Introduction to the Calculus of Variations

Ariel Goodwin

McGill University

January 8th, 2022

Overview

1. Introduction
2. The First Variation and Euler's Equation
3. The Second Variation: Legendre's Condition
4. A Special Quadratic Functional
5. Jacobi's Necessary Condition and a Sufficient Condition

Motivation

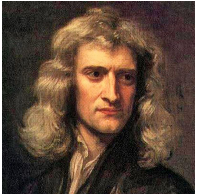
- Calculus teaches you how to optimize nice functions from $\mathbb{R}^n$ to $\mathbb{R}$
- Optimization and convex analysis give us tools for optimizing more general functions
- What about optimizing functions of functions (functionals)?
- Good reason to develop functional analysis
- The **calculus of variations** provides techniques and theory for solving problems of applied and theoretical interest.
- Goal: Introduce the essential definitions/results + historical impact

Historical Perspective

Hadamard, on the calculus of variations (1910)

"A first chapter of functional analysis, whose development will without doubt be one of the first tasks in the analysis of the future."

(a) Euler



(b) Newton

(c) Legendre

(d) Jacobi

Figure: A few contributors to the calculus of variations

The Brachistochrone Problem

- What curve minimizes travel time from A to B for a particle under influence of gravity?
- Associated with the birth of calculus of variations
- Posed by Johann Bernoulli (1696)
- Solved by Newton, Leibniz, and Bernoulli brothers Jacob and Johann
- Newton used such methods even earlier (1694) to solve gun shell problem

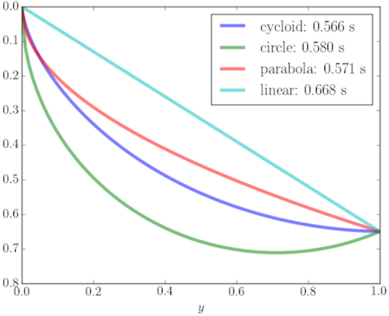


Figure: Curves from A to B

Functionals

The calculus of variations seeks extrema of functionals.
Assumptions:

- X is a normed linear space
- Functionals are maps $J\colon X \to \mathbb{R}$

Given $y \in X$, we can perturb it slightly by adding $h \in X$:

$$\Delta J[h] := J[y+h] - J[y]$$

Definition 1: Variation

Given $y \in X$, if there exists a continuous linear functional φ such that

$$\Delta J[h] = \varphi[h] + \varepsilon \|h\|$$

where $\varepsilon = \varepsilon(\|h\|) \to 0$ as $\|h\| \to 0$ we say J is *differentiable* and $\delta J[h] := \varphi[h]$ is its *variation*.

Extrema

Definition 2: Extrema

A functional J has a *relative extremum* at $\hat{y}$ if $J[y] - J[\hat{y}]$ maintains the same sign in a small neighbourhood of $\hat{y}$.

In what follows, we focus on $X = C^1[a, b]$. Choice of norm?

Definition 3: Weak and Strong Extrema

A point $\hat{y}$ is a *weak extremum* if there exists $\varepsilon > 0$ such that $J[y] - J[\hat{y}]$ maintains the same sign for all y satisfying $\|y - \hat{y}\|_1 < \varepsilon$ where $\|\cdot\|_1$ is the Sobolev norm. A *strong extremum* is characterized by the existence of $\varepsilon > 0$ such that $J[y] - J[\hat{y}]$ maintains the same sign for all y satisfying $\|y - \hat{y}\|_\infty < \varepsilon$.

Strong extrema are weak.

A Simple Necessary Condition

Theorem 1: Vanishing Variation

If $\hat{y}$ is an extremum of a differentiable functional J then $\delta J[h] = 0$ at $\hat{y}$ for all admissible h.

Proof.

WLOG $\hat{y}$ is a minimum, and for a contradiction suppose $\delta J[h_0] \neq 0$ for some h_0. Then

$$\Delta J[h] = \delta J[h] + \varepsilon \|h\|$$

Thus for $\|h\|$ small, $\Delta J[h]$ and $\delta J[h]$ have the same sign. For $\alpha > 0$ sufficiently small we have $\Delta J[\pm \alpha h_0] \geq 0$ by optimality of $\hat{y}$. But since $\delta J[h_0] \neq 0$ we can use linearity of the variation to obtain

$$\delta J[\pm \alpha h_0] = \pm \alpha \delta J[h_0]$$

Thus $\Delta J[\pm \alpha h_0]$, which shares the sign of $\delta J[\pm \alpha h_0]$ for small α, can be made negative which contradicts our assumption. □

Problem of Interest

Find weak extrema of

$$J[y] := \int_a^b F(x, y(x), y'(x))dx$$
$$\text{s.t.} \quad y \in C^1[a,b]$$
$$y(a) = A, y(b) = B$$

for $F \in C^2$. Examples: brachistochrone, minimal surface of rotation, shortest path between two points.
Idea: Perturb y by $h \in C_0^1[a,b]$ to preserve boundary conditions and use Taylor's theorem

$$\Delta J[h] = \int_a^b F_y(x, y, y')h + F_{y'}(x, y, y')h' dx + \cdots$$

where $\cdots$ are $o(\|h\|)$ terms. Then by definition

$$\delta J[h] = \int_a^b F_y(x, y, y')h + F_{y'}(x, y, y')h' dx$$

The Fundamental Lemma of Calculus of Variations

Apply our necessary condition and set equal to zero:

$$\delta J[h] = \int_a^b F_y(x, y, y')h + F_{y'}(x, y, y')h' dx = 0$$

This hold for all admissible h - how to use this?

Lemma 1: du Bois-Reymond

Let $\alpha, \beta \in C[a, b]$. Then

1. If $\int_a^b \alpha(x)h(x)dx = 0$ for all $h \in C_0[a, b]$ then $\alpha = 0$.
2. If $\int_a^b \alpha(x)h'(x)dx = 0$ for all $h \in C_0^1[a, b]$ then α is a constant function.
3. If $\int_a^b \alpha(x)h''(x)dx = 0$ for all $h \in C_0^2[a, b]$ then α is an affine function.
4. If $\int_a^b \alpha(x)h(x) + \beta(x)h'(x)dx = 0$ for all $h \in C_0^1[a, b]$ then β is differentiable on $[a, b]$ with $\beta' = \alpha$.

Euler's Equation

Use Part (4) from Fundamental Lemma on

$$\int_a^b F_y(x, y, y')h + F_{y'}(x, y, y')h' dx = 0$$

Theorem 2: Euler's Equation

Let J be a functional of the form $J[y] = \int_a^b F(x, y, y')dx$ defined on the subset of $C^1[a, b]$ satisfying the boundary conditions $y(a) = A, y(b) = B$. Then a necessary condition for J to have an extremum at y is that y satisfy *Euler's equation*:

$$F_y - \frac{d}{dx} F_{y'} = 0$$

Special Cases: If F does not depend explicitly on one of x, y, y' then the situation simplifies considerably. If F does not depend on x explicitly (autonomous), Euler's equation gives us

$$0 = F_y y' - F_{y'y} y'^2 - F_{y'y'} y' y'' = \frac{d}{dx}\left(F - y' F_{y'}\right)$$

$$\implies F - y' F_{y'} = C$$

Application to Minimal Surfaces

Q: Given points $(a, y_0), (b, y_1)$, find the curve between them which generates the minimum surface area when rotated around the x-axis.

Surface area functional:
$J[y] = 2\pi \int_a^b y\sqrt{1+y'^2}dx$

Integrand is autonomous:
$F - y'F_{y'} = C$

$$2\pi \left(y\sqrt{1+y'^2} - \frac{yy'^2}{\sqrt{1+y'^2}} \right) = C$$

$$\Downarrow$$

$$\frac{y}{\sqrt{1+y'^2}} = B$$

for some constant B.
Separation of variables yields
$y = A\cosh\left(\frac{x+B}{A}\right)$ for constants A, B.

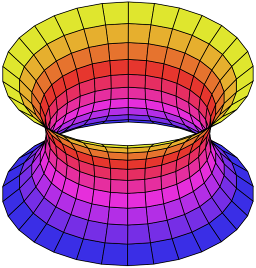


Figure: A catenoid [6].

The Second Variation

Definition 4: Quadratic Functional

Given a bilinear functional $B\colon X \times X \to \mathbb{R}$, a *quadratic functional* A is defined by

$$A[x] := B[x, x]$$

The functional A is said to be *positive definite* if $A > 0$ unless $x = 0$.

Definition 5: Second Variation

If there exists a continuous linear functional φ_1 and a continuous quadratic functional φ_2 such that

$$\Delta J[h] = \varphi_1[h] + \varphi_2[h] + \varepsilon\|h\|^2$$

where $\varepsilon = \varepsilon(\|h\|) \to 0$ as $\|h\| \to 0$ we say that J is *twice differentiable* and we call φ_2 the *second variation* of $J[y]$ and we denote it by $\delta^2 J[h] = \varphi_2[h]$.

Another Necessary Condition

Theorem 2: Nonnegative Second Variation

If $\hat{y}$ is a minimum for the functional J then $\delta^2 J[h] \geq 0$ at $\hat{y}$ for all admissible h. A symmetric statement with the inequality reversed holds in the case of a maximum.

Proof.

By definition

$$\Delta J[h] = \delta J[h] + \delta^2 J[h] + \varepsilon \|h\|^2$$

The first variation $\delta J[h]$ vanishes at an extremum, so

$$\Delta J[h] = \delta^2 J[h] + \varepsilon \|h\|^2$$

This implies $\Delta J[h]$ and $\delta^2 J[h]$ have the same sign for $\|h\|$ small. Suppose for a contradiction that $\delta^2 J[h_0] < 0$ for some admissible h_0. Taking $\alpha > 0$ small we have

$$\delta^2 J[\alpha h_0] = \alpha^2 \delta^2 J[h_0] < 0$$

which contradicts the fact that $\Delta J[\alpha h_0] \geq 0$ by optimality of $\hat{y}$. $\square$

Revisiting the Problem of Interest

Find weak extrema of

$$J[y] := \int_a^b F(x, y(x), y'(x))dx$$
$$\text{s.t.} \quad y \in C^1[a,b]$$
$$y(a) = A, y(b) = B$$

for $F \in C^3$.
As before, perturb y by h and use Taylor's theorem.

$$\Delta J[h] = \int_a^b F_y(x,y,y')h + F_{y'}(x,y,y')h' dx + \frac{1}{2}\int_a^b F_{yy}h^2 + 2F_{yy'}hh' + F_{y'y'}h'^2 dx + \varepsilon$$

The C^3 assumption tells us $\varepsilon = o(\|h\|_1^2)$ and the first term is $\delta J[h]$, hence

$$\delta^2 J[h] = \frac{1}{2}\int_a^b F_{yy}h^2 + 2F_{yy'}hh' + F_{y'y'}h'^2 dx$$

A Special Quadratic Functional

$$\delta^2 J[h] = \frac{1}{2}\int_a^b F_{yy}h^2 + 2F_{yy'}hh' + F_{y'y'}h'^2 dx$$

Use integration by parts and the boundary conditions on h to realize

$$\int_a^b 2F_{yy'}hh'dx = -\int_a^b \frac{d}{dx}\left(F_{yy'}\right)h^2 dx$$

Set $P(x) = \frac{1}{2}F_{y'y'}$ and $Q(x) = \frac{1}{2}\left(F_{yy'} - \frac{d}{dx}\left(F_{yy'}\right)\right)$. Then rewrite the second variation as

$$\delta^2 J[h] = \int_a^b Ph'^2 + Qh^2 dx$$

Theorem 4

If the quadratic functional $\delta^2 J[h] = \int_a^b Ph'^2 + Qh^2 dx$ defined on $C_0^1[a, b]$ is non-negative, then P is non-negative on $[a, b]$.

Proof Idea: Choose h with $\|h'\|_\infty$ large and $\|h\|_\infty$ small so that Ph'^2 term dominates.

Legendre's Condition

Theorem 5: (Legendre)

If the functional

$$J[y] = \int_a^b F(x, y, y')dx, \quad y(a) = A, y(b) = B$$

has a minimum at y then $F_{y'y'} \geq 0$ at every point along the curve y.

Proof.

Follows immediately from Nonnegative Second Variation and the previous theorem. $\square$

Q: Can a strict inequality (*strengthened Legendre condition*) turn this into a sufficient condition?
Legendre "proved" this, but there was a serious flaw for which he was heavily criticized by Lagrange. Need stronger assumptions.

Jacobi's Equation

Ignoring its second variation origins, study the quadratic functional

$$A[h] = \int_a^b Ph'^2 + Qh^2 dx$$

Its Euler equation is known as the *Jacobi equation*:

$$-\frac{d}{dx}\left(Ph'\right) + Qh = 0$$

with constraints $h(a) = h(c) = 0$ for fixed $c \in (a, b]$. Trivial solution: $h = 0$

Definition 6: Conjugate Points

A point $a_0 \in (a, b]$ is *conjugate* to a with respect to the quadratic functional A if the corresponding Jacobi equation has a nontrivial solution that vanishes at a_0 and a.

Conjugate points are another obstacle to sufficiency.

Jacobi's Necessary Condition

Theorem 6

If $P(x) > 0$ on $[a, b]$ and the quadratic functional $\int_a^b Ph'^2 + Qh^2 dx$ is non-negative (resp. positive definite) for h satisfying the zero boundary conditions then there are no points conjugate to a in (a, b) (resp. $(a, b]$).

Proof Idea: Define the family $\int_a^b t(Ph'^2 + Qh^2) + (1-t)h'^2 dx$ for $t \in [0, 1]$. Note that for $t = 0$ there are no conjugate points, argue geometrically that conjugate points cannot arise as $t \to 1$.

Theorem 7: (Jacobi)

If y minimizes the functional $J[y] = \int_a^b F(x, y, y')dx$ and $F_{y'y'} > 0$ at all points of the curve y, then there are no points conjugate to a in (a, b).

Proof.

Nonnegative Second Variation gives $\int_a^b Ph'^2 + Qh^2 dx \geq 0$, and then $F_{y'y'} > 0$ implies $P > 0$ and so Theorem 6 completes the argument. $\square$

A Sufficient Condition

Euler's equation and the Legendre and Jacobi conditions have all been necessary. Our last theorem ties everything together with sufficiency.

Theorem 8

Suppose that for some admissible curve y, the functional

$$J[y] = \int_a^b F(x, y, y')dx, \;\; y(a) = A, y(b) = B$$

satisfies the following conditions:

1. The curve y satisfies Euler's equation, $F_y - \frac{d}{dx}F_{y'} = 0$
2. Along the curve y, $F_{y'y'}(x, y(x), y'(x)) > 0$ (strengthened Legendre condition)
3. The interval $[a, b]$ does not contain any points conjugate to a (strengthened Jacobi condition).

Then J admits a weak minimum at y.

Summary

If you need to optimize functionals, look to the calculus of variations! Its principles lend themselves to:

- Geometry
- Differential Equations
- Optimization and Optimal Control Theory
- Mechanics and Physics

Further reading:

- Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*
- Gelfand and Fomin, *Calculus of Variations*
- Kreyszig, *On the Calculus of Variations and Its Major Influences on the Mathematics of the First Half of Our Century*

References

F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Graduate Texts in Mathematics, Springer-Verlag London, 2013.

I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Prentice-Hall. Inc. Englewood Cliffs, New Jersey, 1963.

E. Kreyszig. *On the Calculus of Variations and Its Major Influences on the Mathematics of the First Half of Our Century*. The American Mathematical Monthly, Vol. 101, No. 7, pp. 674-678, 1994.

Assorted mathematicians. Wikimedia, 2021.

Image of brachistochrone data. `https://scipython.com/blog/the-brachistochrone-problem/`, 2017.

Catenoid image. Wikimedia, 2021.