CALCULUS OF VARIATIONS

In calculus, one studies min-max problems in which one looks for a number or for a point that minimizes (or maximizes) some quantity. The calculus of variations is about min-max problems in which one is looking not for a number or a point but rather for a function that minimizes (or maximizes) some quantity.

For example: given two points \((x_0, y_0)\) and \((x_1, y_1)\), find the shortest curve (that is a graph) joining the two points. That is, find a function \(y(\cdot) : [x_0, x_1] \rightarrow \mathbb{R}\) with \(y(x_0) = y_0\) and \(y(x_1) = y_1\) that makes the arclength

\[ L[y(\cdot)] = \int_{x_0}^{x_1} \sqrt{1 + \dot{y}^2} \, dx \]

as small as possible. (Here \(\dot{y}\) denotes \(y'(x) = \frac{dy}{dx}\).

If we let \(s\) denote arclength, then

\[ ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + (\frac{dy}{dx})^2} \, dx. \]

More generally, given any \(C^2\) function \(L : \mathbb{R}^3 \rightarrow \mathbb{R}\), we can look for a function \(y(\cdot) : [x_0, x_1] \rightarrow \mathbb{R}\) that makes the quantity

\[ L[y(\cdot)] = \int_{x_0}^{x_1} L(x, y(x), \dot{y}(x)) \, dx \]

as small as possible.

In general, the minimum might not exist. However, if the minimum does exist, then it has to satisfy a differential equation called the Euler-Lagrange Equation. If we can solve the Euler-Lagrange Equation, then we can find the minimum (if it exists.)

**Theorem 1.** Suppose \(y(\cdot) : [x_0, x_1] \rightarrow \mathbb{R}\) is a \(C^2\) function that minimizes

\[ L[y(\cdot)] = \int_{x_0}^{x_1} L(x, y(x), \dot{y}(x)) \, dx \]

subject to the boundary conditions \(y(x_0) = y_0\) and \(y(x_1) = y_1\). Then \(y(\cdot)\) is a solution to the differential equation

\[ \frac{\partial L}{\partial y} - \frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}} \right) = 0. \]

**Notation:** Here it’s important to understand the distinction between \(\frac{\partial}{\partial x}\) and \(\frac{d}{dx}\). Note that \(L\) is a function of three variables which we denote \(x, y, \) and \(\dot{y}\). As usual, \(\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y},\) and \(\frac{\partial L}{\partial \dot{y}}\) denote its partials with respect to those variables. The composed function \(L(x, y(x), \dot{y}(x))\) is a function of one variable (namely \(x\)); its derivative is written \(\frac{dL}{dx}\). Thus \((*)\) can be written as

\[ D_2 L(x, y(x), \dot{y}(x)) - \frac{d}{dx} D_3 L(x, y(x), \dot{y}(x)). \]
Proof. Consider a $C^2$ function $u : [x_0, x_1] \to \mathbb{R}$ that vanishes on the endpoints: $u(x_0) = u(x_1) = 0$. Then the function $y(\cdot) + u(\cdot)$ also satisfies the boundary conditions, so

$$\mathcal{L}[y(\cdot)] \leq \mathcal{L}[y(\cdot) + u(\cdot)].$$

More generally,

$$\mathcal{L}[y(\cdot)] \leq \mathcal{L}[y(\cdot) + su(\cdot)]$$

for every $s \in \mathbb{R}$. Thus the function $f(s) := \mathcal{L}[y(\cdot)] \leq \mathcal{L}[y(\cdot) + su(\cdot)]$ has its minimum at 0, so $f'(0) = 0$ if the derivative exists.

In fact, the derivative does exist and we can calculate it as follows:

$$f'(s) = \frac{d}{ds} \int_{x_0}^{x_1} L(x, y(x) + su(x), y'(x) + su'(x)) \, dx$$

Furthermore,

$$f'(s) = \int_{x_0}^{x_1} \frac{d}{ds} L(x, y(x) + su(x), y'(x) + su'(x)) \, dx$$

so

$$f'(0) = \int_{x_0}^{x_1} \left( \frac{\partial L}{\partial y} (x, y(x), y'(x))u(x) + \frac{\partial L}{\partial y'} (x, y(x), y'(x))u'(x) \right) \, dx$$

or simply

$$f'(0) = \int_{x_0}^{x_1} \left( \frac{\partial L}{\partial y} u(x) + \frac{\partial L}{\partial y'} (x, y(x), y'(x))u'(x) \right) \, dx$$

Integrating the second expression by parts gives

$$f'(0) = \int_{x_0}^{x_1} \left( \frac{\partial L}{\partial y} u(x) - \frac{d}{dx} \left( \frac{\partial L}{\partial y'} u(x) \right) \right) \, dx$$

The last expression vanishes since $u(x_0) = u(x_1) = 0$. Thus

$$f'(0) = \int_{x_0}^{x_1} \left( \frac{\partial L}{\partial y} u(x) - \frac{d}{dx} \left( \frac{\partial L}{\partial y'} u(x) \right) \right) \, dx$$

Thus we have shown

$$\int_{x_0}^{x_1} \left( \frac{\partial L}{\partial y} u(x) - \frac{d}{dx} \left( \frac{\partial L}{\partial y'} u(x) \right) \right) \, dx = 0.$$
Indeed, it would be everywhere $> 0$ or everywhere $< 0$ on that interval. Now let $u$ be a $C^2$ function that is $> 0$ on $(a, b)$ and $0$ on $\mathbb{R} \setminus (a, b)$. For example, we could let

$$u(x) = \begin{cases} (x - a)^4(b - x)^4 & \text{if } x \in [a, b], \\ 0 & \text{if } x \notin [a, b]. \end{cases}$$

Then the integral in ($\dagger$) is nonzero, a contradiction.

Note that $y(\cdot)$ being a solution of the Euler-Lagrange equation does not imply that $y(\cdot)$ minimizes $L$. Rather, it means that $y(\cdot)$ passes the first derivative test for being a minimum. However, as in calculus, if we’re lucky, then the first derivative will narrow our search down to a few possibilities.

1. **Example: Shortest Curve**

Let’s try to find a function that minimizes the arclength of its graph

$$L[y(\cdot)] = \int_{x_0}^{x_1} \sqrt{1 + y^\prime(x)^2} \, dx.$$  

Here $L(x, y, y^\prime) = \sqrt{1 + y^\prime(x)^2}$. Thus

$$\frac{\partial L}{\partial y} = 0 \quad \text{and} \quad \frac{\partial L}{\partial y^\prime} = \frac{y^\prime}{\sqrt{1 + y^\prime(x)^2}},$$

so the Euler-Lagrange Equation becomes

$$0 = \frac{\partial L}{\partial y} - \frac{d}{dx} \left( \frac{\partial L}{\partial y^\prime} \right)$$

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

$$= -\frac{d}{dx} \left( \frac{y^\prime}{\sqrt{1 + y^\prime(x)^2}} \right).$$

Thus $y^\prime/\sqrt{1 + y^\prime(x)^2}$ must be constant, and therefore $y^\prime$ must be constant. Thus $y = ax + b$ for constants $a$ and $b$.

From the boundary conditions, we see that

$$y = \frac{y_1 - y_0}{x_1 - x_0} x + y_0.$$

We have not proved that the minimum exists. However, we have proved that if the minimum does exist, it must be the function $[1]$.

2. **Example: Catenoids**

The **Plateau Problem** is the following: given one or more closed curves in $\mathbb{R}^3$, find a surface of least possible area among all surfaces having those curves as boundary. Let us consider a special case of the Plateau Problem: we look for a least area surface whose boundary is a pair of circles, assuming that the minimum exists and is a surface of revolution.

In other words, suppose $0 < x_0 < x_1$ and that $y(\cdot) : [x_0, x_1] \to \mathbb{R}$ is a $C^2$ function. We can rotate the graph of $y(\cdot)$ about the $y$-axis to get a surface $S$ of
revolution in \( \mathbb{R}^3 \). The area of \( S \) is given by
\[
\int_{x_0}^{x_1} 2\pi x \sqrt{1 + (y')^2} \, dx
\]
Let’s try to find a function \( y(\cdot) \) that minimizes this area (subject to specified boundary conditions \( y(x_0) = y_0 \) and \( y(x_1) = y_1 \)). The problem is equivalent to minimizing
\[
\mathcal{L}[y(\cdot)] = \int_{x_0}^{x_1} x \sqrt{1 + (y')^2} \, dx
\]
Here \( L(x, y, \dot{y}) = x \sqrt{1 + (\dot{y})^2} \), so
\[
\frac{\partial L}{\partial y} = 0, \quad \frac{\partial L}{\partial \dot{y}} = x \frac{\dot{y}}{1 + (\dot{y})^2},
\]
so the Euler-Lagrange Equation becomes
\[
0 = \frac{\partial L}{\partial y} - \frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}} \right)
= 0 - \frac{d}{dx} \left( \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} \right)
= -\frac{d}{dx} \left( \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} \right).
\]
Thus
\[
\frac{x \dot{y}}{\sqrt{1 + \dot{y}^2}} = c.
\]
Solving for \( \dot{y} \),
\[
(x^2 - c^2)(\dot{y})^2 = c^2,
\]
or
\[
\frac{dy}{dx} = \frac{c}{\sqrt{x^2 - c^2}},
\]
so
\[
y = \int \frac{c}{\sqrt{x^2 - c^2}} \, dx.
\]
To integrate, let \( x = c \cosh u \). Then \( dx = c \sinh u \) and \( x^2 - c^2 = c^2(\cosh^2 u - 1) = c^2 \sinh^2 u \). Thus
\[
y = \int c \, du = cu + \hat{c},
\]
so
\[
u = \frac{y - \hat{c}}{c}.
\]
Taking cosh of both sides gives
\[
(2) \quad \frac{x}{c} = \cosh \left( \frac{y - \hat{c}}{c} \right).
\]

One can think of \( x = \cosh y \) as the “basic” solution. All other solutions come be dilating the fundamental solution (by \( (x, y) \mapsto (cx, cy) \)) and then translating in the \( y \)-direction (by \( (x, y) \mapsto (x, y + \frac{\hat{c}}{c}) \)).

Of course we try to choose \( c \) and \( \hat{c} \) so that the solution curve passes through the point \((x_0, y_0)\) and \((x_1, y_1)\).
**Definition 2.** The surface of revolution given by \( \sqrt{x^2 + z^2} = \cosh y \) is called a *catenoid*. More generally, if we apply a translation, rotation, and dilation to that surface, the resulting surface is also called a catenoid.

Suppose \( (x_0, y_0) = (1, -h) \) and \( (x_1, y_1) = (0, h) \). (Geometrically, this means that in \( \mathbb{R}^3 \), the boundary of our surface consists of two circles of radius 1, one in the plane the \( x = -h \) and the other in the plane \( x = h \).)

**Exercise 1.** Show that if \( h \) is small, then there are exactly two curves of the form (2) that pass through the points \( (1, -h) \) and \( (1, h) \). Which one has less area?

**Exercise 2.** Show that if \( h \) is large, then there is no curve of the form (2) passing through \( (1, -h) \) and \( (1, h) \). What can you conclude?

***

A curious thing happened in the catenoid example above. We were looking for a function \( y = y(x) \), but we ended up with a function \( x = x(y) \). Suppose for example that \( (x_0, y_0) = (\cosh 1, -1) \) and \( (x_1, y_1) = (\cosh 2, 2) \). Then

\[
x = \cosh y, \quad -1 \leq y \leq 2
\]

is a curve that has the specified endpoints. However, it cannot be written in the form \( y = y(x) \).

So is our analysis valid? It is in the following sense. Suppose we allow curves \( C \) joining \( (x_0, y_0) \) and \( (x_1, y_1) \) that are not necessarily graphs. Then Theorem 1 does apply to each portion that is a graph.

**Conserved Quantities**

Note that if \( L(x, \dot{y}) \) is actually a function of \( x \) and \( \dot{y} \) alone, then \( \frac{\partial L}{\partial y} = 0 \), so the Euler-Lagrange equation simplifies to

\[
\frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}} \right) = 0.
\]

Thus \( y(\cdot) \) is a solution if and only if

\[
\frac{\partial L}{\partial \dot{y}} = c
\]

for some constant \( c \). Thus \( \frac{\partial L}{\partial \dot{y}} \) is a “conserved quantity”; it doesn’t change as \( x \) changes.

Similarly, if \( L = L(y, \dot{y}) \) a a function \( y \) and \( \dot{y} \) alone, there is also a conserved quantity:

**Theorem 3.** Suppose \( L = L(y, \dot{y}) \). Then a nonconstant function \( y(\cdot) \) is solution of the Euler-Lagrange equation if and only if the quantity

\[
Q = \dot{y} \frac{\partial L}{\partial \dot{y}} - L
\]

is constant (i.e., independent of \( x \)).
Proof. Note that
\[ \frac{dQ}{dx} = \ddot{y} \frac{\partial L}{\partial y} + \dot{y} \frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}} \right) - \frac{dL}{dx}. \]

By the chain rule,
\[ \frac{dL}{dx} = \frac{\partial L}{\partial y} \dot{y} + \frac{\partial L}{\partial \dot{y}} \ddot{y}. \]
(Note that if \( L = L(x, y, \dot{y}) \), the right hand side would also include \( \frac{\partial L}{\partial x} \).) Thus
\[ \frac{dQ}{dx} = \dot{y} \left( \frac{\partial L}{\partial y} - \frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}} \right) \right). \]

\[ \Box \]

Vector-Valued Functions

The derivation of the Euler-Lagrange Equation works equally for vector-valued function \( y : [x_0, x_1] \rightarrow \mathbb{R}^n \). Here \( L \) will be a function of \( 2n + 1 \) variables:
\[ L = L(x, y_1, \ldots, y_n, \dot{y}_1, \ldots, \dot{y}_n). \]

In this case, the Euler-Lagrange Equation becomes a system of differential equations:
\[ \frac{\partial L}{\partial y_i} - \frac{d}{dx} \left( \frac{\partial L}{\partial \dot{y}_i} \right) = 0 \quad (1 \leq i \leq n). \]