CSE276C - Calculus of Variation

Henrik I. Christensen

Computer Science and Engineering
University of California, San Diego

October 2021

Introduction

- Going a bit more abstract today
- Calc of variations is tightly coupled to mechanics
- We will only covers the very basics
- Entire courses at UCSD - MATH201C

Applications

- Path Optimization
- Vibrating membranes
- Electrostatics
- Machine vision – reconstruction
- Vision - image flow, ...

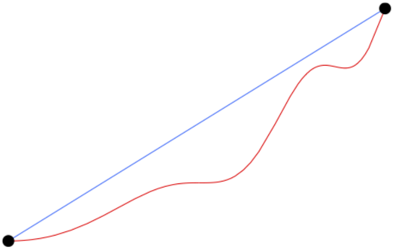
Introduction (cont)

- We have seen the principle
  - To minimize P is to solve P' $= 0$
- So far we have looked at finite dimensional problems
  - f: $\mathcal{R}^n \rightarrow \mathcal{R}$

  Looking at N numbers to minimize f
- In infinite dimensional problems we are considering an continuum
- What about functionals - (functions of functions)?

Example

- Suppose we connect two points in the plane (x_0, y_0) and (x_1, y_1) by a curve of the form $y = y(x)$.



- The length if the curve can be written

$$L(y) = \int_{x_0}^{x_1} \sqrt{1 + (y')^2} dx$$

L is a functional.

- Find the shortest curve between the two points.

Similar problems

- Shortest path connecting a non-planar curve, say sphere
- Minimal surface of revolution generated by a connected curve
- Shortest curve with a given area below it
- Closed curve of a given perimeter that encloses the largest area
- Shape of a string hanging from two points under gravity
- Path of light travelling through an inhomogenous curve

Euler's Equation

- The principle of
  - To minimize P is to solve P' $= 0$
- Rather than solving the integral it is an advantage to consider the differential equation.
- The differential equation is called Euler Equation.
- We will derive it shortly

Consider for a minute

- Suppose $f : \mathcal{R}^n \rightarrow \mathcal{R}$ what does it mean for x^* to be a local extremum of f?

Consider for a minute

- Suppose $f : \mathcal{R}^n \to \mathcal{R}$ what does it mean for x^* to be a local extremum of f?
  1. We must have $f(x) \geq f(x^*)$ for every x in some neighborhood
  2. A necessary condition $\nabla f(x^*) = 0$ i.e., that $\frac{\partial f}{\partial x_i} = 0$ for all i.
- For P the equivalent would be say
  1. $P : C^2(\mathcal{R}^n) \to \mathcal{R}$ and
  2. $f \to P(f)$
- what does it mean for f^* to be an extremum of P?

Optimal functional?

- What would be conditional for a functional?
  1. We need $P(f) \geq P(f^*)$ for every functional close to f^*
     - So what is a neighborhood of a function?
  2. Need a generalized gradient

$$P(f^* + \delta f) \approx P(f^*)$$

- Still very hand wavy

Simplest problem

- Lets start with a simple problem
- Minimize $J(y) = \int_{x_0}^{x_1} F(x, y, y')dx$ with $y, F \in C^2$
- Suppose y^* minimizes J it would then be true
  1. In a neighborhood of y^* then $J(y) \geq J(y^*)$
  2. $\delta J = 0$ for a variation δy is

$$\delta J(y^*) = J(y^* + \delta y) - J(y^*)$$

- What are the necessary conditions for this to be valid

Neighborhood Evaluation

- Lets start by showing optimality in a neighborhood
- Let $y \in C^2[x_0, x_1]$ such that $y(x_0) = y(x_1) = 0$
- Let $\epsilon \in \mathcal{R}$ be a value
- Lets consider a one-parameter family of functions

$$y(x) = y^*(x) + \epsilon y(x)$$

- Where y^* is the (unknown) optimal function
- Define $\Phi : \mathcal{R} \rightarrow \mathcal{R}$ by

$$\Phi(\epsilon) = \int_{x_0}^{x_1} F(x, y, y')dx$$

- If $|\epsilon|$ is small enough then all variants of $y^* + \epsilon y$ lie in a small neighborhood of y^*, therefore Φ attains a local minimum at $\epsilon = 0$
- Thus it must be true that $\Phi'(0) = 0$

So what is Φ'?

- We know that

$$\Phi(\epsilon) = \int_{x_0}^{x_1} F(x, y, y')dx$$

- So it must be true that

$$\Phi'(\epsilon) = \frac{d}{d\epsilon} \int_{x_0}^{x_1} F(x, y, y')dx$$

- Given that we have a C^2 domain we can reverse the order of integration and differentiation, so that

$$\Phi'(\epsilon) = \int_{x_0}^{x_1} \frac{d}{d\epsilon} F(x, y, y')dx$$

So what is Φ'?

- We know that

$$\Phi(\epsilon) = \int_{x_0}^{x_1} F(x, y, y')dx$$

- So it must be true that

$$\Phi'(\epsilon) = \frac{d}{d\epsilon}\int_{x_0}^{x_1} F(x, y, y')dx$$

- Given that we have a C^2 domain we can reverse the order of integration and differentiation, so that

$$\Phi'(\epsilon) = \int_{x_0}^{x_1} \frac{d}{d\epsilon}F(x, y, y')dx$$

  or

$$\Phi'(\epsilon) = \int_{x_0}^{x_1} \left(\frac{\partial}{\partial y}F(x, y^* + \epsilon y, {y^*}' + \epsilon y')y + \frac{\partial}{\partial y'}F(x, y^* + \epsilon y, {y^*}' + \epsilon y')y'\right) dx$$

- We know that

$$\Phi'(0) = 0 = \int_{x_0}^{x_1} \left(\frac{\partial}{\partial y}F(x, y^*, {y^*}')y + \frac{\partial}{\partial y'}F(x, y^*, {y^*}')y'\right) dx$$

Still more Φ'

- We can write this more compactly

$$\Phi'(0) = \int_{x_0}^{x_1} (F_y y + F_{y'} y')\, dx$$

- Using integration by parts we get

$$\begin{aligned} \int_{x_0}^{x_1} F_{y'} y' dx &= F_{y'} y\Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} y \tfrac{d}{dx} F_{y'} dx \\ &= -\int_{x_0}^{x_1} y \tfrac{d}{dx} F_{y'} dx \end{aligned}$$

with this we can rewrite

$$\Phi'(0) = \int_{x_0}^{x_1} \left[F_y - \frac{d}{dx} F_{y'} \right] y dx = 0$$

as this has to apply for any function y it must be true that

$$F_y - \frac{d}{dx} F_{y'} = 0 \text{ on } [x_0, x_1]$$

- This is called Euler's Equation

Side comment

- The Euler Equation is essentially a “directional derivative” in the direction of y
- Going back to earlier - δJ is finding a function y^* where J is stationary.
- We are only considering the basics here.

Shortest path problem

- Remember the initial question of shortest path?
- Recall:

$$L(y) = \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx$$

with $y_0 = y(x_0)$ and $y_1 = y(x_1)$

- So $F(x, y, y') = \sqrt{1 + y'^2}$

$$F_y = 0 \text{ and } F_{y'} = \frac{y'}{\sqrt{1 + y'^2}}$$

- Euler's Equation reduces to

$$\frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}} = 0$$

The shortest path?

- So

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

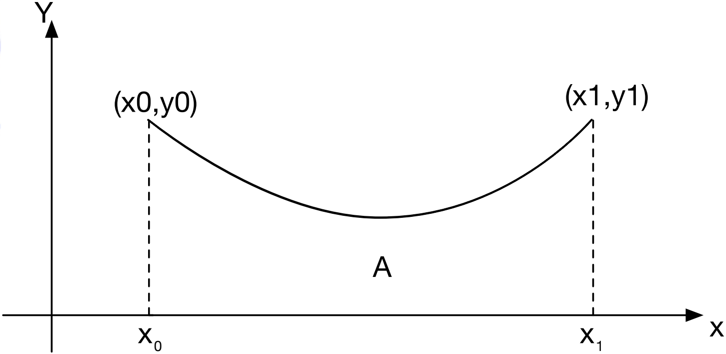
- we can rewrite

$$\begin{aligned} y'^2 &= c^2(1+y'^2) \\ y' &= \pm\frac{c}{\sqrt{1-c^2}} = m \text{ just a constant} \\ y' &= m \\ y &= mx+b \end{aligned}$$

surprise it is the equation for a straight line!

How about contrained optimization?

- Supposed we are supposed to find shortest curve with a fixed area below?



- The area is given to be A and we have end-points?

Constrained optimization

- Our objective is then to optimize

$$
\begin{aligned}
L(y) &= \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx \\
A &= \int_{x_0}^{x_1} y dx
\end{aligned}
$$

- where the second term is our constraint
- An instance of a general class of problems called isoperimetric problems

Isoperimetric problems

- The simplified formulation is

Minimize	$J(y) = \int_{x_0}^{x_1} F(x, y, y')dx$
Subject to	$K(y) = c$
where	$K(y) = \int_{x_0}^{x_1} G(x, y, y')dx$

Constrained Optimization (cont.)

- We can use a combination of variational techniques and Lagrange multipliers to solve such problems
- We can define two functions

$$
\begin{aligned}
\Phi(\epsilon_1, \epsilon_2) &= \int_{x_0}^{x_1} F(x, y^* + \epsilon_1 y + \epsilon_2 \xi, {y^*}' + \epsilon_1 y' + \epsilon_2 \xi')dx \\
\Psi(\epsilon_1, \epsilon_2) &= \int_{x_0}^{x_1} G(x, y^* + \epsilon_1 y + \epsilon_2 \xi, {y^*}' + \epsilon_1 y' + \epsilon_2 \xi')dx
\end{aligned}
$$

- Here y^* is the unknown function and y and ξ are two C^2 functions that vanish at the end-points
- So we want to minimize Φ subject to the constraint Ψ. We know there is a local minimum at $\epsilon_1 = \epsilon_2 = 0$

Constrained Optimization (Cont.)

- Using a Lagrange approach we can form the function

$$E(\epsilon_1, \epsilon_2, \lambda) = \Phi(\epsilon_1, \epsilon_2) + \lambda(\Psi(\epsilon_1, \epsilon_2) - c)$$

- At the local minimum - $\nabla E = 0$
- In other words there is a λ_0 such that

$$\frac{\partial}{\epsilon_1} E(0, 0, \lambda_0) = 0 \qquad \frac{\partial}{\epsilon_2} E(0, 0, \lambda_0) = 0$$
$$\frac{\partial}{\lambda} E(0, 0, \lambda_0) = 0$$

Constrained Optimization - let's compute

- Interchanging differentiation and integration we get

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(F_y y + F_{y'} y' + \lambda_0 G_y y + \lambda_0 G_{y'} y'\right) dx$$

Constrained Optimization - let's compute

- Interchanging differentiation and integration we get

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(F_y y + F_{y'} y' + \lambda_0 G_y y + \lambda_0 G_{y'} y' \right) dx$$

- We can do integration by parts and as y vanishes at end-points we see that

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] \right) y dx$$

Constrained Optimization - let's compute

- Interchanging differentiation and integration we get

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(F_y y + F_{y'} y' + \lambda_0 G_y y + \lambda_0 G_{y'} y' \right) dx$$

- We can do integration by parts and as y vanishes at end-points we see that

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] \right) y dx$$

- Similarly:

$$\frac{\partial}{\partial \epsilon_2} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] \right) \xi dx$$

Constrained Optimization - let's compute

- Interchanging differentiation and integration we get

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(F_y y + F_{y'} y' + \lambda_0 G_y y + \lambda_0 G_{y'} y' \right) dx$$

- We can do integration by parts and as y vanishes at end-points we see that

$$\frac{\partial}{\partial \epsilon_1} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] \right) y dx$$

- Similarly:

$$\frac{\partial}{\partial \epsilon_2} E(0,0,\lambda_0) = \int_{x_0}^{x_1} \left(\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] \right) \xi dx$$

- As before we can conclude

$$\left[F_y - \frac{d}{dx} F_{y'} \right] + \lambda_0 \left[G_y - \frac{d}{dx} G_{y'} \right] = 0$$

Back to our example

- So we can utilize

$$
\begin{array}{cc}
F(x,y,y') = \sqrt{1+y'^2} & G(x,y,y') = y \\
F_y = 0 & G_y = 1 \\
F_{y'} = \frac{y'}{\sqrt{1+y'^2}} & G_{y'} = 0
\end{array}
$$

- We want to satisfy the differential equation

$$
-\frac{d}{dx}\frac{y'}{\sqrt{1+y'^2}} + \lambda_0 = 0
$$

- Or

$$
\begin{array}{rcl}
\frac{y'}{\sqrt{1+y'^2}} & = & \lambda_0 x + c \\
\frac{y'^2}{1+y'^2} & = & (\lambda_0 x + c)^2 \\
y'^2 & = & \frac{(\lambda_0 x+c)^2}{1-(\lambda_0 x+c)^2} \\
y' & = & \pm\frac{\lambda_0 x+c}{\sqrt{1-(\lambda_0 x+c)^2}}
\end{array}
$$

Example (cont.)

- We can do the integration

$$
\begin{aligned}
y(x) &= \pm \int \frac{\lambda_0 x + c}{\sqrt{1-(\lambda_0 x + c)^2}} \\
& \quad \text{substite } u = \lambda_0 x + c \text{ and } du = \lambda_0 dx \\
&= \pm \int \frac{u}{\sqrt{1-u^2}} du = \pm \left[-\sqrt{1-u^2} + k \right] \\
&= \pm \left[-\frac{1}{\lambda} \sqrt{1-(\lambda_0 x + c)^2} - \frac{k}{\lambda_0} \right]
\end{aligned}
$$

- This can be rewritten to

$$
\left(y \pm \frac{k}{\lambda_0} \right)^2 + \left(x + \frac{c}{\lambda_0} \right)^2 = \frac{1}{\lambda_0}
$$

- That is a circle arc!

Extensions

- For multiple variable you can formulate it similar to the simple case
- Ex: Shortest path in a multiple dimensional space
- Ex: Light ray tracing through non-homogenous media
- You would extend Euler's Equation to have more terms

Summary

- Merely broached calculus of variation
- Powerful tool for optimization and derivation of analytical models
- Models for airplane wings, elastic membranes
- Important to consider it part of your toolbox