Variational Principles

Jonathan Gai

5th May 2022

Contents

0.1 The Brachistochrone Problem 1
0.2 Geodesics 2
0.3 Introduction 2

1 Calculus for Functions of $\mathbb{R}^n$ 3
1.1 Constraints and Lagrange Multipliers 5

2 Euler-Lagrange Equations 7
2.1 First Integrals of the Euler-Lagrange Equations 9

Lecture 1

29 Apr. 2022

Motivation

0.1 The Brachistochrone Problem

Problem. Particle slides on a wire under influence of gravity between two fixed points A, B. Which shape of the wire gives the shortest travel time, starting from rest?

The travel time is $T = \int_A^B \frac{\mathrm{d}\ell}{v(x,y)}$, and by energy conservation, and by energy conservation

$$\frac{1}{2}mv^2 + mgy = 0 \implies v = \sqrt{-2gy}.$$

So

$$T[y] = \frac{1}{\sqrt{2g}} \int_0^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{-y}}\,\mathrm{d}x$$

subject to $y(0) = 0$, $y(x_2) = y_2$.

0.2 Geodesics

Problem. What is the shortest path γ between two points A, B on a surface.

Take $\Sigma = \mathbb{R}^2$. The distance along γ is

$$D[y] = \int_A^B \mathrm{d}\ell = \int_{x_1}^{x_2} \sqrt{1+(y')^2}\,\mathrm{d}x,$$

and we want to minimize D by varying γ.

0.3 Introduction

In general, we want to minimize (maximize)

$$F[y] = \int_{x_1}^{x_2} f(x, y, y')\,\mathrm{d}x$$

among all functions s.t. $y(x_1) = y_1$, $y(x_2) = y_2$. The expression is a *functional*. (A function on a space of functions). Functions map numbers to numbers, and functionals map functions to numbers.

Area under the graph is when $f = y$ and the length of a curve is when $f = \sqrt{1+(y')^2}$.

Calculus of variations finds extrema of functionals on space of functions.

Notation.

- $C(\mathbb{R})$ is space of continuous functions on $\mathbb{R}$.
- $C^k(\mathbb{R})$ is the space of functions with continuous kth derivatives.
- $C^k_{(\alpha,\beta)}$ is the space of functions with continuous kth derivatives and $f(\alpha) = f(\beta)$.

We need to specify the function space beforehand. It is a branch of functional analysis—Part III analysis on the space of functions, while Analysis I is analysis on the number line. Variational Principles follows principles in Nature, where the laws follow from extremizing functionals.

Example (Fermat's Principle)**.** Light between two pints travel along paths which require least time.

Example (Principle of Least Action)**.** Let T be the kinetic energy $\frac{m|\dot{\mathbf{x}}|^2}{2}$ and potential energy $V = V(\mathbf{x})$.

$$S[\gamma] = \int_{t_1}^{t_2} (T - V)\,\mathrm{d}t$$

is minimized along paths of motion.

Leibniz commented on this, saying that "we live in the best of all worlds".

Feynman's take on this: "This is wrong. In quantum theorem the motion takes place along all possible path with different possibilities". [See Part III QFT]

In this course, we discuss

1. necessary condition for extrema of the Euler Lagrange Equations;
2. lots of examples (geometry, physics, problems with constraints);
3. second variation (some sufficient condition of extrema).

The following books will be useful

1. Gelfand - Fomin "Calculus of Variations";
2. DAMTP notes (e.g. P. Townsend);
3. Lectures are self-contained.

Lecture 2

2 May 2022

1 Calculus for Functions of $\mathbb{R}^n$

We consider $f \in C^2(\mathbb{R}^n)$, that is, all functions $f : \mathbb{R}^n \to \mathbb{R}^n$ with continuous second derivatives. The point $\mathbf{a} \in \mathbb{R}$ is *stationary* if

$$\boldsymbol{\nabla} f(\mathbf{a}) = \partial_1 f, \ldots, \partial_n f|_{\mathbf{x}=\mathbf{a}} = 0, \quad \partial_i f = \frac{\partial f}{\partial x^i}.$$

If we expand near $\mathbf{a}$, we have

$$f(\mathbf{x}) = f(\mathbf{a}) + (\mathbf{x} - \mathbf{a}) \cdot \nabla|_{\mathbf{a}} + \frac{1}{2}(x_i - a_i)(x_j - a_j)\partial_{ij}^2 f|_{\mathbf{a}} + \mathcal{O}(|\mathbf{x} - \mathbf{a}|^2).$$

The *Hessian matrix* is $H_{ij} = \partial_i \partial_j f = H_{ij}$. We shift the origin to set $\mathbf{a} = \mathbf{0}$, and diagonalize $H(\mathbf{0})$ by orthogonal transformation

$$H' = R^T H(\mathbf{0}) R = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad f(\mathbf{x}') - f(\mathbf{0}) = \frac{1}{2}\sum_i \lambda_i (x_i')^2 + \mathcal{O}(|\mathbf{x}|^2).$$

We consider several cases.

1. If all $\lambda_i > 0$, $f(\mathbf{x}') > f(\mathbf{0})$ in all directions, it is a *local minimum*.

2. If all $\lambda_i < 0$, $f(\mathbf{x}') < f(\mathbf{0})$ in all directions, it is a *local maximum*.

3. If some $\lambda_i > 0$ and some $\lambda_i < 0$, f increases in some directions, and decreases in other directions. It is a *saddle point*.

4. If some $\lambda_i = 0$ and the rest are of the same sign, we need to consider higher order terms in expansion.

The special case when $n = 2$, we have $\det(H) = \lambda_1 \lambda_2$ and $\mathrm{Tr}(H) = \lambda_1 + \lambda_2$. If $\det(H) > 0$, $\mathrm{Tr}(H) > 0$, it is a local minimum. If $\det(H) < 0$, $\mathrm{Tr}(H) < 0$, it is a local maximum. If $\det(H) < 0$, it is a saddle point. If $\det(H) = 0$, we need to look at third and higher derivatives.

Remark.

1. If $f : D \to \mathbb{R}$ maps from a domain. We need to check boundary values too for global extrema.

2. If f is harmonic on $\mathbb{R}^2$, that is $f_{xx} + f_{yy} = 0 \implies \mathrm{Tr}(H) = 0$. If $D \subsetneq \mathbb{R}^2$, all turning points are saddle points, and the extrema must be on the boundary.

Example. Take $f(x, y) = x^3 + Y^3 - 3xy$, and we have $\nabla f = (3x^2 - 3y, 3y^2 - 3x)$. So the critical points are points where $x^2 - y = 0$ and $y^2 - x = 0$. So the stationary points are $x = 0, y = 0$ and $x = 1, y = 1$. The Hessian matrix is

$$H = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

At $(0, 0)$, $\det(H) = -9 < 0$, so it's a saddle point at $(0, 0)$. And at $(1, 1)$, $\det(H) = 27 > 0$, $\mathrm{Tr}(H) = 12 > 0$, it is a local minimum.

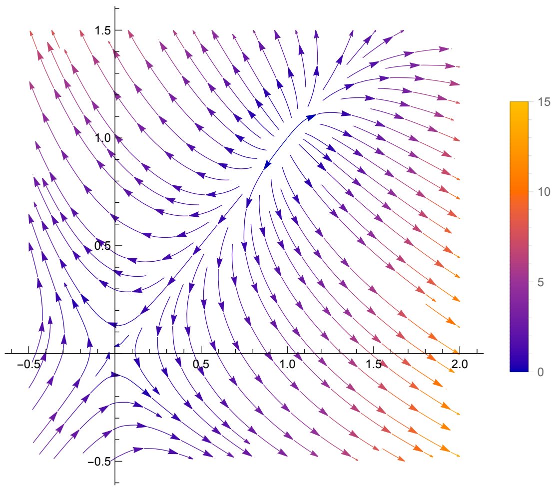

Figure 1: Plot of f showing the critical points

1.1 Constraints and Lagrange Multipliers

Example. Find the circle centered at $(0, 0)$ with the smallest radius, which intersects the parabola $y = x^2 - 1$.

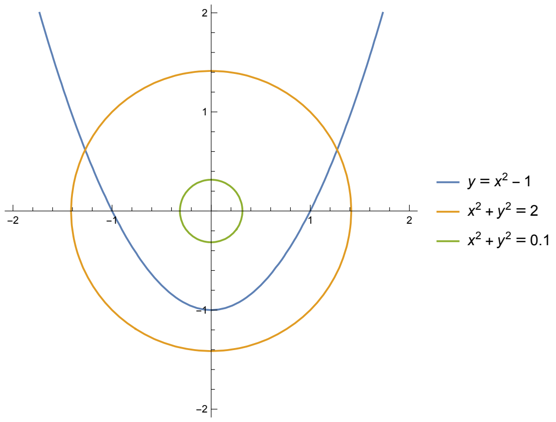

Figure 2: Circles and parabola

The direct method is to solve the constraints

$$f = x^2 + y^2 = x^2 + (x^2 - 1)^2 = x^4 - x^2 + 1.$$

If $\partial f = 0$, we have $4x^3 - 2x = 0$. So the solutions are $(0, -1)$ with radius 1 and $(\pm\frac{1}{\sqrt{2}}, -\frac{1}{2})$ with radius $\frac{\sqrt{3}}{2} < 1$ which is the global minimum.

The other method is *Lagrange Multiplier*. Define $h(x, y, \lambda) = f(x, y) - \lambda g(x, y)$ where $g(x, y) = 0$ is the constraint and λ is the Lagrange multiplier. Now we extremize the function $h(x, y) = x^2 + y^2 - \lambda(y - x^2 + 1)$over 3 variables with no constraints; that is,

$$\begin{cases} \dfrac{\partial h}{\partial x} = 2x + 2\lambda x = 0 \\ \dfrac{\partial h}{\partial y} = 2y - \lambda = 0 \\ \dfrac{\partial h}{\partial \lambda} = y - x^2 + 1 = 0. \end{cases}.$$

Note that the last equation just gives the constraint back. The solutions are $(0, -1)$ and $(\pm\frac{1}{\sqrt{2}}, -\frac{1}{2})$ as before will $\lambda = -2$ and $\lambda = -1$, respectively.

We show that Lagrange Multiplier works by geometry.

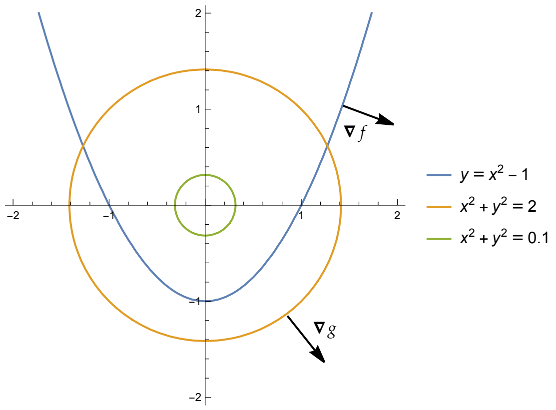

Figure 3: Gradients of the functions

We know that ∇g is perpendicular to $g = 0$ and ∇f is perpendicular to f being constant. They must be parallel so $\nabla f = \lambda \nabla g$ and $\nabla(f - \lambda g) = 0$.

If we have multiple constraints with $f : \mathbb{R}^n \to \mathbb{R}$ subject to $g_a(\mathbf{x}) = 0$ for $a = 1, \ldots, k$, then we let

$$h(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k) = f - \sum_a \lambda_a g_a.$$

And solve for $\frac{\partial h}{\partial x^i} = 0$ and $\frac{\partial h}{\partial \lambda^i} = 0$.

Lecture 3

4 May 2022

2 Euler-Lagrange Equations

If we want to extremize the functional

$$F[y] = \int_{\alpha}^{\beta} f(x, y, y')\, \mathrm{d}x. \tag{2.1}$$

for f given. It depends on y fixing only $y(\alpha) = y_1$ and $y(\beta) = y_2$. We assume that

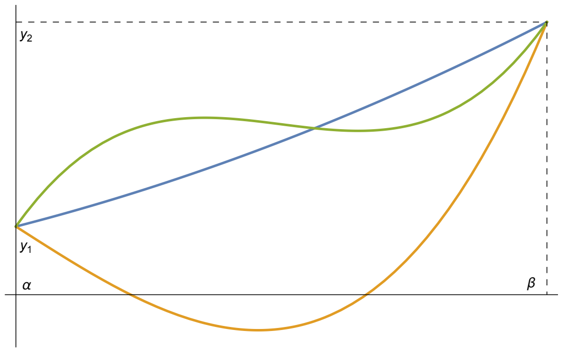

Figure 4: Different possible functions

the extremum exists, $y = y(x)$, and consider small perturbation $y \to y + \epsilon\eta(x)$ with η satisfying $\eta(\alpha) = \eta(\beta) = 0$. We want to compute $F[y + \epsilon\eta]$.

Lemma 2.1: Fundemental Lemma of Calculus of Variations

If $g : [\alpha, \beta] \to \mathbb{R}$ is continuous on $[\alpha, \beta]$ and $\int_{\alpha}^{\beta} g(x) \cdot \eta(x)\, \mathrm{d}x = 0$ for all η continuous on $[\alpha, \beta]$ with $\eta(\alpha) = \eta(\beta) = 0$, then $g = 0$ on $[\alpha, \beta]$.

Proof. Assume otherwise, $\exists \bar{x} \in (\alpha, \beta)$ such that $g(\bar{x}) \neq 0$. Without loss of generality, $g(\bar{x}) > 0$, then there exists interval $[x_1, x_2] \subseteq (\alpha, \beta)$ such that $f(x) > c > 0$ for $x \in [x_1, x_2]$ and some c. Take

$$\eta(x) = \begin{cases} (x - x_1)(x_2 - x), & \text{if } x \in [x_1, x_2] \\ 0, & \text{otherwise} \end{cases}. \tag{2.2}$$

So

$$\int_{\alpha}^{\beta} g(x)\eta(x)\, \mathrm{d}x > c \int_{x_1}^{x_2} (x - x_1)(x_2 - x)\, \mathrm{d}x > 0.$$

■

Remark. The function η is an example of a bump function. There are C^k bump functions

$$\eta(x) = \begin{cases} ((x-x_1)(x_2-x))^{k+1}, & \text{if } x \in [x_1, x_2] \\ 0, & \text{otherwise} \end{cases}.$$

Now we go back to equation (2.1).

$$\begin{aligned} F[y+\epsilon\eta] &= \int_\alpha^\beta f(x, y+\epsilon\eta, y'+\epsilon\eta')\,\mathrm{d}x \\ &= F[y] + \epsilon\int_\alpha^\beta \frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta'\,\mathrm{d}x + \mathcal{O}(\epsilon^2) \\ &= F[y] + \epsilon\left(\int_\alpha^\beta \frac{\partial f}{\partial y}\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\eta\,\mathrm{d}x + \frac{\partial f}{\partial y'}\eta\Big|_\beta^\alpha\right) \end{aligned}$$

At the extremum, we have $F[y+\epsilon\eta] = F[y] + \mathcal{O}(\epsilon^2)$; that is, when $\frac{\partial F}{\partial \epsilon}\Big|_{\epsilon=0} = 0$. So

$$\int_\alpha^\beta \left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\right)\eta\,\mathrm{d}x.$$

By Lemma, we know

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0. \tag{2.3}$$

This is the Euler-Lagrange equation, and it is a necessary condition for an extremum.

Remark.

1. Equation (2.3) is a second order ODE for $y(x)$ with boundary conditions $y(\alpha) = y_1$ and $y(\beta) = y_2$.
2. The left-hand side of equation (2.3) is denoted $\frac{\delta F}{\delta y(x)}$, and called *functional derivative*. Some books use $\delta y = \epsilon\eta(x)$ as small variation.
3. There are other boundary conditions possible. For example, $\frac{\partial f}{\partial y'}\Big|_{\alpha,\beta} = 0$.
4. One needs to be careful with derivatives. $\frac{\partial f}{\partial y} = \left(\frac{\partial f}{\partial y}\right)_{x,y'}$ treating x, y, y' as independent, and the *total derivative* is

$$\begin{aligned} \frac{\mathrm{d}h}{\mathrm{d}x} &= \frac{\partial h}{\partial x} + \frac{\partial h}{\partial y}y' + \frac{\partial h}{\partial y'}y'' \\ \frac{\mathrm{d}}{\mathrm{d}x} &= \frac{\partial}{\partial x} + y'\frac{\partial}{\partial y} + y''\frac{\partial}{\partial y'}. \end{aligned}$$

Example. If $f = x \cdot ((y')^2 - y^2)$, then

$$\partial_x f = (y')^2 - y^2, \quad \partial_y f = -2xy, \quad \partial_{y'} f = 2xy'.$$

So $\frac{\mathrm{d}f}{\mathrm{d}x} = (y')^2 - y^2 - 2xyy' + 2xy'y''$.

2.1 First Integrals of the Euler-Lagrange Equations

In some cases equation (2.3) (second order ODE) can be integrated once to a first order ODE called the *first integral*.

If f does not explicitly depend on y, that is, $\frac{\partial f}{\partial y} = 0$. Then equation (2.3) gives $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = 0$, and

$$\frac{\partial f}{\partial y'} = \text{const.} \tag{2.4}$$

Example. Geodesics on $\mathbb{R}^2$ have the functional

$$F[y] = \int_\alpha^\beta \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_\alpha^\beta \sqrt{1 + (y')^2}\,\mathrm{d}x.$$

By equation (2.4), we must have $\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1+(y')^2}} = \text{const}$. So y' is constant, calling it m. We have $y = mx + c$, a straight line.