontinua superiormente en X si lo es en cada punto de X.

Ejemplo 5.4.4 *la aplicación $F : [0, \infty[\to 2^{\mathbb{R}}$ dada por*

$$F(x) = \begin{cases} [0, \frac{1}{x}] & x > 0 \\ \{0\} & x = 0 \end{cases}$$

no es cerrada en $x = 0$.

En efecto, la sucesión $(x_n = \frac{1}{n})$ converge a 0, pero para cada entero positivo n, $y_n = 1 \in F(x_n)$ y sin embargo $y_n \to 1 \notin F(0)$.

Ejemplo 5.4.5 *La aplicación $F : [0, \infty[\to 2^{[0,+\infty[}$ dada por*

$$F(x) = \begin{cases} [0, x] \cup \{\frac{1}{x}\} & x > 0 \\ \{0\} & x = 0 \end{cases}$$

es cerrada en $x = 0$, pero no es semicontinua superiormente en $x = 0$.

En efecto, si (x_n) es una sucesión convergente a cero, y si (y_n) es una sucesión convergente a cierto y, con $y_n \in F(x_n)$ para cada n, necesariamente (y_n) debe ser acotada, por lo que

$$0 \le y_n \le x_n$$

de donde se sigue que $y_n \to 0$.

Por otra parte, $U = [0, 1)$ es un abierto relativo que contiene a $F(0) = \{0\}$, pero si V es cualquier entorno de 0, contendrá puntos $0 < x < 1$, y por tanto $\frac{1}{x} \in F(x) \subseteq F(V)$ no puede pertenecer a U, pues $\frac{1}{x} > 1$.

Problema 5.4.6 *Estudiar si la aplicación* $F : [0, \infty[\to 2^{\mathbb{R}}$ *definida por*

$$F(x) = \begin{cases} [0, x] & x > 0 \\ \{0\} & x = 0 \end{cases}$$

es semicontinua superiormente y/o cerrada en el 0.

Problema 5.4.7 *Estudiar si la aplicación* $F : \mathbb{R} \to 2^{\mathbb{R}}$ *definida por*

$$F(x) = \begin{cases} [0, x] & x > 0 \\ \{0\} & x = 0 \\ \{\frac{1}{x}\} & x < 0 \end{cases}$$

es semicontinua superiormente y/o cerrada en el 0.

Como hemos visto en el ejemplo anterior, una aplicación cerrada no necesariamente es semicontinua superiormente. Sin embargo, el siguiente resultado muestra que si el espacio de llegada es compacto, entonces la condición de ser cerrada es suficiente para garantizar la semicontinuidad superior.

Lema 5.4.8 *Si* X, Y *son espacios métricos e* Y *es compacto, entonces cualquier aplicación* $F : X \to 2^Y \setminus \{\emptyset\}$ *cerrada en* X, *es semicontinua superiormente en* X.

Prueba.

Sea $x \in X$ y U un subconjunto abierto de Y con $F(x) \subseteq U$.

Tenemos que demostrar la existencia de un subconjunto abierto V de X que contiene a x y de forma que $F(V) \subseteq U$.

Consideremos el conjunto:

$$V := \{z \in X : \ F(z) \subseteq U\}.$$

Es claro que $x \in V$ y además $F(V) = \bigcup_{a \in V} F(a) \subseteq U$. Habremos probado el resultado si vemos que V es abierto o lo que es lo mismo, si vemos que $X \setminus V$ es cerrado.

Observemos que

$$X \setminus V = \{z \in X : \; F(z) \cap (Y \setminus U) \neq \emptyset\}.$$

Tomemos una sucesión (z_n) con $z_n \in X \setminus V$ y con $z_n \to z$, y veamos que $z \in X \setminus V$, lo que demostrará que $X \setminus V$ es cerrado.

Como $F(z_n) \cap (Y \setminus U) \neq \emptyset$, podemos encontrar para cada $n \in \mathbb{N}$ un elemento $y_n \in F(z_n) \cap (Y \setminus U)$.

El hecho de ser Y compacto, nos dice que también lo es $Y \setminus U$, y entonces podemos encontrar una subsucesión (y_{n_k}) con $y_{n_k} \to y \in Y \setminus U$.

Como F es cerrada, las condiciones

$$z_{n_k} \to z, \quad y_{n_k} \to y, \quad y_{n_k} \in F(z_{n_k})$$

aseguran que $y \in F(z)$.

Por lo tanto $y \in F(z) \cap (Y \setminus U)$. Si $F(z) \cap (Y \setminus U)$ es no vacío entonces $z \in X \setminus V$, y la demostración está completa.

El teorema de Kakutani, que es una generalización del teorema de Brouwer al caso multivaluado, fue publicado por Kakutani en 1941. A continuación veremos la versión de dicho teorema en espacios de Banach generales, es decir la extensión del teorema de Schauder que obtuvieron Bohnenblust y Karlin en 1950.

Teorema 5.4.9 *Sea X un espacio de Banach y sea K un compacto convexo no vacío de X.*

Sea $F : K \to 2^K$ una aplicación verificando

(a) Para cada $x \in K$ el conjunto $F(x)$ es convexo y no vacío.

(b) F es cerrada (en K).

Entonces F tiene un punto fijo.

Prueba. Como K es compacto, entonces será totalmente acotado (ver teorema 4.3.3) para todo $\epsilon > 0$ existe un conjunto finito

$$F_\epsilon := \{x_{\epsilon 1}, ..., x_{\epsilon n(\epsilon)}\} \subseteq K$$

de forma que para cada $x \in K$ se cumple $d(x, F_\epsilon) < \epsilon$.

Llamemos para cada $x \in K$

$$\varphi_{\epsilon i}(x) := \text{máx}\{\epsilon - \|x - x_{\epsilon i}\|, 0\}$$

para $i = 1, 2, ..., n(\epsilon)$.

Cada una de estas funciones $\varphi_{\epsilon i}$ es continua y positiva sobre K.

Como para todo $x \in K$ se tiene que $\|x - x_{\epsilon i}\| < \epsilon$ para, al menos, un $i \in \{1, 2, ..., n(\epsilon)\}$ las funciones siguientes están bien definidas:

$$w_{\epsilon i}(x) := \frac{\varphi_{\epsilon i}(x)}{\sum_{j=1}^{n(\epsilon)} \varphi_{\epsilon j}(x)}$$

Fijemos un punto cualquiera, $y_{\epsilon i} \in F(x_{\epsilon i})$, $(i = 1, ..., n(\epsilon))$ y definamos

$$f_\epsilon(x) := \sum_{i=1}^{n(\epsilon)} w_{\epsilon i}(x) y_{\epsilon i}.$$

Como $y_{\epsilon i} \in K$, $w_{\epsilon i}(x) \geq 0$ y $\sum_{i=1}^{n(\epsilon)} w_{\epsilon i} = 1$ vemos que cada $f_\epsilon(x)$ es una combinación convexa de elementos de K, que por hipótesis es convexo, luego $f_\epsilon(x) \in K$.

Además es claro que f_ϵ es una función continua.

Por el teorema de Schauder f_ϵ tiene un punto fijo en K, es decir, existe $x_\epsilon \in K$ tal que $f_\epsilon(x_\epsilon) = x_\epsilon$.

Aplicando la compacidad de K y seleccionando una sucesión (ϵ_n) de números positivos de forma que

1) $\epsilon_n \to 0$.

2) $x_{\epsilon_n} \to x^* \in K$.

3) $f_{\epsilon_n}(x_{\epsilon_n}) = x_{\epsilon_n}$.

Probaremos que x^* es un punto fijo de F.

Llamemos

$$U_\delta := F(x^*) + B(0, \delta).$$

Como por hipótesis $F(x^*)$ es un conjunto convexo, es trivial ver que U_δ es un abierto convexo con $F(x^*) \subset U_\delta$.

Como K es compacto, el lema anterior nos asegura que F es semicontinua superiormente en x^*, es decir que dado el abierto U_δ existe $\epsilon > 0$ tal que si

$$V_\epsilon := \{x \in K : \ \|x - x^*\| < \epsilon\}$$

entonces $F(V_\epsilon) \subset U_\delta$.

Como $x_{\epsilon_n} \to x^*$, y $\epsilon_n \to 0$ existe un entero positivo N de forma que si $n \geq N$ entonces

$$0 < \epsilon_n < \frac{\epsilon}{2}, \ \|x_{\epsilon_n} - x^*\| < \frac{\epsilon}{2}.$$

Los números reales $w_{\epsilon_n i}(x_{\epsilon_n})$ pueden ser nulos o estrictamente positivos. Si sucede que $w_{\epsilon_n i}(x_{\epsilon_n}) > 0$, entonces

$$\epsilon_n - \|x_{\epsilon_n} - x_{\epsilon_n i}\| > 0$$

es decir

$$\|x_{\epsilon_n} - x_{\epsilon_n i}\| < \epsilon_n < \frac{\epsilon}{2}$$

luego

$$\|x^* - x_{\epsilon_n i}\| \leq \|x_{\epsilon_n} - x_{\epsilon_n i}\| + \|x_{\epsilon_n} - x^*\| < \epsilon_n + \frac{\epsilon}{2} < \epsilon$$

Por lo tanto, siempre que $n > N$ y siempre que $w_{\epsilon_n i}(x_{\epsilon_n}) > 0$, $x_{\epsilon_n i} \in V_\epsilon$, con lo cual

$$y_{\epsilon_n i} \in F(x_{\epsilon_n i}) \subset F(V_\epsilon) \subset U_\delta.$$

Por otra parte,

$$x_{\epsilon_n} = f_{\epsilon_n}(x_{\epsilon_n}) := \sum_{i=1}^{n(\epsilon_n)} w_{\epsilon_n i}(x_{\epsilon_n}) y_{\epsilon_n i}.$$

Observando el último término de esta igualdad vemos que x_{ϵ_n} es una combinación convexa de los $y_{\epsilon_n i}$, pero con la particularidad de que, para $n \geq N$, lo es sólamente de puntos $y_{\epsilon_n i}$ que están en U_δ, (puesto que sólo es combinación de los $y_{\epsilon_n i}$ que vayan multiplicados por un coeficiente $w_{\epsilon_n i}(x_{\epsilon_n})$ estrictamente positivo, en cuyo caso hemos visto que $y_{\epsilon_n i} \in U_\delta$).

Como U_δ es convexo, $x_{\epsilon_n} \in U_\delta$ para cada $n \geq N$, y por lo tanto, tomando límites se tendrá que $x^* \in \overline{U_\delta} \subset U_{2\delta}$.

Este razonamiento es válido para todo $\delta > 0$, pero un resultado bien conocido de espacios métricos nos dice que si $F(x^*)$ es un conjunto cerrado (ya que F es cerrada)

$$F(x^*) = \bigcap_{\delta > 0} (F(x^*) + B(0, \delta)) = \bigcap_{\delta > 0} U_\delta.$$

Luego x^* pertenece a $F(x^*)$ como queríamos demostrar.

5.5. Introducción a la Teoría de Juegos.

Juegos de dos personas: Se parte del supuesto de que un jugador I tiene un conjunto de posibles decisiones o estrategias E, y que otro jugador II tiene a su vez el conjunto F de decisiones (supondremos que ambos conjuntos son no vacíos).

La tarea del jugador I es elegir de algún modo un elemento $x \in E$, y la del jugador II hacer lo mismo con un elemento de $y \in F$. Todo par $(x, y) \in E \times F$ se suele llamar biestrategia, coloquialmente jugada.

Una regla de decisión para I es una aplicación (generalmente multivaluada) $T_I : F \to 2^E$ que asocia a cada decisión $y \in F$ el conjunto $T_I(y)$ de las decisiones posibles que puede tomar I. Naturalmente $T_I(y)$ puede depender de las reglas del juego. Hay juegos en que en ocasiones $T_I(y)$ es un sólo elemento.

Análogamente se define $T_{II} : E \to 2^F$.

Una vez fijadas T_I y T_{II}, las biestrategias consistentes serán aquellas $(x, y) \in E \times F$ de forma que

$$x \in T_I(y), \; y \in T_{II}(x).$$

Observemos que el problema de encontrar biestrategias consistentes es un problema de punto fijo, en el siguiente sentido:

Si definimos $\mathcal{T} : E \times F \to 2^{E \times F}$ de forma que $\mathcal{T}(x, y) := T_I(y) \times T_{II}(x)$. Entonces (x, y) es una biestrategia consistente si y sólo si, es un punto fijo de $\mathcal{T}$.

Tradicionalmente la teoría de juegos asume que un elemento esencial definidor de todo juego es una *función ganancia*, que es una aplicación $G : E \times F \to \mathbb{R}^2$, que asocia a cada biestrategia (x, y) un par de números reales $G(x, y) = (G_I(x, y), G_{II}(x, y))$, de forma que $G_I(x, y)$ es la ganancia del jugador I supuesto que se haya producido la bielección (x, y), y $G_{II}(x, y)$ es la ganancia del jugador II en las mismas circunstancias.

Establecida la función de ganancia hay unas reglas de decisión que se denominan canónicas:

Cuando el jugador I elige una estrategia $x \in E$ sin saber nada de la estrategia del jugador II, debe admitir que II eligió la mejor estrategia, es decir la que minimizará la ganancia de I, es decir $\inf_{y \in F} G_I(x, y)$. Por lo tanto, el jugador I debe buscar una estrategia $x_0 \in E$ de forma que maximice sus ganancias y al mismo tiempo minimice las ganancias del contrario:

$$\max_{x \in E} \inf_{y \in F} G_I(x, y) = \inf_{y \in F} G_I(x_0, y).$$

5.5.1. Caso de juegos de suma nula.

Un juego es de suma nula si cada jugador gana lo que el otro pierde, es decir si $-G_I(x, y) = G_{II}(x, y)$.

En este caso, la estrategia para el jugador I será encontrar $x_0 \in E$ de forma que

$$\max_{x\in E} \inf_{y\in F} G_I(x,y) = \inf_{y\in F} G_I(x_0,y). \tag{5.1}$$

Por la otra parte, como el jugador II tiene una ganancia igual a $-G_I(x,y)$ se deduce que debe buscar $y_0 \in F$ tal que

$$\max_{y\in F} \inf_{x\in E} [-G_I(x,y)] = \inf_{x\in E} \{-G_I(x,y_0)\}, \tag{5.2}$$

o lo que es lo mismo,

$$\min_{y\in F} \sup_{x\in E} G_I(x,y) = \sup_{x\in E} G_I(x,y_0).$$

La comparación de las igualdades (5.1) y (5.2) anteriores muestra que una condición necesaria y suficiente para que existan estrategias óptimas (x_0,y_0) es que se cumpla que para cualesquiera $x \in E$, $y \in F$

$$G_I(x,y_0) \le G_I(x_0,y_0) \le G_I(x_0,y). \tag{5.3}$$

Puesto que en este caso se cumple que

$$\sup_{x\in E} G_I(x,y_0) = G_I(x_0,y_0) = \inf_{y\in F} G_I(x_0,y).$$

Luego para obtener una biestrategia óptima es suficiente ver que

$$\min_{y\in F} \sup_{x\in E} G_I(x,y) = \max_{x\in E} \inf_{y\in F} [G_I(x,y)].$$

Todo par de estrategias (x_0,y_0) verificando (5.3) se llama solución del juego, mientras que cada una de las estrategias x_0 y_0 se llama a veces estrategia optimal.

Ejemplo 5.5.1 (Juego de la moneda) *Los jugadores I y II lanzan simultáneamente una moneda. Acuerdan que I gana un euro si las dos monedas muestran el mismo resultado. En otro caso gana II la misma cantidad.*

En este caso $E = F = \{C,F\}$y la función ganancia $K = G_I$ es $K(C,C) = K(F,F) = 1$, $K(C,F) = K(F,C) = -1$.

Aquí tenemos que

$$\min_{y\in F} \sup_{x\in E} K(x,y) = \min_{y\in F} \max\{K(C,y),K(F,y)\} = 1,$$

mientras que

$$\max_{x\in E}\inf_{y\in F} K(x,y) = \max_{x\in E}\min\{K(x,C),K(x,F)\} = -1$$

Por lo tanto K no tiene puntos de ensilladura (ver abajo lo que significa) y no existe una estrategia óptima.

Problema 5.5.2 *Un juego consiste en que el jugador I esconde, a su elección, en su puño bien una o bien dos monedas iguales. El jugador II ha de acertar cuántas monedas ha escogido el jugador I. Si II acierta gana las monedas. Si II no acierta paga (en todos los casos) una moneda.*

Estudiar si este juego tiene una estrategia óptima y si alguno de los dos jugadores tiene una estrategia optimal.

5.6. El teorema de min-max de von Neumann.

Definición 5.6.1 *Dada una función $L : A\times B \to \mathbb{R}$ donde A y B son conjuntos arbitrarios no vacíos, el punto $(x_0,y_0)\in A\times B$ se llama punto de ensilladura de L con respecto a $A\times B$ si, y sólo si,*

$$L(x_0,p) \le L(x_0,y_0) \le L(x,y_0)$$

para cualesquiera $x\in A,\ p\in B$.

Ejemplo 5.6.2 *El punto $(1,0)$ es un punto de ensilladura de la función $K:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ dada por*

$$K(x,y) = (x-1)^2 - y^2.$$

Efectivamente, $K(1,y)\le K(1,0)\le K(x,0)$.

Sean A y B conjuntos arbitrarios no vacíos. Sea $L : A\times B\to\mathbb{R}$ una función. Consideramos los dos problemas siguientes:

Problema primal (P): Determinar $u_0\in A$ tal que

$$\inf_{u\in A}\left(\sup_{p\in B} L(u,p)\right) = \sup_{p\in B} L(u_0,p).$$

Problema dual (P)*: Determinar $p_0\in B$ tal que

$$\sup_{p\in B}\left(\inf_{u\in A} L(u,p)\right) = \inf_{u\in A} L(u,p_0).$$

Lema 5.6.3 *Sean A, B conjuntos arbitrarios no vacíos. Sea $L : A \times B \to \mathbb{R}$ una función. Si llamamos:*

$$\alpha := \inf_{u \in A} \left(\sup_{p \in B} L(u,p) \right), \quad \beta := \sup_{p \in B} \left(\inf_{u \in A} L(u,p) \right)$$

$$F(u) := \sup_{p \in B} L(u,p), \quad G(p) := \inf_{u \in A} L(u,p).$$

Entonces se cumple:

i) $-\infty \leq \beta \leq \alpha \leq \infty$.

ii) Dados $u \in A$ y $p \in B$ se tiene que $G(p) \leq \beta \leq \alpha \leq F(u)$.

iii) Si existe dos puntos $u_0 \in A$ y $p_0 \in B$ de forma que $G(p_0) \geq F(u_0)$ entonces u_0 es solución del problema primal (P), p_0 es solución del correspondiente problema dual (P^*) y además $\alpha = \beta$.

Prueba.

Para todo $v \in A$,

$$\inf_{u \in A} L(u,p) \leq L(v,p)$$

de donde se tiene que

$$\beta = \sup_{p \in B} \left(\inf_{u \in A} L(u,p) \right) \leq \sup_{p \in B} L(v,p)$$

Ahora tomando ínfimos al variar $v \in A$, nos queda:

$$\beta \leq \inf_{v \in A} \left(\sup_{p \in B} L(v,p) \right) = \alpha,$$

lo que demuestra (i).

De (i) se sigue inmediatamente que

$$G(p) \leq \sup_{p \in B} G(p) = \beta \leq \alpha = \inf_{u \in A} F(u) \leq F(u)$$

lo que nos permite concluir (ii).

De la desigualdad anterior, se tiene que

$$G(p_0) \leq \beta \leq \alpha \leq F(u_0)$$

y si se da la desigualdad $G(p_0) \geq F(u_0)$ se concluye que

$$G(p_0) = \beta = \alpha = F(u_0)$$

obteniendo (iii).

Teorema 5.6.4 *Sean A, B conjuntos arbitrarios no vacíos. Sea $L : A\times B \to \mathbb{R}$ una función. Los siguientes estamentos son equivalentes*

(i) (u_0, p_0) es un punto de ensilladura de $A \times B$.

(ii) u_0 es una solución del problema (P), p_0 es una solución del problema dual (P^*), y además

$$\alpha = \inf_{u\in A}\left(\sup_{p\in B} L(u,p)\right) = L(u_0,p_0) = \sup_{p\in B}\left(\inf_{u\in A} L(u,p)\right) = \beta.$$

Prueba.

Veamos que (i) implica (ii).

Como (u_0, p_0) es un punto de ensilladura de la aplicación L se tiene que

$$L(u_0,p) \leq L(u_0,p_0) \leq L(u,p_0)$$

y entonces, para cualesquiera $u \in A$, $p \in B$

$$F(u_0) = \sup_{p\in B} L(u_0,p) \leq L(u_0,p_0) \leq \inf_{u\in A} L(u,p_0) = G(p_0)$$

Así que, por el apartado (ii) del lema anterior se obtiene que $F(u_0) = L(u_0,p_0) = G(p_0)$. Ahora, aplicando el apartado (iii) del lema anterior se obtiene esta implicación.

Veamos que (ii) implica (i).

El enunciado (ii) dice que

$$\inf_{u\in A}\left(\sup_{p\in B} L(u,p)\right) = \sup_{p\in B}\left(\inf_{u\in A} L(u,p)\right) \tag{5.4}$$

y también que u_0 y p_0 son soluciones de los problemas (P) y (P^*)) respectivamente, luego verifican:

$$\begin{cases} \inf_{u\in A}\left(\sup_{p\in B} L(u,p)\right) = \sup_{p\in B} L(u_0,p) \\ \sup_{p\in B}\left(\inf_{u\in A} L(u,p)\right) = \inf_{u\in A} L(u,p_0) \end{cases} \tag{5.5}$$

Por lo tanto, como

$$\inf_{u\in A} L(u,p_0) \leq L(u_0,p_0) \leq \sup_{p\in B} L(u_0,p)$$

de (5.4) y (5.5) se sigue la igualdad

$$\inf_{u\in A} L(u,p_0) = L(u_0,p_0) = \sup_{p\in B} L(u_0,p)$$

Luego para cada $u \in A$, y $p \in B$ se tiene

$$L(u_0, p) \leq \sup_{p \in B} L(u_0, p) = L(u_0, p_0) = \inf_{u \in A} L(u, p_0) \leq L(u, p_0)$$

viendo de esta forma que (u_0, p_0) es un punto de ensilladura.

Para probar la existencia de un punto de ensilladura en condiciones bastante generales estudiaremos el teorema de mini-max (probado por John von Neumann para el caso de $\mathbb{R}^n$ en 1928). Para su demostración utilizaremos el teorema de Kakutani.

Lema 5.6.5 *Sean X, Y dos espacios métricos compactos y sea $K : X \times Y \to \mathbb{R}$ una función continua. Entonces las funciones*

$$\varphi(x) := \max_{y \in Y} K(x, y)$$

y

$$\phi(y) = \min_{x \in X} K(x, y)$$

están bien definidas y son continuas sobre X e Y respectivamente.

Prueba.

Sea (x_n) una sucesión en X con $x_n \to x_0 \in X$. Como K es continua, las funciones $S_n(y) = K(x_n, y)$ son continuas en Y, luego existe $y_n \in Y$ de forma que y_n es un máximo de S_n (Y es compacto) o sea que para $n = 0, 1, 2, ...$ se tiene:

$$S_n(y_n) = K(x_n, y_n) = \max_{y \in Y} K(x_n, y) = \varphi(x_n).$$

Sea $y \in Y$ arbitrario

$$K(x_n, y) \leq \max_{y \in Y} K(x_n, y) = \varphi(x_n) = K(x_n, y_n).$$

Teniendo en cuenta que K es continua se tiene

$$K(x_0, y) \leq \liminf_{n \to \infty} K(x_n, y_n) = \liminf_{n \to \infty} \varphi(x_n).$$

Siendo y arbitrario,

$$\varphi(x_0) = \max_{y \in Y} K(x_0, y) \leq \liminf_{n \to \infty} K(x_n, y_n).$$

Por otra parte, como Y es compacto, existirá una subsucesión (y_{n_p}) tal que $y_{n_p} \to y^* \in Y$ de donde, por ser K continua,

$$\varphi(x_0) \geq K(x_0, y^*) = \lim_{p\to\infty} K(x_{n_p}, y_{n_p}) \geq \liminf_{n\to\infty} K(x_n, y_n).$$

De esta forma hemos visto que

$$\varphi(x_0) = \liminf_{n\to\infty} \varphi(x_n).$$

Si sucediera que $\limsup_n \varphi(x_n) = \lambda > \liminf_n \varphi(x_n) = \varphi(x_0)$, existiría una subsucesión

$$\varphi(x_{n_j}) \to \lambda$$

pero al ser $x_{n_j} \to x_0$ por la parte ya demostrada se debería cumplir

$$\varphi(x_0) = \liminf_{j\to\infty} \varphi(x_{n_j}) = \lambda$$

lo cual es una contradicción.

Así pues

$$\limsup_n \varphi(x_n) = \liminf_n \varphi(x_n) = \varphi(x_0).$$

Lo que demuestra la continuidad de φ.

Para probar la continuidad de ϕ se razona de forma similar.

Teorema 5.6.6 *Sea M un subconjunto compacto y convexo de un espacio de Banach X. Sea N un subconjunto compacto y convexo de otro espacio de Banach Y. Si $K : M \times N \to \mathbb{R}$ es una función continua con las siguientes características:*

a) K es convexa en M para todo $y \in N$, es decir

$$K(\lambda x_1 + (1-\lambda)x_2, y) \leq \lambda K(x_1, y) + (1-\lambda)K(x_2, y)$$

para cualesquiera $x_1, x_2 \in M$, $y \in N$ y $\lambda \in [0,1]$.

b) K es cóncava en N para todo $x \in M$, es decir,

$$K(x, \lambda y_1 + (1-\lambda)y_2) \geq \lambda K(x, y_1) + (1-\lambda)K(x, y_2)$$

para cualesquiera $y_1, y_2 \in N$, $x \in M$ y $\lambda \in [0,1]$.

Entonces K tiene un punto de ensilladura en $M \times N$.

Prueba.

Para cada $x \in M$ definimos

$$B_x := \{y \in N : \; K(x,y) = \max_{z \in N} K(x,z)\}$$

Como B_x es el conjunto de los puntos de N en que la función continua $z \to K(x,z)$ alcanza su máximo global (esto tiene sentido ya que N es un compacto) luego B_x es un subconjunto no vacío. Además, el conjunto de puntos en que una función continua toma un determinado valor siempre es un subconjunto cerrado. Luego B_x es cerrado y por lo tanto se tiene que para cada $x \in M$ el conjunto B_x es no vacío y compacto.

Veamos ahora que B_x es convexo.

Sean $y_1, y_2 \in B_x$ y $\lambda \in [0,1]$. Como N es convexo se tiene que $\lambda y_1 + (1-\lambda)y_2 \in N$, y por la hipótesis (b) se tiene

$$\max_{z \in N} K(x,z) \geq K(x, \lambda y_1 + (1-\lambda)y_2) \geq \lambda K(x,y_1) + (1-\lambda)K(x,y_2)$$

$$= \lambda \max_{z \in N} K(x,z) + (1-\lambda)\max_{z \in N} K(x,z) = \max_{z \in N} K(x,z).$$

Por lo tanto, $K(x, \lambda y_1 + (1-\lambda)y_2) = \max_{z \in N} K(x,z)$ lo que nos permite afirmar que B_x es convexo.

Análogamente se puede ver que el conjunto

$$A_y := \{x \in M : \; K(x,y) = \min_{w \in M} K(w,y)\}$$

es no vacío, compacto y convexo para cada $y \in N$.

Sean $x_1, x_2 \in A_y$ $\lambda \in [0,1]$. Como M es convexo, $\lambda x_1 + (1-\lambda)x_2 \in M$, por tanto,

$$\min_{w \in M} K(w,y) \leq K(\lambda x_1 + (1-\lambda)x_2, y) \leq \lambda K(x_1,y) + (1-\lambda)K(x_2,y)$$

(esto es consecuencia de que $K(.,y)$ es una función convexa).

Ahora teniendo en cuenta que $x_1, x_2 \in A_y$ se desprende que $\lambda x_1 + (1-\lambda)x_2 \in A_y$.

Sea $C := M \times N$. No es dificil probar que C es un compacto y convexo de $X \times Y$.

También es claro que si $(x,y) \in M \times N$ el conjunto $A_y \times B_x$ es un compacto convexo contenido en $M \times N$. Sea F la aplicación multivaluada definida por

$$F(x,y) = A_y \times B_x.$$

Veamos que F es cerrada, para poder aplicar el teorema de Kakutani.

En efecto, sean pues $(x_n, y_n) \in C$, $(x_n, y_n) \to (x, y)$. consideremos $(u_n, v_n) \in F(x_n, y_n)$, $(u_n, v_n) \to (u, v)$. Tenemos que probar que $(u, v) \in F(x, y)$.

Como $u_n \in A_{y_n}$,

$$K(u_n, y_n) = \min_{w \in M} K(w, y_n).$$

Como $v_n \in B_{x_n}$

$$K(x_n, v_n) = \max_{z \in N} K(x_n, z).$$

Siendo K continua y teniendo en cuenta el lema anterior se tiene:

$$K(u, y) = \lim_{n \to \infty} K(u_n, y_n) = \lim_{n \to \infty} \left[\min_{w \in M} K(w, y_n) \right] = \min_{w \in M} K(w, y),$$

luego $u \in A_y$.

Análogamente se prueba que $v \in B_x$.

Por lo tanto F es cerrada y por el teorema de Kakutani, existe $(x_0, y_0) \in C$ tal que

$$(x_0, y_0) \in F(x_0, y_0) = A_{y_0} \times B_{x_0}.$$

Luego:

$$x_0 \in A_{y_0} \Rightarrow K(x_0, y_0) = \min_{x \in M} K(x, y_0)$$

$$y_0 \in B_{x_0} \Rightarrow K(x_0, y_0) = \max_{y \in N} K(x_0, y)$$

de donde, para cualesquiera $x \in M,\ y \in N$ se obtiene:

$$K(x_0, y) \leq \max_{y \in N} K(x_0, y) = K(x_0, y_0) = \min_{x \in M} K(x, y_0) \leq K(x, y_0)$$

lo que nos dice que (x_0, y_0) es un punto de ensilladura de K.

Bibliografía

[1] R. Akerkar, *Nonlinear functional analysis* Narosa Publising House (1999).

[2] R. F. Brown, *A topological introduction to nonlinear analysis* Birkhäuser, (1993).

[3] K. Goebel, W.A. Kirk, *topics in metric fixed point theory* Cambridge studies in advanced mathematics 28,(1990).

[4] M.A. Khamsi, W.A. Kirk, *An introduction to metric spaces and fixed point theory* A Wiley-interscience Publication (2001)

[5] Lawrence C. Evans, *Partial differential equations* American Math. Soc. (Graduate Studies in Math., vol 19).

[6] S. Hu, N.S. Papageorgiou, *Handbook of multivalued analysis* Kluwer Academic Publishers (1997)

[7] E. Llorens Fuster, *Análisis funcional no lineal: una introducción* ed.(Llorens Fuster) (1999)

[8] D. R. Smith, *variational methods in optimization* Pretice- Hall (1974).

[9] John L. Troutman *variational calculus with elementary convexity* Springer (1980)

[10] E. Zeidler, *Nonlinear functional analysis and its applications* Vol I, *Fixed points Theorems.* Springer-Verlag (1986)

[11] E. Zeidler, *Nonlinear functional analysis and its applications. Vol III, Variational methods and optimization.* Springer-Verlag (1986)