

Introducción al cálculo de variaciones

Jesús Oroya

10 de septiembre de 2020

Índice general

Capítulo 1

Introducción al cálculo de variaciones y el control óptimo

En este documento introduciré brevemente el cálculo de variaciones. Este es un conocimiento necesario antes de poder entender la formulación de control óptimo.

El cálculo de variaciones es el análogo a un problema de optimización con variable en $y \in \mathbb{R}^N$. Como bien sabréis en un problema de optimización deberemos definir claramente la dimensión de la variable de optimización, podría ser un variable escalar, un vector de \mathbb{R}^2 o un vector de dimensión más grande. Por otra parte, el cálculo de variaciones aborda un problema de optimización donde la variable y tiene dimensión infinita, es decir es una función en sí.

1.1. La geodésica, el camino más corto entre dos puntos

El problema típico para introducir el cálculo de variaciones es la búsqueda del camino más corto entre dos puntos, la geodésica¹. En general, los libros suelen introducir este problema pasándolo directamente al problema con variables como funciones pero dado que la propuesta es pasar de un problema de optimización en variable discreta a una problema de cálculo variacional empezaremos definiendo el problema de optimización con variable $y \in \mathbb{R}^N$

1.1.1. Problema de optimización convencional

Supongamos que tenemos una secuencia:

$$\{(x_i, y_i)\}_{i=0}^{n+1} = \{(x_0, y_0), (x_1, y_1), \dots, (x_{n+1}, y_{n+1})\} \quad (1.1)$$

De esta secuencia de puntos conocemos todos los valores de las coordenadas $\{x_i\}_{i=0}^{n+1} = \{x_0, x_1, \dots, x_{n+1}\}$ además de los valores de las coordenadas y_0 e y_{n+1} .

Entonces, nos podemos preguntar cuál debería ser los valores de las coordenadas $\{y_i\}_{i=1}^n = \{y_1, \dots, y_n\}$ para que la distancia que define la secuencia $\{x_i, y_i\}_{i=0}^{n+1}$ sea mínima. En la figura (??) se muestra el problema que proponemos.

En un lenguaje más matemático podríamos escribir el problema como la suma de las distancias entre dos puntos consecutivos. :

$$\min_{\mathbf{y} \in \mathbb{R}^n} \left[\sum_{i=1}^{n+1} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \right] \quad (1.2)$$

Donde $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$. Con fines pedagógicos cambiaremos un poco esta expresión para que se pueda comparar fácilmente al problema de cálculo variacional.

Entonces extraemos $(x_i - x_{i-1})$:

¹Obviamente el camino más corto que une a dos puntos es la recta, sin embargo por razones pedagógicas utilizaremos este ejemplo

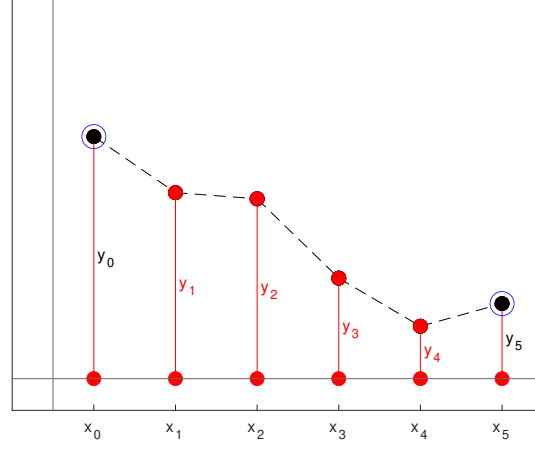


Figura 1.1: Problema de optimización discreto

La variable de optimización en este caso es un vector $\mathbf{y} = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$

$$\min_{\mathbf{y} \in \mathbb{R}^n} \left[\sum_{i=1}^{n+1} (x_i - x_{i-1}) \sqrt{1 + \left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}} \right)^2} \right] \quad (1.3)$$

Entonces si definimos $\Delta x_i = x_i - x_{i-1}$ y $\Delta y_i = y_i - y_{i-1} \mid \forall i \in \{1, \dots, n\}$. Podemos escribir el problema como:

$$\min_{\mathbf{y} \in \mathbb{R}^n} \left[\sum_{i=1}^{n+1} \Delta x_i \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i} \right)^2} \right] \quad (1.4)$$

Para resolver el problema planteado simplemente utilizaremos la condición del gradiente igual a cero para buscar puntos críticos².

Llamaremos $d(\mathbf{y})$ a la función que estamos minimizando, entonces:

$$d(\mathbf{y}) = \sum_{i=1}^{n+1} \Delta x_i \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i} \right)^2} \quad (1.5)$$

Calculamos la derivada con respecto a y_i de la función $d(\mathbf{y})$

$$\frac{\partial d}{\partial y_i} = \frac{1}{(1 + (\Delta x_i / \Delta y_i)^2)^{1/2}} - \frac{1}{(1 + (\Delta x_{i-1} / \Delta y_{i-1})^2)^{1/2}} \quad (1.6)$$

Utilizando la condición necesaria de mínimo $\frac{\partial d}{\partial y_i} = 0 \mid \forall i \in \{1, \dots, N\}$, podemos calcular la solución al problema.

$$\frac{\partial d}{\partial y_i} = \frac{1}{(1 + (\Delta x_i / \Delta y_i)^2)^{1/2}} - \frac{1}{(1 + (\Delta x_{i-1} / \Delta y_{i-1})^2)^{1/2}} = 0 \quad (1.7)$$

Entonces:

$$(\Delta y_i / \Delta x_i) = (\Delta y_{i-1} / \Delta x_{i-1}) \quad (1.8)$$

Por lo tanto la solución al problema es la secuencia que mantenga $(\Delta y_i / \Delta x_i)$ constante. Como notareis este valor es la tangente del ángulo formado entre el eje de las abscisa y el vector definido como la unión de dos puntos consecutivos. Es decir, la solución es la secuencia que este alineada con los puntos x_0 y x_{n+1}

²En este problema no existe más puntos críticos que el mínimo por lo que encontrar un punto crítico es igual a encontrar el mínimo y con esto la solución al problema

1.1.2. Problema de cálculo de variaciones

En el problema anterior la variable de optimización vivía en el espacio vectorial \mathbb{R}^n . En este caso nos podemos preguntar qué curva que pase por los puntos (x_0, y_0) y (x_{n+1}, y_{n+1}) es la que tiene menor longitud. La variable en este problema tiene un grado de libertad infinito, es decir en lugar de ser un vector de \mathbb{R}^n es una función $y(x)$ definida en el intervalo $x \in (x_0, x_f)$. Podemos dibujar cualquier función que queramos y deberemos hallar la que minimice la distancia (figura ??).

Otra pieza fundamental en el cálculo de variaciones es el concepto de funcional. Este es una versión análoga al de función. En el caso de funciones, estas son objetos matemáticas que reciben como entrada un vector $\mathbf{y} \in \mathbb{R}^n$ y devuelven un escalar de $d \in \mathbb{R}$. En el caso de los funcionales estos reciben como entrada una función definida en un intervalo y son capaces de devolver un escalar en \mathbb{R}^n .

Para hallar la longitud de una curva $y(x)$ que empiece en (x_0, y_0) y termine en (x_f, y_f) tomaremos la expresión (??) tomaremos el límite de cuando $\Delta x_i \rightarrow 0$

$$D[y(x)] = \lim_{\Delta x \rightarrow 0} d(\mathbf{y}) = \lim_{\Delta x \rightarrow 0} \sum_{i=1}^{n+1} \Delta x_i \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i} \right)^2} \quad (1.9)$$

Utilizando la definición de integral como suma de Riemman:

$$\lim_{\Delta x \rightarrow 0} \sum_{i=1}^{n+1} \Delta x_i \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i} \right)^2} = \int_{x_0}^{x_f} \sqrt{1 + y'(x)^2} dx \quad (1.10)$$

Entonces el funcional que nos da la longitud de curva es:

$$D[y(x)] = \int_{x_0}^{x_f} \sqrt{1 + y'(x)^2} dx \quad (1.11)$$

Donde $y'(x) = \frac{dy}{dx}$.

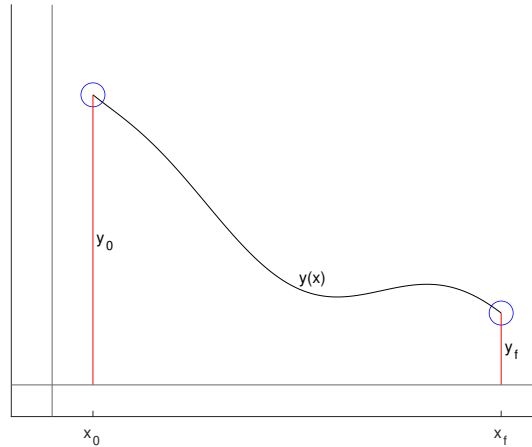


Figura 1.2: Problema de optimización continuo

En calculo de variaciones se le suele llamar lagrangiano $\mathcal{L}(x, y(x), y'(x))$ a la expresión que se encuentra en el interior de la integral. En este caso el lagrangiano sería:

$$\mathcal{L}(x, y(x), y'(x)) = \sqrt{1 + y'(x)^2} \quad (1.12)$$

Al igual que la condición necesaria de mínimo en problemas de optimización convencionales

$$\frac{\partial d}{\partial y_i} = 0 \quad \forall i \in \{1, \dots, n\} \quad (1.13)$$

En cálculo de variaciones existe una condición necesaria para un mínimo, es es llamada la ecuación de Euler-Lagrange:

$$\frac{\partial \mathcal{L}}{\partial y(x)} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial y'(x)} \right) = 0 \quad (1.14)$$

En nuestro caso esta se puede calcular los dos término de la ecuación de Euler-Lagrange fácilmente:

$$\frac{\partial \mathcal{L}}{\partial y(x)} = \frac{\partial}{\partial y(x)} \left(\sqrt{1 + y'(x)^2} \right) = 0 \quad (1.15)$$

$$\frac{\partial \mathcal{L}}{\partial y'(x)} = \frac{\partial}{\partial y'(x)} \left(\sqrt{1 + y'(x)^2} \right) = \frac{y'(x)}{\sqrt{1 + y'(x)^2}} \quad (1.16)$$

Remplazando estos dos términos en la ecuación (??):

$$\frac{d}{dx} \left(\frac{y'(x)}{\sqrt{1 + y'(x)^2}} \right) = 0 \quad (1.17)$$

Es decir:

$$\left(\frac{y'(x)}{\sqrt{1 + y'(x)^2}} \right) = c \quad (1.18)$$

Donde c es una constante. Despejando $y'(x)$ obtenemos:

$$y'(x) = \frac{c^2}{1 - c^2} = m \rightarrow y(x) = mx + b \quad (1.19)$$

Con $m = \frac{c^2}{1 - c^2}$ y b constante. Entonces la solución al problema es una recta $y(x) = mx + b$ que pasa por los puntos (x_0, y_0) y (x_f, y_f) .

1.1.3. Comparación en optimización convencional y cálculo de variaciones

En la tabla (??) se muestra analogías entre el cálculo de variaciones y la optimización convencional.

Objeto	Optimización	Cálculo de Variaciones
Objeto a optimizar	función $d(\mathbf{y})$	funcional $D[y(x)]$
Variable de optimización	vector $\mathbf{y} = \{y_1, y_2, \dots, y_n\} \in \mathbb{R}^n$	función $y(x) \mid x \in (x_0, x_f)$
Condición de optimalidad	$\frac{\partial d}{\partial y_i} = 0$	$\frac{\partial \mathcal{L}}{\partial y(x)} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial y'(x)} \right) = 0$

Cuadro 1.1: Comparación entre la optimización convencional y el cálculo de variaciones

1.2. Control Óptimo

Llegado a este punto podemos definir el control óptimo como un problema de cálculo variacional donde además tenemos una restricción que deben cumplir las variables en cuestión. De manera general un problema de control óptimo se puede escribir como:

$$\min_{u \in \mathcal{U}} \left[\Psi(y(T)) + \int_0^T L(t, y(t), u(t)) \right] \quad (1.20)$$

$$\text{sujeto a: } \begin{cases} \dot{y} = f(t, y(t), u(t)) \\ y(0) = y_0 \end{cases} \quad (1.21)$$

1.2.1. Condiciones de Optimilidad - Principio del mínimo de Pontryagin

Definimos el Hamiltoniano H como:

$$H(y(t), u(t), p(t), t) = p^T(t)f(y(t), u(t)) + L(y(t), u(t)) \quad (1.22)$$

$$\begin{cases} H(y^*(t), u^*(t), p^*(t), t) \leq H(y^*(t), u(t), p^*(t), t) \\ \Psi_T(y(T)) + H(T) = 0 \\ \dot{p}^T(t) = -H_y(y^*(t), u^*(t), p(t), t) \\ p^T(T) = \Psi_y(y(T)) \end{cases} \quad (1.23)$$

Capítulo 2

Problema SHE como un problema de control óptimo

Abordaremos el problema de *SHE* de simetría de media onda y de dos niveles. Suponiendo la simetría de media onda de una onda $f(\omega t)$ se puede descomponer en serie de Fourier:

$$f(\omega t) = \sum_{n=1}^{\infty} \left[a_n \sin((2n-1)\omega t) + b_n \cos((2n-1)\omega t) \right] \quad (2.1)$$

donde:

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(\omega t) \sin((2m-1)\omega t) d(\omega t) \quad (2.2)$$

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(\omega t) \cos((2m-1)\omega t) d(\omega t) \quad (2.3)$$

Para simplificar harémos el cambio de variable $\tau = \omega t$. Entonces las expresiones anteriores se pueden escribir de la siguiente forma:

$$f(\tau) = \sum_{n=1}^{\infty} \left[a_n \sin((2n-1)\tau) + b_n \cos((2n-1)\tau) \right] \quad (2.4)$$

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(\tau) \sin((2m-1)\tau) d\tau \quad (2.5)$$

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(\tau) \cos((2m-1)\tau) d\tau \quad (2.6)$$

Deberemos notar que las expresiones (??) y (??) pueden ser escritas en forma de ecuación diferencial:

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(\tau) \sin((2m-1)\tau) d\tau \Rightarrow \frac{d\alpha_n}{d\tau} = \frac{2}{\pi} f(\tau) \sin((2m-1)\tau) \mid \alpha_n(0) = 0 \quad (2.7)$$

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(\tau) \cos((2m-1)\tau) d\tau \Rightarrow \frac{d\beta_n}{d\tau} = \frac{2}{\pi} f(\tau) \cos((2m-1)\tau) \mid \beta_n(0) = 0 \quad (2.8)$$

Entonces dado una función $f(\tau)$ podemos obtener los coeficiente de Fourier a_n y b_n resolviendo el siguiente sistema de ecuaciones diferenciales hasta tiempo $\tau = \pi$.

$$\frac{d\alpha_n}{d\tau} = \frac{2}{\pi} f(\tau) \sin((2m-1)\tau) \mid \alpha_n(0) = 0 \quad (2.9)$$

$$\frac{d\beta_n}{d\tau} = \frac{2}{\pi} f(\tau) \cos((2m-1)\tau) \mid \beta_n(0) = 0 \quad (2.10)$$

2.1. Problema de control óptimo para SHE

De esta forma si tenemos un vector de objetivo $\mathbf{a}_T = [a_1^T, a_2^T, \dots, a_M^T] \in \mathbb{R}^M$ y $\mathbf{b}_T = [b_1^T, b_2^T, \dots, b_M^T] \in \mathbb{R}^M$, donde M es el número de armónicos impares de interés, podemos plantear el siguiente problema de control óptimo:

$$\min_{f(\tau) \in \mathcal{F}} [||\boldsymbol{\alpha}(T) - \mathbf{a}_T||^2 + ||\boldsymbol{\beta}(T) - \mathbf{b}_T||^2] \quad (2.11)$$

$$\text{sujeto a: } \begin{cases} \frac{d\alpha_m}{d\tau} = \frac{2}{\pi} f(\tau) \sin[(2m-1)\tau] \\ \alpha_m(0) = 0 \end{cases} \quad (2.12)$$

$$\text{sujeto a: } \begin{cases} \frac{d\beta_m}{d\tau} = \frac{2}{\pi} f(\tau) \cos[(2m-1)\tau] \\ \beta_m(0) = 0 \end{cases} \quad (2.13)$$

$$\forall n \in \{1, \dots, M\} \quad (2.14)$$

$$\text{sujeto a: } f_{\min} < f(\tau) < f_{\max} \mid \forall \tau \in (0, \pi) \quad (2.15)$$

Con tiempo final $T = \pi$

2.2. Condiciones de optimalidad para el problema de control óptimo para SHE

Siguiendo el principio de mínimo de Pontryagin construiremos el Hamiltoniano:

$$H = \sum_{m=1}^M \left[p_{\alpha}^m \left(\frac{2}{\pi} f(\tau) \sin((2m-1)\tau) \right) + p_{\beta}^m \left(\frac{2}{\pi} f(\tau) \cos((2m-1)\tau) \right) \right] \quad (2.16)$$

$$H = \frac{2}{\pi} f(\tau) \sum_{m=1}^M \left[p_{\alpha}^m \sin[(2m-1)\tau] + p_{\beta}^m \cos[(2m-1)\tau] \right] \quad (2.17)$$

La ecuación de adjunto se puede calcular de la siguiente manera:

$$\frac{dp_{\alpha}^m}{d\tau} = \frac{\partial H}{\partial \alpha_m} = 0 \quad \text{y} \quad \frac{dp_{\beta}^m}{d\tau} = \frac{\partial H}{\partial \beta_m} = 0 \quad (2.18)$$

Es decir la componente p_{α}^n y p_{β}^n son constantes.

Si llamamos $\Psi(\boldsymbol{\alpha}, \boldsymbol{\beta})$ a:

$$\Psi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = [||\boldsymbol{\alpha}(T) - \mathbf{a}_T||^2 + ||\boldsymbol{\beta}(T) - \mathbf{b}_T||^2] \quad (2.19)$$

Utilizando la condición final del adjunto:

$$p_{\alpha}^m(T) = \frac{d\Psi}{d\alpha_m} = 2(\alpha_m(T) - \mathbf{a}_n^T) \quad (2.20)$$

$$p_{\beta}^m(T) = \frac{d\Psi}{d\beta_m} = 2(\beta_m(T) - \mathbf{b}_n^T) \quad (2.21)$$

Si llamamos $G(\tau, \mathbf{p}_{\alpha}, \mathbf{p}_{\beta})$ a:

$$G(\tau, \mathbf{p}_{\alpha}, \mathbf{p}_{\beta}) = \frac{2}{\pi} \sum_{m=1}^M \left[p_{\alpha}^m \sin[(2m-1)\tau] + p_{\beta}^m \cos[(2m-1)\tau] \right] \quad (2.22)$$

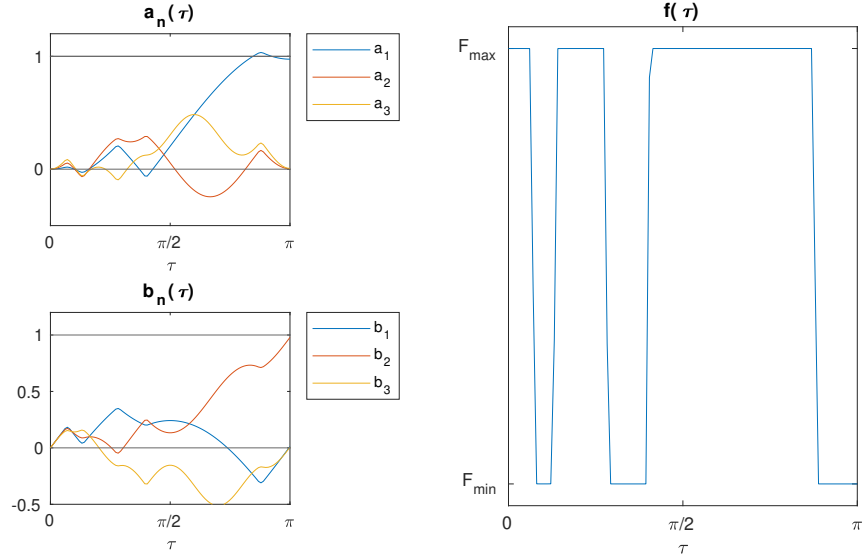


Figura 2.1: Solución numérica del problema de control óptimo. $a_T = [1 \ 0 \ 0]^T$ y $b_T = [0 \ 1 \ 0]^T$

Entonces el Hamiltoniano se puede escribir como:

$$H = G(\tau, \mathbf{p}_\alpha, \mathbf{p}_\beta) f(\tau) \quad (2.23)$$

Aplicando el principio del mínimo de Pontryagin podemos hallar la forma de f^*

$$\min_{f \in \mathcal{F}} H \rightarrow f^* = \begin{cases} f_{\min} & \text{si } G(\tau, \mathbf{p}_\alpha, \mathbf{p}_\beta) > 0 \\ f_{\max} & \text{si } G(\tau, \mathbf{p}_\alpha, \mathbf{p}_\beta) < 0 \end{cases} \quad (2.24)$$

Obtenemos un control *Bang-Bang*, es decir una función que toma solo dos valores (??)