Introducción al Cálculo Variacional

Gonzalo Galiano, 2003

Índice general

Introducción

Tres problemas clásicos

A continuación introducimos tres ejemplos clásicos del Cálculo de Variaciones, en los que se muestran los elementos fundamentales del problema tipo de optimización. Son éstos:

1. Un espacio de funciones, V, tal que $u : \Omega \to \mathbb{R}^q$, donde Ω es una abierto, normalmente acotado, de $\mathbb{R}^n$, de frontera, Γ, regular.
2. Restricciones sobre el conjunto de soluciones, que pueden imponerse bien sobre la frontera Γ, bien sobre el dominio Ω. Por ejemplo $u = 0$ en Γ, $u \geq \psi$ en Ω, etc. El conjunto de funciones que satisfacen estas restricciones es, en general, un subconjunto, U de V.
3. Un funcional $J : V \to \mathbb{R}$ de la forma siguiente:

$$J(u) := \int_\Omega L(x, u(x), u'(x))dx. \tag{1}$$

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

El problema de optimización consiste en hallar el mínimo, $u \in U$, del funcional J.

El problema de la braquistocrona. El problema de la braquistocrona, o curva de descenso más rápido, 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.

Entre todas las curvas que unen los puntos A y B, se desea hallar aquella a lo largo de la cual un punto material, moviéndose bajo la fuerza de la gravedad desde A llega al punto B en el menor tiempo.

Para resolver este problema debemos considerar todas las posibles curvas que unen A y B. A una determinada curva, γ, le corresponderá un valor determinado, T, del tiempo invertido para el descenso del punto material a lo largo de ella. El tiempo, T, dependerá de la elección de γ. De todas las curvas que unen A con B debemos hallar aquella a la que corresponda el menor valor de T. El problema puede plantearse de la siguiente forma.

Tracemos un plano vertical que pase por los puntos A y B. La curva de más rápido descenso debe evidentemente estar en él, así que podemos restringirnos a curvas sobre dicho plano. Tomemos el punto A como el origen de coordenadas, el eje OX apuntando en la dirección de la gravedad y sea $B = (x_1, y_1)$, con $x_1 > 0$ y $y_1 \geq 0$. Consideremos una curva arbitraria descrita por la ecuación

$$y = y(x) \quad 0 \leq x \leq x_1, \tag{2}$$

donde y es una función regular. Como la curva pasa por A y B, la función y debe verificar

$$y(0) = 0, \quad y(x_1) = y_1. \tag{3}$$

El movimiento de la masa puntual puede describirse por medio de la ley de la conservación de la energía, $E_c + E_p = cte.$, del siguiente modo: En el punto A, en el que asumimos que la velocidad inicial es nula, se tiene

$$E_c + E_p = E_p = mgh_A = E,$$

donde $E > 0$ es una constante y h_A es la altura a la que se encuentra el punto A. En cualquier punto por debajo será

$$\frac{1}{2}mv^2 + mgh = E,$$

luego

$$v^2 = 2g(h_A - h),$$

y tomando la coordenada vertical como $x = h_A - h$, deducimos que la velocidad del movimiento del punto material es

$$v \equiv \frac{ds}{dt} = \sqrt{2gx},$$

siendo s una parametrización de la trayectoria del punto material. Deducimos que

$$dt = \frac{ds}{\sqrt{2gx}},$$

y como la longitud de arco de la curva viene dada por

$$ds = \sqrt{1 + y'(x)^2}dx,$$

tenemos que el tiempo empleado a lo largo de la curva y viene dado por

$$J(y) = \int_0^{x_1} \Big(\frac{1 + y'(x)^2}{2gx}\Big)^{1/2} dx. \tag{4}$$

Hallar la braquistocrona es equivalente a resolver el siguiente problema de mínimos: entre todas las posibles funciones (2) que verifiquen las condiciones (3), hallar la que corresponda al menor valor de la integral (4).

El problema de las geodésicas. Las geodésicas son aquellas curvas contenidas en una superficie regular que minimizan la distancia entre dos puntos de la misma. Enunciaremos este problema de dos formas:

1. Consideremos una superficie regular $S \subset \mathbb{R}^3$ definida por la parametrización

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

con $(u, v) \in [u_0, u_1] \times [v_0, v_1]$. Cualquier curva contenida en S puede parametrizarse en la forma $t : [t_1, t_2] \to (u(t), v(t))$.

El elemento de arco de las curvas contenidas en S está determinado por la primera forma fundamental:

$$ds^2 := Eu'^2 + 2Fu'v' + Gv'^{\,2},$$

con

$$E := x_u^2 + y_u^2 + z_u^2, \quad F := x_u x_v + y_u y_v + z_u z_v, \quad G := x_v^2 + y_v^2 + z_v^2.$$

De modo que la longitud del arco entre los puntos correspondientes a los valores t_1 y t_2 es

$$J(u, v) = \int_{t_1}^{t_2} \sqrt{Eu'^2 + 2Fu'v' + Gv'^{\,2}}\,dt,$$

que es el funcional a minimizar.

2. Si la superficie viene dada de forma implícita por $\varphi(x, y, z) = 0$ y representamos una curva sobre ella de forma paramétrica, $(x(t), y(t), z(t))$, debemos minimizar el funcional

$$J(x, y, z) = \int_{t_0}^{t_1} \left(x'(t)^2 + y'(t)^2 + z'(t)^2\right)^{1/2} dt.$$

Además, las funciones x, y y z deben someterse a la condición $\varphi(x(t), y(t), z(t)) = 0$ para $t \in [t_0, t_1]$. Es lo que se llama un problema variacional con restricciones de igualdad.

El problema isoperimétrico. De entre todas las curvas de longitud λ dada, que unen el punto $(0, 0)$ con un punto variable $(\xi, 0)$, encontrar aquella que, junto con el eje OX, encierra una superficie máxima. El problema es, pues, el de hallar una función, u y un número, ξ tales que $u(0) = 0$, $u(\xi) = 0$, $u \geq 0$ y que minimicen el funcional

$$J(u, \xi) := -\int_0^{\xi} u,$$

y satisfagan la restricción

$$\int_0^{\xi} \sqrt{1 + |u'|^2} = \lambda.$$

Métodos de resolución de los problemas variacionales

Existen dos aproximaciones fundamentales a la resolución de los problemas variacionales.

Métodos indirectos. La primera de estas aproximaciones es la heredada de los métodos de minimización de funciones (dimensión finita) vía el cálculo diferencial. Este método proporciona condiciones necesarias y condiciones suficientes que dan lugar a una base metodológica para la resolución de problemas variacionales, la cual está íntimamente ligada a la teoría de ecuaciones diferenciales.

Métodos directos. La idea fundamental es la extensión del Teorema de Weierstrass a funciones definidas en espacios de dimensión infinita, que tendrá un enunciado del tipo:

Teorema. *Sea $J : V \to \mathbb{R}$ un funcional definido en un espacio de funciones V dotado de cierta noción de convergencia para la que V es compacto y J es semicontinuo inferiormente. Entonces existe un mínimo de J en V.*

A partir de este teorema, se procede del siguiente modo:

1. Se elije la clase de funciones V junto con una noción adecuada de convergencia para la que V sea completo.
2. Hay que mostrar que J está bien definido en V y que está acotado inferiormente, de modo que $\inf_{u\in V} J(u)$ sea finito. Esto implica que se puede construir una sucesión minimizante, $u_k \in V$, tal que $J(u_k) \to \inf_{u\in V} J(u)$.
3. Debemos probar que J es semicontinuo inferiormente (secuencialmente), es decir, que $u_k \to u$ implica

$$J(u) \leq \lim_{k\to\infty} J(u_k).$$

4. Finalmente, debemos demostrar que V es compacto (secuencialmente) con respecto a la convergencia considerada en 1.

Las hipótesis del Teorema de Weierstrass 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. La continuidad suele deducirse de un análisis directo de la función a minimizar.

Sin embargo, el Teorema de Riesz 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 nos permitan usar una generalización del Teorema de Weierstrass.

Puesto que los conjuntos cerrados y acotados, en el sentido de la topología fuerte, de un espacio de Banach no son compactos, puede esperarse que si se reduce la cantidad de abiertos mediante la introducción de una nueva topología, la cantidad de cerrados y, por tanto, de compactos, aumente. Esto resulta ser así. En particular, cualquier subconjunto cerrado y acotado de un espacio de Banach es relativamente compacto respecto la *topología débil*.[1]

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 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 cerrado y acotado respecto la topología débil, la verificación práctica de estas propiedades dista de ser sencilla. Por ello, una de las cuestiones centrales es la de la búsqueda de condiciones expresadas respecto a la topología fuerte que impliquen las correspondientes respecto a la topología débil. En este contexto la convexidad de conjuntos y funciones juega un papel fundamental.

[1]Es la topología menos fina que hace continuas a las aplicaciones lineales

CAPÍTULO 1

Analogías entre el Cálculo Diferencial y el Cálculo Variacional

1. Optimización en dimensión finita

Dada una función $f : U \to \mathbb{R}$, con U un abierto de $\mathbb{R}^n$, el programa usual que utiliza el cálculo diferencial para la localización de puntos de mínimo es el siguiente.

En primer lugar, debemos asumir que la función posee cierta regularidad, típicamente que la función posea derivadas parciales segundas continuas, es decir $f \in C^2(U)$.

En segundo lugar, resolvemos la ecuación de los puntos críticos, es decir, hallamos $x_c \in U$ tales que

$$\nabla f(x_c) = 0.$$

En tercer lugar, evaluamos la matriz hessiana de f en los puntos críticos, y comprobamos si dicha matriz es definida positiva, es decir, si los autovalores asociados a la matriz

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_c) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_c) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_c) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_c) \\ \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_c) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x_c) \end{pmatrix}$$

son positivos. En caso afirmativo, x_c es un punto de mínimo local para f, es decir, existe una bola de radio ρ centrada en x_c, $B_\rho(x_c)$, tal que

$$f(x_c) \leq f(x) \quad \text{para todo } x \in B_\rho(x_c).$$

2. Paso a dimensión infinita

La cuestión que surge a continuación es la de extender la metodología de minimización de funciones definidas en espacios de dimensión finita a los funcionales descritos en los ejemplos de la introducción, los cuales se hayan definidos en espacios de funciones (de dimensión infinita). Por ejemplo, en el problema de la braquistocrona se trata de minimizar el funcional

$$J(u) = \int_0^{x_1} \sqrt{\frac{1 + u'(x)^2}{2gx}} dx,$$

con la función $u : [0, x_1] \to \mathbb{R}$ satisfaciendo

$$u(0) = 0, \quad u(x_1) = u_1,$$

y algún requerimiento de regularidad que implique que el funcional J esté bien definido (sea finito).

Siguiendo los pasos del programa de dimensión finita, consideramos un funcional $J : U \to \mathbb{R}$, con $U \subset V$, siendo V un espacio de funciones regulares (hay varias elecciones) y siendo U un subespacio de V, que en el ejemplo de la braquistocrona viene dado por

$$U = \{u \in V : u(0) = 0, \quad u(x_1) = u_1\}.$$

En el caso de dimensión finita asumimos que la función objetivo es dos veces diferenciable con continuidad. La continuidad de funciones expresa el hecho de que a pequeñas variaciones de la variable independiente se siguen pequeñas variaciones del valor de la función. En términos de $\varepsilon - \delta$, f es continua en $x_0 \in \Omega$ si para todo $\varepsilon > 0$ podemos hallar un $\delta > 0$ tal que para todo x satisfaciendo $\|x - x_0\| < \delta$ se consigue que $\|f(x) - f(x_0)\| < \varepsilon$.

Ahora, en el caso de funcionales, ¿qué significan pequeñas variaciones de la variable independiente, siendo ésta una función? Hay distintas elecciones que se pueden hacer a este respecto. Por ejemplo, uno puede llamar próximas a dos funciones continuas si sus ordenadas están próximas. En este sentido podemos introducir la norma del supremo y decir que u_0 y u están próximas respecto esta norma si

$$\|u - u_0\|_{C(\Omega)} \equiv \sup \{|u(x) - u_0(x)| : x \in \Omega\} < \delta,$$

donde Ω es el dominio de definición de u y u_0. Otras elecciones típicas son que tanto las funciones como sus primeras derivadas o, más en general, derivadas de orden k, estén cercanas, dando lugar a los espacios de funciones $C^k(\Omega)$, con norma

$$\|u\|_{C^k(\Omega)} \equiv \sup \left\{ |u(x)| + \ldots + |D^k u(x)| : x \in \Omega \right\}.$$

Otros espacios funcionales que aparecen frecuentemente en las aplicaciones son los espacios $L^p(\Omega)$, en particular, el espacio de Hilbert $L^2(\Omega)$, con norma dada por

$$\|u\|_{L^2(\Omega)} \equiv \Big(\int_\Omega u^2(x) dx \Big)^{1/2},$$

y el espacio de Sobolev $H^1(\Omega)$ (que también es un espacio de Hilbert), con norma

$$\|u\|_{H^1(\Omega)} \equiv \|u\|_{L^2(\Omega)} + \|Du\|_{L^2(\Omega)}.$$

Claramente, los conceptos de proximidad entre funciones dependerán del espacio en que éstas se sitúen, como muestra el siguiente ejemplo.

Ejemplo 1. Consideremos la función de Heavyside

$$H(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x \geq 0, \end{cases}$$

y una sucesión de funciones continuas que la aproximan, dadas por

$$H_n(x) = \begin{cases} 0 & \text{si } x < 0, \\ nx & \text{si } 0 \leq x \leq 1/n, \\ 1 & \text{si } x > 1/n, \end{cases}$$

de modo que

$$|H(x) - H_n(x)| = \begin{cases} 0 & \text{si } x < 0, \\ 1 - nx & \text{si } 0 \leq x \leq 1/n, \\ 0 & \text{si } x > 1/n, \end{cases}$$

y, por tanto, $\|H - H_n\|_{C(\mathbb{R})} = 1$, y ningún término de la sucesión de funciones H_n está arbitrariamente cerca de H en el sentido de la norma del supremo. Sin embargo, respecto la norma de $L^2(\mathbb{R})$, tenemos que

$$\|H - H_n\|_{L^2(\mathbb{R})} = \Big(\int_0^{1/n} (1 - nx)^2 dx \Big)^{1/2} = \frac{1}{\sqrt{3n}},$$

de modo que para n suficientemente grande podemos hacer H y H_n tan próximas como queramos. $\square$

Aunque, a primera vista, parecería lo más natural el buscar los puntos de mínimo de un funcional en espacios *grandes* como $C(\Omega)$ o $L^p(\Omega)$, esto no es así. La razón es que, como puede intuirse, la continuidad de funcionales del tipo variacional

$$J(u) = \int_\Omega L(x, u(x), u'(x))dx,$$

va a depender de la continuidad de u', lo que conduce a la consideración del espacio $C^1(\Omega)$ o, al menos, $H^1(\Omega)$. Puesto que las técnicas que usaremos pasan por métodos analíticos usuales, como el paso al límite, la continuidad del funcional jugará un papel importante.

El siguiente paso en nuestra generalización del método diferencial al caso de funcionales es la introducción de la diferencial para el cálculo de los puntos críticos. En el caso del cálculo diferencial se comienza introduciendo el concepto de derivada parcial o el algo más general de derivada direccional. En el caso de funcionales, comenzamos con el concepto más débil, el de *variación*, que será también el que más utilicemos.

Para motivar su introducción, supongamos que tenemos dado un funcional $J : V \to \mathbb{R}$ y consideremos la siguiente función real

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

con $u_0, v \in V$ fijos. Si J tiene un mínimo local en u_0, es decir,

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

entonces F tiene también un mínimo local en $t = 0$. Suponiendo que F sea dos veces derivable, debe satisfacerse

$$F'(0) = 0 \quad \text{y } F''(0) \geq 0. \tag{5}$$

Si J es una función diferenciable, entonces $F'(0)$ corresponde a la derivada direccional de J en u_0 en la dirección de v. Si J es un funcional, $F'(0)$ corresponde a la *variación primera de J en u_0 en la dirección de v*. Como en el caso de dimensión finita, la anulación de la variación primera será una condición necesaria para la existencia de un extremo.

Definición 1. *La **variación** $n-$**ésima de** J **en un punto** u_0 **en la dirección** $v \in V$ 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.

Si $\delta J(u_0; \cdot) : V \to \mathbb{R}$ define un funcional lineal y continuo, entonces escribimos $\delta J(u_0; \cdot) \equiv \delta_{u_0} J$.

Tenemos, pues, que dado un funcional $J : V \to \mathbb{R}$, si la primera variación existe y es un funcional lineal y continuo en un punto de mínimo, u_0, de J entonces las condiciones (5) se traducen en

$$\delta_{u_0} J = 0 \quad \text{y} \quad \delta_{u_0}^2 J \geq 0. \tag{6}$$

En los capítulos siguientes explotaremos estas propiedades para hallar condiciones necesarias y suficientes sobre J que impliquen la existencia de un mínimo local.

Ejemplo 2. Consideremos el funcional dado por

$$J(u) = \int_\Omega |\nabla u(x)|^2 dx,$$

definido para $u \in U \equiv H^1(\Omega)$. Tenemos que, para $u, v \in U$ fijados, la función

$$F_v(t) = J(u + tv) = \int_\Omega \left(|\nabla u|^2 + 2t\nabla u \cdot \nabla v + t^2 |\nabla v|^2\right) dx$$

tiene derivada en $t = 0$, y por tanto primera variación, dada por

$$\delta_u J(v) = 2 \int_\Omega \nabla u \cdot \nabla v,$$

que es lineal y continua sobre $H^1(\Omega)$. □

Ejercicio 1. Hallar la primera variación de los funcionales asociados al problema de la braquistocrona y al problema isoperimétrico. □

A continuación introducimos una noción más general de diferencial, llamada diferencial Fréchet, la cual generaliza la noción de diferencial de una función en dimensión finita.

Definición 2. *Sea V un espacio de Banach y $U \subset V$ un conjunto abierto. Entonces se dice que $J : U \to \mathbb{R}$ es* ***diferenciable Fréchet en*** $u_0 \in U$ *si existe una aplicación lineal y continua* $D_{u_0}J : V \to \mathbb{R}$ *tal que*

$$J(u_0 + h) - J(u_0) = D_{u_0}J(h) + o(u_0, h) \quad \textit{para todo } h \in V, \quad u_0 + h \in U, \tag{7}$$

donde $o(u_0, \cdot) : V \to \mathbb{R}$ *satisface* $o(u_0, 0) = 0$ *y*

$$\lim_{h \to 0} \frac{|o(u_0, h)|}{\|h\|} = 0. \tag{8}$$

Las propiedades de la derivada Fréchet son análogas a las de la diferencial en dimensión finita, y se demuestran de modo análogo. Recogemos aquí algunas para futuras referencias.

1. La diferencial Fréchet, si existe, es única.
2. La diferencial Fréchet satisface la regla de la cadena.
3. Si J es diferenciable Fréchet en u entonces J es continuo en u.
4. Si J es diferenciable Fréchet entonces se satisface un teorema del valor medio.

Es un ejercicio sencillo el comprobar que la existencia de la diferencial Fréchet implica la existencia de la primera variación. Recíprocamente, si la primera variación define un funcional lineal y continuo sobre V, entonces la primera variación y la diferencial Fréchet coinciden.

Observación 1. 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 los siguientes capítulos. □

El siguiente ejemplo, extraído de la teoría de dimensión finita, muestra un caso en el que la primera variación no es ni lineal ni continua.

Ejemplo 3. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

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 $F(t) = f(th_1, th_2)$ tiene por derivada

$$F'(t) = h_1 \quad \text{si } 0 < t < |h_2|/h_1^2,$$

con lo cual, para tal h, tenemos $\delta f(0)h = h_1$. Por tanto, $\delta f(0)$ es discontinua en todos los puntos del eje OX_1, excepto el origen. Además, puesto que f no es continua en el origen, no es diferenciable Fréchet. □

CAPÍTULO 2

La ecuación de Euler y las condiciones de Legendre

1. Problemas variacionales con fronteras fijas en una variable

Comenzamos con un ejemplo de problema variacional que es de gran importancia en las aplicaciones: minimizar el funcional $J : U \subset V \to \mathbb{R}$, de la forma

$$J(u) = \int_{x_0}^{x_1} L(x, u(x), u'(x))dx,$$

donde, V es el espacio de funciones derivables con continuidad $C^1([x_0, x_1])$ y u toma valores constantes en la frontera, de modo que definimos

$$U = \left\{u \in C^1([x_0, x_1]) : u(x_1) = u_1, \quad u(x_2) = u_2\right\}.$$

Asumiendo que el funcional J posee primera variación y que existe un punto de mínimo, u, sabemos por el capítulo anterior que la variación primera debe anularse en dicho punto, es decir $\delta_u J = 0$. Veamos cómo se traduce este hecho en relación a la función L, llamada *lagrangiana*.

Sea $v \in C_0^1([x_0, x_1]) = \left\{v \in C^1([x_0, x_1]) : v(x_1) = 0, \quad v(x_2) = 0\right\}$, de modo que $u + tv \in U$ para todo $t \in \mathbb{R}$, y definamos la función real

$$F(t) = J(u + tv) = \int_{x_0}^{x_1} L(x, u(x) + tv(x), u'(x) + tv'(x))dx,$$

cuya primera derivada en $t = 0$ es, por hipótesis,

$$F'(0) = \delta_u J(v) = 0.$$

Para $t \in \mathbb{R}$ cualquiera, tenemos que

$$F'(t) = \int_{x_0}^{x_1} \big(L_u(x, u(x) + tv(x), u'(x) + tv'(x))v(x) \\ + L_{u'}(x, u(x) + tv(x), u'(x) + tv'(x))v'(x)\big)dx,$$

de modo que, en $t = 0$

$$0 = F'(0) = \int_{x_0}^{x_1} \big(L_u(x, u(x), u'(x))v(x) + L_{u'}(x, u(x), u'(x))v'(x)\big)dx.$$

Integrando por partes el segundo sumando de la integral y usando que $v \in C_0^1([x_0, x_1])$ obtenemos que la condición necesaria para que u sea un mínimo es

$$(9) \qquad \int_{x_0}^{x_1} \Big(L_u(x, u(x), u'(x)) - \frac{d}{dx} L_{u'}(x, u(x), u'(x))\Big)v(x)dx = 0,$$

para todo $v \in C_0^1([x_0, x_1])$. El siguiente resultado, conocido como el Lema Fundamental del Cálculo de Variaciones, nos proporciona la primera condición necesaria para la existencia de un mínimo:

Lema 1. *Sea $g \in C([x_0, x_1])$ una función continua tal que*

$$\int_{x_0}^{x_1} g(x)v(x)dx = 0$$

para toda $v \in C_0^1([x_0, x_1])$. Entonces $g(x) = 0$ para todo $x \in [x_0, x_1]$.

Demostración. Supongamos que existe $x^* \in (x_0, x_1)$ tal que $g(x^*) > A > 0$. Por continuidad, también existe un $\varepsilon > 0$ tal que $g(x) > A/2$ en $(x^* - \varepsilon, x^* + \varepsilon) \subset (x_0, x_1)$. A continuación, construimos una función regular positiva con soporte contenido en este subintervalo. Para ello, consideramos la función C^∞ definida en $\mathbb{R}$

$$f(x) = \begin{cases} \exp(-(1-x^2)^{-1}) & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso,} \end{cases}$$

y, a partir de ella, construimos la función

$$v(x) = \frac{1}{\varepsilon} f(\frac{x - x^*}{\varepsilon}),$$

con soporte en $(x^* - \varepsilon, x^* + \varepsilon)$. Tenemos entonces la siguiente contradicción

$$0 = \int_{x_0}^{x_1} g(x)v(x)dx = \int_{x^*-\varepsilon}^{x^*+\varepsilon} g(x)v(x)dx \geq \frac{A}{2} \int_{x^*-\varepsilon}^{x^*+\varepsilon} v(x)dx = \frac{A}{2} \int_{-1}^{1} f(x)dx > 0.$$

□

Aplicando el Lema 1 a la identidad (9) deducimos la **Ecuación de Euler**, condición necesaria para la realización de un mínimo (y, en general, de un extremo) en u:

$$L_u(x, u(x), u'(x)) - \frac{d}{dx} L_{u'}(x, u(x), u'(x)) = 0 \quad \text{para todo } x \in (x_1, x_2).$$

La ecuación de Euler juega un papel fundamental en el cálculo de variaciones y es, en general, una ecuación diferencial de segundo orden. A continuación mostramos algunos casos especiales en los que la ecuación de Euler puede reducirse a una ecuación de primer orden o en los que sus soluciones pueden hallarse por medio de cuadraturas.

1. *El integrando no depende de u.* En este caso la ecuación de Euler es

$$\frac{d}{dx} L_{u'} = 0,$$

de modo que obtenemos la ecuación diferencial de primer orden

$$L_{u'}(x, u(x), u'(x)) = C,$$

con C una constante. Resolviendo esta ecuación respecto u' obtendremos una ecuación del tipo

$$u'(x) = F(x; C),$$

que se resuelve mediante el cálculo de una primitiva de $F(x; \cdot)$.

Ejemplo 4. El funcional

$$J(u) = \int_0^{x_1} \frac{\sqrt{1 + u'^2}}{\nu(x)} dx$$

representa el tiempo invertido en el desplazamiento de un punto material que se mueve a velocidad $\nu(x) > 0$, a lo largo de la curva u, desde $(0, 0)$ a (x_1, u_1). La ecuación de Euler es

$$u'(x) = C\nu(x)\sqrt{1 + u'^2} \implies u'(x)^2 = \frac{C^2\nu(x)^2}{1 - C^2\nu(x)^2},$$

de donde

$$u(x) = \pm \int_0^x \frac{C\nu(x)}{\sqrt{1 - C^2\nu(x)^2}} dx.$$

Observemos que si la velocidad es constante entonces la solución es una línea recta. Para $\nu(x) = bx$ la solución es una circunferencia. En efecto, $\nu'(x) = b$, luego

$$u(x) = \pm \frac{1}{Cb} \int_0^x \frac{C^2\nu(x)\nu'(x)}{\sqrt{1 - C^2\nu(x)^2}} dx = \mp \frac{1}{Cb} \sqrt{1 - (Cbx)^2},$$

que es la ecuación de una circunferencia de radio $1/Cb$. Finalmente, para $\nu(x) = \sqrt{x}$, retomamos el problema de la braquistocrona. □

2. *El integrando no depende de x.* La ecuación de Euler es

$$0 = L_u - \frac{d}{dx} L_{u'} = L_u - L_{u'u} u' - L_{u'u'} u''.$$

Multiplicando esta ecuación por u' obtenemos

$$0 = L_u u' - L_{u'u} u'^2 - L_{u'u'} u'' u' = \frac{d}{dx}(L - u' L_{u'}),$$

de modo que la ecuación de Euler se reduce a

$$L - u' L_{u'} = C,$$

donde C es una constante.

Ejemplo 5. Determinar la curva diferenciable, con los puntos extremos fijos, que al girar alrededor del eje de las abscisas forme una superficie de área mínima.

El área de una superficie de revolución viene dada por

$$J(u) = 2\pi \int_{x_1}^{x_2} u\sqrt{1 + u'^2} dx.$$

La ecuación de Euler es

$$u\sqrt{1+u'^2} - \frac{uu'^2}{\sqrt{1+u'^2}} = C,$$

que, simplificando, queda

$$\frac{u}{\sqrt{1+u'^2}} = C \implies u' = \sqrt{\frac{u^2-C^2}{C^2}}.$$

Separando variables, obtenemos

$$dx = \frac{C\,du}{\sqrt{u^2-C^2}} \implies x + C_1 = C \ln \frac{u+\sqrt{u^2-C^2}}{C},$$

de donde se deduce que

$$u(x) = C \cosh \frac{x+C_1}{C}, \tag{10}$$

que es una catenaria ($\cosh t = \frac{1}{2}(e^t + e^{-t})$). Finalmente, las constantes C y C_1 se determinan a partir de las condiciones de contorno $u(x_1) = u_1$ y $u(x_2) = u_2$. Se dan tres casos:

a) Si se puede trazar una única curva de la forma (10) por los puntos (x_1, u_1) y (x_2, u_2), entonces la catenaria da la solución del problema.

b) Si hay dos extremales que puedan trazarse por los puntos dados, entonces uno es solución y el otro no.

c) Si no hay ninguna curva de la forma (10) que pase por (x_1, u_1) y (x_2, u_2), entonces no se alcanza un mínimo en la clase de superficies de revolución regulares.

□

3. *El integrando no depende de* u'. La ecuación de Euler es

$$L_u(x, u(x)) = 0,$$

que es una ecuación algebraica, no diferencial.

Supongamos ahora que J posee también segunda variación, de modo que en un mínimo local, u, debe satisfacerse $\delta_u^2 J \geq 0$, y analicemos las repercusiones sobre el lagrangiano, L. Tenemos

$$\begin{aligned} F''(t) = \int_{x_0}^{x_1} \big(& L_{uu}(x, u(x)+tv(x), u'(x)+tv'(x))v(x)^2 \\ & + 2L_{uu'}(x, u(x)+tv(x), u'(x)+tv'(x))v(x)v'(x) \\ & + L_{u'u'}(x, u(x)+tv(x), u'(x)+tv'(x))v'(x)^2\big)\,dx, \end{aligned}$$

de modo que la condición $F''(0) = \delta_u^2 J(v) \geq 0$ implica

$$\int_{x_0}^{x_1} \big(L_{uu}(x,u(x),u'(x))v(x)^2 + 2L_{uu'}(x,u(x),u'(x))v(x)v'(x) + L_{u'u'}(x,u(x),u'(x))v'(x)^2\big)dx \geq 0. \tag{11}$$

Lema 2. *Sean $F_i : [x_0,x_1] \to \mathbb{R}$, $i = 1,2,3$ funciones continuas y supongamos que la forma cuadrática*

$$Q(v) = \int_{x_0}^{x_1} (F_1(x)v(x)^2 + F_2(x)v(x)v'(x)) + F_3(x)v'(x)^2)dx,$$

definida en $C_0^1([x_0,x_1])$, es no negativa. Entonces $F_3(x) \geq 0$ para todo $x \in (x_0,x_1)$.

Demostración. Supongamos que existe un $x^* \in (x_0,x_1)$ tal que $F_3(x^*) < -A < 0$. Por continuidad, también existe un $\varepsilon > 0$ tal que $F_3(x) < -A/2$ en $(x^* - \varepsilon, x^* + \varepsilon) \subset (x_0,x_1)$. Consideremos nuevamente la función f definida en la demostración del Lema 1 y, a partir de ella, la función

$$v_\varepsilon(x) = f(\frac{x-x^*}{\varepsilon}),$$

con soporte en $(x^* - \varepsilon, x^* + \varepsilon)$. Tenemos

$$\begin{aligned} Q(v_\varepsilon) &= \int_{x^*-\varepsilon}^{x^*+\varepsilon} (F_1(x)v_\varepsilon(x)^2 + F_2(x)v_\varepsilon(x)v_\varepsilon'(x) + F_3(x)v_\varepsilon'(x)^2)dx \\ &= \int_{-1}^{1} (F_1(x^*+\varepsilon y)f(y)^2 + F_2(x^*+\varepsilon y)f(y)\frac{f'(y)}{\varepsilon} + F_3(x^*+\varepsilon y)\frac{f'(y)^2}{\varepsilon^2})\, \varepsilon\, dy \\ &\leq \varepsilon \int_{-1}^{1} F_1(x^*+\varepsilon y)f(y)^2 dy + \int_{-1}^{1} F_2(x^*+\varepsilon y)f(y)f'(y) - \frac{A}{2\varepsilon}\int_{-1}^{1} f'(y)^2 dy, \end{aligned}$$

y como

$$\int_{-1}^{1} f'(y)^2 dy > 0,$$

se obtiene una contradicción tomando ε suficientemente pequeño. □

El Lema 2 aplicado a la identidad (11) nos permite obtener la llamada **Condición de Legendre** sobre el lagrangiano:

$$L_{u'u'}(x,u(x),u'(x)) \geq 0 \quad \text{para todo } x \in (x_0,x_1).$$

Observación 2. Integrando por partes el segundo sumando del miembro derecho de la desigualdad (11) obtenemos

$$\int_{x_0}^{x_1} \big((L_{uu} - \frac{d}{dx}L_{uu'})v^2 + L_{u'u'}v'^{\,2}\big)dx \geq 0.$$

Puesto que $v(x_0) = 0$, si v' es pequeña en todo el intervalo (x_0,x_1) entonces la función v será pequeña en dicho intervalo. El recíproco no es cierto, ya que podemos construir funciones pequeñas pero con una derivada tan grande como queramos. Esto muestra que el término dominante de esta integral es el que involucra a $L_{u'u'}$, y de ahí la condición de Legendre. □

Ejemplo 6. El problema de la braquistocrona. Retomemos el problema de la curva de descenso en tiempo mínimo introducido en la Introducción. Se trata de minimizar el funcional

$$J(u) = \int_0^{x_1} \big(\frac{1+u'(x)^2}{2gx}\big)^{1/2} dx,$$

que representa el tiempo de descenso de una partícula material, con velocidad inicial cero, a lo largo de la curva u que une los puntos $(0,0)$ y (x_1, u_1) debido únicamente a la acción de la gravedad. Aquí asumimos que u es diferenciable con continuidad. Las condiciones necesarias para la existencia de un mínimo vienen dadas por la Ecuación de Euler

$$\frac{d}{dx}\big(\frac{u'(x)}{(2gx(1+u'(x)^2))^{1/2}}\big) = 0 \quad \text{para todo } x \in (0, x_1),$$

y la condición de Lagrange

$$(1+u'(x)^2)^{-3/2}(2gx)^{-1/2} \geq 0 \quad \text{para todo } x \in (0, x_1).$$

La condición de Lagrange se satisface trivialmente. Integrando la Ecuación de Euler obtenemos

$$\frac{u'(x)}{(x(1+u'(x)^2))^{1/2}} = k \quad \text{para todo } x \in (0, x_1), \tag{12}$$

y para cierta constante k, la cual incluye el factor $\sqrt{2g}$. Particularizando en $x = x_1$ obtenemos

$$k^2 x_1 = \frac{u'(x_1)^2}{1+u'(x_1)^2} < 1,$$

con lo que $0 < k^2 x < 1$ para $x \in (0, x_1)$. De (12) obtenemos

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

y usando la parametrización

$$x(t) = \frac{1}{2k^2}(1-\cos t), \qquad y(t) = u(x(t)),$$

tenemos que y debe satisfacer

$$y'(t)^2 = \big(u'(x(t))x'(t)\big)^2 = \frac{k^2 x(t)}{1-k^2 x(t)} x'(t)^2 = \frac{1-\cos t}{1+\cos t}\frac{\operatorname{sen}^2 t}{4k^4} = \frac{(1-\cos t)^2}{4k^4},$$

de donde

$$y'(t) = \pm\frac{1}{2k^2}(1-\cos t).$$

Usando la condición de contorno $y(0) = u(x(0)) = 0$ y que $u \geq 0$ obtenemos

$$y(t) = \frac{1}{2k^2}(t - \operatorname{sen} t).$$

Para que $(x(t), y(t))$ sea una solución de nuestro problema es necesario que satisfaga la segunda condición de contorno, es decir, que exista un t_1 tal que

$$x_1 = x(t_1) = \frac{1}{2k^2}(1-\cos t_1), \qquad u_1 = y(t_1) = \frac{1}{2k^2}(t_1 - \operatorname{sen} t_1).$$

Observemos que, aunque la función y es invertible para todo $t \geq 0$, la función x sólo lo es en $(0, \pi)$, de modo que debemos imponer $t_1 \in (0, \pi)$. Físicamente, x debe ser estrictamente creciente, puesto que la partícula material no puede *subir* en contra de la gravedad. Tomando el cociente de estas dos condiciones obtenemos

$$\frac{y_1}{x_1} = \frac{t_1 - \operatorname{sen} t_1}{1 - \cos t_1}.$$

No es difícil comprobar que la función $f(t) = (t - \sin t)/(1 - \cos t)$ es creciente en $(0, \pi)$, y que su valor máximo es $f(\pi) = \pi/2$. Por tanto, si $u_1/x_1 < \pi/2$ la parametrización construida es una solución del problema. De lo contrario, el problema no tiene solución. □

2. Generalizaciones del problema con fronteras fijas

2.1. El caso de varias variables. Analizaremos el caso de dos variables, por simplicidad en la notación. El caso $n-$dimensional es una extensión directa. Consideremos un funcional de la forma

$$J(u) = \int_\Omega L(x, y, u, u_x, u_y) dxdy$$

con $u : \Omega \to \mathbb{R}$, siendo $\Omega \subset \mathbb{R}^2$. Supondremos, de nuevo, que u es un mínimo de J, con $u \in C^1(\Omega)$ tal que $u(x, y) = u_D(x, y)$ en $\partial\Omega$. Definiendo, para $v \in C_0^1(\Omega)$,

$$F(t) = J(u + tv) = \int_\Omega L(x, y, u + tv, u_x + tv_x, u_y + tv_y) dxdy,$$

obtenemos

$$\begin{aligned} F'(t) = \int_\Omega \big(L_u(x, y, u + tv, u_x + tv_x, u_y + tv_y) v + L_{u_x}(x, y, u + tv, u_x + tv_x, u_y + tv_y) v_x \\ + L_{u_y}(x, y, u + tv, u_x + tv_x, u_y + tv_y) v_y \big) dxdy, \end{aligned}$$

de modo que en $t = 0$ se tiene

$$\begin{aligned} 0 = \delta_u J(v) = F'(0) = \int_\Omega \big(L_u(x, y, u, u_x, u_y) v + L_{u_x}(x, y, u, u_x, u_y) v_x \\ + L_{u_y}(x, y, u_x, u_y) v_y \big) dxdy. \end{aligned}$$

Usando el teorema de la divergencia deducimos

$$\int_\Omega \big(L_u(x, y, u, u_x, u_y) - \frac{\partial}{\partial x} L_{u_x}(x, y, u, u_x, u_y) - \frac{\partial}{\partial y} L_{u_y}(x, y, u_x, u_y) \big) v dxdy = 0,$$

para toda $v \in C_0^1(\Omega)$. Finalmente, una extensión del Lema 1 a dos dimensiones nos permite obtener la Ecuación de Euler

$$L_u(x, y, u, u_x, u_y) - \frac{\partial}{\partial x} L_{u_x}(x, y, u, u_x, u_y) - \frac{\partial}{\partial y} L_{u_y}(x, y, u_x, u_y) = 0.$$

Ejercicio 2. Deducir la condición de Lagrange para el caso de dos variables. □

Ejercicio 3. Deducir las ecuaciones de Laplace y Poisson como las ecuaciones de Euler que minimizan los siguientes funcionales:

$$J(u) = \int_\Omega \big((u_x)^2 + (u_y)^2\big)dxdy$$

$$J(u) = \int_\Omega \big((u_x)^2 + (u_y)^2 + 2uf(x,y)\big)dxdy$$

□

Ejemplo 7. *La ecuación de ondas.* El principio variacional fundamental de la Mecánica es el principio de la acción estacionaria (mínima acción), el cual afirma que entre los movimientos admisibles de un sistema de puntos materiales se efectúa el movimiento que da un valor estacionario a la integral

$$\int_{t_0}^{t_1} (T-U)dt, \tag{13}$$

donde T es la energía cinética y U es la energía potencial del sistema.

Apliquemos este principio para deducir las ecuación del movimiento de una cuerda oscilante, la llamada *ecuación de ondas*. Situemos el origen de coordenadas en uno de los extremos de la cuerda. La cuerda en estado de reposo se encuentra , bajo la acción de la tensión, en cierta línea recta, en la cual situamos el eje de las abscisas. La desviación de la situación de equilibrio, $u(x,t)$, es función de la abscisa, x y del tiempo, t.

La energía potencial, U, de un elemento de una cuerda flexible es proporcional al alargamiento de la misma. El segmento de cuerda dx en estado de deformación tendrá una longitud $ds = \sqrt{1+(u_x)^2}dx$ y, por lo tanto, el alargamiento del elemento es

$$\big(\sqrt{1+(u_x)^2} - 1\big)dx.$$

Usando la fórmula de Taylor, obtenemos la aproximación

$$\sqrt{1+(u_x)^2} \sim 1 + \frac{1}{2}(u_x)^2,$$

de modo que podemos considerar que si u_x es pequeño, entonces la energía potencial del elemento es, aproximadamente, igual a $\frac{1}{2}k(u_x)^2dx$, donde k es un factor de proporcionalidad, un coeficiente de deformación. Tenemos, pues, que la energía potencial de toda la cuerda viene dada por

$$\frac{1}{2}\int_0^l k(u_x)^2dx.$$

Por otra parte, la energía cinética de la cuerda es

$$\frac{1}{2}\int_0^l \rho(u_t)^2dx,$$

donde ρ es la densidad de la cuerda. La integral (13) tiene la forma

$$J(u) = \frac{1}{2}\int_{t_0}^{t_1}\int_0^l \big(\rho(u_t)^2 - k(u_x)^2\big).$$

La ecuación del movimiento de la cuerda viene dada por la ecuación de Euler para el funcional J, que tiene la forma

$$\frac{\partial}{\partial t}(\rho u_t') - \frac{\partial}{\partial x}(k u_x') = 0,$$

que es la ecuación de ondas. □

Ejemplo 8. *Problema de Plateau.* Se trata de hallar la superficie de menor área con un contorno dado. Si el contorno viene dado por $u(x,y) = 0$, podemos introducir la parametrización $(x, y, u(x,y))$ de la superficie buscada, que debe minimizar el funcional

$$J(u) = \int_\Omega \sqrt{1 + u_x^2 + u_y^2}\,dxdy.$$

La ecuación de Euler tiene la forma

$$r(1+q^2) - 2spq + t(1+p^2) = 0, \tag{14}$$

donde

$$p = u_x, \quad q = u_y, \quad r = u_{xx}, \quad s = u_{xy}, \quad t = u_{yy}.$$

La ecuación (14) tiene un significado geométrico importante, que explicaremos a partir de la fórmula para la curvatura media de la superficie

$$M = \frac{1}{2}\Big(\frac{1}{\kappa_1} + \frac{1}{\kappa_2}\Big) = \frac{Eg - 2Ff + Ge}{2(EG - F^2)},$$

donde E, F y G son los coeficientes de la primera forma fundamental y e, f y g los de la segunda. Para la parametrización considerada, la fórmula de la curvatura media queda como

$$M = \frac{r(1+q^2) - 2spq + t(1+p^2)}{\sqrt{1+p^2+q^2}}.$$

Es decir, la curvatura media de cualquier superficie extremal es nula. Las superficies con esta propiedad son llamadas *superficies minimales.* □

2.2. El caso de varias incógnitas. Por simplicidad, nos restringimos al caso de dos incógnitas. Consideremos la minimización del funcional

$$J(u,v) = \int_{x_0}^{x_1} L(x, u, v, u', v')dx,$$

con $u(x_0) = u_0$, $u(x_1) = u_1$, $v(x_0) = v_0$, $v(x_1) = v_1$. Si definimos

$$F(t,\tau) = J(u + t\varphi, v + \tau\psi),$$

y asumimos que J tiene un extremo local en (u, v), entonces también lo tendrá F en $(0, 0)$, de modo que debe satisfacerse $\nabla F(0, 0) = (0, 0)$. Asumiendo la regularidad necesaria sobre L, tenemos que una condición necesaria será

$$0 = F_t(0,0) = \int_{x_0}^{x_1} \big(L_u(x,u,v,u',v')\varphi + L_{u'}(x,u,v,u',v')\varphi'\big)dx,$$
$$0 = F_\tau(0,0) = \int_{x_0}^{x_1} \big(L_v(x,u,v,u',v')\psi + L_{v'}(x,u,v,u',v')\psi'\big)dx,$$

que son ecuaciones análogas a la del caso unidimensional. De modo similar a este caso se deducen las ecuaciones de Euler

$$L_u(x,u,v,u',v') - \frac{d}{dx}L_{u'}(x,u,v,u',v') = 0,$$
$$L_v(x,u,v,u',v') - \frac{d}{dx}L_{v'}(x,u,v,u',v') = 0.$$

Ejemplo 9. *Geodésicas de superficies de revolución*

Consideremos una superficie regular de revolución, S, dada por la parametrización

$$x(u,v) = (f(v)\cos u, f(v)\operatorname{sen} u, g(v)).$$

Cualquier curva regular en S puede describirse mediante una parametrización del tipo $(u, v) \equiv (u(t), v(t))$ con $t \in (0, T)$. Como sabemos, la curva más corta que conecta dos puntos $(u(0), v(0)) = (u_0, v_0)$ y $(u(T), v(T)) = (u_T, v_T)$ se llama una *geodésica*.

La longitud de arco entre estos dos puntos viene dada por

$$J(u,v) = \int_0^T L(t,u,v,u',v')dt = \int_0^T I(t,u,v,u',v')^{1/2}dt,$$

donde I es la primera forma fundamental, que en el caso de superficies de revolución tiene la forma

$$I(u,v,u',v') = f(v)^2u'^2 + (f'(v)^2 + g'(v)^2)v'^2.$$

Las ecuaciones de Euler viene dadas por

$$\frac{d}{dt}\frac{f(v)^2u'}{I(u,v)^{-1/2}} = 0,$$
$$\frac{f'(v)^2u'^2 + (f'(v)f''(v) + g'(v)''(v))v'^2}{I(u,v)^{1/2}} - \frac{d}{dt}\frac{(f'(v)^2 + g'(v))v'}{I(u,v)^{1/2}} = 0.$$

En el sencillo caso del cilindro, podemos tomar $f(v) = 1$ y $g(v) = v$. Las ecuaciones de Euler son

$$\frac{d}{dt}\frac{u'}{I(u,v)^{-1/2}} = 0, \quad \frac{d}{dt}\frac{v'}{I(u,v)^{-1/2}} = 0,$$

es decir

$$u' = C_1\sqrt{u'^2 + v'^2} \quad v' = C_2\sqrt{u'^2 + v'^2},$$

de donde deducimos que $u' = cv'$, o bien, $u(t) = cv(t) + \hat{c}$, es decir, una familia biparamétrica de curvas helicoidales sobre el cilindro. □

2.3. Funcionales que dependen de las derivadas de orden superior. Consideremos el funcional

$$J(u) = \int_{x_0}^{x_1} L(x, u, u', u'')dx,$$

y supongamos que $u \in C^2([x_0, x_1])$ es un extremo local sujeto a las condiciones

$$u(x_0) = u_0, \quad u(x_1) = u_1, \quad u'(x_0) = u_0', \quad u(x_1) = u_1'$$

No es difícil ver que la ecuación de Euler viene dada por

$$L_u - \frac{d}{dx}L_{u'} + \frac{d^2}{dx^2}L_{u''} = 0$$

2.4. Problemas variacionales con restricciones. En las secciones anteriores hemos estudiado las condiciones necesarias para la existencia de un extremo de funcionales definidos en clases de funciones que toman valores constantes en la frontera. Como vimos en la Introducción, existen aplicaciones en las que es natural considerar ciertas restricciones adicionales sobre el conjunto de funciones admisibles. Entre la más importantes se hallan las restricciones de tipo isoperimétrico y las restricciones de igualdad, que estudiamos a continuación.

2.4.1. Restricciones de igualdad. En este problema, se trata de hallar $\mathbf{u} = (u_1, \ldots, u_n)$ para la cual el funcional

$$J(\mathbf{u}) = \int_{x_0}^{x_1} L(x, \mathbf{u}, \mathbf{u}')dx$$

tiene un extremo, con las funciones admisibles satisfaciendo las condiciones usuales de frontera, y tales que

$$\varphi_i(x, \mathbf{u}, \mathbf{u}') = 0, \quad i = 1, \ldots, m, \quad m < n,$$

donde φ_i son ciertas funciones regulares dadas. Aquí asumimos que las restricciones son independientes, es decir, que

$$\frac{\partial(\varphi_1, \ldots, \varphi_m)}{\partial(u_1, \ldots, u_m)} \neq 0.$$

Teorema 1. *Si* $\mathbf{u}$ *realiza un extremo del funcional* J *y satisface las restricciones*

$$\varphi_i(x, \mathbf{u}, \mathbf{u}') = 0, \quad i = 1, \ldots, m, \quad m < n,$$

entonces también satisface las ecuaciones de Euler para el funcional

$$J^*(\mathbf{u}) = \int_{x_0}^{x_1} L^*(x, \mathbf{u}, \mathbf{u}')dx,$$

donde

$$L^*(x, \mathbf{u}, \mathbf{u}') = L(x, \mathbf{u}, \mathbf{u}') + \sum_{i=1}^{m} \lambda_i(x)\varphi_i(x, \mathbf{u}, \mathbf{u}').$$

Las funciones $(\lambda_1, \ldots, \lambda_m)$ *y* $\mathbf{u} = (u_1, \ldots, u_n)$ *se determinan a partir de las ecuaciones de Euler*

$$L^*_{u_j} - \frac{d}{dx} L^*_{u_j} = 0, \quad j = 1, \ldots, n,$$

y de las restricciones

$$\varphi_i = 0, \quad i = 1, \ldots, m.$$

Demostración. Lo demostraremos para el caso particular en el que los enlaces no dependen de $\mathbf{u}'$. Sea $\mathbf{u}$ un extremo restringido de J. La condición fundamental de extremo, $\delta_{\mathbf{u}} J = 0$, tiene la forma habitual

$$\int_{x_0}^{x_1} \big(\sum_{j=1}^{n} L_{u_j} v_j + L_{u'_j} v'_j\big) dx = 0.$$

Sin embargo, no es posible aplicar el lema fundamental para deducir las ecuaciones de Euler ya que las funciones u_j están sometidas a los m enlaces $\varphi_i = 0$ y, por tanto, las variaciones v_j no son arbitrarias. En efecto, asumamos que para $t \in (0, \varepsilon)$ las variaciones satisfacen las restricciones [1], es decir,

$$\varphi_i(x, u_1 + tv_1, \ldots, u_n + tv_n) = 0.$$

Entonces

$$\varphi_i(x, u_1 + tv_1, \ldots, u_n + tv_n) - \varphi_i(x, u_1, \ldots, u_n) = 0, \quad t \in (0, \varepsilon)$$

de donde

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial u_j}(\mathbf{u}) v_j = 0, \quad i =, 1 \ldots, m,$$

y, por tanto, solo puede haber $n - m$ variaciones arbitrarias. Inspirados en el caso de dimensión finita, multiplicamos estas ecuaciones por funciones λ_i, a determinar. Integrando en (x_0, x_1) obtenemos

$$\int_{x_0}^{x_1} \lambda_i \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial u_j} v_j dx = 0,$$

de modo que también se satisface

$$\int_{x_0}^{x_1} \sum_{j=1}^{n} \big(L_{u_j} - \frac{d}{dx} L_{u'_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial u_j}\big) v_j dx = 0. \tag{15}$$

Todavía no podemos aplicar el lema fundamental, puesto que las variaciones siguen siendo dependientes. Consideremos las m funciones $\lambda_1, \ldots, \lambda_m$ determinadas como la solución única del sistema de ecuaciones algebraicas lineales

$$L_{u_j} - \frac{d}{dx} L_{u'_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial u_j} = 0, \quad j = 1, \ldots, m,$$

[1] Es posible, gracias al teorema de la función implícita

que existe debido a que las restricciones son independientes, es decir,

$$\frac{\partial(\varphi_1,\ldots,\varphi_m)}{\partial(u_1,\ldots,u_m)} \neq 0.$$

Con $\lambda_1,\ldots,\lambda_m$ elegidas de este modo, la integral (15) queda como

$$\int_{x_0}^{x_1} \sum_{j=m+1}^{n} \big(L_{u_j} - \frac{d}{dx}L_{u'_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial u_j}\big) v_j dx = 0,$$

donde solo consideramos $n-m$ variaciones, que ya sí pueden tomarse arbitrariamente. Anulando sucesivamente todas las v_j excepto una y aplicando el lema fundamental, se deduce el resultado. □

Ejemplo 10. *Geodésicas.* Sea $\varphi(x,y,z)=0$ la ecuación de una superficie, S, dada y supongamos que toda curva diferenciable definida sobre S admite una parametrización del tipo

$$\alpha:[t_0,t_1]\to S, \quad \alpha(t)=(x(t),y(t),z(t)).$$

Entonces, la longitud de arco viene dada por

$$J(x,y,z)=\int_{t_0}^{t_1} \big(x'(t)^2+y'(t)^2+z'(t)^2\big)^{1/2} dt.$$

Las condiciones de extremo son

$$\frac{d}{dt}\Big(\frac{x'}{\big(x'^2+y'^2+z'^2\big)^{1/2}}\Big)+\lambda\varphi_x=0,$$

con ecuaciones análogas para y y z. Introduciendo el cambio de variable a longitud de arco, s, determinado por

$$\frac{ds}{dt}=\big(x'^2+y'^2+z'^2\big)^{1/2}$$

tenemos que $\frac{d}{dt}=\frac{ds}{dt}\frac{d}{ds}$ y

$$\frac{d}{dt}\Big(\frac{x'}{\big(x'^2+y'^2+z'^2\big)^{1/2}}\Big)=\frac{ds}{dt}\frac{d^2x}{ds^2}.$$

De este modo obtenemos la relación

$$\frac{d^2x/ds^2}{\varphi_x}=\frac{d^2y/ds^2}{\varphi_y}=\frac{d^2z/ds^2}{\varphi_z}=-\frac{\lambda}{ds/dt},$$

la cual expresa que la normal a la curva coincide con la normal a la superficie, que es la definición usual de geodésica en geometría diferencial. □

2.4.2. El problema isoperimétrico. En este problema, se trata de hallar la función $\mathbf{u}$ para la cual el funcional

$$J(\mathbf{u}) = \int_{x_0}^{x_1} L(x, \mathbf{u}, \mathbf{u}')dx$$

tiene un extremo, con las funciones admisibles satisfaciendo las condiciones de frontera $\mathbf{u}(x_0) = \mathbf{u}_0$ y $\mathbf{u}(x_1) = \mathbf{u}_1$, y tales que, para cierta $\mathbf{G} = (G_1, \ldots, G_m)$ se tiene

$$\int_{x_0}^{x_1} G_i(x, \mathbf{u}, \mathbf{u}')dx = \ell_i$$

para ciertas constantes fijadas, ℓ_i, $i = 1, \ldots, m$.

Los problemas isoperimétricos pueden reducirse a problemas con restricciones de igualdad como los vistos en la sección anterior. Para ello, introducimos las funciones

$$z_i(x) = \int_{x_0}^{x} G_i(x, \mathbf{u}, \mathbf{u}')dx,$$

para $i = 1, \ldots, m$, las cuales satisfacen $z_i(x_0) = 0$ y $z_i(x_1) = \ell_i$. Además, $z_i'(x) = G_i(x, \mathbf{u}, \mathbf{u}')$. De este modo, los enlaces isoperimétricos

$$\int_{x_0}^{x_1} G_i(x, \mathbf{u}, \mathbf{u}')dx = \ell_i$$

se transforman en los enlaces diferenciales

$$z_i'(x) = G_i(x, \mathbf{u}, \mathbf{u}').$$

A continuación, definimos el funcional

$$\tilde{J}(\mathbf{u}, \mathbf{z}) = \int_{x_0}^{x_1} L(x, \mathbf{u}, \mathbf{u}')dx,$$

con $\mathbf{z} = (z_1, \ldots, z_m)$. El extremo debe satisfacer las condiciones de frontera

$$(\mathbf{u}(x_0), \mathbf{z}(x_0)) = (\mathbf{u}_0, \mathbf{0}), \quad (\mathbf{u}(x_1), \mathbf{z}(x_1)) = (\mathbf{u}_1, \ell),$$

siendo $\ell = (\ell_1, \ldots, \ell_m)$, y las restricciones de igualdad

$$G_i(x, \mathbf{u}, \mathbf{u}') - z_i' = 0 \quad \text{para } i = 1, \ldots, m.$$

Aplicando lo visto en la sección anterior, consideramos el funcional

$$J^*(\mathbf{u}, \mathbf{z}) = \int_{x_0}^{x_1} \Big(L(x, \mathbf{u}, \mathbf{u}') + \sum_{i=1}^{m} \lambda_i(x)(G_i(x, \mathbf{u}, \mathbf{u}') - z_i') \Big) dx,$$

cuyas ecuaciones de Euler vienen dadas por

$$L_{u_j} + \sum_{i=1}^{m} \big(\lambda_i(x) \frac{\partial G_i}{\partial u_j}\big) - \frac{d}{dx}\Big(L_{u_j'} + \sum_{i=1}^{m} \lambda_i(x) \frac{\partial G_i}{\partial u_j'}\Big) = 0,$$

cuando derivamos respecto las incógnitas u_j, $j = 1, \ldots, n$ y por

$$\frac{d}{dx}\lambda_i(x) = 0,$$

cuando derivamos respecto las incógnitas z_i, $i = 1, \dots, m$. Obviamente, de aquí deducimos que λ_i son constantes. Además, las primeras n ecuaciones son las mismas que las ecuaciones de Euler asociadas al funcional

$$J^{**}(\mathbf{u}) = \int_{x_0}^{x_1} \Big(L(x, \mathbf{u}, \mathbf{u}') + \sum_{i=1}^{m} \lambda_i G_i(x, \mathbf{u}, \mathbf{u}') \Big) dx.$$

De este modo hemos llegado a la siguiente regla: para obtener la condición necesaria fundamental en el problema isoperimétrico sobre la determinación de un extremo del funcional

$$J(\mathbf{u}) = \int_{x_0}^{x_1} L(x, \mathbf{u}, \mathbf{u}') dx,$$

con $\mathbf{u}(x_0) = \mathbf{u}_0$ y $\mathbf{u}(x_1) = \mathbf{u}_1$, y sujeto a las condiciones isoperimétricas

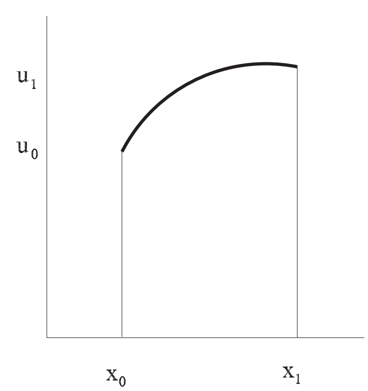
$$\int_{x_0}^{x_1} G_i(x, \mathbf{u}, \mathbf{u}') dx = \ell_i$$

hay que considerar el funcional auxiliar

$$J^{**}(\mathbf{u}) = \int_{x_0}^{x_1} \Big(L(x, \mathbf{u}, \mathbf{u}') + \sum_{i=1}^{m} \lambda_i G_i(x, \mathbf{u}, \mathbf{u}') \Big) dx,$$

donde λ_i son constantes a determinar, y escribir sus ecuaciones de Euler. Las constantes λ_i se determinan a partir de las condiciones isoperimétricas.

Ejemplo 11. Hallar la curva u de longitud ℓ dada para la cual el área del trapecio curvilíneo de la figura es máxima. El funcional a estudiar es



$$J(u) = \int_{x_0}^{x_1} u(x) dx,$$

con $u(x_0) = u_0$ y $u(x_1) = u_1$, y sujeto a la condición isoperimétrica

$$J(u) = \int_{x_0}^{x_1} \sqrt{1 + u'(x)^2} dx = \ell.$$

Las ecuación de Euler del funcional asociado

$$J^{**}(u) = \int_{x_0}^{x_1} \big(u(x) + \lambda\sqrt{1+u'(x)^2}\big)dx$$

viene dada por

$$u + \lambda\sqrt{1+u'^2} - \frac{\lambda u'^2}{\sqrt{1+u'^2}} = C_1,$$

donde hemos usado que el integrando no depende de x. Se sigue que

$$u - C_1 = -\frac{\lambda}{\sqrt{1+u'^2}}.$$

Introduciendo un parámetro t tal que $u' = \tan t$, de la ecuación anterior obtenemos

$$u = C_1 - \lambda \cos t.$$

Además, de $\frac{du}{dx} = \tan t$ se sigue que

$$dx = \frac{du}{\tan t} = \frac{\lambda \operatorname{sen} t dt}{\tan t} = \lambda \cos t dt,$$

de modo que $x = \lambda \operatorname{sen} t + C_2$. Despejando t de las expresiones para u y x obtenemos

$$(x - C_2)^2 + (u - C_1)^2 = \lambda^2.$$

Finalmente, las constantes C_1, C_2 y λ se determinan a partir de las condiciones de frontera y de la condición isoperimétrica. □

Ejemplo 12. *Problema de autovalores.* Hallar el mínimo del funcional

$$J(u) = \int_0^\pi u'(x)^2 dx,$$

con $u(0) = u(\pi) = 0$ y sujeto a

$$\int_0^\pi u(x)^2 dx = 1.$$

El funcional asociado es

$$J(u) = \int_0^\pi (u'(x)^2 + \lambda u(x)^2) dx,$$

cuya ecuación de Euler viene dada por

$$u'' = \lambda u,$$

es decir, es un problema de autovalores. Las raíces del polinomio característico son $\pm\sqrt{\lambda}$. Si $\lambda \geq 0$ entonces la solución general viene dada por

$$u(x) = C_1 \exp(\sqrt{\lambda}x) + C_2 \exp(-\sqrt{\lambda}x)$$

que no puede satisfacer las condiciones de frontera, de modo que no hay solución para $\lambda \geq 0$. En el caso contrario, $\lambda < 0$, se tiene que la solución general es

$$u(x) = C_1 \operatorname{sen}(\sqrt{-\lambda}x) + C_2 \cos(\sqrt{-\lambda}x).$$

Las condiciones de frontera implican que $u(x) = C_1 \sin kx$, con $k = \sqrt{-\lambda}$. La condición isoperimétrica implica

$$1 = \int_0^\pi u(x)^2 dx = C_1^2 \int_0^\pi \sin^2 kx dx = C_1^2 \int_0^{k\pi} \sin^2 s \frac{ds}{k} = \frac{C_1^2 \pi}{2},$$

de donde $C_1 = \pm\sqrt{2/\pi}$, y la solución queda como

$$u(x) = \pm \frac{2}{\pi} \operatorname{sen} kx,$$

es decir, hay una familia uniparamétrica de extremales que, de hecho, también satisfacen la condición de Legendre. □

3. Variación general de un funcional

3.1. Deducción de la fórmula básica. En esta sección deduciremos la fórmula general para la variación de un funcional de la forma

$$J(u) = \int_{x_0}^{x_1} L(x, u, u')dx.$$

Asumiremos que las curvas admisibles son regulares, digamos C^1, pero a diferencia de las hipótesis de las secciones previas, asumiremos que $u(x_0)$ y $u(x_1)$ pueden variar arbitrariamente. Este hecho nos motiva a introducir la siguiente noción de distancia: definimos

$$d(u, u^*) = \text{máx}\,|u - u^*| + \text{máx}\,|u' - u^{*\prime}| + |P_0 - P_0^*| + |P_1 - P_1^*|,$$

donde $P_0, P_0^* \in \mathbb{R}^2$ denotan los puntos correspondientes al extremo izquierdo del intervalo de definición de u y u^*, respectivamente, y P_1, P_1^* al extremo derecho de dicho intervalo. En general, las funciones u y u^* están definidas en intervalos diferentes I e I^*. Así, para que nuestra noción de distancia tenga sentido debemos extender u y u^* (de un modo diferenciable) a un intervalo que contenga a I e I^*.

Supongamos ahora que u y u^* son cercanas en el sentido de la distancia definida y consideremos su diferencia $v = u^* - u$. Sean $P_0 = (x_0, u_0)$ y $P_1 = (x_1, u_1)$, donde $u(x_0) = u_0$, $u(x_1) = u_1$, y

$$P_0^* = (x_0 + \delta x_0, u_0 + \delta u_0), \quad P_1^* = (x_1 + \delta x_1, u_1 + \delta u_1),$$

donde $u^*(x_0 + \delta x_0) = u_0 + \delta u_0$ y $u^*(x_1 + \delta x_1) = u_1 + \delta u_1$. La correspondiente variación, δJ, del funcional J, viene dada como una expresión lineal en términos de v, v', δx_0, δx_1, δu_0, δu_1, y que difiere del incremento

$$\Delta J = J(u + v) - J(u)$$

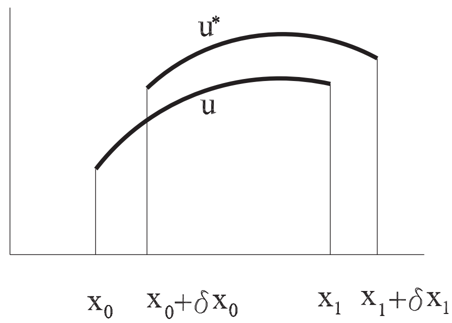


FIGURA 1

en una cantidad de orden menor que uno relativa a la distancia $d(u, u+v)$. Como

$$\begin{aligned}\Delta J &= \int_{x_0+\delta x_0}^{x_1+\delta x_1} L(x, u+v, u'+v')dx - \int_{x_0}^{x_1} L(x, u, u')dx \\ &= \int_{x_0}^{x_1} \big(L(x, u+v, u'+v') - L(x, u, u')\big)dx \\ &+ \int_{x_1}^{x_1+\delta x_1} L(x, u+v, u'+v')dx - \int_{x_0}^{x_0+\delta x_0} L(x, u+v, u'+v')dx,\end{aligned}$$

se sigue de la fórmula de Taylor que

$$\Delta J \sim \int_{x_0}^{x_1} \big(L_u(x, u, u')v - L_{u'}(x, u, u')v'\big)dx + L(x, u, u')\big|_{x=x_1}\delta x_1 - L(x, u, u')\big|_{x=x_0}\delta x_0$$

$$= \int_{x_0}^{x_1} \big(L_u - \frac{d}{dx}L_{u'}\big)vdx + L(x, u, u')\big|_{x=x_1}\delta x_1 + L_{u'}v\big|_{x=x_1} - L(x, u, u')|_{x=x_0}\delta x_0 - L_{u'}v\big|_{x=x_0},$$

donde $\sim$ denota igualdad excepto por términos de orden mayor que uno con respecto a $d(u, u+v)$. Sin embargo, queda claro en la Figura 1 que

$$v(x_0) \sim \delta u_0 - u'(x_0)\delta x_0, \quad v(x_0) \sim \delta u_1 - u'(x_1)\delta x_1,$$

y, por tanto

$$\begin{aligned}\delta_u J(v) = \int_{x_0}^{x_1} \big(L_u - \frac{d}{dx}L_{u'}\big)vdx + L_{u'}\big|_{x=x_1}\delta u_1 + (L - L_{u'}u')\big|_{x=x_1}\delta x_1 - L_{u'}\big|_{x=x_0}\delta u_0 \\ + (L - L_{u'}u')\big|_{x=x_0}\delta x_0,\end{aligned}$$

o, de un modo más conciso

$$\delta_u J(v) = \int_{x_0}^{x_1} \big(L_u - \frac{d}{dx}L_{u'}\big)vdx + L_{u'}\delta y\Big|_{x=x_0}^{x=x_1} + (L - L_{u'}u')\Big|_{x=x_0}^{x=x_1}, \tag{16}$$

donde hemos definido

$$\delta x|_{x=x_i} = \delta x_i, \quad \delta x|_{u=u_i} = \delta u_i.$$

Esta es la fórmula básica para la variación general del funcional J.

Observación 3. La fórmula de la variación general para el caso $n-$dimensional se deduce de un modo análogo, siendo ésta

$$\delta_{\mathbf{u}} J(\mathbf{v}) = \int_{x_0}^{x_1} \sum_{i=1}^{n} \big(L_{u_i} - \frac{d}{dx} L_{u_i'}\big) v_i dx + \sum_{i=1}^{n} L_{u_i'} \delta u_i \Big|_{x=x_0}^{x=x_1} + \big(L - \sum_{i=1}^{n} u_i' L_{u_i'}\big) \delta x \Big|_{x=x_0}^{x=x_1}.$$

□

Ejemplo 13. Estudiamos a continuación el caso en que las condiciones que se imponen en la frontera del intervalo son que las coordenadas estén contenidas en dos curvas dadas. Es decir, se trata de hallar, de entre todas las curvas regulares cuyos puntos frontera están contenidos en dos curvas, φ y ψ, aquellas que realizan un extremo del funcional

$$J(u) = \int_{x_0}^{x_1} L(x, u, u')dx.$$

La variación general de J viene dada por la fórmula (16). Claramente, cualquier extremal de J debe satisfacer la ecuación de Euler, con lo que (16) puede escribirse como

$$\delta_u J(v) = L_{u'}\big|_{x=x_1} \delta u_1 + (L - L_{u'} u')\big|_{x=x_1} \delta x_1 - L_{u'}\big|_{x=x_0} \delta u_0 + (L - L_{u'} u')\big|_{x=x_0} \delta x_0,$$

que debe anularse en todo extremo de J. De acuerdo a la Figura 2,

$$\delta u_0 = (\varphi'(x) + \varepsilon_0)\delta x_0, \quad \delta u_1 = (\psi'(x) + \varepsilon_1)\delta x_1,$$

con $\varepsilon_i \to 0$ cuando $\delta x_i \to 0$. Luego, en un extremo debe satisfacerse

$$0 = \delta_u J(v) = (L_{u'}\psi' + L - u' L_{u'})\big|_{x=x_1} \delta x_1 - (L_{u'}\varphi' + L - u' L_{u'})\big|_{x=x_0} \delta x_0.$$

Como los incrementos δx_0 y δx_1 son independientes, la anterior fórmula implica

$$(L + (\varphi' - u')L_{u'})\big|_{x=x_0} = 0,$$
$$(L + (\psi' - u')L_{u'})\big|_{x=x_1} = 0,$$

que son las llamadas **condiciones de transversalidad**, y que son condición necesaria de extremo para J.

En la resolución de problemas de optimización, a menudo encontramos funcionales del tipo

$$\int_{x_0}^{x_1} f(x, u)\sqrt{1 + u'^2} dx,$$

para los cuales las condiciones de transversalidad tienen una forma particularmente simple. En este caso tenemos

$$L_{u'} = f(x, u) \frac{u'}{\sqrt{1 + u'^2}} = \frac{u' L}{1 + u'^2},$$

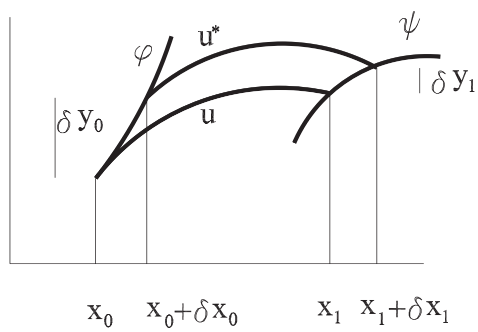


FIGURA 2

de modo que las condiciones de transversalidad son

$$L + (\varphi' - u')L_{u'} = \frac{(1 + u'\varphi')L}{1 + u'^2},$$
$$L + (\psi' - u')L_{u'} = \frac{(1 + u'\psi')L}{1 + u'^2}.$$

Se sigue que

$$u' = -1/\varphi' \quad \text{y } u' = -1/\psi'$$

en los extremos izquierdo y derecho del intervalo, respectivamente. Es decir, en este tipo de funcionales, las condiciones de transversalidad se reducen a condiciones de ortogonalidad. □

CAPÍTULO 3

Las condiciones de Jacobi

1. Introducción

En el capítulo anterior vimos que una condición necesaria para la realización de un mínimo de un funcional del tipo variacional

$$J(u)=\int_{x_0}^{x_1} L(x,u(x),u'(x))dx$$

con condición de frontera fija, $u(x_0)=u_0$ y $u(x_1)=u_1$, es la condición de Legendre

$$L_{u'u'}(x,u(x),u'(x))\geq 0 \quad \text{para todo } x\in(x_0,x_1).$$

Legendre, en analogía al caso de dimensión finita, intentó demostrar, sin éxito, que una condición suficiente para que J tenga un mínimo en u es que se satisfaga la desigualdad estricta

$$L_{u'u'}(x,u(x),u'(x))> 0 \quad \text{para todo } x\in(x_0,x_1). \tag{17}$$

En la Observación 2 hallamos la siguiente expresión para la segunda variación

$$\delta_u^2 J(v)=\frac{1}{2}\int_{x_0}^{x_1} \big((L_{uu}-\frac{d}{dx}L_{uu'})v^2+L_{u'u'}v'^{\,2}\big)dx,$$

con $v\in C_0^1([x_0,x_1])$. Por brevedad, la escribiremos como

$$\delta_u^2 J(v)=\frac{1}{2}\int_{x_0}^{x_1} \big(Pv'^{\,2}+Qv^2\big)dx.$$

La idea de Legendre fue escribir esta expresión en la forma

$$\delta_u^2 J(v)=\frac{1}{2}\int_{x_0}^{x_1} \big(Pv'^{\,2}+2wvv'+(Q+w')v^2\big)dx, \tag{18}$$

siendo w una función derivable arbitraria, y donde hemos usado la relación

$$0=\int_{x_0}^{x_1}\frac{d}{dx}(wv^2)=\int_{x_0}^{x_1}(w'v^2+2wvv')dx,$$

puesto que $v(x_0)=v(x_1)=0$. Seguidamente, observó que la condición $L_{u'u'}>0$ sería suficiente si se pudiera encontrar una función w para la que el integrando de (18) fuera un cuadrado perfecto. Sin embargo, esto no es siempre posible, como el mismo Legendre demostró, puesto que w debería satisfacer la ecuación

$$P(Q+w')=w^2,$$

que no posee, necesariamente, una solución global en todo el intervalo (x_0, x_1).[1] Legendre concluyó que condiciones de tipo *local* como la ecuación de Euler o la condición estricta de Legendre, dada por (17), no podían ser las únicas condiciones suficientes para la realización de un mínimo.

2. Condición necesaria de Jacobi

En esta sección estudiaremos las condiciones bajo las cuales el funcional

$$G(v) = \int_{x_0}^{x_1} \big(Pv'^{\,2} + Qv^2\big)dx, \tag{19}$$

definido para $v \in C_0^1([x_0, x_1])$, es definido positivo. Aunque en la sección anterior obtuvimos este funcional en relación con la función lagrangiana del problema de optimización, en concreto

$$P = \frac{1}{2}L_{u'u'}, \quad Q = \frac{1}{2}\big(L_{uu} - \frac{d}{dx}L_{u'u'}\big),$$

de momento estudiaremos la positividad del funcional (19) como un problema independiente.

Ya vimos que una condición necesaria para la no negatividad del funcional (19) es que

$$P(x) \geq 0 \quad \text{para todo } x \in (x_0, x_1).$$

En esta sección asumiremos la condición

$$P(x) > 0 \quad \text{para todo } x \in (x_0, x_1),$$

y buscaremos condiciones necesarias y suficientes para que dicho funcional sea definido positivo. Comenzamos escribiendo la ecuación de Euler asociada al funcional (19):

$$-\frac{d}{dx}(Pv') + Qv = 0. \tag{20}$$

Esta es una ecuación diferencial lineal de segundo orden. Las condiciones de frontera son $v(x_0) = v(x_1) = 0$. Este problema tiene la solución trivial $v \equiv 0$. Sin embargo, también puede tener soluciones no triviales[2]. En este contexto, introducimos la siguiente definición

Definición 3. *Diremos que el punto $\tilde{x}$ es un punto conjugado de x_0 si la ecuación (20) tiene una solución que se anula en x_0 y $\tilde{x}$ pero que no es idénticamente nula.*

Observación 4. Si v es una solución no idénticamente nula de (20) entonces también lo será Cv, para cualquier $C \neq 0$. Hay varios criterios de normalización que se pueden imponer para forzar la unicidad de solución de este problema. Aquí asumiremos que $v'(x_0) = 1$, que es siempre posible eligiendo C adecuadamente. En efecto, si $v(x_0) = 0$ y v no es idénticamente nula entonces $v'(x_0)$ no puede ser nulo, por el teorema de unicidad para la ecuación lineal (20). □

[1] Por ejemplo, si $P = -1$ y $Q = 1$, la solución viene dada por $w(x) = \tan(c - x)$. Si $x_1 - x_0 > \pi$, no hay solución en todo el intervalo (x_0, x_1), puesto que $\tan(c - x)$ se hace infinita en algún punto de dicho intervalo.

[2] Por ejemplo, para $P = Q = 1$, la función $v(x) = C \operatorname{sen} x$ es solución del problema en el intervalo $(0, \pi)$, para toda constante C.

Teorema 2. *Supongamos que $P(x) > 0$ en $[x_0, x_1]$ y que no hay puntos conjugados en dicho intervalo. Entonces el funcional cuadrático*

$$G(v) = \int_{x_0}^{x_1} \big(Pv'^{\,2} + Qv^2\big)dx$$

es definido positivo para toda $v \in C_0^1([x_0, x_1])$.

Demostración. Para demostrar que el funcional G es definido positivo lo reduciremos a la forma

$$\int_{x_0}^{x_1} P\varphi^2 dx,$$

donde φ^2 es cierta expresión cuya anulación implica que $v \equiv 0$. Para conseguir esto, comenzamos sumando a G la función cero expresada en la forma

$$\int_{x_0}^{x_1} \frac{d}{dx}(wv^2)dx.$$

A continuación, seleccionamos la función diferenciable w de modo que la expresión

$$Pv'^{\,2} + Qv^2 + \frac{d}{dx}(wv^2) = Pv'^{\,2} + 2wvv' + (Q + w')v^2 \tag{21}$$

sea un cuadrado perfecto. Esto será cierto si w es solución de

$$P(Q + w') = w^2, \tag{22}$$

ya que en tal caso (21) puede escribirse como

$$P\big(v' + \frac{w}{P}v\big)^2.$$

De este modo, si (22) tiene una solución definida en todo el intervalo $[x_0, x_1]$, entonces G puede escribirse como

$$G(v) = \int_{x_0}^{x_1} P\big(v' + \frac{w}{P}v\big)^2 dx, \tag{23}$$

y es, por tanto, definido positivo. De hecho, si (23) se anula, entonces debe ser

$$v' + \frac{w}{P}v \equiv 0,$$

puesto que, por hipótesis, es $P > 0$ en $[x_0, x_1]$. Pero esta ecuación de primer orden, con la condición $v(x_0) = 0$ tiene por única solución a $v \equiv 0$.

La demostración se reduce, pues, a comprobar que si no hay puntos conjugados de x_0 en el intervalo $[x_0, x_1]$ entonces la ecuación (22) tiene una solución global en $[x_0, x_1]$. Esta ecuación diferencial es una *ecuación de Ricatti*, , que puede ser reducida a una ecuación lineal de segundo orden mediante el cambio

$$w = -\frac{z'}{z}P, \tag{24}$$

donde z es una nueva incógnita. Con este cambio (22) se transforma en

$$-\frac{d}{dx}(Pz') + Qz = 0, \tag{25}$$

que es, justamente, la ecuación de Euler de G. Ahora, puesto que no hay puntos conjugados en $[x_0, x_1]$ se sigue, por definición, que (25) tiene una solución que no se anula en $[x_0, x_1]$. Por tanto, la ecuación (22) tiene una solución, dada por (24), definida en todo el intervalo $[x_0, x_1]$. □

A continuación veremos que la condición de que no existan puntos conjugados de x_0 en el intervalo $[x_0, x_1]$ no es solo una condición suficiente sino también necesaria. Comenzamos con un lema que usaremos en la demostración.

Lema 3. *Si la función v satisface la ecuación*

$$-\frac{d}{dx}(Pv') + Qv = 0$$

y las condiciones de frontera

$$v(x_0) = v(x_1) = 0,$$

entonces

$$\int_{x_0}^{x_1} \big(Pv'^{\,2} + Qv^2\big)dx = 0.$$

Demostración. El lema es una consecuencia inmediata de la fórmula

$$0 = \int_{x_0}^{x_1} \big(-\frac{d}{dx}(Pv') + Qv\big)v dx = \int_{x_0}^{x_1} \big(Pv'^{\,2} + Qv^2\big)dx,$$

que se obtiene por integración por partes y el uso de las condiciones de frontera. □

Teorema 3. *Supongamos que $P(x) > 0$ para todo $x \in [x_0, x_1]$. Si el funcional cuadrático G es definido positivo para toda $v \in C_0^1([x_0, x_1])$ entonces el intervalo $[x_0, x_1]$ no posee puntos conjugados de x_0.*

Demostración. La idea de la demostración es la de construir una familia de funcionales definidos positivos dependientes de un parámetro, t, tales que para $t = 1$ se reduce al funcional G, mientras que para $t = 0$ nos da el funcional cuadrático

$$\int_{x_0}^{x_1} v'^{\,2} dx,$$

el cual, claramente, no posee puntos conjugados de x_0 en $[x_0, x_1]$. Entonces, demostraremos que cuando variamos el parámetro t en $[0, 1]$, no pueden aparecer puntos conjugados en $[x_0, x_1]$.

Consideremos, pues, el funcional

$$\int_{x_0}^{x_1} \Big[t\big(Pv'^{\,2} + Qv^2\big) + (1-t)v'^{\,2}\Big]dx, \tag{26}$$

que es definido positivo para todo $t \in [0,1]$, puesto que G lo es por hipótesis. La ecuación de Euler correspondiente a este funcional es

$$-\frac{d}{dx}\left[(tP+(1-t))v'\right]+tQv=0. \tag{27}$$

Sea $v(x,t)$ una solución de (27) tal que $v(x_0,t)=0$ y $v_x(x_0,t)=1$ para todo $t \in [0,1]$. Esta solución depende con continuidad del parámetro t, que para $t=1$ se reduce a la solución v de la ecuación (20) con las condiciones $v(x_0,1)=v(x_1,1)=0$, y para $t=0$ se reduce a la solución de $v''=0$ con las mismas condiciones de frontera, es decir, $v(x,0)=x-x_0$.

Supongamos ahora que el intervalo $[x_0,x_1]$ contiene un punto $\tilde{x}$ conjugado de x_0. Necesariamente será $\tilde{x}<x_1$ porque, si $\tilde{x}=x_1$, el Lema 3 implica que existe una v no idénticamente nula tal que $G(v)=0$, contradiciendo la hipótesis de positividad de G. Por tanto, la demostración se reduce a comprobar que no puede haber un punto conjugado, $\tilde{x}$, en el interior de $[x_0,x_1]$.

Para demostrarlo, consideremos el conjunto de los puntos

$$C=\{(x,t)\in[x_0,x_1]\times[0,1]: v(x,t)=0\}.$$

Si mostramos que en los puntos en los que $v(x,t)=0$ no puede tenerse $v_x(x,t)=0$ entonces podremos deducir del teorema de la función implícita que el conjunto C representa una curva diferenciable en el plano xt. Tenemos que, si $v(x^*,t^*)=0$ en algún (x^*,t^*) entonces debe ser $v_x(x^*,t^*)\neq 0$ ya que para cualquier t fijo $v(x,t)$ satisface la ecuación (27), y si se tuviera $v(x^*,t^*)=v_x(x^*,t^*)=0$ entonces debería ser $v(x,t^*)=0$ para todo $x\in[x_0,x_1]$ debido al teorema de unicidad para ecuaciones diferenciales lineales. Pero esto es imposible puesto que v_x es una función continua en $[x_0,x_1]\times[0,1]$ y $v_x(x_0,t)=1$ para todo $t\in[0,1]$.

Tenemos entonces que el teorema de la función implícita nos asegura la existencia de una curva, $x(t)$, tal que $v(x(t),t)=0$ en un entorno de cada punto de C. Por hipótesis, el punto $(\tilde{x},1)$ pertenece a dicha curva. Partiendo de este punto tenemos que (véase Figura 1):

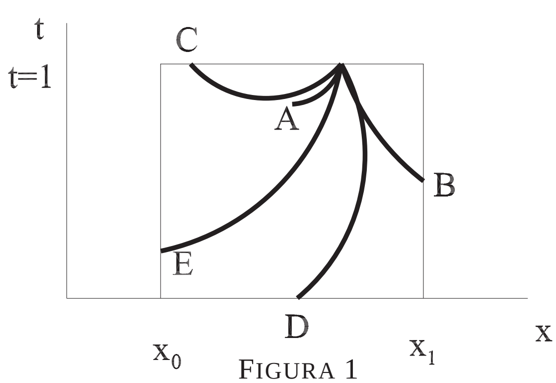


FIGURA 1

A. La curva no puede terminar en el interior de $[x_0,x_1]\times[0,1]$, pues sería una contradicción de la dependencia continua de $v(x,t)$ respecto del parámetro t.

B. La curva no puede cortar el segmento $\{x = x_1, t \in [0,1]\}$, puesto que, por el mismo argumento que el del Lema 3, pero aplicado a la ecuación (27), las condiciones de frontera $v(x_0, t) = v(x_1, t) = 0$ y el funcional (26), se tendría una contradicción de la positividad del funcional para todo t.
C. La curva no puede cortar el segmento $t = 1$, $x \in [x_0, x_1]$, puesto que tendríamos $v(x,t) = v_x(x,t) = 0$ para algún (x,t).
D. La curva no puede cortar el segmento $\{t = 0, x \in [x_0, x_1]\}$, puesto que para $t = 0$ la ecuación (27) se reduce a $v'' = 0$, cuya solución solo se anula en $x = x_0$.
E. La curva no puede aproximarse al segmento $\{x = x_0, t \in [0,1]\}$ puesto que tendríamos $v_x(x_0, t) = 0$ para algún t, contrario a nuestras hipótesis.

Se sigue que tal curva no puede existir, con lo que se concluye la demostración. □

Si reemplazamos la condición de que el funcional G sea definido positivo por la de que sea no negativo obtenemos el siguiente resultado.

Corolario 4. *Si $G(v)$, con $P(x) > 0$ para todo $x \in [x_0, x_1]$, es definido no negativo entonces el intervalo $[x_0, x_1)$ no contiene puntos conjugados de x_0.*

Demostración. La única diferencia con la demostración del teorema anterior es que no podemos asegurar que el funcional auxiliar dado por (26) sea definido positivo en $t = 1$. Así, no se puede excluir la posibilidad de que $\tilde{x} = x_1$. □

Los Teoremas 2 y 3 se combinan en el siguiente enunciado.

Teorema 5. *El funcional cuadrático*

$$\int_{x_0}^{x_1} \left(Pv'^{\,2} + Qv^2\right) dx,$$

con $P(x) > 0$ para todo $x \in [x_0, x_1]$, es definido positivo para toda $v \in C_0^1([x_0, x_1])$ si y solo si el intervalo $[x_0, x_1]$ no contiene puntos conjugados de x_0.

A continuación aplicaremos estos resultados al problema de optimización

$$J(u) = \int_{x_0}^{x_1} L(x, u(x), u'(x)) dx \tag{28}$$

con las condiciones de frontera $u(x_0) = u_0$ y $u(x_1) = u_1$. Recordemos que la segunda variación de este funcional en un extremo viene dada por

$$\int_{x_0}^{x_1} \left(Pv'^{\,2} + Qv^2\right) dx, \tag{29}$$

con

$$P = \frac{1}{2} L_{u'u'}, \quad Q = \frac{1}{2}\left(L_{uu} - \frac{d}{dx} L_{uu'}\right).$$

Definición 4. *La ecuación de Euler*

$$-\frac{d}{dx}(Pv') + Qv = 0$$

del funcional cuadrático (29) es llamada ***ecuación de Jacobi*** *del funcional original dado por (28).*

Definición 5. *Se dice que el punto $\tilde{x}$ es* ***conjugado*** *de x_0 con respecto al funcional (28) si es el conjugado de x_0 con respecto al funcional cuadrático (29) en el sentido de la Definición 3.*

A continuación enunciamos la **condición necesaria de Jacobi**:

Teorema 6. *Si u es un mínimo del funcional*

$$J(u) = \int_{x_0}^{x_1} L(x, u(x), u'(x))dx$$

para el cual $L_{u'u'} > 0$, entonces el intervalo (x_0, x_1) no contiene puntos conjugados de x_0.

Demostración. Ya vimos que una condición necesaria para la realización de un mínimo de J es la no negatividad de la segunda variación evaluada en dicho mínimo. El Corolario 4 asegura que si el funcional cuadrático (29) es no negativo entonces el intervalo (x_0, x_1) no puede contener puntos conjugados de x_0. □

3. Condición de Jacobi. Condiciones suficientes para un mínimo

En esta sección formularemos las condiciones suficientes bajo las cuales un funcional de la forma

$$J(u) = \int_{x_0}^{x_1} L(x, u(x), u'(x))dx \tag{30}$$

con condición de frontera fija, $u(x_0) = u_0$ y $u(x_1) = u_1$ tiene un mínimo en la curva u. Veremos que estas condiciones se parecen mucho a las condiciones necesarias obtenidas en las secciones anteriores. Las condiciones necesarias fueron consideradas separadamente, puesto que cada una de ellas es necesaria en sí misma. Sin embargo, las condiciones suficientes deben considerarse en conjunto puesto que la presencia de un mínimo está asegurada solo si se satisfacen todas las condiciones simultáneamente. Demostraremos previamente un lema que usaremos en la demostración del teorema de suficiencia.

Lema 4. *Sea u una función diferenciable con $u(x_0) = u_0$ y $u(x_1) = u_1$ y $v \in C_0^1([x_0, x_1])$. Entonces si L es tres veces diferenciable con continuidad respecto todos sus argumentos, se tiene*

$$J(u+v) - J(u) = \int_{x_0}^{x_1} (Pu'^2 + Qu)dx + \int_{x_0}^{x_1} (\xi v^2 + \eta v'^{\,2})dx,$$

donde $\xi, \eta \to 0$ cuando $\|v\|_1 \to 0$.

Demostración. El desarrollo de Taylor de orden dos nos proporciona la identidad

$$J(u+v)-J(u)=\int_{x_0}^{x_1}(L_u v+L_{u'}v')dx+\frac{1}{2}\int_{x_0}^{x_1}(L_{uu}v^2+2L_{uu'}vv'+L_{u'u'}v'^{\,2})dx+\varepsilon,$$

donde el resto, ε, puede escribirse como

$$\varepsilon=\int_{x_0}^{x_1}(\varepsilon_1 v^2+\varepsilon_2 vv'+\varepsilon_3 v'^{\,2})dx. \tag{31}$$

Debido a la continuidad de las derivadas L_{uu}, $L_{uu'}$ y $L_{u'u'}$, se sigue que $\varepsilon_i\to 0$ cuando $\|v\|_1\to 0$, para $i=1,2,3$. Ahora, integrando por partes y usando las condición de frontera de h, podemos escribir (31) como

$$\int_{x_0}^{x_1}(\xi v^2+\eta v'^{\,2})dx,$$

con ξ,η satisfaciendo $\xi,\eta\to 0$ cuando $\|v\|_1\to 0$. □

Teorema 7. *Supongamos que el funcional (30) evaluado en u satisface las siguientes condiciones:*

1. *La curva u es un extremo, es decir, satisface la ecuación de Euler*

$$L_u(x,u(x),u'(x))-\frac{d}{dx}L_{u'}(x,u(x),u'(x))=0 \quad \textit{para todo } x\in(x_0,x_1).$$

2. *A lo largo de la curva u se satisface la condición estricta de Legendre, es decir*

$$P(x)=\frac{1}{2}=L_{u'u'}(x,u(x),u'(x))>0 \quad \textit{para todo } x\in(x_0,x_1).$$

3. *El intervalo $[x_0,x_1]$ no contiene puntos conjugados de x_0.*

Entonces el funcional (30) tiene un mínimo en u.

Demostración. Si el intervalo $[x_0,x_1]$ no contiene puntos conjugados de x_0, y si $P(x)>0$ en dicho intervalo, entonces por la continuidad de la solución de la Ecuación de Jacobi y de la función P, tenemos que tampoco habrá puntos conjugados en un intervalo mayor $[x_0,x_1+\varepsilon]$, en el cual también podemos asumir que $P>0$. Consideremos ahora el funcional cuadrático

$$\int_{x_0}^{x_1}(Pv'^{\,2}+Qv)dx-\alpha^2\int_{x_0}^{x_1}v'^{\,2}dx, \tag{32}$$

que tiene por ecuación de Euler

$$-\frac{d}{dx}[(P-\alpha^2)v']+Qv=0. \tag{33}$$

Puesto que $P>0$ en $[x_0,x_1+\varepsilon]$ y, por tanto, tiene una cota inferior positiva en este intervalo, y como la solución de (33) que satisface las condiciones iniciales $v(x_0)=0$, $v'(x_0)=1$ depende con continuidad del parámetro α, se sigue que

1. $P(x)-\alpha^2>0$ para todo $x\in[x_0,x_1]$.
2. La solución de (33) que satisface las condiciones de frontera $v(x_0)=0$, $v'(x_0)=1$ no se anula en $(x_0,x_1]$.

El Teorema 2 implica entonces que el funcional (32) es definido positivo para todo α suficientemente pequeño. Es decir, existe una contante $c > 0$ tal que

$$\int_{x_0}^{x_1} (Pv'^{\,2} + Qv)dx > c \int_{x_0}^{x_1} v'^{\,2} dx. \tag{34}$$

A partir de (34) deducimos que el mínimo es, efectivamente, u. En efecto, sea h una curva tal que $u + h$ está suficientemente próximo a u. Por el Lema 4 tenemos que

$$J(u+h) - J(u) = \int_{x_0}^{x_1} (Pv'^{\,2} + Qv)dx + \int_{x_0}^{x_1} (\xi h^2 + \eta h'^2)dx,$$

donde ξ, η convergen uniformemente a cero en $[x_0, x_1]$ cuando $\|h\|_1 \to 0$. Además, usando la desigualdad de Schwarz obtenemos

$$h^2(x) = \Big(\int_{x_0}^{x} h' dx \Big)^2 \leq (x - x_0) \int_{x_0}^{x} h'^2 dx \leq (x - x_0) \int_{x_0}^{x_1} h'^2 dx,$$

es decir,

$$\int_{x_0}^{x_1} h^2 dx \leq \frac{(x_1 - x_0)^2}{2} \int_{x_0}^{x_1} h'^2 dx,$$

que implica que

$$\Big| \int_{x_0}^{x_1} (\xi h^2 + \eta h'^2) dx \Big| \leq \varepsilon \Big(1 + \frac{(x_1 - x_0)^2}{2} \Big) \int_{x_0}^{x_1} h'^2 dx, \tag{35}$$

si $|\xi| \leq \varepsilon$ y $|\eta| \leq \varepsilon$. Puesto que $\varepsilon > 0$ puede tomarse arbitrariamente pequeño, se sigue de (34) y (35) que $J(u + h) > J(u)$ para toda h con $\|h\|_1$ suficientemente pequeña. □

4. Relación entre la condición de Jacobi y la teoría de formas cuadráticas

De acuerdo al Teorema 5, el funcional cuadrático

$$\int_a^b (Pv'^{\,2} + Qv^2)dx, \tag{36}$$

donde hemos cambiado la notación $[x_0, x_1]$ por $[a, b]$, y donde $P(x) > 0$ para todo $x \in [a, b]$, es definido positivo para toda $v \in C_0^1([a, b])$ si y solo si el intervalo $[a, b]$ no contiene puntos conjugados de a. El funcional (36) es el análogo a una forma cuadrática en dimensión finita. Por tanto, es natural comenzar estudiando las condiciones de este tipo de formas en un espacio n-dimensional y luego tomar el límite $n \to \infty$. Esto puede hacerse de la siguiente manera: introduzcamos la partición

$$a = x_0, x_1, \ldots, x_n, x_{n+1} = b,$$

del intervalo $[a, b]$, en la cual, por comodidad, suponemos los nodos equiespaciados, $\Delta x = (b - a)/(n + 1)$. A continuación, consideremos la forma cuadrática

$$\sum_{i=0}^{n} \Big[P_i \Big(\frac{v_{i+1} - v_i}{\Delta x} \Big)^2 + Q_i v_i^2 \Big] \Delta x, \tag{37}$$

donde P_i, Q_i y v_i son los valores de las funciones P, Q y v en los nodos x_i. Esta forma cuadrática proporciona una aproximación finito dimensional del funcional cuadrático (36). Agrupando términos similares y teniendo en cuenta que $v_0 = v(a) = 0$, $v_{n+1} = v(b) = 0$, podemos escribir (37) como

$$\sum_{i=1}^{n} \Big[\Big(Q_i \Delta x + \frac{P_{i-1} + P_i}{\Delta x}\Big) v_i^2 - 2\frac{P_{i-1}}{\Delta x} v_{i-1} v_i\Big]. \tag{38}$$

En otras palabras, el funcional cuadrático (36) puede aproximarse por una forma cuadrática de n variables cuya matriz viene dada por

$$\begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 & 0 \\ b_1 & a_2 & b_2 & \dots & 0 & 0 & 0 \\ 0 & b_2 & a_3 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & 0 & b_{n-1} & a_n \end{pmatrix}, \tag{39}$$

donde

$$a_i = Q_i \Delta x + \frac{P_{i-1} + P_i}{\Delta x} \quad \text{para } i = 1, \dots, n,$$

y

$$b_i = -\frac{P_i}{\Delta x} \quad \text{para } i = 1, \dots, n-1.$$

Una matriz como (39) en la cual todos los elementos excepto la diagonal principal y sus dos diagonales adyacentes se anulan es llamada *matriz de Jacobi*, y su forma cuadrática asociada es llamada *forma de Jacobi*. Para cualquier matriz de Jacobi existe una fórmula de recurrencia para el cálculo de los menores principales dados por

$$D_i = \begin{vmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 & 0 \\ b_1 & a_2 & b_2 & \dots & 0 & 0 & 0 \\ 0 & b_2 & a_3 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{i-2} & a_{i-1} & b_{i-1} \\ 0 & 0 & 0 & \dots & 0 & b_{i-1} & a_i \end{vmatrix},$$

para $i = 1, \dots, n$. En efecto, expandiendo D_i con respecto a los elementos de la última fila, obtenemos la fórmula

$$D_i = a_i D_{i-1} - b_{i-1}^2 D_{i-2}, \tag{40}$$

que nos permite determinar $D_3, \dots, D_n$ en términos de los dos primeros menores. De hecho, si tomamos $D_0 = 1$ y $D_{-1} = 0$, entonces esta fórmula es válida para todo $i = 1, \dots, n$.

El criterio de Silvester asegura que una forma cuadrática simétrica es definida positiva si y solo si todos los menores D_i son positivos. Podemos así obtener un criterio para que el funcional cuadrático (36) sea definido positivo haciendo $n \to \infty$ en la fórmula (40) . Sustituyendo la expresión de los coeficientes a_i y b_i en dicha fórmula, obtenemos

$$D_i = \Big(Q_i \Delta x + \frac{P_{i-1} + P_i}{\Delta x}\Big) D_{i-1} - \frac{P_{i-1}^2}{(\Delta x)^2} D_{i-2}, \tag{41}$$

para $i = 1, \ldots, n$. Es obviamente imposible pasar directamente al límite $n \to \infty$ en esta expresión, puesto que los coeficientes de D_{i-1} y D_{i-2} se hacen infinitos. Para evitar esta dificultad realizamos el cambio

$$D_i = \frac{P_1 \ldots P_i Z_{i+1}}{(\Delta x)^{i+1}}, \quad D_0 = \frac{Z_1}{\Delta x} = 1, \quad D_{-1} = Z_0 = 0,$$

para $i = 1, \ldots, n$. La fórmula (41) se escribe entonces, en términos de las variables Z_i como

$$\frac{P_1 \ldots P_i Z_{i+1}}{(\Delta x)^{i+1}} = \Big(Q_i \Delta x + \frac{P_{i-1} + P_i}{\Delta x}\Big) \frac{P_1 \ldots P_{i-1} Z_i}{(\Delta x)^i} - \frac{P_{i-1}^2}{(\Delta x)^2} \frac{P_1 \ldots P_{i-2} Z_{i-1}}{(\Delta x)^{i-1}},$$

es decir

$$Q_i Z_i (\Delta x)^2 + P_{i-1} Z_i + P_i Z_i - P_i Z_{i+1} - P_{i-1} Z_{i-1} = 0, \tag{42}$$

o

$$Q_i Z_i - \frac{1}{\Delta x}\Big(P_i \frac{Z_{i+1} - Z_i}{\Delta x} - P_{i-1} \frac{Z_i - Z_{i-1}}{\Delta x}\Big) = 0,$$

para $i = 1, \ldots, n$. Pasando al límite $n \to \infty$ obtenemos la ecuación diferencial

$$-\frac{d}{dx}(PZ') + QZ = 0,$$

que es justamente la ecuación de Jacobi.

La condición de que las cantidades D_i sean positivas es equivalente a la condición de que las cantidades Z_i que satisfacen la ecuación en diferencias (42) sean positivas ya que el factor

$$\frac{P_1 \ldots P_i}{(\Delta x)^{i+1}}$$

es siempre positivo (ya que $P(x) > 0$). De modo que hemos probado que la forma cuadrática (36) es definida positiva si y solo si todas, excepto las primeras $n + 2$ cantidades $Z_0, \ldots, Z_{n+1}$ que satisfacen la ecuación en diferencias (42) son positivas.

Ahora, si consideramos la línea poligonal Π_n con vértices

$$(a_0, Z_0), (x_1, Z_1), \ldots, (b, Z_{n+1}),$$

la condición de que $Z_0 = 0$ y $Z_i > 0$ para $i = 1, \ldots, n + 1$ significa que Π_n no corta el intervalo $[a, b]$ excepto en el punto a. Así, cuando $\Delta x \to 0$, la ecuación en diferencias (42) se transforma en la ecuación de Jacobi, y la línea poligonal Π_n tiende a una solución no trivial de dicha ecuación , la cual satisface la condición inicial

$$Z(a) = Z_0 = 0, \quad Z'(a) = \lim_{x \to 0} \frac{Z_1 - Z_0}{\Delta x} = \lim_{x \to 0} \frac{\Delta x}{\Delta x} = 1,$$

y además, dicha solución no se anula en $(a, b]$. En otras palabras, cuando $n \to \infty$, la forma de Jacobi converge al funcional cuadrático (36), y la condición de que (38) sea definida positiva se traduce en la condición de que (36) sea definida positiva (Teorema 5), que es equivalente a que el intervalo $[a, b]$ no contenga puntos conjugados de a.

CAPÍTULO 4

Introducción a los métodos directos. El método de Ritz

Hasta aquí, el modo en que nos hemos aproximado a la resolución de los problemas variacionales ha sido mediante técnicas que reducen el problema a un problema formulado en términos de ecuaciones diferenciales. Hemos obtenido condiciones necesarias y suficientes para la demostración de la existencia de soluciones, y su posible cálculo mediante la resolución de la ecuación de Euler asociada. Sin embargo, la resolución efectiva de estos problemas dista de ser sencilla y es por ello que la investigación ha derivado a otros métodos, los llamados métodos directos, que permiten la obtención de soluciones aproximadas al problema original.

1. Sucesiones minimizantes

Existen muchas técnicas diferentes agrupadas bajo el nombre de *métodos directos*. Sin embargo, todas ellas contienen una idea común, que es la siguiente.

Consideremos el problema de hallar el mínimo de un funcional, $J(u)$, definido en un espacio $\mathcal{M}$ de funciones admisibles u. Para que el problema tenga sentido, debemos asumir que existen funciones en $\mathcal{M}$ tales que $J(u) < \infty$, y además que

$$\inf_u J(u) = \mu > -\infty, \tag{43}$$

donde el ínfimo se toma sobre todas las funciones admisibles, u. Entonces, por definición de μ, existe una sucesión de funciones $\{u_n\}$, llamada una **sucesión minimizante**, tal que

$$\lim_{n\to\infty} J(u_n) = \mu.$$

Si la sucesión $\{u_n\}$ tiene una función límite, $\hat{u}$, y si es legítimo escribir

$$J(\hat{u}) = \lim_{n\to\infty} J(u_n), \tag{44}$$

es decir,

$$J(\lim_{n\to\infty} u_n) = \lim_{n\to\infty} J(u_n),$$

entonces

$$J(\hat{u}) = \mu,$$

de modo que $\hat{u}$ es solución del problema variacional. Además, las funciones de la sucesión minimizante $\{u_n\}$ pueden tomarse como soluciones aproximadas de nuestro problema. Así, para resolver un problema variacional dado mediante un método directo, debemos

1. Construir una sucesión minimizante $\{u_n\}$.

2. Demostrar que $\{u_n\}$ tiene un límite, $\hat{u}$.
3. Demostrar que se puede tomar el límite (44).

Observación 5. Aunque exista una sucesión minimizante $\{u_n\}$ para un problema variacional dado, no tiene por qué existir el límite. Por ejemplo, consideremos el funcional

$$J(u) = \int_{-1}^{1} x^2 u'^2 dx,$$

donde

$$u(-1) = -1, \quad u(1) = 1.$$

Obviamente, $J(u)$ sólo toma valores positivos, y

$$\inf_u J(u) = 0.$$

Podemos elegir

$$u_n(x) = \frac{\arctan nx}{\arctan n} \tag{45}$$

como la sucesión minimizante, puesto que

$$\int_{-1}^{1} \frac{n^2 x^2 dx}{(\arctan n)^2 (1 + n^2 x^2)^2} < \frac{1}{(\arctan n)^2} \int_{-1}^{1} \frac{dx}{1 + n^2 x^2} = \frac{2}{n \arctan n},$$

y, por tanto, $J(u_n) \to 0$ cuando $n \to \infty$. Sin embargo, cuando $n \to \infty$, la sucesión (45) no tiene límite en la clase de funciones continuas que satisfacen las condiciones de frontera. □

Incluso si la sucesión minimizante tiene un límite en el sentido de las funciones continuas, no es una tarea trivial el justificar el paso al límite (44), puesto que, en general, los funcionales considerados en el cálculo variacional no son continuos en dicha norma. Sin embargo, (44) puede justificarse si la continuidad de $J(u)$ se reemplaza por una condición más débil. En este contexto introducimos la siguiente definición.

Definición 6. *Diremos que el funcional $J(u)$ es* ***semicontinuo inferiormente*** *en $\hat{u} \in \mathcal{M}$ si, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que*

$$J(u) - J(\hat{u}) > -\varepsilon,$$

siempre que $\|u - \hat{u}\| < \delta$.

Tenemos el siguiente resultado

Teorema 8. *Si $\{u_n\}$ es una sucesión minimizante del funcional $J(u)$, con función límite $\hat{u}$, y si $J(u)$ es semicontinuo inferiormente en $\hat{u}$, entonces*

$$J(\hat{u}) = \lim_{n\to\infty} J(u_n).$$

Demostración. Por una parte,

$$J(\hat{u}) \geq \lim_{n\to\infty} J(u_n) = \inf J(u), \tag{46}$$

mientras que, por otra parte, dado cualquier $\varepsilon > 0$,

$$J(u_n) - J(\hat{u}) > -\varepsilon,$$

si n es suficientemente grande. Tomando $n \to \infty$, obtenemos

$$J(\hat{u}) \leq \lim_{n\to\infty} J(u_n) + \varepsilon,$$

o, lo que es lo mismo,.

$$J(\hat{u}) \leq \lim_{n\to\infty} J(u_n), \tag{47}$$

puesto que ε es arbitrario. De (46) y (47), obtenemos la conclusión del teorema. □

2. El método de Ritz

El método de Ritz es uno de los métodos directos más usados en el cálculo de variaciones. Supongamos que estamos buscando el mínimo de un funcional $J(u)$ definido en un espacio de funciones admisibles, $\mathcal{M}$, que por simplicidad asumiremos que es un espacio vectorial normado. Sean

$$\varphi_1, \varphi_2, \ldots \tag{48}$$

una sucesión infinita de funciones de $\mathcal{M}$, y sea $\mathcal{M}_n$ el subespacio lineal $n-$dimensional de $\mathcal{M}$ generado por las primeras n funciones de (48), es decir, el conjunto de todas las combinaciones lineales de la forma

$$\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_n\varphi_n, \tag{49}$$

donde $\alpha_1, \ldots, \alpha_n$ son números reales cualesquiera. Entonces, en cada subespacio $\mathcal{M}_n$, el funcional $J(u)$ da lugar a una función

$$J(\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_n\varphi_n) \tag{50}$$

de n variables.

A continuación, elegimos $\alpha_1, \ldots, \alpha_n$ de tal manera que se minimice (50), y denotamos por μ_n el valor mínimo del funcional y por u_n el mínimo del mismo. Claramente, μ_n no puede crecer con n, es decir

$$\mu_1 \geq \mu_2 \geq \ldots,$$

puesto que cualquier combinación lineal de $\varphi_1, \ldots, \varphi_n$ es automáticamente una combinación lineal de $\varphi_1, \ldots, \varphi_{n+1}$. También, cada subespacio de la sucesión

$$\mathcal{M}_1, \mathcal{M}_2, \ldots$$

está contenido en el siguiente. A continuación obtendremos condiciones que garantizan que la sucesión u_n es una sucesión minimizante.

Definición 7. *La sucesión (48) es* ***completa*** *en $\mathcal{M}$ si dado cualquier $u \in \mathcal{M}$ y cualquier $\varepsilon > 0$, existe una combinación lineal η_n de la forma (49) tal que $\|\eta_n - u\| < \varepsilon$ (donde n depende de ε).*

Teorema 9. *Si el funcional $J(u)$ es continuo (en la norma de $\mathcal{M}$) y si la sucesión (48) es completa entonces*

$$\lim_{n\to\infty} \mu_n = \mu,$$

donde

$$\mu = \inf_u J(u).$$

Demostración. Dado $\varepsilon > 0$, sea u^* tal que

$$J(u^*) < \mu + \varepsilon.$$

Tal u^* existe para todo $\varepsilon > 0$, por definición de μ. Como $J(u)$ es continuo,

$$|J(u) - J(u^*)| < \varepsilon, \tag{51}$$

siempre que $\|u - u^*\| < \delta = \delta(\varepsilon)$. Sea η_n una combinación lineal de la forma (49) tal que $\|\eta_n - u^*\| < \delta$, que existe por ser $\{\varphi_n\}$ completa, y sea u_n una combinación lineal de la misma forma pero para la cual se alcanza el mínimo de (50). Entonces, tenemos que

$$\mu \leq J(u_n) \leq J(\eta_n) < \varepsilon + J(u^*) < \mu + 2\varepsilon.$$

Puesto que ε es arbitrario, se sigue que

$$\lim_{n\to\infty} J(u_n) = \lim_{n\to\infty} \mu_n = \mu.$$

$\square$

Bibliografía

[1] L. ELSGOLTZ, *Ecuaciones diferenciales y cálculo variacional*, Editorial Mir-Rubiños, Madrid, 1969.

[2] C. FOX, *An introduction to the calculus of variations*, Dover Publications Inc., New York, 1987.

[3] I.M. GELFAND and S. V. FOMIN, *Calculus of variations*, Dover Publications Inc., New York, 2000.