Introducción al Cálculo Variacional N-dimensional

Gonzalo Galiano, 2002

Índice general

Capítulo 1

Algunas observaciones sobre la historia y los objetivos del Cálculo Variacional

La expresión *cálculo de variaciones* fue introducida por Leonhard Euler en el año 1756. La usó para describir un nuevo método que J.L. Lagrange había desarrollado el año anterior, del cual quería desentrañar todas sus consecuencias formales. Hoy en día, la expresión *cálculo de variaciones*, o *cálculo variacional*, como también se denomina a menudo, se usa en un sentido mucho más amplio. La materia de la que se ocupa el cálculo variacional es la formulación matemática de las ideas de

- maximizar,
- minimizar y
- *criticalizar.*

En otras palabras, se trata de encontrar el mínimo, máximo y puntos críticos de una función o funcional (funciones cuyos argumentos son, a su vez, otras funciones)

$$J : V \to \mathbb{R}.$$

En la mayor parte de las aplicaciones, V es un conjunto de números, funciones, trayectorias, curvas, superficies, etc.

Al final del siglo XVII, los matemáticos comenzaron a prestar atención a ciertos problemas de valores extremos. La primera publicación de Leibniz sobre cálculo diferencial apareció en el año 1684, bajo el título *Nova methodus pro maximis et minimis itemque tangentibus*.

Los puntos críticos de una función diferenciable J son los puntos en los que la derivada $J'(u)$ de la función $J(u)$ se anula, es decir, aquellos puntos en los que $J'(u) = 0$. La función $J(u)$ no tiene, necesariamente, valores extremos en todos sus puntos críticos; también puede tener *puntos de inflexión* con una tangente horizontal. Los puntos críticos, por tanto, son aquellos puntos que son *posibles* extremos, puesto que la condición $J'(u) = 0$ es meramente una condición necesaria en su definición.

En el siglo XVIII se usaba el término *variación* más que el término *diferenciación*, lo cual explica la aparición de la expresión *cálculo de variaciones*. El cálculo de variaciones es una materia de estudio muy natural para aquellos que piensan que sólo lo mejor es suficientemente bueno. Problemas de máximos y mínimos, o lo mejor y lo peor, aparecen frecuentemente en la vida diaria (uno no tiene más que mirar a los problemas de minimización de costes y de maximización de beneficios). Muchos problemas prácticos pueden representarse en esta forma. Por ejemplo, cómo deberíamos elegir la forma de un avión o un coche para que la resistencia que ejerce al aire sea lo menor posible, cómo debemos plantear el transporte y la distribución de un producto en un conjunto de mercados, cómo organizar un proceso de producción.

El cálculo variacional es, en algún sentido, muy antiguo. Los antiguos egipcios ya sabían que el camino más corto entre dos puntos es una línea recta. También sabían, al menos empíricamente, que el círculo es la figura geométrica que encierra una mayor área para un perímetro fijado.

El nacimiento de las matemáticas modernas tuvo lugar en la Grecia antigua, donde los matemáticos dejaron de preguntarse sólo el *cómo* para preguntarse también el *por qué*: no se contentaban sólo con encontrar reglas. Para los egipcios, las matemáticas no eran más que una simple herramienta de ayuda para la administración y el comercio. Pero los griegos se interesaron, con éxito, en problemas de valores extremos y *demostraron*, por ejemplo, que la trayectoria más corta que une dos puntos es la línea recta (Arquímides usó este hecho para definir la línea recta).

En algún momento entre el 200 a.C. y el 100 d.C., Zenodoros demostró que el área de un polígono es siempre más pequeña que el área de un círculo con el mismo perímetro. En este período también se averiguaron ciertas propiedades de minimalidad de las celdas hexagonales de los panales. Los griegos también determinaron la figura geométrica con mayor área para un perímetro dado, resolviendo así el primer *problema isoperimétrico*.

Otro descubrimiento griego muy importante es atribuído a Herón de Alejandría, en algún momento del siglo I. Era bien conocido que un rayo de luz que parte de un punto P y que incide en un espejo en un punto R será reflejado a un punto Q tal que los ángulos PR y QR son iguales. Si R' es algún otro punto del espejo, entonces la suma de las trayectorias $PR' + R'Q$ es mayor que $PR + RQ$. Herón fue capaz de caracterizar la trayectoria PRQ que la luz realmente sigue como la trayectoria más corta que une P con Q pasando por algún punto del espejo. Este descubrimento es uno de los puntos de inicio de la geometría óptica.

En 1685, Newton (1643-1727) investigó el problema de encontrar la resistencia mínima ofrecida por un sólido de revolución. El pensaba que *este teorema no será completamente inútil en la construcción de barcos*. Su aproximación al problema, sin embargo, no coincidió con posteriores descubrimientos en hidrodinámica. Christian Huygens (1629-1695) también escribió sobre métodos para resolver problemas variacionales en su libro sobre la luz.

No obstante, el nacimiento del moderno cálculo variacional se le atribuye a Johann Bernoulli con la introducción del *problema de la braquistocrona* (*brachristos* el menor *chronos* tiempo) en el Acta Eruditorum Lipsiae. Este problema puede formularse como sigue. Una masa puntual viaja sin fricción a lo largo de una curva que une un punto A con un punto B, el cual está por debajo de A. ¿Qué curva proporciona el tiempo más corto de viaje si la masa se mueve únicamente como consecuencia de la gravedad? El movimiento de la masa es descrito por la Ley de Newton. La fuerza resultante que actúa sobre la masa está compuesta por la gravedad y por la fuerza que restringe a la partícula a no abandonar la trayectoria impuesta. La resultante tiene la misma dirección que la tangente a la curva. Es obvio que tardará distintos tiempos en recorrer distintas trayectorias, y que la solución no es ni una línea recta ni un arco de circunferencia, incluso aunque la primera sea la distancia mínima entre los dos puntos.

En sus *Dialoghi*, Galileo plantea la cuestión de dos esferas iguales que comienza a rodar al mismo tiempo, una a lo largo de un arco circular y la otra a lo largo de su cuerda. ¿Cuál de las esferas alcanza el final de la cuerda antes? Por los experimentos, sabemos que es la esfera que viaja a lo largo del arco. Un cuerpo que se mueve en línea recta acelera relativamente despacio. Sin embargo, si la curva está más inclinada cerca del punto de salida, entonces aunque la longitud sea mayor, también lo será la velocidad.

Johann Bernoulli comienza su tratamiento del problema de la braquistocrona en un tono muy complacido: *Johann Bernouli, Profesor de Matemáticas, manda saludos a los más astutos matemáticos del mundo entero*. Anunció que había encontrado una solución muy elegante al problema, pero sin embargo no quiso hacerla pública, sino que prefirió primero retar a sus contemporáneos a estudiar el problema. Este reto iba particularmente dirigido a su hermano y profesor Jakob, trece años mayor que él, y su más enconado enemigo.

El problema de la braquistocrona concierne al tiempo que tarda la partícula en recorrer toda la trayectoria y es, por tanto, significativamente diferente de otros problemas más simples que dependen de un número finito de variables. Parece que en esos tiempos no se sabía que el problema isoperimétrico es del mismo tipo. A lo largo de los años, muchos matemáticos dieron soluciones a este problema, incluyendo a Newton, Leibniz (1646-1716) y L'Hôpital (1661-1704). Jakob Bernoulli encontró una solución, pero fue solo después de la insistencia de Leibniz, con el que tuvo una larga amistad y una abundante correspondencia científica, que la mandó a su hermano. La solución apareció en Mayo de 1697. Curiosamente, Newton envió su solución a un amigo justo un día después de que apareciera la solución de Jakob Bernoulli. Para la fascinación de los matemáticos, la solución resultó ser una cicloide, una curva que había sido descubierta muy recientemente. La cicloide es una curva simple transcendental que puede ser generada mecánicamente: es la trayectoria seguida por un punto de la circunferencia de un círculo que rueda en línea recta a lo largo de una superficie sin fricción.

Aquí la Figura 1

Ya hemos observado que Herón de Alejandría reconoció que la reflexión de un rayo de luz en un espejo plano podría ser descrita por medio de un principio de minimización. Fermat (1601-1665) observó que la ley de refracción podría ser expresada también como un principio de minimización y generalizó las predicciones de su ley a sistemas ópticos arbitrarios en los que la velocidad de la luz es función de la posición (como es el caso en la atmósfera). El principio de óptica geométrica de Fermat (1662) establece que la trayectoria de la luz a través de un medio no homogéneo es aquella que minimiza el tiempo de viaje. Este principio no es sólo de importancia teórica, sino que forma la base para el diseño práctico de lentes ópticas.

Los diferentes métodos usados para resolver el problema de la braquistocrona fueron creados ad hoc. El siguiente paso en el desarrollo del cálculo variacional fue la creación de una teoría general que permitía la resolución de problemas generales. Los nombres de Lagrange (1736-1813) y Euler (1707-1783) están inextricablemente unidos con este paso. Poco antes de 1732, Euler comenzó un estudio sistemático de problemas de valores extremos. Es poco probable que Euler estudiara cálculo variacional en sus primeros años de formación con Johann Bernoulli, en Basilea. El método que Euler desarrolló está más influenciado por el trabajo de Jakob que por el de Johann. Es muy diferente, sin embargo, del de todos sus predecesores, que trataron solo con problemas particulares. Los cálculos de Euler produjeron una verdadera teoría, elevando el estatus del cálculo de variaciones al de una disciplina matemática por derecho propio. Euler puso el *principio de mínima acción* sobre bases firmes (sin acudir a razonamientos metafísicos, como habían hecho algunos de sus predecesores). El término *acción* es cualquier cantidad con las dimensionses de energía por tiempo. El principio de mínima acción fue usado entonces por Lagrange y, más tarde, por Hamilton (1805-1865). El papel clave que juega la función hamiltoniana en la física matemática moderna pone claro la naturaleza fundamental de la contribución de Euler en este campo.

Una geodésica es el camino más corto entre dos puntos fijados en una superfice. En una esfera, estas líneas son los arcos de círculos máximos. En el mundo curvado de una superficie esférica, las geodésicas juegan el mismo papel que las líneas rectas en un mundo plano. La teoría de la relatividad de Einstein (1915) es una teoría del espacio-tiempo y la gravitación, es decir, una versión moderna de la teoría de Newton. El núcleo de la teoría de la relatividad es la idea de que los efectos gravitacionales pueden estar causados por la geometría de la variedad espacio-temporal. Einstein explicó la fuerza gravitacional vía la curvatura del espacio-tiempo, que es como se explica la influencia de la masa en la relatividad general. Así, una partícula que se mueve bajo los efectos de un campo gravitacional se mueve a lo largo de las geodésicas. Puesto que el espacio cuatridimensional se curva en presencia de un campo gravitacional, el movimiento real de una partícula en el espacio no es, por tanto, una línea recta.

Euler descubrió la primera condición necesaria que debe cumplir una función minimizadora de una clase especial de *funcionales*. Hoy en día esta condición se conoce como ecuación de Euler-Lagrange, y es la condición para funcionales que corresponde a $J'(u) = 0$ para funciones. Euler consideró funcionales de la forma

$$J(u) := \int_a^b F(x, u, u')dx$$

y buscó una función que minimizase este funcional. En aquella época, el tratamiento del cálculo diferencial e integral se basaba en el cálculo con cantidades infinitesimalmente pequeñas. La formulación rigurosa en términos de límites se debe a Cauchy. El método de Euler consiste en dividir el intervalo de integración $[a, b]$ en subintervalos $[x_{i-1}, x_i]$ usando puntos $a \leq x_1 < x_2 < ... < x_k \leq b$, y entonces reemplazar la curva por un polígono. Las esquinas de este polígono son los puntos (x_i, u_i), donde $u_i = u(x_i)$. También reemplazó la derivada $u'(x_i)$ por $\frac{\Delta u_i}{\Delta x_i}$, obteniendo por tanto, en lugar de una integral, una función de k variables $u_1, ..., u_k$. Haciendo que la integral de esta función con respecto u_i se anule y haciendo el número de puntos k tender a infinito sujeto a que la condición de que la ecuación para u se mantenga, obtuvo la bien conocida *condición de Euler*

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'} = 0$$

para la función incógnita u. En general, la ecuación de Euler no es fácil de resolver, puesto que las ecuaciones diferenciales de segundo orden son integrables por funciones elementales sólo en casos muy excepcionales.

El método de Euler fue considerablemente simplificado después por Lagrange, que observó que *no posee la simplicidad que sería de desear en una de las ramas del análisis puro*. En vez de hacer una aproximación poligonal, Lagrange dejó variar la curva y reemplazó $u(x)$ por una función $u(x, \tau)$ con un parámetro extra τ tal que $u(x, 0) = u(x)$. Esta idea fue extendida más adelante, dando lugar al concepto de homotopía.

La variación es entonces la función $\delta u(x) = \frac{(\partial u)}{(\partial \tau)}(x, 0)$ y $\delta(u') = (\delta u')$. Así, es suficiente con reemplazar $u(x)$ por $u(x, \tau)$ en la integral $\int_a^b F(x, u, u')dx$ y requerir que la derivada de esta expresión con respecto a τ se anule en $\tau = 0$. Una integración parcial da entonces la condición de Euler y, adicionalmente, condiciones de frontera que Euler había ignorado casi siempre. En efecto, definiendo

$$\varphi(\tau) = \int_a^b F(x, u(x, \tau), u'(x, \tau))dx,$$

se tiene

$$\begin{aligned}
\varphi'(\tau) &= \int_a^b (\frac{\partial F}{\partial u}\frac{\partial u}{\partial \tau} + \frac{\partial F}{\partial u'}\frac{\partial u'}{\partial \tau})dx \\
&= \int_a^b (\frac{\partial F}{\partial u}\frac{\partial u}{\partial \tau} + \frac{\partial F}{\partial u'}\frac{d}{dx}\frac{\partial u}{\partial \tau})dx \\
&= \int_a^b (\frac{\partial F}{\partial u}\frac{\partial u}{\partial \tau} - \frac{d}{dx}(\frac{\partial F}{\partial u'})\frac{\partial u}{\partial \tau})dx + \left[\frac{\partial F}{\partial u'}\frac{\partial u}{\partial \tau}\right]_a^b \\
&= \int_a^b (\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'})\frac{\partial u}{\partial \tau}dx + \left[\frac{\partial F}{\partial u'}\frac{\partial u}{\partial \tau}\right]_a^,
\end{aligned}$$

de modo que $\varphi'(0) = 0$ si las condiciones de frontera se anulan y

$$\int_a^b (\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'})\frac{\partial u}{\partial \tau}dx = 0.$$

Lagrange añadió, sin justificar, que se sigue directamente del hecho de que la integral

$$\int_a^b (\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'})\delta u(x)dx$$

se anule el que

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'} = 0.$$

Euler se dió cuenta inmediatamente de que hacía falta una demostración para esto. Lagrange escribió a Euler con un intento de demostración que Euler, sin embargo, rechazó justificadamente. El lema necesario no fue probado hasa 1879, por el matemático Du Bois-Reymond, de Tübingen, y más adelante se vió que era el *lema fundamental del cálculo de variaciones*: Si para toda $\psi \in C^2(a,b)$ se tiene

$$\int_a^b f(x)\psi(x)dx = 0,$$

entonces $f(x) = 0$ para todo $x \in (a,b)$. Para demostrarlo, Du Bois-Reymond usó la noción de continuidad que había sido introducida por Cauchy en 1820.

Al final del siglo XIX, el concepto de diferenciación fue generalizado por Volterra, Hadamard y sus dos pupilos Fréchet y Gateaux, Hilbert, y otros, a espacios de dimensión infinita, proporcionando de esta manera una base sólida al cálculo variacional.

La confluencia de métodos e ideas del álgebra, geometría, topología y análisis produjo una nueva rama de las matemáticas, el *análisis funcional*, que abarca una generalización de todos los conceptos del análisis clásico (límites, convergencia, continuidad, diferencial, etc.) al caso de espacios de dimensión infinita. El cálculo variacional o cálculo diferencial de funciones puede ser visto como la parte más antigua del análisis funcional.

Lagrange (1736-1813) desarrolló una teoría analítica del cálculo de variaciones generalizando el método para integrales múltiples, y comenzó a investigar superficies minimales (superficies con curvatura media nula). El problema de encontrar la superficie minimal de una frontera, con restricciones apropiadas (normalmente que la superficie no contenga singularidades) fue también objeto de investigación por parte del físico Plateau (1801-1883) y por ello se le llama el problema de Plateau. Matemáticamente, el problema de Plateau involucra la resolución de una ecuación en derivadas parciales o un sistema de tales ecuaciones.

Mécanique analytique, escrito por Lagrange en 1788, casi cien años después de los *Principia* de Newton, es probablemente su trabajo más importante. Los Principia, que apareció en 1867, fue escrito en el lenguaje de la geometría griega clásica. Usando el lenguaje del cálculo variacional, Lagrange fue capaz de unificar estática y dinámica, que habían estado separadas hasta este momento. El punto de vista geométrico de Newton fue totalmente eliminado. Este fue el triunfo del análisis puro, y Lagrange fue el primer analista puro. Como dijo Hamilton, Lagrange elevó la mecánica a una forma de poesía científica. En su prefacio, Lagrange enfatizó que *en este libro no se encontrarán figuras geométricas sino sólo cálculos algebraicos*. La interacción entre matemáticas y mecánica es característica del siglo XVIII: las matemáticas fueron desarrolladas fundamentalmente como una herramienta para la resolución de problemas de las ciencias naturales (geodesia, astronomía y mecánica).

El conocimiento de la diferencia entre los distintos puntos críticos (máximos, mínimos y puntos de inflexión) fue empírica hasta que Legendre (1752-1833) comenzó sus investigaciones sobre la *segunda variación*. Con la ayuda de este método, en 1786 consiguió deducir la segunda condición necesaria para la existencia de un extremo, aunque Lagrange pronto se dió cuenta de que la demostración de Legendre no estaba completa.

Los métodos del cálculo variacional fueron aplicados satisfactoriamente a la mecánica. Jacobi (1804-1851) coronó con su *Teoría del cálculo variacional y ecuaciones diferenciales* en 1837. Consiguió deducir correctamente la condición de Legendre y demostró la tercera condición necesaria para la existencia de un mínimo. Como ya sabían los griegos, la línea geodésica es aquel arco de un círculo máximo que une dos puntos de la superficie de una esfera. Esta curva es un mínimo en tanto que los puntos no estén diametralmente opuestos. Si lo están, por ejemplo los polos norte y sur, entonces hay una familia entera de arcos semicirculares y los puntos son llamados *conjugados*. Así, hay una familia de curvas que unen dos puntos conjugados: cada miembro de la familia es un arco maximizador o minimizador, pero no es posible distinguir entre ellos examinando sus propiedades de extremos. La condición necesaria de Jacobi elimina la posibilidad de puntos conjugados.

Sin embargo, como ya señalamos, la existencia de una solución de cualquier problema de valores extremos todavía requiere una demostración particular, y eso es usualmente una de las mayores dificultades del cálculo variacional. Hasta bien adentrado el siglo XIX, los grandes matemáticos tales como Gauss (1777-1855), Dirichlet (1805-1859) y Riemann (1826-1866) dieron por hecho la existencia de soluciones de este tipo de problemas. Gauss, en 1839, en sus investigaciones sobre electrostática y Riemann, en 1851, en su tesis doctoral sobre la fundamentación de la teoría de funciones, llevaron acabo investigaciones sobre la resolución del problema de Dirichlet. Este problema consiste en resolver la ecuación de potencial

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{en } \Omega,$$

con

$$U = f \quad \text{en } \partial\Omega,$$

para valores dados sobre la frontera $\partial\Omega$ de un conjunto $\Omega \subset \mathbb{R}^2$. Esta ecuación en derivadas parciales es la ecuación de Euler-Lagrange de la integral de Dirichlet

$$J(u) = \int\int_G \left((\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2 \right) dxdy.$$

Este problema aparece, por ejemplo, en electrostática, donde u es el potencial eléctrico. El sistema está en equilibrio estable cuando la energía $D(u)$ es un mínimo. El significado particular de este problema para la física fue reconocido, en primer lugar, por Thomson (1824-1907).

El integrando en la integral de Dirichlet es no negativo y por tanto tiene una cota mínima que es mayor o igual a cero. Esto llevó a Riemann, el sucesor de Dirichlet en Götingen, a la conclusión de que tenía que haber una función $\hat{u}$ que minimizara la integral y que fuera, por tanto, una solución del problema de Dirichlet. Riemann entonces enunció el *principio de Dirichlet* por el cual este tipo de problemas variacionales siempre tenían una solución, sin tratar siquiera de demostrar esta proposición matemáticamente.

Weierstrass (1815-1897) criticó esta afirmación en 1870 y mostró que la existencia a priori de solución de cualquier problema variacional no está asegurada en absoluto y que, en el caso general, la existencia no puede ser asumida. De hecho, dió muchos ejemplos de casos en los que la cota mínima no se alcanza. Las investigaciones de Weierstrass fueron esenciales y prepararon el camino para Hilbert, que fue, más adelante, el principal responsable de proporcionar una demostración de existencia, y por tanto, de establecer los métodos directos del cálculo de variaciones. Hilbert escribió lo siguiente sobre Weierstrass en 1926: *Si hoy en día reina una completa seguridad en el campo del análisis, es principalmente debido a los esfuerzos científicos de Weierstrass*. Y de hecho, la revisión crítica de Weierstrass y su escuela fueron fundamentales en el desarrollo de todo el análisis.

Un simple ejemplo geométrico de problema variacional sin solución aparece en la Figura 1. Dos puntos A y B de una línea recta se unen por una curva continua C que, además, es perpendicular en los puntos A y B a la línea recta.

Aquí la Figura 2

Este problema no tiene solución porque tal curva es siempre más larga que la línea recta AB, aunque su longitud puede estar arbitrariamente cerca de la longitud de la línea recta. De este modo, hay siempre una cota mínima, pero no hay un mínimo, para una curva de esta familia. En efecto, C puede ser dividida en dos pequeños arcos de radio ε cuyo centro cae en AB y en una línea recta. La longitud de C es entonces $l(C) = \varepsilon\pi + a - 2\varepsilon = a + (\pi - 2)\varepsilon$, donde a es la longitud del segmento AB. Por tanto la diferencia de longitudes, $(\pi - 2)\varepsilon$, es arbitrariamente pequeña. Así, este problema variacional, bien planteado, no tiene solución. La dificultad central del cálculo variacional es determinar cuándo existe una solución y obtener condiciones suficientes para que esto ocurra.

Cincuenta años después de la publicación de los resultados de Riemann, David Hilbert (1862-1943) rescató el principio de Dirichlet en su famosa conferencia *Sobre el principio de Dirichlet* en 1901. Lo que Riemann y Dirichlet habían conjeturado, Hilbert lo demostró. Esta rehabilitación del principio de Dirichlet fue uno de los grandes triunfos en la historia del análisis y de la física matemática. El trabajo de Hilbert, y el de Lebesgue (1875-1941) sobre el mismo tema en 1907, anunciaron la llegada del moderno cálculo variacional abstracto. En el Segundo Congreso Internacional de Matemáticas de Paris, en 1900, Hilbert presentó 23 problemas matemáticos de los cuales el problema 20 trataba de la existencia de un mínimo para todos los funcionales regulares con condiciones de frontera razonables. Hilbert terminó su conferencia con una llamada para ahondar en el desarrollo del cálculo variacional. Trató de describir todas las diferentes ramas de investigación matemática que habían tenido lugar en las décadas previas y de apuntar a trabajos que sería interesante acometer en el futuro. Estaba convencido de que la demostración del teorema de existencia debía ser posible, incluso si había que hacer algún tipo de generalizaciones, como la del concepto de derivada o de solución. Este presentimiento falló en lo que se refiere al problema de superficies mínimas ya que, aunque este funcional es regular en el sentido de Hilbert, Bernstein (1880-1966) produjo una demostración de la no existencia de solución para regiones no convexas del plano en 1912. Sin embargo, dice mucho en favor de Hilbert el hecho de que sus 23 problemas inspiraran a los matemáticos del siglo XX y les llevaran incluso a crear nuevas disciplinas.

Como ya hemos enfatizado, el cálculo variacional siempre ha inspirado a los matemáticos a pensar en nuevos conceptos del análisis: continuidad, semicontinuidad, equicontinuidad, entorno de una función y compacidad son todos merecedores de ser nombrados. Legendre pensó acerca de si una solución u de la ecuación de Euler debía ser necesariamente un extremo relativo de la integral $\int_a^b F(x,u,u')dx$ o no. Su primer intento para resolver el problema fue sustituir u por $u+\varepsilon v$ y requerir a esta función de ε que poseyera un extremo relativo para $\varepsilon = 0$, para cualquier variación $\delta u = \varepsilon v$. Uno se da cuenta enseguida, sin embargo, de que esto no es suficiente para garantizar que el valor de la integral sea mayor (o menor) que el valor que se obtiene cuando se sustituye u por $u + \delta u$ para una variación δu *pequeña*. El punto crucial es la especificación de *pequeño*. Se asumió durante mucho tiempo que el hecho de que la variación δu fuera pequeña también implicaba que la variación $\delta u'$ es pequeña, hasta que Weierstrass y sus estudiantes se dieron cuenta de que esta propiedad, de hecho, significa una condición adicional. Esto llevó, al final del siglo XIX, a la distinción entre dos tipos de extremos: los fuertes y los débiles. Para un extremo fuerte, $|u - u_0|$ debe ser pequeño en un entorno de u_0, mientras que para un extremo débil, también lo tiene que ser $|u' - u_0'|$. Esta noción de diferentes clases de entornos fue claramente expresada por Fréchet (1878-1973), quien introdujo el concepto de distancia en un conjunto arbitrario.

La generalización de conceptos usados normalmente en $\mathbb{R}^n$, tales como cotas, continuidad, entorno, compacidad y conectividad a estas situaciones extremamente generales es remarcable.

Fréchet enfatizó, en particular, que la idea de compacidad está implícita en los métodos directos del cálculo variacional que Hilbert había desarrollado en 1900. Por ejemplo, Hilbert demostró que existe una geodésica en toda superficie regular, compacta y simplemente conexa de $\mathbb{R}^3$. Fréchet observó que es posible definir una distancia en el conjunto de curvas que conectan dos puntos A y B en tales superficies, y que para esta distancia, resulta que el conjunto de arcos cuya longitud está acotada es un conjunto compacto. El resultado de Hilbert entonces se sigue de un teorema general de Baire (1874-1932) que probó que para funciones de $\mathbb{R}^n$, una función semicontinua inferiormente definida en un compacto alcanza siempre un mínimo. La longitud introducida por Fréchet es desde luego semicontinua inferiormente en el conjunto de arcos que unen A y B. El significado que tiene la semicontinuidad para el cálculo variacional fue específicamente enfatizado por Tonelli (1855-1946).

El nombre de *Carathéodory* (1873-1950) está inextricablemente unido con el cálculo variacional. Carathéodory fue descrito por Pringsheim como el *isoperimétrico incomparable*. Su aproximación al cálculo variacional está descrito en su libro *Varational Calculus and First-order Partial Differential Equations* (1935).

La relación entre cálculo variacional y la teoría de ecuaciones en derivadas parciales elípticas ha sido ampliamente estudiada, al igual que la regularidad de las soluciones de los problemas variacionales. Se han desarrollado teorías nuevas, muy cercanas al cálculo variacional, por ejemplo la teoría de las desigualdades variacionales (Lions, Stampacchia), la teoría de operadores monótonos (Browder) y los teoremas de minimax. Los métodos variacionals teóricos de aproximación han sido aplicados a problemas físicos y el trabajo de Rayleigh, Ritz, Galerkin, Courant y Hellinger es especialmente relevante. Las ideas clave del cálculo de variaciones son todavía hoy una importante y fructífera fuente de desarrollo en matemáticas, en particular en el análisis funcional, la geometría diferencial, la topología algebraica y los campos en los que se usan, como la física, la economía, la ingeniería, la biología, etc.

Los métodos varacionales teóricos han jugado siempre un papel importante en la mecánica teórica. Por ejemplo, el principio de Hamilton es frecuentemente visto como una alternativa a las ecuaciones del movimiento de Newton. Por supuesto, el significado de una formulación variacional de las leyes de la física va más allá de ser simplemente una descripción alternativa. Su ventaja práctica reside en la posibilidad que ofrece de incorporar en un único funcional todas las propiedades características de un problema, tales como las condiciones de frontera, las condiciones iniciales y las restricciones. También es muy importante el hecho de que estas formulaciones proporcionan los mejores métodos de aproximación, y en muchos casos puede usarse para construir cotas superiores e inferiores de la solución.

La formulación hamiltoniana es uno de los pilares en los que se basa la teoría cuántica. También es indispensable en la mecánica estadística y resulta de gran utilidad en la teoría clásica de campos, es decir, en sistemas con infinitos grados de libertad, como es el caso de la teoría de

Maxwell . Por otra parte, en los últimos 400 años la mecánica ha probado ser el núcleo de la física teórica.

Debido a la contribución histórica del cálculo variacional al desarrollo de las matemáticas y a las posibilidades que ofrece para la resolución de numerosos problemas matemáticos de la física, ingeniería y economía, el cálculo variacional disfruta de cierto estatus. Nuevos desarrollos en matemáticas puras y aplicadas, física, economía matemática, teoría del control, biología e ingeniería solo sirven para confirmar este punto, y ofrece numerosos ejemplos de la versatilidad y poder de los métodos teóricos del cálculo variacional.

En este punto dejamos la historia del cálculo variacional para pasar a discutir las ideas básicas de sus métodos directos.

Como mencionamos anteriormente, la tarea del cálculo variacional es determinar los extremos y puntos críticos de una función o funcional

$$J : V \to \mathbb{R}$$

bajo las condiciones más generales posibles. Esta generalidad es requerida tanto para el conjunto V como para el funcional J. Si, por ejemplo, V es un intervalo en $\mathbb{R}$ o un subconjunto de $\mathbb{R}^n$, entonces es posible resolver este problema usando únicamente los métodos del análisis clásico: el teorema fundamental de Weiertrass establece que toda función continua $f : K \to \mathbb{R}$ definida en un conjunto compacto $K \subset \mathbb{R}^n$ está acotada y alcanza sus valores máximo y mínimo. En efecto, puesto que un conjunto $K \subset \mathbb{R}^n$ es compacto si y sólo si es cerrado y acotado, y como la continuidad de f implica la compacidad de $f(K)$, se sigue que $f(K)$ es cerrado y acotado, con lo que la aserción se deduce sin más que considerar sucesiones maximizantes y minimizantes.

Pasamos ahora a definir la noción de semicontinuidad, como fue introducida en la teoría de funciones reales por Baire. Se dice que una función $f : V \to \mathbb{R}$ es *semicontinua inferiormente en un punto* x_0 si, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todo x con $|x - x_0| < \delta$ se tiene $f(x) - f(x_0) > -\varepsilon$. Es fácil ver que esta definición es equivalente a requerir que el conjunto

$$\{x \in V : f(x) \leq f(x_0)\}$$

sea cerrado, y que f es semicontinua inferiormente en V si

$$\{x \in V : f(x) \leq \alpha\}$$

es cerrado para todo $\alpha \in \mathbb{R}$. Si, en el teorema de Weierstrass, nos limitamos a demostrar la existencia de un mínimo, entonces podemos ver que la hipótesis de continuidad puede sustituirse por la de semicontinuidad inferior. Demostremos aquí este resultado.

TEOREMA 0.1. *Sea* $f : K \subset \mathbb{R}^n \to \mathbb{R}$ *una función semicontinua inferiormente (sci) y* K *un conjunto compacto. Entonces* f *alcanza un mínimo en* K.

DEMOSTRACIÓN. Sea $\{x_n\} \subset K$ una *sucesión minimizante para* f, es decir, una sucesión tal que

$$f(x_n) \to \inf_{x \in K} f(x).$$

Como f es sci, los conjuntos

$$C_n = \{x \in K : f(x) \leq f(x_n)\}$$

son cerrados, y por tanto también lo es $\cap_{n=1}^{\infty} C_n$. Al estar contenida en el compacto K, esta intersección es un conjunto compacto. Hay dos casos posibles:

1. $\cap_{n=1}^{\infty} C_n = \emptyset$. Como es una intersección de compactos, se sigue que existe una colección finita $\left\{C_{n_j}\right\}$ tal que $C_{n_1} \cap \ldots \cap C_{n_m} = \emptyset$. Pero,

$$C_{n_1} \cap \ldots \cap C_{n_m} = \left\{ x \in K : f(x) \leq \min_{j=1,\ldots,m} f(x_{n_j}) \right\},$$

   es obviamente no vacío.
2. $\cap_{n=1}^{\infty} C_n = \emptyset$. Sea $y \in \cap_{n=1}^{\infty} C_n$. Entonces $f(y) \leq f(x_n)$ para todo $n \in \mathbb{N}$, de donde se sigue que

$$f(y) \leq \inf_{x \in K} f(x),$$

   y por tanto $y \in K$ es un punto de mínimo.

$\square$

En muchas aplicaciones importantes del cálculo variacional, sin embargo, V no es un subconjunto de un espacio vectorial de dimensión finita. En el viejo problema de encontrar el mínimo trayecto entre dos puntos, V es el conjunto de todas las trayectorias que unen esos dos puntos. En este ejemplo y en muchas aplicaciones físicas, V resulta ser un subconjunto de un espacio funcional de dimensión infinita. Lo siguiente parece, por tanto, una buena delimitación del cálculo variacional abstracto.

El cálculo variacional investiga condiciones sobre un espacio topológico X, un subconjunto $V \subset X$ y un funcional $J : V \to \mathbb{R}$ que garantize la existencia de un punto $\hat{u} \in V$ tal que

$$J(\hat{u}) = \inf_{u \in V} J(u)$$

o

$$J(\hat{u}) = \sup_{u \in V} J(u).$$

El punto $\hat{u}$ es llamado un *punto de mínimo* o *punto de máximo*. En su versión abstracta, pues, el cálculo variacional trata de enunciar y demostrar un resultado análogo al del teorema de Weierstrass. Hace más de cien años, Bolzano (1781-1848) demostró que todo conjunto infinito y acotado de $\mathbb{R}$ posee al menos un punto de acumulación y señaló la importancia de este teorema para el análisis real. La idea de elegir una sucesión convergente de entre un conjunto infinito que no consiste de puntos de $\mathbb{R}^n$ sino de funciones o curvas fue y es usado en cálculo variacional para las demostraciones de existencia de soluciones. En general, decimos que un subconjunto de un espacio topológico es *secuencialmente compacto* si toda sucesión de elementos del subconjunto tiene una subsucesión convergente. Esto lleva a la definición general de compacidad de un subconjunto de un espacio topológico abstracto. Se dice que un subconjunto es *compacto* si todo recubrimiento por abiertos posee un subrecubrimiento finito. Observemos aquí que compacidad y compacidad secuencial son equivalentes en espacios normados.

Muchos problemas de optimización tienen lo siguiente en común. Supongamos que J es un funcional lineal continuo en un espacio de Banach V. Entonces nuestro propósito es encontrar un punto $\hat{u}$ de V con $\|\hat{u}\| = 1$ tal que

$$J(\hat{u}) = \|J\| := \sup_{\|u\|=1} |J(u)|.$$

Tenemos pues que encontrar condiciones que garantizen la existencia de tal punto $\hat{u}$. Si la bola cerrada unitaria

$$K_1 = \{u \in V : \|u\| \leq 1\}$$

fuera compacta, entonces la existencia de $\hat{u}$ se deduciría directamente de la continuidad de J. Desafortunadamente, un teorema demostrado por Riesz (1880-1956) en 1918 establece que las bolas cerradas unitarias de espacios de Banach son compactas si y sólo si el espacio es de dimensión finita. De modo que, aunque K_1 es cerrado y acotado, no es compacto en un espacio de Banach de dimensión infinita. En tales espacios se necesita una condición extra, además de ser cerrado y acotado, para caracterizar los conjuntos compactos. Nosotros daremos estas condiciones suplementarias para varios espacios de Banach. Es fácil convencerse del resultado del teorema de Riesz considerando el siguiente contraejemplo. Es imposible elegir una subsucesión convergente de una sucesión $\{x_n\}$ de un espacio de Hilbert de dimensión infinita $\mathcal{H}$, donde $< x_n \, , \, x_m >= \delta_{nm}$ (en otras palabras, $\{x_n\}$ es un sistema ortonormal de $\mathcal{H}$), porque

$$\|x_n - x_m\| =< x_n - x_m \, , \, x_n - x_m >= \sqrt{2} \qquad \text{para todo } n \neq m.$$

Un camino obvio de superar esta dificultad sería introducir una métrica en el espacio de Banach que hiciera que la bola cerrada unitaria fuera compacta al mismo tiempo que conservara la continuidad de la función J. Desafortunadamente, no es posible realizar esto en el caso general. Así, no tenemos más elección que pasar a un concepto más general de espacio que el de espacio métrico. Esto conduce a la *topología débil*: una sucesión $\{u_n\}$ converge débilmente a $\hat{u}$ si, para todo funcional continuo J de V (es decir, $J \in V'$), tenemos $\lim_n J(u_n) = J(\hat{u})$. Es posible elegir de toda sucesión acotada una subsucesión débilmente convergente, y entonces a esta subsucesión se le llama *débil secuencialmente convergente*. La bola cerrada unidad K_1 (y todo subconjunto cerrado, acotado y convexo de V) a menudo tiene una propiedad muy útil: es débil secuencialmente compacto. Estas propiedades de compacidad son de gran importancia en las distintas versiones del teorema fundamental del cálculo variacional. Como en el caso del análisis clásico, el segundo ingrediente de estos teoremas es un debilitamiento natural del concepto de continuidad para funciones de espacios de dimensión infinita, es decir, el concepto de semicontinuidad. Observemos que en, por ejemplo, un espacio de Banach, la continuidad respecto la topología engendrada por la norma no es la misma que la que engendra la topología débil (que es menos fina). Para funciones convexas se puede, al menos, probar semicontinuidad inferior débil.

Una solución completa de un problema variacional comprende

- existencia,
- unicidad, y
- cálculo

del punto de mínimo $\hat{u} \in V$. Los problemas de existencia y unicidad pueden ser resueltos con cierto grado de generalidad. Llevaremos a cabo este trabajo en el Capítulo ?? El problema del cálculo de la solución es considerablemente más complicado y requiere mucha más base. Para resolver este problema, introduciremos en el Capítulo ?? el cálculo diferencial e integral en espacios de Banach y prestaremos especial atención al concepto de derivada de Fréchet J' de un funcional $J : V \to \mathbb{R}$, donde V es un subconjunto abierto de un espacio de Banach. Podremos entonces obtener una solución si somos capaces de encontrar el punto de mínimo $\hat{u}$ en el conjunto de puntos críticos de J, es decir, en $\{u \in V : J'(u) = 0\}$.

Capítulo 2

Introducción y ejemplos

1. Definición de un problema de optimización

Un problema de optimización puede presentarse de modo esquemático de la siguiente manera: una variable física, de decisión o de control debe ser elegida de modo óptimo, es decir, de modo que se optimize (minimize o maximize, según el caso) un criterio físico (acción, energía, ...), un criterio técnico (precisión, estabilidad, duración,...) o económico (coste, rentabilidad, productividad,...), todo respetando restricciones intrínsecas a la situación que se considera.

A continuación introducimos la terminología y la notación que usaremos para este tipo de problemas. Un *problema de optimización* $\mathcal{P}$ está formado por

1. Un *criterio*, J, función o funcional que aplica el espacio V de variables de decisión en $\mathbb{R}$,

$$J : V \to \mathbb{R}.$$

2. Las *restricciones*. En general, cualquier elemento del espacio V no es admisible como solución del problema $\mathcal{P}$, ya que debe satisfacer algunas restricciones que determinan el *espacio de soluciones* de $\mathcal{P}$. Estas restricciones aparecen en las aplicaciones de diversas formas, que pueden intervenir de forma simultánea:

   a) Restricciones de *conjunto* del tipo

$$v \in C,$$

   donde $C \subset V$ es un conjunto dado.

   b) *Restricciones funcionales de igualdad*

$$F(v) = 0, \tag{1}$$

   donde $F : V \to \mathbb{R}^m$. Diremos que la solución está sujeta a m *restricciones de igualdad*

$$F_i(v) = 0, \quad i = 1, ..., m.$$

   c) *Restricciones funcionales de desigualdad*

$$G(v) \leq 0, \tag{2}$$

   donde $G : V \to \mathbb{R}^p$ y (2) tiene el siguiente sentido

$$G_j(v) \leq 0 \quad j = 1, ..., p.$$

Decimos entonces que tenemos p *restricciones de desigualdad*. Para una solución admisible, v, diremos que las restricciones de desigualdad están *saturadas* si $G_j(v) = 0$, y que no lo están si $G_j(v) < 0$.

d) *Restricciones de tipo ecuación o desigualdad funcional*. Aquí distinguiremos los casos en que (1) o (2) son ecuaciones o desigualdades integrales, diferenciales, o en derivadas parciales.

En todo caso, el conjunto de restricciones define un subconjunto $U \subset V$, al que llamaremos *conjunto de soluciones admisibles*:

$$U := \{v : v \quad \text{satisface las restricciones}\} . \tag{3}$$

El problema de optimización $\mathcal{P}$ consiste pues en encontrar $u \in U$ tal que

$$J(u) \leq J(v) \quad \text{para toda } v \in U, \qquad \text{(problema de minimización)},$$

o

$$J(u) \geq J(v) \quad \text{para toda } v \in U, \qquad \text{(problema de maximización)}.$$

Si tal u existe, diremos que es una *solución optimal* y que $J(u)$ es un *valor optimal* del problema $\mathcal{P}$.

Puesto que el problema de determinar el máximo de un funcional J es equivalente al de determinar el mínimo de $-J$, a partir de ahora trataremos únicamente el problema de minimización. En lo que sigue, escribiremos de forma abreviada la formulación del problema $\mathcal{P}$ como sigue:

$$(\mathcal{P}) \begin{cases} \inf J(v) \\ v \in C, \quad F(v) = 0, \quad G(v) \leq 0. \end{cases} \tag{4}$$

En el caso más habitual en el que V es un espacio topológico, podemos introducir la noción más débil de *mínimo local o relativo* (la noción definida anteriormente lleva entonces el nombre de *mínimo global o absoluto*): $\bar{u} \in U$ es un *mínimo local* si existe un entorno B de $\bar{u}$ tal que

$$J(\bar{u}) \leq J(v) \quad \text{para toda } v \in U \cap B. \tag{5}$$

Claramente, un mínimo global es siempre un mínimo local, mientras que el recíproco es falso, en general (veremos en el Capítulo **??** que bajo ciertas hipótesis sobre la convexidad de J y U se da la equivalencia entre ambas nociones).

2. Ejemplos

Los ejemplos que presentamos a continuación servirán para ilustrar los métodos que presentaremos más adelante. Los clasificamos en

- Problemas de dimensión finita, es decir, $J : \mathbb{R}^n \to \mathbb{R}$, y
- Problemas en dimensión infinita, es decir, $J : V \to \mathbb{R}$, los cuales a su vez dividimos en
  - Problemas del Cálculo de Variaciones,
  - Problemas isoperimétricos, y

2.1. Ejemplos en dimensión finita.

2.1.1. Ejemplo 1. Programación lineal. Muchos problemas económicos o técnicos se presentan en la forma de un *programa lineal*:

$$\begin{cases} \inf <c\,,\,v> \\ v \in \mathbb{R}^n, \quad Av \le b, \end{cases} \tag{6}$$

donde A, b y c son matrices de orden $m \times n$, $m \times 1$ y $n \times 1$, respectivamente.

El primer programa lineal (tratado en 1944) que fue el origen de lo que hoy se denomina *investigación operativa*, es un problema de *ración alimentaria*. Se dispone de n tipos de alimentos y tomamos en consideración m parámetros alimentarios (proteínas, vitaminas, etc). Desiganamos por

- a_{ij} la cantidad del parámetro i contenida en una unidad del alimento j,
- b_j la cantidad mínima necesaria que debe ser introducida en cada ración, y
- c_j el coste de la unidad de alimento j.

Así, la ración alimentaria de coste mínimo, conteniendo v_j unidades del alimento j, es solución del programa lineal

$$\begin{cases} \inf \sum\limits_{j=1}^{n} c_j v_j \\ v_j \ge 0, \quad j = 1, ..., n, \quad \sum\limits_{j=1}^{n} a_{ij} v_j \ge b_i, \quad i = 1, ..., m. \end{cases} \tag{7}$$

Otro problema típico de programación lineal es el *problema del transporte*. Se consideran p almacenes en los que los niveles de stocks de un producto son s_i, $i = 1, ..., p$ y q destinos donde se desea disponer de una cantidad mínima del producto r_j, $j = 1, ..., q$. El coste por unidad del transporte desde el almacén i al destino j se designa por c_{ij}. Las variables de decisión son entonces v_{ij}, cantidad del stock i entregado en el destino j, y la decisión óptima es la que minimiza los costes de distribución, la cual es solución del problema

$$\begin{cases} \inf \sum\limits_{i,j} c_{ij} v_{ij} \\ v_{ij} \ge 0, \quad \sum\limits_{j=1}^{q} v_{ij} \le s_i, \quad \sum\limits_{i=1}^{p} v_{ij} \ge b_i, \quad i = 1, ..., p, \quad j = 1, ..., q. \end{cases} \tag{8}$$

Este ejemplo ilustra la clase más general de problemas típicos de investigación operativa, donde el coste y las restricciones, de igualdad o desigualdad, son funciones afines de las variables de decisión.

2.1.2. Ejemplo 2. Programación no lineal. Llamamos *programación no lineal* a un problema en dimensión finita del siguiente tipo

$$\begin{cases} \inf J(v) \\ F(v) \le 0, \quad \text{(es decir, } F_i(v) \le 0, \quad i = 1, ..., n), \end{cases} \tag{9}$$

donde las funciones J y $F_i : \mathbb{R}^n \to \mathbb{R}$ no son lineales. En este caso tendremos que hacer hipótesis sobre la diferenciabilidad de estas funciones para poder caracterizar las soluciones optimales.

2.1.3. Ejemplo 3. Programación convexa. Un problema del tipo

$$\begin{cases} \text{ínf } J(v) \\ v \in C \quad F(v) \leq 0, \end{cases} \tag{10}$$

donde las funciones $J, F_i : V \to \mathbb{R}$ son convexas, V es un espacio vectorial (real) y $C \subset V$ es un subconjunto convexo, se llama *programa convexo*. Las hipótesis realizadas aseguran que el conjunto de soluciones admisibles

$$U := \{v : v \in C, \quad F(v) \leq 0\}$$

es un conjunto convexo (o vacío).

2.1.4. Ejemplo 4. Programa cuadrático y mínimos cuadrados. Un caso particular de programa convexo es aquel en el que J es un funcional cuadrático definido por una matriz simétrica y definida positiva, Q, y en el que F es afín. En este caso nos referimos al problema como un *programa cuadrático*

$$\begin{cases} \text{ínf } \frac{1}{2} < v\,, Qv > + < c\,, v > \\ Av \leq b. \end{cases} \tag{11}$$

Una aplicación importante es la que concierne a la resolución numérica de sistemas lineales subdeterminados de la forma

$$Av = b,$$

donde A es una matriz $m \times n$ de rango $m < n$, en el cual se busca una *solución generalizada*, definida como solución del programa

$$\begin{cases} \text{ínf } \frac{1}{2} < v\,, Qv > \\ Av = b, \end{cases}$$

donde Q tiene un sentido físico determinado por el origen del problema. De manera similar planteamos la búsqueda de soluciones de sistemas sobre-determinados de la forma

$$Av = b,$$

donde A es una matriz $m \times n$ de rango $m > n$, y llamamos *solución de mínimos cuadrados* a la solución del programa cuadrático

$$\text{ínf } < Av - b\,, R(Av - b) >,$$

donde la matriz simétrica definida positiva, R, tiene un sentido físico.

2.2. Ejemplos en dimensión infinita: problemas del Cálculo Variacional. Muchos de los problemas en dimensión infinita pueden ser formulados en la forma de un *problema del cálculo de variaciones*. En éstos tenemos que

1. V es un espacio funcional a precisar, de funciones $v : \Omega \to \mathbb{R}^q$, donde Ω es una abierto, normalmente acotado, de $\mathbb{R}^n$, de frontera, Γ, regular.
2. Las restricciones sobre el conjunto de soluciones pueden imponerse bien sobre la frontera Γ, bien sobre el dominio Ω. Por ejemplo $u = 0$ en Γ, $u \geq \psi$ en Ω, etc.
3. Se considera un operador diferencial $P(D)$ bien definido sobre V (o, al menos, sobre el conjunto de soluciones admisibles, U). Normalmente $P(D)v = v'$.
4. El criterio $J : V \to \mathbb{R}$ es de la forma siguiente

$$J(v) := \int_\Omega L(x, v(x), v'(x))dx. \tag{12}$$

Naturalmente, las hipótesis sobre V y L deben asegurar la existencia de J sobre V, o al menos sobre U.

El problema es entonces

$$\begin{cases} \text{ínf } J(v) \\ v \in U \subset V. \end{cases} \tag{13}$$

A continuación daremos varios ejemplos, los cuales dividiremos atendiendo al carácter unidimensional (ejemplos 5-9) o multidimensional (ejemplos 6-14) del problema.

2.2.1. Ejemplo 5. Trayectoria en tiempo mínimo. Se trata de encontrar la trayectoria que conecta en un tiempo mínimo a los puntos (a, v_0) y (b, v_1), teniendo en cuenta que en cada punto (x, y) el módulo de la velocidad viene dado por la función $c(x, y) \geq 0$. El problema es entonces

$$\begin{cases} \text{ínf } J(v) := \displaystyle\int_a^b \frac{\sqrt{1 + |v'(x)|^2}}{c(x, v(x))} dx \\ v(a) = v_0, \quad v(b) = v_1. \end{cases} \tag{14}$$

Conviene precisar el espacio en el que buscaremos la solución v. Más adelante justificaremos por qué el espacio adecuado se trata del espacio de Sobolev

$$V := H^1(a, b).$$

A continuación veremos varios problemas correspondientes a la elección de la función $c(x, y)$.

5.1 Ley de refracción de la luz (Principio de Fermat-Huygens). Se trata de un problema de transmisión con discontinuidad entre dos medios. Supondremos que $a < 0$ y $b > 0$, correspondiendo $x = 0$ a la frontera entre ambos medios, y definimos

$$c(x, y) := \begin{cases} c_1 & \text{si } x < 0, \\ c_2 & \text{si } x > 0. \end{cases}$$

5.2 Braquistocrona. Supongamos que tenemos un cuerpo situado inicialmente en el punto $(a, v_0) := (0, 0)$ y consideremos la fuerza gravitatoria dirigida en el sentido positivo del eje OX. Sabemos que la velocidad de un cuerpo en caída libre sin rozamiento viene dada por

$$c(x, y) := \sqrt{2gx},$$

siendo g el módulo de la aceleración gravitatoria. El problema de la braquistocrona consiste en encontrar el perfil $v : [a, b] \to \mathbb{R}_+$ tal que el cuerpo llega al punto (b, v_1) en tiempo mínimo.

5.3 Superficie de revolución mínima. La superficie de revolución en torno al eje OX engendrada por la curva v y que une los puntos (a, v_0) y (b, v_1) tiene como superficie

$$J(v) := 2\pi \int_a^b v(x)\sqrt{1 + |v'(x)|^2}dx.$$

El perfil de la superficie de revolución mínima corresponde por tanto al problema (14) con

$$c(x, y) := \frac{1}{2\pi y}.$$

2.2.2. Ejemplo 6. Geodésicas. Consideremos una superficie regular de $\mathbb{R}^3$ dada por la carta

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v).$$

El elemento de arco es entonces

$$ds^2 := dx^2 + dy^2 + dz^2 = edu^2 + 2fdudv + gdv^2,$$

con

$$e := x_u^2 + y_u^2 + z_u^2, \quad f := x_u x_v + y_u y_v + z_u z_v, \quad g := x_v^2 + y_v^2 + z_v^2.$$

La geodésica, es decir, la curva de longitud mínima que une el punto (u_0, v_0) con el punto (u_1, v_1) viene dada por la función $v : [u_0, u_1] \to \mathbb{R}$ solución del problema de optimización

$$\left\{ \begin{array}{l} \text{ínf } J(v) := \displaystyle\int_{u_0}^{u_1} \sqrt{e + 2fv' + g|v'|^2}du \\ v(u_0) = v_0, \quad v(u_1) = v_1. \end{array} \right. \tag{15}$$

2.2.3. Ejemplo 7. Dinámica de sistemas con un número finito de grados de libertad. Un sistema mecánico con un número finito de grados de libertad, es decir, en el que la posición de una partícula en un instante dado viene dada por un vector $\mathbf{q} \in \mathbb{R}^n$, está caracterizado por una función L llamada *lagrangiana*, $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ tal que la trayectoria seguida para unir los puntos $(\mathbf{q}^0, t_0)$ y $(\mathbf{q}^1, t_1)$ es solución del problema siguiente

$$\left\{ \begin{array}{l} \text{ínf } J(\mathbf{q}) := \displaystyle\int_{t_0}^{t_1} L(\mathbf{q}(t), \mathbf{q}'(t), t)dt \\ \mathbf{q}(t_0) = \mathbf{q}^0, \quad \mathbf{q}(t_1) = \mathbf{q}^1. \end{array} \right. \tag{16}$$

La lagrangiana tiene la siguiente forma

$$L(\mathbf{q}, \mathbf{q}', t) := T(\mathbf{q}, \mathbf{q}', t) - U(\mathbf{q}, t),$$

con

$$T(\mathbf{q}, \mathbf{q}', t) := \frac{1}{2} \sum_{ij} a_{ij}(\mathbf{q}, t) \mathbf{q}'_i \mathbf{q}'_j,$$

donde T es una forma cuadrática definida positiva respecto $\mathbf{q}'$, que representa la energía cinética del sistema, y donde U representa la energía potencial del mismo.

2.2.4. Ejemplo 8. Criterio cuadrático. Un caso importante es el de un funcional cuadrático en un problema unidimensional

$$\left\{ \begin{array}{l} \text{ínf } J(v) := \frac{1}{2} \int_a^b (|v'|^2 + Mv) - \int_a^b fv \\ v \in H^1(a, b), \quad v(a) = v_0, \quad v(b) = v_1, \end{array} \right. \tag{17}$$

con $M := M(x) \geq 0$ en (a, b).

2.2.5. Ejemplo 9. Problema elemental del cálculo de variaciones. Finalmente, escribimos el problema típico de cálculo de variaciones (el Ejemplo 8 es un caso particular):

$$\left\{ \begin{array}{l} \text{ínf } J(v) := \int_a^b L(v(x), v'(x), x) dx \\ v(a) = v_0, \quad v(b) = v_1, \end{array} \right. \tag{18}$$

donde, en muchas aplicaciones, $v \in H^1(a, b)$, aunque también pueden imponerse restricciones más fuertes.

Acabamos esta sección con la introducción de algunos problemas multidimensionales.

2.2.6. Ejemplo 10. Problema de Plateau. Sea $\Omega \subset \mathbb{R}^n$ un dominio acotado de frontera, Γ, regular. Buscamos la superficie de $\mathbb{R}^{n+1}$, es decir, la aplicación $v : \Omega \to \mathbb{R}$ de superficie mínima y tomando en Γ unos valores predeterminados.

$$\left\{ \begin{array}{l} \text{ínf } J(v) := \int_\Omega \sqrt{1 + |\nabla v|^2} \\ v \in H^1(\Omega), \quad v = g_0 \quad \text{en } \Gamma. \end{array} \right. \tag{19}$$

Este problema proporciona por ejemplo, en el caso $n = 2$, la forma de equilibrio de una pompa de jabón que se apoya sobre el contorno (Γ, g_0).

2.2.7. Ejemplo 11. Equilibrio en mecánica de medios continuos. Consideremos sobre una membrana elástica homogénea e isótropa sobre Ω y sea $v : \Omega \to \mathbb{R}$ el desplazamiento de dicha membrana. La energía potencial de deformación es, salvo una constante (coeficiente de elasticidad), igual a la variación de la superficie, digamos

$$\int_\Omega \sqrt{1 + |\nabla v|^2}.$$

Si consideramos solamente deformaciones pequeñas, la energía potencial de deformación viene dada por

$$U_1 := \frac{1}{2}\int_\Omega |\nabla v|^2.$$

Si la membrana está sometida en todo punto a una fuerza de recuperación elástica, la energía potencial correspondiente es

$$U_2 := \frac{1}{2}\int_\Omega |v|^2.$$

Finalmente, si existe una fuerza exterior transversal, constante en tiempo y de densidad f aplicada a la membrana, entonces la energía potencial asociada es

$$U_3 := -\int_\Omega fv.$$

La membrana estará en equilibrio cuando las restricciones se verifiquen y cuando el criterio energético correspondiente sea minimizado. Según el caso, este criterio será

$$J_1 := \frac{1}{2}\int_\Omega (|\nabla v|^2 + |v|^2) - \int_\Omega fv,$$

o

$$J_2 := \frac{1}{2}\int_\Omega |\nabla v|^2 - \int_\Omega fv.$$

Las condiciones de frontera pueden ser muy variadas. Veamos algunos ejemplos en los que supondremos $f \in L^2$.

11.1 Problema de Dirichlet. Es el problema siguiente

$$\left\{\begin{array}{l} \text{ínf } J_2(v) \\ v \in H_0^1(\Omega). \end{array}\right. \tag{20}$$

Corresponde a las condiciones de frontera $v = 0$ en Γ.

11.2 Problema de Neumann.

$$\left\{\begin{array}{l} \text{ínf } J_1(v) \\ v \in H^1(\Omega). \end{array}\right. \tag{21}$$

11.3 Elasticidad con restricciones unilaterales.

$$\left\{\begin{array}{l} \text{ínf } J_1(v) \\ v \in H^1(\Omega), \quad v \geq 0 \quad \text{ctp en } \Gamma. \end{array}\right. \tag{22}$$

11.3 Equilibrio con obstáculo.

$$\left\{\begin{array}{l} \text{ínf } J_2(v) \\ v \in H_0^1(\Omega), \quad v \geq \psi \quad \text{ctp en } \Omega. \end{array}\right. \tag{23}$$

Para que las restricciones sean compatibles (es decir, $U \neq \emptyset$) hay que suponer, por ejemplo, que $\psi \in L^2(\Omega)$ y $\psi \leq 0$ en Γ.

2.3. Problemas isoperimétricos. El origen histórico de esta apelación viene del siguiente ejemplo.

2.3.1. Ejemplo 12. Problema isoperimétrico elemental (problema de Pappus). De entre todas las curvas de longitud ℓ dada, que unen el punto $(0,0)$ con un punto variable $(0,\xi)$, encontrar aquella que, junto con el eje OX, encierra una superficie máxima.

$$\begin{cases} \text{ínf } J(v,\xi) := -\displaystyle\int_0^\xi v \\ v(0)=0, \quad v(\xi)=0, \quad v \geq 0, \quad \displaystyle\int_0^\xi \sqrt{1+|v'|^2} = \ell. \end{cases} \tag{24}$$

Generalizando, llamaremos un *problema isoperimétrico* a todo problema del tipo

$$\begin{cases} \text{ínf } J(v) \\ v \in U \subset V, \quad F_i(v) = c_i, \quad i = 1, ..., m, \end{cases} \tag{25}$$

donde J y F_i son funcionales de la misma forma que en (12) y donde $U \subset V$ expresa las condiciones de frontera que restringen a v.

2.3.2. Ejemplo 13. Equilibrio de una cadena. Consideremos una cadena colgante homogénea de longitud ℓ suspendida de los puntos (a, v_0) y (b, v_1). Su posición de equilibrio $v : [a,b] \to \mathbb{R}$ corresponde a la altitud mínima de su centro de gravedad (energía potencial mínima). El problema es

$$\begin{cases} \text{ínf } J(v) := \displaystyle\int_a^b v\sqrt{1+|v'|^2} \\ v(a) = v_0, \quad v(b) = v_1 \quad \displaystyle\int_a^b \sqrt{1+|v'|^2} = \ell. \end{cases} \tag{26}$$

Este problema es similar al formulado en el Ejemplo 5.3, pero con una restricción adicional de tipo isoperimétrico.

2.3.3. Ejemplo 14. Autovalores y autofunciones. Sea $J(v) := a(v,v)$, donde a es una forma bilineal, simétrica, continua y definida positiva sobre el espacio de Hilbert V. El siguiente problema es de tipo isoperimétrico.

$$\begin{cases} \sup J(v) \\ v \in V, \quad < v, v >= 1. \end{cases} \tag{27}$$

El valor óptimo resulta ser el mayor autovalor de a y la solución óptima es la autofunción correspondiente.

Capítulo 3

Métodos directos del Cálculo Variacional

En este capítulo vamos a estudiar las técnicas básicas que nos permiten asegurar la existencia y unicidad del mínimo de un funcional (función definida sobre un espacio de dimensión infinta con valores a $\mathbb{R}$). La idea fundamental es la extensión del Teorema de Weierstrass a funciones definidas en espacios de dimensión infinita.

Las hipótesis del Teorema de Weiertrass atañen a la función que se desea minimizar (semicontinuidad inferior) y al conjunto en el cual se busca el mínimo (compacto). En espacios de dimensión finita estas hipótesis son relativamente fáciles de comprobar dado que la compacidad de un conjunto es equivalente a que el mismo sea cerrado y acotado, y la continuidad puede deducirse frecuentemente de un análisis directo de la función a minimizar.

Sin embargo, el Teorema de Riesz (véase Teorema ???) establece que la bola unidad cerrada de un espacio de Banach es compacta si y solo si la dimensión del espacio es finita. Puesto que este criterio de compacidad falla en el caso de dimensión infinita, se impone la investigación de nuevas condiciones sobre los subconjuntos de espacios de dimensión infinita y sobre los funcionales definidos en estos espacios que, de algún modo, nos permitan enunciar un resultado que generalize el Teorema de Weierstrass.

Las aplicaciones clásicas del cálculo variacional conducen de un modo natural al estudio de la topología débil asociada a un espacio de Banach de dimensión infinita. En efecto, puesto que los conjuntos cerrados y acotados en la topología fuerte no son compactos, uno puede esparar que si reduce la cantidad de abiertos mediante la introducción de una nueva topología, la cantidad de cerrados y, por tanto, de compactos, aumente. Y esto resulta ser así. En particular, cualquier subconjunto acotado de un espacio de Banach reflexivo es relativamente compacto respecto la topología débil.

El problema que surge a continuación es el de la continuidad (respecto la topología débil) del funcional a minimizar. Claramente, al introducir una topología con menos abiertos, la cantidad de funciones continuas también disminuye y, así, por ejemplo la norma asociada a la topología fuerte no es una función continua respecto la topología débil. Cobra especial importancia en este contexto la noción de semicontinuidad inferior.

Finalmente, observemos que aunque, como veremos, la introducción de la topología débil y de los funcionales semicontinuos inferiormente respecto dicha topología nos permiten asegurar la existencia de un mínimo sobre cualquier conjunto acotado y cerrado respecto la topología débil, la verificación práctica de estas propiedades dista de ser sencilla, como lo es en el caso de la topología fuerte. Por ello, una de las cuestiones centrales de este capítulo será la búsqueda de condiciones expresadas respecto la topología fuerte que impliquen las correspondientes respecto la topología débil. En este contexto la convexidad de conjuntos y funciones jugará un papel fundamental.

1. Topología débil

1.1. Definiciones y resultados básicos sobre la topología débil. En todo espacio de Banach, V, puede definirse una topología menos fina que la topología engendrada por la norma, y que, por tanto, posee más compactos. Es la *topología débil* $\sigma(V, V')$.

Sea V un espacio de Banach y sea $J \in V'$ (es decir, $J : V \to \mathbb{R}$ lineal y continua). Se designa por $\varphi_J : V \to \mathbb{R}$ a la aplicación dada por $\varphi_J(u) :=< J\,,\, u>$. Cuando J recorre V' se obtiene una familia $(\varphi_J)_{J\in V'}$ de aplicaciones de V en $\mathbb{R}$.

DEFINICIÓN 1.1. *La **topología débil** $\sigma(V, V')$ sobre V es la topología menos fina sobre V que hace continuas a todas las aplicaciones $(\varphi_J)_{J\in V'}$.*

A continuación veremos algunas propiedades importantes de la topología débil.

PROPOSICIÓN 1.2. *Sea $u_0 \in V$. Se obtiene una base de entornos de u_0 para la topología $\sigma(V, V')$ al considerar todos los conjuntos de la forma*

$$U := \{u \in V : | < J_i\,,\, u - u_0 > | < \varepsilon, \quad \text{para todo } i \in I\}\,, \tag{28}$$

donde I es finito, $J_i \in V'$ y $\varepsilon > 0$.

DEMOSTRACIÓN. Véase Brezis [**?**], Proposición III.4. □

Notación. Dada una sucesión $\{u_n\} \subset V$, se designa por

$$u_n \rightharpoonup u$$

a la convergencia débil de u_n a u en la topología débil de $\sigma(V, V')$. A menudo escribiremos u_n en vez de $\{u_n\}$ para denotar una sucesión.

TEOREMA 1.3. *Sea u_n una sucesión en V. Se tiene:*

1. *$u_n \rightharpoonup u$ en $\sigma(V, V')$ si y sólo si $< J\,,\, u_n > \to < J\,,\, u >$ para todo $J \in V'$.*
2. *Si $u_n \to u$ fuertemente entonces $u_n \rightharpoonup u$ débilmente en $\sigma(V, V')$.*
3. *Si $u_n \rightharpoonup u$ débilmente en $\sigma(V, V')$ entonces $\|u_n\|$ está acotada y $\|u\| \leq \liminf \|u_n\|$.*
4. *Si $u_n \rightharpoonup u$ débilmente en $\sigma(V, V')$ y $J_n \to J$ fuertemente en V'(es decir, $\|J_n - J\|_{V'} \to 0$) entonces $< J_n\,,\, u_n > \to < J\,,\, u >$.*

DEMOSTRACIÓN. 1. Por definición de topología débil, $\varphi_J(u) :=< J\,,\, u >$ es continua para toda $J \in V'$, de modo que si $u_n \rightharpoonup u$ se sigue $< J\,,\, u_n > \to < J\,,\, u >$ para toda $J \in V'$.

Recíprocamente, sea U un entorno de u de la forma (28). Para cada $i \in I$ existe un entero N_i tal que $| < J_i\,,\, u_n > | < \varepsilon$, para $n \geq N_i$. Sea $N = \max_{i\in I} N_i$. Resulta entonces que $u_n \in U$ para cada $n \geq N$.

2. Se sigue de 1. ya que $| < J\,,\, u_n > - < J\,,\, u > | \leq \|J\| \|u_n - u\|$.
3. Véase Brezis [**?**], Proposición III.5.
4. Se tiene

$$| < J_n\,,\, u_n > - < J\,,\, u > | \leq | < J_n - J\,,\, u_n > | + | < J\,,\, u_n - u > |$$
$$\leq \|J_n - J\| \|u_n\| + | < J\,,\, u_n - u > |,$$

y se concluye por 1. y 3. □

PROPOSICIÓN 1.4. *V es de dimensión finita si y sólo si la topología fuerte y la débil coinciden.*

DEMOSTRACIÓN. Supongamos que V es de dimensión infinita. Tenemos entonces que el conjunto $S = \{u \in V : \|u\| = 1\}$ nunca es cerrado en la topología débil $\sigma(V, V')$ (véase Ejemplo (1.6)). Sin embargo, S es claramente cerrado en la topología fuerte. Por tanto, $V\S$ es abierto en la topología fuerte pero no en la débil. De hecho, la topología débil **siempre** tiene menos abiertos que la topología fuerte.

Recíprocamente, supongamos que V es de dimensión finita. Hemos de comprobar que todo abierto fuerte es un abierto débil. Sea $u_0 \in V$ y sea U un entorno de u_0 en la topología fuerte. Hemos de construir un entorno W de u_0 en la topología débil tal que $W \subset U$. O dicho de otro modo, encontrar un subconjunto finito $(J_i)_{i \in I}$ de V' y un $\varepsilon > 0$ tales que

$$W := \{u \in V : | < J_i\,,\, u - u_0 > | < \varepsilon \quad \text{para todo } i \in I\} \subset U.$$

Supongamos que $B(u_0, r) \subset U$. Tomemos una base $\{e_1, e_2, \ldots, e_n\}$ de V con $\|e_i\| = 1$ (siendo n la dimensión de V). Para todo $u \in V$ se tiene la descomposición $u = \sum_{i=1}^{n} u_i e_i$; las aplicaciones $u \to u_i$ definen n formas lineales continuas sobre V que denotamos por J_i. Se tiene entonces

$$\|u - u_0\| \leq \sum_{i=1}^{n} | < J_i\,,\, u - u_0 > | < n\varepsilon$$

para $u \in W$. Eligiendo $\varepsilon := r/n$ se obtiene $W \subset U$. □

OBSERVACIÓN 1.5. *Los abiertos de la topología débil son también abiertos de la topología fuerte. Cuando V es de dimensión infinita, la topología débil es estrictamente menos fina que la fuerte, es decir, existen abiertos en la topología fuerte que no son abiertos en la topología débil. He aquí un ejemplo:*

EJEMPLO 1.6. *El conjunto $S = \{u \in V : \|u\| = 1\}$ **nunca** es cerrado en la topología débil $\sigma(V, V')$. Más exactamente, demostraremos que*

$$\bar{S}^\sigma = \{u \in V : \|u\| \leq 1\}\,,$$

siendo $\bar{S}^\sigma$ el cierre de S en la topología $\sigma(V, V')$.

Sea $u_0 \in V$ *con* $\|u_0\| < 1$. *Comprobemos que* $u_0 \in \bar{S}^\sigma$. *Dado un entorno* U *de* u_0 *en* $\sigma(V, V')$ *vamos a probar que* $U \cap S \neq \emptyset$. *Siempre se puede suponer que* U *es de la forma*

$$U := \{u \in V : | < J_i \, , \, u - u_0 > | < \varepsilon, \quad i = 1, 2, \ldots, n\} \, ,$$

con $\varepsilon > 0$ *y* $J_i \in V'$, $i = 1, 2, \ldots, n$. *Fijemos* $v_0 \in V$, $v_0 \neq 0$ *tal que*

$$< J_i \, , \, v_0 >= 0 \quad \textit{para todo } i = 1, 2, \ldots, n.$$

Un tal v_0 *existe ya que de lo contrario, la aplicación* $\varphi : V \to \mathbb{R}^n$ *definida por*

$$\varphi(w) := (< J_1 \, , \, w >, \ldots, < J_n \, , \, w >)$$

sería inyectiva y φ *sería un isomorfismo de* V *en* $\mathbb{R}^m$, *de donde* V *sería de dimensión* n.

La función $g(t) := \|u_0 + tv_0\|$ *es continua en* $[0, \infty)$, *con*

$$g(0) < 1 \quad y \quad \lim_{t \to \infty} g(t) = +\infty.$$

Así pues, existe $t_0 > 0$ *tal que* $\|u_0 + tv_0\| = 1$. *Por consiguiente,* $u_0 + tv_0 \in U \cap S$. *Con lo que hemos probado que*

$$S \subset \{u \in V : \|u\| \leq 1\} \subset \bar{S}^\sigma.$$

De donde se deduce el resultado si se sabe que $\{u \in V : \|u\| \leq 1\}$ *es cerrado en la topología débil, lo cual demostraremos en el Teorema 1.9.*

La interpretación geométrica de esta construcción es la siguiente. En dimensión infinita, todo entorno U *de* u_0 *de la topología débil contiene una recta que pasa por* u_0 *(e incluso un* enorme *espacio afín que contiene a* u_0*).*

EJEMPLO 1.7. *Supongamos ahora que* V *es un espacio de Hilbert de dimensión infinita y sea* u_n *un sistema numerable ortonormal de* V. *Claramente* u_n *no tiene ninguna subsucesión fuertemente convergente, pero sin embargo*

$$u_n \rightharpoonup 0.$$

En efecto, por la desigualdad de Bessel tenemos que

$$\sum_{n=1}^{\infty} | < u \, , \, u_n > |^2 \leq \|u\|^2 \quad \textit{para toda } u \in V,$$

y por ser la serie convergente deducimos que el término general tiende a cero.

EJEMPLO 1.8. ***Convergencia de medias.*** *Consideremos el espacio de Banach* $L^p(\Omega)$, *con* $1 < p < \infty$ *y sea* u_n *una sucesión de* $L^p(\Omega)$ *tal que*

$$u_n \rightharpoonup u \quad \textit{débilmente en } L^p(\Omega). \tag{29}$$

Entonces, si $E \subset \Omega$ *es un conjunto medible y acotado, la función indicatriz de* E, 1_E *es un elemento de* $L^{p'}(\Omega)$, *con lo que*

$$\int_E u_n = \int_\Omega u_n 1_E \to \int_\Omega u 1_E = \int_E u,$$

es decir, las medias de las funciones u_k sobre cualquier conjunto medible y acotado E convergen a la media del límite débil, u, sobre E. Un problema al que nos enfrentamos continuamente en las aplicaciones es que la convergencia en media no implica la convergencia en norma y, ni siquiera, la convergencia en casi todo punto. Puede pasar, por ejemplo, que la sucesión u_n converja débilmente a u a través de oscilaciones irregulares, de muy alta frecuencia o incluso no acotadas.

En particular, observemos que (29) no implica $J(u_n) \rightharpoonup J(u)$ para cualquier funcional no lineal J. Para comprobar esto, tomemos $a, b, \lambda \in \mathbb{R}$, con $a < b$ y $0 < \lambda < 1$, de modo que

$$J(\lambda a + (1-\lambda)b) \neq \lambda J(a) + (1-\lambda)J(b).$$

Sea $\Omega = (0,1)$ y

$$u_n(x) = \begin{cases} a & \textit{si} \quad \dfrac{j}{n} \leq x \leq \dfrac{j+\lambda}{n}, \quad j = 0, \dots, n-1 \\ b & \textit{en otro caso.} \end{cases}$$

*Entonces $u_n \overset{*}{\rightharpoonup} u \equiv \lambda a + (1-\lambda)b$ en $L^\infty(\Omega)$, es decir, (**definición de convergencia débil** $*$), para toda $\varphi \in L^1(\Omega)$ se tiene*

$$\int_\Omega u_n \varphi \to \int_\Omega u \varphi.$$

En efecto, para toda $\varphi \in C_0(\Omega)$ se tiene, por el Teorema del Valor Medio, que existen $\hat{x}_j \in [j/n, (j+\lambda)/n]$ para $j = 0, \dots, n-1$ tales que

$$\int_\Omega u_n \varphi = \sum_{j=0}^{n-1} \int_{j/n}^{(j+\lambda)/n} \varphi(x)dx = \sum_{j=0}^{n-1} \frac{\lambda}{n} \varphi(\hat{x}_j),$$

donde hemos fijado $a = 1$ y $b = 0$, por comodidad. Consideremos ahora la partición de $[0,1]$ dada por $x_j = j/n$, para $j = 0, \dots, n$. La integral de Riemann de φ viene dada por

$$\int_\Omega \varphi = \lim_n \sum_{j=0}^{n-1} \varphi(\bar{x}_j)|x_{j+1} - x_j| = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(\bar{x}_j),$$

donde $\bar{x}_j \in [x_j, x_{j+1}]$. Puesto que $[j/n, (j+\lambda)/n] \subset [x_j, x_{j+1}]$, podemos tomar $\bar{x}_j = \hat{x}_j$, de modo que resulta

$$\lim_n \int_\Omega u_n \varphi = \lambda \lim_n \sum_{j=0}^{n-1} \frac{1}{n} \varphi(\hat{x}_j) = \lambda \int_\Omega \varphi,$$

para toda $\varphi \in C_0(\Omega)$. Un argumento de densidad permite probar la misma propiedad para toda $\varphi \in L^1(\Omega)$. Por otra parte,

$$J(u_n(x)) = \begin{cases} J(1) & \textit{si} \quad \dfrac{j}{n} \leq x \leq \dfrac{j+\lambda}{n}, \quad j = 0, \dots, n-1 \\ J(0) & \textit{en otro caso,} \end{cases}$$

y, procediendo como para la sucesión u_n, hallamos que

$$J(u_n) \overset{*}{\rightharpoonup} \lambda J(1) \neq J(\lambda) \equiv J(u).$$

1.2. Algunas propiedades que relacionan la topología fuerte y débil. Puesto que la topología débil es difícil de manejar, a continuación exponemos algunos resultados que permiten, por medio de la topología fuerte, deducir ciertas propiedades sobre la débil.

Todo conjunto cerrado en la topología débil $\sigma(V, V')$ es cerrado en la topología fuerte, pero el recíproco no es cierto en general, véase Proposición 1.4. Sin embargo, hay un caso importante, el de los conjuntos convexos, para el que ambas nociones coinciden.

TEOREMA 1.9. *Sea $C \subset V$ convexo. Entonces C es débilmente cerrado en $\sigma(V, V')$ si y sólo si es fuertemente cerrado.*

DEMOSTRACIÓN. Véase Teorema III.7 de Brezis [**?**]. □

OBSERVACIÓN 1.10. *El teorema anterior implica que un convexo cerrado coincide con la intersección de los semiespacios cerrados que lo contienen. Por otra parte, también implica que si una sucesión u_n converge débilmente a u, entonces existe una sucesión de combinaciones convexas de los u_n que converge fuertemente a u, véase el Corolario 1.18.*

Introducimos a continuación la definición de funcional semicontinuo inferiormente.

DEFINICIÓN 1.11. *Sea V un espacio de Hausdorff. Se dice que $J : V \to \mathbb{R}$ es **semicontinua inferiormente (sci)** en $u \in V$ si, para todo $\varepsilon > 0$ existe un entorno U de u tal que*

$$J(v) \geq J(u) - \varepsilon \quad \textit{para todo } v \in U.$$

O, equivalentemente,

$$J(u) \leq \liminf_{v \to u} J(v).$$

Diremos que J es sci *en V si es sci en todo punto de V.*

OBSERVACIÓN 1.12. *Es fácil ver que J es sci en V si y sólo si el conjunto*

$$\{v \in V : J(v) \leq \alpha\}$$

es cerrado para cualquier $\alpha \in \mathbb{R}$.

Un ejemplo de función que no es continua en $x = 0$ pero que es sci es

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

véase Figura 1.

FIGURA 1. La función $(1+x^2)sen(1/x)$

COROLARIO 1.13. *Sea $J : V \to (-\infty, \infty]$ una función convexa y sci para la topología fuerte. Entonces J es sci para la topología débil $\sigma(V, V')$. En particular, si $u_n \rightharpoonup u$ entonces*

$$J(u) \leq \liminf_n J(u_n).$$

DEMOSTRACIÓN. Es suficiente comprobar que para todo $\alpha \in \mathbb{R}$ el conjunto

$$A := \{u \in V : J(u) \leq \alpha\}$$

es cerrado en $\sigma(V, V')$. A es convexo (por ser J convexa) y A es fuertemente cerrado (ya que J es sci para la topología fuerte). Según el Teorema 1.9 el conjunto A es también cerrado para la topología débil. □

OBSERVACIÓN 1.14. *Se obtiene en particular que si $u_n \rightharpoonup u$ en $\sigma(V, V')$ entonces*

$$\|u\| \leq \liminf_n \|u_n\|.$$

En efecto, el funcional $J(u) := \|u\|$ es convexo y continuo para la topología fuerte (luego fuertemente sci). Por consiguiente, J es débilmente sci.

DEFINICIÓN 1.15.

1. *Se dice que un espacio métrico es **separable** si posee un subconjunto numerable y denso.*
2. *Sea V un espacio de Banach y sea I la inyección canónica de V en V''. Se dice que V es **reflexivo** si $I(V) = V''$.*

EJEMPLO 1.16.

1. *El espacio L^p, con $1 < p < \infty$ es reflexivo y separable. El espacio dual correspondiente es $L^{p'}$, con $p' = p/(p-1)$*
2. *L^1 es un espacio separable, pero no reflexivo. Su espacio dual es L^∞.*
3. *L^∞ no es ni reflexivo ni separable. El espacio dual de L^∞ contiene estrictamente a L^1*

TEOREMA 1.17. *Sea V un espacio de Banach reflexivo. Sea $K \subset V$ un subconjunto convexo, cerrado y acotado. Entonces K es compacto en la topología débil $\sigma(V, V')$. En particular, si u_n es una sucesión acotada en V entonces existe una subsucesión u_{n_j} que converge débilmente a un elemento de V.*

DEMOSTRACIÓN. Para el caso en el que V es un espacio de Banach reflexivo, véase [**?**], Corolario III.19. Demostraremos aquí que si u_n es una sucesión acotada en un espacio de Hilbert separable, V, entonces existe una subsucesión u_{n_j} que converge débilmente a un elemento de V.

Tomemos un conjunto, v_k, numerable y denso en V. Como $| < u_n \, , \, v_1 > | \leq \|u_n\| \|v_1\|$ para todo n, la sucesión de números reales $< u_n \, , \, v_1 >$ está acotada y, por tanto, existe una subsucesión de u_n, que denotamos por u_{1n}, que converge a algún real a_1. Análogamente podemos probar que existe una subsucesión u_{2n} de u_{1n} tal que $< u_{2n} \, , \, v_2 > \to a_2$ cuando $n \to \infty$. Continuando este proceso, tenemos que

$$\begin{aligned} < u_{11} \, , \, v_1 >, < u_{12} \, , \, v_1 >, < u_{13} \, , \, v_1 >, \ldots \to & \quad a_1, \\ < u_{21} \, , \, v_2 >, < u_{22} \, , \, v_2 >, < u_{23} \, , \, v_2 >, \ldots \to & \quad a_2, \\ & \quad \vdots \end{aligned}$$

La sucesión diagonal w_n definida por $w_n := u_{nn}$ tiene la propiedad de que

$$< w_n \, , \, v_k > \to a_k \quad \text{cuando } n \to \infty \quad \text{para todo } k. \tag{30}$$

Además, existen números $a(v)$ tales que

$$< w_n \, , \, v > \to a(v) \quad \text{cuando } n \to \infty \quad \text{para todo } v \in V. \tag{31}$$

En efecto, esto se sigue de que la sucesión w_n es una sucesión de Cauchy:

$$| < w_n - w_m \, , \, v > | \leq | < w_n - w_m \, , \, v - v_k > | + | < w_n - w_m \, , \, v_k > | < \varepsilon$$

para ciertos v_k y para todos $n, m \geq n_0(\varepsilon)$. Para esta acotación hemos usado que la sucesión w_n está acotada, que el conjunto v_k es denso en V y la propiedad (30).

Obviamente, la aplicación $v \to a(v)$ es lineal y por (31)

$$|a(v)| \leq \|v\| \sup_n \|w_n\| \quad \text{para todo } v \in V,$$

luego $a(v)$ es también continua. Aplicando el Teorema de Riesz[1] deducimos la existencia de un $w \in V$ tal que

$$a(v) = < w \, , \, v > \quad \text{para todo } v \in V.$$

Por (31), $< v \, , \, w_n > \to < v \, , \, w >$ cuando $n \to \infty$ para todo $v \in V$. Por tanto

$$w_n \rightharpoonup w \quad \text{cuando } n \to \infty.$$

[1]Teorema de Riesz: Sea V un espacio de Hilbert. $J \in V'$ si y sólo si existe $u \in V$ tal que $J(v) = < u \, , \, v >$ para todo $v \in V$.

□

COROLARIO 1.18. *El límite, u, de cada subsucesión débilmente convergente de u_n es un elemento del cierre de la envolvente convexa del conjunto* $\{u_1, u_2, \ldots\}$.

DEMOSTRACIÓN. Lo demostraremos en el caso de espacios de Hilbert. Sea u_n una sucesión acotada en el espacio de Hilbert V tal que $u_n \rightharpoonup u$. Nos es suficiente con probar que existe un subconjunto de índices $n_1 < n_2 < \ldots$ tal que

$$k^{-1}(u_{n_1} + \ldots + u_{n_k}) \to u \quad \text{cuando } n \to \infty.$$

Sustituyendo u_n por $u_n - u$, podemos suponer que $u = 0$. Pongamos primero $n_1 := 1$. Por la convergencia débil tenemos que

$$| < u_{n_1}, u_n > | \to 0 \quad \text{cuando } n \to \infty,$$

de modo que existe un índice $n_2 > n_1$ tal que $| < u_{n_1}, u_{n_2} > | \leq 2^{-1}$. Continuando de esta manera, obtenemos un conjunto de índices tal que

$$| < u_{n_1}, u_{n_k} > | \leq (k-1)^{-1}, \ldots, | < u_{n_{k-1}}, u_{n_k} > | \leq (k-1)^{-1},$$

para todo $k \geq 3$. Como $\|u_n\| \leq C$ para todo n, siendo C una constante positiva, deducimos

$$\begin{aligned}
\|k^{-1}(u_{n_1} + \ldots + u_{n_k})\|^2 &\leq k^{-2}\Big\{\sum_{j=1}^{k} \|u_{n_j}\|^2 + 2\sum_{j=1}^{k-1} | < u_{n_j}, u_{n_k} > | \\
&\quad + 2\sum_{j=1}^{k-2} | < u_{n_j}, u_{n_{k-1}} > | + \ldots\Big\} \\
&\leq k^{-2}(kC^2 + 2(k-1)(k-1)^{-1} + 2(k-2)(k-2)^{-1} + \ldots + 2) \\
&\leq k^{-1}(C^2 + 2) \to 0 \quad \text{cuando } k \to \infty.
\end{aligned}$$

□

2. El Teorema de Weierstrass Generalizado a Espacios de Banach

El siguiente resultado es una generalización del Teorema de Weierstrass.

DEFINICIÓN 2.1. *Sea $J : U \subset V \to \mathbb{R}$ un funcional sobre un subconjunto U de un espacio real normado V. Entonces*

1. *J es **débil secuencialmente semicontinuo inferiormente** si para todo $u \in U$ y toda sucesión $u_n \in V$ tales que $u_n \rightharpoonup u$ cuando $n \to \infty$ se tiene*

$$J(u) \leq \liminf_n J(u_n).$$

2. *J es **coercivo** si*

$$J(u) \to \infty \quad \textit{cuando } \|u\| \to \infty \quad \textit{en } U.$$

para todo $\alpha \in (0,1)$ y para todo $u, v \in U$.

TEOREMA 2.2. *Sea V un espacio de Banach reflexivo y $U \subset V$ un subconjunto cerrado, convexo y no vacío. Supongamos que el funcional $J : U \to \mathbb{R}$ es débil secuencialmente semicontinuo inferiormente y que si U no está acotado, entonces J es coercivo.*

Entonces existe un punto de mínimo $u \in U$ del funcional J. Si, además, J es estrictamente convexo entonces el punto de mínimo es único.

DEMOSTRACIÓN.
Paso 1. Supongamos que U está acotado y sea

$$\gamma := \inf_{u \in U} J(u).$$

Se tiene $-\infty \leq \gamma < \infty$. Existe una sucesión (*minimizante*) u_n en U tal que

$$J(u_n) \to \gamma \quad \text{cuando } n \to \infty.$$

Como U está acotado, la sucesión u_n está acotada, y por el Teorema 1.17, existe una subsucesión débilmente convergente, que seguimos denotando por u_n, tal que $u_n \rightharpoonup u$. El Corolario 1.18 asegura que u pertenece al cierre de la envolvente convexa de $\{u_1, u_2, \ldots\}$. Por tanto, $u \in U$. Finalmente, como J es débil secuencialmente sci,

$$J(u) \leq \liminf_n J(u_n) = \gamma,$$

lo cual implica que $J(u) = \gamma$.
Paso 2. Supongamos que U no está acotado y fijemos $v \in U$. Como $J(u) \to \infty$ cuando $\|u\| \to \infty$, existe un $r > 0$ tal que

$$J(u) > J(v) \quad \text{para todo } u \in V \text{ tal que } \|u\| > r, \tag{32}$$

y el conjunto $M_r := \{u \in U : \|u\| \leq r\}$ es no vacío. Por (32), toda solución u del problema

$$\min_{u \in M_r} J(u)$$

es también una solución del problema de minimización en todo U. Como M_r es cerrado, convexo y acotado, basta con aplicar el Paso 1 de esta demostración para conluir.

Para probar la unicidad del mínimo cuando J es estrictamente convexo, supongamos que existen dos soluciones u y v. Entonces $\frac{1}{2}(u+v) \in U$, y por tanto,

$$J(\frac{1}{2}(u+v)) < \frac{1}{2}J(u) + \frac{1}{2}J(v) = J(u),$$

que contradice el hecho de que u sea un punto de mínimo para J en U. $\square$

Los siguientes lemas dan una caracterización de la débil secuencialmente sci de un funcional J en términos de los conjuntos $\{u : J(u) \leq \lambda\}$.

LEMA 2.3. *Sea V un espacio de Banach y $U \subset V$ un subconjunto débil (secuencialmente) cerrado. Un funcional $J : U \to \mathbb{R}$ es débil (secuencialmente) sci en U si y sólo si los conjuntos*

$$M_\lambda = \{u \in U : J(u) \leq \lambda\}$$

son débil (secuencialmente) cerrados para todo $\lambda \in \mathbb{R}$.

DEMOSTRACIÓN. Demostraremos el resultado para convergencia secuencial y cierre secuencial. Para el caso general, véase [**?**].

Sea J débil secuencialmente sci. Para $\lambda \in \mathbb{R}$, sea u_n una sucesión débilmente convergente en M_λ. Entonces, por el Teorema 1.17 $u_n \rightharpoonup u$, para algún $u \in U$, puesto que U es débil secuencialmente cerrado. Se sigue del Corolario 1.13 que

$$J(u) \leq \liminf_n J(u_n) \leq \lambda,$$

y, por tanto, $u \in M_\lambda$, es decir, M_λ es débil secuencialmente cerrado.

Recíprocamente, supongamos que los conjuntos M_λ son débil secuencialmente cerrados para todo $\lambda \in \mathbb{R}$. Si J no fuera débil secuancialmente sci en U, existiría un punto $u \in U$ y una sucesión u_n en U tal que

1. $u_n \rightharpoonup u$ cuando $n \to \infty$,
2. $\liminf_n J(u_n) < J(u)$.

Entonces existe también un $\lambda \in \mathbb{R}$ con

$$\liminf_n J(u_n) < \lambda < J(u)$$

y, además, una subsucesión u_{n_j} en M_λ que, por (i) converge débilmente a u. Como M_λ es débil secuencialmnete cerrado, $u \in M_\lambda$, es decir, $J(u) \leq \lambda$, lo cual es una contradicción. $\square$

DEFINICIÓN 2.4. *Sea V un espacio de Banach y $U \subset V$. Un funcional $J : U \to \mathbb{R}$ es* ***cuasi-convexo*** *en U si M_λ es convexo para todo $\lambda \in \mathbb{R}$.*

Observemos que si J es convexo entonces J es cuasi-convexo.

LEMA 2.5. *Sea V un espacio de Banach reflexivo y $U \subset V$ un subconjunto convexo y cerrado. Si $J : U \to \mathbb{R}$ es un funcional sci y cuasi-convexo entonces J es débil secuencialmente sci.*

DEMOSTRACIÓN. Procedemos por reducción al absurdo. Supongamos que existe un elemento $u \in U$ y una sucesión u_n débilmente convergente a u tal que

$$J(u) > \lim_n J(u_n).$$

Entonces, existe un $r \in \mathbb{R}$ tal que $r < J(u)$ y $u_n \in M_r$ para $n \geq n_0$. Como M_r es cerrado y convexo, se sigue de que $u_n \rightharpoonup u$ cuando $n \to \infty$ que $u \in M_r$, y por tanto, que $J(u) \leq r$, lo cual es una contradicción. $\square$

3. El método de aproximación de Ritz

Sea $U \subset V$ un subconjunto de un espacio topológico y $f : U \to \bar{\mathbb{R}}$ una función definida en U. Toda sucesión u_n de U para la que

$$\lim_n J(u_0) = \inf \{J(u) : u \in U\}$$

se denomina **sucesión minimizante**. Esta definición no asume que la sucesión converje.

Para que una sucesión minimizante exista basta con que se cumplan dos condiciones:

1. Que haya puntos $u \in U$ tales que $J(u) < \infty$,
2. Que J esté acotada inferiormente, es decir, que

$$\text{ínf}\,\{J(u) : u \in U\} = \mu > -\infty.$$

Sin embargo, esto no resuelve el problema de minimización. En efecto, una solución del problema de minimización es un punto $u_0 \in U$ tal que

$$J(u_0) = \mu = \operatorname*{ínf}_{u \in M} J(u).$$

Asumiendo que las condiciones (i) y (ii) se cumplen y que, por lo tanto, existe al menos una sucesión minimizante u_n, esta sucesión no converge necesariamente e, incluso aunque lo haga, no está claro que

$$\lim_n J(u_n) = J(\lim_n u_n).$$

De modo que para resolver el problema de minimización necesitamos dar los siguientes pasos:

1. Construir una sucesión minimizante u_n.
2. Demostrar que esta sucesión converge a un punto $u_0 \in U$.
3. Justificar la ecuación

$$\lim_n J(u_n) = J(\lim_n u_n).$$

El teorema de Weierstrass generalizado nos proporciona condiciones suficientes bajo las cuales nos garantizan que los tres pasos pueden llevarse a cabo. Sin embargo, no nos dice cómo hacerlo, es decir, no nos proporciona un método práctico para encontrar el punto en el que se realiza el mínimo (o, al menos, una aproximación al mismo).

El método de aproximación de Ritz viene a solventar este problema en ciertas condiciones particulares: en espacios de Hilbert separables. El método consiste en una aproximación al problema infinito-dimensional mediante una sucesión de problemas finito-dimensionales, que pueden ser resueltos sin dificultad. Obtenemos así una sucesión de funciones que son puntos minimizantes de los problemas finito-dimensionales. Se trata entonces de ver que esta sucesión es una sucesión minimizante y que converge a un punto de mínimo del problema infinito-dimensional.

El método de Ritz tiene varias versiones, entre las que es especialmente importante el método de Galerkin, ampliamente usado para demostrar la existencia de soluciones de problemas de ecuaciones diferenciales en derivadas parciales, así como para la aproximación de la solución mediante métodos numéricos tales como el método de los elementos finitos.

La que formulamos aquí es una versión simple en espacios de Hilbert separables. Sin embargo, esto puede ser generalizado a espacios de Banach reflexivos y separables.

TEOREMA 3.1. *Sea H un espacio de Hilbert separable y $J : H \to \mathbb{R}$ un funcional continuo, débilmente sci y coercivo. Puede entonces construirse una sucesión minimizante u_u de J*

resolviendo los problemas

$$J(u_n) = \min_{u \in H_n} J(u),$$

donde H_n *es una sucesión de subespacios de* H *tales que*

1. $\dim H_n = n$,
2. $H_n \subset H_{n+1}$,
3. $H = \overline{\cup_{n=1}^{\infty} H_n}$.

DEMOSTRACIÓN. Puesto que H es un espacio de Hilbert separable, la sucesión de espacios H_n existe. En efecto, si $\{e_1, e_2, \ldots\}$ es una base de H, basta con considerar los subespacios H_n engendrados por $\{e_1, \ldots, e_n\}$.

El Teorema 2.2 implica la existencia de puntos de mínimo tanto de los problemas finito dimensionales

$$J(u_n) = \min_{u \in H_n} J(u),$$

como del problema infinito-dimensional

$$J(u_0) = \min_{u \in H} J(u).$$

Consideremos la proyección ortogonal $P_n : H \to H_n$. La continuidad de f implica que

$$\lim_n J(P_n u_0) = J(\lim_n P_n u_0),$$

y las hipótesis (ii) y (iii) implican que

$$\lim_n P_n u = u \quad \text{para todo } u \in H.$$

Así deducimos que $\lim_n J(P_n u_0) = J(u_0)$. Por otra parte, como $J(P_n u_0) \geq J(u_n) \geq J(u_0)$, entonces se tiene

$$lim_n J(u_n) = J(u_0),$$

es decir, u_n es una sucesión minimizante. □

El Teorema 3.1 no asegura que la sucesión minimizante u_n converja. Sólo asegura que la sucesión minimizante construida proporciona una buena aproximación del valor mínimo del funcional J si la dimensión n del espacio H_n en el cual se aproxima el mínimo de J es suficientemente grande. Tampoco proporciona ningún criterio sobre la precisión de la aproximación. El tamaño de n dependerá tanto de la forma concreta del funcional J como de la adecuada elección de la sucesión de espacios H_n.

EJEMPLO 3.2. ***Funcionales cuadráticos.*** *Sea* $J : H \to \mathbb{R}$ *dado por*

$$J(u) = \frac{1}{2} Q(u,u) - < w, u\,, \,>$$

para cierta forma bilineal, Q, que es además continua, coerciva y simétrica, y donde $w \in H$ *es un punto fijado. Sea* $\{e_1, e_2, \ldots\}$ *una base ortonormal de* H *y definamos* H_n *como el subespacio* *generado por* $\{e_1, \ldots, e_n\}$*. Entonces, como veremos en el siguiente capítulo, si* u_n *es un punto de mínimo para* J *en* H_n *entonces*

$$J'(u_n)(v) = 0 \quad \textit{para todo } v \in H_n,$$

que es cierto si y sólo si

$$J'(u_n)(e_k) = 0 \quad \textit{para } k = 1, \ldots, n.$$

Como $J'(u_n)(v) = Q(u_n, v) - < w \, , \, v >$*, para determinar*

$$u_n = \sum_{j=1}^{n} \alpha_j^n v_j$$

basta con resolver el sistema de ecuaciones lineales

$$\sum_{j=1}^{n} \alpha_j^n Q(v_j, v_k) = < w \, , \, v_k > \qquad \textit{para } k = 1, \ldots, n.$$

Capítulo 4

Cálculo diferencial en espacios de Banach

1. La derivada Fréchet

En esta sección veremos que los resultados del cálculo diferencial en espacios de dimensión finita pueden ser generalizados de un modo natural al caso de funcionales defindos sobre un espacio de Banach.

En el caso de una función diferenciable $f : U \subset \mathbb{R} \to \mathbb{R}$, la derivada $f'(x_0)$ de f en un punto x_0 se interpreta como la pendiente de la tangente al grafo de f (es decir, el conjunto $\{(x, f(x)) : x \in U\}$) en el punto $(x_0, f(x_0))$. La tangente viene dada por

$$T_{f,x_0}(x) = f(x_0) + f'(x_0)(x - x_0) \quad \text{para } x \in \mathbb{R}.$$

Por definición, la tangente es la *mejor* aproximación lineal al grafo de f en un entorno de x_0. La idea de esta aproximación está expresada por la siguiente ecuación:

$$f(x_0 + h) - T_{f,x_0}(x_0 + h) = o(x_0, h),$$

donde h denota la distancia entre x y x_0. La función o, que determina la aproximación, satisface $o(x_0, 0) = 0$ y

$$\lim_{|h|\to 0} \frac{|o(x_0, h)|}{|h|} = 0.$$

Esta idea, que puede ser fácilmente generalizada al caso de espacios de Banach, consiste en interpretar el producto $f'(x_0)(x - x_0)$ como una aplicación lineal del espacio de Banach $\mathbb{R}$ sobre el espacio de Banach $\mathbb{R}$. Por ello introducimos la siguiente definición.

DEFINICIÓN 1.1. *Sean $(V_1, \|\cdot\|_1)$ y $(V_2, \|\cdot\|_2)$ dos espacios de Banach y $U \subset V_1$ un conjunto abierto. Entonces se dice que una función*

$$J : U \to V_2$$

*es **diferenciable Fréchet en un punto** $u_0 \in U$ si existe una aplicación lineal y continua $D_{u_0}(f) : V_1 \to V_2$ tal que*

$$J(u_0 + h) - (J(u_0) + D_{u_0}J(h)) = o(u_0, h) \quad \text{para todos } h \in V_1, \quad u_0 + h \in U, \tag{33}$$

donde la función $o(u_0, \cdot) : V_1 \to V_2$ satisface $o(u_0, 0) = 0$ y

$$\lim_{h\to 0} \frac{\|o(u_0, h)\|_2}{\|h\|_1} = 0. \tag{34}$$

A veces usaremos las notaciones $J'(u_0)$ o $DJ(u_0)$ en vez de $D_{u_0}J$. Si J es diferenciable en todos los puntos de U entonces diremos que J es ***diferenciable en*** *U y, entonces, la aplicación*

$$DJ : U \to \mathcal{L}(V_1, V_2),$$

que asocia todo punto $u \in U$ con la derivada de J en U, es llamada la ***derivada Fréchet de*** *J. Si DJ es una aplicación continua, diremos que J* ***es diferenciable con continuidad en*** *U o que es de clase C^1 (es decir, $J \in C^1(U, V_2)$).*

En el caso en que $V_2 = \mathbb{R}$, la derivada Fréchet DJ es un elemento de $\mathcal{L}(V_1, \mathbb{R}) = V_1'$, es decir, del espacio dual de V_1.

Observemos que la derivada Fréchet, si existe, es única. En efecto, si hubiera dos aplicaciones lineales y continuas, digamos T_1 y T_2 satisfaciendo (33) y llamamos $T = T_1 - T_2$, entonces restando las identidades correspondientes obtenemos $T(h) = o_1(u_0, h) - o_2(u_0, h)$. Dividiendo por $\|h\|_1$ y tomando el límite $\|h\|_1 \to 0$ concluimos que $T = 0$.

La diferenciabilidad Fréchet de J en U_0 implica la continuidad. De hecho, tenemos un resultado más fuerte.

PROPOSICIÓN 1.2. *Sea $J : U \subset V_1 \to V_2$ diferenciable Fréchet en u_0. Entonces, para todo $\varepsilon > 0$ existe un entorno $B \subset U$ de u_0 tal que*

$$\|J(u) - J(u_0)\|_2 \leq (\|DJ(u_0)\| + \varepsilon)\|u - u_0\|_1 \quad \textit{para todo } u \in B.$$

DEMOSTRACIÓN. Debido a (33) tenemos que, para $\varepsilon > 0$ existe un entorno B tal que

$$\|J(u) - J(u_0) - DJ(u_0)(u - u_0)\|_2 \leq \varepsilon\|u - u_0\|_1 \quad \text{para todo } u \in B,$$

y por tanto

$$\|J(u) - J(u_0)\|_2 \leq \|DJ(u_0)(u - u_0)\|_2 + \varepsilon\|u - u_0\|_1 \leq (\|DJ(u_0)\| + \varepsilon)\|u - u_0\|_1.$$

□

Otra propiedad interesante de la derivada Fréchet que se hereda de la derivación en espacios de dimensión finita es la **regla de la cadena**:

PROPOSICIÓN 1.3. *Sean V_i, i=1,2,3 espacios de Banach y $U_i \subset V_i$ abiertos. Sean $J_1 : U_1 \to U_2$ y $J_2 : U_2 \to U_3$ aplicaciones diferenciables. Entonces $J_2 \circ J_1 : U_1 \to U_3$ es una aplicación diferenciable y, para todo $u \in U_1$ se tiene*

$$D(J_2 \circ J_1)(u) = DJ_2(J_1(u)) \circ DJ_1(u).$$

La demostración se basa en el hecho de que expresiones del tipo $DJ_2(J_1(u)) \circ o(u, h)$ tienen la propiedad (34).

El **Teorema del valor medio Lipschitzianidad derivadas de orden superior**

EJEMPLO 1.4. ***El caso de dimensión finita***

EJEMPLO 1.5. ***Aplicaciones lineales entre espacios de Banach***

EJEMPLO 1.6. ***Funcionales lineales sobre espacios de Hilbert***

EJEMPLO 1.7. ***Funcionales definidos por un núcleo integral***

EJEMPLO 1.8. ***Funcionales del tipo variacional***

2. La derivada Gateaux

Así como la generalización de diferencial de una función definida en un espacio de dimensión infinita dió lugar a la derivada Fréchet, la generalización de la derivada parcial conduce a la derivada Gateaux.

DEFINICIÓN 2.1. *Sean $(V_1, \|\cdot\|_1)$ y $(V_2, \|\cdot\|_2)$ dos espacios normados y $J : V_1 \to V_2$ una aplicación. Si, para algún punto $u_0 \in V_1$ existe una aplicación $\delta J(u_0, \cdot) : V_1 \to V_2$ tal que*

$$\lim_{t\to 0} \|\frac{J(u_0 + tv) - J(u_0)}{t} - \delta J(u_0, v)\| = 0 \quad \textit{para todo } v \in V_1,$$

entonces llamamos a $\delta J(u_0, v)$ la ***diferencial o variación Gateaux de** J **en el punto** u_0 **en la dirección** v.*

Observemos que $\delta J(u_0, \cdot)$ no es lineal ni continua en general, aunque siempre es homogénea en el sentido de que

$$\delta J(u_0, \alpha v) = \alpha \delta J(u_0, v) \quad \text{para todo } \alpha \in \mathbb{R}.$$

En efecto, si $\alpha \neq 0$,

$$\delta J(u_0, \alpha v) = \lim_{t\to 0} \frac{J(u_0 + \alpha tv) - J(u_0)}{t} = \alpha \lim_{t\to 0} \frac{J(u_0 + \alpha tv) - J(u_0)}{\alpha t} = \alpha \delta J(u_0, v).$$

El siguiente ejemplo muestra un caso en el que la variación Gateaux no es ni lineal ni continua.

EJEMPLO 2.2. *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x_1, x_2) = \begin{cases} x_1 & \textit{si } |x_2| \geq x_1^2, \\ |x_2|/x_1 & \textit{en otro caso.} \end{cases}$$

Tenemos que f vale 0 en los ejes, con lo cual $\delta f(0)h$ existe y es igual a cero cuando h está en alguno de los ejes. Sin embargo, si $h = (h_1, h_2)$ no está en ninguno de los ejes, entonces

$$\frac{f(th_1, th_2) - f(0, 0)}{t} = h_1 \quad \textit{si } 0 < t < |h_2|/h_1^2,$$

con lo cual, para tal h, tenemos $\delta f(0)h = h_1$. Por tanto, $\delta f(0)$ es no lineal y discontinua en todos los puntos del eje OX_1, excepto el origen.

Si $\delta J(u_0, \cdot)$ es lineal y continuo entonces escribiremos

$$\delta J(u_0, v) \equiv \delta_{u_0} J(v),$$

y nos referiremos a $\delta_{u_0} J$ como la **derivada Gateaux de** J **en** u_0 (la unicidad de la deivada Gateaux se deduce de un modo similar a la de la derivada Fréchet). Observemos finalmente que si J es un funcional sobre V_1 entonces $\delta_{u_0} J$ es un elemento del dual, V_1'.

Los siguientes resultados establecen la relación existente entre la derivada Fréchet y la derivada Gateaux. En primer lugar, observemos que una función puede tener derivada Gateaux en un punto y no tener derivada Fréchet.

EJEMPLO 2.3. *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x_1, x_2) = \begin{cases} 1 & \textit{si } x_2 = x_1^2 \neq 0, \\ 0 & \textit{en otro caso.} \end{cases}$$

Entonces $\delta f(0)h$ existe y es igual a cero para todo $h \in \mathbb{R}^2$, luego es una aplicación linela y continua. Sin embargo, f no es continua en el origen, luego no tiene derivada Fréchet en tal punto.

Sin embargo, el anterior ejemplo no contradice el siguiente resultado, el cual es inmediato.

LEMA 2.4. *Si J es diferenciable Gateaux en u_0 entonces J es continua en u_0 en cualquier dirección v, es decir*

$$\lim_{t \to 0} \|J(u_0 + tv) - J(u_0)\| = 0.$$

LEMA 2.5. *Si J es diferenciable Fréchet en el punto u_0 entonces J es diferenciable Gateaux en tal punto y ambas diferenciales coinciden.*

DEMOSTRACIÓN. Basta observar que, para $t \neq 0$, la definición de derivada de Fréchet implica que

$$\frac{J(u_0 + tv) - J(u_0)}{t} - \frac{D_{u_0} J(tv)}{t} = \frac{o(u_0, tv)}{t}.$$

Como $D_{u_0} J$ es un operador lineal, se sigue que

$$\lim_{t \to 0} \frac{J(u_0 + tv) - J(u_0)}{t} = D_{u_0} J(v).$$

De modo que J es diferenciable Gateaux y ambas derivadas coinciden. □

Para la demostración de un resultado recíproco necesitamos establecer el Teorema del Valor Medio.

LEMA 2.6. *Sean $u_0, u_0 + v \in V_1$ y $[u_0, u_0 + v] \subset V_1$ el segmento cuyos extremos son u_0 y $u_0 + v$. Supongamos que J es Gateaux diferenciable en dicho segmento. Entonces existe al menos un $\hat{u} \in [u_0, u_0 + v]$ tal que*

$$J(u_0 + v) - J(u_0) = \delta_{\hat{u}v}.$$

La demostración se basa en la aplicación del Teorema de Rolle. Primero consideramos la función derivable $\psi : [0,1] \to V_2$ dada por

$$\psi(t) = J(t(u_0 + v) + (1-t)u_0) - (tJ(u_0 + v) + (1-t)J(u_0)).$$

Observemos que, por la regla de la cadena,

$$\psi'(t) = \delta_{t(u_0+v)+(1-t)u_0} J(v) - (J(u_0 + v) - J(u_0),$$

luego nos basta con comprobar que ψ tiene un punto crítico $\hat{t} \in [0,1]$. Observemos que en el caso en que J es un funcional, es decir $V_2 = \mathbb{R}$, el resultado es inmediato, pues ψ es derivable y $\psi(0) = \psi(1) = 0$.

LEMA 2.7. *Si la derivada Gateaux $\delta_{u_0} J$ existe y es continua en un entorno U de u_0 entonces*

$$\delta_{u_0} J = D_{u_0} J.$$

DEMOSTRACIÓN. Para todo $\varepsilon > 0$ podemos encontrar una bola abierta $u_0 + B_r$ contenida en U tal que $\delta_u J$ está definida y satisface

$$\|\delta_u J - \delta_{u_0} J\| \leq \varepsilon \quad \text{para todo } u \in u_0 + B_r. \tag{35}$$

Consideremos la aplicación, definida en B_r,

$$\omega(v) = J(u_0 + v) - J(u_0) - \delta_{u_0} J(v).$$

Por una parte, el Teorema del Valor Medio implica

$$\omega(v) = \omega(v) - \omega(u_0) = \delta_{\hat{u}} \omega(v - u_0) \quad \text{para cierto } \hat{u} \quad \text{del segmento } [u_0, u_0 + v].$$

Por otra, un sencillo cálculo nos permite obtener

$$\delta_{\hat{u}} \omega(v) = \delta_{u_0 + \hat{u}} J(v) - \delta_{u_0} J(v).$$

Por tanto,

$$J(u_0 + v) - J(u_0) - \delta_{u_0} J(v) = \delta_{u_0 + \hat{u}} J(v) - \delta_{u_0} J(v) = o(u_0, v),$$

cuando $\varepsilon \to 0$, debido a (35). □

EJEMPLO 2.8. ***Diferenciabilidad de la norma L^p, con $1 < p < 2$.*** *Demostraremos la diferenciabilidad del funcional*

$$J(u) = \|u\|_p^p = \int_{\mathbb{R}^n} |u(x)|^p dx,$$

en $M = L^p(\mathbb{R}^n)$, con $1 < p < 2$. Para ello veremos que la derivada Gateaux existe y es continua en todo punto de $L^p(\mathbb{R}^n)$. La diferenciabilidad de J se sigue entonces del Lema ???

Sean $u, v \in L^p(\mathbb{R}^n)$, $t \neq 0$ y consideremos la expresión

$$\frac{J(u + tv) - J(u)}{t}.$$

Consideremos también la función real

$$\varphi(\xi) = \frac{|1+\xi|^p - 1 - p\xi}{|\xi|^p}.$$

Se tiene que φ está acotada en $\mathbb{R}$, de donde se deduce la existencia de constantes c_1 y c_2 tales que

$$c_1|\xi|^p \le |1+\xi|^p - 1 - p\xi \le c_2|\xi|^p.$$

Sustituyendo $\xi = tv(x)/u(x)$ (en los $x \in \mathbb{R}^n$ tales que $u(x) \neq 0$) y multiplicando por $|u|^p$ obtenemos

$$c_1|tv(x)|^p \le |u(x)+tv(x)|^p - |u(x)|^p - ptv(x)|u(x)|^{p-1}\operatorname{sign} u(x) \le c_2|tv(x)|^p.$$

Dividiendo por t e integrando en $\mathbb{R}^n$ se deduce que

$$\delta_u J(v) = p\int_{\mathbb{R}^n} v(x)|u(x)|^{p-1}\operatorname{sign} u(x)dx,$$

Finalmente, como

$$|u|^{p-1}\operatorname{sign} u \in L^{p'}(\mathbb{R}^n),$$

se sigue de la desigualdad de Hölder que la aplicación $v \to \delta_u J(v)$ es un funcional lineal y continuo sobre $L^p(\mathbb{R}^n)$, y por tanto la diferencial Gateaux coincide con la Fréchet.

3. Variación $n-$ésima

Frecuentemente resulta útil para el estudio de un funcional $J : V_1 \to \mathbb{R}$ el considerar la siguiente función real

$$F_v(t) = J(u_0 + tv),$$

con $u_0, v \in V_1$ fijados. Si F_h es N veces diferenciable, podemos desarrollarla como una serie de Taylor

$$F_v(t) = \sum_{n=1}^{N} F_v^{(n)}(0)\frac{t^n}{n!} + R_N \quad \text{para todo } t \in (-t_0, t_0)$$

para cierto $t_0 > 0$. El resto R_N viene dado por

$$R_N = \frac{t^N}{N!}(F_v^{(N)}(\sigma(t,v)t) - F_v^{(N)}(0)) = o(t^N) \quad 0 < \sigma < 1,$$

es decir $\frac{R_N}{t^N} \to 0 \quad$ cuando $t \to 0$.

DEFINICIÓN 3.1. *La **variación $n-$ésima de J en un punto u_0 en la dirección** $v \in V_1$ viene dada por*

$$\Delta^n J(u_0; v) \equiv F_v^{(n)}(0) \equiv \left.\frac{d^n J(u_0 + tv)}{dt^n}\right|_{t=0},$$

si esta derivada existe.

El siguiente lema describe la relación entre la primera variación y la derivada de Gateaux.

LEMA 3.2. *La derivada de Gateaux $\delta_{u_0}J$ existe si y sólo si la primera variación $\Delta J(u_0; v)$ existe para todo $v \in V_1$ y si la aplicación $v \to \Delta J(u_0; v)$ es un funcional lineal y continuo sobre V_1. En tal caso,*

$$\Delta J(u_0; v) = \delta_{u_0} J(v).$$

OBSERVACIÓN 3.3. *Si J tiene un mínimo local en u_0, es decir,*

$$J(u) \geq J(u_0) \quad \text{para todo } u \in U(u_0),$$

entonces F_v tiene también un mínimo local en $t = 0$, esto es

$$F_v'(0) = 0 \quad \text{y } F_v''(0) \geq 0,$$

siempre que estas derivadas existan. Esto muestra la importancia que puede tener el estudio de las variaciones $n-$ésimas para la obtención de resultados en problemas de valores extremos.

OBSERVACIÓN 3.4. *Hemos visto la siguiente cadena de implicaciones: la existencia de la derivada Fréchet implica la existencia de la derivada Gateaux, que a su vez implica la existencia de la primera variación.*

En muchas aplicaciones la existencia de la $n-$ésima variación puede ser verificada más fácilmente que, por ejemplo, la $n-$ésima derivada Fréchet. Por ello es una ventaja el poder obtener condiciones sobre la $n-$ésima variación que implique la existencia de un extremo. Estas condiciones las estudiaremos en el siguiente capítulo.

En las aplicaciones, muchas veces sucede que el dominio M del funcional J no es un subconjunto o un entorno abierto. En esta situación, la clase de curvas $u(t)$ que hay que considerar para investigar el funcional $J(u(t))$ debe ser restringido apropiadamente. Por ejemplo, si M es un convexo, entonces normalmente eligiremos $u(t) = tu_0 + (1-t)v$, con $u_0, v \in M$ y $t \in [0, 1]$.

4. El Teorema de Weierstrass generalizado revisitado

Con la ayuda del cálculo diferencial, vamos a deducir nuevos criterios que nos ayuden a verificar el cumplimiento de las hipótesis del Teorema de Weierstrass generalizado.

Veremos una condición suficiente para que el funcional $J : U \to \mathbb{R}$ sea débilmente sci. Supongamos que se satisface

$$J(u) - J(u_0) - D_{u_0}J(u - u_0) \geq 0 \tag{36}$$

para $u \in u_0 + \bar{B}_r$. Entonces, si $u_n \rightharpoonup u_0$, tendremos

$$J(u_n) - J(u_0) \geq J'(u_0)(u_n - u_0),$$

y en el límite obtenemos

$$\liminf_n J(u_n) \geq J(u_0),$$

con lo que J resulta ser débilmente sci en u_0. Por otra parte, veamos que (36) implica que

$$D^2J(u)(v, v) \geq 0.$$

En efecto, para $\tau, \tau' \in [0,1]$ tenemos

$$\begin{aligned} J(u) - J(u_0) &= D_{u_0+\tau(u-u_0)} J(u-u_0) \\ &= D_{u_0}(u-u_0) + (D_{u_0+\tau(u-u_0)} J(u-u_0) - D_{u_0} J(u-u_0)) \\ &= D_{u_0} J(u-u_0) + \tau D^2_{u_0+\tau'(u-u_0)} J(u-u_0, u-u_0). \end{aligned}$$

Tenemos, pues, el siguiente resultado.

LEMA 4.1. *Sea V un spacio de Banach y $J : V \to \mathbb{R}$ un funcional C^2 sobre $\bar{B}_r$ tal que*

$$D^2 J(u)(v,v) \geq 0 \quad \textit{para todo } u \in \bar{B}_r, \quad v \in V.$$

Entonces J es débilmente sci en $\bar{B}_r$.

Con la ayuda del cálculo diferencial también podemos encontrar criterios de coercividad.

LEMA 4.2. *Sea J un funcional C^1 sobre un espacio de Banach V. Supongamos que existe una función continua $g : (0,\infty) \to \mathbb{R}$ con*

1. $$\int_{r_0}^{\infty} \frac{g(t)}{t} dt = +\infty \quad \textit{para algún } r_0 > 0,$$
2. $= D_u J(u) \geq g(\|u\|)$.

Además, supongamos que J está acotado inferiormente en $S_{r_0}(V) = \{u \in V : \|u\| = r_0\}$. Entonces J es coercivo en V.

DEMOSTRACIÓN. Tenemos que

$$J(sv) - J(r_0 v) = \int_{r_0}^{s} D_{tv} J(v) dt \geq \int_{r_0}^{s} \frac{g(t)}{t} dt,$$

uniformemente en $v \in V$, con $\|v\| = 1$, para $s > r_0$. Se sigue que, para $\|u\| \to \infty$, teniendo (i) en cuenta,

$$J(u) \geq J(r_0 \frac{u}{\|u\|}) + \int_{r_0}^{\|u\|} \frac{g(t)}{t} dt \to \infty \quad \text{cuando } \|u\| \to \infty,$$

lo que concluye la demostración. □

Finalmente, tenemos el siguiente resultado para funcionales débilmente sci, el cual es el correspondiente al Teorema de Rolle en análisis en espacios de dimensión finita.

LEMA 4.3. *Sea V un espacio de Banach reflexivo y $U \subset V$ un subconjunto abierto y acotado. Sea $\bar{U}^w$ el cierre de U en la topología débil y ∂U su frontera. Supongamos que $J : V \to \mathbb{R}$ es un funcional débilmente sci en $\bar{U}^w$ tal que*

$$J(u) \geq J(u_0) \quad \textit{para todo } u \in \partial U$$

y para algún $u_0 \in U$. Entonces J tiene un punto crítico en U.

DEMOSTRACIÓN. Se sigue del Teorema de Weierstrass generalizado que J tiene un mínimo en $\bar{U}^w$. Se sigue que u_0 es un punto interior de mínimo para J. Por tanto $J'(u_0) = 0$. □

5. Convexidad de J y monotonía de J'

Sea V un espacio de Banach y $T : V \to V'$ tal que

$$<Tx - Ty\,,\, x - y> \geq 0 \quad \text{para todos } x, y \in V.$$

Entonces decimos que T **es un operador monótono**. (Si la desigualdad es estricta, diremos que es un **es un operador estrictamente monótono**.

TEOREMA 5.1. *Sea V un espacio de Banach y J un funcional C^1 sobre V. Entonces J es (estrictamente) convexo en V si y solo si J' es (estrictamente) monótono en V.*

DEMOSTRACIÓN. Supongamos que J es convexo. Entonces

$$J(v + \lambda(u - v)) - J(v) \leq \lambda(J(u) - J(v)) \quad \text{para todos } u, v \in V, \quad \lambda \in [0, 1].$$

Se sigue que

$$<J'_v\,,\, u - v> \leq J(u) - J(v),$$

y análogamente,

$$<J'_u\,,\, v - u> \leq J(v) - J(u).$$

Sumando estas desigualdades obtenemos

$$<J'_v - J'_u\,,\, u - v> \leq 0,$$

Luego J' es monótona.

Supongamos ahora que J' es monótona. Para $u, v \in V$, consideremos la función

$$p(\lambda) = J(\lambda u + (1 - \lambda)v) - \lambda J(u) - (1 - \lambda)J(v),$$

con $\lambda \in [0, 1]$. Para demostrar que J es convexa basta con verificar que $p(\lambda) \leq 0$ en $[0, 1]$. Claramente, p es una función derivable con $p(0) = p(1) = 0$. Si existe un λ^* tal que $p(\lambda^* > 0$, entonces existe un punto de máximo $\lambda_0 \in (0, 1)$ tal que $p'(\lambda_0) = 0$. Sea $\lambda \in [0, \lambda_0)$. Entonces

$$\begin{aligned} p'(\lambda) - p'(\lambda_0) &= \; <J'(\lambda u + (1 - \lambda)v) - J'(\lambda_0 u + (1 - \lambda_0))\,,\, u - v> \\ &= \; (\lambda - \lambda_0)^{-1} <J'(\hat{u}) - J'(\hat{u}_0)\,,\, \hat{u} - \hat{u}_0> \geq 0, \end{aligned}$$

donde $\hat{u} = \lambda u + (1 - \lambda)v$ y $\hat{u}_0 = \lambda_0 u + (1 - \lambda_0)v$. De modo que p no puede decrecer para $\lambda > \lambda_0$. Se sigue la contradicción $p(\lambda_0) \leq p(1) = 0$. □

Capítulo 5

Extremos de funcionales diferenciables

En este capítulo estudiaremos cómo calcular los puntos de mínimo de un funcional cuya existencia y, a veces, unicidad, asumiremos que ha sido previamente demostrada. Observemos que con los métodos introducidos en el Capítulo ?? no nos esposible calcular las soluciones de los problemas variacionales que nos planteamos en la Introducción, y que es precisamente en este punto en el que entra en juego el cálculo diferencial de un modo análogo al caso bien conocido de funciones reales de variable real.

1. Extremos y valores críticos

Comenzamos con la siguiente definición.

DEFINICIÓN 1.1. *Sea $U \subset V$ un subconjunto abierto de un espacio de Banach, V, y sea $\varphi : U \to \mathbb{R}$ una aplicación diferenciable. Entonces $u_0 \in U$ es un* ***punto crítico*** *de φ si la derivada Fréchet $D\varphi(u_0) = 0$. En caso contrario, diremos que u_0 es un* ***punto regular*** *de φ.*

EJEMPLO 1.2. *Sea $V = H$ un espacio de Hilbert y supongamos que B es un operador acotado y simétrico en H con inversa acotada, $B^{-1} \in \mathcal{L}(H, H)$. Sea $W : H \to \mathbb{R}$ una aplicación diferenciable. La derivada Fréchet del funcional φ definido por*

$$\varphi(u) = \frac{1}{2} < u\,,\, Bu > - W(u),$$

está dada por

$$D\varphi(u) = Bu - DW(u).$$

Esto significa que $u_0 \in H$ es un punto crítico de φ si es solución de la ecuación

$$u_0 = B^{-1} DW(u_0).$$

En el caso particular en el que $W(u) = \frac{1}{2}\lambda < u\,,\, u >= \frac{1}{2}\lambda \|u\|^2$, entonces $DW(u) = \lambda u$, con lo cual u_0 sería un punto crítico de φ si y solo si es un autovector de B asociado al autovalor λ.

2. Condiciones necesarias para un extremo

El siguiente teorema establece que los puntos de extremos de un funcional son necesariamente puntos críticos del mismo.

TEOREMA 2.1. ***Condición necesaria de Euler-Lagrange.*** *Sea U un subconjunto abierto de un espacio de Banach, V, y sea $J : U \to \mathbb{R}$ un funcional diferenciable Fréchet en U. Entonces todo punto de extremo de J es un punto crítico de J.*

DEMOSTRACIÓN. Sea $u_0 \in U$ un punto de extremo para J. Existe entonces un $r > 0$ tal que $B_r(u_0) \subset U$ y u_0 es un extremo de J restringido a $B_r(u_0)$. Sea $h \in V$, $h \neq 0$ y $\delta > 0$ tal que $\|h\|\delta < r$. Consideremos la función real dada por

$$F_h(t) = J(u_0 + th), \quad t \in (-\delta, \delta).$$

Por la regla de la cadena, F_h es derivable, y su derivada viene dada por

$$\frac{d}{dt} F_h(t) = DJ(u_0 + th)(h).$$

Como u_0 es un extremo de J en $B_r(u_0)$ tenemos que $t = 0$ es un punto de extremo de F_h, y por tanto, es un punto crítico de esta función real. Por tanto

$$0 = \frac{dF_h}{dt}(0) = DJ(u_0)(h).$$

Puesto que $h \in V$ es arbitrario, se sigue que $DJ(u_0) = 0$, es decir, que u_0 es un punto crítico de J. □

EJEMPLO 2.2. *Consideremos dos puntos P, Q de la esfera que no están diametralmente opuestos. Hay entonces dos arcos (de distinta longitud) del círculo máximo que contiene a P y Q que conecta a estos dos puntos. El arco más corto es la geodésica (es decir, el arco más corto entre dos puntos de la superficie), mientras que el arco más largo que los conecta, aunque es un* punto crítico*, no es el arco ni de mínima ni de máxima longitud que los conecta. Este ejemplo muestra que las condiciones de Euler-Lagrange no son suficientes.*

Antes de pasar a discutir las condiciones suficientes que debe cumplir un funcional para tener un extremo, veamos las condiciones que han de cumplir los funcionales que son dos veces diferenciables en el sentido de Fréchet. Estos funcionales tienen la siguiente propiedad de desarrollo en serie de Taylor.

LEMA 2.3. *Sea V un espacio de Banach y $U \subset V$ un subconjunto abierto. Supongamos que $J \in C^2(U, \mathbb{R})$. Entonces, para todo $u_0 \in U$ existe un $r \equiv r(u_0) > 0$ tal que $B_r(u_0) \subset U$ y, para todo $h \in B_r(0)$ se tiene*

$$J(u_0 + h) = J(u_0) + DJ(u_0)(h) + \frac{1}{2} D^2 J(u_0)(h, h) + R_2(u_0, h)(h, h).$$

Aquizá, el funcional $R_2(u_0, \cdot)$ satisface la condición

$$\lim_{\|h\| \to 0} R_2(u_0, h) = 0,$$

en $\mathcal{L}(V \times V, \mathbb{R})$.

Para la demostración véase [?]. La primera consecuencia de este lema es el siguiente teorema.

TEOREMA 2.4. *Sea $U \subset V$ un subconjunto abierto de un espacio de Banach V y $J \in C^2(U, \mathbb{R}^2)$. Entonces, si J tiene un mínimo relativo en $u_0 \in U$, se tiene:*

1. $DJ(u_0) = 0$,
2. $D^2J(u_0) \in \mathcal{L}(U \times U, \mathbb{R})$ *es no negativa, es decir, para todo* $h \in U$, $D^2J(u_0)(h,h) \geq 0$.

DEMOSTRACIÓN. Sea $u_0 \in U$ un, por ejemplo, mínimo relativo. Entonces, existe un $r > 0$ tal que

1. $B_r(u_0) \subset U$,
2. $J(u_0) \leq J(u)$ para todo $u \in B_r(u_0)$.

Aplicando el Lema 2.3 tenemos que, para todo $h \in B_r$ se tiene

$$J(u_0) \leq J(u_0 + h) = J(u_0) + DJ(u_0)(h) + \frac{1}{2}D^2(u_0)(h,h) + R_2(u_0,h)(h,h).$$

Por el Teorema 2.1 se tiene $DJ(u_0) = 0$, luego,

$$0 \leq \frac{1}{2}D^2(u_0)(h,h) + R_2(u_0,h)(h,h) \quad \text{para todo } h \in B_r.$$

Esto es cierto, en particular, para εh, donde $\varepsilon \in (0,1]$ y $h \in B_r$, con lo cual se tiene

$$0 \leq \frac{1}{2}D^2(u_0)(\varepsilon h, \varepsilon h) + R_2(u_0, \varepsilon h)(\varepsilon h, \varepsilon h).$$

Entonces,

$$0 \leq \varepsilon^2 \big(\frac{1}{2}D^2(u_0)(h,h) + R_2(u_0, \varepsilon h)(h,h)\big).$$

De aquí obtenemos, para todos $\varepsilon \in (0,1]$ y $h \in B_r$,

$$0 \leq D^2(u_0)(h,h) + 2R_2(u_0, \varepsilon h)(h,h),$$

y teniendo en cuenta que $R_2(u_0, \varepsilon h) \to 0$ cuando $\varepsilon \to 0$ obtenemos

$$0 \leq D^2(u_0)(h,h) \quad \text{para todo } h \in B_r.$$

Finalmente, si $h \in U$, tenemos que para cierto $\lambda > 0$ se tiene $h \in B_{\lambda r}$ por lo que

$$0 \leq D^2(u_0)(\frac{h}{\lambda}, \frac{h}{\lambda}) = \lambda^{-2} D^2(u_0)(h,h),$$

con lo que se concluye la demostración. □

OBSERVACIÓN 2.5. *Si el funcional $J : V \to \mathbb{R}$ no es dos veces Fréchet diferenciable en V, todavía podemos deducir condiciones necesarias para la existencia de un extremo relativo en tírminos de la segunda variación. En efecto, si J tiene un mínimo relativo en u_0 entonces*

$$\Delta J(u_0, h) = 0 \quad \text{y } \Delta^2 J(u_0, h) \geq 0 \quad \text{para todo } h \in U,$$

si estas aplicaciones existen. Como ya mencionamos anteriormente, normalmente es más fácil comprobar la existencia y calcular la $n-$ésima variación que la n-ésima derivada de Fréchet.

3. Condiciones suficientes para un extremo

El Teorema 2.4 tiene casi un recï¿½proco.

TEOREMA 3.1. *Sea V un espacio de Banach y $U \subset V$ un subconjunto abierto. Supongamos que $J \in C^2(U, \mathbb{R})$ y que*

1. $DJ(u_0) = 0$,
2. $D^2J(u_0)$ *es estrictamente positivo, es decir,*

$$\inf \left\{ D^2J(u_0)(h,h) : h \in U, \|h\| = 1 \right\} > 0,$$

para cierto $u_0 \in U$. Entonces J tiene un mínimo relativo estricto en u_0.

DEMOSTRACIÓN. Tomemos $r_1 > 0$ como en el Lema 2.3. Entonces, para todo $h \in B_{r_1}$,

$$J(u_0 + h) = J(u_0) + \frac{1}{2} D^2J(u_0)(h,h) + R_2(u_0,h)(h,h).$$

Susutituyendo h por εh, con $\varepsilon \in (0,1]$ obtenemos

$$J(u_0 + \varepsilon h) - J(u_0) = \varepsilon^2 \big(\frac{1}{2} D^2J(u_0)(h,h) + R_2(u_0,\varepsilon h)(h,h) \big).$$

Debido a la hipótesis 2. del Teorema y al hecho de que $R_2(u_0,\varepsilon h) \to 0$ cuando $\varepsilon \to 0$, se tiene que para ε suficientemente pequeño el signo de $J(u_0 + \varepsilon h) - J(u_0)$ viene determinado por el de $D^2J(u_0)(h,h)$, de donde se sigue el resultado. □

Como ya mencionamos anteriormente, en el caso en que J es convexo se tienen resulados más fuertes.

COROLARIO 3.2. *Supongamos que $U \subset V$ es un conjunto convexo y que $J \in C^1(U, \mathbb{R})$ es un funcional convexo. Entonces todo punto crítico de J es un punto de mínimo.*

DEMOSTRACIÓN. Sea u_0 un punto crítico de J, de modo que $DJ(u_0) = 0$. Consideremos la función auxiliar definida sobre U

$$g(u) = J(u) - J(u_0).$$

Esta función también es convexa y satisface $g(u_0) = 0$ y $Dg(u_0) = 0$. Supongamos que J no tiene un mínimo relativo en u_0. Entonces existe un $u_1 \in B_{r_1}(u_0) \subset U$ tal que $g(u_1) < 0$. Consideremos otra función auxiliar $G : [0,1] \to \mathbb{R}$ dada por

$$G(t) = g(u_0 + t(u_1 - u_0)).$$

Por la regla de la cadena, G es diferenciable en $(0,1)$. Su derivada viene dada por

$$G'(t) = Dg(u_0 + t(u_1 - u_0))(u_1 - u_0),$$

con lo que

$$G'(0) = Dg(u_0)(u_1 - u_0) = 0.$$

Pero, por otra parte, la convexidad de g implica que

$$g(u_0 + t(u_1 - u_0)) \leq tg(u_1) + (1-t)g(u_0) = tg(u_1),$$

con lo que

$$G'(0) = \lim_{t\to 0} \frac{G(t) - G(0)}{t} = \lim_{t\to 0} \frac{g(u_0 + t(u_1 - u_0))}{t} = g(u_1) < 0.$$

□

EJEMPLO 3.3. *Sea H un espacio de Hilbert y $Q : H \times H \to \mathbb{R}$ una forma bilineal, continua y coerciva. Entonces el funcional*

$$J(u) = Q(u,u) - T(u)$$

tiene, para todo $T \in H'$, un único punto de mínimo u_0 (Corolario 3.2). Se sigue del Teorema 2.1 que

$$DJ(u_0)(h) = 2Q(u_0, h) - T(h) = 0 \quad \textit{para todo } h \in H.$$

En el caso $Q(u,u) =< Au\,, u >$ y $T(u) = 2 < u\,, v >$, la condición necesaria y suficiente para la existencia de un mínimo u_0 es, por tanto,

$$Au_0 = v.$$

EJEMPLO 3.4. ***El problema de la braquistocrona*** *Como mencionamos en la intoducción, el problema de la braquistocrona es uno de los problemas más antiguos del cálculo de variaciones. La primera solución fue dada por Johann Bernoulli en 1696, aunque también dieron soluciones algunos contemporáneos suyos como Jacob Bernoulli, Leibniz y Newton.*

El problema es determinar la curva C que une dos puntos A y B situados en un plano perpendicular a la superficie terrestre tal que un punto de masa $m = 1$ que se mueve a lo largo de la curva C bajo la ï¿½nica influencia de la fuerza de la gravedad, recorre el trayecto entre A y B en el mínimo tiempo posible.

Para nuestro tratamiento variacional del problema elijamos un sistema coordenado en este plano, con el semieje positivo de OX apuntando en la dirección de la fuerza gravitatoria. Tomemos $A = (0,0)$ y $B = (b, b')$, con $b > 0$ y $b' \geq 0$. El movimiento de esta masa puntual puede describirse por medio de las leyes de la mecánica clásica. Como ya vimos en la Introducción, el funcional a ser minimizado es

$$J(u) = \int_0^b \big(\frac{1 + u'(x)^2}{2gx}\big)^{1/2} dx,$$

donde $u \in C^1([0,b]; 0, b')$. La primera derivada Fréchet de J viene dada por

$$DJ(u)(h) = \int_0^b \frac{u'(x)}{\big(2gx(1 + u'(x)^2)\big)^{1/2}} h'(x)dx, \tag{37}$$

mientras que la segunda derivada Fréchet es

$$D^2J(u)(h,h) = \int_0^b \frac{1}{\big(2gx(1+u'(x)^2)\big)^{1/2}} \frac{h'(x)^2}{1+u'(x)^2} dx,$$

donde $h \in C_0^1([0,b];0,b')$. *La segunda derivada es claramente definida positiva por lo que cualquier punto crítico de* J *sería un mínimo. De (37) deducimos, tras una integración por partes, que la condición de Euler-Lagrange es*

$$0 = DJ(u)(h) = -\int_0^b \frac{d}{dx}\Big(\frac{u'(x)}{(2gx(1+u'(x)^2))^{1/2}}\Big)h(x)dx,$$

para todo $h \in C_0^1([0,b];0,b')$. *Por tanto, todo mínimo de* J *debe satisfacer*

$$\frac{d}{dx}\Big(\frac{u'(x)}{(2gx(1+u'(x)^2))^{1/2}}\Big) = 0 \quad \textit{para todo } x \in (0,b).$$

Tenemos, para cierta constante k

$$\frac{u'(x)}{(2gx(1+u'(x)^2))^{1/2}} = k \quad \textit{para todo } x \in (0,b),$$

y entonces, particularizando en $x = b$, *obtenemos*

$$2gk^2b = \frac{u'(b)^2}{1+u'(b)^2} \leq 1,$$

con lo que $k^2 \leq 1/2gb$. *Definiendo* $c = 1/4gk^2$ *obtenemos*

$$u'(x)^2(1-\frac{x}{2c}) = \frac{x}{2c},$$

donde $b/2c \leq 1$. *Como estamos buscando soluciones regulares, podemos eliminar el caso en que* $b/2c = 1$. *Cuando* $b/2c < 1$, *existe un único* $t_0 \in (0,\pi)$ *tal que* $b = c(1-\cos t_0)$. *Podemos entonces introducir la parametrización*

$$x \equiv x(t) = 2c(1-\cos t) \quad \textit{para todo } t \in [0,t_0],$$

con la cual $y(t) = u(x(t))$ *debe satisfacer*

$$y'(t) = u'(x(t))x'(t) = u'(x(t))c\sin t.$$

Se sigue que

$$y'(t)^2 = c^2\frac{1-\cos t}{1+\cos t}\sin^2 t = c^2(1-\cos t)^2.$$

Por tanto, $y'(t) = +-c(1-\cos t)$, *y usando la condición incial* $y(0) = 0$ *obtenemos la solución*

$$y(t) = +-c(t-\sin t).$$

Hemos obtenido así la solución de la ecuación de Euler para el problema de la braquistocrona en forma paramétrica:

$$\begin{aligned} x(t) &= 2c(1-\cos t) \\ y(t) &= +-c(t-\sin t) \quad \textit{para } t \in [0,t_0]. \end{aligned}$$

Estas son las ecuaciones de parte de una cicloide. Como debemos imponer que la curva pase por el punto $B = (b, b')$, las siguientes condiciones son necesarias

$$\begin{aligned} b &= x(t_0) = c(1 - \cos t_0) \\ b' &= y(t_0) = c(t_0 - \sin t_0). \end{aligned}$$

Podemos determinar t_0 del cociente entre b' y b:

$$\frac{b'}{b} = \frac{t_0 - \sin t_0}{1 - \cos t_0} \quad \text{para } t_0 \in (0, \pi). \tag{38}$$

Una vez resueltas estas ecuaciones la constante c se determina fácilmente. Para resolverlas, observemos que la función $f(t) = (t - \sin t)/(1 - \cos t)$ es creciente en $(0, \pi)$, y que su máximo valor es $f(\pi) = \pi/2$. Existe por tanto una ï¿½nica solución de (38) tal que $b'/b < \pi/2$ y por tanto una ï¿½nica solución del problema de la braquistocrona. Observemos finalmente que si $b'/b \geq \pi/2$ entonces el problema no tiene solución.

4. El Principio de Mínima Acción

Uno de los principios de la mecánica clásica asume que el comportamiento dinámico de un sistema físico Σ puede ser descrito en tírminos del *lagrangiano* asociado al sistema, el cual tiene la forma

$$L = L_\Sigma = T - V,$$

donde T es la energï¿½a cinï¿½tica y V la energï¿½a potencial. Para ello, se usa el Principio de Hamilton que establece que el movimiento del sistema Σ, que comienza en el instante $T = a$ en el punto x_i y que llega en el instante $t = b$ al punto x_f es tal que el *funcional de la acción* $W = W_\Sigma$, donde

$$W = \int_I L dt, \quad I = [a, b],$$

es estacionario. Por tanto, la determinación del comportamiento de un sistema clásico es un problema variacional. El Principio de Hamilton se conoce también como el *Principio de Mínima Acción*, puesto que hay casos en el que el funcional W no es, simplemente, estacionario sino que alcanza un mínimo.

En esta sección veremos quizá resultados podemos obtener acerca de este problema de la mecánica clásica a partir de los resultados generales que hemos deducido a lo largo del capítulo. Por claridad y simplicidad, nos restringiremos al caso unidimensional, aunque el caso $N-$dimensional puede tratarse de un modo análogo, véase [**?**].

Sea $L = L(t, x, \dot{x})$ el lagrangiano de un sistema físico con un grado de libertad. Para comenzar asumiremos que L es dos veces diferenciable con continuidad. Sea $I = [a, b]$ un intervalo compacto y $C^1(I)$ el espacio de Banach de funciones diferenciablesc on continuidad $x : I \to \mathbb{R}$ cuya norma es la del supremo

$$\|x\| = \|x\|_{I,1} = \sup_{t \in I} \{|x(t)|, |\dot{x}(t)|\},$$

donde $\dot{x}$ es la derivada de x con respecto a t. Denotemos por $E^1(I)$ al espacio de trayectorias diferenciables de I en $\mathbb{R}$ que comienzan en x_i y terminan en x_f, es decir, con $x(a) = x_i$ y $x(b) = x_f$. El funcional de acción W del sistema está dado por $W : E^1(I) \to \mathbb{R}$

$$W(x) = \int_I L(t, x(t), \dot{x}(t))dt.$$

Si definimos

$$E_0^1 = \left\{h \in C^1(I) : h(a) = h(b) = 0\right\},$$

entonces, para cualquier $x \in E^1(I)$ y para todo $h \in E_0^1(I)$ se tiene que $x + h \in E^1(I)$. Así podremos investigar la dependencia funcional de W.

4.1. Condiciones necesarias. A continuación discutiremos las condiciones necesarias sobre L para que una trayectoria $x \in E^1(I)$ sea un mínimo local del funcional de acción W. Si éste es el caso, entonces por el Teorema ??? se tiene que, si $D^2W(x) \neq 0$ entonces

$$DW(x)(h) = 0 \quad \text{para todo } h \in E_0^1, \tag{39}$$

$$D^2W(x)(h,h) \geq 0 \quad \text{para todo } h \in E_0^1. \tag{40}$$

La derivada Fréchet de W fue calculada en el Ejemplo ???. Usando la regla de la cadena podemos calcular también la segunda derivada Fréchet de W. Tenemos

$$DW(x)(h) = \int_I (L_x(t)h(t) + L_{\dot{x}}(t)\dot{h}(t))dt \tag{41}$$

y

$$D^2W(x)(h,h) = \int_I (L_{xx}(t)h(t)^2 + 2L_{x\dot{x}}(t)h(t)\dot{h}(t) + L_{\dot{x}\dot{x}}\dot{h}^2)dt, \tag{42}$$

donde hemos usado la notación

$$L_x(t) = \frac{\partial L}{\partial x}(t, x(t), \dot{x}(t)), \quad L_{xx}(t) = \frac{\partial^2 L}{\partial x^2}(t, x(t), \dot{x}(t)),$$

y expresiones similares para L_{xd}, $L_{x\dot{x}}$ y $L\dot{x}\dot{x}$.

Para deducir de (39)-(42) las condiciones necesarias sobre L y x para obtener un mínimo local de W necesitaremos unos resultados previos.

LEMA 4.1. ***Lema Fundamental del Cálculo de Variaciones (Lagrange)***. *Sea $g : I \to \mathbb{R}$ una función continua tal que*

$$\int_I g(t)h(t) = 0$$

para toda $h \in E_0^1$. Entonces $g(t) = 0$ para todo $t \in I$.

DEMOSTRACIÓN. □

LEMA 4.2. ***(Du Boys-Reymond)*** *Sea $g : I \to \mathbb{R}$ una función continua tal que*

$$\int_I g(t)\dot{h}(t) = 0$$

para toda $h \in E_0^1$. Entonces $g(t) = g(a)$ para todo $t \in I$.

DEMOSTRACIÓN. □

OBSERVACIÓN 4.3. *En las demostraciones anteriores hemos probado realmente resultados más fuertes que los de los enunciados de los respectivos lemas, puesto que solo hemos usado que las hipótesis son ciertas para toda función C^∞. Estos lemas se usan de un modo habitual en la teoría de distribuciones sobre $\mathbb{R}$, y en el lenguaje de esta teoría se enuncian como*

- *si una función continua se anula en el sentido de las distribuciones entonces se anula como función*
- *si la derivada, en el sentido de las distribuciones, de una función continua se anula, entonces la función es constante.*

Ya podemos comenzar a investigar cuál es la implicación de la condición (40) demostrando el siguiente lema sobre formas cuadráticas sobre E_0^1.

LEMA 4.4. *Sean $F_i : I \to \mathbb{R}$, $i = 1, 2, 3$ funciones continuas y supongamos que la forma cuadrática*

$$Q(h) = \int_I (F_1(t)h(t)^2 + F_2(t)h(t)\dot{h}(t) + F_3(t)\dot{h}(t)^2)dt,$$

definida en E_0^1, es no negativa. Entonces $F_3(t) \geq 0$ para todo $t \in I$.

DEMOSTRACIÓN. □

TEOREMA 4.5. ***Ecuación de Euler y condición de Legendre***. *Supongamos que el lagrangiano L satisface las condiciones usuales y supongamos que $D^2W(x) \neq 0$. Si una trayectoria $x \in E^1$ es un mínimo local de W entonces se tiene*

1. *La función $L_{\dot{x}}(t, x(t), \dot{x}(t))$ es diferenciable y se satisface la Ecuación de Euler*

$$\frac{d}{dt}L_{\dot{x}}(t, x(t), \dot{x}(t)) = L_x(t, x(t), \dot{x}(t)) \quad \text{para todo } t \in I. \tag{43}$$

2. *Se satisface la condición de Legendre*

$$L_{\dot{x}\dot{x}}(t, x(t), \dot{x}(t)) \geq 0 \quad \text{para todo } t \in I.$$

DEMOSTRACIÓN. □

LEMA 4.6. *Si L satisface las hipótesis usuales y $x \in E^1$ es una trayectoria para la cual $L_{\dot{x}}(t, x(t), \dot{x}(t))$ es diferenciable con continuidad con respecto a t entonces x es dos veces diferenciable con continuidad en todos aquellos puntos $t \in I$ tales que $L_{\dot{x}\dot{x}}(t) \neq 0$.*

DEMOSTRACIÓN. □

TEOREMA 4.7. *Si L satisface las hipótesis usuales y $x \in E^1$ es un mínimo local de W tal que $D^2W(x) \neq 0$ entonces x es dos veces diferenciable con continuidad en todos aquellos puntos $t \in I$ tales que $L_{\dot{x}\dot{x}}(t) > 0$.*

DEMOSTRACIÓN. □

El Teorema 4.7 muestra que todo mínimo local gana un orden de derivabilidad en el conjunto en que $L_{\dot{x}\dot{x}}(t) > 0$. Este es, pues, un resultado de regularidad adicional de la solución de un problema de minimización que, de hecho, es el primer resultado de esta clase que se obtuvo en el cálculo variacional.

En muchas aplicaciones el lagrangiano L tiene mayor regularidad que la que hemos asumido en esta sección (C^2). En el caso, por ejemplo, en el que $L \in C^3$ y $L_{\dot{x}\dot{x}} \geq \delta > 0$, entonces la trayectoria que minimiza la acción, x es dos veces diferenciable con continuidad, y por tanto, $L_{x\dot{x}} \in C^1$. Integrando por partes obtenemos

$$\int_I 2L_{x\dot{x}}(t)h(t)\dot{h}(t)dt = \int_I L_{x\dot{x}}(t)\frac{d}{dt}h(t)^2dt == \int_I h(t)^2\frac{d}{dt}L_{x\dot{x}}(t)dt.$$

Esto permite expresar la segunda derivada Fréchet de W como

$$D^2W(x)(h,h) = \int_I \big[\big(L_{xx} - \frac{d}{dt}L_{x\dot{x}}\big)h(t)^2 + L_{\dot{x}\dot{x}}\dot{h}(t)^2\big]dt, \tag{44}$$

que es en la forma en la que es normalmente estudiada.

4.2. Condiciones suficientes. A continuación pasamos al estudio de las condiciones suficientes sobre el lagrangiano L y la trayectoria x para que esta última minimize la acción. Como veremos, estas condiciones son considerablemente más difáciles de obtener.

El Teorema ??? establece que las condiciones suficientes para que $x \in E^1$ sea un mínimo local de W son:

$$DW(x) = 0, \tag{45}$$

$$D^2W(x) \quad \text{estrictamente positivo en } E_0^1 \times E_0^1. \tag{46}$$

Nuestro objetivo es traducir estas dos condiciones sobre W a condiciones, presumiblemente, similares a las que hemos obtenido como condiciones necesarias sobre L.

Legendre trató de demostrar que

$$L_{\dot{x}\dot{x}}(t) > 0 \quad \text{para todo } t \in I, \tag{47}$$

era la condición suficiente, correspondiente a la condición necesaria expresada en el Teorema ???. Si el lagrangiano es suficientemente regular, podemos escribir

$$J(h) = D^2W(x)(h,h) = \int_I (A(t)h^2(t) + B(t)\dot{h}(t)^2)dt, \tag{48}$$

donde

$$A(t) = L_{xx}(t) - \frac{d}{dt}L_{x\dot{x}}(t) \quad \text{y} \quad B(t) = L_{\dot{x}\dot{x}}(t). \tag{49}$$

La conjetura de Legendre fue entonces que, suponiendo que $B(t) > 0$ para todo $t \in I$ (es decir, que se satisface (47)), entonces se puede diagonalizar la forma cuadrática J en E_0^1 o, en otras palabras, que J puede ser escrito como una integral de suma de cuadrados de funciones de E_0^1.

Legendre usó una segunda expresión de J que se deduce del siguiente modo. Dado $\varphi \in C^1(I)$ y para toda $h \in E_0^1(I)$ se tiene

$$0 = \int_I \frac{d}{dt}(\varphi h^2)dt = \int_I (\dot{\varphi} h^2 + 2\varphi h \dot{h})dt,$$

y por tanto

$$J(h) = \int_I \big[B(t)\dot{h}(t)^2 + 2\varphi(t)h(t)\dot{h}(t) + (A(t) + \dot{\varphi}(t))h(t)^2\big]dt. \tag{50}$$

Se sigue que, si $B(t) > 0$ para todo $t \in I$ entonces

$$J(h) = \int_I B(t)\big[\big(\dot{h} + \frac{\varphi}{B}h\big)^2 + \big(\frac{h}{B}\big)^2(B(A + \dot{\varphi}) - \varphi^2)\big]dt. \tag{51}$$

Se sigue de esta expresión que J puede expresarse como una integral de suma de cuadrados de funciones si la *ecuación de Ricatti*

$$B(A + \dot{\varphi}) = \varphi^2 \tag{52}$$

tiene una solución $\varphi \in C^1(I)$. (Ejemplo de dificultad: $A = 0$, $B = 1$).

Para estudiar esta ecuación no lineal, que en principio no tiene por quizá tener solución en todo el intervalo I, observemos la siguiente relación entre las soluciones de (52) y las soluciones de la ecuación de Euler asociada al funcional J sobre E_0^1, es decir

$$-\frac{d}{dt}(B\dot{u}) + Au = 0. \tag{53}$$

Es fácil comprobar que si $\varphi \in C^1$ es una solución de (52) entonces

$$u(t) = u(a)\exp\big(-\int_a^t \frac{\varphi(\tau)}{B(\tau)}d\tau\big), \quad u(a) \neq 0, \tag{54}$$

es una solución de (53) que no se anula en I. Recíprocamente, si u es una solución de (53) que no se anula en I, entonces

$$\varphi = -\frac{\dot{u}}{u}B \tag{55}$$

es una solución de (52). Asumiendo que $B \in C^1$, la ecuación (53) es una ecuación diferencial ordinaria de segundo orden cuyos coeficientes son funciones continuas en I, cuya existencia de soluciones en todo I es un resultado bien conocido.

En otras palabras, el problema de determinar la existencia de soluciones en todo el intervalo I de la ecuación (no lineal) de Ricatti puede ser reducido al estudio de condiciones que hacen que la ecuación lineal (53) tenga una solución que no se anule en I. Obtenemos así una condición suficiente para la positividad estricta de J.

TEOREMA 4.8. *Supongamos que $A \in C^0(I)$ y $B \in C^1(I)$ y que*

1. $B(t) > 0$ *para todo* $t \in I$,
2. *la ecuación diferencial (53) tiene una solución* $u \in C^2(I)$ *que no se anula en ningï¿½n punto de* I *(condición de Jacobi).*

Entonces el funcional

$$J(h) = \int_I (A(t)h^2(t) + B(t)\dot{h}(t)^2)dt, \tag{56}$$

es estrictamente positivo en $E_0^1(I)$, *es decir,* $J(h) > 0$ *para todo* $h \in E_0^1(I)$ *con* $h \neq 0$.

DEMOSTRACIÓN. □

TEOREMA 4.9. *Si las funciones* A *y* B *satisfacen las hipótesis del Teorema 4.8 entonces las siguientes condiciones son equivalentes:*

1. A *y* B *satisfacen la condición de Jacobi en* $I = [a, b]$.
2. *La solución* u *de (53) asociada a los datos iniciales*

$$u(a) = 0 \quad \dot{u}(a) = 1, \tag{57}$$

no tiene un cero en $(a, b]$

DEMOSTRACIÓN. □

DEFINICIÓN 4.10. *Si la solución de la ecuación de Ricatti (52) asociada a los datos iniciales (57) tiene un cero* $c \in (a, b]$, *entonces llamaremos a* c *un* ***punto conjugado*** *de* a *con respecto a la ecuación de Ricatti.*

LEMA 4.11. *Las siguientes condiciones son equivalentes*

1. A *y* B *satisfacen la condición de Jacobi en* $I = [a, b]$.
2. *No existe un punto conjugado de* a *en* $(a, b]$ *con respecto a la ecuación diferencial (53).*

TEOREMA 4.12. *Supongamos que* $A \in C^0(I)$ *y* $B \in C^1(I)$ *y que*

1. $B(t) > 0$ *para todo* $t \in I$,
2. *El funcional dado por (56) es estrictamente positivo en* $E_0^1(I)$.

Entonces estas funciones satisfacen la condición de Jacobi en $I = [a, b]$.

DEMOSTRACIÓN. □

COROLARIO 4.13. *Supongamos que* $A \in C^0(I)$ *y* $B \in C^1(I)$ *y que*

1. $B(t) > 0$ *para todo* $t \in I$,
2. *El funcional dado por (56) es estrictamente positivo en* $E_0^1(I)$.

Entonces estas funciones satisfacen la condición de Jacobi en (a, b).

DEMOSTRACIÓN. □

TEOREMA 4.14. *Si $x \in E^1(I)$ es una trayectoria para la cual se minimiza la acción W y si la segunda derivada Fréchet de W no se anula en x entonces*

1. *Se satisface la ecuación de Euler en I:*
$$\frac{d}{dt}L_{\dot{x}}(t,x(t),\dot{x}(t)) = L_x(t,x(t),\dot{x}(t)).$$
2. *Se satisface la condición de Legendre en I:*
$$L_{\dot{x}\dot{x}}(t,x(t),\dot{x}(t)) \geq 0.$$
3. *Si $A^x \in C^0(I)$ y $B^x \in C^1(I)$, entonces estas funciones satisfacen la condición de jacobi en $(a,b]$.*

Recíprocamente, si la trayectoria $x \in E^1(I)$ satisface las siguientes condiciones,

1. *Se satisface la ecuación de Euler en I:*
$$\frac{d}{dt}L_{\dot{x}}(t,x(t),\dot{x}(t)) = L_x(t,x(t),\dot{x}(t)).$$
2. *Se satisface*
$$L_{x\dot{x}}(t,x(t),\dot{x}(t)) > 0 \quad \text{para todo} t \in I.$$
3. *$A^x \in C^0(I)$, $B^x \in C^1(I)$ y además satisfacen la condición de jacobi en $[a,b]$.*

Entonces el funcional de acción W tiene un mínimo local estricto en x.