Variational Methods and Optimal Control

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods and Optimal Control

A/Prof. Matthew Roughan

July 26, 2012

lecture01 (37 pages)

- **Introduction**

  What is the point of this course?

  Example 1: The money pit. (slide 5)

  Example 2: Catenary: shape of a hanging wire (defined). (slide 6)

  Example 3: Brachystochrone: curve of quickest descent (defined). (slide 7)

  Example 4: Dido's problem (defined) (slide 13)

- **Revision**

  Extrema of functions of one variable.

  "Nothing takes place in the world whose meaning is not that of some maximum or minimum."
  L.Euler

- **Extra bits**

  Some notation and definitions

lecture02 (29 pages)

- **Revision, part ii**

  Extrema of functions of multiple variables. Taylor's theorem and the chain rule in N-D. Hessians and classification of extrema.

  Example 1: $f(x_1,x_2) = x_1^2 - x_2^2 + x_1^3$ (slide 6)

  Example 2: $f(x_1,x_2) = r - 1/2r^2$, *where* $r^2 = x_1^2 + x_2^2$ (slide 7)

  Example 3: $f(x_1,x_2) = x_2^3 - 3x_1^2 x_2$ (slide 8)

- **Extra bits**

lecture03 (36 pages)

- **Revision, part iii**

  Constrained extrema and Lagrange multipliers.

  Example 1: Rectangle of fixed perimeter with maximal area. (slide 5)

  Example 2: Largest area rectangle inscribed in a circle. (slide 8)

  Example 3: Largest area rectangle inscribed in a ellipse. (slide 10)

  Example 4: Maximize $f(x_1,x_2,x_3) = x_1x_2x_3$ *subject to* $x_1x_2 + x_1x_3 + x_2x_3 = 1$, *and* $x_1 + x_2 + x_3 = 3$ (slide 12)

  Example 5: Inequality constraint: largest area rectangle inscribed in a unit circle. (slide 15)

  Example 6: Maximize $3x$ *subject to* $x \leq 10$. (slide 17)

- **Revision, part iv**

  Vector space notation.

- **Functionals**

  In CoV we are not maximizing the value of a simple function, we want to find a "curve" that maximizes (or minimizes) a **functional**. Think of functionals as a generalization of a function, except we can think of it as an ∞-dimensional max. problem.

  Example 1: Catenary: the shape of a hanging wire (the functional). (slide 30)

  Example 2: Brachystochrone: curve of quickest descent (the functional). (slide 31)

  Example 3: Bent elastic beam (defined). (slide 32)

  Example 4: Stimulated plant growth (defined). (slide 34)

  Example 5: Parking a car (defined). (slide 35)

lecture04 (40 pages)

- **Fixed-end point problems**

  We'll start with the simplest functional maximization problem, and show how to solve by finding the first variation and deriving the Euler-Lagrange equations:

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

  Example 1: Catenary: a hanging wire (without a length constraint). (slide 3)

  Example 2: Geodesics in a plane. (slide 22)

- **Special cases**

  Now that we know the Euler-Lagrange (E-L) equations, we can use them directly, but there are some special cases for which the equations simplify, and make our life easier:

  - f depends only on y'
  - f has no explicit dependence on x (autonomous case)
  - f has no explicit dependence on y
  - $f = A(x,y)y' + B(x,y)$ (degenerate case)

- **Special case 1**

  When f depends only on y' the E-L equations simplify to

$$\frac{\partial f}{\partial y'} = const$$

  An example of this is calculating geodesics in the plane (which we all know are straight lines).

  Example 1: $F\{y\} = \int_0^1 \alpha y'^4 - \beta y'^2, dx$ (slide 28)

  Example 2: Fermat's principle of geometrical optics. (slide 29)

  Example 3: Fermat's principle and Snell's law (slide 39)

lecture05 (25 pages)

- **Special case 2**

  When f has no dependence on x we call this an autonomous problem, and we can replace the E-L equations with

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y') = const$$

  We will see H again later – it often turns out to be a conserved quantity like energy, and so arises naturally in computing the shape of a catenary.

  Example 1: Catenary: the shape of a hanging wire. (slide 7)

lecture06 (32 pages)

- **Special case 2: autonomous problems continued**

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y') = const$$

  We will see H again later – it often turns out to be a conserved quantity like energy, and so arises naturally in computing the shape of the brachystochrone.

  Example 1: Brachystochrone: curve of quickest descent. (slide 5)

  Example 2: Newton's aerodynamical problem (part i). (slide 15)

lecture07 (22 pages)

- **Special case 3**

  When f has no explicit dependence on y the E-L equations simplify to give

$$\frac{\partial f}{\partial y'} = const$$

  An example where we might use this is in calculating geodesics on non-planar objects such as the sphere.

  Example 1: Geodesics on the unit sphere. (slide 6)

  Example 2: Geodesics on other surfaces in $\mathbb{R}^3$. (slide 20)

lecture08 (26 pages)

- **Invariance of the E-L equations**

  We side-track here to note that extremals found using the E-L equations don't depend on the coordinate system! This can be very useful – a change of co-ordinates can often simplify a problem dramatically.

  Example 1: Polar (circular) coordinates. (slide 11)

- **Special case 4**

  When $f = A(x,y)y' + B(x,y)$ we call this a degenerate case, because the E-L equations reduce to

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

  but we can't necessarily solve these, and when they are true, the functional's value only depends on the end-points, not the actual shape of the curve.

  Example 1: $f(x,y,y') = (x^2 + 3y^2)y' + 2xy$ (slide 24)

lecture09 (23 pages)

- **Extensions**

  Now we consider extensions to the simple E-L equations presented so far:

- when f includes higher-order derivatives, e.g., $f(x,y,y',y'')$, e.g., the shape of a bent bar.
- when there are several dependent variables (i.e., y is a vector), e.g., calculating a particles trajectory.
- when there are several independent variables (i.e., x is a vector), e.g. calculating extremal surface.

- **Extension 1: higher-order derivatives**

  When f includes higher-order derivatives then the E-L equations can be extended, e.g., if the function includes a y'' term, i.e., $f(x,y,y',y'')$, then

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 0$$

  but now we now need extra edge conditions. A simple example we will consider is the shape of a bent bar.

  Example 1: $F\{y\} = \int_0^1 (1+y''^2)\,dx$. (slide 12)

  Example 2: $F\{y\} = \int_0^{\pi/2} \left(y''^2 - y^2 + x^2\right)\,dx$. (slide 14)

  Example 3: Bent elastic beam. (slide 18)

lecture10 (42 pages)

- **Extension 2: several dependent variables**

  When there are several dependent variables, i.e., y is a vector, then the E-L equations generalize to give one DE per dependent variable. A simple example is when we calculate the trajectory of a particle in 3D. This section introduces a number of physics ideas/principles: potentials, Lagrangians, Hamilton's principle, Newton's laws of motion, and conservations laws.

  Example 1: $F\{\mathbf{q}\} = \int_0^1 \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2\right)\,dt$. (slide 12)

  Example 2: Movement of a particle. (slide 17)

  Example 3: Simple pendulum. (slide 21)

  Example 4: Kepler's problem of planetary motion. (slide 22)

  Example 5: Brachystochrone in 3D. (slide 29)

lecture11 (35 pages)

- **Extension 3: several independent variables**

  When there are several independent variables, e.g., (x,y) and the extremal we wish to find represents, for instance, a surface $z(x,y)$, and f is a function $f(x,y,z(x,y),z_x,z_y)$, then the E-L equation generalizes to give

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

  Example 1: Laplace's equation. (slide 14)

  Example 2: A vibrating string and the wave equation. (slide 16)

  Example 3: Minimal area surface (Plateau's problem). (slide 23)

lecture12 (27 pages)

- **Numerical Solutions**

  The E-L equations may be hard to solve

  Natural response is to find numerical methods

  - Numerical solution of E-L DE

    - ∗ we won't consider these here (see other courses)
  - Euler's finite difference method
  - Ritz (Rayleigh-Ritz)
    - ∗ In 2D: Kantorovich's method

- **Euler's finite difference method**

  We can approximate our function (and hence the integral) onto a finite grid. In this case, the problem reduces to a standard multivariable maximization (or minimization) problem, and we find the solution by setting the derivatives to zero. In the limit as the grid gets finer, this approximates the E-L equations.

  Example 1: Euler's FDM on a simple problem (slide 6)

- **Ritz's method**

  In Ritz's method (called Kantorovich's methods where there is more than one independent variable), we approximate our functions (the extremal in particular) using a family of simple functions. Again we can reduce the problem into a standard multivariable maximization problem, but now we seek coefficients for our approximation.

  Example 1: $F\{y\} = \int_0^1 \left[y'^2/2 + y^2/2 - y\right] dx$ (slide 21)

lecture13 (22 pages)

- **Numerical solutions continued**

  Ritz applied to the catenary gives additional insights and Kantorovich's method generalizes Ritz to 2D functions..

  Example 1: Catenary approximation using Ritz's method. (slide 3)

  Example 2: $F\{z(x,y)\} = \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z\right) dx\, dy.$ (slide 15)

lecture14 (37 pages)

- **Constraints**

  We now include additional constraints into the problems:

  - Integral constraints of the form
    $$\int g(x,y,y')dx = const$$
    e.g., the Isoperimetric problem.
  - Holonomic constraints, e.g., $g(x,y) = 0$
  - Non-holonomic constraints, e.g., $g(x,y,y') = 0$
  - We won't consider inequality constraints until later.

- **Integral Constraints**

  Integral constraints are of the form
  $$\int g(x,y,y')dx = const$$
  The standard example of such a problem is Dido's problem, leading to us referring to such constraints as **isoperimetric**. We solve these by introducing the functional analogy of a Lagrange multiplier.

  Example 1: Dido's problem: simplified (slide 8)

  Example 2: Catenary of fixed length (slide 24)

  Example 3: Rigid extremals (slide 33)

lecture15 (19 pages)

- **Isoperimetric constraints (continued)**

  We solve the more general case of Dido's problem: a general shape, without a coast, so that the perimeter must be parametrically described.

  Example 1: Dido's problem: traditional form (slide 4)

lecture16 (47 pages)

- **Holonomic Constraints**

  Constraints of the form $g(x,y) = 0$, or $g(t,\mathbf{q}) = 0$, which don't involve derivatives of $y(x)$ or $\mathbf{q}$ can also be handled using a Lagrange multiplier technique, but we have to introduce a Lagrange multiplier function $\lambda(x)$, not just a single value λ. Effectively we introduce one Lagrange multiplier at each point where the constraint is enforced.

  Example 1: Geodesics on the sphere. (slide 12)

- **Non-Holonomic Constraints**

  Constraints of the form $g(x,y,y') = 0$, or $g(t,\mathbf{q},\dot{\mathbf{q}}) = 0$, which involve derivatives. They are effectively additional DEs which we need to solve, but we can once again use Lagrange multipliers.

  Example 1: A simple solution for $F\{y\} = \int_a^b f(x,y,y',y'')\,dx$. (slide 19)

  Example 2: Newton's aerodynamical problem (part ii). (slide 24)

- **Intro to Optimal Control**

  One way we see non-holonomic constraints is when we consider control problems. In these we seek to control a system described by a DE (the constraint) subject to some input which we can control (optimize).

  Example 1: Stimulated plant growth. (slide 42)

lecture17 (38 pages)

- **Non-fixed end point problems**

  What happens when we don't fix the end-points of an extremal? In this case **natural boundary conditions** are automatically introduced, and these can allow us to solve the E-L equations.

- **Free end points: Fixed x, Free y and/or y'**

  First we'll consider what happens when we allow y or y' to vary at the end-points, but we still keep the x values of the end-points fixed at x_0 and x_1.

  Example 1: Freely supported elastic beam. (slide 5)

  Example 2: Elastic beam fixed at one end point. (slide 6)

- **Intro to Optimal Control (part II)**

  Often in optimal control problems we may specify the initial state, but not the final state. However, there may be a cost associated with the final state, and we include this in the functional to be minimized (or maximized). We call this a terminal cost.

  Example 1: Stimulated plant growth with a free end-point. (slide 30)

lecture18 (33 pages)

- **Free end points: Free x, y and y'**

  We now allow x to vary as well, although we may apply some condition on the relationship between x and y, for instance that the end point must lie on a curve. In these cases we often rename our extremals, and call them **transversals**.

  Example 1: Shortest-path between two curves. (slide 6)

  Example 2: Orbit transfer problem. (slide 7)

Example 3: Fixed x, ffree y (slide 26)

Example 4: Free x, fixed y (slide 26)

Example 5: $F\{y\} = \int_0^{x_1} 1 + y'^2\, dx$ *subject to* $y(0) = 1$ *and* $y(x_1) = L > 1$, *but with* x_1 *unspecified.* (slide 27)

Example 6: $F\{\mathbf{q}\} = \int_0^1 \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2\right) dt$. (slide 30)

lecture19 (18 pages)

- **Transversals**

  When we consider an extremal joining a curve to a point (or two curves) then we often call the extremal a transversal. The free-end-point condition simplifies in many such cases, for instance, in many situations we look for a transversal that joins the proscribed curve at right angles.

  Example 1: Shortest path from the origin to a curve. (slide 7)

  Example 2: Generalized shortest path between two curves. (slide 12)

  Example 3: Shape of a wire hanging between two curves. (slide 14)

  Example 4: Curve of fastest descent from a point to line (Brachystochrone to a wall). (slide 16)

  Example 5: Shortest-path from a point to a surface. (slide 18)

lecture20 (32 pages)

- **Broken Extremals**

  Until now we have required that extremal curves have at least two well-defined derivatives. Obviously this is not always true (see for instance Snell's law). In this lecture we consider the alternatives.

  Example 1: $F\{y\} = \int_{-1}^{1} y^2(1-y')^2\, dx$ (slide 4)

  Example 2: Newton's aerodynamical problem (part iii). (slide 26)

lecture21 (38 pages)

- **Inequality Constraints and Optimal Control**

  Earlier we didn't consider inequalities as constraints, but these are needed particularly in control. For instance, often there is a maximum force we can apply to an object. The resulting extremals either (i) satisfy the E-L equations, or (ii) lie along the edge of the constraint. We also get boundary conditions between these two types of regions.

  Example 1: Parking a car. (slide 3)

  Example 2: Shortest-path avoiding an obstacle. (slide 13)

lecture22 (26 pages)

- **More Optimal Control Examples**

  First we'll cover a bit more terminology, and then some examples primarily focussed on planned growth strategies in economics.

  Example 1: Dynamic production control. (slide 13)

  Example 2: Optimal economic growth. (slide 18)

lecture23 (35 pages)

- **More Optimal Control Examples**

  An aerospace example: a rocket launch profile.

  Example 1: Rocket launch. (slide 3)

lecture24 (26 pages)

- **Hamilton's formulation**

  We've seen the Hamiltonian H earlier on, but haven't explored its full power. Firstly, using H can often result in a simpler approach than solving the E-L equations, e.g., where f has no dependence on x, or where there is more than one dependent variable. More importantly though, this formulation can lead to an understanding of how symmetries in the problem of interest lead to conservation laws. Finally, we will use the Hamiltonian in the Pontryagin Maximum Principle, which we will study soon.

  Example 1: Simple pendulum: Hamilton's formulation. (slide 11)

  Example 2: Simple pendulum: Hamilton-Jacobi approach. (slide 25)

lecture25 (29 pages)

- **Conservation Laws**

  One of the more exciting things we can derive relates to fundamental physics laws: conservation of energy, momentum, and angular momentum. We can now derive all of these from an underlying principle: Noether's theorem.

  Example 1: Invariance under translations in x is $\equiv$ conservation of H (energy). (slide 22)

  Example 2: Invariance under translations in y is $\equiv$ conservation of p (momentum). (slide 23)

  Example 3: Invariance under rotations is $\equiv$ conservation of angular momentum. (slide 26)

lecture26 (37 pages)

- **Pontryagin Maximum Principle**

  Modern optimal control theory often starts from the PMP. It is a simple, concise condition for an optimal control.

  Example 1: Stimulated plant growth (revisited). (slide 10)

  Example 2: Stimulated plant growth with a free end-point (revisited). (slide 16)

  Example 3: Optimal treatment of gout (slide 20)

  Example 4: Lunar lander. (slide 25)

lecture27 (40 pages)

- **Bang-Bang controllers and other related issues**

  Here we consider more generally what conditions result in a bang-bang controller.

  Example 1: Optimal fish harvesting (slide 6)

  Example 2: Time minimization problem. (slide 11)

  Example 3: Singular control example. (slide 33)

lecture28 (15 pages)

- **Feedback control systems**

  In all of our previous examples, we solve optimization problem "all at once", i.e., we plan the shape of the curve y to optimize the functional. However, sometimes, we need a control that reacts continuously to perturbations in a system. Such controllers typically utilize feedback.

  Example 1: Liquid level control. (slide 4)

lecture29 (28 pages)

- **Classification of extrema**

  We have so far typically ignored the issue of classification of extrema, but remember that for simple stationary points we need to look at higher derivatives to see if a stationary point is a maximum, minimum or point of inflection. We need an analogous process for extremal curves as well.

lecture30 (16 pages)

Variational Methods & Optimal Control

lecture 01

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 01

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Did you bring your duck?

Suddenly, Professor Liebowitz realizes he has come to the seminar without his duck.

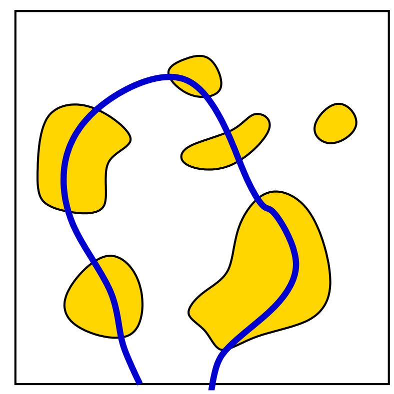
Larson, 1989

Introduction

What is the point of this course?

Motivation

- Imagine a field containing patches of gold.
- Collect the most gold
- We want to choose best **path**
- But the path length is limited.

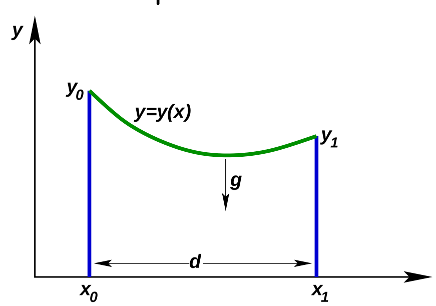


Gold example (part ii)

- The gold collected on the path is the **integral** of the gold at each point.
- The length of the path is **fixed**.
- We are maximizing an integral over a path for **all possible paths**.
- Maximizing a function of a function (a **functional**).

The catenary

Consider a thin, uniformly-heavy, flexible cable suspended from the top of two poles of height y_0 and y_1 spaced a distance d apart. What is the shape of the cable between the two poles?



What is the difference if the cable is coiled at the base of the poles and is free to move up and down via a pulley?

Brachystochrone problem

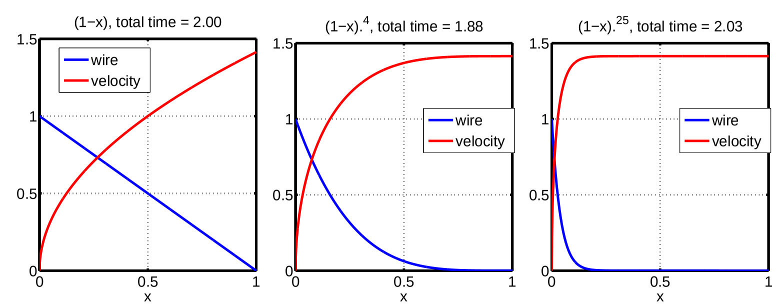
"Did Bernoulli sleep before he found the curves of quickest descent? ", Peter Parker, Spiderman II

Find the shape of a wire along which a bead, initially at rest, slides from one end to the other as quickly as possible under the influence of gravity.

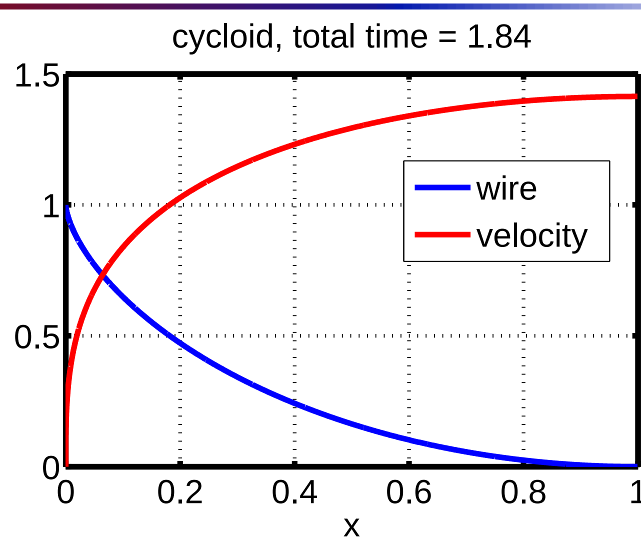
- endpoints are fixed
- motion is frictionless

Can think of as the "optimal slippery dip"

Brachystochrone problem



Brachystochrone solution

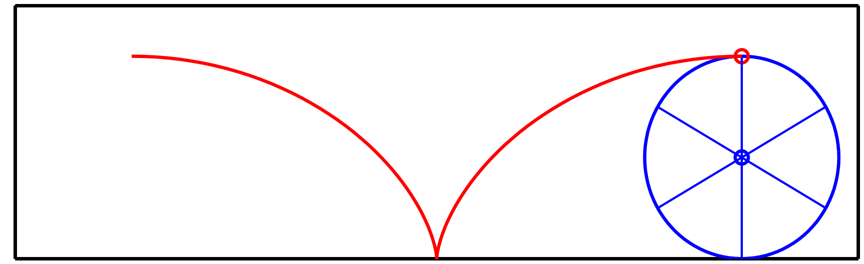


Brachystochrone history

- problem posed by Johann Bernoulli (1696)
- Newton, Liebnitz, Huygens, Bernoulli's
- Euler developed method to solve it that was generalizable
- Jacob first to solve?
- Johann, "Ah, I recognize the paws of a lion"
- Christiaan Huygens discovered cycloid property

  A bead sliding down a cycloid generated by a circle of radius ρ under gravity g reaches the bottom after $\pi\sqrt{\rho/g}$ regardless of where the bead starts. Hence **cycloid = isochrone**

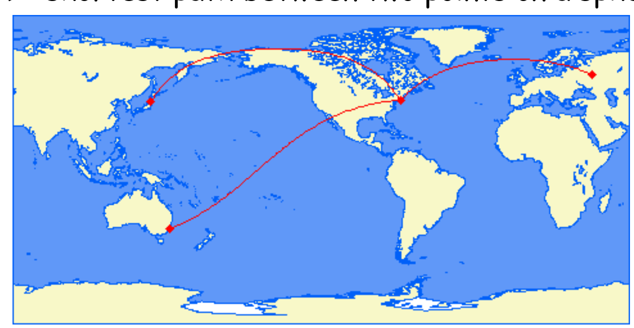
Cycloid generation



Geodesics

Geodesic = shortest path

- shortest path between two points on a plane
- shortest path between two points on a sphere



- shortest path on an arbitrary manifold on $\mathbb{R}^n$

Dido's problem

Isoperimetric problem: what shaped curve encompasses the largest area given a fixed perimeter.

- 200 B.C. proof by Zendorus (but flawed)
- Steiner proved that "if it exists" its a circle
- Weierstraβ proved using **Calculus of Variations**

Control problems

Control of systems is critical in modern life

- Mech.Eng: Design of active suspension
- Medicine: Drug delivery to minimize harmful side-effects
- Aerospace: optimize rocket thrust (to minimize fuel consumption)
- Economics: maximize utility of consumption (vs savings)
- Environment: optimal harvesting (say of fish)
- Minimizing cost of A/C

Optimal control is the best (cheapest, fastest, smoothest, ...) we can do.

Other examples

- Design of vehicle profile that minimizes drag
- Finding shapes of soap bubbles

Revision

Extrema of functions of one variable.

"Nothing takes place in the world whose meaning is not that of some maximum or minimum."
L.Euler

Revision

Calculus of variations is concerned with maximization (minimization)

We are going to maximize (minimize) functionals, not functions

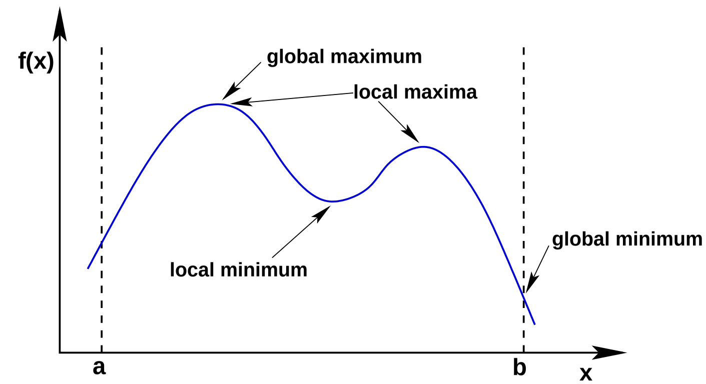
Let us first revise maximization (minimization) of function

Maxima and minima

Functions of one variable:

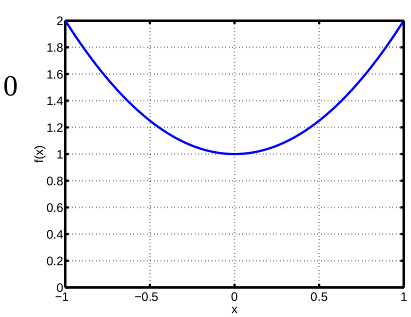
- Let $x \in [a,b]$ and $f(x) : [a,b] \to \mathbb{R}$
- If there is a point $x_{\min}$ such that $f(x_{\min}) \leq f(x)$ for all $x \in [a,b]$, then $x_{\min}$ is called a **global minima** of $f(x)$ in $[a,b]$.
- The set of points x such that $f(x) = f(x_{\min})$ is called the **minimal set**.
- If there is an interior point $x \in (a,b)$ such that there exists a $\delta > 0$ with $f(x) \leq f(\hat{x})$ for all $\hat{x} \in (x-\delta, x+\delta)$, then x is called a **local minimum** of $f(\cdot)$.
- similar definitions apply for maxima, note maxima of $f(x)$ are the minima of $-f(x)$

Maxima and minima: example 1



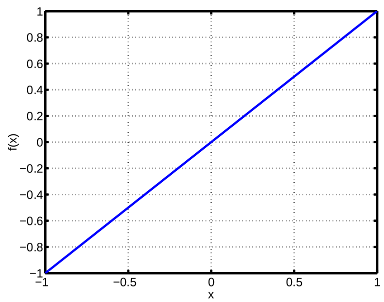
Maxima and minima: example 2

- $f(x) = 1 + x^2$ on $[-1,1]$
- global minimum at $x = 0$
- local minimum at $x = 0$
- maximal set $\{-1,1\}$



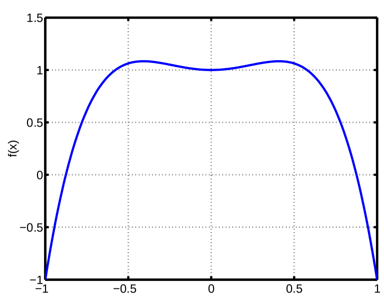
Maxima and minima: example 3

- $f(x) = x$ on $[-1, 1]$
- global minimum at $x = -1$
- not a local min. because not an interior point
- global maximum at $x = 1$



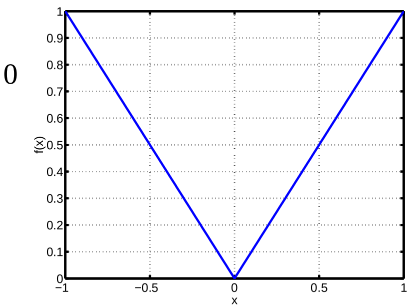
Maxima and minima: example 4

- $f(x) = 1 + x^2 - x^4$ on $[-1, 1]$
- global minimum at $x = -1$ and $x = 1$
- local minimum at $x = 0$.



Maxima and minima: example 5

- $f(x) = |x|$ on $[-1, 1]$
- global minimum at $x = 0$
- local minimum at $x = 0$



How to find maxima and minima

Theorem 1: Let $f(x) : [a, b] \to I\!R$ be differentiable in (a, b). If $f(\cdot)$ has a local extrema at x then

$$\frac{df}{dx} = f'(x) = 0$$

Proof: The derivative is given by

$$f'(x) = \lim_{\hat{x} \to x} \frac{f(\hat{x}) - f(x)}{\hat{x} - x}$$

Suppose x is a local minima, then $\exists \delta > 0$ such that $\hat{x} \in (x - \delta, x + \delta) \Rightarrow f(\hat{x}) > f(x)$, hence the numerator > 0. The denominator changes sign at $\hat{x} = x$. Differentiability implies the left and right hand limits exist and are equal, and hence $f'(x) = 0$.

Sufficient conditions

Theorem 2: Let $f(x) : [a,b] \to I\!R$ be twice differentiable in (a,b). Sufficient conditions for a local minimum at x are

$$f'(x) = 0 \quad \text{and} \quad f''(x) > 0$$

Proof: see following.

Some useful theorems

- **Mean Value Theorem:** Let $x_0 < x_1$, and $f(\cdot)$ be a continuous function in $[x_0, x_1]$, and differentiable in (x_0, x_1), then $\exists \xi \in (x_0, x_1)$ such that

$$f(x_1) = f(x_0) + (x_1 - x_0) f'(\xi)$$

- **Taylor's theorem:** Let $f(\cdot)$ be a function whose first n derivatives exist and are continuous in the interval $[x_0, x_1]$, and $f^{(n+1)}(x)$ exists for all $x \in (x_0, x_1)$, then $\exists \xi \in (x_0, x_1)$

$$f(x_1) = f(x_0) + (x_1 - x_0) f'(x_0) + \frac{(x_1 - x_0)^2}{2} f''(x_0) + \cdots + \frac{(x_1 - x_0)^n}{n!} f^{(n)}(x_0) + \frac{(x_1 - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

Sufficient conditions

Theorem 3: Let $f(x) : [a,b] \to I\!R$ have derivatives of all orders, then a necessary and sufficient condition for a local minima is that for some n

$$f'(x) = f''(x) = \cdots = f^{(2n-1)}(x) = 0 \quad \text{and} \quad f^{(2n)}(x) > 0$$

Proof: Taylor's theorem, where $\hat{x} - x = \varepsilon$

$$f(\hat{x}) = f(x) + \varepsilon f'(x) + \cdots + \frac{\varepsilon^{2n-1}}{(2n-1)!} f^{(2n-1)}(x) + \frac{\varepsilon^{2n}}{(2n)!} f^{(2n)}(x) + O(\varepsilon^{2n+1})$$

Then

$$\begin{aligned} f(\hat{x}) - f(x) &= \frac{\varepsilon^{2n}}{(2n)!} f^{(2n)}(x) + O(\varepsilon^{2n+1}) \\ &> 0 \quad \text{for small enough } \varepsilon \end{aligned}$$

Classifying extrema

Assume that $f'(x) = 0$

- local maxima $f''(x) < 0$
- local minima $f''(x) > 0$
- turning point $f''(x) = 0$, and $f^{(3)}(x) \neq 0$
- $+$ a lot of higher order conditions

Call all points with $f'(x) = 0$ the set of **stationary** points

Conclusion

We have looked at 1D local maxima and minima
We need to generalize this

- next lecture, to functions of N variables
- then, to functions of functions (∞ variables)

Extra bits

Some notation and definitions

Notation

- $[a,b]$ is the closed interval, i.e. the set $\{x \in \mathbb{R} | a \leq x \leq b\}$
- (a,b) is the open interval, i.e. the set $\{x \in \mathbb{R} | a < x < b\}$
- $(a,b]$ is the set $\{x \in \mathbb{R} | a < x \leq b\}$
- $f(x) : [a,b] \to \mathbb{R}$ denotes a function that maps the set $[a,b]$ to a real number.
- $\frac{d^n f}{dx^n} = f^{(n)}(x)$ denotes the nth derivative of $f(x)$.

Synonyms

- the global minimum is sometimes called a strong minimum
- a local minimum is sometimes called a weak minimum
- the local extrema are the collection of local minima and maxima
  We sometimes abuse notation to include stationary points in the set of extrema.

Useful Definitions: continuity

- a function $f(x)$ is **continuous** at x_0 iff the left and right limits at x_0 exist and are equal, i.e.,

$$\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x)$$

otherwise it is said to have a **discontinuity**.

- We say a function is continuous on an interval if it is continuous at every point inside the interval and the limits exist at the boundaries.
- A function is **piecewise continuous** on an interval if it has at most finite number of discontinuities.

Useful Definitions: differentiability

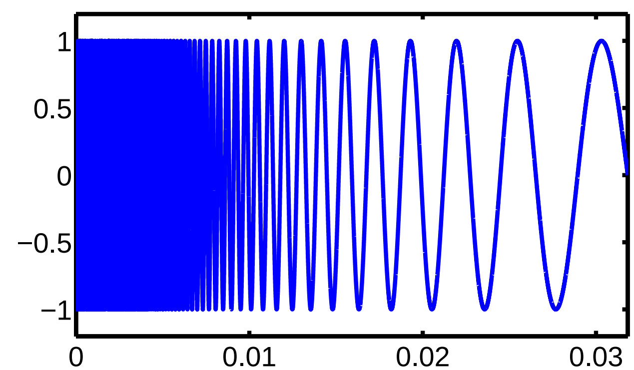
- A function is **differentiable** at x_0 if its derivative exists, and is continuous at x_0, i.e., the following limit exists and is the same from both directions

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$$

- We say a function is differentiable on an interval if it is differentiable at every point inside the interval and the limits exist at the boundaries.
- A function is **piecewise differentiable** if the derivative has at most a finite number of discontinuities.
- A function is **twice differentiable** if its second derivative exists and is continuous.

Useful Definitions

- We also eliminate from consideration functions whose derivative changes sign an infinite number of times in a finite interval.
  - e.g. $\sin(1/x)$



Notation

We define the **del** or **grad** operator by

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

So, given a scalar function $\phi(x,y,z)$, then $\nabla\phi$ is a vector function

$$\nabla\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right)$$

Given a vector function $\mathbf{f}(x,y,z) = (f_1, f_2, f_3)$ then we define the **div** operator $\operatorname{div}\mathbf{f} = \nabla\cdot\mathbf{f}$, e.g.

$$\nabla\cdot\mathbf{f} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\cdot(f_1, f_2, f_3) = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

Notation

We can also use del to define the **curl** operator using a cross-product $\mathrm{curl} = \mathrm{del}\times$, e.g.

$$\mathrm{curl}\,\mathbf{f} = \nabla \times \mathbf{f}$$

The **Laplacian operator**, or del-squared operator of a scalar function (of (x,y,z)) is defined by

$$\nabla^2\phi = \nabla \cdot (\nabla\phi) = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$

Variational Methods & Optimal Control

lecture 02

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 02

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

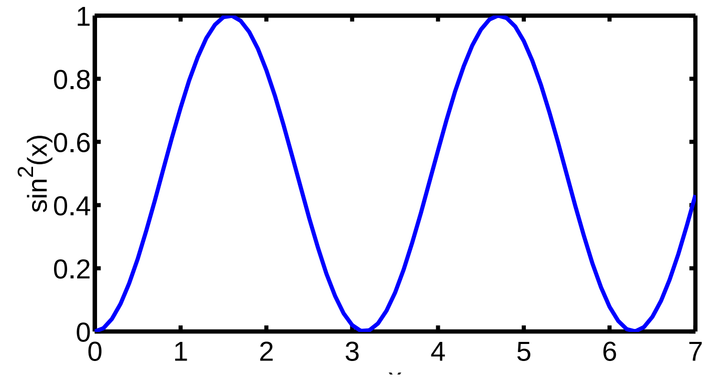
July 26, 2012

Revision, part ii

Extrema of functions of multiple variables. Taylor's theorem and the chain rule in N-D. Hessians and classification of extrema.

Extrema of functions of one variable

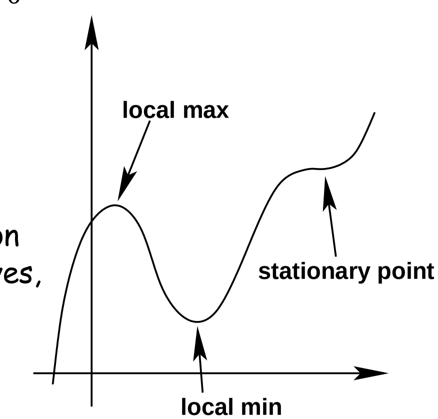
Local extrema have $f'(x) = 0$
includes maxima, minima, and stationary points of inflection



Classification of extrema

Local extrema have $f'(x) = 0$

- $f''(x) > 0$ local minima
- $f''(x) < 0$ local maxima
- $f''(x) = 0$ it might be a stationary point of inflection, depending on higher order derivatives, e.g. x^4.



Functions of n variables

- Let Ω be a closed region of $\mathbb{R}^n$, i.e. $\Omega \subset \mathbb{R}^n$
- Let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \Omega$
- Let $f : \Omega \to \mathbb{R}$
- A local minima if $f(\cdot)$ is point $\mathbf{x}$ such that there exists $\delta > 0$ where

$$f(\hat{\mathbf{x}}) \geq f(\mathbf{x})$$

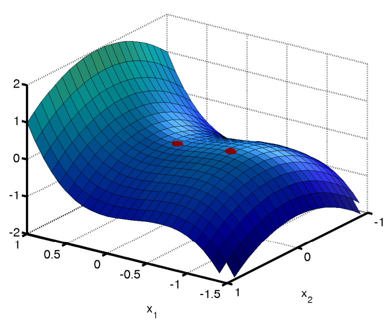
  for any $\hat{\mathbf{x}} \in B(\mathbf{x}; \delta)$.
- A global minima of $f(\cdot)$ on Ω is point $\mathbf{x}$ such that

$$f(\hat{\mathbf{x}}) \geq f(\mathbf{x})$$

  for any $\hat{\mathbf{x}} \in \Omega$.

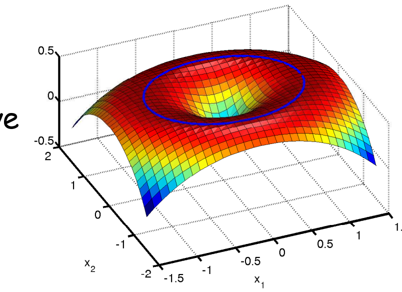
2D example 1

- $f(x_1, x_2) = x_1^2 - x_2^2 + x_1^3$
- local maximum at $(-2/3, 0)$
- saddle point at $(0, 0)$



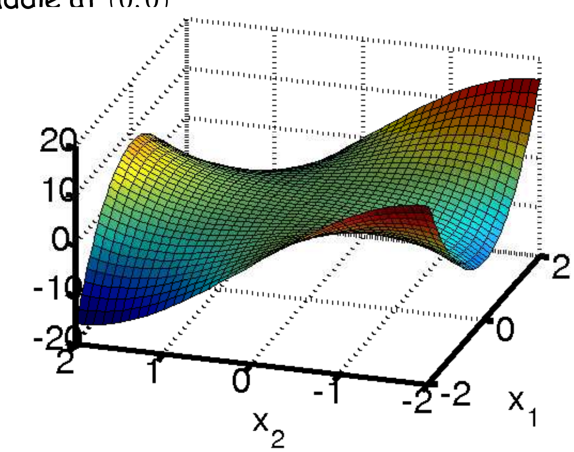
2D example 2

- $f(x_1, x_2) = r - 1/2r^2$, where $r = \sqrt{x_1^2 + x_2^2}$
- global maxima on curve $r = 1$
- local minima at $r = 0$



2D example 3

- $f(x_1, x_2) = x_2^3 - 3x_1^2 x_2$
- Monkey saddle at $(0, 0)$



The Chain rule

The derivative of a function $f(x_1, x_2)$ along a line described parametrically by $(x_1(t), x_2(t))$

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt}$$

Another way to think of this is as the directional derivative formed from the dot product of grad and the direction of the line, e.g.,

$$\begin{aligned}\frac{df}{dt} &= \nabla f \cdot \frac{d\mathbf{x}}{dt} \\ &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) \cdot \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}\right)\end{aligned}$$

Chain Rule for more variables

The chain rule (for a function of more than one variable) $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$, where we want to find the derivative of a function $f(\mathbf{x})$ along a line described parametrically by $(x_1(t), x_2(t), \ldots, x_n(t))$ then we take

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

or alternatively

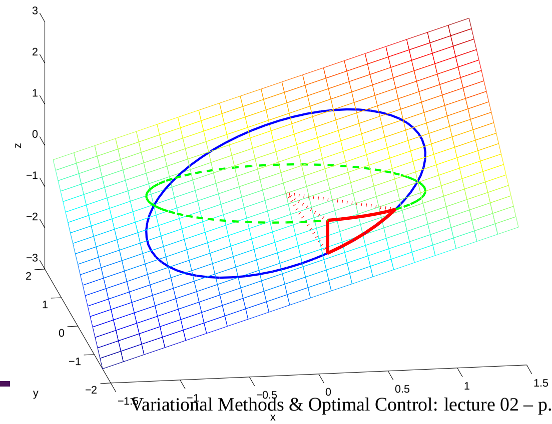
$$\begin{aligned}\frac{df}{dt} &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) \cdot \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \ldots, \frac{dx_n}{dt}\right) \\ &= \nabla f \cdot \frac{d\mathbf{x}}{dt}\end{aligned}$$

A graphical example

For a function of two variables $f(x,y)$ we get

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

$$\begin{aligned} f(x,y) &= x + y \\ x &= \cos t \\ y &= \sin t \end{aligned}$$



Chain Rule Derivation

By the definition

$$\frac{df}{dt} = \lim_{\varepsilon \to 0} \frac{f(x(t+\varepsilon), y(t+\varepsilon)) - f(x(t), y(t))}{\varepsilon}$$

But note that from Taylor's theorem

$$x(t+\varepsilon) = x(t) + \varepsilon x'(t) + O(\varepsilon^2).$$

As we consider the limit as $\varepsilon \to 0$ we may ignore the $O(\varepsilon^2)$ term, to get

$$\frac{df}{dt} = \lim_{\varepsilon \to 0} \frac{f(x(t) + \varepsilon x'(t), y(t) + \varepsilon y'(t)) - f(x(t), y(t))}{\varepsilon}$$

Chain Rule Derivation

$$\begin{aligned}\frac{df}{dt} &= \lim_{\varepsilon\to 0}\frac{f(x(t)+\varepsilon x'(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t))}{\varepsilon}\\ &= \lim_{\varepsilon\to 0}\frac{f(x(t)+\varepsilon x'(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t)+\varepsilon y'(t))}{\varepsilon}\\ &\quad +\lim_{\varepsilon\to 0}\frac{f(x(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t))}{\varepsilon}\\ &= x'(t)\lim_{\varepsilon\to 0}\frac{f(x(t)+\varepsilon x'(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t)+\varepsilon y'(t))}{\varepsilon x'(t)}\\ &\quad +y'(t)\lim_{\varepsilon\to 0}\frac{f(x(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t))}{\varepsilon y'(t)}\end{aligned}$$

Chain Rule Derivation

$$\begin{aligned}\frac{df}{dt} &= x'(t)\lim_{\varepsilon\to 0}\frac{f(x(t)+\varepsilon x'(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t)+\varepsilon y'(t))}{\varepsilon x'(t)}\\ &\quad +y'(t)\lim_{\varepsilon\to 0}\frac{f(x(t),y(t)+\varepsilon y'(t)) - f(x(t),y(t))}{\varepsilon y'(t)}\\ &= x'(t)\lim_{\varepsilon_x\to 0}\frac{f(x(t)+\varepsilon_x,y(t)+\varepsilon y'(t)) - f(x(t),y(t)+\varepsilon y'(t))}{\varepsilon_x}\\ &\quad +y'(t)\lim_{\varepsilon_y\to 0}\frac{f(x(t),y(t)+\varepsilon_y) - f(x(t),y(t))}{\varepsilon_y}\\ &= x'(t)\frac{\partial f}{\partial x}+y'(t)\frac{\partial f}{\partial y}\end{aligned}$$

which is the chain rule!

Chain Rule special case

When we only have one variable, we simple want to calculate the derivative of a function f of another function x, e.g.

$$\frac{d}{dt}f(x(t)) = \frac{df}{dx}\frac{dx}{dt}$$

Another way of writing this is

$$\frac{d}{dt}f(x(t)) = f'[x(t)]\,x'(t),$$

which is the form you learnt in 1st year.

Taylor's theorem in 2D

$$\begin{aligned}f(x_1+\delta x_1,x_2+\delta x_2) &= f(x_1,x_2)+\delta x_1\frac{\partial f}{\partial x_1}+\delta x_2\frac{\partial f}{\partial x_2}\\ &\quad +\frac{1}{2}\left[\delta x_1^2\frac{\partial^2 f}{\partial x_1^2}+2\delta x_1\delta x_2\frac{\partial^2 f}{\partial x_1\partial x_2}+\delta x_2^2\frac{\partial^2 f}{\partial x_2^2}\right]+\cdots\end{aligned}$$

Write $(\delta x_1,\delta x_2)=\varepsilon\times(\eta_1,\eta_2)$

$$\begin{aligned}f(\mathbf{x}+\varepsilon\eta) &= f(\mathbf{x})+\varepsilon\left(\eta_1\frac{\partial f}{\partial x_1}+\eta_2\frac{\partial f}{\partial x_2}\right)\\ &\quad +\frac{\varepsilon^2}{2}\left[\eta_1^2\frac{\partial^2 f}{\partial x_1^2}+2\eta_1\eta_2\frac{\partial^2 f}{\partial x_1\partial x_2}+\eta_2^2\frac{\partial^2 f}{\partial x_2^2}\right]+O(\varepsilon^3)\end{aligned}$$

Taylor's theorem in N-D

$$f(\mathbf{x}+\delta\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{n} \delta x_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i x_j}\delta x_i \delta x_j + O(\delta\mathbf{x}^3)$$

$$f(\mathbf{x}+\delta\mathbf{x}) = f(\mathbf{x}) + \delta\mathbf{x}^T \nabla f(\mathbf{x}) + \frac{1}{2}\delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} + O(\delta\mathbf{x}^3)$$

Where $H(\mathbf{x})$ is the Hessian matrix

$$H(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Maxima of N variables

If a smooth function $f(\mathbf{x})$ has a local extrema at $\mathbf{x}$ then

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)^T = \mathbf{0}$$

A sufficient condition for the extrema $\mathbf{x}$ to be a local minimum is for the quadratic form

$$Q(\delta x_1, \ldots, \delta x_n) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\delta x_i \delta x_j$$

to be strictly positive definite.

Quadratic forms

A quadratic form

$$Q(\mathbf{x}) = \sum_{i,j} a_{ij} x_i x_j = \mathbf{x}^T A\mathbf{x}$$

is said to be positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$.

A quadratic form is positive definite iff every eigenvalue of A is greater than zero.

A quadratic form is positive definite if all the principal minors in the top-left corner of A are positive, in other words

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \cdots$$

Notes on maxima and minima

- maxima of $f(x)$ are minima of $-f(x)$.
- haven't said anything about non-differentiable functions
- if continuous in the interval, must achieve maximum (minimum) in the interval

Calculus of variations

- We are not maximizing the value of a function.
- We are maximizing a **functional** — a function of a function
- Can think of it as an ∞-dimensional max. problem.
  - can choose between different functions
  - function sits in ∞-dimensional vector space
- This might take some effort.

Functionals

A **Functional** maps an element of a vector space (e.g. a space containing functions) to a real number, e.g. $F : S \to \mathbb{R}$.

Example Functionals

$$
\begin{aligned}
F\{y(x)\} &= |y(0)| \\
F\{y(x)\} &= \max_x \{y(x)\} \\
F\{y(x)\} &= \left.\frac{dy}{dx}\right|_{x=1} \\
F\{y(x)\} &= y(0) + y(1) \\
F\{y(x)\} &= \sum_{n=0}^{N} a_n y(n)
\end{aligned}
$$

Integral functionals

- Previous functionals not very interesting.
- Easy to find $y(x)$ which minimizes these.
- Integral functionals are more interesting.
- Example integral functionals

$$
\begin{aligned}
F\{y\} &= \int_a^b y(x)dx \\
F\{y\} &= \int_a^b f(x)y(x)dx \\
F\{y\} &= \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx
\end{aligned}
$$

Some simple cases

$$
\begin{aligned}
F\{y\} &= \int_{a(\varepsilon)}^{b(\varepsilon)} y(x,\varepsilon)dx \\
\frac{dF}{d\varepsilon} &= y(b,\varepsilon)\frac{db}{d\varepsilon} - y(a,\varepsilon)\frac{da}{d\varepsilon} + \int_{a(\varepsilon)}^{b(\varepsilon)} \frac{\partial y(x,\varepsilon)}{\partial \varepsilon} dx
\end{aligned}
$$

If a and b are fixed then

$$
\begin{aligned}
\frac{da}{d\varepsilon} &= 0 \\
\frac{db}{d\varepsilon} &= 0
\end{aligned}
$$

and so the derivative of the integral becomes the integral of the derivative.

Crude Brachystochrone

Brachystochrone involves the functional

$$F\{y\} = \int_{x_0}^{x_1} \sqrt{\frac{1+y'^2}{y}}\,dx$$

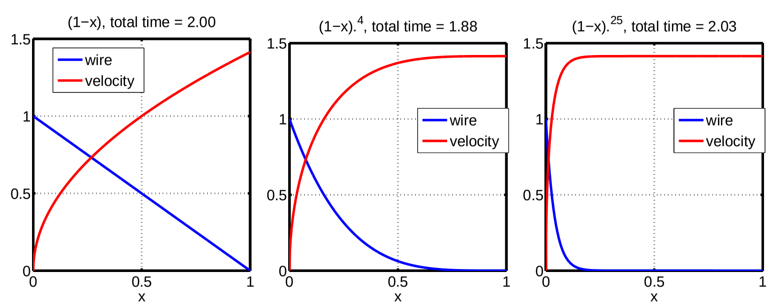
Let us guess that the brachystochrone takes the form

$$y(x,\varepsilon) = (1-x)^{\varepsilon}$$

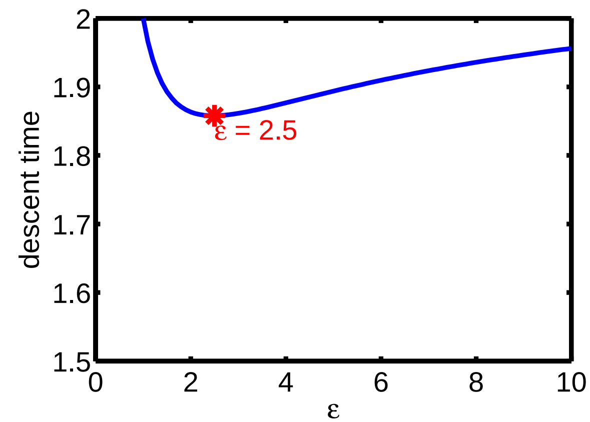
We could calculate the derivative WRT ε as above and compute the stationary points by finding

$$\frac{dF}{d\varepsilon} = 0$$

Crude Brachystochrone



Crude Brachystochrone



- but what if the family of curves doesn't contain the maximum?

Extra bits

Notation

- $f(\mathbf{x}) : S \to \mathbb{R}$ denotes a function that maps the set $S \subset \mathbb{R}^n$ to a real number.
- $\frac{\partial^n f}{\partial x_i^n}$ denotes the nth partial derivative of $f(\mathbf{x})$, with respect to x_i.
- the ε-neighborhood under the Euclidean norm is $B(\mathbf{x};\varepsilon) = \{\hat{\mathbf{x}} \in \mathbb{R}^n \mid \|\hat{\mathbf{x}} - \mathbf{x}\|_2 < \varepsilon\}$
- The Euclidean norm in $\mathbb{R}^n$ is $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$
- $F\{y\}$ denotes a functional of the function $y(x)$.

Variational Methods & Optimal Control

lecture 03

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 03

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Revision, part iii

Constrained extrema and Lagrange multipliers.

Constrained maxima and minima

Problem: find the minimum (or maximum) of $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ subject the the constraints

$$g_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m < n$$

The conditions define a subset of $\mathbf{x} \in \mathbb{R}^n$ called a manifold.
Solution requires **Lagrange Multipliers**. Minimize (or maximize) a new function (of $m+n$ variables)

$$h(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}),$$

where λ_i are the undetermined Lagrange multipliers.

Why Lagrange multipliers?

Maximize $f(\mathbf{x})$ subject to $g(\mathbf{x}) = 0$

$$h(\mathbf{x}) = f(\mathbf{x}) + \lambda g(\mathbf{x}).$$

So $\partial h/\partial x_i = 0$ implies that $\partial f/\partial x_i = -\lambda \partial g/\partial x_i$
Assume $\mathbf{x}$ is an extremal, which satisfies the constraint, consider all of the $\mathbf{x} + \delta\mathbf{x}$ in the neighborhood of $\mathbf{x}$ that also satisfy the constraint (i.e. $g(\mathbf{x}+\delta\mathbf{x}) = g(\mathbf{x}) = 0$), we also know from Taylor's theorem that

$$g(\mathbf{x}+\delta\mathbf{x}) = g(\mathbf{x}) + \delta\mathbf{x}^T \nabla g + O(\delta\mathbf{x}^2)$$

which implies that for small $\delta\mathbf{x}$

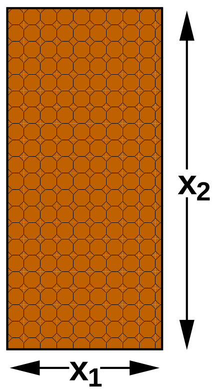
$$\delta\mathbf{x}^T \nabla g = 0$$

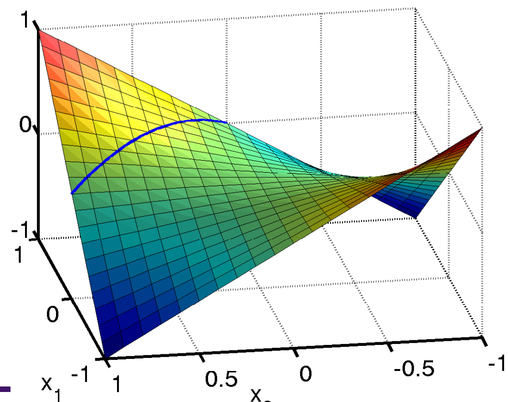
If we take $\partial f/\partial x_i = -\lambda \partial g/\partial x_i$ then

$$\delta\mathbf{x}^T \nabla f = 0$$

Constrained maxima example 1

Find the rectangle with fixed perimeter, and max. area.
E.G. the maximum of $f(x_1,x_2) = x_1 x_2$ subject to $x_1 + x_2 = 1$, $x_1, x_2 > 0$.





Constrained maxima ex. 1, solution

Maximize $h(x_1,x_2,\lambda) = x_1 x_2 + \lambda(x_1 + x_2 - 1)$
Set partial derivatives to be zero

$$\frac{\partial h}{\partial x_1} = \frac{\partial h}{\partial x_2} = \frac{\partial h}{\partial \lambda} = 0$$

$$\begin{aligned}
\frac{\partial h}{\partial x_1} &= x_2 + \lambda &= 0 \\
\frac{\partial h}{\partial x_2} &= x_1 + \lambda &= 0 \\
\frac{\partial h}{\partial \lambda} &= x_1 + x_2 - 1 &= 0
\end{aligned}$$

Solution $x_1 = x_2 = 1/2$, $\lambda = -1/2$.

Constrained maxima ex. 1, a bit more

Maximize $h(x_1,x_2,\lambda) = x_1 x_2 + \lambda(x_1 + x_2 - 1)$

$$\begin{bmatrix} \frac{\partial^2 h}{\partial x_1^2} & \frac{\partial^2 h}{\partial x_2 \partial x_1} \\ \frac{\partial^2 h}{\partial x_1 \partial x_2} & \frac{\partial^2 h}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

This is not positive definite!
However, note that $x_1 + x_2 = 1$, so the only possible perturbation vectors have the form $(\delta x, -\delta x)^T$.

$$(\delta x, -\delta x) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \delta x \\ -\delta x \end{bmatrix} = -2\delta x^2 < 0$$

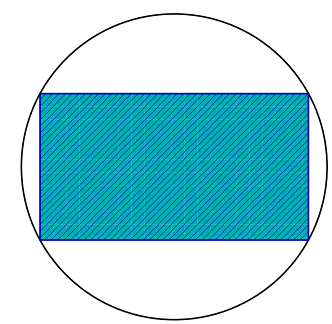
Hence, given the constraints on $\delta \mathbf{x}$, for all **possible** $\delta \mathbf{x}$, $f(\mathbf{x} + \delta \mathbf{x}) < f(\mathbf{x})$, and we have a local maximum.

Constrained maxima example 2

Largest area rectangle inscribed in a circle diameter 1.
Maximize $f(x_1,x_2) = x_1 x_2$ subject to $x_1^2 + x_2^2 = 1$, $x_1, x_2 > 0$.

$$h = x_1 x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

$$\begin{aligned}
\frac{\partial h}{\partial x_1} &= x_2 + 2\lambda x_1 \\
\frac{\partial h}{\partial x_2} &= x_1 + 2\lambda x_2 \\
\frac{\partial h}{\partial \lambda} &= x_1^2 + x_2^2 - 1
\end{aligned}$$



Constrained maxima ex. 2, solution

Subtract $2\lambda\times$ (1) from (2) and we get

$$x_1(1-4\lambda^2)=0$$

So $\lambda=\pm1/2$. To satisfy $x_1,x_2>0$, $\lambda=-1/2$, and hence $x_1=x_2$. To satisfy the constraint

$$x_1=x_2=1/\sqrt{2}.$$

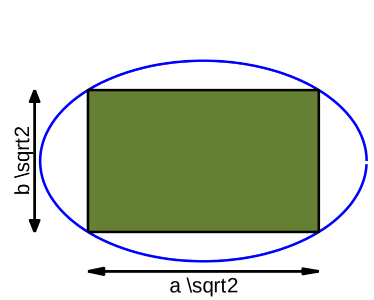
Solution is a square.

Constrained maxima example 3

Largest area rectangle inscribed in an ellipse.
Maximize $f(x,y)=xy$ subject to $x^2/a^2+y^2/b^2=1$, $x,y>0$.

$$h=xy+\lambda(x^2/a^2+y^2/b^2-1)$$

$$\begin{aligned}\frac{\partial h}{\partial x}&=y+2\lambda x/a^2\\ \frac{\partial h}{\partial y}&=x+2\lambda y/b^2\\ \frac{\partial h}{\partial \lambda}&=x^2/b^2+y^2/a^2-1\end{aligned}$$



Constrained maxima ex. 3, solution

Subtract $2\lambda/a^2\times$ (5) from (4) and we get

$$y\left(1-4\frac{\lambda^2}{a^2b^2}\right)=0$$

So $\lambda=\pm ab/2$. To satisfy $x,y>0$, $\lambda=-ab/2$, and hence $x=(a/b)y$. To satisfy the constraint

$$x=a/\sqrt{2},\quad y=b/\sqrt{2}.$$

Solution is now a rectangle.

Constrained maxima example 4

Maximize $f(x_1,x_2,x_3)=x_1x_2x_3$ subject to
$x_1x_2+x_1x_3+x_2x_3=1$, and $x_1+x_2+x_3=3$
$h=x_1x_2x_3+\lambda(x_1x_2+x_1x_3+x_2x_3-1)+\mu(x_1+x_2+x_3-3)$

$$\begin{aligned}\frac{\partial h}{\partial x_1}&=x_2x_3+\lambda(x_2+x_3)+\mu&&=0\\ \frac{\partial h}{\partial x_2}&=x_1x_3+\lambda(x_1+x_3)+\mu&&=0\\ \frac{\partial h}{\partial x_3}&=x_1x_2+\lambda(x_1+x_2)+\mu&&=0\\ \frac{\partial h}{\partial \lambda}&=x_1x_2+x_1x_3+x_2x_3-1&&=0\\ \frac{\partial h}{\partial \mu}&=x_1+x_2+x_3-3&&=0\end{aligned}$$

Constrained maxima and minima

Problem: find the minimum (or maximum) of $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ subject to the constraints

$$g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m < n$$

The conditions define a subset of $\mathbf{x} \in \mathbb{R}^n$ called a manifold.
Solution requires **Lagrange Multipliers**. Minimize (or maximize) a new function (of $m+n$ variables)

$$h(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}),$$

where λ_i are the undetermined Lagrange multipliers.

Inequality constraints

What if we have a constraint, say $g(x) \geq 0$?

For one constraint, its easy, we just find the max (min), and then check the constraint. If its satisfied, then we are OK, but if not, the global max (min) is on the boundary $g(x) = 0$, so now solve the constrained problem.

Example: inequality constraints

Example: the largest area square we can inscribe in a unit circle.

Earlier, we assumed that $x^2 + y^2 = 1$, but really the constraint says that $x^2 + y^2 \leq 1$.

However, the max area square (without the constraint) is clearly unbounded, and so doesn't satisfy the constraint, so we look for the square that lies on the boundary $g(x,y) = x^2 + y^2 - 1 = 0$, which we solve (as before) to get $x = y = 1/\sqrt{2}$.

Slack variables

An alternative is to introduce slack variables.

For each inequality constraint, rewrite as $g_i(x) \geq 0$. We introduce a slack variable α_i, and rewrite the constraint as

$$g_i(x) - \alpha_i^2 = 0$$

The α^2 term is automatically positive.

Then add in a standard Lagrange multiplier for this constraint, but note that in our maximization problem we now have the variables $\mathbf{x}$, α, and λ.

Example

Maximize $3x$ subject to $x \leq 10$.

Introduce slack variable α, and set the constraint to be $10 - x - \alpha^2 = 0$.

Now add a standard Lagrange multiplier, to maximize

$$h(x,\alpha,\lambda) = 3x + \lambda(10 - x - \alpha^2)$$

$$\begin{aligned}
\frac{\partial h}{\partial x} &= 3 - \lambda &= 0 &\quad \Rightarrow \lambda &= 3 \\
\frac{\partial h}{\partial \alpha} &= -2\lambda\alpha &= 0 &\quad \Rightarrow \alpha &= 0 \\
\frac{\partial h}{\partial \lambda} &= 10 - x - \alpha^2 &= 0 &\quad \Rightarrow x &= 10
\end{aligned}$$

Solution $(x,\alpha,\lambda) = (10,0,3)$

Revision, part iv

Vector space notation.

Vector spaces and function spaces

A **Vector Space** S is a collection of objects (vectors) $X, Y, \ldots$, along with two operators (addition, and scalar multiplication) that is

- closed under addition, e.g.

  For all $X, Y \in S$ we have $X + Y \in S$

- closed under scalar multiplication, e.g.

  For all $X \in S$, and $k \in \mathbb{R}$ we have $kX \in S$

Vector spaces and function spaces

The operators have to satisfy various properties

commutivity of addition	$X + Y = Y + X$
associativity of addition	$X + (Y + Z) = (Y + X) + Z$
additive identity	$\exists 0$ such that $X + 0 = X$
additive inverse	$\forall X, \exists -X$ such that $X + (-X) = 0$
distributivity	$\alpha(X + Y) = \alpha X + \alpha Y$
distributivity	$(\alpha + \beta)X = \alpha X + \beta X$
associativity of scalar mult.	$(\alpha\beta)X = \alpha(\beta X)$
multiplicative identity	$\exists 1$ such that $1.X = X$

Examples

- **example 1:** the set of vectors $\mathbf{x} \in \mathbb{R}^n$, with the standard vector addition and scalar multiplication.
- **example 2:** the set of all continuous functions on the interval $[x_0, x_1]$, denoted $C[x_0,x_1] = \{ f : [x_0,x_1] \to \mathbb{R} \mid f \text{ is continuous} \}$, with addition and scalar multiplication defined by

$$(f+g)(x) = f(x) + g(x), \qquad (\alpha f)(x) = \alpha f(x)$$

  for any $\alpha \in \mathbb{R}$, and $f, g \in C[x_0, x_1]$.
- **example 3:** The set of square integrable functions L^2 is the set of functions $f : \mathbb{R} \to \mathbb{R}$ for which $\int_{-\infty}^{\infty} f(x)^2 dx$ exists and is finite, with the same definition of sum and scalar product as for C.

Normed spaces

More structure is needed, in particular a way of measuring distances. A **norm** on a vector space S is a real-valued function(al) whose value at $x \in S$ is denoted $\|x\|$, and has the properties

$\|x\| \geq 0$
$\|x\| = 0$ iff $x = 0$
$\|\alpha x\| = \alpha \|x\|$
$\|x+y\| \leq \|x\| + \|y\|$ (the triangle inequality)

A vector space equipped with a norm is called a **normed vector space**.

Examples

- **example 1:** the vector space $\mathbb{R}^n$ can be equipped with the Euclidean norm defined by $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$. Alternatively we could use the norm defined by $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$
- **example 2:** the vector space $C[x_0,x_1]$ can be equipped with norms
  $\|f\|_\infty = \sup_{x \in [x_0,x_1]} |f(x)|$
  $\|f\|_1 = \int_{x_0}^{x_1} |f(x)| \, dx$
  $\|f\|_2 = \sqrt{\int_{x_0}^{x_1} f(x)^2 \, dx}$
- **example 3:** L^2 can be equipped with norm $\|f\|_2 = \sqrt{\int_{-\infty}^{\infty} f(x)^2 \, dx}$

Examples (cont.)

- **example 4:** Define $C^n[x_0,x_1]$ to be the set of functions that have at least n continuous derivatives on $[x_0,x_1]$. Note

$$C^n[x_0,x_1] \subset C^{n-1}[x_0,x_1] \subset \cdots \subset C^1[x_0,x_1] \subset C[x_0,x_1]$$

  $C^n[x_0,x_1]$ is a vector space, and $\|f\|_\infty$, $\|f\|_1$, and $\|f\|_2$ are all possible norms on this space. Other norms

$$\|f\|_{\infty,j} = \sum_{k=0}^{j} \sup_{x \in [x_0,x_1]} |f^{(k)}(x)|$$

  for $j \leq n$ on $C^n[x_0,x_1]$.

Norms

- denote a normed vector space $(S, \|\cdot\|)$.
- Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are said to be equivalent if there exists positive numbers α and β such that for all $x \in S$
$$\alpha\|x\|_a \le \|x\|_b \le \beta\|x\|_a$$
- In finite dimensional spaces all norms are equivalent, but not in infinite dimensional spaces.
- Norms define **distances** between elements of space
$$d(f,g) = \|f-g\|$$
- Distance defines the **ε-neighborhood** on $(S, \|\cdot\|)$
$$B(f,\varepsilon,\|\cdot\|) = \{g \in S \mid \|f-g\| \le \varepsilon\}$$

Inner products

An **inner product** is a function $\langle\cdot,\cdot\rangle : S \times S \to \mathbb{R}$, i.e. it maps two elements from a vector space S to a real number, such that for any $f,g,h \in S$ and $\alpha \in \mathbb{R}$.

$$\begin{aligned}
&\langle f,f\rangle \ge 0\\
&\langle f,f\rangle = 0 \text{ iff } f=0\\
&\langle f+g,h\rangle = \langle f,g\rangle + \langle f,h\rangle\\
&\langle f,g\rangle = \langle g,f\rangle\\
&\langle \alpha f,g\rangle = \alpha\langle f,g\rangle
\end{aligned}$$

A vector space with an inner product is called an **inner product space**.
We can use $\sqrt{\langle f,f\rangle}$ as a norm.

Functionals

In CoV we are not maximizing the value of a simple function, we want to find a "curve" that maximizes (or minimizes) a **functional**. Think of functionals as a generalization of a function, except we can think of it as an ∞-dimensional max. problem.

Functionals

- A **Functional** maps an element of a vector space to a real number, e.g. $F : S \to \mathbb{R}$.
- Typically in CoV S is a space of functions, e.g. $y(x)$
- Example Functionals

$$\begin{aligned}
F\{y(x)\} &= |y(0)|\\
F\{y(x)\} &= \max_x\{y(x)\}\\
F\{y(x)\} &= \left.\frac{dy}{dx}\right|_{x=1}\\
F\{y(x)\} &= y(0)+y(1)\\
F\{y(x)\} &= \sum_{n=0}^{N} a_n y(n)
\end{aligned}$$

Integral functionals

- Previous functionals not very interesting.
- Easy to find $y(x)$ which minimizes these.
- Integral functionals are more interesting.
- Example integral functionals

$$F\{y\} = \int_a^b y(x)dx$$

$$F\{y\} = \int_a^b f(x)y(x)dx$$

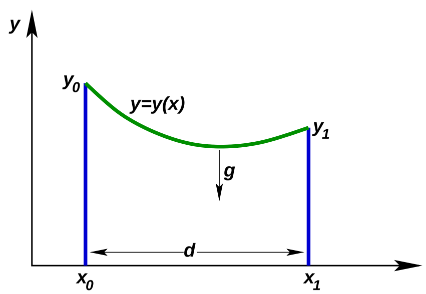
$$F\{y\} = \int_a^b \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx$$

Example: a hanging wire

The potential energy of the cable is

$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable



$$m = \text{mass per unit length}$$
$$g = \text{gravitational constant}$$

The system will seek to minimize $W_p\{y\}$

Example: the Brachystochrone

The time taken is

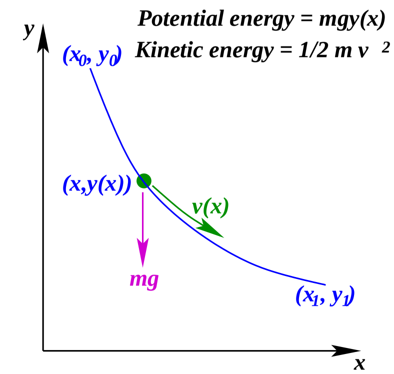
$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$

The energy of a body is the sum of potential and kinetic energy

$$E = \frac{1}{2}mv(x)^2 + mgy(x)$$

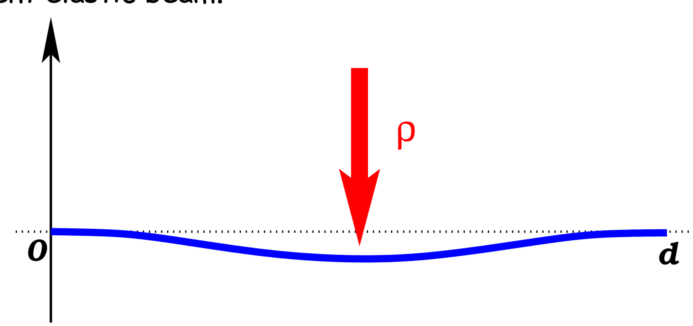
and a simple conservation law says this is constant, so

$$v(x) = \sqrt{\frac{2E}{m} - 2gy(x)}$$



Example: a bent beam

Bent elastic beam.



Two end-points are fixed, and clamped so that they are level, e.g. $y(0) = 0$, $y'(0) = 0$, and $y(d) = 0$ and $y'(d) = 0$. The load (per unit length) on the beam is given by a function $\rho(x)$.

Example: a bent beam

Let $y : [0,d] \to \mathbb{R}$ describe the shape of the beam, and $\rho : [0,d] \to \mathbb{R}$ be the load per unit length on the beam. For a bent elastic beam the potential energy from elastic forces is

$$V_1 = \frac{\kappa}{2} \int_0^d y''^2 \, dx, \qquad \kappa = \text{flexural rigidity}$$

The potential energy is

$$V_2 = -\int_0^d \rho(x) y(x) \, dx$$

Thus the total potential energy is

$$V = \int_0^d \frac{\kappa y''^2}{2} - \rho(x) y(x) \, dx$$

Control Example: plant growth

Stimulated plant growth problem:

- market gardener wants her plants' height x to reach 2 within a fixed window of time $[0,1]$
- can supplement natural growth with lights (at night)
- growth rate of the plants

$$\dot{x} = 1 + u$$

- cost of lights

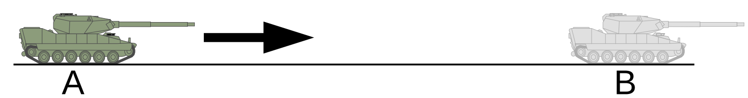
$$F\{u\} = \int_0^1 \frac{1}{2} u^2 \, dt$$

Minimize cost, subject to the constraints.

Example: parking a car

Classic problem: from *Craggs*, p.55

We want to drive a car/tank from point A to point B as quickly as possible, and at point B the car should be stationary.



Example: parking a car

We want to drive a car/tank from point A to point B as quickly as possible, and at point B the car should be stationary.

Newton's law

$$\text{force} = F = m\ddot{x}$$

Choose force F that minimizes the time subject to $\dot{x} = 0$ at $t = 0$ and $t = T$, where T is not specified, but rather given by

$$T\{F\} = \int_A^B dt$$

and it is this functional we wish to minimize.

Variational Methods & Optimal Control

lecture 04

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 04

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Fixed-end point problems

We'll start with the simplest functional maximization problem, and show how to solve by finding the first variation and deriving the Euler-Lagrange equations:

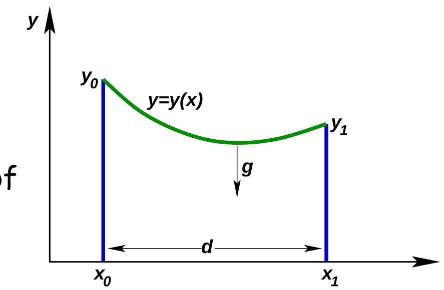
$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

The Catenary

The potential energy of the cable is

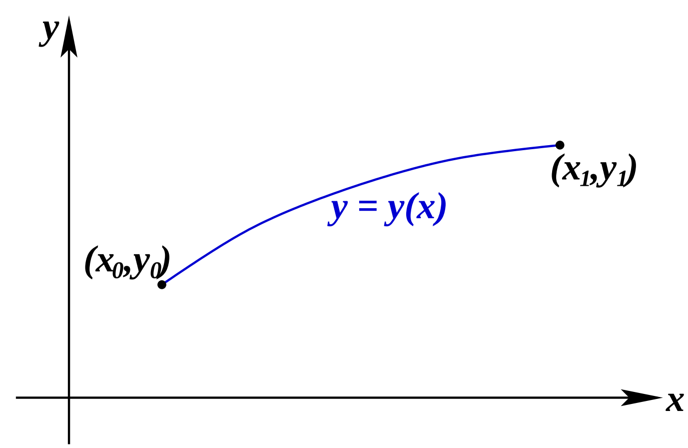
$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable



Catenary problem where we have pullies on top of each pylon, and a large amount of cable. Under appropriate conditions it will reach an equilibrium shape. The critical features of this problem are that the end-points are fixed but the length L of the cable is unconstrained.

Fixed end-point variational problem



Formulation

Define the functional $F : C^2[x_0, x_1] \to \mathbb{R}$

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

where f is assumed to be function with (at least) continuous second-order partial derivatives, WRT x, y, and y'.

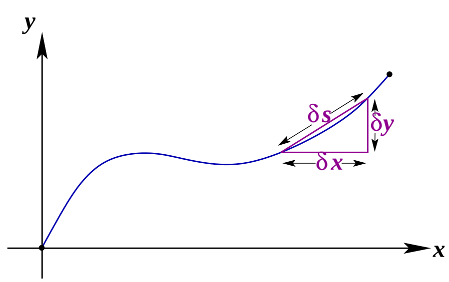
Problem: determine $y \in C^2[x_0, x_1]$ such that $y(x_0) = y_0$ and $y(x_1) = y_1$, such that F has a local extrema.

The Catenary

$$W_p\{y\} = \int_0^L mgy(s)ds$$

But I don't know how to evaluate this integral directly. Lets do a simple change of variables. The length of a line segment from (x,y) to $(x+\delta x, y+\delta y)$ is

$$\begin{aligned} \delta s &\simeq \sqrt{\delta x^2 + \delta y^2} \\ &= \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}\,\delta x \\ ds &= \sqrt{1 + y'^2}dx \end{aligned}$$



The Catenary

$$W_p\{y\} = \int_0^L mgy(s)ds$$

Change of variables $ds = \sqrt{1+y'^2}dx$. So the functional of interest (the potential energy) is

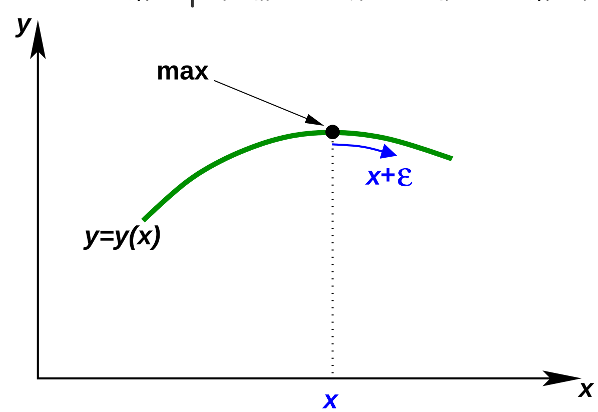
$$\begin{aligned} W_p\{y\} &= mg\int_{x_0}^{x_1} y\sqrt{1+y'^2}\,dx, \\ &= mg\int_{x_0}^{x_1} f(x,y,y')\,dx, \end{aligned}$$

where

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

How do we tackle these problems

look at small perturbations about the max/min.

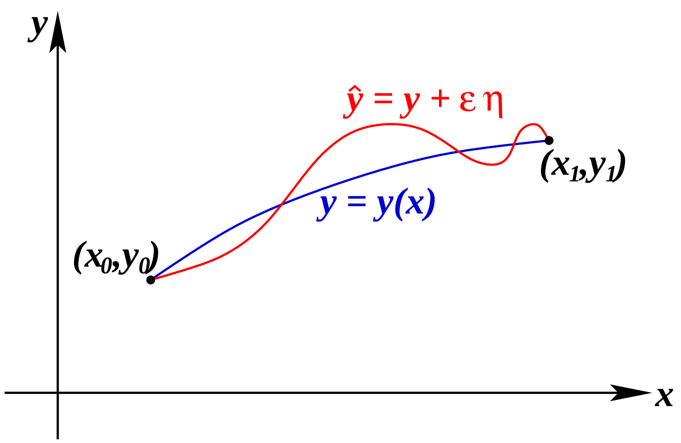


For a local maximum

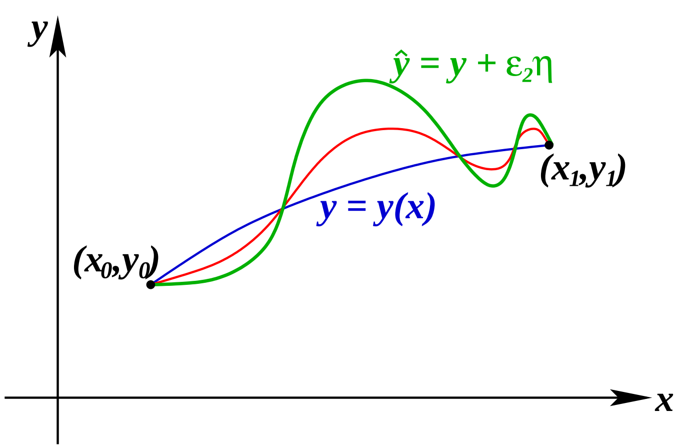
$$f(x+\varepsilon) \le f(x)$$

$\Rightarrow$ Conditions for extremals, i.e., $f'(x) = 0$

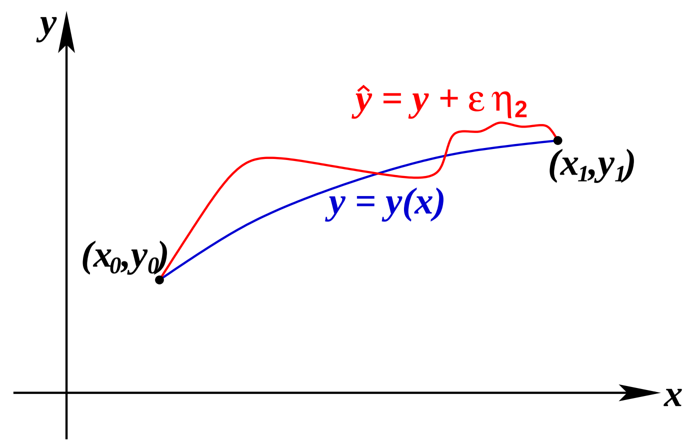
Perturbations of functions



Perturbations of functions



Perturbations of functions



The Functional of interest.

Define the functional $F : C^2[x_0,x_1] \to I\!R$

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

where f is assumed to be function with continuous second-order partial derivatives, WRT x, y, and y'.
Problem: determine $y \in C^2[x_0,x_1]$ such that $y(x_0) = y_0$ and $y(x_1) = y_1$, such that F has a local extrema.
The space of possible curves is

$$S = \{ y \in C^2[x_0,x_1] \,|\, y(x_0) = y_0, y(x_1) = y_1 \}$$

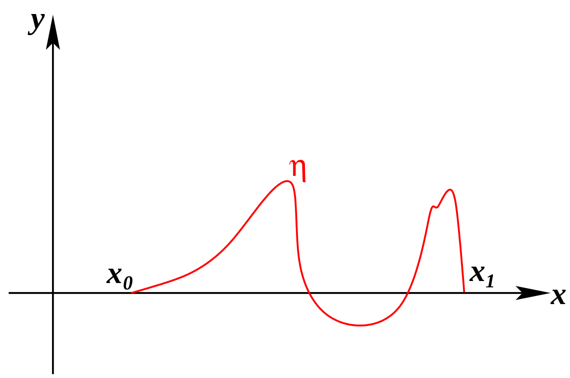
$\Rightarrow$ The vector space of allowable perturbations is

$$\mathcal{H} = \{ \eta \in C^2[x_0,x_1] \,|\, \eta(x_0) = 0, \eta(x_1) = 0 \}$$

Perturbation functions

The vector space of allowable perturbations is

$$\mathcal{H} = \{\eta \in C^2[x_0, x_1] \,|\, \eta(x_0) = 0, \eta(x_1) = 0\}$$



What to do

Regard f as a function of 3 independent variables: x, y, y'
Take $\hat{y}(x) = y(x) + \varepsilon\eta(x)$, where $y \in S$ and $\eta \in \mathcal{H}$.
Taylor's theorem (note x is kept constant below)

$$f(x,\hat{y},\hat{y}') = f(x,y,y') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right] + O(\varepsilon^2)$$

So

$$\begin{aligned} F\{\hat{y}\} - F\{y\} &= \int_{x_0}^{x_1} f(x,\hat{y},\hat{y}')dx - \int_{x_0}^{x_1} f(x,y,y')dx \\ &= \varepsilon\int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right]dx + O(\varepsilon^2) \end{aligned}$$

The first variation

For small ε the quantity

$$\delta F(\eta, y) = \lim_{\varepsilon\to 0}\frac{F\{y+\varepsilon\eta\} - F\{y\}}{\varepsilon} = \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right]dx$$

is called **the First Variation**.

For $F\{y\}$ to be a minimum, for small ε, $F\{\hat{y}\} \geq F\{y\}$, so the sign of $\delta F(\eta, y)$ is determined by ε.

As before, we can vary the sign of ε, so for $F\{y\}$ to be a local minima it must be the case that

$$\delta F(\eta, y) = 0, \qquad \forall \eta \in \mathcal{H}$$

Analogy to functions

This condition on the first variation is analogous to all partial derivatives being zero!

For a function of N variables to have a local extrema

$$\frac{\partial f}{\partial x_i} = 0, \qquad \forall i = 1, \ldots, n$$

For a functional to be an extrema

$$\delta F(\eta, y) = \frac{d}{d\varepsilon}F(y+\varepsilon\eta)\Big|_{\varepsilon=0} = 0, \qquad \forall \eta \in \mathcal{H}$$

Note now that we have to minimize over an infinite dimensional space $\mathcal{H}$, instead of $\mathbb{R}^n$.

Simplification

Integrate the second term by parts

$$\begin{aligned}\delta F(\eta,y) &= \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y}+\eta'\frac{\partial f}{\partial y'}\right]dx \\ &= \left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1}+\int_{x_0}^{x_1}\eta\left[\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]dx\end{aligned}$$

But note that by the problem definition $\eta \in \mathcal{H}$, and so $\eta(x_0)=\eta(x_1)=0$, and so the first term is zero.
The function inside the integral exists, and is continuous by our assumption that f has two continuous derivatives, so we may write

$$\delta F(\eta,y) = \langle \eta(x), E(x)\rangle^2 = \int_{x_0}^{x_1}\eta(x)E(x)dx = 0$$

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0)=y_0 \text{ and } y(x_1)=y_1 \right\},$$

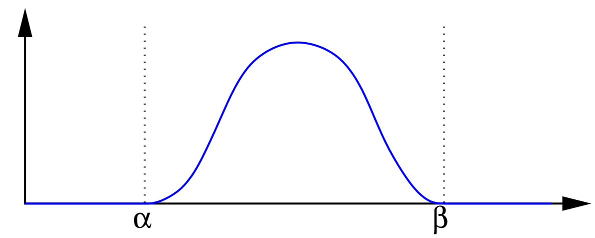
where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0,x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0} \Leftarrow \text{the Euler-Lagrange equation}$$

A useful lemma

Lemma 2.2.1: Let $\alpha,\beta \in \mathbb{R}$, such that $\alpha < \beta$. Then there is a function $\nu \in C^2(\mathbb{R})$, such that $\nu(x) > 0$ for all $x \in (\alpha,\beta)$ and $\nu(x) = 0$ otherwise.
Proof: by example

$$\nu(x) = \begin{cases} (x-\alpha)^3(\beta-x)^3, & \text{if } x \in (\alpha,\beta) \\ 0, & \text{otherwise.}\end{cases}$$



A second useful lemma

Lemma 2.2.2: Suppose $\langle \eta, g\rangle = 0$ for all $\eta \in \mathcal{H}$. If $g : [x_0,x_1] \to \mathbb{R}$ is a continuous function then $g(x) = 0$ for all $x \in [x_0,x_1]$.
Proof: Suppose $g(x) > 0$ for $x \in [\alpha,\beta]$. Choose ν as in Lemma 2.2.1.

$$\langle \nu(x), g(x)\rangle^2 = \int_{x_1}^{x_2}\nu(x)g(x)dx = \int_{\alpha}^{\beta}\nu(x)g(x)dx > 0$$

Hence a contradiction.
Similar proof for $g(x) < 0$.

Proof of Euler-Lagrange equation

As noted earlier, at an extremal the first variation

$$\delta F(\eta, y) = \langle \eta(x), E(x) \rangle^2 = \int_{x_0}^{x_1} \eta(x) E(x) dx = 0$$

for all $\eta(x) \in \mathcal{H}$. From Lemma 2.2.2, we can therefore state that

$$E(x) = \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] = 0,$$

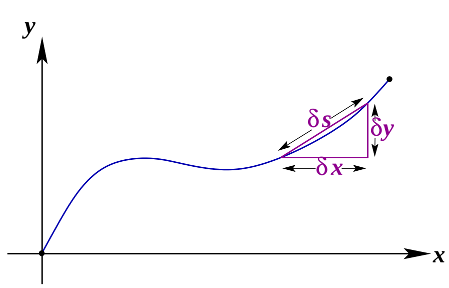
the Euler-Lagrange equation.□

Example: geodesics in a plane

Let $(x_0, y_0) = (0,0)$ and $(x_1, y_1) = (1,1)$, find the shortest path between these two points.

The length of a line segment from x to $x + \delta x$ is

$$\begin{aligned} \delta s &= \sqrt{\delta x^2 + \delta y^2} \\ &= \sqrt{1 + \left(\frac{\delta y}{\delta x} \right)^2} \delta x \\ ds &= \sqrt{1 + y'^2} dx \end{aligned}$$



So the total path length is $F\{y\} = \int_{x=0}^{x=1} ds = \int_0^1 \sqrt{1+y'^2}\, dx$

Example: geodesics in a plane

The arclength of a curve described by $y(x)$ will be

$$F\{y\} = \int_0^1 \sqrt{1 + y'^2} dx$$

Then

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

So $\frac{y'}{\sqrt{1+y'^2}}$ is a constant, implying $y' = const$. Hence $y(x) = c_1 x + c_2$, the equation of a straight line.

Special cases

Now that we know the Euler-Lagrange (E-L) equations, we can use them directly, but there are some special cases for which the equations simplify, and make our life easier:

- f depends only on y'
- f has no explicit dependence on x (autonomous case)
- f has no explicit dependence on y
- $f = A(x,y)y' + B(x,y)$ (degenerate case)

Special case 1

When f depends only on y' the E-L equations simplify to

$$\frac{\partial f}{\partial y'} = const$$

An example of this is calculating geodesics in the plane (which we all know are straight lines).

f depends only on y'

Geodesics in the plane are a special case of $f = f(y')$, with no explicit dependence on y. Apply the chain rule to the E-L equation and we get

$$\begin{aligned} \frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\ \frac{d^2 f(y')}{dy'^2}\frac{dy'}{dx} &= 0 \\ \frac{d^2 f(y')}{dy'^2} y'' &= 0 \end{aligned}$$

so one of the two following must be true

$$\begin{aligned} f''(y') &= 0 \\ y'' &= 0 \end{aligned}$$

f depends only on y'

- If $f''(y') = 0$, then $f(y') = ay' + b$. We will later see that problems in this form are "degenerate", and solutions don't depend on the curve's shape.
- If $y'' = 0$, then trivially

$$y = c_1 x + c_2.$$

So for non-trivial problems with only y' dependence the extremals are straight lines (e.g. geodesics in the plane).

Example f depends only on y'

Consider finding the extremals of

$$F\{y\} = \int_0^1 \alpha y'^4 - \beta y'^2, dx$$

such that $y(0) = 0$ and $y(1) = b$.
The Euler-Lagrange equation is

$$\frac{d}{dx}\left[4\alpha y'^3 - 2\beta y'^2\right] = 0$$

We could play around with this for a while to solve, but we already know the solutions are straight lines, so the extremal will be

$$y = bx$$

Fermat's principle

Fermat's principle of geometrical optics:

> Light travels along a path between any two points such that the time taken is minimized

Take the speed of light to be dependent on the media, e.g. $c = c(x,y)$, the time taken by light along a path $y(x)$ is

$$T\{y\} = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{c(x,y)}\,dx$$

Fermat's principle says the actual path of light will be a minima of this functional.

Speed of light

The speed of light (EM radiation) is only constant in a vacuum

medium	speed (km/s)	refractive index
vacuum	300,000	1.0
water	231,000	$\sim$ 1.3
glass	200,000	$\sim$ 1.5
diamond	125,000	$\sim$ 2.4
silicon	75,000	$\sim$ 4.0

Refractive index = c/v

Example

Consider $c(x,y) = 1/g(x)$

$$T\{y\} = \int_{x_0}^{x_1} g(x)\sqrt{1+y'^2}\,dx$$

$$f(x,y,y') = g(x)\sqrt{1+y'^2}$$

f has no explicit dependence on y so

$$\begin{aligned} \frac{\partial f}{\partial y'} &= const \\ g(x)\frac{y'}{\sqrt{1+y'^2}} &= const \end{aligned}$$

Example (ii)

$$\begin{aligned} g(x)\frac{y'}{\sqrt{1+y'^2}} &= c_1 \\ \frac{y'^2}{1+y'^2} &= \frac{c_1^2}{g(x)^2} \qquad \text{implies } c_1^2 \le g(x)^2 \\ y'^2 &= \frac{c_1^2}{g(x)^2}(1+y'^2) \\ y'^2\left(1-\frac{c_1^2}{g(x)^2}\right) &= \frac{c_1^2}{g(x)^2} \\ y' &= \sqrt{\frac{c_1^2}{g(x)^2-c_1^2}} \end{aligned}$$

Example (iii)

$$y' = \sqrt{\frac{c_1^2}{g(x)^2 - c_1^2}}$$

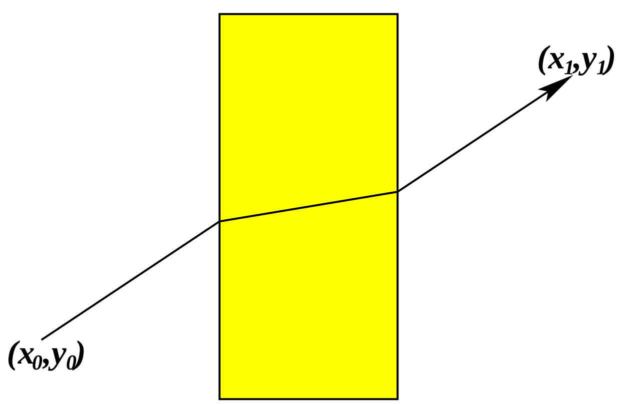
$$y = c_1 \int \frac{1}{\sqrt{g(x)^2/c_1^2 - 1}} dx + c_2$$

The constants, c_1 and c_2 are determined by the fixed end points.

- so not all extremals are straight lines
- we had to include an x term here to make it more interesting

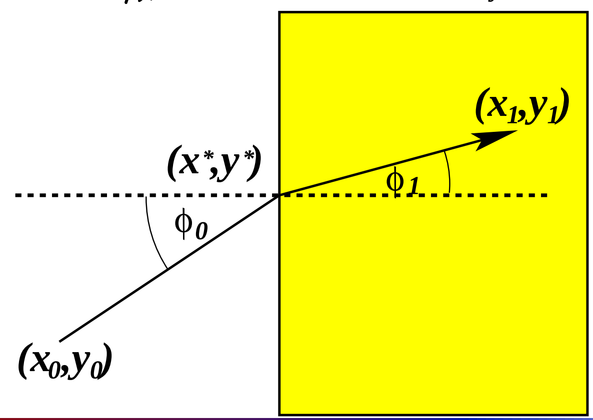
What we can't do (yet)

Remember, f must have at least two continuous derivatives. If the speed of light $c(x,y)$ has discontinuities, then we are in trouble.



How we might solve

Break into two problems, with a boundary point (x^*, y^*), which has a fixed value of x^* (the location of the boundary), but a movable value for y^*.



The functional

$$F\{y\} = \int_{x_0}^{x^*} \frac{\sqrt{1+y'^2}}{c_0} dx + \int_{x^*}^{x_1} \frac{\sqrt{1+y'^2}}{c_1} dx$$

Separate into two problems, as if we knew (x^*, y^*). Each is a geodesic in the plane problem. So the solutions are straight lines

$$y(x) = \begin{cases} (x-x_0)\frac{y^*-y_0}{x^*-x_0} + y_0 & x \leq x^* \\ (x-x^*)\frac{y_1-y^*}{x_1-x^*} + y^* & x \geq x^* \end{cases}$$

Now we can explicitly compute $F\{y\}$ as a function of x, by differentiating y, and then we can treat it as a minimization problem in one variable y^*.

The total time taken

We can simplify the integrals by noting from Pythagoras that the lengths of the two lines are

$$\sqrt{(x^*-x_0)^2+(y^*-y_0)^2} \quad \text{and} \quad \sqrt{(x^*-x_1)^2+(y^*-y_1)^2}$$

and that the time take to traverse the pair of line segments will be

$$T\{y\} = \frac{\sqrt{(x^*-x_0)^2+(y^*-y_0)^2}}{c_0} + \frac{\sqrt{(x^*-x_1)^2+(y^*-y_1)^2}}{c_1}$$

$$\frac{dT}{dy^*} = \frac{(y^*-y_0)}{c_0\left[(x^*-x_0)^2+(y^*-y_0)^2\right]^{1/2}} - \frac{(y_1-y^*)}{c_1\left[(x^*-x_1)^2+(y^*-y_1)^2\right]^{1/2}}$$

The result

$$\frac{dT}{dy^*} = \frac{(y^*-y_0)}{c_0\left[(x^*-x_0)^2+(y^*-y_0)^2\right]^{1/2}} - \frac{(y_1-y^*)}{c_1\left[(x^*-x_1)^2+(y^*-y_1)^2\right]^{1/2}}$$
$$= \frac{\sin\phi_0}{c_0} - \frac{\sin\phi_1}{c_1}$$

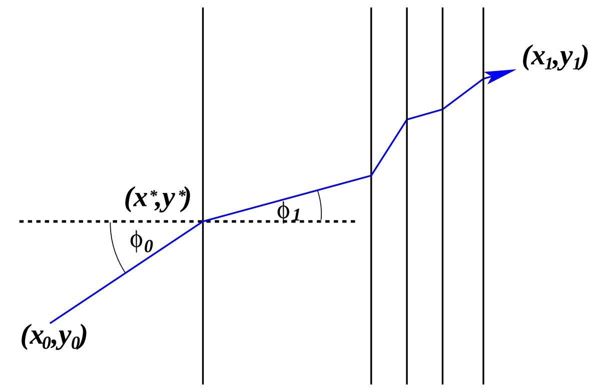
which we require to be zero to find the minimum. Hence

$$\frac{\sin\phi_0}{c_0} = \frac{\sin\phi_1}{c_1} \Leftarrow$$ **Snell's law for refraction**

Hence there are often ways around discontinuities, though it may involve some pain
(e.g. what about internal reflection)

More than one boundary

Snell's law applies at each boundary



Dealing with "kinks"

- We'll spend a fair bit of time later on dealing with "kinks" in curves
- Underlying point
  - The integral can still be well defined even if extremal isn't "smooth"
  - But the Euler-Lagrange equations don't work at the kinks
  - Use the Euler-Lagrange equations everywhere except the kinks
  - Do something else at the kinks

Variational Methods & Optimal Control

lecture 05

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 05

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Special case 2

When f has no dependence on x we call this an autonomous problem, and we can replace the E-L equations with

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y') = const$$

We will see H again later – it often turns out to be a conserved quantity like energy, and so arises naturally in computing the shape of a catenary.

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0,x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0}$$

Autonomous case

The autonomous case is where f has no explicit dependence on x, so $\partial f/\partial x = 0$.

Theorem 2.3.1: Let J be a functional of the form

$$J\{y\} = \int_{x_1}^{x_2} f(y,y')dx$$

and define the function H by

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y')$$

Then H is constant along any extremal of y.

Proof of Theorem 2.3.1

$$\begin{aligned}\frac{d}{dx}H(y,y') &= \frac{d}{dx}\left(y'\frac{\partial f}{\partial y'} - f(y,y')\right),\\ &= y''\frac{\partial f}{\partial y'} + y'\frac{d}{dx}\frac{\partial f}{\partial y'} - y'\frac{\partial f}{\partial y} - y''\frac{\partial f}{\partial y'}\\ &= y'\left(\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}\right)\\ &= 0\end{aligned}$$

So

$$H(y,y') = const$$

□

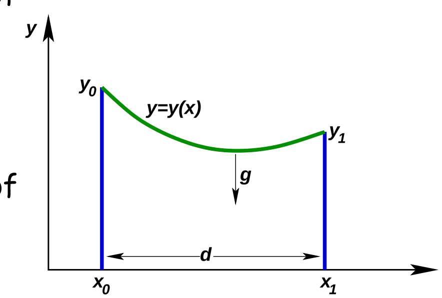
NB: this is a first order differential equation for the extremal y.

The Catenary

The potential energy of the cable is

$$W_p\{y\} = \int_0^L mgy(s)ds$$

Where L is the length of the cable



$$\begin{aligned} m &= \text{mass}\\ g &= \text{gravitational constant}\end{aligned}$$

The Catenary, reformulation

As with geodesic in the plan

$$ds = \sqrt{1+y'^2}dx$$

So the functional of interest (the potential energy) is

$$W_p\{y\} = mg\int_{x_0}^{x_1} y\sqrt{1+y'^2}dx$$

which does not contain x explicitly.

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f = const.$$

where $f(y,y') = y\sqrt{1+y'^2}$.

The Catenary (iii)

$$\begin{aligned} c_1 &= H(y,y')\\ &= y'\frac{\partial f}{\partial y'} - f\\ &= y'\frac{yy'}{\sqrt{1+y'^2}} - y\sqrt{1+y'^2}\\ c_1\sqrt{1+y'^2} &= yy'^2 - y(1+y'^2)\\ c_1^2(1+y'^2) &= y^2\\ \frac{y^2}{1+y'^2} &= c_1^2\end{aligned}$$

The Catenary (iv)

If $c_1 = 0$ the only solution is $y = 0$.
If $c_1 \neq 0$ then, rearrange to get

$$\frac{dy}{dx} = \sqrt{\frac{y^2}{c_1^2} - 1}$$

$$dx = \frac{1}{\sqrt{\frac{y^2}{c_1^2} - 1}} dy$$

$$\int dx = \int \frac{1}{\sqrt{\frac{y^2}{c_1^2} - 1}} dy$$

$$x - c_2 = \int \frac{1}{\sqrt{\frac{y^2}{c_1^2} - 1}} dy$$

The Catenary (v)

Now

$$\frac{d}{dx}\left(\cosh^{-1} u\right) = \frac{1}{\sqrt{u^2 - 1}} \frac{du}{dx},$$

So taking $u = y/c_1$ we get

$$\frac{d}{dx} \cosh^{-1}(y/c_1) = \frac{1}{\sqrt{y^2/c_1^2 - 1}} \frac{1}{c_1},$$

So, the integral above results in

$$x - c_2 = c_1 \cosh^{-1}(y/c_1).$$

The Catenary (vi)

The extremals are thus given by

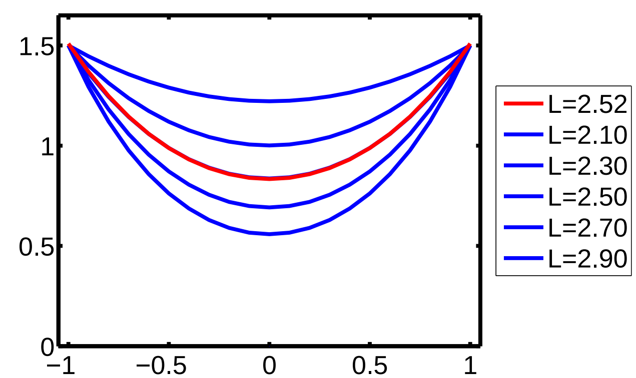
$$y = c_1 \cosh\left(\frac{x - c_2}{c_1}\right)$$

In particular, the minimal potential energy occurs when y takes this form, a **catenary**.

The constants c_1 and c_2 are determined by the end conditions, the heights of the poles, e.g. $y(x_0) = x_0$ and $y(x_1) = x_1$.

Notice I didn't specify L anywhere here.

Catenaries of different L



Finding the constants

cosh is an even function so if we take $x_0 = -1$ and $x_1 = 1$, then the constant $c_2 = 0$. So we can rewrite this as

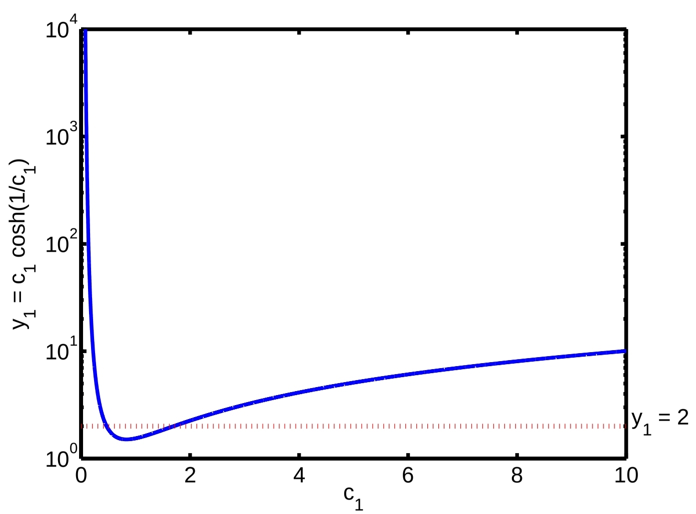
$$y(x) = c_1 \cosh\left(\frac{x}{c_1}\right)$$

which we solve for $y(1)$ to get c_1.

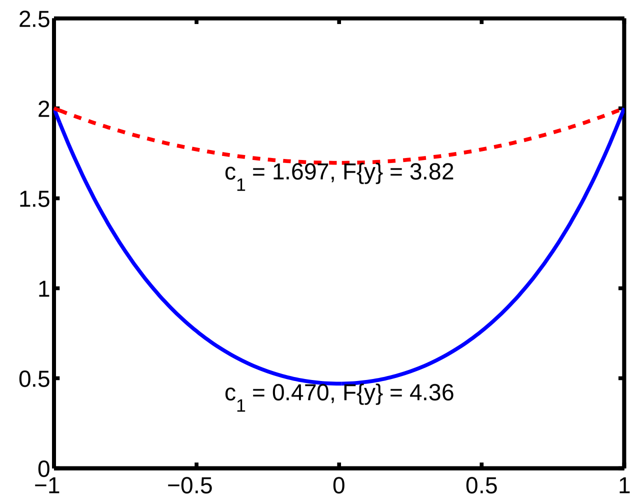
For instance $y(1) = 2$ we get two possible values $c_1 = 0.47$ and $c_1 = 1.697$

- they don't have to both be minima
- one could be a maxima, or a stationary point

Finding the constants



Finding the constants



Existence of a solution

In the above solution, note that for some values of y_0 and y_1, we can get multiple solution, but in some cases there may be a unique solution, or no solutions!!!

Calculating the functional

Once we know y, it is (in principle) easy to calculate $F\{y\}$, e.g., for the catenary note the following identities

$$\begin{aligned}\frac{d}{dx}c_1\cosh(x/c_1) &= \sinh(x/c_1)\\ 1+\sinh^2(x/c_1) &= \cosh^2(x/c_1)\end{aligned}$$

and so

$$\begin{aligned}F\{y\} &= \int_{-1}^{1} y\sqrt{1+y'^2}\,dx\\ &= \int_{-1}^{1} c_1\cosh(x/c_1)\sqrt{1+\sinh^2(x/c_1)}\,dx\\ &= \int_{-1}^{1} c_1\cosh^2(x/c_1)\,dx\end{aligned}$$

Calculating the functional

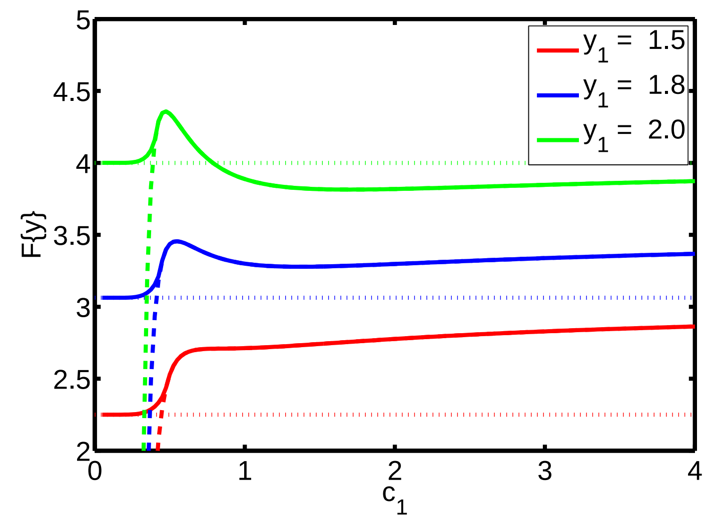
Now note that

$$\cosh^2(x) = (\cosh(2x)+1)/2$$

so that

$$\begin{aligned}F\{y\} &= \frac{c_1}{2}\int_{-1}^{1}(\cosh(2x/c_1)+1)\,dx\\ &= \frac{c_1}{2}\int_{-1}^{1}dx+\frac{c_1}{2}\int_{-1}^{1}\cosh(2x/c_1)\,dx\\ &= c_1+\frac{c_1^2}{4}\left[\sinh(2x/c_1)\right]_{-1}^{1}\\ &= c_1+\frac{c_1^2}{2}\sinh(2/c_1)\end{aligned}$$

Calculating the functional



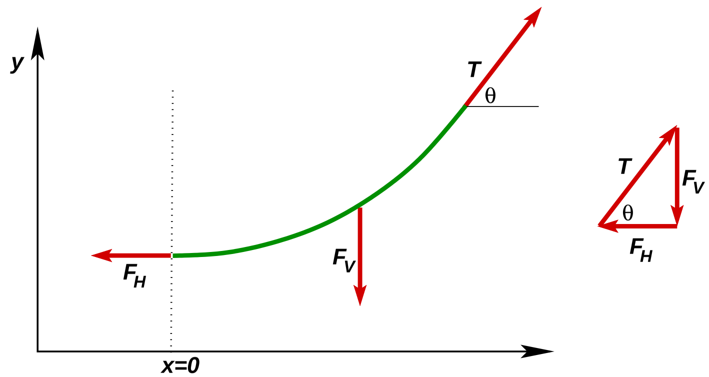
The length of the Catenary

$$\begin{aligned}L\{y\} &= \int_{-1}^{1}\sqrt{1+y'^2}\,dx\\ &= \int_{-1}^{1}\cosh(x/c_1)\,dx\\ &= c_1\left[\sinh(x/c_1)\right]_{-1}^{1}\\ &= 2c_1\sinh(1/c_1)\end{aligned}$$

But note that in this version of the problem we can't **set** the length, it is an output. Later on we will constrain the length so it is an input to the problem.

Catenary addendum

The usual explanation for the shape of the catenary is based on a simple physical argument: **forces must be balance in equilibrium.**



Catenary addendum

forces must be balance in equilibrium so tension in the cable (which must be in the direction of the cable) must balance the horizontal force F_H at the lowest point, and the downwards force F_V. The results is

$$\begin{aligned} \tan\theta &= \frac{F_V}{F_H} \\ \frac{dy}{dx} &= \frac{gms}{F_H} \end{aligned}$$

where ms is the mass of the cable integrated from $[0,s]$ along the cable, and F_H is constant.

Catenary addendum

Taking derivatives with respect to x we get

$$\begin{aligned} \frac{d}{dx}\frac{dy}{dx} &= \frac{d}{dx}\frac{m(x)g}{F_H} \\ y'' &= \frac{mg}{F_H}\frac{ds}{dx} \end{aligned}$$

where we know that $\frac{ds}{dx} = \sqrt{1+y'^2}$ so

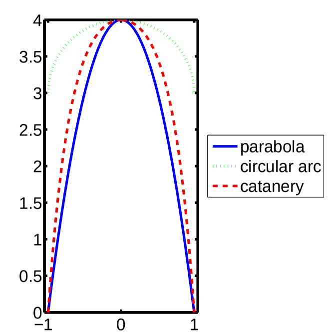
$$\frac{y''}{\sqrt{1+y'^2}} = \frac{mg}{F_H}$$

which has the same solution, but now c_1 has a meaning

$$y(x) = \frac{F_H}{mg}\cosh\left(\frac{mg}{F_H}x\right).$$

The shape of an arch

Flip a catenary upside down, and the above argument shows simply that the strongest form of an arch is an inverted catenary. This balances the forces at each point, so that the arch is under the least possible stress.



Note that F_H must be applied to the edges or the arch will collapse outwards.

The shape of an arch

However, this argument assumes that the arch's own weight is all that matters. Commonly, an arch supports a wall above, and so the forces are not so simply described. The shape that is optimal is closer to the shape of a suspension bridge, which we shall see in tutorials is a parabola.

- BTW, the *Gateway Arch* in St Louis isn't strictly a catenary as is sometimes claimed.

  `http://www.springerlink.com/content/u7734w06700776x0/`
- the optimal form changes if the "arch" isn't a pure curve, but has shape.

Variational Methods & Optimal Control

lecture 06

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 06

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Special case 2: autonomous problems continued

$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y') = const$$

We will see H again later – it often turns out to be a conserved quantity like energy, and so arises naturally in computing the shape of the brachystochrone.

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0,x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0}$$

Autonomous case

The autonomous case is where f has no explicit dependence on x, so $\partial f/\partial x = 0$.

Theorem 2.3.1: Let J be a functional of the form

$$J\{y\} = \int_{x_1}^{x_2} f(y,y')dx$$

and define the function H by

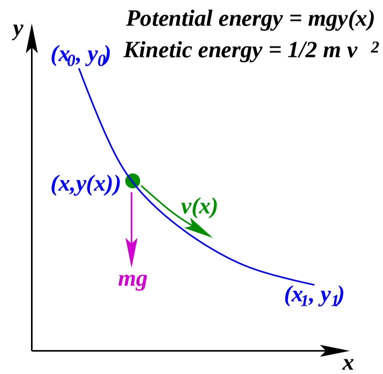
$$H(y,y') = y'\frac{\partial f}{\partial y'} - f(y,y')$$

Then H is constant along any extremal of y.

Example: Brachystochrone

The time taken is

$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$



The energy of a body is the sum of potential and kinetic energy

$$E = \frac{1}{2}mv(x)^2 + mgy(x)$$

and a simple conservation law says this is constant, so

$$v(x) = \sqrt{\frac{2E}{m} - 2gy(x)}$$

Example: Brachystochrone (ii)

As for the geodesic in the plane

$$ds = \sqrt{1+y'^2}dx$$

So the functional of interest (the time taken) is

$$T\{y\} = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{\frac{2E}{m} - 2gy(x)}}dx$$

We can perform a substitution

$$w(x) = \frac{1}{2g}\left(\frac{2E}{m} - 2gy(x)\right)$$

And note that $w'^2 = y'^2$, so (ignoring the constant factor of $-1/2g$) we look for extremals of

Example: Brachystochrone (iii)

Look for extremals of

$$T\{w\} = \int_{x_0}^{x_1} \sqrt{\frac{1+w'^2}{w}}dx$$

which does not contain x explicitly.

$$\begin{aligned} H(w,w') = w'\frac{\partial f}{\partial w'} - f &= \frac{w'^2}{w}\left(\frac{1+w'^2}{w}\right)^{-1/2} - \sqrt{\frac{1+w'^2}{w}} \\ &= \frac{w'^2}{\sqrt{w(1+w'^2)}} - \sqrt{\frac{1+w'^2}{w}} \\ &= \frac{-1}{\sqrt{w(1+w'^2)}} \end{aligned}$$

Example: Brachystochrone (iv)

$$H(w,w') = const$$

So we can write

$$w(1+w'^2) = c_1$$

Let $w' = \tan\phi$, then $1+w'^2 = \sec^2\phi$ and for $\kappa_1 = c_1/2$

$$w = \frac{c_1}{\sec^2\phi} = c_1\cos^2\phi = \kappa_1[1+\cos(2\phi)]$$

$$\frac{dw}{d\phi} = -2\kappa_1\sin(2\phi) = -4\kappa_1\cos(\phi)\sin(\phi)$$

Example: Brachystochrone (v)

Also $dw/dx = \tan\phi$, which means

$$\frac{dx}{dw} = \frac{1}{\tan\phi} = \cot\phi$$

Also

$$\frac{dx}{d\phi} = \frac{dx}{dw}\frac{dw}{d\phi} = -4\kappa_1 \cos^2\phi = -2\kappa_1(1+\cos(2\phi))$$

Integrating

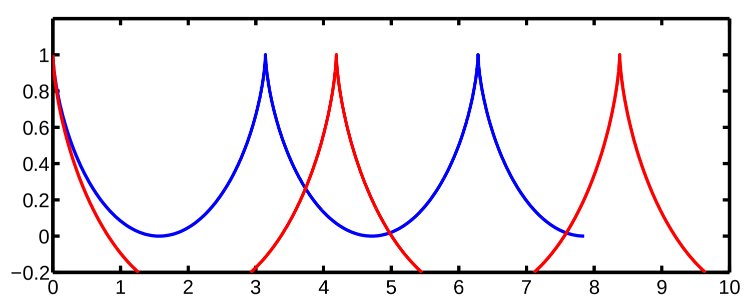
$$x = \kappa_2 - \kappa_1(2\phi + \sin(2\phi))$$

Along with

$$w = \kappa_1[1+\cos(2\phi)]$$

we have a parametric form of the solution.

Cycloids



Example: Brachystochrone solution

Take $\theta + \pi = 2\phi$ and we get

$$\begin{aligned} x &= \kappa_2 + \kappa_1(\theta - \sin(\theta)) \\ w &= \kappa_1[1-\cos(\theta)] \end{aligned}$$

Lets change back to y, remembering $w(x) = \frac{1}{2g}\left(\frac{2E}{m} - 2gy(x)\right)$, and that $E = \frac{1}{2}mv^2 + mgy = const$ and $v(x_0) = 0$, so that $E = mgy_0$, hence

$$y = y_0 - w$$

Note that $y(x)$ doesn't depend on g or m!

Now $y(x_0) = y_0$ and so $w(\theta_0) = 0$, which we get when $\theta_0 = 0$.

Now $x(\theta_0) = x_0$ and so $\kappa_2 = x_0$, so the solution is

Example: Brachystochrone solution

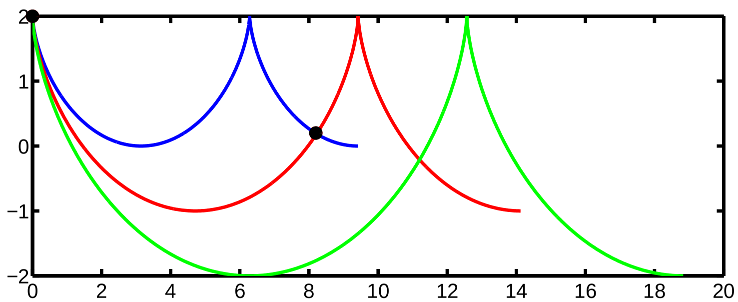
Take $\theta + \pi = 2\phi$ and we get

$$\boxed{\begin{aligned} x &= x_0 + \kappa_1(\theta - \sin(\theta)) \\ y &= y_0 - \kappa_1[1-\cos(\theta)] \end{aligned}}$$

Now, note that $y(x_1) = y_1$. We find θ_1 first by solving

$$\begin{aligned} y_1 &= y_0 - \kappa_1[1-\cos(\theta_1)] \\ [1-\cos(\theta_1)] &= \frac{y_0 - y_1}{\kappa_1} \\ \cos(\theta_1) &= 1 - \frac{y_0 - y_1}{\kappa_1} \\ \theta_1 &= \arccos\left(1 - \frac{y_0 - y_1}{\kappa_1}\right) \end{aligned}$$

Cycloids



More than one possible solution!!
We need to find the fastest one!

Meaning of H

- H is a **conserved** quantity.
- In physics often see such, e.g. the energy
  H is not energy in Brachystochrone problem
- Can derive conservation laws mathematically.
  rather than deriving them as physical laws
- later on we consider Noether's theorem

Newton's aerodynamic problem

"If in a rare medium, consisting of equal particles freely disposed at equal distances from each other, a globe and a cylinder described on equal diameter move with equal velocities in the direction of the axis of the cylinder, the resistance of the globe will be half as great as that of the cylinder ... I reckon that this proposition will be not without application in the building of ships".

Isaac Newton, Principia Mathematica

Newton's aerodynamical problem

Consider finding the optimal shape of a rocket's nose cone in order that it creates the least resistance when passing through air.
Assumptions:

- Air is thin, and composed of perfectly elastic particles:
  - particles will bounce off the nose cone with equal speed, and equal angle of reflection and incidence.
  - We ignore tangential friction.
  - We ignore "non-Newtonian" affects such as those from compression of the air.

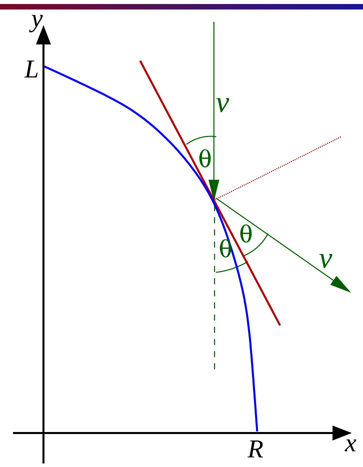
Realistic for high-altitude, supersonic flight

Newton's aerodynamical problem

Consider finding the optimal shape of a rocket's nose cone in order that it creates the least resistance when passing through air.
Assumptions:

- As the rocket may rotate along its length, the nose cone must be circularly symmetric, and so we reduce the problem to one of determining the optimal profile of the nose cone.
- The rocket's nose cone must have radius R at its base, and length L, and its shape should be convex
  - its profile must be concave and non-increasing
  - ratio $L/2R$ is called the **fineness ratio**
  - bigger is better, though little gain for $> 5:1$

Newton's aerodynamical problem



Newton's aerodynamical problem

$$Force = ma$$

- $m =$ mass
- $a =$ acceleration = change in velocity

$$a = v - s = 2v\sin^2\theta.$$

Scale constants so that

$$2vm = 1,$$

and then

$$Force = \sin^2\theta = \frac{1}{1+\cot^2\theta} = \frac{1}{1+y'^2}.$$

Newton's aerodynamical problem

- Previous calculation gives force per particle $= 1/(1+y'^2)$
- Need to integrate over surface area
- Surface area at radius x is

$$2\pi x\, dx.$$

- Scaling to remove irrelevant constants, the functional describing the resistance

$$F\{y\} = \int_0^R \frac{x}{1+y'^2}\, dx,$$

- subject to $y(0) = L$ and $y(R) = 0$ and $y' \leq 0$ and $y'' \geq 0$

Newton's aerodynamical problem

The Euler-Lagrange equations are

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = \frac{d}{dx}\frac{2xy'}{(1+y'^2)^2} = 0$$

So for a given constant c, we get

$$\frac{2xy'}{(1+y'^2)^2} = c.$$

Rearranging we get

$$2xy' \;=\; c(1+y'^2)^2$$

Newton's aerodynamical problem

We'll solve this when we get to optimal control.
For now here is the parametric solution without explanation

$$\begin{aligned} x(u) &= c\left(\frac{1}{u} + 2u + u^3\right) = \frac{c}{u}(1+u^2)^2 \\ y(u) &= L - c\left(-\ln u - \frac{7}{4} + u^2 + \frac{3}{4}u^4\right) \end{aligned}$$

But notice that

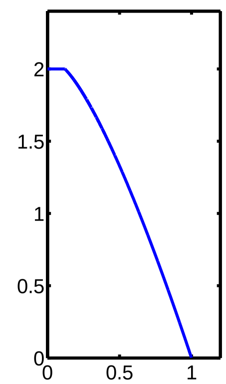
$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{dy}{du}\Big/\frac{dx}{du} = -u$$

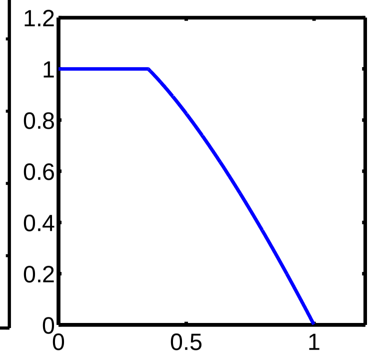
from which it is relatively clear that this is a solution.

Newton's aerodynamical problem

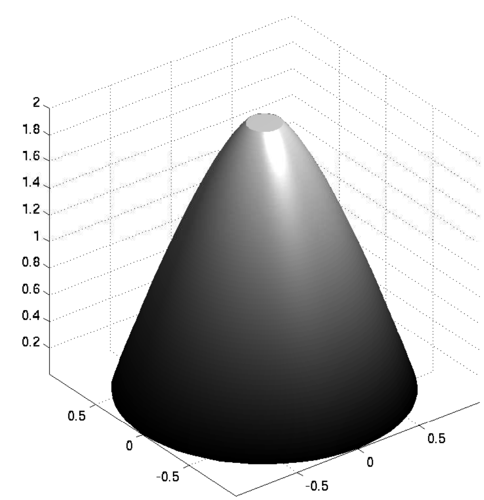
Solution looks almost like a blunted cone

- perhaps that seems counter-intuitive?





Newton's aerodynamical problem



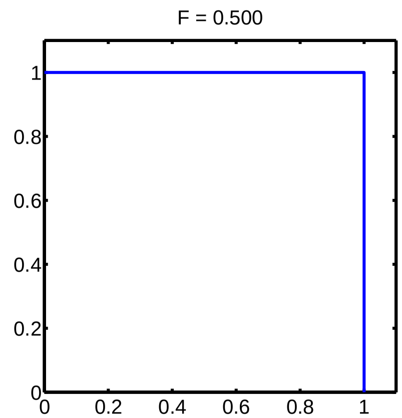
Alternatives: cylinder

Cylinder:

$$\begin{aligned} y' &= 0 \\ F\{y\} &= \int_0^R x\,dx \\ &= \frac{R^2}{2} \end{aligned}$$

For $R = 1$

$$F = 1/2$$



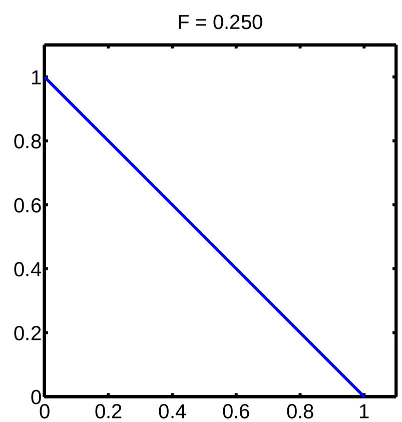
Alternatives: cone

Cone:

$$\begin{aligned} y' &= -L/R \\ F\{y\} &= \int_0^R \frac{x}{1+(L/R)^2}\,dx \\ &= \frac{R^2}{2(1+(L/R)^2)} \end{aligned}$$

For $R = L = 1$

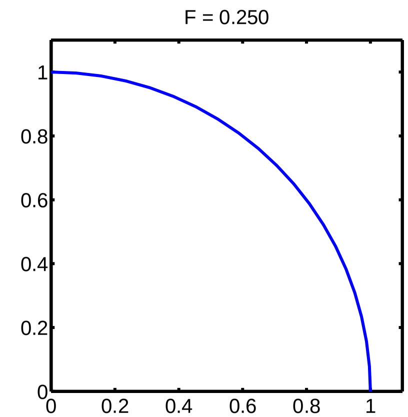
$$F = 1/4$$



Alternatives: sphere

Sphere: $R = L = 1$

$$\begin{aligned} x^2 + y^2 &= 1 \\ y' &= -x/y \\ &= -x/\sqrt{1-x^2} \\ F\{y\} &= \int_0^1 \frac{x}{1+y'^2}\,dx \\ &= \int_0^1 x(1-x^2)\,dx \\ &= \frac{1}{4} \end{aligned}$$



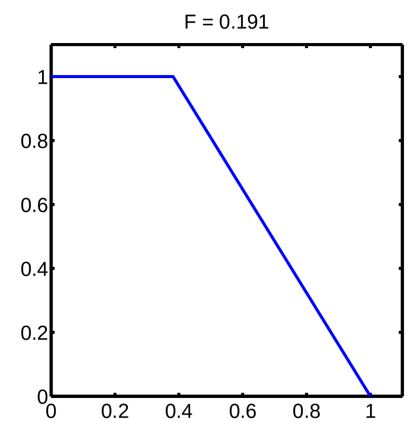
Alternatives: frustum of cone

Frustum of cone: corner at a

$$\begin{aligned} y' &= \begin{cases} 0 & x \le a \\ -L/(R-a) & x \ge a \end{cases} \\ F\{y\} &= \int_0^a x\,dx + \int_a^R \frac{x}{1+y'^2}\,dx \\ &= \frac{a^2L^2 + R^2(R-a)^2}{2(L^2+(R-a)^2)} \end{aligned}$$

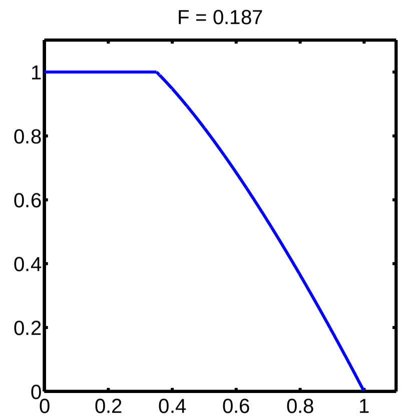
Optimal value of a:

$$a = \frac{(L^2+2R^2) - L\sqrt{L^2+4R^2}}{2R}$$



Alternatives: optimal

Optimal profile:
F computed numerically

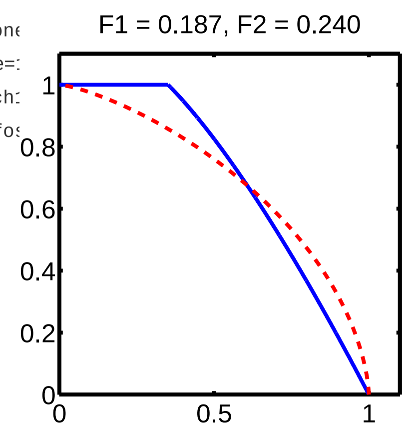


Alternatives: Haack series

Haack series:

http://en.wikipedia.org/wiki/Nose_cone
http://www.info-central.org/?article=:
http://mcfisher.0catch.com/other/mach:
http://www.if.sc.usp.br/~projetosulfo:

F computed numerically



Typical shapes

- Note that the frustum of a cone isn't much worse than the optimal shape.
- other shapes: ogive, Haack, ...
- In the context of bullets a flattened end is called a **meplat**.
  - ▷ typically justified by
    - ⋆ making all bullets precise
      - tips are hard to get just right
    - ⋆ impact damage
  - ▷ but they wouldn't do it if it wasn't working

Bullets

Variational Methods & Optimal Control

lecture 07

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 07

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Special case 3

When f has no explicit dependence on y the E-L equations simplify to give

$$\frac{\partial f}{\partial y'} = const$$

An example where we might use this is in calculating geodesics on non-planar objects such as the sphere.

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0, x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0, x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0, x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0}$$

No explicit y dependence

Suppose the function is of the form

$$J\{y\} = \int_{x_0}^{x_1} f(x, y') dx$$

where y does not appear explicitly.

The Euler-Lagrange equation reduces to

$$\frac{\partial f}{\partial y'} = c_1$$

where c_1 is a constant.

Solving

$\frac{\partial f}{\partial y'}$ is a known function of x and y',
so this is a first order DE for y.
In principle for $\frac{\partial^2 f}{\partial y'^2} \neq 0$ can recast

$$\frac{\partial f}{\partial y'} = c_1 \quad \text{as} \quad y' = g(x, c_1)$$

for some g.

Geodesics on the unit sphere

Find the shortest path between two points on the unit sphere.

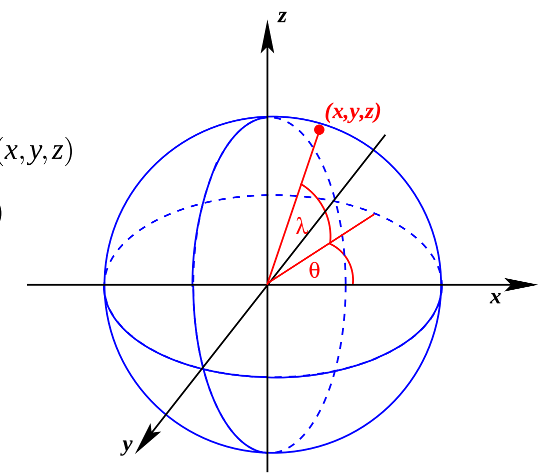
Spherical co-ordinates

Define

$$\begin{aligned} \lambda &= \text{latitude} \\ \theta &= \text{longitude} \end{aligned}$$

Cartesian co-ordinates (x, y, z)

$$\begin{aligned} x &= \cos(\theta)\cos(\lambda) \\ y &= \sin(\theta)\cos(\lambda) \\ z &= \sin(\lambda) \end{aligned}$$



Transformation to spherical co-ord.

$$\begin{aligned} x &= \cos(\theta)\cos(\lambda) \\ y &= \sin(\theta)\cos(\lambda) \\ z &= \sin(\lambda) \end{aligned}$$

Chain rule

$$\begin{aligned} dx &= \frac{\partial x}{\partial \theta} d\theta + \frac{\partial x}{\partial \lambda} d\lambda &= -\sin(\theta)\cos(\lambda) d\theta - \cos(\theta)\sin(\lambda) d\lambda \\ dy &= \frac{\partial y}{\partial \theta} d\theta + \frac{\partial y}{\partial \lambda} d\lambda &= \cos(\theta)\cos(\lambda) d\theta - \sin(\theta)\sin(\lambda) d\lambda \\ dz &= \frac{\partial z}{\partial \theta} d\theta + \frac{\partial z}{\partial \lambda} d\lambda &= \cos(\lambda) d\lambda \end{aligned}$$

$$ds^2 = dx^2 + dy^2 + dz^2 = d\lambda^2 + \cos^2(\lambda) d\theta^2$$

Geodesics on the unit sphere

$$\int_{(x(s_0),y(s_0),z(s_0))}^{(x(s_1),y(s_1),z(s_1))} 1 ds = \int_{\lambda_0}^{\lambda_1} \left[1+\cos^2(\lambda)\left(\frac{d\theta}{d\lambda}\right)^2\right]^{\frac{1}{2}} d\lambda$$

θ is like y, λ is like x, $\frac{d\theta}{d\lambda}=\theta'$ is like y', hence EL eqn:

$$\begin{aligned}\frac{\partial}{\partial\theta'}\left[1+\cos^2(\lambda)\theta'^2\right]^{\frac{1}{2}} &= c_1\\ \frac{\cos^2(\lambda)\theta'}{[1+\cos^2(\lambda)\theta'^2]^{\frac{1}{2}}} &= c_1\\ \frac{\cos^4(\lambda)\theta'^2}{1+\cos^2(\lambda)\theta'^2} &= c_1^2\end{aligned}$$

The constant

$$\frac{\cos^4(\lambda)\theta'^2}{1+\cos^2(\lambda)\theta'^2} = c_1^2$$

Now

$$\theta'^2\cos^4(\lambda) \le \theta'^2\cos^2(\lambda) \le 1+\theta'^2\cos^2(\lambda)$$

So

$$c_1 \in [-1,1]$$

So we can replace c_1 with

$$c_1 = \cos(\alpha)$$

Geodesics on the unit sphere

Re-arrange

$$\cos^4(\lambda)\theta'^2 = c_1^2\left(1+\cos^2(\lambda)\theta'^2\right)$$

Re-arrange some more

$$\begin{aligned}\theta'^2 &= \frac{c_1^2}{\cos^4(\lambda)-c_1^2\cos^2(\lambda)}\\ \theta' &= \left\{\frac{c_1^2}{\cos^2(\lambda)(\cos^2(\lambda)-c_1^2)}\right\}^{\frac{1}{2}}\\ \theta' &= g(\lambda,c_1)\end{aligned}$$

Analogous to $y'=g(x,c_1)$.

Solving the DE

Insert $c_1=\cos(\alpha)$

$$\begin{aligned}\theta' &= \frac{\cos(\alpha)}{\cos(\lambda)\left[\cos^2(\lambda)-\cos^2(\alpha)\right]^{\frac{1}{2}}}\\ \theta &= \int \frac{\cos(\alpha)}{\cos(\lambda)\left[\cos^2(\lambda)-\cos^2(\alpha)\right]^{\frac{1}{2}}}\,d\lambda\\ &= \int \frac{\sec^2(\lambda)}{\left[\sec^2(\alpha)-\sec^2(\lambda)\right]^{\frac{1}{2}}}\,d\lambda\\ &= \int \frac{\sec^2(\lambda)}{\left[\tan^2(\alpha)-\tan^2(\lambda)\right]^{\frac{1}{2}}}\,d\lambda \quad \text{as } \sec^2 x = 1+\tan^2 x\end{aligned}$$

Solving the DE (part ii)

$$\theta = \int \frac{\sec^2(\lambda)}{[\tan^2(\alpha) - \tan^2(\lambda)]^{\frac{1}{2}}} d\lambda = \frac{1}{\tan(\alpha)} \int \frac{\sec^2(\lambda)}{[1 - \tan^2(\lambda)/\tan^2(\alpha)]^{\frac{1}{2}}} d\lambda$$

Substitute $u = \tan(\lambda)/\tan(\alpha)$
Then $d\lambda = \frac{\tan(\alpha)}{\sec^2(\lambda)} du$

$$\begin{aligned} \theta &= \int \frac{1}{[1-u^2]^{\frac{1}{2}}} du \\ &= \sin^{-1}\left(\frac{\tan(\lambda)}{\tan(\alpha)}\right) - \beta \end{aligned}$$

As $\frac{d}{du} \sin^{-1}(u) = \frac{1}{\sqrt{1-u^2}}$

The solution

$$\sin(\theta + \beta) = \frac{\tan(\lambda)}{\tan(\alpha)}$$

Note we can write this

$$\begin{aligned} \sin(\theta+\beta) &= \frac{1}{\tan(\alpha)} \frac{\sin(\lambda)}{\cos(\lambda)} \\ \tan(\alpha)\cos(\lambda)\sin(\theta+\beta) &= \sin(\lambda) \\ \tan(\alpha)\cos(\lambda)\,[\sin(\theta)\cos(\beta) + \cos(\theta)\sin(\beta)] &= \sin(\lambda) \end{aligned}$$

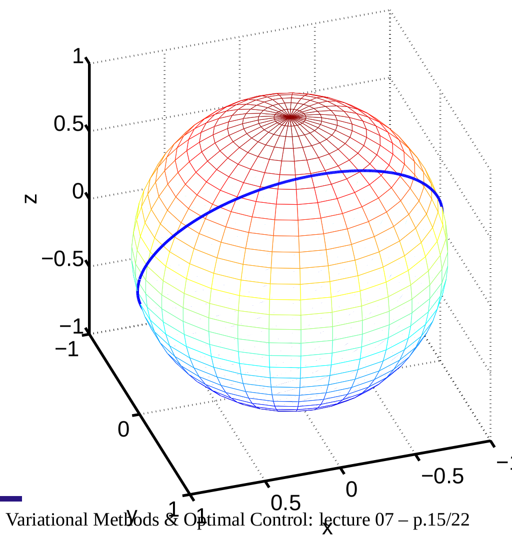
Convert back to Cartesian co-ordinates,

$$\tan(\alpha)\sin(\beta)x + \tan(\alpha)\cos(\beta)y = z$$

which is the equation of a plane, through the origin.
Hence, solution is a **great circle**, the intersection of plane (through the origin) and the sphere.

Example

We can find the solution because three points (the origin plus the start and end point of the curve) define a plane, and therefore the solution is the intersection of this plane with the sphere.



Co-ordinate transformation

More generally, spherical co-ordinates

$$\begin{aligned} x &= r\cos(\theta)\cos(\lambda) \\ y &= r\sin(\theta)\cos(\lambda) \\ z &= r\sin(\lambda) \end{aligned}$$

And

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = J \begin{pmatrix} d\theta \\ d\lambda \\ dr \end{pmatrix}, \quad J = \begin{pmatrix} \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \lambda} & \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \lambda} & \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \lambda} & \frac{\partial z}{\partial r} \end{pmatrix}$$

Where J is the Jacobian matrix

Jacobians

If

$$\mathbf{y} = \begin{pmatrix} y_1(\mathbf{x}) \\ y_2(\mathbf{x}) \\ \vdots \\ y_n(\mathbf{x}) \end{pmatrix}$$

Then the Jacobian matrix is

$$J(\mathbf{x}) = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}$$

The Jacobian determinant

Then the determinant of the Jacobian matrix is also sometimes called the Jacobian

$$|J(\mathbf{x})| = \left| \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \right|$$

This gives the ratios of n-dimensional volumes between the two co-ord. systems, i.e.

$$d\mathbf{y} = |J(\mathbf{x})| \, d\mathbf{x}$$

Transforms and integrals

Substitution in 1D: $y = \phi(x)$

$$\int_{x_0}^{x_1} f(\phi(x)) \frac{d\phi}{dx} dx = \int_{\phi(x_0)}^{\phi(x_1)} f(y) dy$$

In 2D

$$\int_R f(x,y) \, dx \, dy = \int_{R^*} f(x(u,v), y(u,v)) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| du \, dv$$

Geodesics

Can we find a geodesic on other surfaces in $\mathbb{R}^3$?

Consider a surface parameterized by $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$, and minimize the arc length

$$L = \int ds = \int \sqrt{dx^2 + dy^2 + dz^2}$$

$$\begin{aligned} dx &= \frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv \\ dx^2 &= \left(\frac{\partial x}{\partial u} \right)^2 du^2 + 2 \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} du \, dv + \left(\frac{\partial x}{\partial v} \right)^2 dv^2 \end{aligned}$$

and likewise for dy^2 and dz^2.

Geodesics

So we can write the path length as

$$\begin{aligned} L &= \int \sqrt{P + 2Qv' + Rv'^2}\, du \\ &= \int \sqrt{Pu'^2 + 2Qu' + R}\, dv \end{aligned}$$

where $u' = du/dv$ and $v' = dv/du$ and

$$\begin{aligned} P &= \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 \\ Q &= \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\frac{\partial z}{\partial v} \\ R &= \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2 \end{aligned}$$

Geodesics

Then the Euler-Lagrange equations become

$$\frac{\frac{\partial P}{\partial v} + 2v'\frac{\partial Q}{\partial v} + v'^2\frac{\partial R}{\partial v}}{2\sqrt{P + 2Qv' + Rv'^2}} - \frac{d}{du}\left(\frac{Q + Rv'}{\sqrt{P + 2Qv' + Rv'^2}}\right) = 0$$

References:
`http://mathworld.wolfram.com/GreatCircle.html`
`http://mathworld.wolfram.com/Geodesic.html`

Variational Methods & Optimal Control

lecture 08

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 08

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Invariance of the E-L equations

We side-track here to note that extremals found using the E-L equations don't depend on the coordinate system! This can be very useful – a change of co-ordinates can often simplify a problem dramatically.

Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0,x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0}$$

Invariance of the E-L equations

The extremals found using the E-L equations don't depend on the coordinate system!

For instance take co-ordinate transform

$$\begin{aligned} x &= x(u,v) \\ y &= y(u,v) \end{aligned}$$

- **smooth:** if functions x and y have continuous partial derivatives.
- **non-singular:** if Jacobian is non-zero

For example, the path of a particle does not depend on the coordinate system used to describe the path!

Notation

Use the notation

$$x_u = \frac{\partial x}{\partial u}$$

For example, the Jacobian for transform $x = x(u,v)$ and $y = y(u,v)$ can be written

$$J = \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix} = x_u y_v - x_v y_u$$

Note that if $J \neq 0$ the transform is invertible.

- treat u like the independent variable (like x)
- treat v like the dependent variable (like y)

Transforming dy/dx

Treat v like a function $v(u)$. The chain rule says for $x = x(u,v)$

$$\frac{dx}{du} = \frac{du}{du}\frac{\partial x}{\partial u} + \frac{dv}{du}\frac{\partial x}{\partial v}$$

so

$$\begin{aligned} \frac{dx}{du} &= x_u + x_v v' \\ \frac{dy}{du} &= y_u + y_v v' \end{aligned}$$

where $v' = dv/du$. So

$$\frac{dy}{dx} = \frac{dy/du}{dx/du} = \frac{y_u + y_v v'}{x_u + x_v v'}$$

Transforming functional

Transforming the functional, we get

$$\begin{aligned} F\{y\} &= \int_{x_0}^{x_1} f(x,y,y')\,dx \\ &= \int_{u_0}^{u_1} f\left(x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v')\,du \\ &= \int_{u_0}^{u_1} \tilde{f}(u,v,v')\,du \end{aligned}$$

Relabel the functional to get

$$\tilde{F}\{v\} = \int_{u_0}^{u_1} \tilde{f}(u,v,v')\,du$$

Fixed end-point problem

Find extremals of functional $F : C^2[x_0,x_1] \to \mathbb{R}$ given by

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

and the extremal is in the set S

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

Becomes, find extremals of $\tilde{F} : C^2[u_0,u_1] \to \mathbb{R}$ given by

$$\tilde{F}\{v\} = \int_{u_0}^{u_1} \tilde{f}(u,v,v')\,du$$

and the extremal is in the set S

$$\tilde{S} = \left\{ v \in C^2[u_0,u_1] \,\middle|\, v(u_0) = v_0 \text{ and } v(u_1) = v_1 \right\},$$

Relation between extremals

Theorem: Let $y \in S$ and $v \in \tilde{S}$ be two functions that satisfy the smooth, non-singular transformation $x = x(u,v)$, and $y = y(u,v)$, then y is an extremal for F if and only if v is an extremal for $\tilde{F}$.

Proof Sketch: The proof needs to show that the Euler-Lagrange equations for both problems produce the same extremals.
We can do so, by noting that

$$\frac{d}{du}\left(\frac{\partial \tilde{f}}{\partial v'}\right) - \frac{\partial \tilde{f}}{\partial v} = J\left[\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y}\right]$$

As the transform is non-singular $J \neq 0$, so if either side is zero, the Euler-Lagrange equation is satisfied for both problems.

Some of the details

$$\begin{aligned}
\tilde{f}(u,v,v') &= f\left(x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v') \\
\frac{\partial \tilde{f}}{\partial v} &= \left(\frac{\partial f}{\partial x}x_v + \frac{\partial f}{\partial y}y_v + \frac{\partial f}{\partial y'}\frac{\partial}{\partial v}\left(\frac{y_u + y_v v'}{x_u + x_v v'}\right)\right)(x_u + x_v v') \\
&\quad + f\frac{\partial}{\partial v}(x_u + x_v v') \\
\frac{\partial \tilde{f}}{\partial v'} &= \frac{\partial f}{\partial y'}(x_u + x_v v')\frac{\partial}{\partial v'}\left(\frac{y_u + y_v v'}{x_u + x_v v'}\right) + x_v f \\
J &= x_u y_v - x_v y_u
\end{aligned}$$

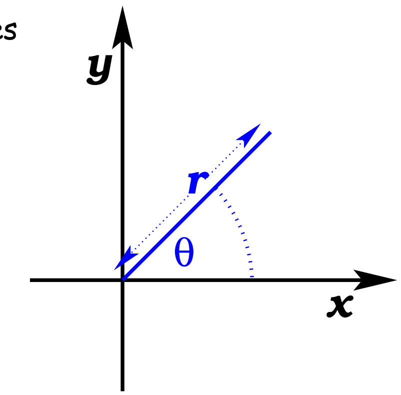
Example

Polar (circular) coordinates have

$$\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta
\end{aligned}$$

and inverse transform

$$\begin{aligned}
r &= \sqrt{x^2 + y^2} \\
\theta &= \arctan\left(\frac{y}{x}\right)
\end{aligned}$$



Find extremals of $F\{r\} = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + r'^2}\, d\theta$

Example

For the inverse transform

$$\begin{aligned}
r_x &= x/\sqrt{x^2 + y^2} \\
r_y &= y/\sqrt{x^2 + y^2} \\
\theta_x &= (-y/x^2)/(1 + (y/x)^2) = -y/(x^2 + y^2) \\
\theta_y &= (1/x)/(1 + (y/x)^2) = x/(x^2 + y^2)
\end{aligned}$$

using $\boxed{\frac{d}{dz}\arctan(z) = \frac{1}{1+z^2}}$

Example

The Jacobian

$$\begin{aligned} J &= \det\begin{pmatrix} r_x & \theta_x \\ r_y & \theta_y \end{pmatrix} \\ &= \det\begin{pmatrix} x/\sqrt{x^2+y^2} & -y/(x^2+y^2) \\ y/\sqrt{x^2+y^2} & x/(x^2+y^2) \end{pmatrix} \\ &= \frac{x^2+y^2}{(x^2+y^2)^{3/2}} \\ &= 1/\sqrt{x^2+y^2} \end{aligned}$$

$J \neq 0$ everywhere except $(x,y)=(0,0)$, where it is undefined.

Example

$$\begin{aligned} \frac{dr}{d\theta} &= \frac{r_x + r_y y'}{\theta_x + \theta_y y'} \\ &= \frac{x/\sqrt{x^2+y^2} + yy'/\sqrt{x^2+y^2}}{-y/(x^2+y^2) + xy'/(x^2+y^2)} \\ &= \sqrt{x^2+y^2}\,\frac{x+yy'}{-y+xy'} \\ r^2 + \left(\frac{dr}{d\theta}\right)^2 &= (x^2+y^2) + (x^2+y^2)\left(\frac{x+yy'}{-y+xy'}\right)^2 \\ &= (x^2+y^2)\left[1 + \left(\frac{x+yy'}{-y+xy'}\right)^2\right] \end{aligned}$$

Example

$$\begin{aligned} r^2 + \left(\frac{dr}{d\theta}\right)^2 &= (x^2+y^2)\left[1 + \left(\frac{x+yy'}{-y+xy'}\right)^2\right] \\ &= (x^2+y^2)\left[1 + \frac{x^2+2xyy'+y^2y'^2}{y^2-2xyy'+x^2y'^2}\right] \\ &= (x^2+y^2)\left[\frac{y^2-2xyy'+x^2y'^2+x^2+2xyy'+y^2y'^2}{y^2-2xyy'+x^2y'^2}\right] \\ &= (x^2+y^2)\left[\frac{x^2+y^2+(x^2+y^2)y'^2}{y^2-2xyy'+x^2y'^2}\right] \\ &= \frac{(x^2+y^2)^2(1+y'^2)}{(-y+xy')^2} \end{aligned}$$

Example

Now

$$\begin{aligned} \frac{d\theta}{dx} &= \frac{\partial\theta}{\partial x} + \frac{\partial\theta}{\partial y}\frac{dy}{dx} \\ &= -\frac{y}{(x^2+y^2)} + \frac{x}{(x^2+y^2)}y' \\ &= \frac{-y+xy'}{(x^2+y^2)} \\ \frac{dx}{d\theta} &= \frac{(x^2+y^2)}{-y+xy'} \\ r^2 + \left(\frac{dr}{d\theta}\right)^2 &= (1+y'^2)\left(\frac{dx}{d\theta}\right)^2 \end{aligned}$$

Example

Given that

$$r^2 + \left(\frac{dr}{d\theta}\right)^2 = (1+y'^2)\left(\frac{dx}{d\theta}\right)^2$$

The functional can be rewritten

$$\begin{aligned} F\{r\} &= \int_{\theta_0}^{\theta_1} \sqrt{r^2 + r'^2}\, d\theta \\ &= \int_{\theta_0}^{\theta_1} \sqrt{1+y'^2}\, \frac{dx}{d\theta}\, d\theta \\ \tilde{F}\{y\} &= \int_{\mathbf{x}_0(r_0,\theta_0)}^{\mathbf{x}_1(r_1,\theta_1)} \sqrt{1+y'^2}\, dx \end{aligned}$$

which is just the functional for finding shortest paths in the plain!

Example

Given that $f(r,r') = \sqrt{r^2 + r'^2}$ does not depend explicitly on θ we can construct the constant function

$$H(r,r') = r'\frac{\partial f}{\partial r'} - f = \frac{r'^2}{\sqrt{r^2+r'^2}} - \sqrt{r^2+r'^2} = const$$

which we can rearrange to get $r' = r\sqrt{c_1^2 r^2 - 1}$ which we can rearrange to get

$$\theta = \int \frac{dr}{c_1 r^2 \sqrt{1 - 1/c_1^2 r^2}}$$

and integrate to get

$$\theta + c_2 = -\sin^{-1}\left(\frac{1}{c_1 r}\right) \quad \text{or} \quad Ar\cos(\theta) + Br\sin(\theta) = C$$

Special case 4

When $f = A(x,y)y' + B(x,y)$ we call this a degenerate case, because the E-L equations reduce to

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

but we can't necessarily solve these, and when they are true, the functional's value only depends on the end-points, not the actual shape of the curve.

Degenerate cases

Take $f = A(x,y)y' + B(x,y)$, so that the functional (for which we are looking for extrema) is

$$F\{y\} = \int_{x_0}^{x_1} A(x,y)y' + B(x,y)\, dx$$

Then the Euler-Lagrange equation can be written as

$$\begin{aligned} \frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} &= 0 \\ \frac{d}{dx}A(x,y) - \left[y'\frac{\partial A}{\partial y} + \frac{\partial B}{\partial y}\right] &= 0 \\ \frac{\partial A}{\partial x} + y'\frac{\partial A}{\partial y} - \left[y'\frac{\partial A}{\partial y} + \frac{\partial B}{\partial y}\right] &= 0 \end{aligned}$$

Degenerate cases

So the extremals for

$$F\{y\} = \int_{x_0}^{x_1} A(x,y)y' + B(x,y)\,dx$$

satisfy

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

This is not even a differential equation!

- may or may not have solutions depending on A and B
- no arbitrary constants, so can't impose conditions
- maybe true everywhere?

Degenerate cases

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 0$$

Where there is a solution, there exists a function $\phi(x,y)$ such that

$$\frac{\partial \phi}{\partial y} = A$$
$$\frac{\partial \phi}{\partial x} = B$$

Thus,

$$\frac{\partial A}{\partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial B}{\partial y}$$

Degenerate cases

In this case, the integrand $f(x,y)$ can be written

$$f = \frac{\partial \phi}{\partial y} y' + \frac{\partial \phi}{\partial x} = \frac{d\phi}{dx}$$

So the functional can be written

$$\begin{aligned} F\{y\} &= \int_{x_0}^{x_1} f(x,y,y')\,dx \\ &= \int_{x_0}^{x_1} \frac{d\phi}{dx}\,dx \\ &= [\phi(x,y)]_{x_0}^{x_1} \\ &= \phi(x_1,y(x_1)) - \phi(x_0,y(x_0)) \end{aligned}$$

So the functional depends only on the end-points!

Example

Let $f(x,y,y') = (x^2+3y^2)y' + 2xy$ so the functional is

$$F\{y\} = \int_{x_0}^{x_1} \left[(x^2+3y^2)y' + 2xy\right]\,dx$$

Then $A(x,y) = (x^2+3y^2)$ and $B(x,y) = 2xy$, so the E-L equation reduces to

$$\frac{\partial A}{\partial x} - \frac{\partial B}{\partial y} = 2x - 2x = 0$$

which is always true, for any curve y!

this is what we mean by an identity

Hence the Euler-Lagrange equation is always satisfied.

Example

If we choose $\phi(x,y) = x^2y + y^3 + k$ then

$$\begin{aligned}\frac{\partial\phi}{\partial y} &= x^2 + 3y^2 = A \\ \frac{\partial\phi}{\partial x} &= 2xy = B\end{aligned}$$

So the functional is determined by the end-points, e.g.

$$F\{y\} = x_1^2y_1 + y_1^3 - x_0^2y_0 - y_0^3$$

and this does not depend on the curve between the two end points.

Theorem

Suppose that the functional F satisfies the conditions of such that its extremals satisfy the Euler-Lagrange equation, which in this case reduces to an identity. Then the integrand must be linear in y', and the value of the functional is independent of the curve y (except through the end-points).

Basically this says that the degenerate case above only occurs for $f(x,y,y') = A(x,y)y' + B(x,y)$.

Variational Methods & Optimal Control

lecture 09

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 09

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Extensions

Now we consider extensions to the simple E-L equations presented so far:

- when f includes higher-order derivatives, e.g., $f(x,y,y',y'')$, e.g., the shape of a bent bar.
- when there are several dependent variables (i.e., y is a vector), e.g., calculating a particles trajectory.
- when there are several independent variables (i.e., x is a vector), e.g. calculating extremal surface.

Extension 1: higher-order derivatives

When f includes higher-order derivatives then the E-L equations can be extended, e.g., if the function includes a y'' term, i.e., $f(x,y,y',y'')$, then

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 0$$

but now we now need extra edge conditions. A simple example we will consider is the shape of a bent bar.

Standard Euler-Lagrange equation

Theorem 2.2.1: Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0,x_1]$

$$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0}$$

Higher-order derivatives

Let $F : C^2[x_0,x_1] \to \mathbb{R}$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, y', and y'', and $x_0 < x_1$. As before, the necessary condition for the extremum is that the first variation be zero, e.g.

$$\delta F(\eta,y) = 0$$

Taylor's theorem

As before we perturb y to get $\hat{y} = y + \varepsilon\eta$
Once again we apply Taylor's theorem to derive

$$f(x,y+\varepsilon\eta,y'+\varepsilon\eta',y''+\varepsilon\eta'') = f(x,y,y',y'') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''}\right] + O(\varepsilon^2)$$

and hence that

$$F\{y+\varepsilon\eta\} = \int_{x_0}^{x_1} f(x,y,y',y'') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''}\right] dx + O(\varepsilon^2)$$

First Variation

So, now the first variation will be given by

$$\begin{aligned}
\delta F(\eta,y) &= \lim_{\varepsilon\to 0}\frac{F\{y+\varepsilon\eta\} - F\{y\}}{\varepsilon} \\
&= \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \eta''\frac{\partial f}{\partial y''}\right] dx \\
&= \left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} + \left[\eta'\frac{\partial f}{\partial y''}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} - \eta\frac{d}{dx}\frac{\partial f}{\partial y'} - \eta'\frac{d}{dx}\frac{\partial f}{\partial y''}\right] dx \\
&= \left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} + \left[\eta'\frac{\partial f}{\partial y''}\right]_{x_0}^{x_1} - \left[\eta\frac{d}{dx}\frac{\partial f}{\partial y''}\right]_{x_0}^{x_1} \\
&\quad + \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} - \eta\frac{d}{dx}\frac{\partial f}{\partial y'} + \eta\frac{d^2}{dx^2}\frac{\partial f}{\partial y''}\right] dx
\end{aligned}$$

New boundary conditions

We require new fixed-end point conditions

$$\begin{aligned}
y(x_0) &= y_0 & \quad y(x_1) &= y_1 \\
y'(x_0) &= y'_0 & \quad y'(x_1) &= y'_1
\end{aligned}$$

which implies that

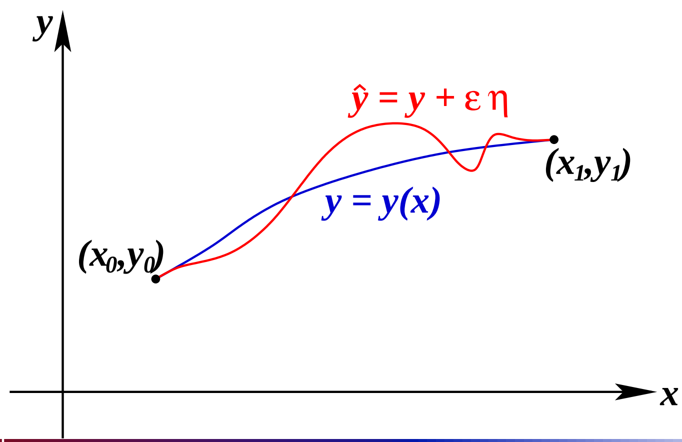
$$\begin{aligned}
\eta(x_0) &= 0 & \quad \eta(x_1) &= 0 \\
\eta'(x_0) &= 0 & \quad \eta'(x_1) &= 0
\end{aligned}$$

Which gives

$$\delta F(\eta,y) = \int_{x_0}^{x_1} \eta\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''}\right] dx$$

Fixing the end-points

We now fix the derivative and value of y at the end points.



4th Order Euler-Lagrange equation

$\delta F(\eta, y) = 0$ for arbitrary η satisfying the boundary conditions, so the result is the 4th order Euler-Lagrange equation

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 0$$

This is a 4th order differential equation.

Generalization

Let $F : C^2[x_0, x_1] \to I\!R$ be a functional of the form

$$F\{y\} = \int_{x_0}^{x_1} f(x, y, y', \ldots, y^{(n)})\, dx,$$

where f has continuous partial derivatives of second order with respect to $x, y, y', \ldots, y^{(n)}$, and $x_0 < x_1$, and the values of $y, y', \ldots, y^{(n-1)}$ are fixed at the end-points, then the extremals satisfy the condition

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} + \cdots + (-1)^n \frac{d^n}{dx^n}\frac{\partial f}{\partial y^{(n)}} = 0$$

This is sometimes called the **Euler-Poisson Equation**.

Example 1

$$F\{y\} = \int_0^1 (1 + y''^2)\, dx$$

subject to $y(0) = 0, y(1) = 1, y'(0) = 1, y'(1) = 1$

$$\begin{aligned} \frac{\partial f}{\partial y} &= 0 \\ \frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\ \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} &= \frac{d^2}{dx^2} 2y'' = 2\frac{d^4 y}{dx^4} \end{aligned}$$

Example 1 (cont)

The E-P equation gives

$$\frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 2\frac{d^4y}{dx^4} = 0$$

The solution is

$$y(x) = c_1 + c_2x + c_3x^2 + c_4x^3$$

Given the end-points

$$y(0) = 0 \Rightarrow c_1 = 0$$
$$y'(0) = 1 \Rightarrow c_2 = 1$$
$$y(1) = 1 \Rightarrow c_2 + c_3 + c_4 = 1$$
$$y'(1) = 1 \Rightarrow c_2 + 2c_3 + 3c_4 = 1$$

Final solution is $y(x) = x$

Example 2

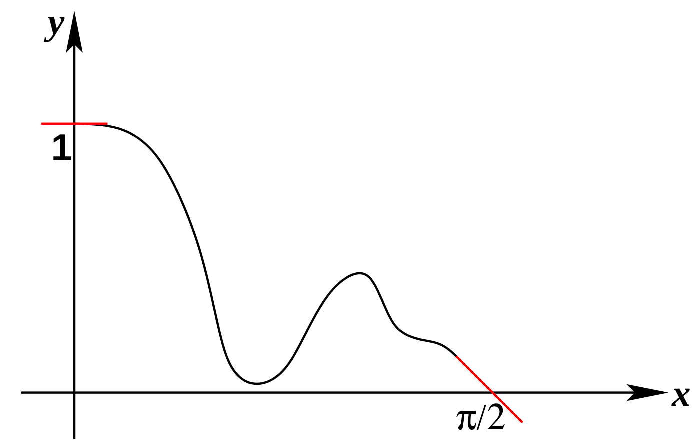
$$F\{y\} = \int_0^{\pi/2} \left(y''^2 - y^2 + x^2\right)\,dx$$

subject to $y(0) = 1, y(\pi/2) = 0, y'(0) = 0, y'(\pi/2) = -1$

$$\begin{aligned} \frac{\partial f}{\partial y} &= -2y \\ \frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\ \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} &= 2\frac{d^4y}{dx^4} \end{aligned}$$

Notice the x^2 doesn't influence the form of extremal!

Example 2 (cont)



Example 2 (cont)

The E-P equation gives

$$\frac{\partial f}{\partial y} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = -2y + 2\frac{d^4y}{dx^4} = 0$$

The solution is

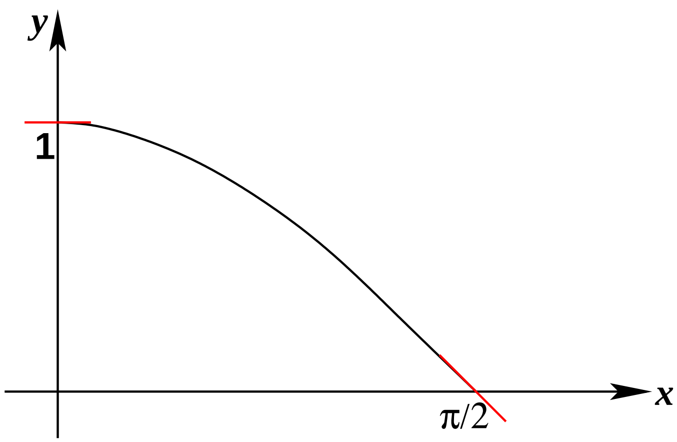
$$y(x) = Ae^x + Be^{-x} + C\sin x + D\cos x$$

Given the end-points

$$y(0) = 1 \Rightarrow A + B + D = 1$$
$$y'(0) = 0 \Rightarrow A - B + C = 0$$
$$y(\pi/2) = 0 \Rightarrow Ae^{\pi/2} + Be^{-\pi/2} + C = 0$$
$$y'(\pi/2) = -1 \Rightarrow Ae^{\pi/2} - Be^{-\pi/2} - D = -1$$

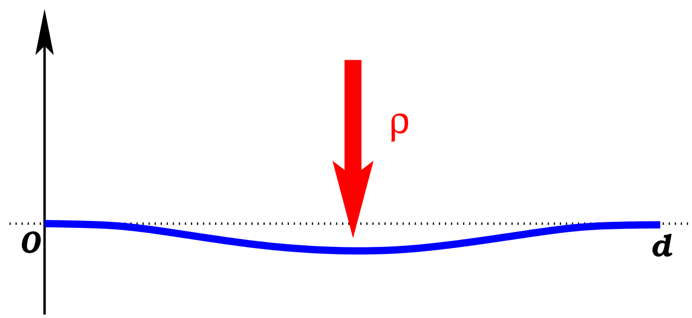
Example 2 (solution)

$y(x) = \cos(x)$



Example 3

Bent elastic beam.



Two end-points are fixed, and clamped so that they are level, e.g. $y(0) = 0, y'(0) = 0$, and $y(d) = 0$ and $y'(d) = 0$. The load (per unit length) on the beam is given by a function $\rho(x)$.

Example 3

Let $y : [0, d] \to \mathbb{R}$ describe the shape of the beam, and $\rho : [0, d] \to \mathbb{R}$ be the load per unit length on the beam. For a bent elastic beam the potential energy from elastic forces is

$$V_1 = \frac{\kappa}{2} \int_0^d y''^2 \, dx, \qquad \kappa = \text{flexural rigidity}$$

The potential energy is

$$V_2 = -\int_0^d \rho(x) y(x) \, dx$$

Thus the total potential energy is

$$V = \int_0^d \frac{\kappa y''^2}{2} - \rho(x) y(x) \, dx$$

Example 3

The Euler-Lagrange equation is

$$\begin{aligned} \frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} &= 0 \\ -\rho(x) + \kappa y^{(4)} &= 0 \\ y^{(4)} &= \frac{\rho(x)}{\kappa} \end{aligned}$$

This DE has solution

$$y(x) = P(x) + c_3 x^3 + c_2 x^2 + c_1 x + c_0$$

where the c_k's are the constants of integration, and $P(x)$ is a particular solution to $P^{(4)}(x) = \rho(x)/\kappa$.

Example 3: uniform load

If the beam is uniformly loaded, then $\rho(x) = \rho$ and so

$$y(x) = \frac{\rho x^4}{4!\kappa} + c_3 x^3 + c_2 x^2 + c_1 x + c_0$$

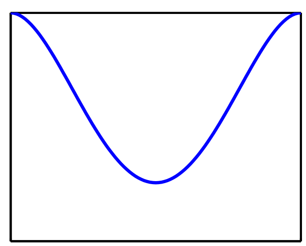
The end-conditions imply

$$\begin{aligned}
y(0) &= 0 \quad \Rightarrow c_0 = 0 \\
y'(0) &= 0 \quad \Rightarrow c_1 = 0 \\
y(d) &= 0 \quad \Rightarrow \frac{\rho d^4}{4!\kappa} + c_0 + c_1 d + c_2 d^2 + c_3 d^3 = 0 \\
y'(d) &= 0 \quad \Rightarrow \frac{\rho d^3}{3!\kappa} + c_1 + 2c_2 d + 3c_3 d^2 = 0
\end{aligned}$$

Example 3: uniform load

Choose a solution of the form

$$y(x) = \frac{\rho(d-x)^2 x^2}{24\kappa}$$



Then the derivative

$$y'(x) = \frac{2\rho(d-x)x^2}{12\kappa} + \frac{\rho(d-x)^2 x}{12\kappa}$$

We can see that the constraints are satisfied

$$\begin{aligned}
y(0) &= 0 \\
y'(0) &= 0 \\
y(d) &= 0 \\
y'(d) &= 0
\end{aligned}$$

Example 3: uniform load

$$\tilde{y}(x) = -\frac{\rho(d-x)^2 x^2}{24\kappa}$$

Maximum displacement occurs at $x = d/2$, and is given by

$$\tilde{y}(d/2) = -\frac{\rho d^4}{384\kappa}$$

Contrast this with the catenary.

$$\tilde{y}(x) = c_1 \cosh\left(\frac{x - c_2}{c_1}\right)$$

where c_1 and c_2 are determined by the end-points (there are no physical values such as m or g in the solution).

Variational Methods & Optimal Control

lecture 10

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 10

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Extension 2: several dependent variables

When there are several dependent variables, i.e., y is a vector, then the E-L equations generalize to give one DE per dependent variable. A simple example is when we calculate the trajectory of a particle in 3D. This section introduces a number of physics ideas/principles: potentials, Lagrangians, Hamilton's principle, Newton's laws of motion, and conservations laws.

Extension

Several dependent variables

- in prior problem formulations, we have only one dependent variable y, which is dependent on x, e.g. $y = y(x)$.
- we can extend this to many dependent variables q_i
- a typical example might be the position of a particle in 3D space with respect to time, e.g. $(x(t), y(t), z(t))$
- the particle has three dependent variables x, y and z

Definitions

Define $\mathbf{C}^2[t_0,t_1]$ to denote the set of vector functions $\mathbf{q} : [t_0,t_1] \to \mathbb{R}^n$, such that for $\mathbf{q} = (q_1, q_2, \ldots, q_n)$ its component functions $q_k \in C^2[t_0,t_1]$ for $k = 1, 2, \ldots, n$.

- i.e. take a set of n functions $q_k(t)$, with two continuous derivatives with respect to t, and put them into a vector $\mathbf{q}(t)$
- dot notation:

$$\dot{q}_k = \frac{dq_k}{dt}, \quad \ddot{q}_k = \frac{d^2 q_k}{dt^2} \quad \text{and} \quad \dot{\mathbf{q}} = \left(\frac{dq_1}{dt}, \frac{dq_2}{dt}, \ldots, \frac{dq_n}{dt} \right)$$

- we can define norms on the space $\mathbf{C}^2[t_0,t_1]$, e.g.

$$\|\mathbf{q}\| = \max_{k=1,\ldots,n} \sup_{t \in [t_0,t_1]} |q_k(t)|$$

Functionals

We can define functionals, for example

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

where we choose the function L to have continuous 2nd-order derivatives with respect to t, q_k and $\dot{q}_k$, for $k = 1, \ldots, n$.

For the fixed end-point problem, we look for $\mathbf{q} \in S$, where

$$S = \{\mathbf{q} \in \mathbf{C}_2^n[t_0, t_1] | \mathbf{q}(t_0) = \mathbf{q}_0, \mathbf{q}(t_1) = \mathbf{q}_1\}$$

Extremals

As before, we look for extremals by examining perturbations of $\mathbf{q}$, and seeing their effect on the functional, e.g. take the perturbation

$$\hat{\mathbf{q}} = \mathbf{q} + \varepsilon \mathbf{n}$$

where $\mathbf{n} \in \mathcal{H}^n$, where

$$\mathcal{H} = \left\{ n_i \in \mathbf{C}^2[t_0, t_1] | n_i(t_0) = 0, n_i(t_1) = 0 \right\}$$

For instance, for a local minima, we require

$$F\{\mathbf{q} + \varepsilon \mathbf{n}\} \geq F\{\mathbf{q}\}$$

for all $\mathbf{n} \in \mathcal{H}^n$ and $\mathbf{q} + \varepsilon \mathbf{n}$ in a small neighborhood of $\mathbf{q}$ with respect to some distance metric.

Applying Taylor's theorem

Taylor's theorem (again)

$$f(\mathbf{x} + \delta \mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{n} \delta x_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i x_j} \delta x_i \delta x_j + O(\delta \mathbf{x}^3)$$

Applying with $\mathbf{x} = (t, \mathbf{q}, \dot{\mathbf{q}})$, and $\delta \mathbf{x} = (0, \varepsilon \mathbf{n}, \varepsilon \dot{\mathbf{n}})$

$$L(t, \mathbf{q} + \varepsilon \mathbf{n}, \dot{\mathbf{q}} + \varepsilon \dot{\mathbf{n}}) = L(t, \mathbf{q}, \dot{\mathbf{q}}) + \varepsilon \sum_{k=1}^{n} \left(n_k \frac{\partial L}{\partial q_k} + \dot{n}_k \frac{\partial L}{\partial \dot{q}_k} \right) + O\left(\varepsilon^2\right)$$

Deriving the Euler-Lagrange eq.s

As before the **First Variation** is

$$\begin{aligned}
\delta F(\mathbf{n}, \mathbf{q}) &= \frac{F\{\mathbf{q} + \varepsilon \mathbf{n}\} - F\{\mathbf{q}\}}{\varepsilon} \\
&= \frac{1}{\varepsilon} \int_{t_0}^{t_1} L(t, \mathbf{q} + \varepsilon \mathbf{n}, \dot{\mathbf{q}} + \varepsilon \dot{\mathbf{n}}) - L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt \\
&= \int_{t_0}^{t_1} \sum_{k=1}^{n} \left(n_k \frac{\partial L}{\partial q_k} + \dot{n}_k \frac{\partial L}{\partial \dot{q}_k} \right) dt + O(\varepsilon) \\
&= 0
\end{aligned}$$

for all $\mathbf{n} \in \mathcal{H}^n$ as $\varepsilon \to 0$.

This is still a little too hard for us

Deriving the Euler-Lagrange eq.s

Note the above must be true for all $\mathbf{n} \in \mathcal{H}^n$.

We can simplify by choosing: $\mathbf{n}_1 = (n_1, 0, 0, \ldots, 0)$.

Then the First Variation simplifies

$$\begin{aligned} \delta F(\mathbf{n}_1, \mathbf{q}) &= \int_{t_0}^{t_1} \sum_{k=1}^{n} \left(n_k \frac{\partial L}{\partial q_k} + \dot{n}_k \frac{\partial L}{\partial \dot{q}_k} \right) dt \\ &= \int_{t_0}^{t_1} \left(n_1 \frac{\partial L}{\partial q_1} + \dot{n}_1 \frac{\partial L}{\partial \dot{q}_1} \right) dt \end{aligned}$$

We integrate the term $\dot{n}_1 \frac{\partial L}{\partial \dot{q}_1}$ by parts as in the derivation of the simple Euler-Lagrange equation and we get

Deriving the Euler-Lagrange eq.s

$$\delta F(\mathbf{n}_1, \mathbf{q}) = \int_{t_0}^{t_1} n_1 \left(\frac{\partial L}{\partial q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} \right) dt$$

For an extremal we want $\delta F(\mathbf{n}_1, \mathbf{q}) = 0$
for all $n_1 \in \mathcal{H} = \{C^2[t_0, t_1] | n_1(t_0) = 0, n_1(t_1) = 0\}$
Applying Lemma 2.2.2 gives

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} = 0$$

This is directly analogous to the original Euler-Lagrange equation.

Deriving the Euler-Lagrange eq.s

We can do likewise for

$$\mathbf{n}_k = (0, 0, \ldots, 0, n_k, 0, \ldots, 0)$$

in exactly the same fashion to obtain a set of equations

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} &= 0 \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} - \frac{\partial L}{\partial q_2} &= 0 \\ &\vdots \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_n} - \frac{\partial L}{\partial q_n} &= 0 \end{aligned}$$

The result is analogous to maximizing a function of several variables, where we must set all of the partial derivatives $\partial f / \partial x_k = 0$.

Simple example

Find extremals of

$$F\{\mathbf{q}\} = \int_0^1 \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2 \right) dt$$

for $\mathbf{q}(0) = \mathbf{q}_0$ and $\mathbf{q}(1) = \mathbf{q}_1$

The Euler-Lagrange equations are

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} &= 0 \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} - \frac{\partial L}{\partial q_2} &= 0 \end{aligned}$$

Simple example

$$L = \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2\right)$$

So

$$\begin{aligned} \frac{\partial L}{\partial q_1} &= 2q_1 + q_2, & \frac{\partial L}{\partial q_2} &= q_1 \\ \frac{\partial L}{\partial \dot{q}_1} &= 2\dot{q}_1, & \frac{\partial L}{\partial \dot{q}_2} &= 2(\dot{q}_2 - 1) \end{aligned}$$

So the E-L equations are

$$\begin{aligned} 2\ddot{q}_1 - 2q_1 - q_2 &= 0 \\ 2\ddot{q}_2 - q_1 &= 0 \end{aligned}$$

Simple example

Differentiate the second equation twice with respect to t to get

$$2q_2^{(4)} - \ddot{q}_1 = 0$$

which we rearrange to get $\ddot{q}_1 = 2q_2^{(4)}$, which we can substitute (along with the second equation $q_1 = 2\ddot{q}_2$) into the first equation to get a 4th order DE for q_2, e.g.

$$4q_2^{(4)} - 4\ddot{q}_2 - q_2 = 0$$

Simple example

The forth order linear ODE

$$2q_2^{(4)} - 2\ddot{q}_2 - \frac{1}{2}q_2 = 0$$

has characteristic equation

$$2\mu^4 - 2\mu^2 - 1/2 = 0$$

which has roots

$$\begin{aligned} \mu_1, \mu_2 &= \pm\sqrt{\frac{1}{2} + \frac{1}{\sqrt{2}}} \\ \mu_3, \mu_4 &= \pm\sqrt{\frac{1}{2} - \frac{1}{\sqrt{2}}} = \pm im \end{aligned}$$

Simple example

The solution is

$$q_2(t) = c_1 e^{\mu_1 t} + c_2 e^{\mu_2 t} + c_3 \cos(mt) + c_4 \sin(mt)$$

where c_1, c_2, c_3 and c_4 are determined by the 4 end-point conditions $\mathbf{q}(0) = \mathbf{q}_0$ and $\mathbf{q}(1) = \mathbf{q}_1$.

We can determine q_1 from

$$q_1 = 2\ddot{q}_2 = 2c_1\mu_1^2 e^{\mu_1 t} + 2c_2\mu_2^2 e^{\mu_2 t} - 2c_3 m^2 \cos(mt) - 2c_4 m^2 \sin(mt)$$

Example: movement of a particle

The **kinetic energy** of a particle is

$$T = \frac{1}{2}mv^2(t) = \frac{1}{2}m\left(\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t)\right)$$

where $v(t)$ is the speed of the particle at time t.

Assume there exists a scalar function of time and position $V(t,x,y,z)$, such that the forces acting on the particle are

$$f_x = -\frac{\partial V}{\partial x},\ f_y = -\frac{\partial V}{\partial y},\ f_z = -\frac{\partial V}{\partial z}$$

Then V is called the **potential energy** of the particle.

The Lagrangian

The function $L(t,x,y,x,\dot{x},\dot{y},\dot{z})$

$$L = T - V$$

is called the **Lagrangian**

The path of a particle is given by $\mathbf{r}(t) = (x(t),y(t),z(t))$ over the time interval $[t_0,t_1]$.

We can define the **action integral** by

$$F\{\mathbf{r}\} = \int_{t_0}^{t_1} L(t,\mathbf{r},\dot{\mathbf{r}})\,dt$$

Hamilton's principle

The path of a particle $\mathbf{r}(t)$ is such that the functional

$$F\{\mathbf{r}\} = \int_{t_0}^{t_1} L(t,\mathbf{r},\dot{\mathbf{r}})\,dt$$

is stationary.

- could be a saddle point (not just minima)
- note, Hamilton's principle is far more general
  - multiple particles
  - non-Cartesian coordinates
  - remember changing coordinates shouldn't change extremal curves

Generalized coordinates

We can describe the mechanical system by generalized coordinates $\mathbf{q}(t)$.

- The kinetic energy is given by $T(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\sum_{j,k=1}^{n} C_{j,k}(\mathbf{q})\dot{q}_j\dot{q}_k$
- The potential energy is given by $V(t,\mathbf{q})$
- The Lagrangian is $L(t,\mathbf{q},\dot{\mathbf{q}}) = T(\mathbf{q},\dot{\mathbf{q}}) - V(t,\mathbf{q})$

Hamilton's principle states that the path of the particle $\mathbf{q}(t)$ will be such that the functional

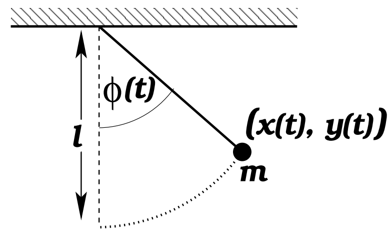
$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t,\mathbf{q},\dot{\mathbf{q}})\,dt$$

is stationary.

Example: a simple pendulum

Kinetic energy

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\phi}^2$$



Potential energy

$$V = mg(l-y) = mgl(1-\cos\phi)$$

The Lagrangian is

$$L(\phi,\dot{\phi}) = \frac{1}{2}ml^2\dot{\phi}^2 - mgl(1-\cos\phi)$$

and the action integral is

$$F\{\phi\} = \int_{t_0}^{t_1}\left(\frac{1}{2}ml^2\dot{\phi}^2 - mgl(1-\cos\phi)\right)dt$$

Hamilton's principle and EL eq.s

Hamilton's principle states we should look for curves along which the function

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t,\mathbf{q},\dot{\mathbf{q}})\,dt$$

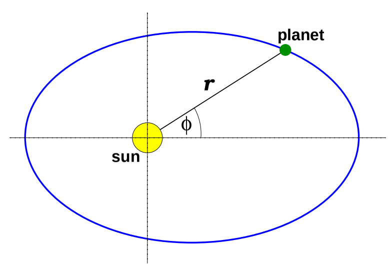
is stationary. The Euler-Lagrange equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0$$

for all $k = 1,\ldots,n$, and so for mechanical systems, the Lagrangian satisfies these equations.

Kepler's problem of planetary motion

Single planet orbiting the sun.



Kinetic energy

$$\begin{aligned} T &= \frac{1}{2}m\left(\dot{x}^2(t) + \dot{y}^2(t)\right) \\ &= \frac{1}{2}m\left(\dot{r}^2(t) + r^2(t)\dot{\phi}^2(t)\right) \end{aligned}$$

Potential energy

$$V(r) = -\int f(r)\,dr = -\frac{GmM}{r(t)}$$

where the force $f = -\frac{dV}{dr} = -\frac{GmM}{r^2}$ (from Newton)

Newton's laws

Often the potential V depends only on location and time, and the kinetic energy depends only on the derivatives of the position, then the Euler-Lagrange equations reduce to

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_k} + \frac{\partial V}{\partial q_k} = 0$$

Given kinetic energy of the form $T(\dot{\mathbf{q}}) = \frac{1}{2}m\sum_i \dot{q}_i^2$, then the EL equations become

$$m\ddot{q}_k = -\frac{\partial V}{\partial q_k} = f_k = \text{ the force in direction } k$$

We have **derived** Newton's laws of motion, i.e. $\mathbf{f} = m\mathbf{a}$ from a more general principle.

Conservation laws

If the potential does not depend on time, the Lagrangian does not explicitly depend on t and so we may form $H(\mathbf{q},\dot{\mathbf{q}})$ as before, i.e.

$$H(\mathbf{q},\dot{\mathbf{q}}) = \sum_{k=1}^{n} \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L = const$$

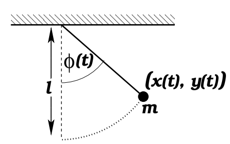
Given kinetic energy of the form $T(\dot{\mathbf{q}}) = \frac{1}{2} m \sum_i \dot{q}_i^2$, this becomes

$$H(\mathbf{q},\dot{\mathbf{q}}) = 2T - L = T + V = const$$

Thus energy is conserved in such a system.

Example: a simple pendulum

$$F\{\phi\} = \int_{t_0}^{t_1} \left(\frac{1}{2} m l^2 \dot{\phi}^2 - mgl(1-\cos\phi) \right) dt$$



The kinetic energy is in the appropriate form, and the potential does not depend on time, so the pendulum system conserves energy, e.g.

$$\frac{1}{2} m l^2 \dot{\phi}^2 + mgl(1-\cos\phi) = const$$

Removing constant terms (where possible), we get

$$\dot{\phi}^2 - \frac{2g}{l}\cos\phi = c_1$$

Example: a simple pendulum

Given conservation of energy

$$\dot{\phi}^2 - \frac{2g}{l}\cos\phi = c_1$$

To solve, differentiate with respect to t

$$2\dot{\phi}\left[\ddot{\phi} + \frac{g}{l}\sin\phi\right] = 0$$

Assume that $\dot{\phi} \neq 0$, and multiply by m, and we get

$$m\ddot{\phi} + \frac{gm}{l}\sin\phi = 0$$

which is an equation relating torque to the rate of change of angular momentum

Example: a simple pendulum

$$\ddot{\phi} + \frac{g}{l}\sin\phi = 0$$

Motion is quite complicated. Small oscillations approximation $\sin\phi \simeq \phi$ we get

$$\ddot{\phi} + \frac{g}{l}\phi = 0$$

and so

$$\phi(t) = A\sin\left(\sqrt{\frac{g}{l}}t\right) + \phi_0$$

which has period $2\pi\sqrt{\frac{l}{g}}$

Brachystochrone in 3D

Find the curve of fastest descent between the points (x_0, y_0, z_0) and (x_1, y_1, z_1) where z is height, and x and y are spatial. Consider y and z to be functions of x. The time for the descent is

$$\sqrt{2g}T\{y,z\} = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2+z'^2}}{\sqrt{z_0-z}}\,dx$$

The Euler-Lagrange equations are

$$\begin{aligned}
\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}}\right) &= 0 \\
\frac{d}{dx}\left(\frac{z'}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}}\right) - \frac{\sqrt{1+y'^2+z'^2}}{2(z_0-z)^{3/2}} &= 0
\end{aligned}$$

Brachystochrone in 3D

We can transform the first to get

$$\frac{y'}{\sqrt{1+y'^2+z'^2}} = c_1\sqrt{z_0-z}$$

but the second EL equation is a mess. Instead, note that the function f is not explicitly dependent on x, and so we may derive a function $H(y,y',z,z') = const$ as before. In this case

$$-H(y,y',z,z') = f - y'\frac{\partial f}{\partial y'} - z'\frac{\partial f}{\partial z'} = c_2$$

Brachystochrone in 3D

$$\begin{aligned}
-H(y,y',z,z') &= f - y'\frac{\partial f}{\partial y'} - z'\frac{\partial f}{\partial z'} \\
&= \frac{\sqrt{1+y'^2+z'^2}}{\sqrt{z_0-z}} - \frac{y'^2}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}} - \frac{z'^2}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}} \\
&= \frac{1+y'^2+z'^2-y'^2-z'^2}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}} \\
&= \frac{1}{\sqrt{1+y'^2+z'^2}\sqrt{z_0-z}} = c_2
\end{aligned}$$

Brachystochrone in 3D

The two parts we have derived are

$$\begin{aligned}
\frac{y'}{\sqrt{1+y'^2+z'^2}} &= c_1\sqrt{z_0-z} \\
\frac{1}{\sqrt{1+y'^2+z'^2}} &= c_2\sqrt{z_0-z}
\end{aligned}$$

Divide the first, by the second, and we get

$$y' = \frac{c_1}{c_2} = const$$

from which we derive $y = \frac{c_1}{c_2}(x-x_1)+y_1$, which is the equation of a vertical plane. Thus the solutions in 3D can be reduced to the solution to the Brachystochrone in a 2D vertical plane (which is physically obvious).

Kepler's problem of planetary motion

Single planet orbiting the sun.

$$L = T - V = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right) + \frac{GmM}{r}$$

Hamilton's principle says we have to find stationary curves of the integral of L, so we can jump straight to the E-L equations

$$\begin{aligned} \frac{\partial L}{\partial r} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}} &= 0 \\ \frac{\partial L}{\partial \phi} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} &= 0 \end{aligned}$$

Kepler's problem of planetary motion

E-L equations $L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right) + \frac{GmM}{r}$

$$\begin{aligned} \frac{\partial L}{\partial r} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}} &= 0 \\ \frac{\partial L}{\partial \phi} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} &= 0 \end{aligned}$$

give

$$\begin{aligned} mr\dot{\phi}^2 - \frac{GmM}{r^2} - m\frac{d}{dt}\dot{r} &= 0 \\ m\frac{d}{dt}r^2\dot{\phi} &= 0 \end{aligned}$$

Equations of planetary motion

Simplify (assuming $m \neq 0$ and $r \neq 0$)

$$\begin{aligned} mr\dot{\phi}^2 - \frac{GmM}{r^2} - m\frac{d}{dt}\dot{r} &= 0 \\ m\frac{d}{dt}r^2\dot{\phi} &= 0 \end{aligned}$$

to get

$$\begin{aligned} \ddot{r} - r\dot{\phi}^2 &= -\frac{GM}{r^2} \\ \dot{\phi}r^2 &= c \end{aligned}$$

Interesting aside

The equation $\dot{\phi}r^2 = c$, gives the angular velocity $\dot{\phi}$ in terms of distance from the sun, but also allows us to determine the velocity at right angles to the direction of the sun as

$$v_r = r\dot{\phi} = c/r$$

So we can calculate the angular momentum

$$p_a = rm\dot{\phi} = cm$$

which is constant (as you might expect).
The law also allows one to derive Kepler's second law (the arc of an orbit over equal periods of time traverse equal areas).

Solving the equations

First equation, including the condition $\dot{\phi} = c/r^2$ gives

$$\begin{aligned} \ddot{r} - r\dot{\phi}^2 &= -\frac{GM}{r^2} \\ \ddot{r} - \frac{c^2}{r^3} &= -\frac{GM}{r^2} \end{aligned}$$

Now instead of calculating this in terms of derivatives with respect to time, lets convert to derivatives with respect to ϕ. Denote such derivatives using, e.g., r'

$$\dot{r} = \frac{dr}{d\phi}\frac{d\phi}{dt} = r'\dot{\phi}$$

Solving the equations

From the chain rule and $\dot{\phi} = c/r^2$ we get

$$\begin{aligned} \dot{r} &= \frac{dr}{d\phi}\frac{d\phi}{dt} = r'\dot{\phi} \\ \ddot{r} &= \frac{d}{d\phi}\left(r'\dot{\phi}\right)\frac{d\phi}{dt} \\ &= \frac{d}{d\phi}\left(\frac{cr'}{r^2}\right)\dot{\phi} \\ &= \left[\frac{cr''}{r^2} - \frac{2cr'^2}{r^3}\right]\dot{\phi} \\ &= \frac{c^2}{r^2}\left[\frac{r''}{r^2} - \frac{2r'^2}{r^3}\right] \end{aligned}$$

Solving the equations

Substitute the above form of $\ddot{r}$ into the first DE and we get

$$\begin{aligned} \ddot{r} - \frac{c^2}{r^3} &= -\frac{GM}{r^2} \\ \frac{c^2}{r^2}\left[\frac{r''}{r^2} - \frac{2r'^2}{r^3}\right] - \frac{c^2}{r^3} &= -\frac{GM}{r^2} \end{aligned}$$

Once again note that $r \neq 0$, and $\dot{\phi} \neq 0$ for all but degenerate orbits (straight lines through the origin), so that we can multiply by r^2/c^2 to get

$$\frac{r''}{r^2} - \frac{2r'^2}{r^3} - \frac{1}{r} = -\frac{GM}{c^2}$$

Solving the equations

Take the substitution $u = p/r$ and then

$$\begin{aligned} u' &= -\frac{pr'}{r^2} \\ u'' &= -\frac{pr''}{r^2} + \frac{2pr'^2}{r^3} \end{aligned}$$

Now note that in our equation for r' we get

$$\begin{aligned} \frac{r''}{r^2} - \frac{2r'^2}{r^3} - \frac{1}{r} &= -\frac{GM}{c^2} \\ -\frac{u''}{p} - \frac{u}{p} &= -\frac{GM}{c^2} \\ u'' + u &= \frac{GMp}{c^2} \end{aligned}$$

Solving the equations

The equation

$$u'' + u = k$$

has a simple solution. The homogeneous form has the solution

$$u = A\cos(\phi - \omega)$$

for some constants A and ω and the particular solution is

$$u = k$$

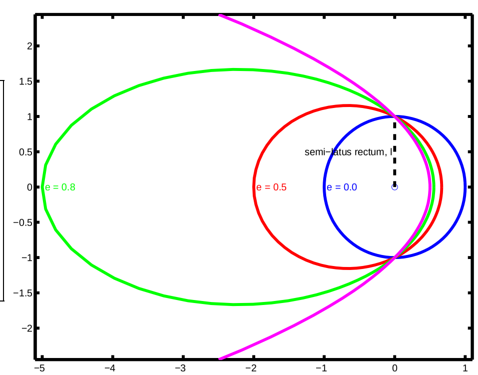
So the final solution can be scaled to give

$$\frac{L}{r} = 1 + e\cos(\phi - \omega)$$

This is just the equation of a conic section.

Possible trajectories

- $e = 0$: circle
- $0 < e < 1$: ellipse
- $e = 1$: parabola
- $e > 1$: hyperbola



L is the semi-latus rectum (dashed line), e is the eccentricity, and ω gives the angle of the perihelion (point of closest approach) which is zero in the above figure.

Variational Methods & Optimal Control

lecture 11

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 11

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Extension 3: several independent variables

When there are several independent variables, e.g., (x,y) and the extremal we wish to find represents, for instance, a surface $z(x,y)$, and f is a function $f(x,y,z(x,y),z_x,z_y)$, then the E-L equation generalizes to give

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

Several independent variables

Consider a surface minimization problem. We have a surface in 3D that is a function of (x,y), e.g. $z = z(x,y)$ then x and y are both independent variables.

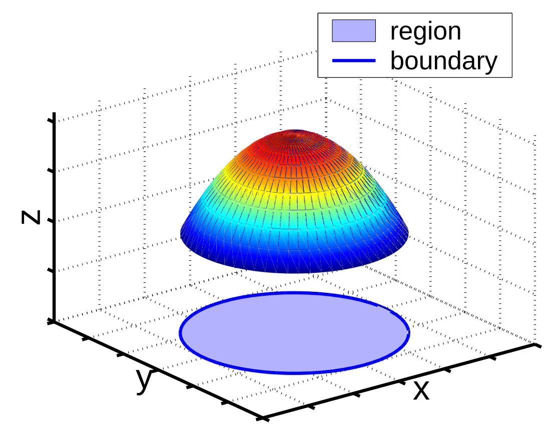
Examples:

- minimal area surfaces
  - soap films and bubbles
  - for construction
- problems of the form, minimize

$$F\{z\} = \iint_\Omega z_x^2 + z_y^2 \, dx\, dy$$

Notation

$$\begin{aligned} \text{region} &= \Omega \\ \text{boundary} &= \delta\Omega \\ \text{surface} &= z(x,y) \end{aligned}$$



Formalisms

Ω is a simply connected, bounded region of $\mathbb{R}^2$

$\delta\Omega$ is the boundary of Ω

$\bar{\Omega} = \Omega \cup \delta\Omega$ is the closure of Ω

$C^2(\bar{\Omega}) = \{z : \bar{\Omega} \to \mathbb{R} \mid z$ has 2 continuous derivatives$\}$

$C^2(\delta\Omega) = \{z_0 : \delta\Omega \to \mathbb{R} \mid z_0$ has 2 continuous derivatives$\}$

$\iint_\Omega f(x,y)\,dx\,dy$ is an area integral of f over the region Ω

$\oint_{\delta\Omega} f(x,y)\,dx$ is a contour integral around the boundary $\delta\Omega$.

The problem

Find extremals for the functional

$$F\{z\} = \iint_\Omega f(x,y,z(x,y),z_x,z_y)\,dx\,dy$$

Analogy of fixed end points is a fixed boundary, e.g.

$$z(x,y) = z_0(x,y) \text{ for all } (x,y) \in \delta\Omega$$

for some specified function $z_0 \in C^2(\delta\Omega)$.

Solution

As before we consider perturbations, though in this case they are perturbations to a surface, with fixed edge, e.g.

$$\hat{z}(x,y) = z(x,y) + \varepsilon\eta(x,y)$$

where $\eta(x,y) = 0$ for all $(x,y) \in \delta\Omega$.

Taylor's theorem gives

$$\begin{aligned} & f(x,y,z+\varepsilon\eta, z_x+\varepsilon\eta_x, z_y+\varepsilon\eta_y) \\ = \; & f(x,y,z,z_x,z_y) + \varepsilon\left[\eta\frac{\partial f}{\partial z} + \eta_x\frac{\partial f}{\partial z_x} + \eta_y\frac{\partial f}{\partial z_y}\right] + O(\varepsilon^2) \end{aligned}$$

The First Variation

As before we demand that at an extremal, the First Variation $\delta F(\eta,z) = 0$ for all possible η, and small ε

$$\begin{aligned} \delta F(\eta,z) &= \lim_{\varepsilon\to 0} \frac{F\{z+\varepsilon\eta\} - F\{z\}}{\varepsilon} \\ &= \iint_\Omega \left[\eta\frac{\partial f}{\partial z} + \eta_x\frac{\partial f}{\partial z_x} + \eta_y\frac{\partial f}{\partial z_y}\right] dx\,dy \end{aligned}$$

We next need to do the equivalent of integration by parts, but its a bit more complicated — we need to use Green's theorem.

Green's theorem

One form of Green's theorem states

$$\iint_\Omega \left(\frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y} \right) dx\,dy = \oint_{\delta\Omega} \phi\, dy - \oint_{\delta\Omega} \psi\, dx$$

for $\phi, \psi : \bar{\Omega} \to I\!R$ such that ϕ, ψ, ϕ_x and ψ_y are continuous.

This converts an area integral over a region into a line integral around the boundary.

Green's theorem in use

Green's theorem: $\iint_\Omega \left(\frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y} \right) dx\,dy = \oint_{\delta\Omega} \phi\, dy - \oint_{\delta\Omega} \psi\, dx$

For instance, take

$$\phi = \eta \frac{\partial f}{\partial z_x} \quad \text{and} \quad \psi = \eta \frac{\partial f}{\partial z_y}$$

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \eta_x \frac{\partial f}{\partial z_x} + \eta \frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} \\ \frac{\partial \psi}{\partial y} &= \eta_y \frac{\partial f}{\partial z_y} + \eta \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} \end{aligned}$$

Green's theorem in use

Green's theorem: $\iint_\Omega \left(\frac{\partial \phi}{\partial x} + \frac{\partial \psi}{\partial y} \right) dx\,dy = \int_{\delta\Omega} \phi\, dy - \int_{\delta\Omega} \psi\, dx$

So

$$\begin{aligned} \iint_\Omega \left(\eta_x \frac{\partial f}{\partial z_x} + \eta_y \frac{\partial f}{\partial z_y} + \eta \frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} + \eta \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} \right) dx\,dy \\ = \oint_{\delta\Omega} \eta \frac{\partial f}{\partial z_x} dy - \oint_{\delta\Omega} \eta \frac{\partial f}{\partial z_y} dx \end{aligned}$$

Notice that $\eta(x,y) = 0$ for all $(x,y) \in \delta\Omega$, and so the right hand side integrals are both zero.

Given the RHS of the equation was zero, we can rearrange to get

$$\iint_\Omega \left(\eta_x \frac{\partial f}{\partial z_x} + \eta_y \frac{\partial f}{\partial z_y} \right) dx\,dy = - \iint_\Omega \eta \left[\frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} + \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} \right] dx\,dy$$

With the result that the First Variation can be written

$$\delta F(\eta, z) = \iint_\Omega \eta \left[\frac{\partial f}{\partial z} - \frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} \right] dx\,dy$$

This step is the analogy of integration by parts in the derivation of the standard Euler-Lagrange equation.

Euler-Lagrange equation

Given that $F(\eta, z) = 0$ for all allowable η, Lemma 2.2.2 (see last page) can be extended directly to the 2D case, with the result that

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

This is also called the Euler-Lagrange equation.

The general case of the Euler-Lagrange equations with 2 independent variables (and the boundary conditions) produces a Dirichlet boundary value problem. these can be very hard to solve.

Simple example

Let Ω be the disk defined by $x^2+y^2<1$, and the functional of interest be

$$F\{z\} = \iint_\Omega 1 + \frac{1}{2}z_x^2 + \frac{1}{2}z_y^2 \, dx\, dy$$

with boundary conditions

$$z_0(x,y) = 2x^2 - 1$$

for all (x,y) such that $x^2+y^2=1$.

Simple example: solution

The Euler-Lagrange equation is

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

Note that in this example, f has no explicit dependence on x, y or z, and so we get

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

This equation is called Laplace's equation.

Consider the function $z = x^2 - y^2$. This satisfies Laplace's equation, and on the boundary $y^2 = 1 - x^2$, so $z = 2x^2 - 1$, which satisfies our boundary condition.

Example: vibrating string

- Imagine a taut string
  - flexible
  - uniform mass
  - small deflections

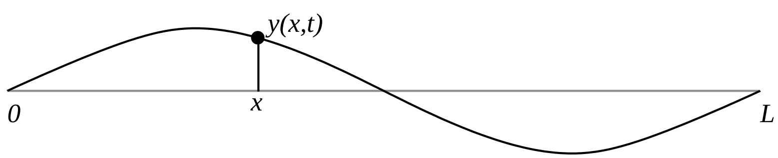
- Equilibrium solution
  - the string sits in a straight line
  - consider small perturbations

Example: vibrating string

Model:

- length of string is L
- position along the string is $x \in [0, L]$
- constant tension τ
- points on string move up/down perpendicular to x-axis
- displacement at x at time t is $w(x,t) \ll L$
- no friction or other damping
- only force occurs to stretch string
- constant density σ along the string's length

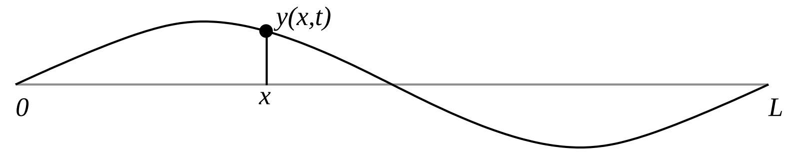
Example: vibrating string



- end points are fixed so $w(0,t) = w(L,t) = 0$
- velocity of particle is $w_t = \frac{\partial w}{\partial t}$
  - kinetic energy of string

$$T = \frac{\sigma}{2} \int_0^L w_t^2 \, dx$$

Example: vibrating string



- slope of string $\frac{\partial w}{\partial x}$
  - potential energy of the string depends on how much it is stretch from its original length L
    - ⋆ length at time t is given by

$$J(t) = \int_0^L \sqrt{1 + w_x^2} \, dx$$

Example: vibrating string

- potential is $V = \tau(J - L)$, so

$$V(t) = \tau \int_0^L \sqrt{1 + w_x^2} - 1 \, dx$$

- we assumed that w is small, so we can approximate

$$\sqrt{1 + w_x^2} \simeq 1 + \frac{1}{2} w_x^2$$

- so we use

$$V(t) = \frac{\tau}{2} \int_0^L w_x^2 \, dx$$

Example: vibrating string

The system is conservative so we apply the "principle of least action" (Hamilton's principle), which says the shape will be an extremum with respect to

$$F\{w\} = \int_{t_1}^{t_2} (T - V)\, dt = \frac{1}{2}\int_{t_1}^{t_2}\int_0^L \sigma w_t^2 - \tau w_x^2\, dt$$

The Euler-Lagrange equation is

$$\frac{\partial f}{\partial w} - \frac{\partial}{\partial x}\frac{\partial f}{\partial w_x} - \frac{\partial}{\partial t}\frac{\partial f}{\partial w_t} = 0$$

which gives

$$\frac{\partial}{\partial x}\tau w_x = \frac{\partial}{\partial t}\sigma w_t$$

Example: vibrating string

$$\frac{\partial}{\partial x}\tau w_x = \frac{\partial}{\partial t}\sigma w_t$$

or

$$\frac{\partial^2 w}{\partial x^2} = \frac{\sigma}{\tau}\frac{\partial^2 w}{\partial t^2}$$

which is the classic wave equation, which you have no doubt seen solved in other contexts.

Example: Plateau's problem

We want to find the surface with minimal area stretched between a boundary.

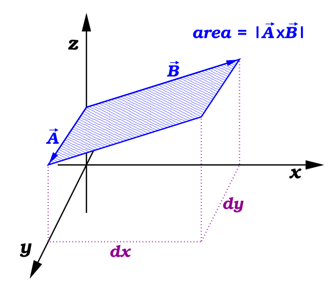
- this is what a soap film does
- architecture influenced by minimal surfaces
  - ▷ architect Frei Otto
  - ▷ Munich Olympic Stadium

Surface area minimization

The functional of interest is the surface area

$$F\{z\} = \int_\Omega dS$$

As before, we can't compute this integral, so we must convert it to a convenient form:



$$\begin{aligned}\mathbf{A} &= (0, dy, z_y\, dy)\\ \mathbf{B} &= (dx, 0, z_x\, dx)\\ \mathbf{A}\times\mathbf{B} &= (z_x\, dx\, dy, z_y\, dx\, dy, -dx\, dy)\end{aligned}$$

$$\begin{aligned}dS &= |\mathbf{A}\times\mathbf{B}| = \sqrt{(z_x\, dx\, dy)^2 + (z_y\, dx\, dy)^2 + (-dx\, dy)^2}\\ &= dx\, dy\sqrt{1 + z_x^2 + z_y^2}\end{aligned}$$

Surface area minimization

So we may rewrite the functional as

$$F\{z\} = \iint_{\Omega} \sqrt{1+z_x^2+z_y^2}\,dx\,dy$$

The Euler-Lagrange equation is

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial z_y} = 0$$

Which in this context is

$$-\frac{\partial}{\partial x}\left[\frac{z_x}{\sqrt{1+z_x^2+z_y^2}}\right] - \frac{\partial}{\partial y}\left[\frac{z_y}{\sqrt{1+z_x^2+z_y^2}}\right] = 0$$

Surface area minimization

Continuing the derivation

$$\begin{aligned}
\frac{\partial}{\partial x}\left[\frac{z_x}{\sqrt{1+z_x^2+z_y^2}}\right] &= \frac{z_{xx}}{\sqrt{1+z_x^2+z_y^2}} - \frac{z_x(z_x z_{xx}+z_y z_{yx})}{(1+z_x^2+z_y^2)^{3/2}} \\
&= \frac{z_{xx}(1+z_x^2+z_y^2) - z_x(z_x z_{xx}+z_y z_{yx})}{(1+z_x^2+z_y^2)^{3/2}} \\
&= \frac{z_{xx}(1+z_y^2) - z_x z_y z_{yx}}{(1+z_x^2+z_y^2)^{3/2}} \\
\frac{\partial}{\partial y}\left[\frac{z_y}{\sqrt{1+z_x^2+z_y^2}}\right] &= \frac{z_{yy}(1+z_x^2) - z_x z_y z_{yx}}{(1+z_x^2+z_y^2)^{3/2}}
\end{aligned}$$

Surface area minimization

Add the two terms above to get the E-L equation

$$2\mathcal{C} = \frac{z_{xx}(1+z_y^2) - 2z_x z_y z_{yx} + z_{yy}(1+z_x^2)}{(1+z_x^2+z_y^2)^{3/2}} = 0$$

where we call $\mathcal{C}$ the mean curvature (which is 0 on the extremals).

We multiply both sides of the E-L equation by the denominator to get

$$z_{xx}(1+z_y^2) - 2z_x z_y z_{yx} + z_{yy}(1+z_x^2) = 0$$

This is a hard PDE in general.

Approximate solutions

If the surfaces are almost planes (e.g. if z is small), then we can take squared derivate terms like z_x^2, z_y^2 and $z_x z_y$ to be zero. In this case the general equation

$$z_{xx}(1+z_y^2) - 2z_x z_y z_{yx} + z_{yy}(1+z_x^2) = 0$$

simplifies to give us

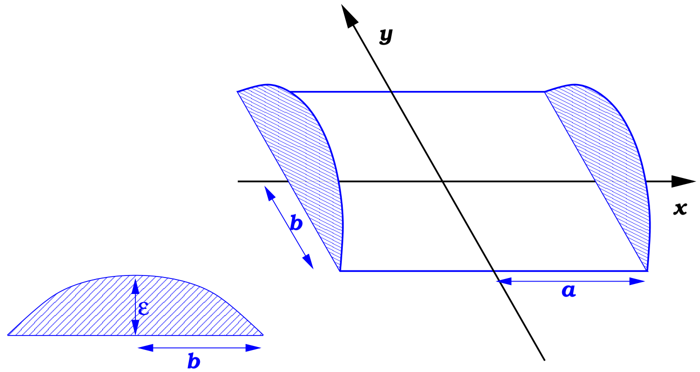
$$z_{xx} + z_{yy} = 0$$

the Laplace equation again. We know from the previous example that this is equivalent to approximating

$$f(x,y,z,z_x,z_y) = \sqrt{1+z_x^2+z_y^2} \simeq 1 + \frac{1}{2}z_x^2 + \frac{1}{2}z_y^2$$

Example

Design a surfaces of minimum surface area over a stadium with small curved walls, of shape $z = \varepsilon \cos\left(\frac{\pi}{2}\frac{y}{b}\right)$, located at $x = \pm a$, and with no end walls at $y = \pm b$.



Example

Use the approximation, so we wish to solve

$$\begin{aligned} z_{xx} + z_{yy} &= 0 \\ z(\pm a, y) &= \varepsilon \cos\left(\frac{\pi}{2}\frac{y}{b}\right) \\ z(x, \pm b) &= 0 \end{aligned}$$

Assume a solution with separation of variables, e.g. $z(x,y) = X(x)Y(y)$, then the DE implies that

$$z \propto \begin{matrix}\cosh \\ \sinh\end{matrix}(\lambda y) \times \begin{matrix}\cos \\ \sin\end{matrix}(\lambda y)$$

Choose $\cos$ with $\lambda = \frac{\pi}{2b}$ to match the boundary conditions, and choose $\cosh$ because we expect the solution to be even.

Example: solution

So the solution is

$$z(x,y) = A\cos\left(\frac{\pi y}{2b}\right)\cosh\left(\frac{\pi x}{2b}\right)$$

Determine A using the end-points, e.g.

$$\varepsilon\cos\left(\frac{\pi y}{2b}\right) = A\cos\left(\frac{\pi y}{2b}\right)\cosh\left(\frac{\pi a}{2b}\right)$$

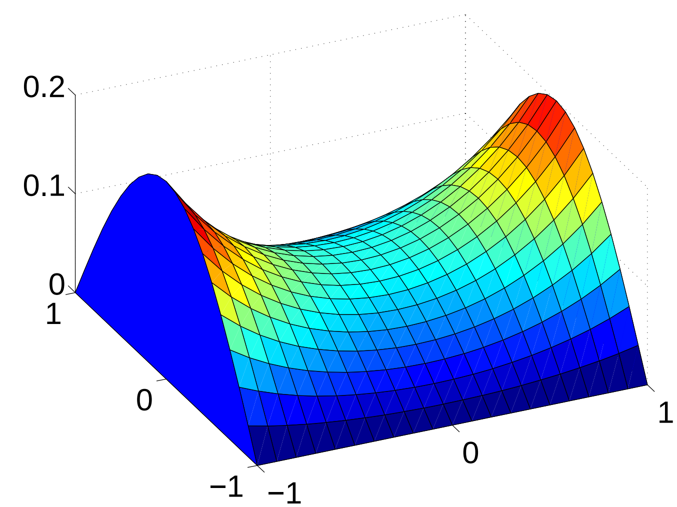
So

$$A = \varepsilon / \cosh\left(\frac{\pi a}{2b}\right)$$

and

$$z(x,y) = \varepsilon\cos\left(\frac{\pi y}{2b}\right)\cosh\left(\frac{\pi x}{2b}\right) / \cosh\left(\frac{\pi a}{2b}\right)$$

Example: solution



Example: solution

In fact, once we realize it will have a cosine cross-section, we know that the "area" of the curve for any given x will be proportional to the height, so we are in fact solving a problem that looks a lot like that of the catenary. So we should be surprised to see that the result has the same $\cosh$ function.

But this is hard...

Solving the PDE form of the EL equations can be very hard. What can we do to make it easier? Surely computers can help?

Plateau's laws

A little bit extra:

- Soap films are made of entire smooth surfaces
- The average curvature of a portion of a soap film is always constant on any point on the same piece of soap film
- Soap films always meet in threes, and they do so at an angle of $cos^{-1}(-1/2) = 120$ degrees forming an edge called a Plateau Border.
- Plateau Borders meet in fours at an angle of $\cos^{-1}(-1/3) \simeq 109.47$ degrees (the tetrahedral angle) to form a vertex.

Variational Methods & Optimal Control

lecture 12

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 12

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Numerical Solutions

The E-L equations may be hard to solve

Natural response is to find numerical methods

- Numerical solution of E-L DE
  - we won't consider these here (see other courses)
- Euler's finite difference method
- Ritz (Rayleigh-Ritz)
  - In 2D: Kantorovich's method

Euler's finite difference method

We can approximate our function (and hence the integral) onto a finite grid. In this case, the problem reduces to a standard multivariable maximization (or minimization) problem, and we find the solution by setting the derivatives to zero. In the limit as the grid gets finer, this approximates the E-L equations.

Numerical Approximation

Numerical approximation of integrals:

- use an arbitrary set of mesh points $a = x_0 < x_1 < x_2 < \cdots < x_n = b$.
- approximate

$$y'(x_i) = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \frac{\Delta y_i}{\Delta x_i}$$

- rectangle rule

$$F\{y\} = \int_a^b f(x,y,y')\,dx \simeq \sum_{i=0}^{n-1} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x_i = \bar{F}(\mathbf{y})$$

$\bar{F}(\cdot)$ is a function of the vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$.

Finite Difference Method (FDM)

Treat this as a maximization of a function of n variables, so that we require

$$\frac{\partial \bar{F}}{\partial y_i} = 0$$

for all $i = 1, 2, \ldots, n$.

Typically use uniform grid so $\Delta x_i = \Delta x = (b-a)/n$.

Simple Example

Find extremals for

$$F\{y\} = \int_0^1 \left[\frac{1}{2}y'^2 + \frac{1}{2}y^2 - y\right] dx$$

with $y(0) = 0$ and $y(1) = 0$.

E-L equations $y'' - y = 1$.

Simple Example: direct solution

E-L equations $y'' - y = -1$
Solution to homogeneous equations $y'' - y = 0$ is given by $e^{\lambda x}$ giving characteristic equation $\lambda^2 - 1 = 0$, so $\lambda = \pm 1$.
Particular solution $y = 1$
Final solution is

$$y(x) = Ae^x + Be^{-x} + 1$$

The boundary conditions $y(0) = y(1) = 0$ constrain $A + B = -1$ and $Ae + Be^{-1} = -1$, so $Ae + (1-A)e^{-1} = 1$, so $A = \frac{e^{-1}-1}{e-e^{-1}}$ and $B = \frac{1-e}{e-e^{-1}}$.
Then the exact solution to the extremal problem is

$$y(x) = \frac{e^{-1}-1}{e-e^{-1}}e^x + \frac{1-e}{e-e^{-1}}e^{-x} - 1$$

Simple Example: Euler's FDM

Find extremals for

$$F\{y\} = \int_0^1 \left[\frac{1}{2}y'^2 + \frac{1}{2}y^2 - y\right] dx$$

Euler's FDM.

- Take the grid $x_i = i/n$, for $i = 0, 1, \ldots, n$ so
  - end points $y_0 = 0$ and $y_n = 0$
  - $\Delta x = 1/n$
  - $\Delta y_i = y_{i+1} - y_i$
- So
  - $y'_i = \Delta y_i / \Delta x = n(y_{i+1} - y_i)$
  - and

$$y_i'^2 = n^2\left(y_i^2 - 2y_i y_{i+1} + y_{i+1}^2\right)$$

Simple Example: Euler's FDM

Find extremals for

$$F\{y\} = \int_0^1 \left[\frac{1}{2}y'^2 + \frac{1}{2}y^2 - y\right] dx$$

Its FDM approximation is

$$\begin{aligned}\bar{F}(\mathbf{y}) &= \sum_{i=0}^{n-1} f(x_i, y_i, y_i')\,\Delta x \\ &= \sum_{i=0}^{n-1} \frac{1}{2}n^2\left(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2\right)\Delta x + (y_i^2/2 - y_i)\Delta x \\ &= \sum_{i=0}^{n-1} \frac{1}{2}n\left(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2\right) + \frac{y_i^2/2 - y_i}{n}\end{aligned}$$

Simple Example: end-conditions

- We know the end conditions $y(0) = y(1) = 0$, which imply that

$$y_0 = y_n = 0$$

- Include them into the objective using Lagrange multipliers

$$\bar{H}(\mathbf{y}) = \sum_{i=0}^{n-1} \frac{1}{2}n\left(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2\right) + \frac{y_i^2/2 - y_i}{n} + \lambda_0 y_0 + \lambda_n y_n$$

Simple Example: Euler's FDM

Taking derivatives, note that y_i only appears in two terms of the FDM approximation

$$\begin{aligned}\bar{H}(\mathbf{y}) &= \sum_{i=0}^{n-1} \frac{1}{2}n\left(y_i^2 - 2y_iy_{i+1} + y_{i+1}^2\right) + \frac{y_i^2/2 - y_i}{n} + \lambda_0 y_0 + \lambda_n y_n \\ \frac{\partial \bar{H}(\mathbf{y})}{\partial y_i} &= \begin{cases} n(y_0 - y_1) + \frac{y_0 - 1}{n} + \lambda_0 & \text{for } i = 0 \\ n(2y_i - y_{i+1} - y_{i-1}) + \frac{y_i - 1}{n} & \text{for } i = 1, \ldots, n-1 \\ n(y_n - y_{n-1}) + \lambda_n & \text{for } i = n \end{cases}\end{aligned}$$

We need to set the derivatives to all be zero, so we now have $n+3$ linear equations, including $y_0 = y_n = 0$, and $n+3$ variables including the two Lagrange multipliers. We can solve this system numerically using, e.g., matlab.

Simple Example: Euler's FDM

Example: $n = 4$, solve

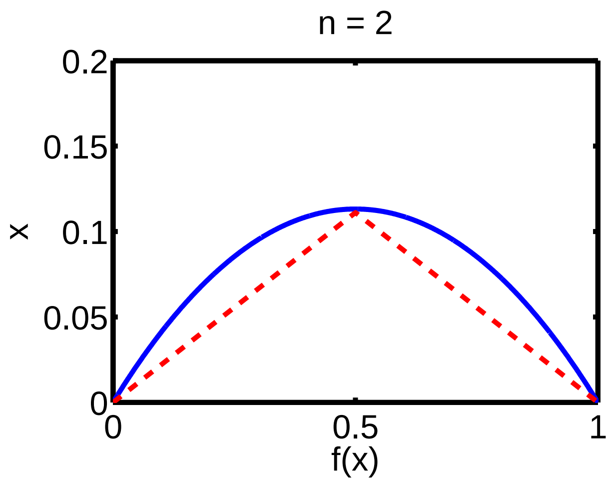
$$A\mathbf{z} = \mathbf{b}$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.062 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & -2.062 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2.062 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2.062 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2.062 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 0 \\ -0.062 \\ -0.062 \\ -0.062 \\ -0.062 \\ 0 \\ 0 \end{pmatrix}$$

- first $n+1$ terms of $\mathbf{z}$ give $\mathbf{y}$
- last two terms given the Lagrange multipliers λ_0 and λ_n

Simple example: results



Convergence of Euler's FDM

$$\bar{F}(\mathbf{y}) = \sum_{i=0}^{n-1} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)\Delta x \quad \text{and} \quad \Delta y_i = y_{i+1} - y_i$$

Only and two terms in the sum involve y_i, so

$$\begin{aligned}
\frac{\partial \bar{F}}{\partial y_i} &= \frac{\partial}{\partial y_i} f\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right) + \frac{\partial}{\partial y_i} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) \\
&= \frac{1}{\Delta x}\frac{\partial f}{\partial y_i'}\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right) \\
&\quad + \frac{\partial f}{\partial y_i}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) - \frac{1}{\Delta x}\frac{\partial f}{\partial y_i'}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) \\
&= \frac{\partial f}{\partial y_i}(x_i, y_i, y_i') - \frac{\frac{\partial f}{\partial y_i'}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) - \frac{\partial f}{\partial y_i'}\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right)}{\Delta x}
\end{aligned}$$

Convergence of Euler's FDM

The condition for a stationary point becomes

$$\frac{\partial \bar{F}}{\partial y_i} = \frac{\partial f}{\partial y_i}(x_i, y_i, y_i') - \frac{\frac{\partial f}{\partial y_i'}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) - \frac{\partial f}{\partial y_i'}\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right)}{\Delta x} = 0$$

In limit $n \to \infty$, then $\Delta x \to 0$, and so we get

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$$

which are the Euler-Lagrange equations.

- i.e., the finite difference solution converges to the solution of the E-L equations

Comments

- There are lots of ways to improve Euler's FDM
  - ▷ use a better method of numerical quadrature (integration)
    - ⋆ trapezoidal rule
    - ⋆ Simpson's rule
    - ⋆ Romberg's method
  - ▷ use a non-uniform grid
    - ⋆ make it finer where there is more variation
- We can use a different approach that can be even better

Ritz's method

In Ritz's method (called Kantorovich's methods where there is more than one independent variable), we approximate our functions (the extremal in particular) using a family of simple functions. Again we can reduce the problem into a standard multivariable maximization problem, but now we seek coefficients for our approximation.

Ritz's method

Assume we can approximate $y(x)$ by

$$y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x) + \cdots + c_n\phi_n(x)$$

where we choose a convenient set of functions $\phi_j(x)$ and find the values of c_j which produce an extremal.

For fixed end-point problem:

- Choose $\phi_0(x)$ to satisfy the end conditions.
- Then $\phi_j(x_0) = \phi_j(x_1) = 0$ for $j = 1, 2, \ldots, n$

The ϕ can be chosen from standard sets of functions, e.g. power series, trigonometric functions, Bessel functions, etc. (but must be linearly independent)

Ritz's method

- select $\{\phi_j\}_{j=0}^n$
- Approximate
  $y_n(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x) + \cdots + c_n\phi_n(x)$
- Approximate $F\{y\} \simeq F\{y_n\} = \int_{x_0}^{x_1} f(x, y_n, y_n')\, dx$
- Integrate to get $F\{y_n\} = F_n(c_1, c_2, \ldots, c_n)$
- F_n is a known function of n variables, so we can maximize (or minimize) it as usual by

$$\frac{\partial F_n}{\partial c_i} = 0$$

for all $i = 1, 2, \ldots, n$.

Upper bounds

Assume the extremal of interest is a minimum, then for the extremal

$$F\{y\} < F\{\hat{y}\}$$

for all $\hat{y}$ within the neighborhood of y. Assume our approximating function y_n is close enough to be in that neighborhood, then

$$F\{y\} < F\{y_n\} = F_n(\mathbf{c})$$

so the approximation provides an **upper bound** on the minimum $F\{y\}$.

Simple Example

Find extremals for

$$F\{y\} = \int_0^1 \left[\frac{1}{2}y'^2 + \frac{1}{2}y^2 - y\right] dx$$

with $y(0) = 0$ and $y(1) = 0$.

E-L equations $y'' - y = 1$, but we shall bypass the E-L equations to use Ritz's method.

$$y_n(x) = \phi_0(x) + \sum_{i=1}^{n} c_i \phi_i(x)$$

where we take $\phi_0(x) = 0$ and $\phi_i(x) = x^i(1-x)^i$.

Simple Example

Simple approximation $y_1 = c_1\phi_1(x)$ we get

$$F_1(c_1) = F\{y_1\} = \int_0^1 \left[\frac{1}{2}c_1^2\phi_1'^2 + c_1^2\frac{1}{2}\phi_1^2 - c_1\phi_1\right] dx$$

Now $\phi(x) = x(1-x)$ so $\phi_1' = 1 - 2x$, and

$$\begin{aligned}
F_1(c_1) &= \int_0^1 \left[\frac{c_1^2}{2}(1-2x)^2 + \frac{c_1^2}{2}x^2(1-x)^2 - c_1 x(1-x)\right] dx \\
&= \frac{c_1^2}{2}\int_0^1 \left[1 - 4x + 5x^2 - x^4\right] dx + c_1 \int_0^1 \left[-x + x^2\right] dx \\
&= \frac{c_1^2}{2}\left[x - 2x^2 + 5x^3/3 - x^5/5\right]_0^1 + c_1\left[-x^2/2 + x^3/3\right]_0^1 \\
&= \frac{c_1^2}{2}\frac{11}{30} - \frac{c_1}{6}
\end{aligned}$$

Simple Example

We solve for c_1 by setting

$$\frac{dF_1}{dc_1} = \frac{11c_1}{30} - \frac{1}{6} = 0$$

to get $c_1 = 5/11$, so the approximate extremal is

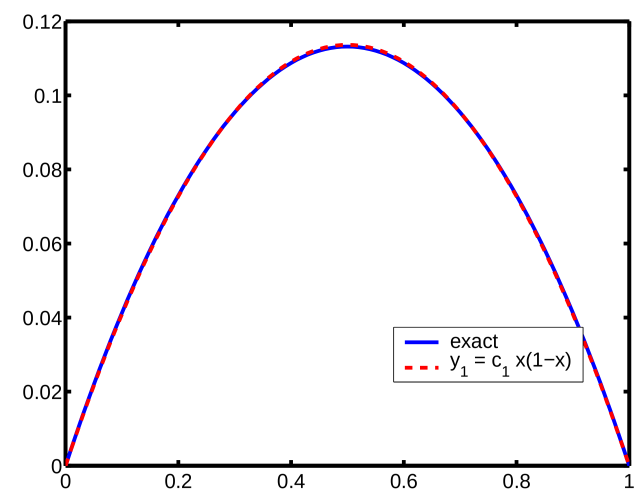
$$y_1(x) = \frac{5}{11}x(1-x)$$

The value of the approximate functional at this point is

$$F_1(5/11) = \frac{c_1^2}{2}\frac{11}{30} - \frac{c_1}{6} = -0.37879$$

which is an upper bound on the true value of the functional on the extremal.

Simple example: results



Alternate approach

Choose $\phi_1(x) = \sin(\pi x)$ (use the first element of a trigonometric series to approximate y). Then, $\phi'(x) = \pi\cos(\pi x)$, and so the functional is

$$F_1(c_1) = F\{c_1\phi_1\} = \int_0^1 \left[\frac{1}{2}c_1^2\phi_1'^2 + c_1^2\frac{1}{2}\phi_1^2 - c_1\phi_1\right] dx$$
$$= \int_0^1 \left[\frac{c_1^2\pi^2}{2}\cos^2(\pi x) + \frac{c_1^2}{2}\sin^2(\pi x) - c_1\sin(\pi x)\right] dx$$

Now $\int_0^1 \cos^2(\pi x) = \int_0^1 \sin^2(\pi x) = 1/2$,
and $\int_0^1 \sin(\pi x) = [-\frac{1}{\pi}\cos(\pi x)]_0^1 = -2/\pi$, so

$$F(c_1) = \frac{c_1^2}{2}\frac{1}{2}\left[\pi^2+1\right] - \frac{2}{\pi}c_1$$

Alternate approach

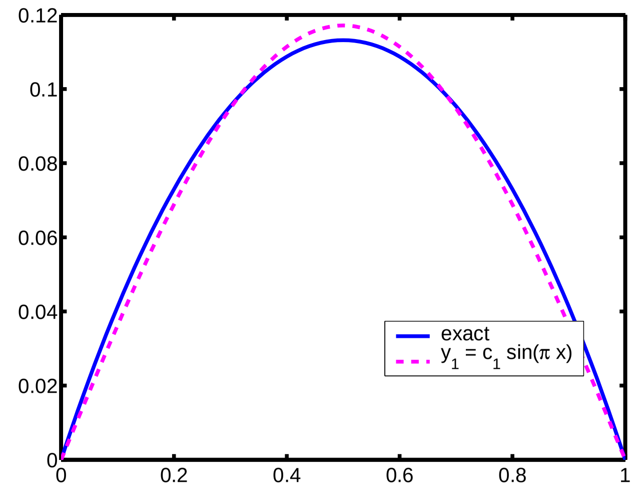
Once again we solve for c_1 by setting

$$\frac{dF_1}{dc_1} = c_1\frac{1}{2}\left[\pi^2+1\right] - \frac{2}{\pi} = 0$$

to get $c_1 = \frac{4}{\pi(\pi^2+1)}$, so the approximate extremal is

$$y_1(x) = \frac{4}{\pi(\pi^2+1)}\sin(\pi x)$$

example: alternative results



Variational Methods & Optimal Control

lecture 13

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 13

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Numerical solutions continued

Ritz applied to the catenary gives additional insights and Kantorovich's method generalizes Ritz to 2D functions..

Example: the Catenary, again

The functional of interest (the potential energy) is

$$W_p\{y\} = mg \int_{x_0}^{x_1} y\sqrt{1+y'^2}\,dx$$

Take symmetric problem with fixed end points

$$y(-1) = a \text{ and } y(1) = a$$

and we know the solution looks like

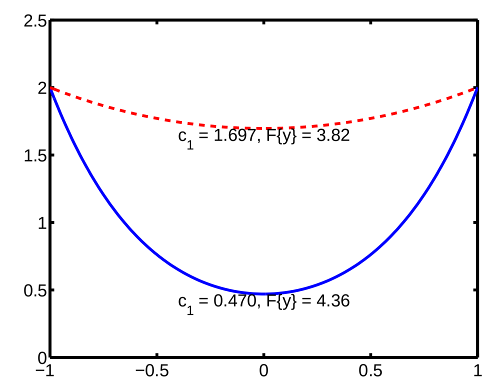
$$y(x) = c_1 \cosh\left(\frac{x}{c_1}\right)$$

where c_1 is chosen to match the end points.

Example: the Catenary, again

$y(1) = 2$ gives $c_1 = 0.47$ or $c_1 = 1.697$

- are they both local minima?



Ritz and the Catenary

Lets try approximating the curve by a polynomial

$$y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots$$

Note that symmetry of problem implies y is an even function, and hence the odd terms $a_1 = a_3 = \cdots = 0$. So, to second order we can approximate

$$y(x) \simeq a_0 + a_2 x^2$$

We have fixed $y(1) = y_1$, so we can simplify to get

$$y(x) \simeq a_0 + (y_1 - a_0)x^2$$

Ritz and the Catenary

$$\begin{aligned} y &\simeq a_0 + (y_1 - a_0)x^2 \\ y' &\simeq 2(y_1 - a_0)x \end{aligned}$$

We can substitute into the functional

$$W_p\{y\} = mg \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx$$

and integrate to get a function $W_p(a_1)$ with respect to a_0.

But this function is pretty complicated

Ritz and the Catenary

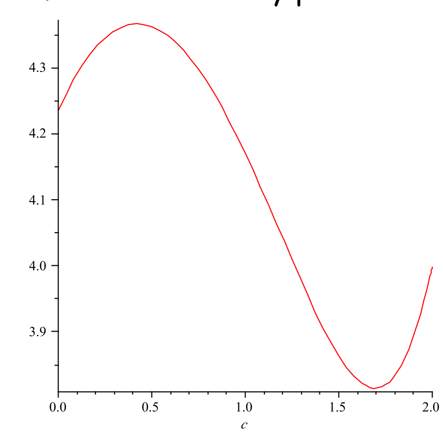
From Maple

$$\begin{aligned} W_p(a_0) = \; & -1/4\, a_0 (-8\sqrt{\pi}(4 - 4a_0 + a_0{}^2) + (-4\ln(2) - 1 - \ln(4 - 4a_0 + a_0{}^2))\sqrt{\pi} \\ & -\sqrt{\pi}(4 - 4a_0 + a_0{}^2)(-(4 - 4a_0 + a_0{}^2)^{-1} - 8) \\ & -8\sqrt{\pi}(4 - 4a_0 + a_0{}^2) sqrt(1 + (16 - 16a_0 + 4a_0{}^2)^{-1}) \\ & -1/16\, \frac{\sqrt{\pi}(128 - 128a_0 + 32a_0{}^2)\ln(1/2 + 1/2\, sqrt(1 + (16 - 16a_0 + 4a_0{}^2)^{-1}))}{4 - 4a_0 + a_0{}^2})(\sqrt{\pi})^{-1}(sqrt(4 - 4a_0 + a_0{}^2))^{-1} \\ & -1/16\,(2 - a_0)(-16\sqrt{\pi}(4 - 4a_0 + a_0{}^2)^2 - 4\sqrt{\pi}(4 - 4a_0 + a_0{}^2) \\ & -1/4\,(1/2 - 4\ln(2) - \ln(4 - 4a_0 + a_0{}^2))\sqrt{\pi} \\ & +2\sqrt{\pi}(4 - 4a_0 + a_0{}^2)^2(1/16\,(4 - 4a_0 + a_0{}^2)^{-2} + 2\,(4 - 4a_0 + a_0{}^2)^{-1} + 8) \\ & +2\sqrt{\pi}(4 - 4a_0 + a_0{}^2)^2(-(4 - 4a_0 + a_0{}^2)^{-1} - 8) sqrt(1 + (16 - 16a_0 + 4a_0{}^2)^{-1}) \\ & +1/32\, \frac{\sqrt{\pi}(64 - 64a_0 + 16a_0{}^2)\ln(1/2 + 1/2\, sqrt(1 + (16 - 16a_0 + 4a_0{}^2)^{-1}))}{4 - 4a_0 + a_0{}^2})(4 - 4a_0 + a_0{}^2)^{-3/2}\sqrt{\pi}^{-1} \end{aligned}$$

Its a pain to find the zeros of dW/da_0, but its easy to plot, and find them numerically.

Ritz and the Catenary

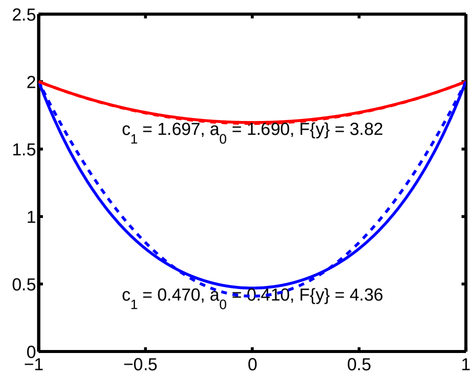
Its a function, and I can plot it, or use simple numerical techniques to find its stationary points.



Ritz and the Catenary

Stationary points

- local max: $a_0 \simeq 0.41$
- local min: $a_0 \simeq 1.69$



Ritz and the Catenary

Doesn't just give us an approximation to the extremal curves, its also gives us some insight into the nature of these extremals. If

- approximations are near to the actual extrema
- There are no other extrema so close by
- The functional is smooth (it can't have jumps either)

Then the type of extrema we get for the approximation will be the same for the real extrema, i.e.,

- local max: $a_0 \simeq 0.41 \Rightarrow$ local max for $c_1 = 0.47$
- local min: $a_0 \simeq 1.69 \Rightarrow$ local min for $c_1 = 1.697$

More than one indep. var

2D case: we are approximating a surface with series of functions, e.g.

$$z(x,y) \simeq z_n(x,y) = \phi_0(x,y) + \sum_{i=1}^{n} c_i \phi_i(x,y)$$

where $\phi_0(x,y)$ satisfies the boundary conditions, e.g. $\phi_0(x,y) = z_0(x,y)$ for $(x,y) \in \delta\Omega$, the boundary of the region on interest Ω, and the $\phi_i(x,y)$ satisfy the homogeneous boundary conditions $\phi_i(x,y) = 0$ for $(x,y) \in \delta\Omega$.

More than one indep. var

As before, we approximate the functional by

$$F\{z\} \simeq F\{z_n\} = F_n(c_1, \ldots, c_n)$$

As before we determine the c_j by requiring that the partial derivatives are zero, e.g.

$$\frac{\partial F_n}{\partial c_i} = 0$$

for all $i = 1, 2, \ldots, n$

Kantorovich's method

Approximate with

$$z(x,y) \simeq z_n(x,y) = \phi_0(x,y) + \sum_{i=1}^{n} c_i(x)\phi_i(x,y)$$

Again the ϕ_i are suitably chosen, but the c_i are no longer constants, but rather functions of one independent variable. This allows a larger class of functions to be used.

Kantorovich's method

Note that the integral function

$$F\{z_n\} = \iint_{\Omega} z_n(x,y)\,dx\,dy = \sum_{i=0}^{n} \int c_i(x) \left[\int_{y_0(x)}^{y_1(x)} \phi_i(x,y)\,dy \right] dx$$

We integrate the inner integral, and get

$$F\{z_n\} = \sum_{i=0}^{n} \int c_i(x)\Phi_i(x)\,dx$$

Now we just have a function of x, and so we may apply the Euler-Lagrange machinery.

The method approx. separates the variables x and y.

Example

Find the extremals of

$$F\{z(x,y)\} = \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z\right)\,dx\,dy$$

with $z = 0$ on the boundary.

The Euler-Lagrange equation reduces to the Poisson equation, e.g.

$$\begin{aligned}
\frac{d}{dx}\frac{\partial f}{\partial z_x} + \frac{d}{dx}\frac{\partial f}{\partial z_x} &= \frac{\partial f}{\partial z} \\
\frac{d}{dx}2z_x + \frac{d}{dx}2z_y &= -2 \\
\nabla^2 z(x,y) &= -1
\end{aligned}$$

Example

Approximate

$$z_1(x,y) = c(x)(b^2 - y^2)$$

Note $z_1(x,\pm b) = 0$ (as required) and

$$\begin{aligned}
\left(\frac{\partial z_1}{\partial x}\right)^2 &= \left(c'(x)(b^2-y^2)\right)^2 \\
&= c'(x)^2(b^4 - 2b^2y^2 + y^4) \\
\left(\frac{\partial z_1}{\partial y}\right)^2 &= (c(x)2y)^2 \\
&= 4c(x)^2y^2
\end{aligned}$$

Example

Hence, we approximate

$$
\begin{aligned}
F\{z(x,y)\} &\simeq F\{z_1(x,y)\} \\
&= \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z\right)\, dx\, dy \\
&= \int_{-a}^{a}\left[\int_{-b}^{b} \left[c'(x)^2(b^2-y^2)^2 + 4c(x)^2y^2 - 2c(x)(b^2-y^2)\right] dy\right] dx \\
&= \int_{-a}^{a}\left[c'(x)^2(b^4y - 2b^2y^3/3 + y^5/5) + 4c(x)^2y^3/3 - \right. \\
&\qquad \left. 2c(x)(b^2y - y^3/3)\right]_{-b}^{b} dx \\
&= \int_{-a}^{a}\left[\frac{16}{15}b^5c'(x)^2 + \frac{8}{3}b^3c(x)^2 - \frac{8}{3}b^3c(x)\right] dx
\end{aligned}
$$

Example

So we can write

$$F\{z(x,y)\} \simeq F\{z_1(x,y)\} = F\{c(x)\} = \int_{-a}^{a} f(x,c,c')\, dx$$

We can use the simple Euler-Lagrange equations, where

$$
\begin{aligned}
f(x,c,c') &= \frac{16}{15}b^5c'(x)^2 + \frac{8}{3}b^3c(x)^2 - \frac{8}{3}b^3c(x) \\
\frac{\partial f}{\partial c} &= \frac{16}{3}b^3c(x) - \frac{8}{3}b^3 \\
\frac{\partial f}{\partial c'} &= \frac{32}{15}b^5c'(x) \\
\frac{d}{dx}\frac{\partial f}{\partial c'} &= \frac{32}{15}b^5c''(x)
\end{aligned}
$$

Example

Euler-Lagrange equations

$$
\begin{aligned}
\frac{d}{dx}\frac{\partial f}{\partial c'} - \frac{\partial f}{\partial c} &= 0 \\
\frac{32}{15}b^5c''(x) - \frac{16}{3}b^3c(x) + \frac{8}{3}b^3 &= 0 \\
c''(x) - \frac{5}{2b^2}c(x) &= -\frac{5}{4b^2}
\end{aligned}
$$

Solutions

$$c(x) = k_1\cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + k_2\sinh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + \frac{1}{2}$$

Example

Note that the function must be zero on the boundary so $z(\pm a, y) = 0$, and so we look for an even function $c(x)$, and so $k_2 = 0$, and also $c(\pm a) = 0$, so

$$
\begin{aligned}
c(a) &= k_1\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right) + \frac{1}{2} \\
-\frac{1}{2} &= k_1\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right) \\
k_1 &= -\frac{1}{2\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}
\end{aligned}
$$

Example

Solution

$$z_1(x,y) = \frac{1}{2}(b^2 - y^2)\left(1 - \frac{\cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right)}{\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}\right)$$

If we wanted a more exact approximation, we could try

$$z_2(x,y) = (b^2 - y^2)c_1(x) + (b^2 - y^2)^2 c_2(x)$$

Lower bounds

- Obviously, quality of solution depends on
  - family of functions chosen
  - number of terms used, n
- Could test convergence by increasing n and seeing the difference in $|F\{y_{n+1}\} - F\{y_n\}|$, but this is not guaranteed to be a good indication.
- A better way to assess convergence is to have a lower-bound

$$\text{lower bound } \leq F\{y\} \leq \text{ upper bound}$$

- use **complementary variation principle**
- but its a bit complicated for us to cover here.

Variational Methods & Optimal Control

lecture 14

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 14

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Constraints

We now include additional constraints into the problems:

- Integral constraints of the form

$$\int g(x,y,y')dx = const$$

  e.g., the Isoperimetric problem.
- Holonomic constraints, e.g., $g(x,y) = 0$
- Non-holonomic constraints, e.g., $g(x,y,y') = 0$
- We won't consider inequality constraints until later.

Integral Constraints

Integral constraints are of the form

$$\int g(x,y,y')dx = const$$

The standard example of such a problem is Dido's problem, leading to us referring to such constraints as **isoperimetric**. We solve these by introducing the functional analogy of a Lagrange multiplier.

Dido's problem

- Dido (Carthaginian queen) fled to North Africa, where a local chief offered her as much land as an oxhide could contain.
- Cut the oxhide into thin strips, and then use them to surround a patch of ground (in which to found Carthage).
- Obviously, she wanted to contain the largest possible land area
- Given a fixed length of oxhide, what shape would encompass the largest area?
- Hengist and Horsa had the same problem (semi-mythological rulers in southern England around Vortigen, preceding Arthur)

Isoperimetric problems

Dido's problem falls into the class of **isoperimetric** problems.

- iso- (from same) and perimetric (from perimeter), roughly meaning "same perimeter".
- in general, such problems involve a constraint
  - e.g. the length of the oxhide strip
  - But the constraint is not always to fix the perimeter length,
  - sometimes the constraint does not even involve a length,
  - but the term isoperimetric is still used.

Isoperimetric problems formulation

We can write the isoperimetric problems as the problem of finding extremals of the functional $F : C^2[x_0, x_1] \to \mathbb{R}$ given by

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

with all the usual conditions (e.g. on end points, and continuous derivatives) but in addition we must satisfy the extra functional constraint

$$G\{y\} = \int_{x_0}^{x_1} g(x, y, y')\, dx = L$$

A simplified form of Dido's problem

Imagine that the two end-points are fixed, along the coast (Carthage was a great sea power), and we wish to encompass the largest possible area inland with a fixed length L. We can write this problem as maximize the area

$$F\{y\} = \int_{x_0}^{x_1} y\, dx$$

encompassed by the curve y, such that the the curve y has fixed length L, e.g. as before the length of the curve is

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

subject to the end-point conditions $y(x_0) = 0$ and $y(x_1) = 0$.

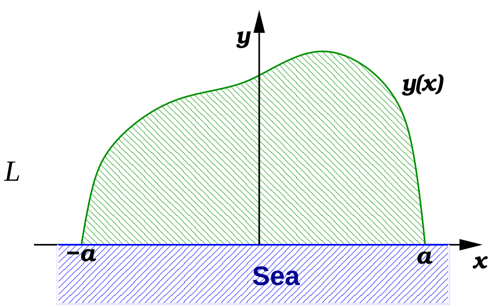
A simplified form of Dido's problem

$$F\{y\} = \int_{x_0}^{x_1} y\, dx$$

subject to

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

$y(-a) = 0$ and $y(a) = 0$.



For simplicity take $2a < L \leq \pi a$

Approach

As before

- we perturb the curve, and consider the first variation,
- but we cannot perturb by an arbitrary function $\varepsilon\eta$, because then the constraint $G\{y+\varepsilon\eta\} = L$ might be violated.
- solution: use the same approach as we did earlier with constrained maximization, e.g. use Lagrange multipliers

Lagrange multiplier refresher

Problem: find the minimum (or maximum) of $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ subject the constraints

$$g_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m < n$$

Solution requires **Lagrange Multipliers**. Minimize (or maximize) a new function (of $m+n$ variables)

$$h(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}),$$

where λ_i are the undetermined Lagrange multipliers.

Lagrange multipliers in functionals

To maximize

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

subject to

$$G\{y\} = \int_{x_0}^{x_1} g(x,y,y')\, dx = L$$

we instead consider the problem of finding extremals of

$$H\{y\} = \int_{x_0}^{x_1} h(x,y,y')\, dx = \int_{x_0}^{x_1} f(x,y,y') + \lambda g(x,y,y')\, dx$$

Euler-Lagrange equations

The Euler-Lagrange equations become

$$\boxed{\frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) - \frac{\partial h}{\partial y} = 0}$$

where $h = f + \lambda g$, and λ is the unknown Lagrange multiplier.

Dido's problem

$$H\{y\} = \int_{x_0}^{x_1} y + \lambda\sqrt{1+y'^2}\,dx$$

so

$$\frac{\partial h}{\partial y} = 1$$

$$\frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) = \frac{d}{dx}\frac{\lambda y'}{\sqrt{1+y'^2}}$$

and the Euler-Lagrange equations are

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

Dido's problem

Integrating WRT x we get

$$\frac{y'}{\sqrt{1+y'^2}} = (x+c_1)/\lambda$$

writing for the moment $\tilde{x} = (x+c_1)/\lambda$

$$\begin{aligned} y' &= \tilde{x}\sqrt{1+y'^2} \\ y'^2 &= \tilde{x}^2(1+y'^2) \\ y'^2 - \tilde{x}^2 y'^2 &= \tilde{x}^2 \\ y'^2 &= \frac{\tilde{x}^2}{1-\tilde{x}^2} \\ y' &= \tilde{x}\left(\frac{1}{1-\tilde{x}^2}\right)^{1/2} \end{aligned}$$

Dido's problem

Integrating WRT x again

$$y = \int \tilde{x}\sqrt{\frac{1}{1-\tilde{x}^2}}\,dx$$

Change variables to $\tilde{x} = (x+c_1)/\lambda = \sin(\theta)$, then

$$\begin{aligned} y &= \int \sin(\theta)\frac{1}{\cos(\theta)}\frac{dx}{d\theta}\,d\theta \\ &= \lambda\int \sin(\theta)\,d\theta \\ &= -\lambda\cos(\theta) + c_2 \end{aligned}$$

where λ, c_1 and c_2 are determined by the two end-points, and the length of the curve L.

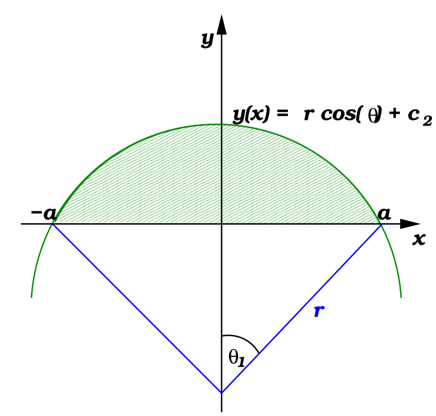
Dido's problem: constants

From the solution

$$\begin{aligned} x + c_1 &= \lambda\sin(\theta) \\ y &= -\lambda\cos(\theta) + c_2 \end{aligned}$$

we may draw a sketch of the solution, and clearly we can identify $-\lambda = r$ the radius of a circle, of which our region is a segment.

Note we deliberately started with $2a < L \leq \pi a$



Dido's problem: constants

We can see that the arc length of the enclosing curve will be

$$L = 2\theta_1 r$$

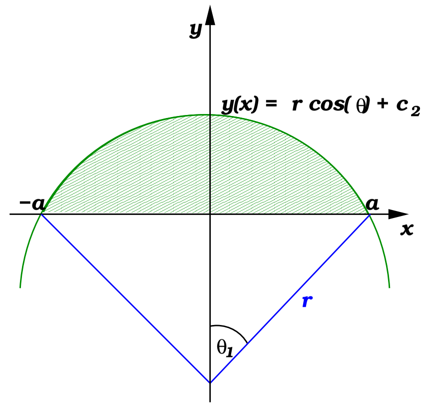
and that the value $x_1 = a$ determines that

$$r = a/\sin\theta_1$$

which combined give

$$L = 2a\theta_1/\sin\theta_1$$

from which we may determine θ_1.



Dido's problem: constants

We determine θ_1 from

$$\sin\theta_1 = \frac{2a}{L}\theta_1$$

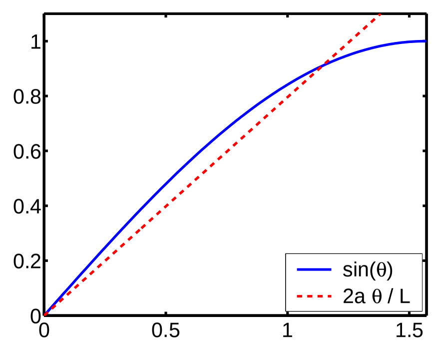
then we may compute

$$r = a/\sin(\theta_1)$$

and, we can easily see that

$$c_2 = -\cos(\theta_1)$$

from the condition that $y_1 = 0$.



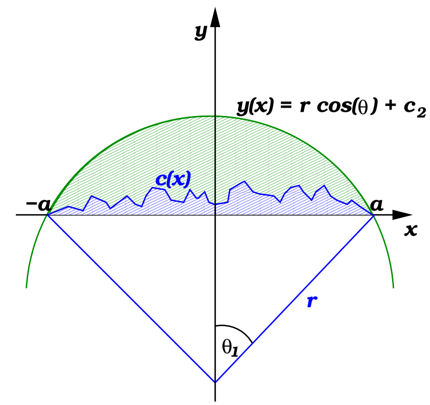
A realistic coast

What effect would a realistic coastline have?

Coast $c(x)$.

$$\text{Area} = \int_{x_0}^{x_1} y - c\,dx$$

But note that c doesn't depend on y or y', so the Euler-Lagrange equations are unchanged, provided $c(x) < y(x)$ for the extremal.



A realistic coast

Note the caveat: $c(x) < y(x)$. If this is not satisfied then the area integral includes negative components, so the problem we are maximizing is not really Dido's problem any more (she can't own negative areas).

We really want to maximize

$$\text{Area} = \int_{x_0}^{x_1} [y - c]^+ \, dx$$

where

$$[x]^+ = \begin{cases} x, & \text{for } x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

But this function does not have a derivative at $x = 0$.

The real Carthage (modern Tunis)

36.858197, 10.322342 - Google Maps http://maps.google.com.au/maps?ie=UTF8&z=5&ll=-25...

Catenary redux

We previously computed the shape of a suspended wire, when we put no constraints on the length of the wire.

What happens to the shape of the suspended wire when we fix the length of the wire?

As before we seek a minimum for the potential energy

$$W_p\{y\} = mg \int_{x_0}^{x_1} y\sqrt{1+y'^2}dx$$

but now we include the constraint that the length of the wire is L, e.g.

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

Catenary of fixed length

We instead seek extremals of

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

Notice that the above has no explicit dependence on x, and so we may compute

$$H(y,y') = y'\frac{\partial h}{\partial y'} - h = const$$

(apologies for reusing the notation H here).

Catenary of fixed length

$$H(y,y') = \frac{(y+\lambda)y'^2}{\sqrt{1+y'^2}} - (y+\lambda)\sqrt{1+y'^2} = const$$

Perform the change of variables $u = y+\lambda$, and note that $u' = y'$ so that the above can be rewritten as

$$\frac{uu'^2}{\sqrt{1+u'^2}} - u\sqrt{1+u'^2} = c_1$$

and as before (as with the earlier catenary example), this reduces to

$$\frac{u^2}{1+u'^2} = c_1^2$$

Equation of Catenary of Fixed Length

This is exactly the same equation (in u) as we had previously for the catenary in y. So the result is a catenary also, but shifted up or down by an amount such that the length of the wire is L.

$$
\begin{aligned}
y &= u - \lambda \\
&= c_1 \cosh\left(\frac{x - c_2}{c_1}\right) - \lambda
\end{aligned}
$$

so we have three constants to determine.

- we have two end points
- we have the length constraint

The length of the Catenary

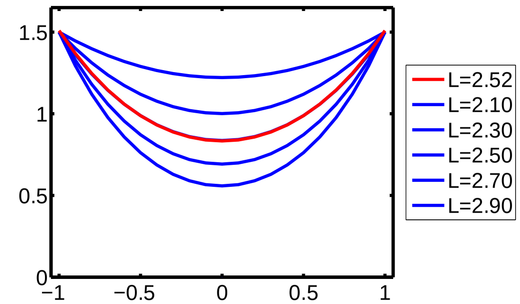
As before (taking the even solution with $y_0 = y_1$)

$$
\begin{aligned}
L\{y\} &= \int_{-1}^{1} \sqrt{1 + y'^2}\, dx \\
&= \int_{-1}^{1} \cosh(x/c_1)\, dx \\
&= c_1 \left[\sinh(x/c_1)\right]_{-1}^{1} \\
&= 2c_1 \sinh(1/c_1)
\end{aligned}
$$

But now we can use this as a constraint to calculate c_1 given L. Once we know c_1 we can calculate λ to satisfy end heights $y_0 = y_1$.

Catenary of fixed length

All catenaries are valid, but one is **natural**



The red curve shows the natural catenary (without length constraints), and the blue curves show other catenaries with different lengths

Calculating the functional

As before its easy to calculate $F\{y\}$,

$$
\begin{aligned}
F\{y\} &= \int_{-1}^{1} (c_1 \cosh(x/c_1) - \lambda) \sqrt{1 + \sinh^2(x/c_1)}\, dx \\
&= \int_{-1}^{1} c_1 \cosh^2(x/c_1) - \lambda \cosh(x/c_1)\, dx \\
&= c_1 + \frac{c_1^2}{2} \sinh(2/c_1) - 2\lambda c_1 \sinh(1/c_1)
\end{aligned}
$$

Note however, that this assumes that $y < 0$ is possible. If not, then we have to truncate y and calculate the integral numerically.

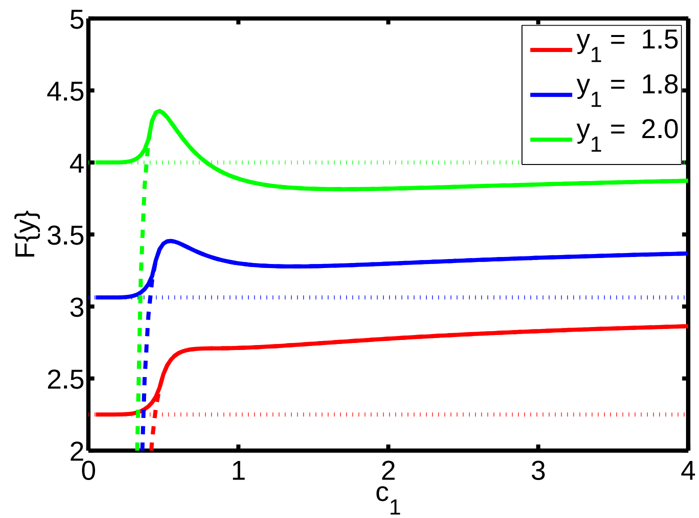
Degenerate solution

The degenerate solution has the wire lying on the ground, but we have to add in the energy of the wire leading from the pole to the ground at each end

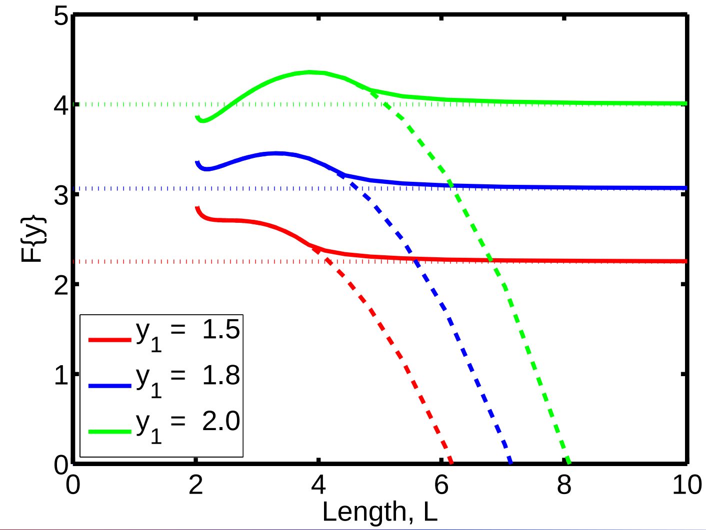
$$\begin{aligned} F\{y\} &= \int_0^{y_0} s\,ds + \int_0^{y_1} s\,ds \\ &= \frac{y_0^2}{2} + \frac{y_1^2}{2} \end{aligned}$$

- This is the energy of the degenerate solution
- It isn't necessarily the minimum energy configuration

Calculating the functional (energy)



Energy as a function of length



Pathologies

Notice in both cases above

- the approach only works for certain ranges of L.
- if L is too small, there is no physically possible solution.
  - e.g. if wire length $L < x_1 - x_0$
  - e.g. if oxhide length $L < x_1 - x_0$
- if L is too large in comparison to y_1, the solution may have our wire dragging on the ground.

Rigid Extremals

A particular problem to watch for are **rigid extremals**

- Extremals that cannot be perturbed, and still satisfy the constraint.
- For example

$$G\{y\} = \int_0^1 \sqrt{1+y'^2}\,dx = \sqrt{2}$$

with the boundary constraints $y(0) = 0$ and $y(1) = 1$.

The only possible y to satisfy this constraint is $y(x) = x$, so we cannot perturb around this curve to find conditions for viable extremals.

Rigid Extremals

Rigid extremal cases have some similarities to maximization of a function, where the constraints specify a single point:

- e.g. maximize $f(x,y) = x + y$, under the constraint that $x^2 + y^2 = 0$.

In the extremal case above, the constraint, and the end-points leave only one choice of function, $y(x) = x$

Interpretation of λ

- Consider finding extremals for

$$H\{y\} = F\{y\} + \lambda G\{y\},$$

where we iclude G to meet an isoperimetric constraint $G\{y\} = L$

- One way to think about λ is to think of the above as trying to minimuze $F\{y\}$ and $G\{y\} - L$
  - λ is a tradeoff between F and G
  - if λ is big, we give a lot of weight to G
  - if λ is small, then we give most weight to F
- So λ might be thought of as how hard we have to "pull" towards the constraint in order to make it

Interpretation of λ

- For example:
  - in the catenary problem, the size of λ is the amount we have to shift the cosh function up or down to get the right length.
  - when $\lambda = 0$ we get the natural catenary
    - i.e., in this case, we didn't need to change anything to get the right shape, so the constraint had no affect

Interpretation of λ

Write the problem (including the constant) as minimise

$$H\{y\} = \int f + \lambda(g-k)\,dx,$$

for constant $k = L/\int 1\,dx$, then

$$\frac{\partial H}{\partial k} = \lambda,$$

- we can also think of λ as the rate of change of the value of the optimum with respect to k
- when $\lambda = 0$, the functional H has a stationary point
  - e.g., in the catenary problem this is a local minimum corresponding to the natural catenary

Variational Methods & Optimal Control

lecture 15

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 15

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

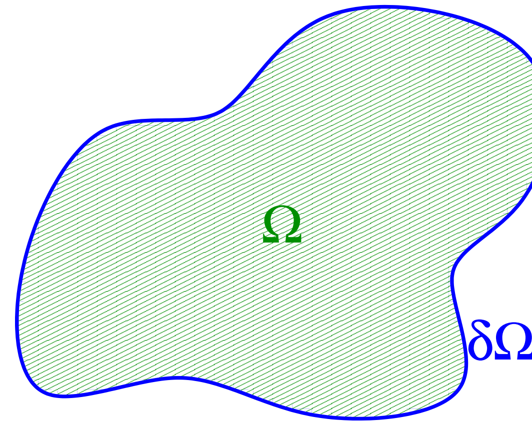
Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Isoperimetric constraints (continued)

We solve the more general case of Dido's problem: a general shape, without a coast, so that the perimeter must be parametrically described.

Isoperimetric problems



Dido's problem - traditional

Dido's problem is usually posed as follow

Find the curve of length L which encloses the largest possible area, i.e. maximize

$$\text{Area} = \iint_{\Omega} 1 \, dx \, dy$$

subject to the constraint

$$\oint_{\delta\Omega} 1 \, ds = L$$

Of course the problem is not yet in a convenient form.

Green's theorem

Green's theorem converts an integral over the area Ω to a contour integral around the boundary $\delta\Omega$.

$$\iint_\Omega \left(\frac{\partial\phi}{\partial x} + \frac{\partial\psi}{\partial y}\right) dx\, dy = \oint_{\delta\Omega} \phi\, dy - \psi\, dx$$

for $\phi, \psi : \bar{\Omega} \to \mathbb{R}$ such that ϕ, ψ, ϕ_x and ψ_y are continuous.

This converts an area integral over a region into a line integral around the boundary.

Geometric representation of area

The area of a region is given by

$$\text{Area} = \iint_\Omega 1\, dx\, dy$$

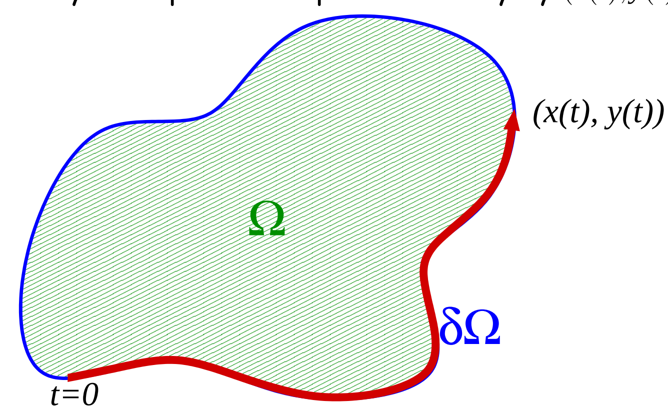
In Green's theorem choose $\phi = x/2$ and $\psi = y/2$, so that we get

$$\text{Area} = \iint_\Omega 1\, dx\, dy = \frac{1}{2}\oint_{\delta\Omega} x\, dy - y\, dx$$

Previous approach to Dido, was to use $y = y(x)$, but in more general case where the boundary must be closed, we can't define y as a function of x (or visa versa). So we write the boundary curve parametrically as $(x(t), y(t))$.

Parametric description of boundary

Boundary $\delta\Omega$ represented parametrically by $(x(t), y(t))$



Dido's problem

If the boundary $\delta\Omega$ is represented parametrically by $(x(t), y(t))$ then

$$\begin{aligned} \text{Area} &= \iint_\Omega 1\, dx\, dy \\ &= \frac{1}{2}\oint_{\delta\Omega} x\, dy - y\, dx \\ &= \frac{1}{2}\oint_{\delta\Omega} x\dot{y} - y\dot{x}\, dt \end{aligned}$$

So now the problem is written in terms of

$$\begin{aligned} \text{one independent variable} &= t \\ \text{two dependent variables} &= (x, y) \end{aligned}$$

Isoperimetric constraint

Previously we wrote the isoperimetric constraint as

$$G\{y\} = \int 1 ds = \int_{x_0}^{x_1} \sqrt{1+y'^2}\, dx = L$$

but now we must also modify this using

$$\frac{ds}{dt} = \sqrt{\frac{dx}{dt}^2 + \frac{dy}{dt}^2}$$

to get

$$G\{x,y\} = \oint 1 ds = \oint \sqrt{\dot{x}^2 + \dot{y}^2} dt = L$$

Dido's problem: Lagrange multiplier

Hence, we look for extremals of

$$H\{x,y\} = \oint \frac{1}{2}(x\dot{y} - y\dot{x}) + \lambda\sqrt{\dot{x}^2 + \dot{y}^2}\, dt$$

So $h(t,x,y,\dot{x},\dot{y}) = \frac{1}{2}(x\dot{y} - y\dot{x}) + \lambda\sqrt{\dot{x}^2+\dot{y}^2}$, and there are two dependent variables, with derivatives

$$\frac{\partial h}{\partial x} = \tfrac{1}{2}\dot{y} \qquad \frac{\partial h}{\partial \dot{x}} = -\frac{1}{2}y + \frac{\lambda \dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}$$

$$\frac{\partial h}{\partial y} = -\tfrac{1}{2}\dot{x} \qquad \frac{\partial h}{\partial \dot{y}} = \frac{1}{2}x + \frac{\lambda \dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}$$

Dido's problem: EL equations

Leading to the 2 Euler-Lagrange equations

$$\frac{d}{dt}\left[-\frac{1}{2}y + \frac{\lambda\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right] = \frac{1}{2}\dot{y}$$

$$\frac{d}{dt}\left[\frac{1}{2}x + \frac{\lambda\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right] = -\frac{1}{2}\dot{x}$$

Integrate

$$-\frac{1}{2}y + \frac{\lambda\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = \frac{1}{2}y + A$$

$$\frac{1}{2}x + \frac{\lambda\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}} = -\frac{1}{2}x - B$$

Dido's problem: solution

$$\frac{\lambda\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = y + A$$

$$\frac{\lambda\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}} = -x - B$$

Now square the two, and add them to get

$$\lambda^2 \frac{\dot{x}^2+\dot{y}^2}{\dot{x}^2+\dot{y}^2} = (y+A)^2 + (x+B)^2$$

or, more simply $(y+A)^2 + (x+B)^2 = \lambda^2$, the equation of a circle with center $(-A,-B)$, radius $|\lambda|$

End-conditions

Note, we can't set value at end points arbitrarily.

- if $x(t_0) = x(t_1)$, and $y(t_0) = y(t_1)$, then we get a closed curve, obviously a circle.
  - these conditions only amount to setting one constant, λ
  - there are many valid circles through (x_0, y_0), with centered along a circle of radius $|\lambda|$ about (x_0, y_0).
- on the other hand, if we specify different end-points, we are really solving a problem such as the simplified problem considered last week.

Why does it work?

Why does the Lagrange multiplier approach work here?

Consider Euler's finite difference method on a uniform grid for approximation of the functional

$$F\{y\} = \int_a^b f(x,y,y')\,dx \simeq \sum_{i=1}^{n} f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) \Delta x = \bar{F}(\mathbf{y})$$

where $\Delta x = (b-a)/n$, and $\Delta y_i = y_i - y_{i-1}$. The problem of finding an extremal curve now becomes one of finding stationary points of the function $\bar{F}(y_1, y_2, \ldots, y_n)$.

- we solve this by looking for $\partial \bar{F}/\partial y_i = 0$ for all $i = 1, 2, \ldots, n$.

Why does it work?

The constraint can be likewise approximated to give

$$G\{y\} \simeq \sum_{i=1}^{n} g\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) \Delta x = \bar{G}(\mathbf{y}) = L$$

Under our usual conditions on F and G, the limit as $n \to \infty$ gives

$$\begin{aligned} \bar{F}(\mathbf{y}) &\to F\{y\} \\ \bar{G}(\mathbf{y}) &\to G\{y\} \end{aligned}$$

That is, the **functions** of the approximation $\mathbf{y}$ converge to the **functionals** of the curve $y(x)$.

Why does it work?

In the finite dimensional case the constraint is

$$\bar{G}(y_1, y_2, \ldots, y_n) - L = 0$$

we use a standard Lagrange multiplier

$$\bar{H}(y_1, y_2, \ldots, y_n, \lambda) = \bar{F}(y_1, y_2, \ldots, y_n) + \lambda \left[\bar{G}(y_1, y_2, \ldots, y_n) - L\right]$$

- we solve this by looking for

$$\frac{\partial \bar{H}}{\partial y_i} = 0, \ \forall i = 1, 2, \ldots, n, \quad \text{and} \quad \frac{\partial \bar{H}}{\partial \lambda} = 0$$

- last equation just gives you back your constraint

Why does it work?

In our formulation of the isoperimetric problem we take

$$H\{y\} = F\{y\} + \lambda G\{y\}$$

and we also have

$$\bar{H}(\mathbf{y},\lambda) = \bar{F}(\mathbf{y}) + \lambda\left[\bar{G}(\mathbf{y}) - L\right]$$

In the limit as $n \to \infty$ we find that

$$\bar{H}(\mathbf{y},\lambda) \to H\{y\} - \lambda L$$

The E-L equations for $H\{y\} - \lambda L$ and $H\{y\}$ are the same, so they have the same extremals!

Why does it work?

See van Brunt, pp.83–87 for a more rigorous explanation of Lagrange multipliers in this context.

Multiple constraints

We can also handle multiple constraints via multiple Lagrange multipliers. For instance, given we wish to find extremals of

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

with the m constraints

$$G_k\{y\} = \int_{x_0}^{x_1} g_k(x,y,y')\,dx = L_k$$

we would look for extremals of

$$H\{y\} = \int_{x_0}^{x_1} h(x,y,y')\,dx = \int_{x_0}^{x_1} f(x,y,y') + \sum_{k=1}^{m} \lambda_k g_k(x,y,y')\,dx$$

Variational Methods & Optimal Control

lecture 16

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 16

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Non-integral constraints

It is relatively easy to adapt the Lagrange multiplier technique to the case with non-integral constraints.

- **Holonomic constraints** [a] are of the form

$$g(t, \mathbf{q}) = 0$$

- **Non-Holonomic constraints** are of the form

$$g(t, \mathbf{q}, \dot{\mathbf{q}}) = 0$$

The former is simpler, and we consider this first.

[a] Holonomic comes from the greek "holos", for "whole". In this context it refers to integrability of the constraint. Notice that non-holonomic constraints are really DEs

Holonomic Constraints

Constraints of the form $g(x,y) = 0$, or $g(t,\mathbf{q}) = 0$, which don't involve derivatives of $y(x)$ or $\mathbf{q}$ can also be handled using a Lagrange multiplier technique, but we have to introduce a Lagrange multiplier function $\lambda(x)$, not just a single value λ. Effectively we introduce one Lagrange multiplier at each point where the constraint is enforced.

Holonomic constraints

Consider the problem of finding extremals of

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

subject to the constraint

$$g(x,y) = 0$$

In this case we introduce a function $\lambda(x)$ (also called a Lagrange multiplier), and look for extremals of

$$H\{y,\lambda\} = F\{y\} + \int_{x_0}^{x_1} \lambda(x) g(x,y) dx$$

Why does it work

Go back to the finite approximation Consider Euler's finite difference method on a uniform grid for approximation of the functional

$$F\{y\} = \int_a^b f(x,y,y')\,dx \simeq \sum_{i=1}^n f\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right)\Delta x = \bar{F}(\mathbf{y})$$

The constraint applies a condition on each (x_i, y_i), so in the approximation there are n constraints,

$$g(x_i, y_i) = 0 \text{ for } i = 1, \ldots, n$$

Why does it work

There are n constraints,

$$g(x_i, y_i) = 0 \text{ for } i = 1, \ldots, n$$

For optimization problems with n constraints, we introduce n Lagrange multipliers, and maximize

$$H(\mathbf{y}) = F(\mathbf{y}) + \sum_{k=1}^n \lambda_k g(x_k, y_k)$$

In the limit as $n \to \infty$

$$\Delta x \sum_{k=1}^n \lambda_k g(x_k, y_k) \to \int_a^b \lambda(x) g(x,y)\,dx$$

and hence the choice of $H\{y,\lambda\} = F\{y\} + \int_a^b \lambda(x) g(x,y) dx$.

Holonomic constraints

$$\begin{aligned} H\{y,\lambda\} &= F\{y\} + \int_{x_0}^{x_1} \lambda(x) g(x,y)\,dx \\ &= \int_{x_0}^{x_1} f(x,y,y') + \lambda(x) g(x,y)\,dx \end{aligned}$$

So we can apply the Euler-Lagrange equations to

$$h(x,y,y',\lambda) = f(x,y,y') + \lambda(x) g(x,y)$$

To get the Euler-Lagrange equations

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} - \lambda(x)\frac{\partial g}{\partial y} = 0$$

Multiple dependent variables

With multiple dependent variables holonomic constraints are of the form

$$g(t, \mathbf{q}) = 0$$

and they don't involve derivatives.

Example: find geodesics on a cylinder, e.g. minimize

$$F\{x,y,z\} = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\,dt$$

subject to $x^2 + y^2 - r^2 = 0$, the equation of a right circular cylinder with radius r.

Multiple dependent variables

$$H\{\mathbf{q},\lambda\} = F\{\mathbf{q}\} + \int_{x_0}^{x_1} \lambda(t) g(t,\mathbf{q}) dx$$

So we can apply the Euler-Lagrange equations to

$$h(t,q,\dot{q},\lambda) = f(t,q,\dot{q}) + \lambda(t) g(t,q)$$

To get the Euler-Lagrange equations

$$\frac{d}{dx}\frac{\partial f}{\partial \dot{q}_k} - \frac{\partial f}{\partial q_k} - \lambda(t)\frac{\partial g}{\partial q_k} = 0$$

for all k.

General geodesic problem

General geodesic problem can be written as minimize

$$F\{x,y,z\} = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt$$

subject to

$$g(x,y,z) = 0$$

where $g(x,y,z) = 0$ is the equation describing the surface of interest.

We instead minimize

$$H\{x,y,z,\lambda\} = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} + \lambda(t) g(x,y,z)\, dt$$

General geodesic problem

Given this formulation of the geodesic problem, the E-L equations become

$$\begin{aligned}
\frac{d}{dt}\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} - \lambda(t)\frac{\partial g}{\partial x} &= 0 \\
\frac{d}{dt}\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} - \lambda(t)\frac{\partial g}{\partial y} &= 0 \\
\frac{d}{dt}\frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} - \lambda(t)\frac{\partial g}{\partial z} &= 0
\end{aligned}$$

which may be easier to solve in some cases.

Example: Geodesics on the sphere

Find the geodesics on the sphere: e.g. we wish to find a parametric curve $(x(t),y(t),z(t))$ to minimize distance

$$F\{x,y,z\} = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt$$

subject to being on the surface of a sphere

$$x^2 + y^2 + z^2 = a^2$$

We get

$$h(t,x,y,z,\dot{x},\dot{y},\dot{z}) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} + \lambda(t)(x^2 + y^2 + z^2)$$

and there are three dependent variables (x,y,z)

Example: Geodesics on the sphere

$$\frac{\partial h}{\partial x} = 2\lambda x \qquad \frac{\partial h}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}$$

$$\frac{\partial h}{\partial y} = 2\lambda y \qquad \frac{\partial h}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}$$

$$\frac{\partial h}{\partial z} = 2\lambda z \qquad \frac{\partial h}{\partial \dot{z}} = \frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}$$

Example: Geodesics on the sphere

There are 3 dependent variables (x,y,z), and so 3 E-L equations, e.g.

$$\begin{aligned} 2\lambda x &= \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right) \\ &= \frac{\ddot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}} - \frac{\dot{x}(\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z})}{(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{3/2}} \end{aligned}$$

Due to symmetry, the equation

$$2\lambda u = \frac{\ddot{u}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}} - \frac{\dot{u}(\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z})}{(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{3/2}}$$

holds for $u = x, y$ and z.

Example: Geodesics on the sphere

Now

$$2\lambda u = \ddot{u}\frac{1}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}} - \dot{u}\frac{(\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z})}{(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{3/2}}$$

is a second order linear DE in u, and so it has only 2 linearly independent solutions, but the DE holds for $u = x,y$ and z

Therefore x, y, and z are linearly dependent, and so we can write them as

$$Ax + By + Cz = 0$$

but this is the equation of a plane through the origin. Once again we have shown that geodesics on the sphere are great circles

Misc

- Note, sometimes a constraint involving derivatives may be integrated to get a holonomic constraint, so we refer to these constraints as integrable.
- In general, though, we will also need to deal with constraints involving derivatives as these may describe an entire systems behaviour, and be very difficult to integrate out of the problem.
  - e.g., when we want to describe a "controlled" system

Non-Holonomic Constraints

Constraints of the form $g(x,y,y')=0$, or $g(t,\mathbf{q},\dot{\mathbf{q}})=0$, which involve derivatives. They are effectively additional DEs which we need to solve, but we can once again use Lagrange multipliers.

Non-holonomic constraints

Example non-holonomic constraints:

$$g(x,y,y')=0 \quad \text{or} \quad g(t,\mathbf{q},\dot{\mathbf{q}})=0$$

Instances:

- $y=\dot{x}$
- $y'^2=\log x$

Solution technique is just as for holonomic constraints, e.g.,

$$H\{y,\lambda\}=F\{y\}+\int_{x_0}^{x_1}\lambda(x)g(x,y,y')dx$$

and the argument for why it works is almost identical.

Example

Using such constraints to avoid higher derivatives

Imagine the functional

$$F\{y\}=\int_a^b f(x,y,y',y'')\,dx$$

we have already see that we can derive a new form of the E-L (Euler-Poisson) equations for this case, e.g.

$$\frac{\partial f}{\partial y}-\frac{d}{dx}\frac{\partial f}{\partial y'}+\frac{d^2}{dx^2}\frac{\partial f}{\partial y''}=0$$

but constraints give us an alternative approach to this problem.

Example

Introduce the new variable $z=y'$, and rewrite the functional as

$$F\{y\}=\int_a^b f(x,y,z,z')\,dx$$

Now there is more than one dependent variable, but no second order derivatives, however, we must also introduce the constraint that

$$z-y'=0$$

and so we look for stationary curves of the functional

$$H\{y,z,\lambda\}=\int_a^b f(x,y,z,z')+\lambda(x)(z-y')\,dx$$

Example

The Euler-Lagrange equations for y and z are

$$\frac{d}{dx}\frac{\partial h}{\partial y'} - \frac{\partial h}{\partial y} = 0$$
$$\frac{d}{dx}\frac{\partial h}{\partial z'} - \frac{\partial h}{\partial z} = 0$$

note that $h(x,y,z,z') = f(x,y,z,z') + \lambda(x)(z-y')$ so the E-L equations become

$$\frac{d}{dx}[-\lambda(x)] - \frac{\partial f}{\partial y} = 0$$
$$\frac{d}{dx}\frac{\partial f}{\partial z'} - \frac{\partial f}{\partial z} - \lambda(x) = 0$$

Example

The first Euler-Lagrange equation can be rewritten

$$\frac{d\lambda}{dx} = -\frac{\partial f}{\partial y}$$

Differentiating the second E-L equation WRT x we get

$$\frac{d^2}{dx^2}\frac{\partial f}{\partial z'} - \frac{d}{dx}\frac{\partial f}{\partial z} - \frac{d\lambda}{dx} = 0$$

Note from above that $\lambda' = -f_y$ and that $z = y'$ and $z' = y''$ we get (as before) the Euler-Poisson equation:

$$\frac{d^2}{dx^2}\frac{\partial f}{\partial y''} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{\partial f}{\partial y} = 0$$

Some intuition

- Earlier we derived the Euler-Lagrange equations assuming treating y and y' as if they were independent variables.
  - In reality they are related along the extremal
  - Lets get some motivation for this
- Start by taking a new variable $u(x) = y'(x)$, and put this into our optimization problem

$$H\{y,u,\lambda\} = \int_a^b f(x,y,u) + \lambda(x)(u-y')\,dx$$

  - we can use same trick as in previous slides to get the Euler-Lagrange equations

Newton's aerodynamical problem

Find extremal of "air resistance"

$$F\{y\} = \int_0^R \frac{x}{1+y'^2}\,dx,$$

subject to $y(0) = L$ and $y(R) = 0$ and $y' \leq 0$ and $y'' \geq 0$
The Euler-Lagrange equations are

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = \frac{d}{dx}\frac{2xy'}{(1+y'^2)^2} = 0$$

Rearranging we get

$$2xy' = c(1+y'^2)^2$$

which isn't much fun to solve directly.

Newton's aerodynamical problem

Alternative: define a new variable u, and constrain it

$$u = -y'$$

Add Lagrange multiplier $\lambda(x)$ to the functional

$$H\{y,u,\lambda\} = \int_0^R \frac{x}{1+u^2} + \lambda(y'+u)\,dx,$$

Now solve as you would for a problem with three dependent variables (y,u,λ) of x.

- We expect three Euler-Lagrange equations
- One equation in each dependent variable
  - but we already know the λ equation

Newton's aerodynamical problem

Euler-Lagrange equations

$$\begin{aligned}
\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} &= 0 \\
\frac{d}{dx}\frac{\partial f}{\partial u'} - \frac{\partial f}{\partial u} &= 0 \\
\frac{d}{dx}\frac{\partial f}{\partial \lambda'} - \frac{\partial f}{\partial \lambda} &= 0
\end{aligned}$$

give the DEs

$$\begin{aligned}
\lambda &= const \\
-\frac{2xu}{(1+u^2)^2} - \lambda &= 0 \\
y' + u &= 0
\end{aligned}$$

Newton's aerodynamical problem

$$\begin{aligned}
\lambda &= const \\
-\frac{2xu}{(1+u^2)^2} - \lambda &= 0 \\
y' + u &= 0
\end{aligned}$$

If $\lambda = 0$, then for $x > 0$ we get $u = 0$, and hence $y = const$.
If $\lambda \neq 0$ then the second equation implies

$$x(u) = \frac{c}{u}(1+u^2)^2 = c\left(\frac{1}{u} + 2u + u^3\right).$$

for c constant.

Newton's aerodynamical problem

From the last equation (which we insisted on at the start), we get

$$\frac{dy}{dx} = -u$$

Now note that from the chain rule

$$\begin{aligned}
\frac{dy}{du} &= \frac{dy}{dx}\frac{dx}{du} = -u\frac{dx}{du} \\
&= c\left(-\frac{1}{u} + 2u + 3u^3\right)
\end{aligned}$$

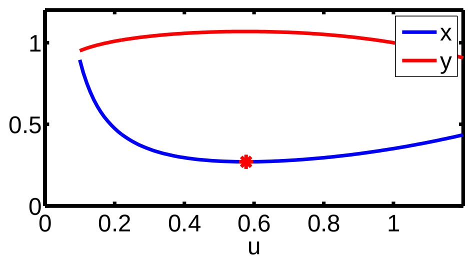
which we can integrate with respect to u to get

$$y(u) = const - c\left(-\ln u + u^2 + \frac{3}{4}u^4\right)$$

Newton's aerodynamical problem

Some notes about the solution

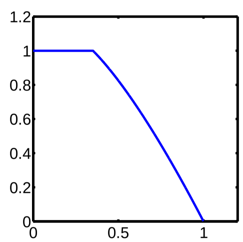
- $x(u) = \frac{c}{u}(1+u^2)^2 > 0$ for all u



- hence the part of the curve near $x = 0$ must have $y = L$
- but $y(x) = L$ for all $x \in [0, R]$ can't be the minimum because we know better profiles (e.g. a cone).

Newton's aerodynamical problem

So we know the solution must look like something like



A simple example is the **frustum of a cone**

- the part of a cone between two parallel planes
- but we can do better by making the sloped part follow E-L equations
- still need to work out where the "corner" goes

Newton's aerodynamical problem

$$y(u) = const - c\left(-\ln u + u^2 + \frac{3}{4}u^4\right)$$

At u_1 we have $y(u_1) = L$. For convenience we write

$$y(u) = L - c\left(-A - \ln u + u^2 + \frac{3}{4}u^4\right)$$

so at u_1 we get

$$\begin{aligned} L &= L - c\left(-\ln u_1 - A + u_1^2 + \frac{3}{4}u_1^4\right) \\ 0 &= -c\left(-\ln u_1 - A + u_1^2 + \frac{3}{4}u_1^4\right) \\ A &= -\ln u_1 + u_1^2 + \frac{3}{4}u_1^4 \end{aligned}$$

Newton's aerodynamical problem

$$\begin{aligned} y(u) &= L - c\left(-A - \ln u + u^2 + \frac{3}{4}u^4\right) \\ x(u) &= \frac{c}{u}(1+u^2)^2 \end{aligned}$$

Now at u_2 we have $x(u_2) = R$ and $y(u_2) = 0$ so

$$\begin{aligned} L &= c\left(-A - \ln u_2 + u_2^2 + \frac{3}{4}u_2^4\right) \\ R &= \frac{c}{u_2}(1+u_2^2)^2 \end{aligned}$$

divide the first equation by the second and we get ...

Newton's aerodynamical problem

$$\frac{L}{R} = u_2\left(-A - \ln u_2 + u_2^2 + \frac{3}{4}u_2^4\right)(1+u_2^2)^{-2}$$

The function on the RHS is increasing so we can solve this equation (numerically (e.g., using matlab's `fsolve`), and we obtain a value for u_2.
We can find c using $x(u_2) = R$

$$\begin{aligned} R &= \frac{c}{u_2}(1+u_2^2)^2 \\ c &= \frac{u_2 R}{(1+u_2^2)^2} \end{aligned}$$

All we need to know now is u_1, which gives us A and $x(u_1)$, which gives us u_2, which gives us c.

Newton's aerodynamical problem

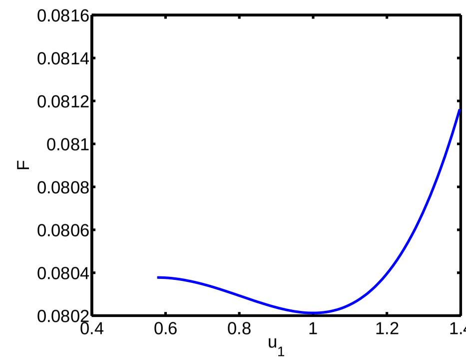
Take $x(u_1) = x_1$

$$\begin{aligned} F\{y\} &= \int_0^{x_1} x\,dx + \int_{x_1}^{R} \frac{x}{1+y'^2}\,dx \\ &= \frac{x_1^2}{2} + \int_{u_1}^{u_2} \frac{c(1+u^2)^2/u}{1+u^2}\frac{dx}{du}\,du \\ &= \frac{x_1^2}{2} + c^2\int_{u_1}^{u_2} \frac{(1+u^2)^2(3u^2-1)}{u^3}\,du \\ &= \frac{x_1^2}{2} + c^2\left[\frac{3u^4}{4} + \frac{5u^2}{2} + \ln(u) + \frac{1}{2u^2}\right]_{u_1}^{u_2} \end{aligned}$$

and note that c and u_2 are effectively functions of u_1.

Newton's aerodynamical problem

Numerical evaluation of the integral F for different values of if u_1



Minimum occurs for $u_1 = 1$, we'll prove this later on.

Intro to Optimal Control

One way we see non-holonomic constraints is when we consider control problems. In these we seek to control a system described by a DE (the constraint) subject to some input which we can control (optimize).

Systems

system = machine + controller

e.g. vehicle

- **machine:** engine, body, seating
- **control:** accelerator, brakes, steering (driver)

Problems:

- **Control problems:** how do we set, say the steering and acceleration of a car to get it from point A to point B.
- **Optimal Control problems:** same as above, but do it in minimum time, or using minimum fuel.

Solution Philosophy

Solve however you can

- often easier approach that CoV
  - systems of DEs just need to be solved
  - a lot is about whether a control exists!
  - e.g. see "Optimal Control: an Introduction to the Theory with Applications", Leslie M. Hocking, Clarendon Press, Oxford, 1991.
- on the other hand, we have a powerful set of tools now, so we shall use them here
  - all it takes is a shift in perspective
  - then all of the CoV work from earlier is applicable

CoV for Optimal Control

Optimal control is just a switch in our perspective:

- previous problems, mainly concerned with modeling and analysis of physical (often mechanical systems), e.g. catenary
  - take a system, and find an extremal which minimizes, say potential energy, and this describes the system
- now we can set part of the systems (e.g. force) to create a particular curve which minimizes some quantity
  - e.g. set force to minimize fuel usage of a rocket (changing orbits)

Formulation of control problems

We break a control problem into two parts

- **The system state:** $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^t$
  The system state describes the system (e.g. position and velocity of the car in car parking example)
- **The control:** $\mathbf{u}(t) = (u_1(t), \ldots, u_m(t))^t$
  We apply the control to the system (e.g. force applied to the car).

The evolution of the system is governed by a DE

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$

In a control problem we control the system to get it to a particular state $\mathbf{x}(t_1)$ at time t_1, given initial state $\mathbf{x}(t_0)$.

Optimal control problems

In an optimal control problem we have

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$

and once again we wish to get to state $\mathbf{x}(t_1)$ given initial state $\mathbf{x}(t_0)$, but now we wish to do so while minimizing a functional

$$F\{\mathbf{u}\} = \int_{t_0}^{t_1} f(t, \mathbf{x}, \mathbf{u})\, dt$$

That is, we wish to choose a function $\mathbf{u}(t)$ which minimizes the functional $F\{\mathbf{u}\}$, while satisfying the end-point conditions $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{x}(t_1) = \mathbf{x}_1$, and the non-holonomic constraint $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$.

Example: stimulated plant growth

Plant growth problem:

- market gardener wants to plants to grow to a fixed height 2 within a fixed window of time $[0,1]$
- can supplement natural growth with lights with "brightness" $u(t)$
- growth rate dictates

$$\dot{x} = 1 + u$$

- cost of lights

$$F\{u\} = \int_0^1 \frac{1}{2} u(t)^2\, dt$$

Plant growth problem statement

Minimize

$$F\{u\} = \int_0^1 \frac{1}{2} u^2\, dt$$

Subject to $x(0) = 0$, and $x(1) = 2$ and

$$\dot{x} = 1 + u$$

- we effectively have two dependent variables x and u
- though we can only control one of these

Plant growth: Lagrange multiplier

We can include the non-holonomic constraint into the problem via a Lagrange multiplier, e.g., we seek to minimize

$$\begin{aligned} H\{u, x, \lambda\} &= \int_0^1 \frac{1}{2} u^2 + \lambda(t) \left[\dot{x} - 1 - u \right]\, dt \\ &= \int_0^1 h(x, u, \dot{x}, \lambda)\, dt \end{aligned}$$

We might think of λ as a third variable, but the E-L equations in λ will just give us the constraint $\dot{x} = 1 + u$ back again.

Plant growth: E-L equations

2 dependent variables, so E-L equations

$$\begin{aligned}\frac{\partial h}{\partial u} - \frac{d}{dt}\frac{\partial h}{\partial \dot{u}} &= 0\\ \frac{\partial h}{\partial x} - \frac{d}{dt}\frac{\partial h}{\partial \dot{x}} &= 0\end{aligned}$$

These are

$$\begin{aligned}u - \lambda &= 0\\ \dot{\lambda} &= 0\end{aligned}$$

Simplifying we see the solution is

$$u = const$$

Plant growth solution

Going back to the DE $\dot{x} = 1 + u$, and taking $u = c$ we get

$$x = (c+1)t + k$$

The end-point constraints require that $x(0) = 0$ and $x(1) = 2$ so

$$x = 2t$$

Clearly $u = 1$ is the optimal control.
We can also derive the optimal cost

$$F\{u\} = \int_0^1 \frac{1}{2} u(t)^2 \, dt = \frac{1}{2}$$

Optimal Control

We will consider optimal control much more thoroughly later in this course. There are many approaches one can adopt to such problems, and we shall come back to this problem in particular, later in the course. First we need to know some more CoV, especially how to deal with

- free end points
  - say there isn't a fixed time window
  - perhaps the final state isn't pre-determined
- costs other than integrals
  - e.g., costs associated with end states

Variational Methods & Optimal Control

lecture 17

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 17

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Non-fixed end point problems

What happens when we don't fix the end-points of an extremal? In this case **natural boundary conditions** are automatically introduced, and these can allow us to solve the E-L equations.

Non-fixed end point problems

What happens when we don't fix the boundary points?

There are real problems like this, for instance

- a freely supported beam
  end points fixed, but not derivatives
- a beam supported at only one end
  one end point and derivative fixed, other free
- shortest path between two curves
  end points lie of curves, but not fixed
- rocket changing between two orbits
  end points lie on curves, and path is tangent to the two orbits.

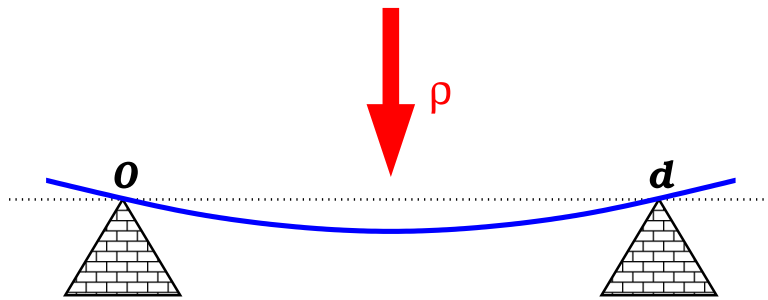
We then get **natural boundary conditions**

Free end points: Fixed x, Free y and/or y'

First we'll consider what happens when we allow y or y' to vary at the end-points, but we still keep the x values of the end-points fixed at x_0 and x_1.

Example: freely supported beam

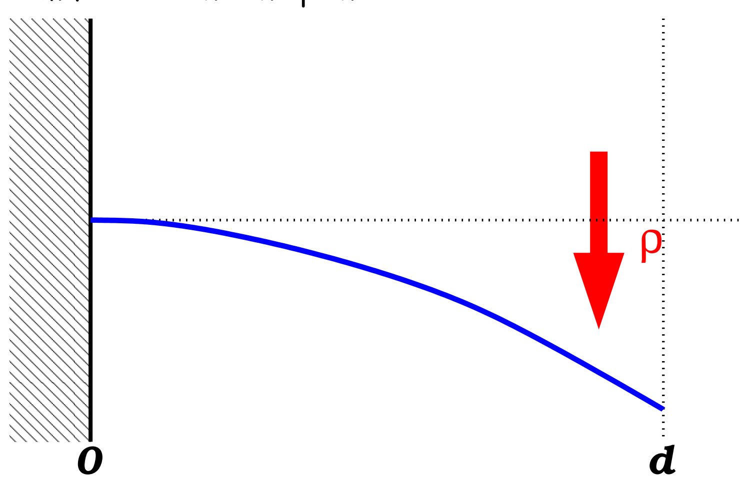
Freely supported beam



For the beam problems considered before, we had to specify the derivative at the boundary, but here it can vary.

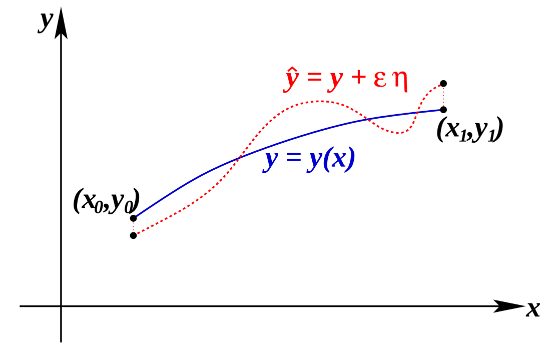
Example: beam fixed at one end point

Beam fixed at one end point



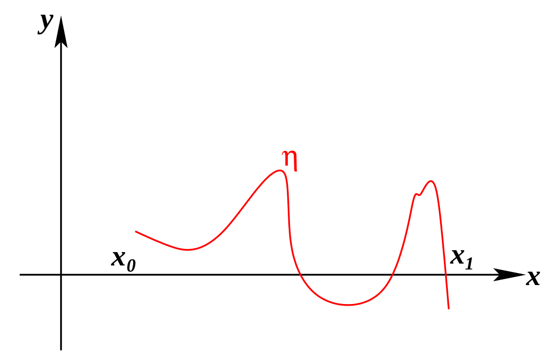
Perturbation again

We approach this the same way we did with all other variational problems, we perturb the curve and examine the First Variation, but this time, we allow $y(x_0)$ and $y(x_1)$ to vary as well.



Space of Perturbations

Now the space $\mathcal{H}$ of perturbations η contains functions whose value at x_0 and x_1 is no longer zero.



Same derivation of the first variation

Simple case where $F\{y\} = \int_{x_0}^{x_1} f(x,y,y')\,dx$

$$\begin{aligned}
f(x,\hat{y},\hat{y}') &= f(x,y,y') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right] + O(\varepsilon^2) \\
F\{\hat{y}\} - F\{y\} &= \int_{x_0}^{x_1} f(x,\hat{y},\hat{y}')dx - \int_{x_0}^{x_1} f(x,y,y')dx \\
&= \varepsilon\int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right]dx + O(\varepsilon^2) \\
\delta F(\eta,y) &= \lim_{\varepsilon\to 0}\frac{F\{y+\varepsilon\eta\} - F\{y\}}{\varepsilon} \\
&= \int_{x_0}^{x_1}\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right]dx
\end{aligned}$$

The first variation

As before, we can vary the sign of ε, so for $F\{y\}$ to be a local minima it must be the case that

$$\delta F(\eta,y) = 0, \qquad \forall \eta \in \mathcal{H}$$

however, now $\mathcal{H}$ allows curves with arbitrary end-points, so that $\eta(x_0) \neq 0$, and $\eta(x_1) \neq 0$ are possible.

Hence when we integrate by parts we get

$$\delta F(\eta,y) = \left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1}\eta\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]dx$$

But now the first term $\left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1}$ is not necessarily zero.

The first variation

However, $\delta F(\eta,y) = 0$ for all η, which includes cases where $\eta(x_0) = \eta(x_1) = 0$, and so the Euler-Lagrange equation must still be satisfied for such and extremal.

Given the E-L equation is satisfied by an extremal, the condition $\delta F(\eta,y) = 0$ next implies that

$$\left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} = 0$$

and we can likewise choose curves η such that $\eta(x_0) \neq 0$ and $\eta(x_1) = 0$, or visa versa, so that we must have

$$\left.\frac{\partial f}{\partial y'}\right|_{x_0} = 0 \quad \text{and} \quad \left.\frac{\partial f}{\partial y'}\right|_{x_1} = 0$$

Euler-Lagrange again

Hence, as before, the extremal must satisfy the E-L equations

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$

but now that the boundary conditions were not specified as part of the problem, we get natural boundary conditions

$$\left.\frac{\partial f}{\partial y'}\right|_{x_0} = 0 \quad \text{and} \quad \left.\frac{\partial f}{\partial y'}\right|_{x_1} = 0$$

which specify that the derivative at the end-points will be zero.

Extensions (i)

What happens if we fix one end point, e.g. $y(x_0) = y_0$.

The result is we cannot vary this end-point when perturbing, so $\eta(x_0) = 0$, and therefore the condition

$$\left[\eta \frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} = 0$$

collapses to give just one extra condition

$$\left.\frac{\partial f}{\partial y'}\right|_{x_1} = 0$$

Hence the boundary conditions are **modular** in the sense that when we remove one, we replace it automatically with the natural boundary condition.

Extensions (ii)

The above results can be extended as before, in particular, consider a functional containing higher order derivatives:

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

$$\delta F(\eta, y) = \left[\eta\left(\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\right)\right]_{x_0}^{x_1} + \left[\eta' \frac{\partial f}{\partial y''}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} - \eta \frac{d}{dx}\frac{\partial f}{\partial y'} + \eta \frac{d^2}{dx^2}\frac{\partial f}{\partial y''}\right] dx$$

where we see integration by parts introduces terms including η and η'.

Extensions (ii)

The Euler-Lagrange equations are

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 0$$

where the natural boundary conditions are

$$\left.\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\right|_{x_0} = 0 \quad \text{and} \quad \left.\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\right|_{x_1} = 0$$

$$\left.\frac{\partial f}{\partial y''}\right|_{x_0} = 0 \quad \text{and} \quad \left.\frac{\partial f}{\partial y''}\right|_{x_1} = 0$$

where the first two replace absent conditions on the value of y at the end-points, and the second two replace absent conditions on y' at the end points.

Bent beam

Let $y : [0, d] \to \mathbb{R}$ describe the shape of the beam, and $\rho : [0, d] \to \mathbb{R}$ be the load per unit length on the beam. For a bent elastic beam the potential energy from elastic forces is

$$V_1 = \frac{\kappa}{2}\int_0^d y''^2\, dx, \qquad \kappa = \text{flexural rigidity}$$

The potential energy is

$$V_2 = -\int_0^d \rho(x) y(x)\, dx$$

Thus the total potential energy is

$$V = \int_0^d \frac{\kappa y''^2}{2} - \rho(x) y(x)\, dx$$

Bent Beam: see earlier

The Euler-Lagrange equation is

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} = 0$$

$$y^{(4)} = \frac{\rho(x)}{\kappa}$$

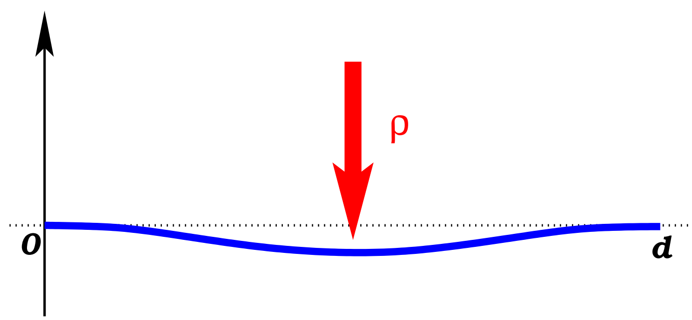
This DE has solution

$$y(x) = P(x) + c_3x^3 + c_2x^2 + c_1x + c_0$$

where the c_k's are the constants of integration, and $P(x)$ is a particular solution to $P^{(4)}(x) = \rho(x)/\kappa$.

Bent Beam: Example 1

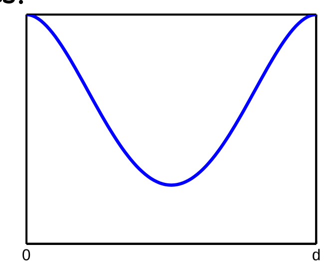
Doubly clamped: see earlier lectures.



Two end-points are fixed, and clamped so that they are level, e.g. $y(0) = 0, y'(0) = 0$, and $y(d) = 0$ and $y'(d) = 0$.

Bent Beam: Example 1

Doubly clamped: see earlier lectures.
Choose a solution of the form



$$y(x) = \frac{\rho(d-x)^2x^2}{24\kappa}$$

Then the derivative

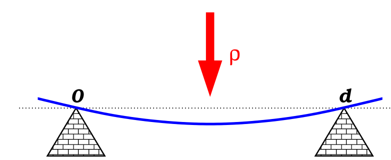
$$y'(x) = \frac{2\rho(d-x)x^2}{12\kappa} + \frac{\rho(d-x)^2x}{12\kappa}$$

We can see that the constraints are satisfied

$$y(0) = 0 \quad \text{and} \quad y(d) = 0$$
$$y'(0) = 0 \quad \text{and} \quad y'(d) = 0$$

Bent Beam Example 2

Freely supported, uniform load
The natural constraints are



$$\left.\frac{\partial f}{\partial y''}\right|_{x_0} = \kappa y''(x_0) = 0$$

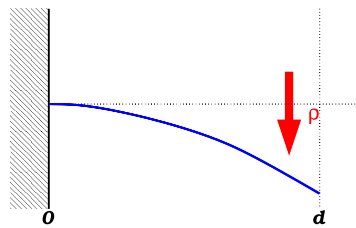
$$\left.\frac{\partial f}{\partial y''}\right|_{x_1} = \kappa y''(x_1) = 0$$

The fixed end-points are $y(0) = y(d) = 0$, so uniform load solution looks like

$$y(x) = \frac{\rho x\left(d^3 - 2dx^2 + x^3\right)}{24\kappa}$$

Bent Beam Example 3

One end-point fixed, and clamped.
Called a **Cantilever**
The natural constraints are



$$\begin{aligned} \left.\frac{\partial f}{\partial y''}\right|_{x_1} &= \kappa y''(x_1) = 0 \\ \left.\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\right|_{x_1} &= \left.-\frac{d}{dx}\kappa y''\right|_{x_1} \\ &= \kappa y'''(x_1) = 0 \end{aligned}$$

The clamped end-point introduces constraints $y(0) = 0$ and $y'(0) = 0$ so the solution for uniform load is

$$y(x) = \frac{\rho x^2(6d^2 - 4dx + x^2)}{24\kappa} \quad \text{and} \quad y(d) = \frac{\rho d^4}{8\kappa}$$

Bent beam, end-points conditions

End-point conditions are modular: i.e., we can use different end-point conditions at each end of the beam.

- **clamped**: specifies the position, and the derivative.
- **freely supported**: specifies the position. Natural boundary condition is that the second derivative is zero at the end point.
- **no condition**: neither position, nor end-point are specified, so the natural boundary conditions fix the second and third derivatives at the end point to be zero.

Bent Beam Example 4

One end-point fixed, but not clamped.

In this case the beam just collapses, and lies vertical.

The approach doesn't work, but this is a failure of the **model**, not the **method**.

In this case, the cantilever approximation (that x_1 is fixed) no longer works, and we need to consider a more general model that allows x_1 to vary as well.

Intro to Optimal Control (part II)

Often in optimal control problems we may specify the initial state, but not the final state. However, there may be a cost associated with the final state, and we include this in the functional to be minimized (or maximized). We call this a terminal cost.

Optimal control with terminal costs

In an optimal control problem we again have a non-holonomic constraint

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$

given initial state $x(t_0)$, but now the final state will be free and we wish to minimize a functional

$$F\{\mathbf{u}\} = \phi(t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} f(t, \mathbf{x}, \mathbf{u})\, dt$$

the term $\phi(t_1, \mathbf{x}(t_1))$ is called the **terminal cost**.

Terminal costs reformulation

Note that

$$\phi(t_1, \mathbf{x}(t_1)) = \phi(t_0, \mathbf{x}(t_0)) + \int_{t_0}^{t_1} \frac{d}{dt}\phi(t, \mathbf{x})\, dt$$

so we can rewrite

$$\begin{aligned} F\{\mathbf{u}\} &= \phi(t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} f(t, \mathbf{x}, \mathbf{u})\, dt \\ &= \phi(t_0, \mathbf{x}(t_0)) + \int_{t_0}^{t_1} \left[f(t, \mathbf{x}, \mathbf{u}) + \frac{d}{dt}\phi(t, \mathbf{x}) \right] dt \end{aligned}$$

where note that the first term is fixed by the starting point, and so we can drop it from the problem.

Terminal costs: example

Imagine the problem we wish to solve is to minimize the time, i.e. t_1. ,We could write this as a terminal cost problem, e.g. minimize

$$F\{\mathbf{u}\} = t_1$$

So $\phi(t) = t$, and $\frac{d}{dt}\phi = 1$ and therefore, we can write the minimum time problem in the form

$$F\{\mathbf{u}\} = \int_{t_0}^{t_1} 1\, dt$$

Terminal costs and E-L equations

Given a problem like

$$F\{\mathbf{u}\} = \int_{t_0}^{t_1} \left[f(t, \mathbf{x}, \mathbf{u}) + \frac{d}{dt}\phi(t, \mathbf{x}) \right] dt$$

Note that

$$\frac{d}{dt}\phi(t, \mathbf{x}) = \frac{\partial \phi}{\partial t} + \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}\dot{x}_i$$

E-L equations:

$$\begin{aligned} \frac{d}{dt}\frac{\partial f}{\partial \dot{x}_k} - \frac{\partial f}{\partial x_k} + \frac{d}{dt}\frac{\partial \phi}{\partial x_k} - \frac{\partial^2 \phi}{\partial x_k \partial t} - \sum_{i=1}^{n} \frac{\partial^2 \phi}{\partial x_k \partial x_i}\dot{x}_i &= 0 \\ \frac{d}{dt}\frac{\partial f}{\partial \dot{x}_k} - \frac{\partial f}{\partial x_k} &= 0 \end{aligned}$$

Terminal costs and E-L equations

Hence terminal costs play no part in the Euler-Lagrange equations, which makes sense

- fixed end-point problem
  - terminal cost is fixed (by the end-point)
  - so Euler-Lagrange equations unchanged
- free end-point problem
  - Euler-Lagrange equations aren't effected by freeing up the end-points
  - natural boundary conditions may be effected, e.g. in the case with free x_k (but fixed t) we would get a natural boundary condition like

$$\left. \frac{\partial f}{\partial \dot{x}_k} + \frac{\partial}{\partial \dot{x}_k} \frac{d\phi}{dt} \right|_{t=t_1} = \left. \frac{\partial f}{\partial \dot{x}_k} + \frac{\partial \phi}{\partial x_k} \right|_{t=t_1} = 0$$

Example: stimulated plant growth

Plant growth problem:

- market gardener wants to plants to grow as much as possible within a fixed window of time $[t_0, t_1] = [0, 1]$
- supplement natural growth with lights as before
- growth rate dictates $\dot{x} = 1 + u$
- cost of lights

$$F\{u\} = \int_0^1 \frac{1}{2} u(t)^2 \, dt$$

- value of crop is proportional to the height at $t_1 = 1$

$$\phi(t_1, x(t_1)) = kx(1)$$

Plant growth problem statement

Minimize (equivalent to maximizing the profit)

$$F\{u,x\} = -kx(1) + \int_0^1 \frac{1}{2} u^2 \, dt = \int_0^1 \frac{1}{2} u^2 - k \, dt$$

Subject to $x(0) = 0$,

$$\dot{x} = 1 + u$$

- note that the extra constant in F will not effect the E-L equations, so the solution must still have the same form, i.e., $u = const$
- but the end conditions have changed

Plant growth

Including the Lagrange multiplier $\lambda(t) \left[\dot{x} - 1 - u \right]$

$$H\{u,x\} = \int_0^1 h(t,u,\dot{x}) + \frac{d}{dt} \phi(x) \, dt$$

where

$$\begin{aligned} h(t,u,\dot{x}) &= \frac{1}{2} u^2 + \lambda(t) \left[\dot{x} - 1 - u \right] \\ \phi(x) &= -kx \end{aligned}$$

Now the independent variable is t, and there are three dependent variables x, u, λ.

Plant growth: E-L equations

Three dependent variables, so three E-L equations

$$\frac{d}{dt}\frac{\partial h}{\partial \dot{x}} + \frac{\partial h}{\partial x} = 0 \quad (1)$$

$$\frac{d}{dt}\frac{\partial h}{\partial \dot{u}} + \frac{\partial h}{\partial u} = 0 \quad (2)$$

$$\frac{d}{dt}\frac{\partial h}{\partial \dot{\lambda}} + \frac{\partial h}{\partial \lambda} = 0 \quad (3)$$

Notice that $d\phi/dt$ is a constant, so it plays no part.

- h is linear in $\dot{x}$ so equation (1) is degenerate
- equation (2) gives us the E-L equation we had before
- equation (3) just gives us back the constraint

Plant growth: natural boundary cond.

Natural boundary conditions at $t_1 = 1$.

$$\frac{\partial h}{\partial \dot{x}} + \left.\frac{\partial \phi}{\partial x}\right|_{t_1} = 0$$

$$\frac{\partial h}{\partial \dot{u}} + \left.\frac{\partial \phi}{\partial u}\right|_{t_1} = 0$$

The second is trivial, i.e., $0 = 0$, so consider the first:

$$\frac{\partial h}{\partial \dot{x}} + \frac{\partial \phi}{\partial x} = \lambda - k = 0$$

We already know from the E-L equations that $\lambda = u$, and $u = const$, so the end result is that $u = k$.

Plant growth solution

The solution is $u = k$, and so

$$x(1) = 1 + k$$

When $k = 1$ we get the same solution we got before, but that isn't a general rule.

Also the optimization objective will be

$$F\{u,x\} = -1 - k + 1.5k^2$$

written in terms of profit we get

$$\text{profit} = 1 + k - 1.5k^2$$

Plant growth solution

Another way to see how the end-point conditions work

- The E-L equations still apply
- So u is still a constant
- $x(1) = 1 + u$ is the solution to the system DE

The height at $t_1 = 1$ would be $1 + u$ and so the profit would be

$$F\{u,x\} = 1 + ku - \int_0^1 \frac{1}{2}u^2\,dt = 1 + ku - \frac{1}{2}u^2$$

Clearly, the maximum here occurs for $u = k$.

Optimal Control

We will continue with optimal control later in the course when we have considered a bit more theory, but consider the following problem:

> Replace the previous plant growth problem by a similar problem, but instead of a terminal cost (related to value of plant), we aim to get the plants to height 2 in time that minimizes the cost.

Now t_1 is also a free variable – how can we deal with this?

Freeing up the independent variable

We can deal with both the optimal control problem and the collapsing beam by freeing up the value of the dependent variable.

Variational Methods & Optimal Control

lecture 18

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 18

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

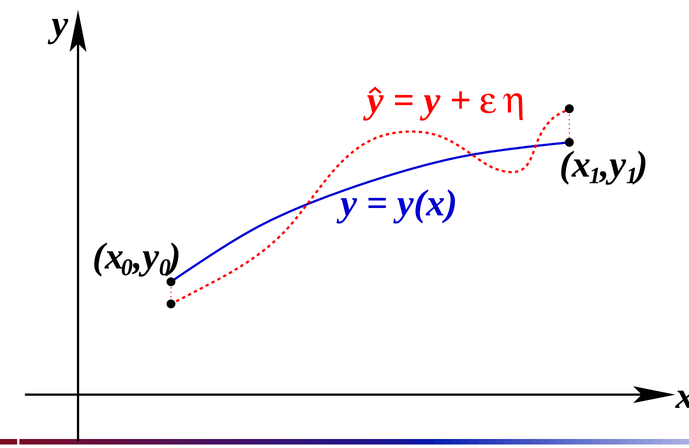
July 26, 2012

Free end points: Free x, y and y'

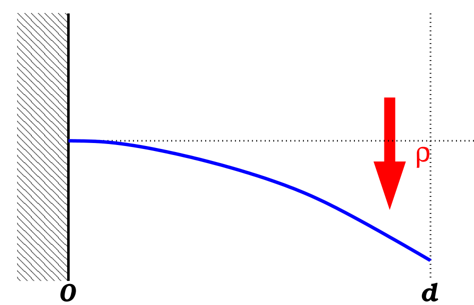
We now allow x to vary as well, although we may apply some condition on the relationship between x and y, for instance that the end point must lie on a curve. In these cases we often rename our extremals, and call them **transversals**.

Free end points

In previous problem, we allow $y(x_0)$ and $y(x_1)$ to vary but kept x_0 and x_1 fixed.



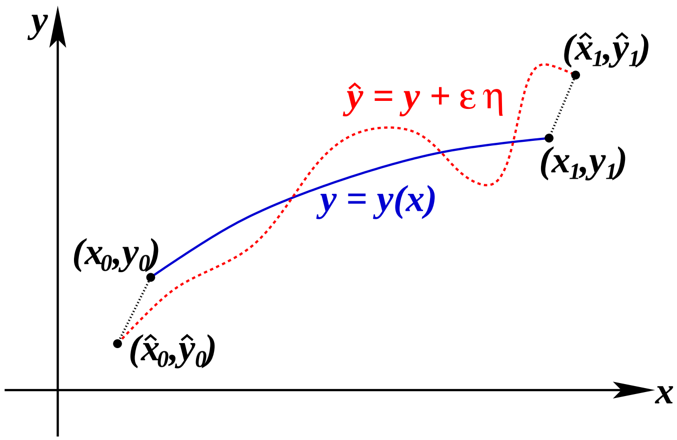
Example: Cantilever



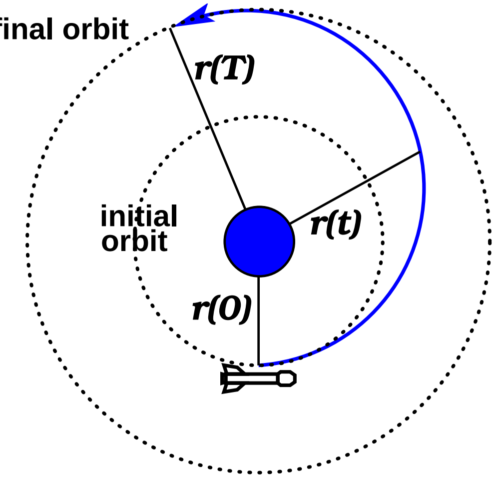
But this can fail in some cases, for instance, if the left end of the cantilever isn't clamped (to have zero slope) then the right end can swing freely, and x_1 won't be fixed.

Free end points

In some problems we even want to allow x_0 and x_1 to vary.

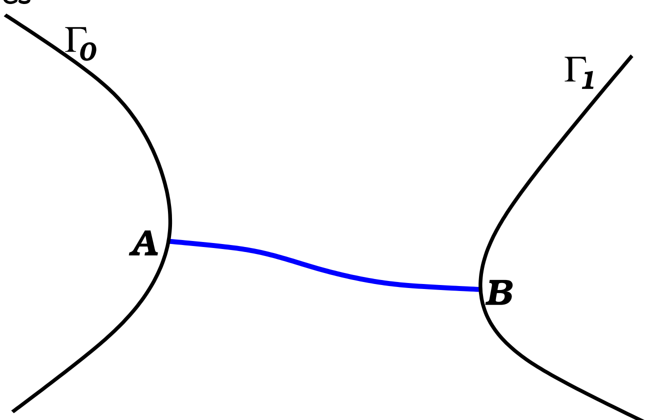


Example: Orbit Transfer Problem

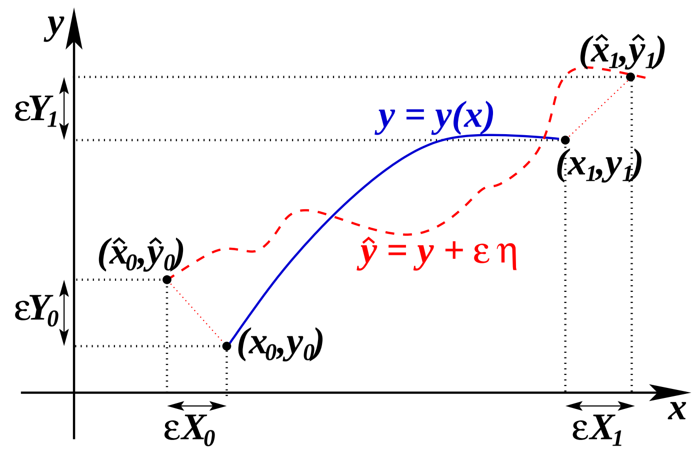


Example: shortest path

There may still be some constraints on the possible positions of end-points: e.g., shortest path between two curves



Approach



Extension of y

Define $\tilde{x}_0 = \min(x_0, \hat{x}_0)$ and $\tilde{x}_1 = \max(x_1, \hat{x}_1)$
We can use Taylor's theorem to extend y onto the interval $[\tilde{x}_0, \tilde{x}_1]$, e.g.

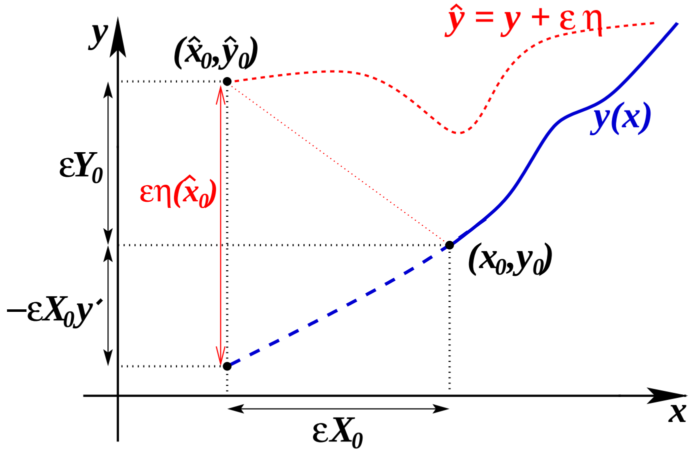
$$y(x) = \begin{cases} y(x) & \text{if } x \in [x_0, x_1] \\ y(x_1) + (x - x_1)y'(x_1) + \frac{(x-x_1)^2}{2} y''(x_1) + \cdots & \text{if } x \in (x_1, \tilde{x}_1] \\ y(x_0) + (x_0 - x)y'(x_0) + \frac{(x_0-x)^2}{2} y''(x_0) + \cdots & \text{if } x \in [\tilde{x}_0, x_0) \end{cases}$$

For instance, if the perturbed end-point $\hat{x}_0 < x_0$, we get

$$y(\hat{x}_0) = y(x_0) - \varepsilon X_0 y'(x_0) + O(\varepsilon^2)$$

We can likewise extend the perturbed curve $\hat{y}$.

Extension of y



Distance

However, we can no longer define distance as simply

- previous definition

$$d(y, \hat{y}) = \|y - \hat{y}\|$$

where the norm could be defined in a number of ways, but an example might be

$$\|y - \hat{y}\| = \int_{x_0}^{x_1} |y(x) - \hat{y}(x)|\, dx$$

- x_0 and x_1 can vary now, so the range of integral is not well defined anymore
- if we just extend y to new interval, we don't take proper account of distortion from difference in x end-points

New distance

New distance metric

$$d(y, \hat{y}) = \|y - \hat{y}\| + |\mathbf{p}_0 - \hat{\mathbf{p}}_0| + |\mathbf{p}_1 - \hat{\mathbf{p}}_1|$$

where we define

$$|\mathbf{p}_k - \hat{\mathbf{p}}_k| = \sqrt{(x_k - \hat{x}_k)^2 + (y_k - \hat{y}_k)^2}$$

We want allowed perturbations to be close to y (according to the distance defined above), but don't specify the end-points except to require they be $O(\varepsilon)$ apart, e.g.

$$\begin{aligned} \hat{x}_k &= x_k + \varepsilon X_k \\ \hat{y}_k &= y_k + \varepsilon Y_k \end{aligned}$$

so that $|\mathbf{p}_k - \hat{\mathbf{p}}_k| = \varepsilon\sqrt{X_k^2 + Y_k^2}$, for $k = 0, 1$

Forming the first variation

$$
\begin{aligned}
F\{\hat{y}\} - F\{y\} &= \int_{\hat{x}_0}^{\hat{x}_1} f(x,\hat{y},\hat{y}')\,dx - \int_{x_0}^{x_1} f(x,y,y')\,dx \\
&= \int_{x_0+\varepsilon X_0}^{x_1+\varepsilon X_1} f(x,\hat{y},\hat{y}')\,dx - \int_{x_0}^{x_1} f(x,y,y')\,dx \\
&= \int_{x_0}^{x_1} f(x,\hat{y},\hat{y}') - f(x,y,y')\,dx \\
&\quad + \int_{x_1}^{x_1+\varepsilon X_1} f(x,\hat{y},\hat{y}')\,dx - \int_{x_0}^{x_0+\varepsilon X_0} f(x,\hat{y},\hat{y}')\,dx
\end{aligned}
$$

Forming the first variation

From earlier arguments

$$\int_{x_0}^{x_1} f(x,\hat{y},\hat{y}') - f(x,y,y')\,dx = \varepsilon \left[\eta \frac{\partial f}{\partial y'}\bigg|_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} \right) dx \right]$$

and as ε is small

$$
\begin{aligned}
\int_{x_1}^{x_1+\varepsilon X_1} f(x,\hat{y},\hat{y}')\,dx &= \varepsilon X_1\, f(x,y,y')|_{x_1} + O(\varepsilon^2) \\
\int_{x_0}^{x_0+\varepsilon X_0} f(x,\hat{y},\hat{y}')\,dx &= \varepsilon X_0\, f(x,y,y')|_{x_0} + O(\varepsilon^2)
\end{aligned}
$$

Forming the first variation

Therefore the first variation is

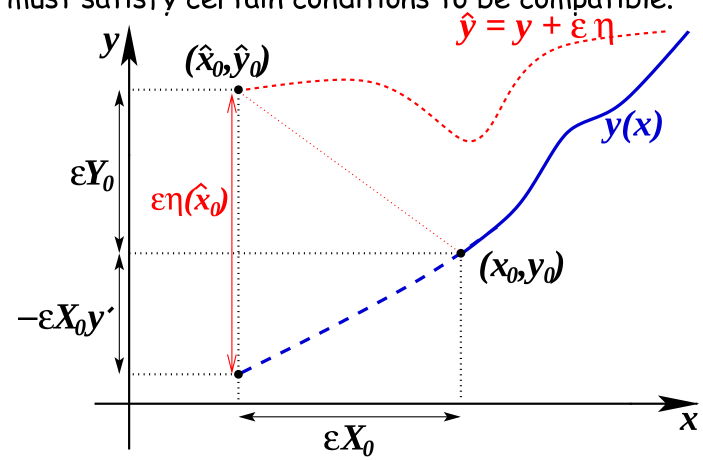
$$
\begin{aligned}
\delta F(\eta,y) &= \left[\eta \frac{\partial f}{\partial y'} \right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} \right) dx \\
&\quad + X_1\, f(x,y,y')|_{x_1} - X_0\, f(x,y,y')|_{x_0} + O(\varepsilon)
\end{aligned}
$$

But note that $\left[\eta \frac{\partial f}{\partial y'} \right]_{x_0}^{x_1}$ is no longer simple to calculate because we don't fix x_0 or x_1.

- how can we learn x_0 and x_1?
- we need a new natural boundary condition that will give us this.

End-point compatibility

The perturbed end-points, and perturbation function η must satisfy certain conditions to be compatible.



End-point compatibility

Remember that

$$\begin{aligned}\hat{x}_0 &= x_0 + \varepsilon X_0 \\ \hat{y}_0 &= y_0 + \varepsilon Y_0\end{aligned}$$

Notice that

$$\hat{y}_0 = \hat{y}(\hat{x}_0) = \hat{y}(x_0+\varepsilon X_0) = y(x_0+\varepsilon X_0) + \varepsilon\eta(x_0+\varepsilon X_0)$$

From Taylor's theorem, for small ε

$$\begin{aligned}y(x_0+\varepsilon X_0) &= y(x_0) + \varepsilon X_0 y'(x_0) + O(\varepsilon^2) \\ &= y_0 + \varepsilon X_0 y'(x_0) + O(\varepsilon^2) \\ \varepsilon\eta(x_0+\varepsilon X_0) &= \varepsilon\eta(x_0) + O(\varepsilon^2)\end{aligned}$$

End-point compatibility

So

$$\begin{aligned}y_0 + \varepsilon Y_0 &= y_0 + \varepsilon X_0 y'(x_0) + \varepsilon\eta(x_0) + O(\varepsilon^2) \\ \varepsilon Y_0 &= \varepsilon X_0 y'(x_0) + \varepsilon\eta(x_0) + O(\varepsilon^2) \\ \eta(x_0) &= Y_0 - X_0 y'(x_0) + O(\varepsilon)\end{aligned}$$

Similarly

$$\eta(x_1) = Y_1 - X_1 y'(x_1) + O(\varepsilon)$$

The First Variation

Substituting the end-point compatibility constraints into the first variation we get

$$\begin{aligned}\delta F(\eta,y) &= \left[\eta\frac{\partial f}{\partial y'}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right) dx \\ &\quad + X_1\, f(x,y,y')|_{x_1} - X_0\, f(x,y,y')|_{x_0} + O(\varepsilon) \\ &= \int_{x_0}^{x_1} \eta\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right) dx \\ &\quad + Y_1 \left.\frac{\partial f}{\partial y'}\right|_{x_1} - Y_0 \left.\frac{\partial f}{\partial y'}\right|_{x_0} \\ &\quad + X_1\left.\left(f - y'\frac{\partial f}{\partial y'}\right)\right|_{x_1} - X_0\left.\left(f - y'\frac{\partial f}{\partial y'}\right)\right|_{x_0} + O(\varepsilon)\end{aligned}$$

Deriving Euler-Lagrange equations

The end-points are free, but this includes the case where they sit on the extremal, i.e. we can always choose the end-points so that $X_k = Y_k = 0$, for $k = 0,1$. For instance, when $X_0 = X_1 = Y_1 = Y_2 = 0$, then the first variation collapses to

$$\delta F(\eta,y) = \int_{x_0}^{x_1} \eta\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right) dx$$

And so the E-L equations hold here.
Likewise, when $X_1 = Y_1 = Y_2 = 0$, but $X_0 \neq 0$ we can see that this creates one of the natural boundary condition

$$X_0\left.\left(f - y'\frac{\partial f}{\partial y'}\right)\right|_{x_0} = 0$$

Notation

Some notation

- Hamiltonian
$$H = y'\frac{\partial f}{\partial y'} - f$$
we saw the Hamiltonian H earlier.
- p is often identified with momentum of a particle, but we can use it for other systems as well.
$$p = \frac{\partial f}{\partial y'}$$
- we'll replace the notations X_k and Y_k for $k = 0,1$ with
$$\delta x(x_k) = X_k \quad \text{and} \quad \delta y(y_k) = Y_k$$

The Euler-Lagrange equations

As before, we can always choose the end-points so that $X_k = Y_k = 0$, for $k = 0,1$, so that the Euler-Lagrange equation

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0$$

must be satisfied plus the additional constraints:

$$\Big[p\delta y - H\delta x\Big]_{x_0}^{x_1} = 0$$

Including constraints

Typically the end-points satisfy some set of constraints, in the most general form $g(x_0,y_0,x_1,y_1) = 0$, but often these constraints separate to constraint a single end-point, e.g. we have constraints

$$g_k(x_j,y_j) = 0$$

for $j = 0,1$, and some number of constraints, typically $k < 4$.

For example, the fixed end-point problem has constraints that specify the values of (x_0,y_0) and (x_1,y_1) precisely.

Separable constraints

Where the constraints for one end point are not linked to those of the other, we may separate the conditions to get

$$\begin{aligned} p\delta y - H\delta x\Big|_{x_0} &= 0 \\ p\delta y - H\delta x\Big|_{x_1} &= 0 \end{aligned}$$

Note not all possible end constraints make sense!

Simple example: fixed x

We have already considered this condition:

- $\delta x = 0$ and $\delta y \neq 0$
- conditions

$$p\delta y - H\delta x\Big|_{x_i} = 0$$

reduce down to

$$p = \frac{\partial f}{\partial y'}\Big|_{x_i} = 0$$

at the relevant end points.

- that is just the natural boundary conditions we derived earlier

Simple example: fixed y

Imagine a problem where we have to get to a fixed state y, but the point at which that happens is variable, so that

- $\delta y = 0$ and $\delta x \neq 0$
- conditions

$$p\delta y - H\delta x\Big|_{x_i} = 0$$

reduce down to

$$H\Big|_{x_i} = 0$$

at the relevant end points.

Simple example: fixed y

Minimise

$$F\{y\} = \int_0^{x_1} 1 + y'^2 \, dx$$

subject to $y(0) = 1$ and $y(x_1) = L > 1$, but with x_1 unspecified.

- We could derive the E-L equations, but note that this problem is autonomous (no x dependence) so

$$H = const$$

- The free end point at x_1 means that

$$H\Big|_{x_1} = 0$$

- Hence for all $x \in [0, x_1]$ we have $H = 0$

Simple example: fixed y

Minimise

$$F\{y\} = \int_0^{x_1} 1 + y'^2 \, dx$$

So

$$H = y'\frac{\partial f}{\partial y'} - f = 2y'^2 - y'^2 - 1 = y'^2 - 1 = 0$$

Hence

$$y' = \pm 1$$

subject to $y(0) = 1$ and $y(x_1) = L > 1$ so we take $y' = 1$

$$y = x + 1$$

Extension to several dep. var.s

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

If F is stationary at $\mathbf{q}$ then it can be shown that the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0$$

for $k = 1, \ldots, n$ and that at the end points t_0 and t_1

$$\sum_{k=1}^{n} p_k \delta q_k - H\delta t = 0 \text{ where } p_k = \frac{\partial L}{\partial \dot{q}_k} \text{ and } H = \sum_{k=1}^{n} \dot{q}_k p_k - L$$

Simple Example

Find extremals of

$$F\{\mathbf{q}\} = \int_0^1 \left(\dot{q}_1^2 + (\dot{q}_2 - 1)^2 + q_1^2 + q_1 q_2\right) dt$$

for $\mathbf{q}(0) = \mathbf{q}_0$ and $\mathbf{q}(1)$ free, i.e., we can finish anywhere on the plane $t = 1$.

The Euler-Lagrange equations are

$$\begin{aligned} 2\ddot{q}_1 - 2q_1 - q_2 &= 0 \\ 2\ddot{q}_2 - q_1 &= 0 \end{aligned}$$

Simple example

As earlier we can combine the E-L equations to get

$$4q_2^{(4)} - 4\ddot{q}_2 - q_2 = 0$$

which has solutions in the form

$$q_2(t) = c_1 e^{\mu_1 t} + c_2 e^{\mu_2 t} + c_3 \cos(mt) + c_4 \sin(mt)$$

where

$$\begin{aligned} \mu_1, \mu_2 &= \pm\sqrt{\frac{1}{2} + \frac{1}{\sqrt{2}}} \\ \mu_3, \mu_4 &= \pm\sqrt{\frac{1}{2} - \frac{1}{\sqrt{2}}} = \pm im \end{aligned}$$

Simple example

Natural boundary conditions

$$\left.\sum_{k=1}^{n} p_k \delta q_k - H\delta t\right|_{t=1} = 0$$

but $t = 1$ is fixed at the RHS, so $\delta t = 0$, and we can vary q_k independently, so we can take any combination of $\delta q_k = 0$, and hence all of the $p_k = 0$ at $t = 1$, i.e.,

$$p_k|_{t=1} = \frac{\partial L}{\partial \dot{q}_k} = 0$$

Simple example

$$p_k|_{t=1} = \frac{\partial L}{\partial \dot{q}_k} = 0$$

So

$$\begin{aligned} p_1 &= 2\dot{q}_1 = 0 \\ p_2 &= 2(\dot{q}_2 - 1) = 0 \end{aligned}$$

The natural boundary conditions reduce to

$$\begin{aligned} \dot{q}_1 &= 0 \\ \dot{q}_2 &= 1 \end{aligned}$$

Combine with the conditions at the start point we have enough constraints to find the constants of integration.

Variational Methods & Optimal Control

lecture 19

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 19

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

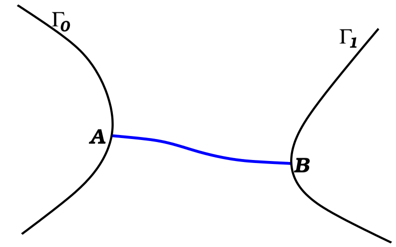
July 26, 2012

Transversals

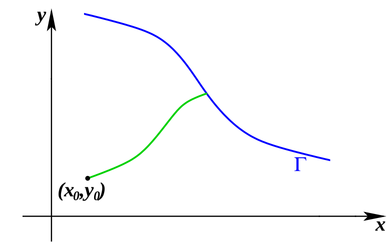
When we consider an extremal joining a curve to a point (or two curves) then we often call the extremal a transversal. The free-end-point condition simplifies in many such cases, for instance, in many situations we look for a transversal that joins the proscribed curve at right angles.

Transversal examples

Find the shortest path between two curves



Find the shortest path from a point (x_0, y_0) to a curve Γ.



Transversals

Specify the curve Γ parametrically by $(x_\Gamma(\xi), y_\Gamma(\xi))$, then the end-points must lie on this line, and so we can write

$$\delta x = \delta\xi \frac{dx_\Gamma}{d\xi}$$

$$\delta y = \delta\xi \frac{dy_\Gamma}{d\xi}$$

and then the condition is

$$\left. p\delta y - H\delta x \right|_{x_1} = 0$$

$$\left. p\frac{dy_\Gamma}{d\xi} - H\frac{dx_\Gamma}{d\xi} \right|_{x_1} = 0$$

Transversal condition

Note that the vector $(\frac{dx_\Gamma}{d\xi}, \frac{dy_\Gamma}{d\xi})$ is a tangent to the curve Γ.

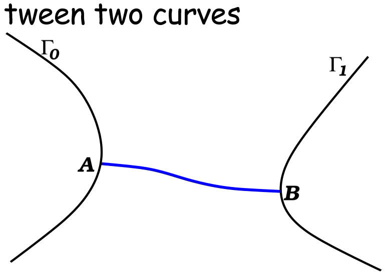
The **Transversality Condition** is that the vector $\mathbf{v} = (-H, p)$ is orthogonal to the tangent vector.

i.e.

$$\left(\frac{dx_\Gamma}{d\xi}, \frac{dy_\Gamma}{d\xi}\right) \cdot (-H, p) = p\frac{dy_\Gamma}{d\xi} - H\frac{dx_\Gamma}{d\xi} = 0$$

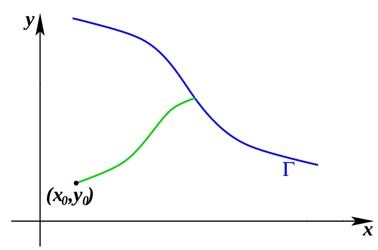
Transversals

Find the shortest path between two curves



$$p\frac{dy_{\Gamma_0}}{d\xi} - H\frac{dx_{\Gamma_0}}{d\xi}\bigg|_{x_0} = 0$$
$$p\frac{dy_{\Gamma_1}}{d\xi} - H\frac{dx_{\Gamma_1}}{d\xi}\bigg|_{x_1} = 0$$

Find the shortest path from a point (x_0, y_0) to a curve Γ.



$$p\frac{dy_\Gamma}{d\xi} - H\frac{dx_\Gamma}{d\xi}\bigg|_{x_1} = 0$$

Example 1

Shortest path from the origin to a curve $\mathbf{r}_\Gamma = (x_\Gamma(\xi), y_\Gamma(\xi))$. The path length is given by

$$F\{y\} = \int_0^{x_1} \sqrt{1+y'^2}\, dx$$

Then

$$\begin{aligned} p &= \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1+y'^2}} \\ H &= y'\frac{\partial f}{\partial y'} - f = \frac{y'^2}{\sqrt{1+y'^2}} - \sqrt{1+y'^2} = \frac{-1}{\sqrt{1+y'^2}} \end{aligned}$$

Example 1

Thus the transversality condition becomes

$$\begin{aligned} p\frac{dy_\Gamma}{d\xi} - H\frac{dx_\Gamma}{d\xi}\bigg|_{x_1} &= 0 \\ \frac{y'}{\sqrt{1+y'^2}}\frac{dy_\Gamma}{d\xi} + \frac{1}{\sqrt{1+y'^2}}\frac{dx_\Gamma}{d\xi}\bigg|_{x_1} &= 0 \end{aligned}$$

Now $\sqrt{1+y'^2} \neq 0$, so we can multiply through by $\sqrt{1+y'^2}$ to give

$$\frac{dx_\Gamma}{d\xi} + y'\frac{dy_\Gamma}{d\xi}\bigg|_{x_1} = 0$$

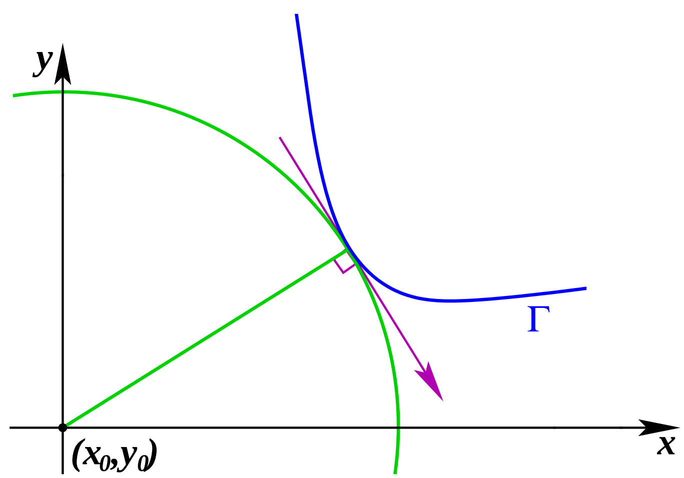
Alternatively stated, $(\frac{dx_\Gamma}{d\xi}, \frac{dy_\Gamma}{d\xi}) \cdot (1, y') = 0$

Example 1

We can interpret $(\frac{dx_\Gamma}{d\xi}, \frac{dy_\Gamma}{d\xi}) \cdot (1, y') = 0$ geometrically

- the condition means that the tangent to the extremal must be orthogonal to the tangent to the curve Γ where they connect.
- E-L equations still imply $y(x)$ will be straight line
- this makes perfect sense!
  - find the distance of curve Γ from the origin.
  - do this by creating expanding circles, and the one that touched the curve would give us the distance.
  - it would touch so the circle was tangent
  - the (straight line) radius would be perpendicular to the tangent.

Example 1



Example 1

Sometimes, there will be many possible solutions, for instance if the curve Γ was a circle around the origin!. But now we know how to find them, it would be easy.

Example 2

Consider the general functional

$$F\{y\} = \int_0^{x_1} K(x,y)\sqrt{1+y'^2}\,dx$$

for which we wish to find stationary paths between two curves $\mathbf{r}_{\Gamma_0} = (x_{\Gamma_0}(\xi), y_{\Gamma_0}(\xi))$, and $\mathbf{r}_{\Gamma_1} = (x_{\Gamma_1}(\xi), y_{\Gamma_1}(\xi))$. The path length is given by Then

$$p = \frac{\partial f}{\partial y'} = \frac{y' K(x,y)}{\sqrt{1+y'^2}}$$

$$H = y'\frac{\partial f}{\partial y'} - f = \frac{y'^2 K(x,y)}{\sqrt{1+y'^2}} - K(x,y)\sqrt{1+y'^2} = \frac{-K(x,y)}{\sqrt{1+y'^2}}$$

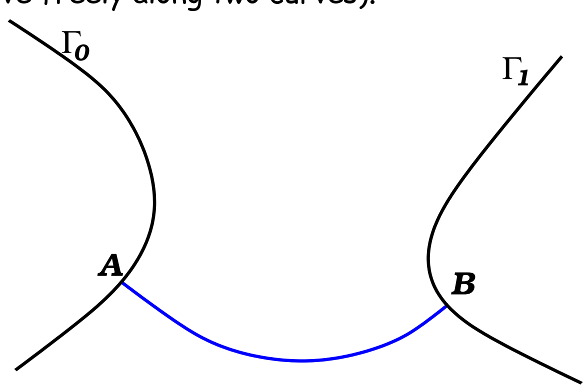
Example 2

Once again the transversality conditions at each end will revert to the extremal being orthogonal to the tangent to the curves at either end.

However, in this case, the curve joining the two could be distorted by the factor of $K(x,y)$ so that it is no longer a straight line. Its shape can be determined from the E-L equations.

Example 3

Find the shape of a fixed length chain hanging between two curves (similar to catenary problem, but end-points can move freely along two curves).



Example 3

This problem is a special case of Example 2, and so

- from transversality constraints that the chain will join the two curves at a right angle.
- E-L equations imply the curve will be a catenary (see earlier lectures for the derivation of the catenary)

$$y(x) + \lambda = c_1 \cosh\left(\frac{x - c_2}{c_1}\right)$$

- we need simply to use the perpendicularity (and fixed length) constraints to derive the values of the constant of integration, and the Lagrange multiplier.

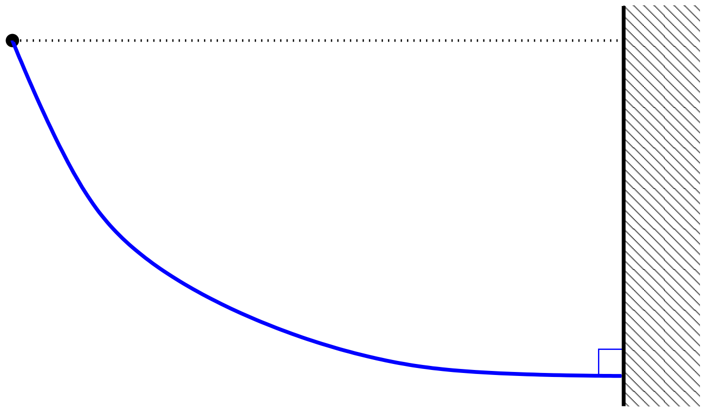
Example 4

A variant of the Brachystochrone: find the curve of fastest descent from a point to line.

$$T\{y\} = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{\sqrt{\frac{2E}{m} - 2gy(x)}}\,dx = \int_{x_0}^{x_1} K(y)\sqrt{1+y'^2}\,dx$$

- E-L equations show that the curve must be a cycloid
- Transversality constraints (see Example 2) show that, at the point of contact, the extremal will be perpendicular to the line.

Example 4



Example 5

Shortest path from a point to a surface.

- E-L equations show that the curve must be a straight line
- Transversality constraints show that, at the point of contact, the extremal will be normal to the surface.
  (see CE 5 solutions for an example that shows this).

Variational Methods & Optimal Control

lecture 20

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 20

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

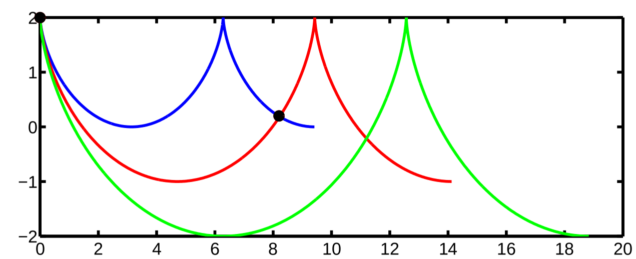
Broken Extremals

Until now we have required that extremal curves have at least two well-defined derivatives. Obviously this is not always true (see for instance Snell's law). In this lecture we consider the alternatives.

Broken extremals

Broken extremals are continuous extremals for which the gradient has a discontinuity at one of more points.

If a variational problem has a smooth extremal (that therefore satisfies the E-L equations), this will be better than a broken one, e.g. Brachystochrone.



Broken extremals

But some problems don't admit smooth extremals

Example: Find $y(x)$ to minimize

$$F\{y\} = \int_{-1}^{1} y^2 (1-y')^2 \, dx$$

subject to $y(-1) = 0$ and $y(1) = 1$.

Broken extremals example

There is no explicit x dependence inside the integral, so we can find

$$\begin{aligned} H(y,y') = y'\frac{\partial f}{\partial y'} - f &= const \\ y'y^2(-2)(1-y') - y^2(1-y')^2 &= -c_1 \\ y^2(1-y')(-1+y'-2y') &= -c_1 \\ y^2(1-y')(-1-y') &= -c_1 \\ y^2(1-y'^2) &= c_1 \end{aligned}$$

If $c_1 = 0$ we get the singular solutions $y = 0$ and $y = \pm x + B$
None of these satisfies the end-points conditions
$y(-1) = 0$ and $y(1) = 1$, so $c_1 \neq 0$

Broken extremals example

Given $c_1 \neq 0$

$$\begin{aligned} y^2(1-y'^2) &= c_1 \\ y'^2 &= \frac{y^2-c_1}{y^2} \\ \frac{dy}{dx} &= \pm\frac{1}{y}\sqrt{y^2-c_1} \\ dx &= \pm\frac{y}{\sqrt{y^2-c_1}}dy \\ x &= \pm\sqrt{y^2-c_1}+c_2 \\ (x-c_2)^2 &= y^2-c_1 \end{aligned}$$

The solution is a rectangular hyperbola

Broken extremals example

Find c_1 and c_2 from

$$(x-c_2)^2 = y^2 - c_1$$

using the end-points.

$$\begin{aligned} y(-1) &= 0 \;\Rightarrow\; (-1-c_2)^2 = -c_1 \\ y(1) &= 1 \;\Rightarrow\; (1-c_2)^2 = 1-c_1 \end{aligned}$$

Combine the two equations

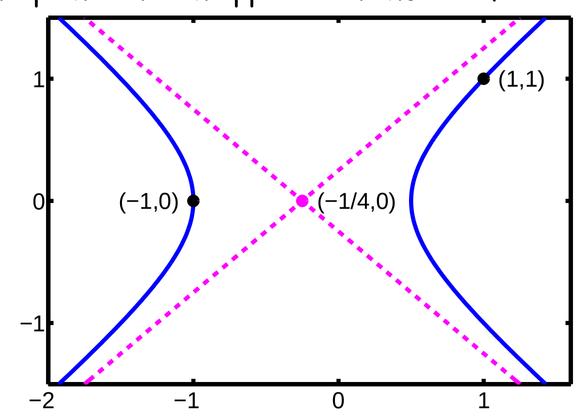
$$(1-c_2)^2 = 1 + (1+c_2)^2$$

which has solutions $c_2 = -1/4$, and so $c_1 = -9/16$

$$y^2 = (x+1/4)^2 - 9/16$$

Broken extremals example

The end-points are on opposite branches of the hyperbola!



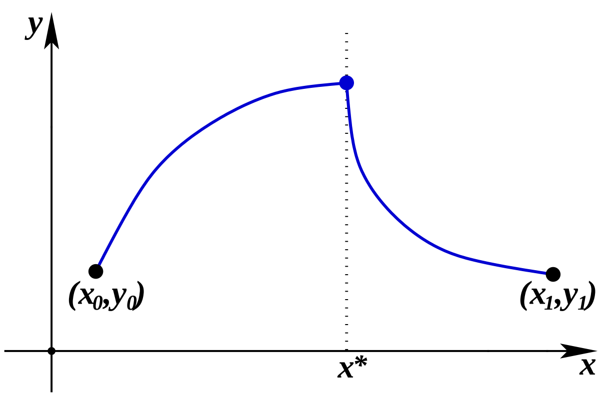
There is **NO** smooth extremal curve that connects $(-1,0)$ and $(1,1)$

Broken extremal

- sometimes there is no smooth extremal
- we must seek a broken extremal
- still want a continuous extremal
- what should we do?
  - previous smoothness results suggest that we should use a smooth extremal when we can, and so we will try to minimize the number of corners.
  - We'll start by looking for curves with one corner
  - But can we apply E-L equations?

Broken extremal

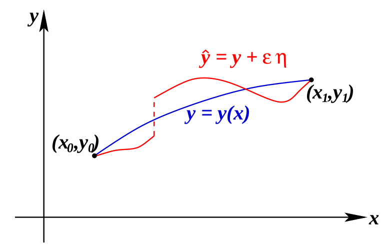
If we have an extremal like this, can we use E-L equations?



Smoothness theorem

Theorem: If the smooth curve $y(x)$ gives an extremal of a functional $F\{y\}$ over the class of all admissible curves in some ε neighborhood of y, then $y(x)$ also gives an extremal of a functional $F\{y\}$ over the class of all piecewise smooth curves in the same neighborhood.

Meaning: we can extend our results to piecewise smooth curves (where a smooth result exists), not just curves with 2 continuous derivatives.



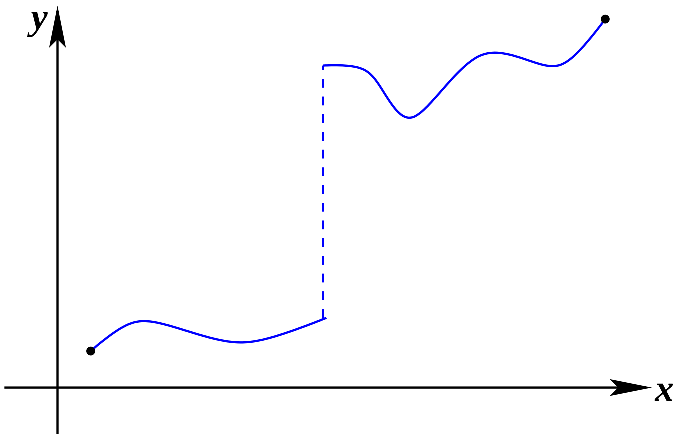
Proof sketch

The theorem assumes that there exists a smooth extremal (in this case a minimum for the purpose of illustration) y, then for any other smooth curve $\hat{y} \in B_\varepsilon(y)$ we know $F\{\hat{y}\} > F\{y\}$.

Assume for the moment that for a piecewise smooth function $\tilde{y} \in B_\varepsilon(y)$ that $F\{\tilde{y}\} < F\{y\}$. We can approximate $\tilde{y}$ by a smooth curve $\hat{y}_\delta \in B_\varepsilon(y)$ by rounding off the edges of the discontinuity.

Given that we can approximate the curve $\tilde{y}$ arbitrarily closely by a smooth curve $\hat{y}_\delta$, for which we already know $F\{\hat{y}_\delta\} > F\{y\}$, we get a contradiction with $F\{\tilde{y}\} < F\{y\}$, and so no such alternative extremal can exist.

Proof sketch

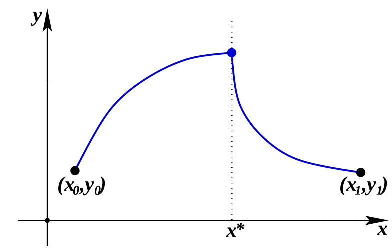


So what do we do?

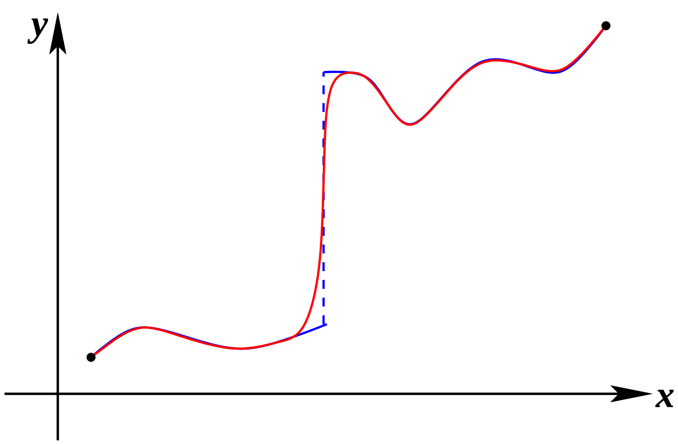
Break the functional into two parts:

$$F\{y\} = F_1\{y\} + F_2\{y\} = \int_{x_0}^{x^*} f(x, y_1, y_1')\,dx + \int_{x^*}^{x_1} f(x, y_2, y_2')\,dx$$

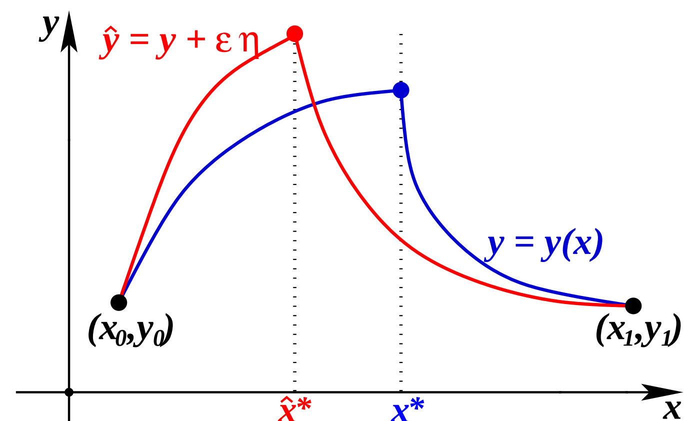
where we require y to have two continuous derivatives everywhere except at x^*, and $y_1(x^*) = y_2(x^*)$



Proof sketch



Possible perturbations



The location of the "corner" can also be perturbed.

The First Variation: part 1

We get first component of the first variation by considering a problem with only one fixed end-point, and allowing x^* to vary, so that

$$\delta F_1(\eta, y) = \lim_{\varepsilon \to \infty} \frac{1}{\varepsilon} \left[\int_{x_0}^{\hat{x}^*} f(x, \hat{y}_1, \hat{y}_1')\, dx - \int_{x_0}^{x^*} f(x, y_1, y_1')\, dx \right]$$

And as with transversals, we get an integral term which results in the E-L equation, plus the additional term

$$p_1 \delta y - H_1 \delta x \Big|_{x^*}$$

where

$$\delta x(x^*) = X^* \quad \text{and} \quad \delta y(y_1^*) = Y^*$$
$$H_1 = y_1' \frac{\partial f}{\partial y_1'} - f \quad \text{and} \quad p_1 = \frac{\partial f}{\partial y_1'}$$

The First Variation: part 2

Note that, for the second component of the First Variation we get a similar extra term, e.g. $\delta F_2(\eta, y)$ introduces the term

$$-p_2 \delta y + H_2 \delta x \Big|_{x^*}$$

the sign is reversed because it corresponds to the x_0 term in the transversal problem (as opposed to the x_1 term for δF_1.

The combined second variation (minus the terms that result from the E-L equation which must be zero) is

$$\delta F(\eta, y) = \delta F_1(\eta, y) + \delta F_2(\eta, y) = p_1 \delta y - H_1 \delta x - p_2 \delta y + H_2 \delta x \Big|_{x^*}$$

Conditions

We rearrange to give

$$\delta F(\eta, y) = (p_1 - p_2)\delta y - (H_1 - H_2)\delta x \Big|_{x^*}$$

Note that the point of discontinuity may vary freely, so we may independently vary δx and δy or set one or both to zero. Hence, we can separate the condition to get two conditions

$$p_1 - p_2 \Big|_{x^*} = 0$$
$$H_1 - H_2 \Big|_{x^*} = 0$$

Weierstrass-Erdman

We can write the conditions as

$$p_1 \Big|_{x^*} = p_2 \Big|_{x^*}$$
$$H_1 \Big|_{x^*} = H_2 \Big|_{x^*}$$

Called the Weierstrass-Erdman Corner Conditions

Rather than separating y into y_1 and y_2 we may write the corner conditions in terms of limits from the left and right, e.g.

$$p \Big|_{x^{*-}} = p \Big|_{x^{*+}}$$
$$H \Big|_{x^{*-}} = H \Big|_{x^{*+}}$$

Solution

So the broken extremal solution must satisfy

- the E-L Equations
- the Weierstrass-Erdman Corner Conditions

$$p\Big|_{x^{*-}} = p\Big|_{x^{*+}}$$
$$H\Big|_{x^{*-}} = H\Big|_{x^{*+}}$$

must hold at any 'corner'

Example 1

In the example considered,

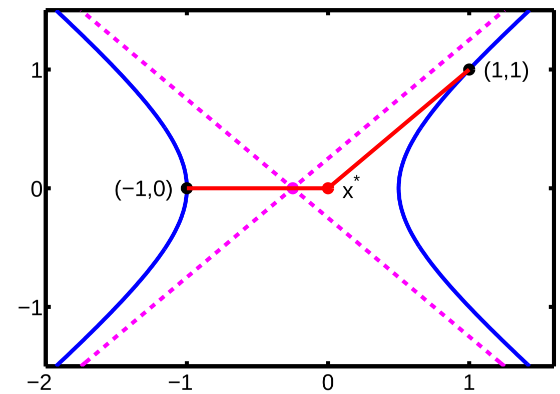
$$p = \frac{\partial f}{\partial y'} = -2y^2(1-y')$$
$$H = y'\frac{\partial f}{\partial y'} - f = y^2(1-y'^2)$$

Remember that $y=0$ and $y=x+A$ are valid solutions to the E-L equations, and that for both of these solutions $p=H=0$, so we can put a 'corner' where needed.

The solution must also satisfy the end-point conditions, so $y(-1)=0$ and $y(1)=1$, and therefore, as valid solution has $x^*=0$ and

$$y_1 = 0 \text{ for } x \in [-1, x^*]$$
$$y_2 = x \text{ for } x \in [x^*, 1]$$

Example 1



Example 2

Example: See Fermat's principle, and Snell's law in earlier lectures.

- sometimes the discontinuity arise from the problem itself, e.g. a discontinuous boundary, such as in refraction

General strategy

- solve E-L equations
- look for solutions for each end condition
- match up the solutions at a corner x^* so that
  - $y_1(x^*) = y_2(x^*)$
  - the Weierstrass-Erdman Corner Conditions are satisfied
- in theory can allow more than one corner, but this would get very painful!

Newton's aerodynamical problem

Find extremal of "air resistance"

$$F\{y\} = \int_0^R \frac{x}{1+y'^2}\, dx,$$

subject to $y(0) = L$ and $y(R) = 0$ with solutions

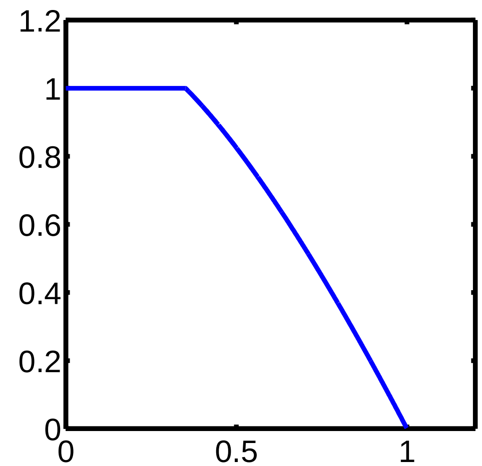
1. $y = const$ for $x \in [0, x_1]$
2. $u \in [u_1, u_2]$

$$x(u) = \frac{c}{u}(1+u^2)^2 = c\left(\frac{1}{u} + 2u + u^3\right).$$

$$y(u) = L - c\left(-\ln u - A + u^2 + \frac{3}{4}u^4\right)$$

Tricky bit is working out u_1 which sets the location of the "corner", and fixes A, c and u_2.

Newton's aerodynamical problem



Newton's aerodynamical problem

- we could find u_1 by trying to minimize F as a function of u_1, but this is hard because we only have a numerical solution to get u_2.
- alternative is to use corner conditions
  1. at the corner
     (a) $x^* = x(u_1)$ is free
     (b) $y = L$ is fixed
  2. corner condition of interest is

$$H\Big|_{x^{*-}} = H\Big|_{x^{*+}}$$

Newton's aerodynamical problem

Calculating H

$$\begin{aligned} H &= y'\frac{\partial f}{\partial y'} - f \\ &= \frac{-2y'^2 x}{(1+y'^2)^2} - \frac{x}{(1+y'^2)} \\ &= \frac{-x}{(1+y'^2)^2}\left[2y'^2 + (1+y'^2)\right] \\ &= \frac{-x}{(1+y'^2)^2}\left[3y'^2 + 1\right] \end{aligned}$$

Newton's aerodynamical problem

Corner condition

$$H = \frac{-x}{(1+y'^2)^2}\left[2y'^2 + 1\right]$$

Now on the LHS of $x_1 = x^*$ we have $y' = 0$, so

$$H\Big|_{x^{*-}} = -x^*$$

On the RHS

$$H\Big|_{x^{*+}} = \frac{-x^*}{(1+u^2)^2}\left[3u^2 + 1\right]$$

Newton's aerodynamical problem

$$\begin{aligned} H\Big|_{x^{*-}} &= H\Big|_{x^{*+}} \\ -x^* &= \frac{-x^*}{(1+u^2)^2}\left[3u^2+1\right] \\ (1+u^2)^2 &= 3u^2+1 \\ u^4 - u^2 &= 0 \\ u^2(u^2-1) &= 0 \\ u &= 0 \text{ or } \pm 1 \end{aligned}$$

but $-y' = u > 0$ so $u = 1$ is the only valid solution, hence

$$u_1 = 1$$

and the rest of the solution follows from there.

Newton's aerodynamical problem

- real rockets don't look like this
  1. resistance functional is only approximate
     (a) ignores friction
     (b) ignores shock waves
  2. rockets must pass through multiple layers of atmosphere, at varying speeds
- additional constraints:
  1. nose cone is tangent to rocket at joint

     $$y'(R) = -\infty$$

  2. nose is easy to build
- really, we need to do CFD++

Variational Methods & Optimal Control

lecture 21

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 21

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

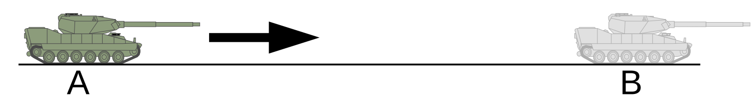
Inequality Constraints and Optimal Control

Earlier we didn't consider inequalities as constraints, but these are needed particularly in control. For instance, often there is a maximum force we can apply to an object. The resulting extremals either (i) satisfy the E-L equations, or (ii) lie along the edge of the constraint. We also get boundary conditions between these two types of regions.

Example: parking a car

Classic problem: from *Craggs*, p.55

We want to drive a car/tank from point A to point B as quickly as possible, and at point B the car should be stationary.



Example

Parking a car seems like a trivial problem:

- in fact this problem appears in other contexts, e.g.
  - automatic positioning of components on a circuit board
  - has to be done frequently (so has to be fast)
  - speed limited by robot, and how delicate the components are
- shortest-time problems are a case of a more general type of problem as well.
- further, this type of controller appears often
  - we can make some general statements about when a bang-bang controller is a good idea

Example

http://www.expo21xx.com/automation77/news/2085_robot_mitsubishi/news_default.htm

Example: parking a car

We want to drive a car/tank from point A to point B as quickly as possible, and at point B the car should be stationary.

Newton's law

$$\text{force} = u = m\ddot{x}$$

Choose force u that minimizes the time subject to $\dot{x} = 0$ at $t = 0$ and $t = T$, where T is not specified, but rather given by

$$T\{u\} = \int_A^B dt$$

and it is this functional we wish to minimize.

Example: parking a car

As before, note $\dot{x}(t) = dx/dt$ is the car's velocity, so we can write

$$T\{x\} = \int_A^B dt = \int_{x_A}^{x_B} \frac{1}{\dot{x}} dx$$

We wish to maximize this extremal, subject to the DE constraint that

$$\ddot{x} = \frac{u(t)}{m}$$

where $u(t)$ is the control (force) that we exert, and also subject to

$$\dot{x}(0) = \dot{x}(T) = 0$$

i.e., the car is stationary at the start and finish.

Example: parking a car

Take $y = \dot{x}$, and we can rewrite the problem as minimize

$$T\{y\} = \int_A^B dt = \int_{x_A}^{x_B} \frac{1}{y} dx$$

We wish to minimize this extremal, subject to the DE constraint that

$$\dot{y} = \frac{u(t)}{m}$$

where $u(t)$ is the control (force) that we exert, and also subject to

$$y(x_A) = y(x_B) = 0$$

Example: parking a car

Including the non-holonomic constraint into the problem using a Lagrange multiplier we get

$$H\{y,u\} = \int_{x_A}^{x_B} \frac{1}{y} + \lambda \left(\dot{y} - \frac{u(t)}{m} \right) \, dx$$

subject to

$$y(x_A) = y(x_B) = 0$$

The E-L equations are

$$\begin{aligned} \frac{d}{dt}\frac{\partial h}{\partial \dot{y}} - \frac{\partial h}{\partial y} &= 0 \\ \frac{d}{dt}\frac{\partial h}{\partial \dot{u}} - \frac{\partial h}{\partial u} &= 0 \end{aligned}$$

Example: parking a car

$$\begin{aligned} \frac{d}{dt}\lambda + \frac{1}{y^2} &= 0 \\ \frac{\lambda}{m} &= 0 \end{aligned}$$

From the second equation $\lambda = 0$, and so we see that
So the only viable solutions are $y = \pm\infty$

Example: parking a car

E-L solutions:

- solutions are $y = \pm\infty$
- this requires $u = \pm\infty$ at some points in time
- but in reality we can't exert infinite force
  - i.e., force is bounded

$$|u| \leq u_{\max}$$

- need to consider optimizing functionals with inequality constraints.
  - similar (in some respects) to min/max functions with inequality constraints
  - min/max is in the interior, or on the boundary

Inequality constraints

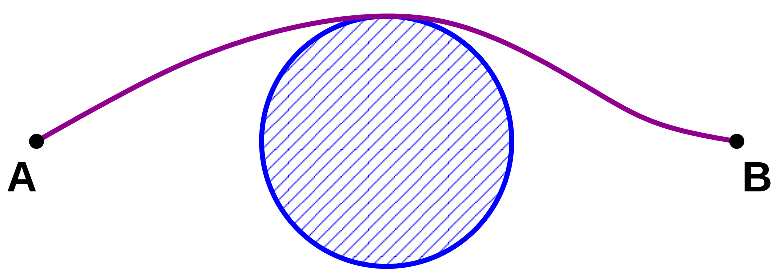
We have considered problems with

- integral constraints (Dido's problem)
- holonomic constraints (geodesics formulation)
- non-holonomic constraints (problems with higher derivatives)

But we have not considered inequality constraints

A problem

What is the shortest path, between A and B, avoiding an obstacle



E.G. what is the shortest path around a lake?

Formulation

Find extremals of

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

subject to $y(0) = y_0$ and $y(1) = y_1$ and

$$y(x) \geq g(x)$$

Enforce the constraint by taking

$$y(x) = g(x) + z(x)^2$$

In other words introduce a "slack function" $z(x)$, and note that

$$y(x) - g(x) = z(x)^2 \geq 0$$

Formulation

We have slack function $z(x)$, and constraint $y(x) \geq g(x)$ and

$$\begin{aligned} y &= z^2 + g \\ y' &= 2zz' + g' \end{aligned}$$

Substitute these into the functional and we can change the original functional $F\{y\}$ for a new one in terms of $F\{z\}$

$$\begin{aligned} F\{y\} &= \int_{x_0}^{x_1} f(x,y,y')\,dx \\ F\{z\} &= \int_{x_0}^{x_1} f(x,z^2+g,2zz'+g')\,dx \end{aligned}$$

Euler-Lagrange equations

Given we look for the extremals of

$$F\{z\} = \int_{x_0}^{x_1} f(x,z^2+g,2zz'+g')\,dx$$

the Euler-Lagrange equations are

$$\begin{aligned} \frac{d}{dx}\frac{\partial f}{\partial z'} - \frac{\partial f}{\partial z} &= 0 \\ \frac{d}{dx}\left[2z\frac{\partial f}{\partial y'}\right] - 2z\frac{\partial f}{\partial y} - 2z'\frac{\partial f}{\partial y'} &= 0 \\ 2z\frac{d}{dx}\frac{\partial f}{\partial y'} + 2z'\frac{\partial f}{\partial y'} - 2z\frac{\partial f}{\partial y} - 2z'\frac{\partial f}{\partial y'} &= 0 \\ z\left[\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}\right] &= 0 \end{aligned}$$

Euler-Lagrange equations

The Euler-Lagrange equations give

$$z\left[\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}\right] = 0$$

for which there are two solutions

- **Euler areas:** The E-L equations are satisfied
- **Boundary areas:** $z(x) = 0$, so $y(x) = g(x)$ and the curve lies on the boundary.

Analogy: a global minima of function on an interval can happen at stationary point, or at the edges.
But we can mix the two along the curve y.

Example

Find the shortest path around a circular lake (radius a, centered at the origin), between the points $(b,0)$ and $(-b,0)$ (for $b > a$).

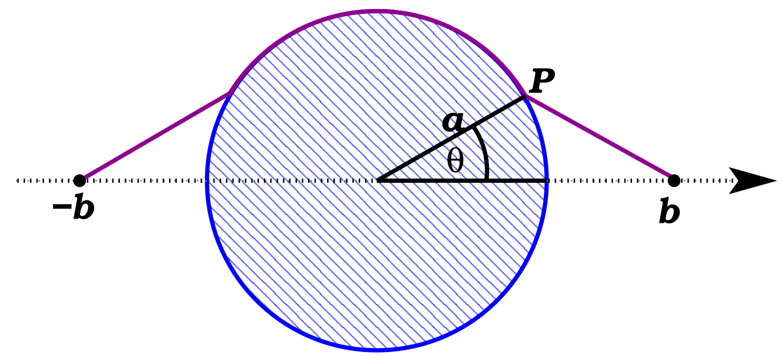
The conditions are

- **Euler areas:** The E-L equations are satisfied, so the curve is a straight line.
- **Boundary areas:** $z(x) = 0$, so $y(x) = g(x)$ and the curve lies on the boundary of the circle.

We can mix the two along the curve y.

Example

Given the conditions, the solution must look like



i.e. straight lines joining the end-points to a circular arc, where P, the point of intersection of the right-hand straight line, and the circle is at $(a\cos\theta, a\sin\theta)$.

Example

The total distance of such a line is

$$\begin{aligned} d(\theta) &= 2\sqrt{(b - a\cos\theta)^2 + a\sin^2\theta} + a(\pi - 2\theta) \\ &= 2\sqrt{b^2 - 2ab\cos\theta + a^2} + a(\pi - 2\theta) \end{aligned}$$

We find the minimum of $d(\theta)$, by differentiating WRT θ, to get

$$\begin{aligned} d' &= \frac{2ab\sin\theta}{\sqrt{b^2 - 2ab\cos\theta + a^2}} - 2a \\ &= 0 \end{aligned}$$

So

$$2ab\sin\theta = 2a\sqrt{b^2 - 2ab\cos\theta + a^2}$$

Example

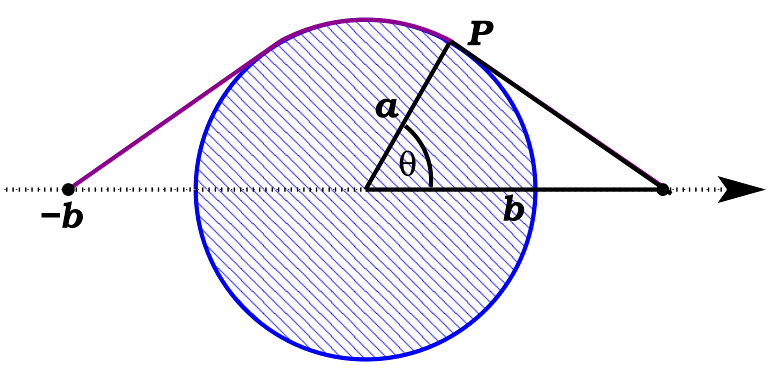
Dividing both sides by $2a$ we get the condition

$$\begin{aligned} b\sin\theta &= \sqrt{b^2 - 2ab\cos\theta + a^2} \\ b^2\sin^2\theta &= b^2 - 2ab\cos\theta + a^2 \\ b^2 - b^2\cos^2\theta &= b^2 - 2ab\cos\theta + a^2 \\ 0 &= b^2\cos^2\theta - 2ab\cos\theta + a^2 \\ 0 &= (b\cos\theta - a)^2 \end{aligned}$$

So the result is

$$\cos\theta = a/b$$

Example: solution



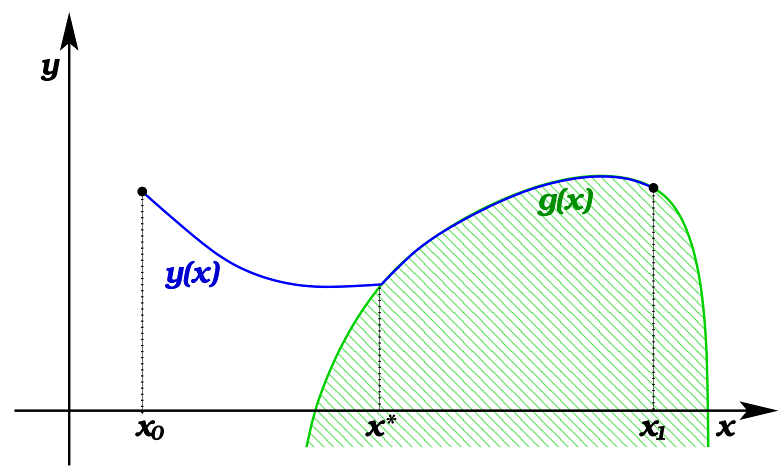
Think of what we would get if we stretch an elastic band between the two points.

General result

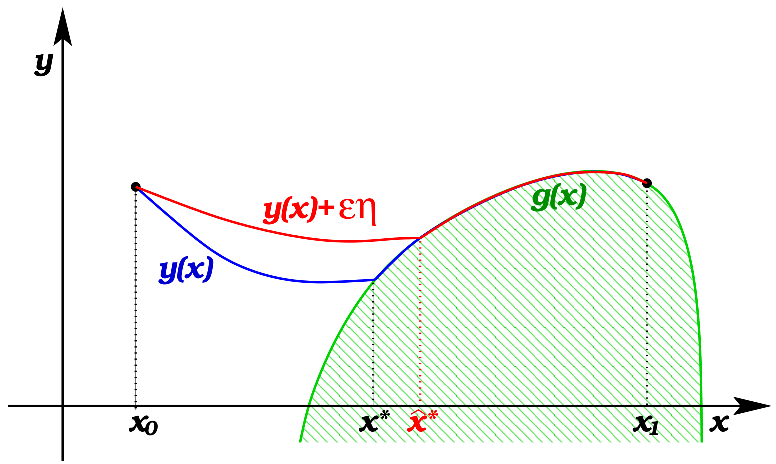
If $f_{y'}$ depends on y', then at the point where the extremal transfers from the Euler-Lagrange curve to the domain boundary the tangent varies continuously.

The problem is similar to that of the broken extremal. Here, the break is imposed by the change from one solution to the other (Euler-Lagrange to domain boundary). However, the condition can be seen in the same way, e.g. by perturbing the possible corner, along the boundary.

General result: proof



General result: proof



General result: proof

Similarly to the Weierstrass-Erdman Corner Conditions proof, we break the integral into two parts:

$$F\{y\} = F_1\{y\} + F_2\{y\} = \int_{x_0}^{x^*} f(x,y,y')\,dx + \int_{x^*}^{x_1} f(x,y,y')\,dx$$

but we will assume the shape of the curve on the RHS of x^* fits the boundary, e.g. $y(x) = g(x)$, and the LHS follows the E-L equations

$$F\{y\} = F_1\{y\} + F_2\{y\} = \int_{x_0}^{x^*} f(x,y,y')\,dx + \int_{x^*}^{x_1} f(x,g,g')\,dx$$

General result: proof

The first component of the first variation is, as with transversals, and corners conditions,

$$\delta F_1(\eta,y) = \lim_{\varepsilon\to\infty} \frac{1}{\varepsilon}\left[\int_{x_0}^{\hat{x}^*} f(x,\hat{y}_1,\hat{y}_1')\,dx - \int_{x_0}^{x^*} f(x,y_1,y_1')\,dx\right]$$

And as with corners, we get an integral term which results in the E-L equation, plus the additional constraint

$$\left[p\delta y - H\delta x\right]_{x^{*-}} - \left[p\delta y - H\delta x\right]_{x^{*+}} = 0$$

where

$$H = y'\frac{\partial f}{\partial y'} - f \quad \text{and} \quad p = \frac{\partial f}{\partial y'}$$

General result: proof

As with other potential corners, we get a corner condition

$$\left[p\delta y - H\delta x\right]_{x^{*-}} - \left[p\delta y - H\delta x\right]_{x^{*+}} = 0$$

but note that curve $y(x^*) = g(x^*)$, which constrains the end-point, so we cannot consider arbitrary variations $(\delta x, \delta y)$. In fact, we can only consider variations where

$$\delta y = g'\delta x$$

Assuming that dg/dx is in fact defined the above is

$$\left[pg'\delta x - H\delta x\right]_{x^{*-}} - \left[pg'\delta x - H\delta x\right]_{x^{*+}} = 0$$

General result: proof

The condition

$$\left[pg'\delta x - H\delta x\right]_{x^*-} - \left[pg'\delta x - H\delta x\right]_{x^*+} = 0$$

which can be simplified to

$$\left[pg' - H\right]_{x^*-} - \left[pg' - H\right]_{x^*+} = 0$$

Substituting H and p, and $y' = g'$ on the RHS of x^* we get

$$\left[g'\frac{\partial f}{\partial y'} - y'\frac{\partial f}{\partial y'} + f\right]_{x^*-} - \left[g'\frac{\partial f}{\partial y'} - g'\frac{\partial f}{\partial y'} + f\right]_{x^*+} = 0$$

General result: proof

Simplifying we get

$$\left[(g' - y')\frac{\partial f}{\partial y'} - f\right]_{x^*-} + [f]_{x^*+} = 0$$

or

$$\left[(g' - y')\frac{\partial f}{\partial y'}\right]_{x^*-} - [f]_{x^*-} + [f]_{x^*+} = 0$$

- Consider the term $-\{[f]_{x^*-} - [f]_{x^*+}\}$
- Note that at the "join" $y(x^*) = g(x^*)$, so if the two limits of f differ it is because of a difference in y' on either side of the join
- Treat f as a function of just y', i.e, $f(x,y,y') = q_{x,y}(y')$

General result: proof

Taking $q_{x,y}(y') = f(x,y,y')$ where

- on the left side of x^*, we have y' determined by E-L equations
- on the right side of x^* we have $y' = g'$

So

$$\begin{aligned}[f]_{x^*-} - [f]_{x^*+} &= \lim_{x\to x^*-} f(x,y,y') - \lim_{x\to x^*+} f(x,g,g') \\ &= q_{x^*,y^*}[y'(x^*)] - q_{x^*,y^*}[g'(x^*)]\end{aligned}$$

Given its all the same, I won't keep writing the subscripts of q, and will just use

$$q(z) = q_{x^*,y^*}(z)$$

General result: proof

The Mean Value Theorem states: if a function $q(z)$ is continuous on the closed interval $[a,b]$ and differentiable on the open interval (a,b), then there exists a point c in (a,b) such that

$$q(b) - q(a) = (b-a)q'(c)$$

So we get

$$\begin{aligned}[f]_{x^*-} - [f]_{x^*+} &= q(y'(x^*)) - q(g'(x^*)) \\ &= [y'(x^*) - g'(x^*)]\, q'(c)\end{aligned}$$

for some c between $g'(x^*)$ and $y'(x^*)$

General result: proof

Taking $q(y') = f(x,y,y')$ we get

$$\frac{d}{dz}q(z) = \frac{\partial f}{\partial y'}(x,y,y')\Big|_{y'=z}$$

So

$$q'(c) = \frac{\partial f}{\partial y'}(x^*,y^*,c)$$

and hence

$$\begin{aligned}[f]_{x^*-} - [f]_{x^*+} &= q(y'(x^*)) - q(g'(x^*)) \\ &= [y'(x^*) - g'(x^*)]\, q'(c) \\ &= [y'(x^*) - g'(x^*)]\, \frac{\partial f}{\partial y'}(x^*,y^*,c)\end{aligned}$$

General result: proof

So the condition from before can be rewritten as follows:

$$\begin{aligned}\left[(g'-y')\frac{\partial f}{\partial y'} - f\right]_{x^*-} + [f]_{x^*+} &= 0 \\ \left[(g'-y')\left(\frac{\partial f}{\partial y'}(x,y,y') - \frac{\partial f}{\partial y'}(x,y,c)\right)\right]_{x^*} &= 0\end{aligned}$$

for some c between $g'(x^*)$ and $y'(x^*)$

General result: proof

$$(g'(x^*) - y'(x^*))\left(\frac{\partial f}{\partial y'}(x^*,y(x^*),y'(x^*)) - \frac{\partial f}{\partial y'}(x^*,y(x^*),c)\right) = 0$$

So there are two possibilities:

- $g'(x^*) = y'(x^*)$, which means that y meets the boundary at a tangent to the boundary.
- $\frac{\partial f}{\partial y'}(x,y,y') - \frac{\partial f}{\partial y'}(x,y,c) = 0$. This latter condition holds when $\frac{\partial f}{\partial y'}$ is constant with respect to y', i.e.,

$$\frac{\partial^2 f}{\partial y'^2} = 0$$

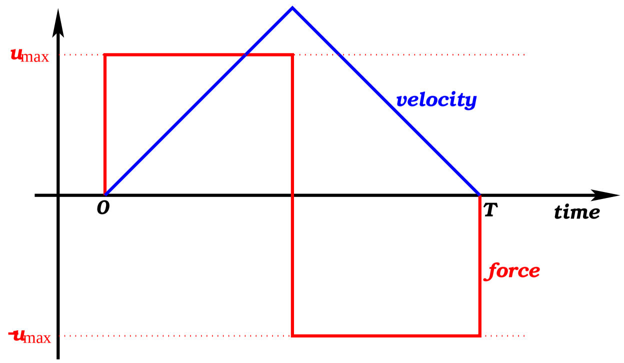
In the lake example, $\frac{\partial^2 f}{\partial y'^2} \neq 0$

Example: parking a car

- Revisit the problem of parking a car.
- If we think about the problem, it makes no sense unless there is maximum force $u_{\max}$.
  - otherwise we move from A to B arbitrarily fast.
- There are no valid E-L equation solutions.
- We must end-up in the boundary domain, e.g. $u = \pm u_{\max}$
  - obvious solution is to accelerate as fast as possible until we get half-way, and then to decelerate as fast as possible.
  - $\dfrac{\partial f}{\partial \dot{u}} = 0$, so we don't have to stress about continuity (u is not continuous either)

Example: parking a car

- Our solution is in the boundary domain, *e.g.* $u = \pm u_{\max}$



- called a **bang-bang controller**

Bang-bang controllers

Bang-bang controllers appear in a number of other contexts, and we will consider them in more generality later.

Variational Methods & Optimal Control

lecture 22

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 22

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

More Optimal Control Examples

First we'll cover a bit more terminology, and then some examples primarily focussed on planned growth strategies in economics.

Formulation of control problems

We break a control problem into two parts

- **The system state:** $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^t$
  The system state describes the system (e.g. position and velocity of the car in car parking example)
- **The control:** $\mathbf{u}(t) = (u_1(t), \ldots, u_m(t))^t$
  We apply the control to the system (e.g. force applied to the car).

The evolution of the system is governed by the set of DEs

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$

In a control problem we want to get the system to a particular state $\mathbf{x}(t_1)$ at time t_1, given initial state $\mathbf{x}(t_0)$.

Optimal control problems

In an **optimal** control problem we have still have the system equations $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$ and we might wish to get to state $\mathbf{x}(t_1)$ given initial state $\mathbf{x}(t_0)$, but now we wish to do so while minimizing a functional

$$F\{\mathbf{x}, \mathbf{u}\} = \int_{t_0}^{t_1} f(t, \mathbf{x}, \mathbf{u})\, dt$$

That is, we wish to choose a function $\mathbf{u}(t)$ which minimizes the functional $F\{\mathbf{x}, \mathbf{u}\}$, while satisfying the end-point conditions $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{x}(t_1) = \mathbf{x}_1$, and the non-holonomic constraints $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$.

Optimal control problems

Optimization functional

$$F\{\mathbf{x},\mathbf{u}\} = \int_{t_0}^{t_1} f(t,\mathbf{x},\mathbf{u})\,dt$$

Note that

- $f(t,\mathbf{x},\mathbf{u})$ has no dependence on $\dot{\mathbf{u}}$: this is typically because costs depend on the control, not how we change the control, but there might be counter-examples
- $f(t,\mathbf{x},\mathbf{u})$ has no dependence on $\dot{\mathbf{x}}$: this is common in control problems, but not universal (we have seen at least one counter example).

Terminal costs

Sometimes in optimal control we don't fix the end-point $\mathbf{x}(t_1)$, but rather we assign a cost $\phi(t_1,\mathbf{x}(t_1))$ to particular end-points.

So now we wish to choose a control $\mathbf{u}(t)$ which minimizes the functional

$$F\{\mathbf{x},\mathbf{u}\} = \phi(t_1,\mathbf{x}(t_1)) + \int_{t_0}^{t_1} f(t,\mathbf{x},\mathbf{u})\,dt$$

while satisfying the single end-point condition $\mathbf{x}(t_0) = \mathbf{x}_0$, and the non-holonomic constraint $\dot{\mathbf{x}}(t) = \mathbf{g}(t,\mathbf{x},\mathbf{u})$.

- $\phi(t_1,\mathbf{x}(t_1))$ is called the **terminal cost**.

System Terminology

- **linear:** the state equations are a set of linear DEs.
- **autonomous:** time doesn't appear explicitly in the state equations (e.g. in $g(\mathbf{x},\mathbf{u})$, or $f(\mathbf{x},\mathbf{u})$).
  - also called time-invariant
- **terminal cost:** the term $\phi(t_1,\mathbf{x}(t_1))$ is called the terminal cost.
- **controllable:** a solution to the control problem exists.
- **stable:** a stable equilibrium solution to the system DEs exists.
  - often we are interested in problems that are unstable, or we wouldn't really need a control

Control Terminology

- control (driver or automatic)
  - **planned** (open loop)
  - **feedback** (closed loop) control depends on current state
- type of control
  - movement from A to B
  - continuous operations (maintain equilibrium)
- type of cost function
  - minimum time
  - minimum fuel
  - quadratic costs
- admissible controls
  - unbounded/bounded/bang-bang

Cost function

- **minimum time:** choose the fastest possible control

$$F\{x,u\} = \int_{t_0}^{t_1} dt$$

- **minimum fuel:** fuel is expended by the controller, and we wish to minimize this

$$F\{x,u\} = \int_{t_0}^{t_1} |u(t)|\,dt$$

- **quadratic costs:**

$$F\{x,u\} = \int_{t_0}^{t_1} x^2(t) + \alpha u^2(t)\,dt$$

Boundary conditions

- End time t_1: can be fixed or free
- End position $\mathbf{x}(t_1)$: can be fixed or free

In the cases with free boundary conditions, we introduce natural, or transversal boundary conditions.

Example: dynamic production

- A producer in purely competitive market
  - A large numbers of independent producers
  - Standardized product, e.g. potatoes
  - Firms are "price takers", i.e. they have no significant control over product price
  - Free entry and exit
  - Free flow of information
- wants to find optimal production path $x(t)$, $0 \leq t \leq T$.
- production target $x(T) = x_T$
- profit at time t is $\pi(x, \dot{x}, t)$
- maximize profit functional $F\{x\} = \int_0^T \pi(x, \dot{x}, t)\,dt$

Example: dynamic production

Profit calculation

- quadratic production costs $C_1 = a_1 x^2 + b_1 x + c_1$
  - labor
  - raw materials
- production increase costs $C_2 = a_2 \dot{x}^2 + b_2 \dot{x} + c_2$
  - new buildings
  - recruiting and training costs
- revenue $r = px$ where p is the constant price per unit
  - $p = const$ due to purely competitive market
- profit at time t is

$$\pi(x, \dot{x}, t) = px - C_1(x) - C_2(\dot{x})$$

Example: dynamic production

Problem formulation: maximize total profit

$$F\{x\} = \int_0^T px - C_1(x) - C_2(\dot{x})\, dt$$

subject to $x(0) = 0$ and $x(T) = x_T$.

- notice that the control, and state are the same
- autonomous problem
- the control is planned, and has quadratic costs
- admissible controls are unbounded

Example: dynamic production

Euler-Lagrange equations

$$\begin{aligned}
\frac{d}{dt}\frac{\partial \pi}{\partial \dot{x}} - \frac{\partial \pi}{\partial x} &= 0 \\
-\frac{d}{dt}\frac{\partial C_2}{\partial \dot{x}} - p + \frac{\partial C_1}{\partial x} &= 0 \\
-\frac{d}{dt}\left[2a_2\dot{x} + b_2\right] - p + 2a_1 x + b_1 &= 0 \\
-2a_2\ddot{x} - p + 2a_1 x + b_1 &= 0 \\
\ddot{x} - \frac{a_1}{a_2}x &= \frac{-p + b_1}{2a_2}
\end{aligned}$$

for $a_2 \neq 0$

Example: dynamic production

Solution (for $a_1, a_2 \neq 0$)

$$x(t) = Ae^{\sqrt{\frac{a_1}{a_2}}t} + Be^{-\sqrt{\frac{a_1}{a_2}}t} + \frac{b_1 - p}{2a_2}$$

where A and B are determined by the fixed end points $x(0) = x_0$ and $x(T) = X_T$.

This gives the optimal production schedule.

- no dependence on c_1 or c_2 (these are constant costs and so shouldn't effect production strategy)
- no dependence on b_2 because this is a linear cost in increasing production, and so occurs regardless of how we increase over time (to get to the final production target $x(T) = X_T$).

Example: dynamic production

What happens if we make the end point $x(T)$ free, i.e. we don't have a production target at time T?

Then we get a natural boundary condition

$$\left.\frac{\partial \pi}{\partial \dot{x}}\right|_{t=T} = \left.\frac{\partial C_2}{\partial \dot{x}}\right|_{t=T} = 2a_2\dot{x} + b_2\big|_{t=T} = 0$$

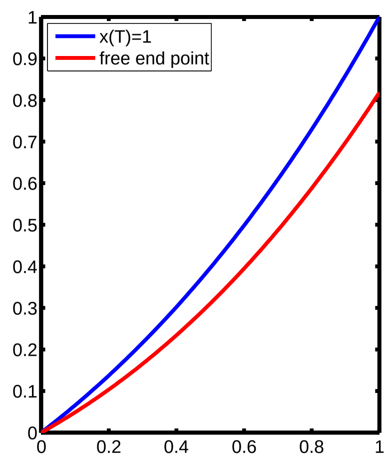
So, rearranging, we get

$$\dot{x}(T) = -\frac{b_2}{2a_2}$$

- constants A and B are determined by end-point conditions $x(0) = x_0$ and $\dot{x}(T) = -\frac{b_2}{2a_2}$

Example: dynamic production

- production costs
  $C_1 = x^2 + 5x$
- production increase costs
  $C_2 = 2\dot{x}^2 + 5\dot{x}$
- $p = 10$
- $T = 1$
- $x_0 = 0$, $x_T = 1$



Example: optimal economic growth

How much should be consumed, and how much invested for future consumption?

- optimal theory of saving (Ramsey, 1928)
- Total capital at time t is $K(t)$
- Total population (labor force) $L(t)$, which grows at exogenous rate n, e.g. $\dot{L} = nL$
- Homogeneous quantity called GDP denoted $Y(t)$
- GDP can either be consumed $C(t)$ or invested to get $\dot{K}(t)$, or used to replace depreciated capital $\mu K(t)$.

$$Y(t) = C(t) + \dot{K}(t) + \mu K(t)$$

Example: optimal economic growth

- GDP $Y(t)$ is a function of labor $L(t)$, and capital $K(t)$
- The production function $Y(t) = f_2(K, L)$ is homogeneous of degree one, e.g.

$$Y(t) = L(t)\, f_2(K/L, 1) = L(t)\, f(K/L)$$

- Hence we normalize all quantities by population L

$$\begin{aligned} y &= Y/L \quad \text{GDP per capita} \\ k &= K/L \quad \text{Capital investment per capita} \\ c &= C/L \quad \text{Consumption per capita} \end{aligned}$$

and write $y(t) = f(k)$ where f is assumed to be a strictly concave, monotonically increasing function, with slope decreasing from ∞ at 0, to 0 at ∞.

Example: optimal economic growth

Consider the rate of per capita investment

$$\dot{k} = \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{\dot{K}}{L} - \left(\frac{K\dot{L}}{L^2}\right) = \frac{\dot{K}}{L} - n\frac{K}{L} = \frac{\dot{K}}{L} - nk$$

using the fact that $\dot{L}/L = n$. Now we assumed that GDP could be expended in one of three ways, leading to

$$Y = C + \dot{K} + \mu K$$

which we also divide by L to obtain

$$y = c + \dot{k} + (\mu + n)k$$

which, when we substitute $y = f(k)$ gives

$$c(t) = f(k) - \dot{k} - (\mu + n)k(t)$$

Example: optimal economic growth

- We want to maximize the total **utility**
- Utility of per capita consumption is $U(c)$. This would also be a strictly concave, monotonically increasing function (according to the law of diminishing marginal utility, i.e. $U''(c) < 0 < U'(c)$).
- Utility in the future is discounted by rate r, e.g. is given by $U(c)e^{-rt}$
- Our control is how much we consume (and hence what is left to invest $\dot{k}$), and the state is the per capita investment $k(t)$.

Example: optimal economic growth

We want to maximize the total **utility** over time, e.g.

$$F\{c\} = \int_0^T U(c)e^{-rt}\,dt$$

subject to

$$c(t) = f(k) - \dot{k} - (\mu+n)k(t)$$

with $k(0) = k_0$, and $k(T) = k_T$.

Substitute c into the functional and we get

$$F\{k\} = \int_0^T U\left(f(k) - \dot{k} - (\mu+n)k(t)\right)e^{-rt}\,dt$$

Example: optimal economic growth

The E-L equations are

$$\frac{d}{dt}\frac{\partial \psi}{\partial \dot{k}} - \frac{\partial \psi}{\partial k} = 0$$

where $\psi(k,\dot{k}) = U\left(f(k) - \dot{k} - (\mu+n)k(t)\right)e^{-rt}$, so

$$\begin{aligned}
-\frac{d}{dt}e^{-rt}\frac{dU}{dc} - e^{-rt}\frac{dU}{dc}\left[\frac{df}{dk} - (\mu+n)\right] &= 0 \\
-e^{-rt}\frac{d}{dt}\frac{dU}{dc} + e^{-rt}\frac{dU}{dc}\left[r - \frac{df}{dk} + (\mu+n)\right] &= 0 \\
-e^{-rt}\frac{d^2U}{dc^2}\frac{dc}{dt} + e^{-rt}\frac{dU}{dc}\left[r - \frac{df}{dk} + (\mu+n)\right] &= 0
\end{aligned}$$

Example: optimal economic growth

We know $e^{-rt} \neq 0$, so we divide it out, and rearrange to get

$$\frac{dc}{dt} = \left[r + \mu + n - \frac{df}{dk}\right]\frac{U'}{U''}$$

which together with

$$\dot{k} = f(k) - c(t) - (\mu+n)k(t)$$

determines the optimal solution of the system. Remember we are given

- U the utility
- f the per capita production as a function of capital

Example: optimal economic growth

Example, $U(c) = \log(c)$, then $U' = 1/c$ and $U'' = -1/c^2$, so

$$\frac{dc}{dt} = \alpha c \text{ where } \alpha = -\left[r + \mu + n - \frac{df}{dk}\right]$$

so

$$c(t) = Ae^{\alpha t}$$

To solve for k, take linear production model, e.g. $y = \beta k$, and then

$$\dot{k} = \gamma k(t) - c(t) \text{ where } \gamma = (\beta - \mu - n)$$

So

$$k(t) = Be^{\gamma t} + \frac{c(t)}{\gamma - \alpha} = Be^{\gamma t} + \frac{c(t)}{r}$$

with A and B determined by $k(0) = k_0$, and $k(T) = k_T$.

Example: optimal economic growth

To maintain constant consumption $c(t)$ we require $\dot{c} = 0$, and so we must have

$$\frac{df}{dk} = r + \mu + n$$

To maintain constant investment, we require

$$\dot{k} = f(k) - c(t) - (\mu + n)k(t) = 0$$

which together determine a solution (c^*, k^*), where the system is in equilibrium.

For the example $y = \beta k$

$$k = \frac{r + \mu + n}{\beta} \quad \text{and} \quad c = (\beta - \mu - n)k$$

Variational Methods & Optimal Control

lecture 23

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 23

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

More Optimal Control Examples

An aerospace example: a rocket launch profile.

Example: launching a rocket

Launch a rocket (with one stage) to deliver its payload into Low-Earth Orbit (LEO) at some height h above the Earth's surface.
Assumptions:

- ignore drag, and curvature and rotation of Earth
- LEO so assume gravitational force at ground and orbit are approximately the same
- thrust will generate acceleration a, which is predefined by rocket parameters
- we thrust for some time T, then follow a ballistic trajectory until (hopefully) we reach height h, at zero vertical velocity, and with horizontal velocity matching the required orbital injection speed.

Example: launching a rocket

Notation:

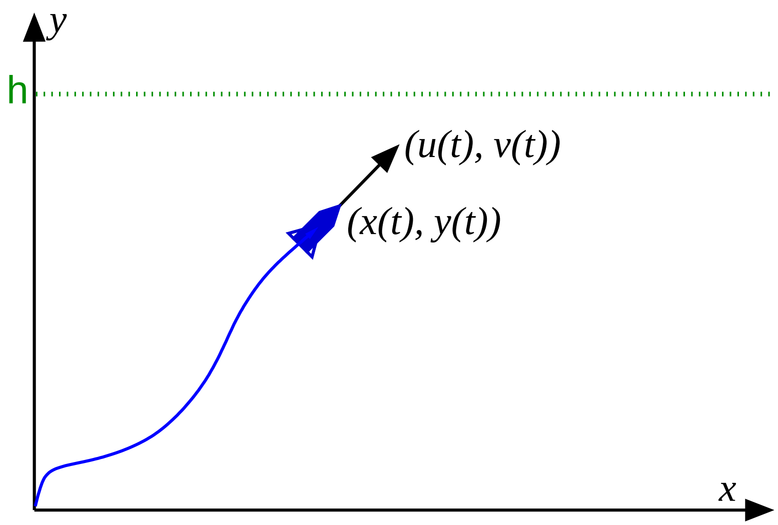
$$\begin{aligned} x &= \text{horizontal position} \\ y &= \text{vertical position} \\ u &= \text{horizontal velocity} \\ v &= \text{vertical velocity} \end{aligned}$$

Initial conditions $x(0) = y(0) = u(0) = v(0) = 0$. Thrust stops at time T, and then at some later time S, we reach the peak of the trajectory where

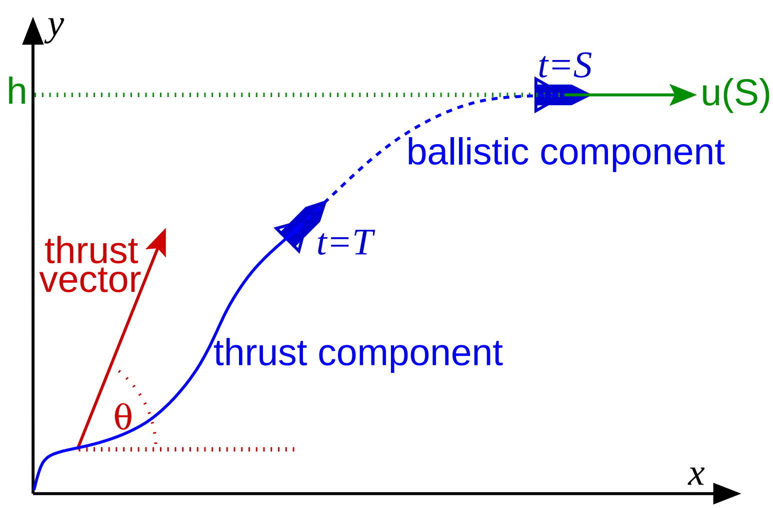
$$\begin{aligned} y(S) &= h \\ u(S) &= u_o, \text{ orbital velocity} \\ v(S) &= 0 \end{aligned}$$

We don't actually care about the final position $x(S)$

Example: launching a rocket



Example: launching a rocket



Example: launching a rocket

- Control: thrust profile is pre-determined. Only thing we can control (in this problem) is the **angle** of thrust.
  - Measure the angle of thrust $\theta(t)$ relative to horizontal.
- want to minimize fuel
  - but this is equivalent to minimizing time, e.g.,

$$F = \int_0^t a\,dt = a\int_0^T 1\,dt$$

- need to get to height h
- need to get to horizontal velocity u_o to enter orbit

Example: constraint equations

Thrust component

$$\begin{aligned}\dot{x} &= u\\ \dot{y} &= v\\ \dot{u} &= a\cos\theta\\ \dot{v} &= a\sin\theta - g\end{aligned}$$

Initial point:
$x(0) = y(0) = u(0) = v(0) = 0.$

Final point: $free$

Ballistic component

$$\begin{aligned}\dot{x} &= u\\ \dot{y} &= v\\ \dot{u} &= 0\\ \dot{v} &= -g\end{aligned}$$

Initial point: fixed
$x(T), y(T), u(T), v(T)$

Final point:
$x(S)$ $free$,
$y(S) = h$, $v(S) = 0$, $u(S) = u_o$

Example: ballistic component

For $t \in [T,S]$ the trajectory is ballistic, so

$$\begin{aligned} \dot{x} &= u \\ \dot{y} &= v \\ \dot{u} &= 0 \\ \dot{v} &= -g \end{aligned}$$

we can calculate the top of the resulting parabola as

$$\begin{aligned} u(S) &= u(T) \\ v(S) &= 0 \\ y(S) &= y(T) + v(T)^2/2g \end{aligned}$$

and we don't care about $x(T)$ or $x(S)$.

Example: co-ordinate transform

So we can change variables: make the final point $t = T$, and take variables u, v as before, and

$$z = y + v^2/2g.$$

We can differentiate this and combine with previous results to get the new **system DEs**

$$\begin{aligned} \dot{u} &= a\cos\theta \\ \dot{v} &= a\sin\theta - g \\ \dot{z} &= \dot{y} + v\dot{v}/g \\ &= v(1+\dot{v}/g) \\ &= \frac{av}{g}\sin\theta \end{aligned}$$

Example: optimization functional

Time minimization problem

$$T = \int_0^T 1\, dt$$

Including Lagrange multipliers for the 3 system constraints we aim to minimize

$$J\{\theta\} = \int_0^T 1 + \lambda_u(\dot{u} - a\cos\theta) + \lambda_v(\dot{v} - a\sin\theta + g) + \lambda_z\left(\dot{z} - \frac{av}{g}\sin\theta\right) dt$$

subject to

$$\begin{aligned} u(0) &= 0, & u(T) &= u_o \\ v(0) &= 0, & v(T) &= free \\ z(0) &= 0, & z(T) &= h \\ \theta(0) &= free, & \theta(T) &= free \end{aligned}$$

Example: Euler-Lagrange equations

E-L equations

$$\begin{aligned} u: \quad \frac{\partial h}{\partial u} - \frac{d}{dt}\frac{\partial h}{\partial \dot{u}} &= 0 \Rightarrow & \dot{\lambda}_u &= 0 \\ v: \quad \frac{\partial h}{\partial v} - \frac{d}{dt}\frac{\partial h}{\partial \dot{v}} &= 0 \Rightarrow & \dot{\lambda}_v &= -\lambda_z \tfrac{a}{g}\sin\theta \\ z: \quad \frac{\partial h}{\partial z} - \frac{d}{dt}\frac{\partial h}{\partial \dot{z}} &= 0 \Rightarrow & \dot{\lambda}_z &= 0 \\ \theta: \quad \frac{\partial h}{\partial \theta} - \frac{d}{dt}\frac{\partial h}{\partial \dot{\theta}} &= 0 \Rightarrow & & \end{aligned}$$

$$a\lambda_u\sin\theta - \lambda_v a\cos\theta - \lambda_z\frac{av}{g}\cos\theta = 0$$

(λ equations give back systems DEs)

Example: solving the E-L equations

Take the v equation, and noting that $\dot{v} = a\sin\theta - g$

$$\begin{aligned}\dot{\lambda}_v &= -\lambda_z \frac{a}{g}\sin\theta \\ &= -\frac{\lambda_z}{g}(\dot{v}+g) \\ \lambda_v &= -\frac{\lambda_z}{g}(v+gt+c) \\ &= -\frac{\lambda_z v}{g} - \lambda_z t + b\end{aligned}$$

Example: solving the E-L equations

Substitute

$$\lambda_v = -\frac{\lambda_z v}{g} - \lambda_z t + b$$

into the θ E-L equation (dropping the common factor a)

$$\lambda_u \sin\theta - \lambda_v \cos\theta - \lambda_z \frac{v}{g}\cos\theta = 0$$

and we get

$$\begin{aligned}\lambda_u \sin\theta + \left(\frac{\lambda_z v}{g} + \lambda_z t - b\right)\cos\theta - \lambda_z \frac{v}{g}\cos\theta &= 0 \\ \lambda_u \sin\theta + (\lambda_z t - b)\cos\theta &= 0 \\ \tan\theta &= -(\lambda_z t - b)/\lambda_u\end{aligned}$$

Example: solution

Remember that λ_u and λ_v and b are all constants, so the equation

$$\boxed{\tan\theta = -(\lambda_z t - b)/\lambda_u}$$

- angle of thrust now specified

$$\theta = \tan^{-1}\left(-(\lambda_z t - b)/\lambda_u\right)$$

- but we need to determine constants

Example: end-point conditions

Final end-points conditions

$$\begin{aligned}z(T) &= h \\ u(T) &= u_o, \text{ orbital velocity} \\ v(T) &= free \\ \theta(T) &= free \\ \lambda_u &= free \\ \lambda_v &= free \\ \lambda_z &= free\end{aligned}$$

Example: natural boundary conditions

The free-end point boundary condition for

$$F\{t,\mathbf{q},\dot{\mathbf{q}}\} = \int L(t,\mathbf{q},\dot{\mathbf{q}})\,dt$$

is

$$\sum_{k=1}^{n} p_k \delta q_k - H\delta t = 0 \text{ where } p_k = \frac{\partial L}{\partial \dot{q}_k} \text{ and } H = \sum_{k=1}^{n} \dot{q}_k p_k - L$$

In this problem

$$\frac{\partial L}{\partial \dot{\lambda}_k} = 0, \quad \frac{\partial L}{\partial \dot{\theta}} = 0, \quad \frac{\partial L}{\partial \dot{u}} = \lambda_u, \quad \frac{\partial L}{\partial \dot{v}} = \lambda_v, \quad \frac{\partial L}{\partial \dot{z}} = \lambda_z$$

Example: natural boundary conditions

Consider δq_k for each co-ordinate:

- for fixed co-ordinates $\delta q_k = 0$
- its free for θ, λ_u, λ_v, λ_z, but in each case the corresponding $p_k = 0$, so we can ignore these.
- only case where it matters is δv, which we can vary, and for which $p_v = \lambda_v$.

Also δt is free, so we get two end-point conditions at $t = T$.

$$\begin{aligned} H &= 0 \\ p_v = \lambda_v(T) &= 0 \end{aligned}$$

Example: natural boundary conditions

Given $\lambda_v(T) = 0$, and from previous work

$$\lambda_v = -\frac{\lambda_z v}{g} - \lambda_z t + b$$

we get

$$\begin{aligned} \lambda_z v(T)/g &= -\lambda_z T + b \\ &= \lambda_u \tan\theta(T) \\ v(T) &= \frac{\lambda_u g}{\lambda_z} \tan\theta(T) \end{aligned}$$

Example: natural boundary conditions

$$\frac{\partial L}{\partial \dot{\lambda}_k} = 0, \quad \frac{\partial L}{\partial \dot{\theta}} = 0, \quad \frac{\partial L}{\partial \dot{u}} = \lambda_u, \quad \frac{\partial L}{\partial \dot{v}} = \lambda_v, \quad \frac{\partial L}{\partial \dot{z}} = \lambda_z$$

So H is given by

$$H = \lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} - L$$

Substitute L, and the system DEs, and we get

$$H = \lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} - 1$$

The end-point condition at $t = T$ is therefore

$$\lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} = 1$$

Example: natural boundary conditions

Substitute

$$\begin{aligned} \lambda_v &= -\lambda_z v/g - \lambda_z t + b \\ &= -\lambda_z v/g + \lambda_u \tan\theta \\ \dot{u} &= a\cos\theta \\ \dot{v} &= a\sin\theta - g \\ \dot{z} &= \frac{av}{g}\sin\theta \end{aligned}$$

Into

$$\lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} = 1$$

and we get

Example: natural boundary conditions

We get

$$\begin{aligned} \lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} &= 1 \\ \lambda_u a\cos\theta + (-\lambda_z v/g + \lambda_u \tan\theta)(a\sin\theta - g) + \lambda_z \frac{av}{g}\sin\theta &= 1 \\ \lambda_u a\cos\theta + \lambda_z v + \lambda_u a \tan\theta\sin\theta - g\lambda_u\tan\theta &= 1 \\ \lambda_u a(\cos\theta + \tan\theta\sin\theta) + \lambda_z v - g\lambda_u \tan\theta &= 1 \\ \lambda_u a\left(\frac{\cos^2\theta + \sin^2\theta}{\cos\theta}\right) + \lambda_z v - g\lambda_u\tan\theta &= 1 \\ \lambda_u a\sec\theta + \lambda_z v - g\lambda_u \tan\theta &= 1 \end{aligned}$$

all evaluated at $t = T$. Combine with $g\lambda_u \tan\theta = \lambda_z v$ and

$$\lambda_z = \cos(\theta(T))/a$$

Example: natural boundary conditions

Another way to get the same result is to note

$$H = \lambda_u \dot{u} + \lambda_v \dot{v} + \lambda_z \dot{z} - L$$

and

$$L = 1 + \lambda_u(\dot{u} - a\cos\theta) + \lambda_v(\dot{v} - a\sin\theta + g) + \lambda_z\left(\dot{z} - \frac{av}{g}\sin\theta\right)$$

so

$$H = \lambda_u a\cos\theta + \lambda_v[a\sin\theta - g] + \frac{av\lambda_z}{g}\sin\theta - 1$$

which is what we got near the start of the previous slide before substituting $\lambda_v = -\lambda_z v/g + \lambda_u \tan\theta$.

Example: natural boundary conditions

At the starting point, all of the co-ordinates are fixed (except for θ, and the Lagrange multipliers), so the only free-end points condition at this point is

$$H = 0$$

as before. In fact, if $a = const$ the problem is not time-dependent, so H is conserved, i.e.

$$H(t) = 0$$

for the entire rocket flight. Note though, that for this system, H is not "energy" as this is not conserved (unless you include the chemical energy stored in the rocket).

Example: acceleration profile

The next steps depend on the acceleration profile $a(t)$, but lets take a simple case $a = const$.

First we can solve the DEs, with respect to θ using the chain rule

$$\frac{dX}{dt} = \frac{dX}{d\theta}\frac{d\theta}{dt} = -\cos^2\theta \frac{\lambda_z}{\lambda_u}\frac{dX}{d\theta}$$

e.g. from the system DE $\dot{u} = a\cos\theta$

$$\begin{aligned} \dot{u} &= -\cos^2\theta \frac{\lambda_z}{\lambda_u}\frac{du}{d\theta} \\ \frac{du}{d\theta} &= -\frac{\lambda_u}{\lambda_z \cos^2\theta}\dot{u} \\ &= -\frac{a\lambda_u}{\lambda_z\cos\theta} \end{aligned}$$

Example: acceleration profile

$$\frac{dX}{d\theta} = \frac{dX}{dt}\Big/\frac{d\theta}{dt} = \frac{dX}{dt}\Big/\left(-\cos^2\theta\frac{\lambda_z}{\lambda_u}\right)$$

The complete set of system DEs becomes

$$\begin{aligned} \frac{du}{d\theta} &= -\frac{a\lambda_u}{\lambda_z\cos\theta} \\ \frac{dv}{d\theta} &= -\frac{a\lambda_u}{\lambda_z}\frac{\sin\theta}{\cos^2\theta} + \frac{g\lambda_u}{\lambda_z\cos^2\theta} \\ \frac{dz}{d\theta} &= -\frac{a\lambda_u}{g\lambda_z}\frac{\sin\theta}{\cos^2\theta}v(\theta) \end{aligned}$$

These can just be integrated with respect to θ

Example: acceleration profile

The system DEs can be directly integrated (with respect to θ) including initial conditions $u(0) = v(0) = z(0) = 0$ to get

$$\begin{aligned} u(\theta) &= \frac{a\lambda_u}{\lambda_z}\log\left(\frac{\sec\theta_0 + \tan\theta_0}{\sec\theta + \tan\theta}\right) \\ v(\theta) &= \frac{a\lambda_u}{\lambda_z}(\sec\theta_0 - \sec\theta) - \frac{g\lambda_u}{\lambda_z}(\tan\theta_0 - \tan\theta) \\ z(\theta) &= \frac{a^2\lambda_u^2}{g\lambda_z^2}\sec\theta_1(\sec\theta_0 - \sec\theta) - \frac{a^2\lambda_u^2}{2g\lambda_z^2}(\tan^2\theta_0 - \tan^2\theta) \\ &\quad + \frac{a\lambda_u^2}{2\lambda_z^2}\left[\tan\theta_0\sec\theta_0 - \tan\theta\sec\theta + \log\left(\frac{\sec\theta_0 + \tan\theta_0}{\sec\theta + \tan\theta}\right)\right] \\ \theta &= \tan^{-1}\left(-(\lambda_z t - b)/\lambda_u\right) \end{aligned}$$

Example: calculating the constants

There are five constants to calculate:

- θ_0 the initial angle of thrust
- θ_1 the final angle of thrust
- λ_u
- λ_z
- b

and we also need to calculate T.

Solving for end-point conditions is non-trivial, but a method that works well (from Lawden) follows.

Example: calculating the constants

Take the equation for v at time T, and substitute $\lambda_z v(T) = g\lambda_u \tan\theta_1$ to get

$$\begin{aligned} v(\theta_1) &= \frac{a\lambda_u}{\lambda_z}(\sec\theta_0 - \sec\theta_1) - \frac{g\lambda_u}{\lambda_z}(\tan\theta_0 - \tan\theta_1) \\ \frac{g\lambda_u}{\lambda_z}\tan\theta_1 &= \frac{a\lambda_u}{\lambda_z}(\sec\theta_0 - \sec\theta_1) - \frac{g\lambda_u}{\lambda_z}(\tan\theta_0 - \tan\theta_1) \\ \sec\theta_1 &= \sec\theta_0 - \frac{g}{a}\tan\theta_0 \end{aligned}$$

which gives us a way to calculate θ_1 from θ_0. Once we know θ_1 we can calculate λ_u using $\lambda_u a = \cos\theta_1$, and b from $\tan\theta = (-(\lambda_z t - b)/\lambda_u)$ at $t = 0$. Then we can calculate λ_z from $u(\theta_1) = u_o$, the orbital injection velocity

Example: calculating the constants

So the only remaining question is how to calculate θ_0. We do so numerically, by

- take a range of θ_0
- calculate all of the above
- use this to calculate $z(T) = z_1$ as a function of θ_0
- look for the point where $z_1(\theta_0) = h$ the orbit height.

That gives us the θ_0, from which we can derive everything else. There are good numerical methods to search for such a solution, particularly if we start with a clear range over which to look.

Example: restricting choice of θ_0

Calculating the range of θ_0 to search

- The maximum (reasonable) value for θ_0 is $\pi/2$.
- The minimum value of θ_0 will be determined by the minimum possible value of θ_1, i.e., $\theta_1 = 0$

$$\begin{aligned} \sec\theta_1 &= \sec\theta_0 - \frac{g}{a}\tan\theta_0 \\ \sec 0 &= \sec\theta_0 - \frac{g}{a}\tan\theta_0 \\ 1 &= \sec\theta_0 - \frac{g}{a}\tan\theta_0 \\ 1 &= \frac{1+\tan^2\theta_0/2}{1-\tan^2\theta_0/2} - \frac{g}{a}\frac{2\tan\theta_0/2}{1-\tan^2\theta_0/2} \\ 1-\tan^2\theta_0/2 &= 1+\tan^2\theta_0/2 - \frac{2g}{a}\tan\theta_0/2 \end{aligned}$$

Example: restricting choice of θ_0

$$\begin{aligned} 1-\tan^2\theta_0/2 &= 1+\tan^2\theta_0/2 - \frac{2g}{a}\tan\theta_0/2 \\ 2\tan^2\theta_0/2 - \frac{2g}{a}\tan\theta_0/2 &= 0 \\ \tan\theta_0/2\left(\tan\theta_0/2 - \frac{g}{a}\right) &= 0 \end{aligned}$$

Now θ_0 can't be zero, so the last step implies that the minimum value of θ_0 is

$$\theta_0 = 2\tan^{-1}(g/a)$$

Note the existence of a minimum critical h below which we can't find a trajectory of this type.

Example: parameters

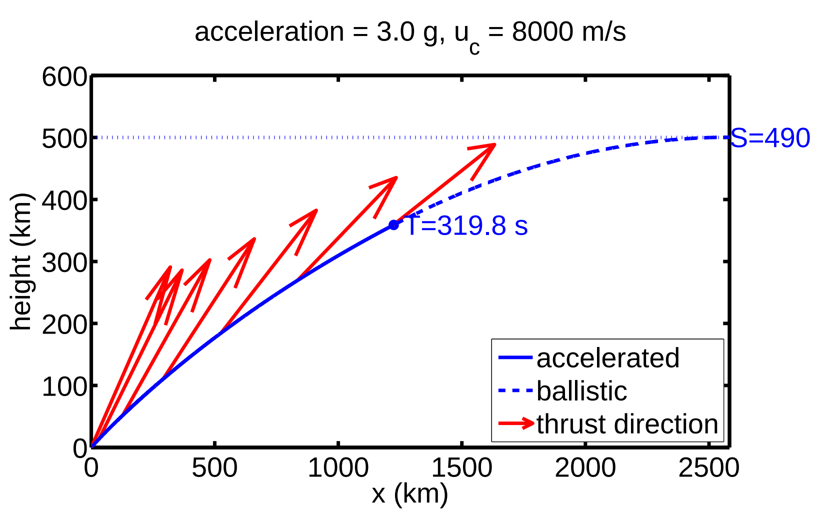
Parameters of previous example consistent with a LEO.

$$\begin{aligned} h &= 500 \text{ km} \\ u_o &= 8000 \text{ m/s} \\ g &= 9.8 \text{ m/s}^2 \\ a &= 3g \end{aligned}$$

Derived constants

$$\begin{aligned} \theta_0 &= 0.2349\pi & \theta_1 &= 0.0973\pi \\ \lambda_u &= 0.0324 & \lambda_z &= 6.0257e-05 \\ b &= -0.0295 \\ T &= 319.8 \text{ seconds} \\ S &= 489.6 \text{ seconds} \end{aligned}$$

Example: trajectory



Example: generalizations

More realistic assumptions

- non-zero drag (depends on velocity and height)
- thrust is constant, but rocket mass changes, so that acceleration isn't constant
- multiple stages
- centripetal forces

For more examples, and discussion see Lawden, "Optimal Trajectories for Space Navigation", Butterworths, 1963 (which is incidentally where the above example comes from).

Variational Methods & Optimal Control

lecture 24

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 24

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Hamilton's formulation

We've seen the Hamiltonian H earlier on, but haven't explored its full power. Firstly, using H can often result in a simpler approach than solving the E-L equations, e.g., where f has no dependence on x, or where there is more than one dependent variable. More importantly though, this formulation can lead to an understanding of how symmetries in the problem of interest lead to conservation laws. Finally, we will use the Hamiltonian in the Pontryagin Maximum Principle, which we will study soon.

Legendre transformation

- Contact transformation
  (as opposed to point transformation)
- transformation that depends on the derivatives of a variable
- simple one variable Legendre transform of $y : [x_0, x_1] \to I\!R$, by defining new variable p, by
$$p(x) = y'(x)$$
- provided $y''(x) \neq 0$ we can define x in terms of p, by introducing the Hamiltonian
$$H(p) = px - y(x)$$

Legendre transformation

Assume for convenience that y is convex, e.g. $y'' > 0$ for $x \in [x_0, x_1]$. Then

$$\begin{aligned}
\frac{dH}{dp} &= \frac{d}{dp}(xp) - \frac{dy}{dp} \\
&= p\frac{dx}{dp} + x - \frac{dy}{dp} \\
&= p\frac{dx}{dp} + x - \frac{dy}{dx}\frac{dx}{dp} \\
&= \left(p - \frac{dy}{dx}\right)\frac{dx}{dp} + x \\
&= x
\end{aligned}$$

and also note $px - H = y$, so from the pair (p, H) we can recover the original pair (x, y), by a Legendre transform.

Example Legendre transformation

Let $f(x) = x^4/4$, then

$$\begin{aligned} p &= \frac{df}{dx} &&= x^3 \\ H(p) &= px - \tfrac{1}{4}x^4 &&= \tfrac{3}{4}p^{4/3} \end{aligned}$$

Note that we can reverse with another Legendre transform

$$\begin{aligned} \frac{dH}{dp} &= p^{1/3} &&= x \\ px - H &= x^4 - \tfrac{3}{4}x^4 &&= f(x) \end{aligned}$$

Hamilton's formulation

Refer back to problems with more than one dependent variable, or where f has no dependence on x.

Define **generalized coordinates** $\mathbf{q} : [t_0, t_1] \to \mathbb{R}^n$.

- i.e. take a set of n functions $q_k(t)$, with two continuous derivatives with respect to t, and put them into a vector $\mathbf{q}(t)$
- dot notation:

  $$\dot{q}_k = \frac{dq_k}{dt}, \quad \ddot{q}_k = \frac{d^2 q_k}{dt^2} \quad \text{and} \quad \dot{\mathbf{q}} = \left(\frac{dq_1}{dt}, \frac{dq_2}{dt}, \ldots, \frac{dq_n}{dt} \right)$$

- Lagrangian $L(t, \mathbf{q}, \dot{\mathbf{q}})$

Hamilton's formulation

The extremals of the functional

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

satisfy the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0$$

for all k.

Hamilton's formulation

Legendre transform introduces the **conjugate** variables

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

Suppose these equations can be solved to write $\dot{q}_i$ as a function of (t, q_i, p_i), then the **Hamiltonian** is

$$H(t, q_1, \ldots, q_n, p_1, \ldots, p_n) = \sum_{i=1}^{n} p_i \dot{q}_i - L(t, \mathbf{q}, \dot{\mathbf{q}})$$

We've seen p_i and H before, for instance in transversality conditions.

- the p_i are called **generalized momenta**

Hamilton's formulation

$$H(t, q_1, \ldots, q_n, p_1, \ldots, p_n) = \sum_{i=1}^{n} p_i \dot{q}_i - L(t, \mathbf{q}, \dot{\mathbf{q}})$$

So

$$\begin{aligned} \frac{\partial H}{\partial p_i} &= \dot{q}_i \\ \frac{\partial H}{\partial q_i} &= -\frac{\partial L}{\partial q_i} \end{aligned}$$

Given the E-L equations, the second equation gives

$$\frac{\partial H}{\partial q_i} = -\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = -\frac{dp_i}{dt}$$

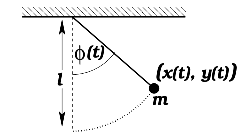
Canonical Euler-Lagrange equations

$$\begin{aligned} \frac{\partial H}{\partial p_i} &= \frac{dq_i}{dt} \\ \frac{\partial H}{\partial q_i} &= -\frac{dp_i}{dt} \end{aligned}$$

- called Hamilton's equations, or Canonical Euler-Lagrange equations
- The n E-L DEs converted into $2n$ first-order DEs
- derivatives are now uncoupled
  - therefore maybe easier to solve

Harmonic oscillator example

Simple pendulum



$$F\{\phi\} = \int_{t_0}^{t_1} \left(\frac{1}{2} m l^2 \dot{\phi}^2 - mgl(1 - \cos\phi) \right) dt$$

E-L equations

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} &= 0 \\ \frac{d}{dt} m l^2 \dot{\phi} - mgl \sin\phi &= 0 \\ m\ddot{\phi} - \frac{mg}{l} \sin\phi &= 0 \end{aligned}$$

standard pendulum equations, solve for small ϕ

Harmonic oscillator example

Generalized momentum (in this case angular momentum)

$$p = \frac{\partial L}{\partial \dot{\phi}} = m l^2 \dot{\phi} \quad \Rightarrow \quad \dot{\phi} = \frac{p}{ml^2}$$

Hamiltonian is

$$H(\phi, p) = p\dot{\phi} - L = \frac{p^2}{2ml^2} + mgl(1 - \cos\phi)$$

Hamilton's equations are

$$\begin{aligned} \frac{\partial H}{\partial p} &= \frac{d\phi}{dt} & \Rightarrow \quad \dot{\phi} &= \frac{p}{ml^2} \\ \frac{\partial H}{\partial \phi} &= -\frac{dp}{dt} & \Rightarrow \quad \dot{p} &= mgl \sin\phi \end{aligned}$$

Harmonic oscillator example

Hamilton's equations (2 first order DEs)

$$\begin{aligned}\dot{\phi} &= \frac{p}{ml^2}\\ \dot{p} &= mgl\sin\phi\end{aligned}$$

Differentiate the first equation and we get

$$\ddot{\phi} = \frac{\dot{p}}{ml^2}$$

Substitute the value of $\dot{p}$ from the second of Hamilton's equations and we get

$$\ddot{\phi} = \frac{g}{l}\sin\phi$$

the Euler-Lagrange equation.

Canonical Euler-Lagrange equations

We can get the same Canonical E-L equations from finding extremals of the functional of $2n$ variables

$$\tilde{F}\{q_1,\ldots,q_n,p_1,\ldots,p_n\} = \int_a^b \left[\sum_{i=1}^n p_i\dot{q}_i - H\right] dx$$

E.G.

$$\begin{aligned}\left(\frac{\partial}{\partial q_i} - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i}\right)\left[\sum_{i=1}^n p_i\dot{q}_i - H\right] &= 0\\ \left(\frac{\partial}{\partial p_i} - \frac{d}{dt}\frac{\partial}{\partial \dot{p}_i}\right)\left[\sum_{i=1}^n p_i\dot{q}_i - H\right] &= 0\end{aligned}$$

Hamilton's formulation

- F and $\tilde{F}$ are equivalent under the Legendre transformation
  - make q and p independent, whereas before it was a bit of a trick to pretend q and $\dot{q}$ were independent
- If L does not depend on t, then it should be clear from the Legendre transformation that H won't depend on t.
  - the system will be **conservative**
  - i.e. H is a conserved (constant) quantity

Hamilton-Jacobi equation

Find stationary points of

$$F\{y\} = \int_{x_0}^{x_1} f(x,y,y')\,dy$$

given particular fixed end points (x_0,y_0) and (x_1,y_1).

Now vary the second end-point. We can consider that the value of $F\{y\}$ along the extremal is now a function of (x_1,y_1), e.g.

$$F\{y\} = S(x_1,y_1)$$

Hamilton-Jacobi equation

Make a small variation in the end-point $(\delta x, \delta y)$. We know that the first variation will consist of an E-L component, plus a (free end-point) term like

$$p\delta y - H\delta x$$

but we are only considering extremal curves here, so the E-L component must be zero. Hence, we can write

$$\delta S = S(x+\delta x, y+\delta y) - S(x,y) = p\delta y - H\delta x$$

Keep x fixed, and vary only y, and we get

$$\frac{\delta S}{\delta y} = p$$

where the LHS is $\partial S/\partial y$ in the limit as $\delta y \to 0$

Hamilton-Jacobi equation

Similarly keeping y fixed and varying x we get an expression for $\partial S/\partial x$, which together with the previous expressions give

$$\begin{aligned} \frac{\partial S}{\partial y} &= p \\ \frac{\partial S}{\partial x} &= -H(x,y,p) \end{aligned}$$

Substitute the former equation into the latter, and we get

$$\frac{\partial S}{\partial x} + H\left(x, y, \frac{\partial S}{\partial y}\right) = 0$$

This is the **Hamilton-Jacobi** equation

Hamilton-Jacobi equation

Given a solution $S(x,y,\alpha)$ to the Hamilton-Jacobi equations (where α is a constant of integration), the extrema lie along the curves

$$\frac{\partial S}{\partial \alpha} = const$$

Proof: see

- Arthurs, Thm 8.1, p. 32
- van Brunt, Thm 8.4.1, p. 177

Simple example

Find extrema of

$$F\{y\} = \int_a^b y'^2\, dx$$

The conjugate variable and Hamiltonian are given by

$$\begin{aligned} p &= \frac{\partial f}{\partial y'} \\ &= 2y' \\ H(x,y,p) &= y'\frac{\partial f}{\partial y'} - f \\ &= y'^2 \\ &= \frac{1}{4}p^2 \end{aligned}$$

Simple example

So the Hamilton-Jacobi equation is

$$\frac{\partial S}{\partial x} + H\left(x, y, \frac{\partial S}{\partial y}\right) = 0$$

$$\frac{\partial S}{\partial x} + \frac{1}{4}\left(\frac{\partial S}{\partial y}\right)^2 = 0$$

To solve we take $S(x,y) = u(x) + v(y)$ which gives

$$\frac{du}{dx} + \frac{1}{4}\left(\frac{dv}{dy}\right)^2 = 0$$

As u doesn't depend on y, and v doesn't depend on x, the above equation implies that du/dx is a constant, hence we can write

$$u(x) = -\alpha^2 x + \gamma$$

Simple example

Then, the Hamilton-Jacobi equation becomes

$$-\alpha^2 + \frac{1}{4}\left(\frac{dv}{dy}\right)^2 = 0$$

Or

$$\frac{dv}{dy} = 2\alpha$$

So

$$v(x) = 2\alpha y + \beta$$

So we now have

$$S(x,y) = -\alpha^2 x + 2\alpha y + \gamma + \beta$$

Simple example

Taking the derivative of S WRT to β and γ just gives an identity, and so nothing new.

Taking the derivative of S WRT to α gives

$$2y - 2\alpha x = const$$

which is the equation of a straight line.

Simple example

The functional is

$$F\{y\} = \int_a^b y'^2\, dx$$

The E-L equation is

$$\frac{d}{dt}\frac{\partial f}{\partial y'} = \frac{d}{dt} 2y' = y'' = 0$$

which obviously has straight lines as solutions. So the Hamilton-Jacobi equations gave us the same result (in the end).

Pendulum example

$$\begin{aligned}\frac{\partial S}{\partial \phi} &= p = ml^2\dot{\phi} \\ \frac{\partial S}{\partial t} &= -H(t,\phi,p) = -\frac{p^2}{2ml^2} - mgl(1-\cos\phi)\end{aligned}$$

So the Hamilton-Jacobi equation is

$$\frac{\partial S}{\partial t} + \frac{1}{2ml^2}\left(\frac{\partial S}{\partial \phi}\right)^2 + mgl(1-\cos\phi) = 0$$

Hamilton-Jacobi equation

Where there are multiple dependent variables, we write the Hamilton-Jacobi equation as

$$\frac{\partial S}{\partial t} + H\left(t, q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}\right) = 0$$

- Note this is a first order **partial** DE
- May be easier to solve in some cases, but often partial DEs are harder
- Helps if we can separate the variables.

Variational Methods & Optimal Control

lecture 25

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 25

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Conservation Laws

One of the more exciting things we can derive relates to fundamental physics laws: conservation of energy, momentum, and angular momentum. We can now derive all of these from an underlying principle: Noether's theorem.

Hamilton's principle

We now have a group of equivalent methods

- Euler-Lagrange equations
- Hamilton's equations
- Hamilton-Jacobi equation

We saw earlier that these can give us other methods

- Hamilton's principle $\Rightarrow$ Newton's laws of motion
- When L is not explicitly dependent on t, then the Hamiltonian H is constant in time.
  - ▹ conservation of energy
  - ▹ this is an illustration of a symmetry in the problem appearing in the Hamiltonian

Conservation laws

Given the functional

$$F\{y\} = \int_{x_0}^{x_1} f(x,y,y',\ldots,y^{(n)})\,dx$$

if there is a function $\phi(x,y,y',\ldots,y^{(k)})$ such that

$$\frac{d}{dx}\phi(x,y,y',\ldots,y^{(k)}) = 0$$

for all extremals of F, then this is called a **kth order conservation law**

- use obvious extension for functionals of several dependent variables.

Conservation law example

Given the functional

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

where f is not explicitly dependent on t, we know that the Hamiltonian

$$H = y'\frac{\partial f}{\partial y'} - f$$

is constant, and so

$$\frac{dH}{dx} = 0$$

is a first order conservation law for the system.

Several independent variables

For functionals of several independent variables, e.g.

$$F\{z\} = \iint_\Omega z(x,y)\,dx\,dy$$

the equivalent conservation law is

$$\nabla \cdot \phi = 0$$

For some function $\phi(x,y,z,z',\ldots,z^{(k)})$.

- Results here can be extended to these cases, but we won't look at them here.

Conservation laws

- physically interesting
  - tell you about system of interest
- can simplify solution
  - $\phi(x,y,y',\ldots,y^{(k)}) = const$ is an order k DE, rather than E-L equations which are order $2n$
- $\phi(x,y,y',\ldots,y^{(k)}) = const$ is often called the **first integral** of the E-L equations
  - RHS is a constant of integration (determined by boundary conditions)
- how do we find conservation laws?
  - Noether's theorem

Variational symmetries

The key to finding conservation laws lies in finding symmetries in the problem.

- "symmetries" are the result of transformations under which the functional is invariant
- E.G. time invariance symmetry results in constant H
- more generally, take a parameterized family of smooth transforms

  $$X = \theta(x,y;\varepsilon), \quad Y = \phi(x,y;\varepsilon)$$

  where

  $$x = \theta(x,y;0), \quad y = \phi(x,y;0)$$

  e.g. we get the identity transform for $\varepsilon = 0$
- examples **translations** and **rotations**

Jacobian

The Jacobian is

$$J = \begin{vmatrix} \theta_x & \theta_y \\ \phi_x & \phi_y \end{vmatrix} = \theta_x\phi_y - \theta_y\phi_x$$

- **smooth:** if functions x and y have continuous partial derivatives.
- **non-singular:** if Jacobian is non-zero (and hence an inverse transform exists)

Now for $\varepsilon = 0$, we require the identity transform, so $J = 1$. Also, we require a smooth transform, so J is a smooth function of ε, and so for sufficiently small $|\varepsilon|$, the transform is non-singular.

Example transformations

- **translations** (ε is the translation distance)

$$X = x+\varepsilon \qquad Y = y$$
or
$$X = x \qquad Y = y+\varepsilon$$

both have Jacobian

$$J = 1$$

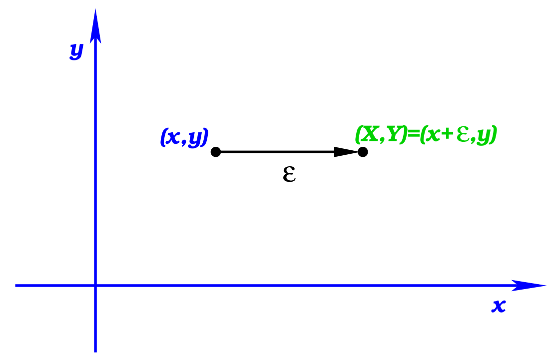
and inverse transformations

$$x = X-\varepsilon \qquad y = Y$$
or
$$x = X \qquad y = Y-\varepsilon$$

Example transformations

- **translations** (ε is the translation distance)

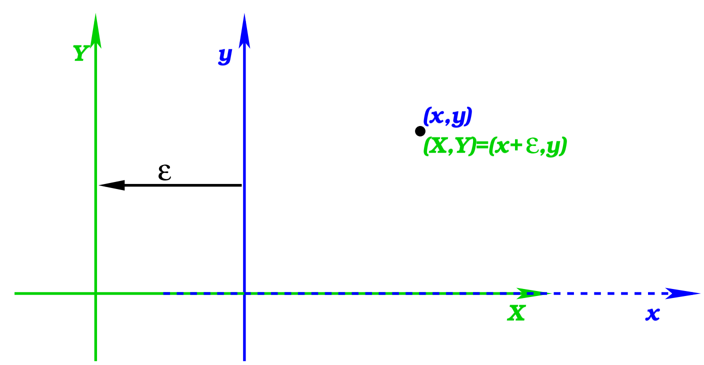
$$X = x+\varepsilon \qquad Y = y$$



Example transformations

- **translations** (ε is the translation distance)

$$X = x+\varepsilon \qquad Y = y$$



Example transformations

- **rotations** (ε is the rotation angle)

$$X = x\cos\varepsilon + y\sin\varepsilon \qquad Y = -x\sin\varepsilon + y\cos\varepsilon$$

has Jacobian

$$J = \cos^2\varepsilon + \sin^2\varepsilon = 1$$

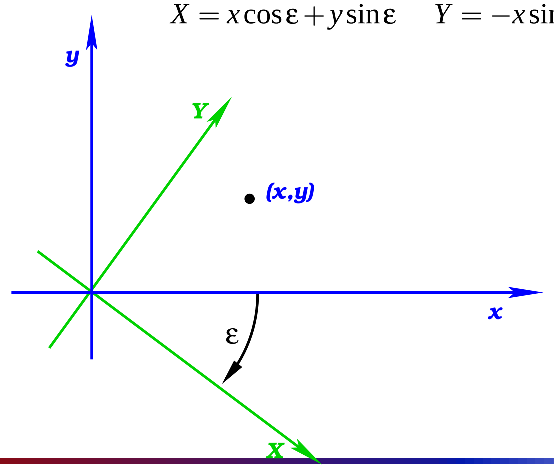
and inverse

$$x = X\cos\varepsilon - Y\sin\varepsilon \qquad y = X\sin\varepsilon + Y\cos\varepsilon$$

Example transformations

- **rotations** (ε is the rotation angle)

$$X = x\cos\varepsilon + y\sin\varepsilon \qquad Y = -x\sin\varepsilon + y\cos\varepsilon$$



Example transformations

- **rotations** (ε is the rotation angle)

$$X = x\cos\varepsilon + y\sin\varepsilon \qquad Y = -x\sin\varepsilon + y\cos\varepsilon$$

To derive this, change coordinates to polar coordinates

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta)$$

Under a rotation by ε, the new coordinates (X,Y) are

$$X = r\cos(\theta-\varepsilon) \quad \text{and} \quad Y = r\sin(\theta-\varepsilon)$$

Use trig. identities $\cos(u-v) = \cos u\cos v + \sin u\sin v$ and $\sin(u-v) = \sin u\cos v - \cos u\sin v$, to get

$$\begin{aligned} X &= r\cos(\theta)\cos(\varepsilon) + r\sin(\theta)\sin(\varepsilon) = x\cos(\varepsilon) + y\sin(\varepsilon) \\ Y &= r\sin(\theta)\cos(\varepsilon) - r\cos(\theta)\sin(\varepsilon) = y\cos(\varepsilon) - x\sin(\varepsilon) \end{aligned}$$

Transformation of a function

Given a function $y(x)$, we can rewrite $Y(X)$ using the inverse transformation, e.g.

$$\phi^{-1}(X,Y(X);\varepsilon) = y(x) = y(\theta^{-1}(X,Y;\varepsilon))$$

For example, taking the curve $y = x$ under rotations

$$X\sin\varepsilon + Y\cos\varepsilon = X\cos\varepsilon - Y\sin\varepsilon$$

which we rearrange to get

$$Y(X) = \frac{\cos\varepsilon - \sin\varepsilon}{\cos\varepsilon + \sin\varepsilon}X$$

Similarly we can derive $Y'(X)$

Transform invariance

If

$$\int_{x_0}^{x_1} f(x,y,y'(x))\,dx = \int_{X_0}^{X_1} f(X,Y,Y'(X))\,dX$$

for all smooth functions $y(x)$ on $[x_0,x_1]$ then we say that the functional in invariant under the transformation.

- also called **variational invariance**
- The transform is called a **variational symmetry**
- Related to conservation laws

Also note that the E-L equations are invariant under such a transform, e.g. they produce the same extremal curves.

Infinitesimal generators

For small ε we can use Taylor's theorem to write

$$\begin{aligned} X &= \theta(x,y;0) + \varepsilon \left.\frac{\partial\theta}{\partial\varepsilon}\right|_{(x,y;0)} + O(\varepsilon^2) \\ Y &= \phi(x,y;0) + \varepsilon \left.\frac{\partial\phi}{\partial\varepsilon}\right|_{(x,y;0)} + O(\varepsilon^2) \end{aligned}$$

Define the infinitesimal generators

$$\xi(x,y) = \left.\frac{\partial\theta}{\partial\varepsilon}\right|_{(x,y;0)} \qquad \eta(x,y) = \left.\frac{\partial\phi}{\partial\varepsilon}\right|_{(x,y;0)}$$

and then for small ε

$$\begin{aligned} X &\simeq x + \varepsilon\xi \\ Y &\simeq y + \varepsilon\eta \end{aligned}$$

Examples

- **translations:**

$$\begin{aligned} &(X,Y) = (x+\varepsilon,y) &&\Rightarrow &&(\xi,\eta) = (1,0) \\ \text{or}\quad &(X,Y) = (x,y+\varepsilon) &&\Rightarrow &&(\xi,\eta) = (0,1) \end{aligned}$$

- **rotations:**

$X = \theta(x,y;\varepsilon) = x\cos\varepsilon + y\sin\varepsilon \quad Y = \phi(x,y;\varepsilon) = -x\sin\varepsilon + y\cos\varepsilon$

So

$$\begin{aligned} \xi &= \left.\frac{\partial\theta}{\partial\varepsilon}\right|_{\varepsilon=0} = -x\sin\varepsilon + y\cos\varepsilon|_{\varepsilon=0} = y \\ \eta &= \left.\frac{\partial\phi}{\partial\varepsilon}\right|_{\varepsilon=0} = -x\cos\varepsilon - y\sin\varepsilon|_{\varepsilon=0} = -x \end{aligned}$$

Emmy Noether

Noether's theorem

Suppose the $f(x,y,y')$ is variationally invariant on $[x_0,x_1]$ under a transform with infinitesimal generators ξ and η, then

$$\eta p - \xi H = const$$

along any extremal of

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

Example (i)

Invariance in translations in x, i.e.

$$\begin{aligned}(X,Y) &= (x+\varepsilon,y)\\ (\xi,\eta) &= (1,0)\end{aligned}$$

So, a system with such invariance has

$$H = const$$

which is what we showed earlier regarding functionals with no explicit dependence on x.

Example (ii)

Invariance in translations in y, i.e.

$$\begin{aligned}(X,Y) &= (x,y+\varepsilon)\\ (\xi,\eta) &= (0,1)\end{aligned}$$

So, a system with such invariance has

$$p = const$$

which is what we showed earlier regarding functionals with no explicit dependence on y.

More than one dependent variable

Transforms with more than one dependent variable

$$\begin{aligned}T &= \theta(t,\mathbf{q};\varepsilon)\\ Q_k &= \phi_k(t,\mathbf{q};\varepsilon)\end{aligned}$$

and the infinitesimal generators are

$$\begin{aligned}\xi &= \left.\frac{\partial\theta}{\partial\varepsilon}\right|_{\varepsilon=0}\\ \eta_k &= \left.\frac{\partial\phi_k}{\partial\varepsilon}\right|_{\varepsilon=0}\end{aligned}$$

More than one dependent variable

Noether's theorem: Suppose $L(t,\mathbf{q},\dot{\mathbf{q}})$ is variationally invariant on $[t_0,t_1]$ under a transform with infinitesimal generators ξ and η_k. Given

$$p = \frac{\partial L}{\partial \dot{q}_k}, \qquad H = \sum_{k=1}^{n} p_k \dot{q}_k - L$$

Then

$$\sum_{k=1}^{n} p_k \eta_k - H\xi = const$$

along any extremal of

$$F\{\mathbf{q}\} = \int_{t_0}^{t_1} L(t,\mathbf{q},\dot{\mathbf{q}})\, dt$$

Example: rotations

Invariance in rotations, i.e.

$$\begin{aligned}(T,Q_1,Q_2) &= (t, q_1\cos\varepsilon + q_2\sin\varepsilon, -q_1\sin\varepsilon + q_2\cos\varepsilon)\\ (t,q_1,q_2) &= (T, Q_1\cos\varepsilon - Q_2\sin\varepsilon, Q_1\sin\varepsilon + Q_2\cos\varepsilon)\end{aligned}$$

The infinitesimal generators are

$$\begin{aligned}\xi &= 0\\ \eta_1 &= -q_1\sin\varepsilon + q_2\cos\varepsilon\big|_{\varepsilon=0} = q_2\\ \eta_2 &= -q_1\cos\varepsilon - q_2\sin\varepsilon\big|_{\varepsilon=0} = -q_1\end{aligned}$$

So, a system with such invariance has

$$\sum_{i=1}^{2} p_i\eta_i - H\xi = p_1 q_2 - p_2 q_1 = const$$

So **angular momentum** in conserved.

Common symmetries

Given a system in 3D with Kinetic Energy $T(\dot{\mathbf{q}}) = \frac{1}{2}m\left(\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2\right)$, and Potential Energy $V(t,\mathbf{q})$.

- invariance of L under time translations corresponds to conservation of Energy
- invariance of L under spatial translations corresponds to conservation of momentum
- invariance of L under rotations corresponds to conservation of angular momentum

Finding symmetries

Testing for non-trivial symmetries can be tricky.
Useful result is the Rund-Trautman identity:
It leads also to a simple proof of Noether's theorem

More advanced cases

- Laplace-Runge-Lenz vector in planetary motion corresponds to rotations of 3D sphere in 4D
- symmetries in general relativity
- symmetries in quantum mechanics
- symmetries in fields

Variational Methods & Optimal Control

lecture 26

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 26

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Pontryagin Maximum Principle

Modern optimal control theory often starts from the PMP. It is a simple, concise condition for an optimal control.

General control problem

Minimize functional

$$F = \int_{t_0}^{t_1} f_0(t, \mathbf{x}, \mathbf{u})\, dt$$

subject to constraints $\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u})$, or more fully,

$$\dot{x}_i = f_i(t, \mathbf{x}, \mathbf{u})$$

- notice no dependence on $\dot{\mathbf{x}}$ in f_0
  - this differs from many CoV problems
- no dependence on $\dot{\mathbf{x}}$ in f_i because we rearrange the equations so that derivatives are on the LHS

Pontryagin Maximum Principle (PMP)

Let $\mathbf{u}(t)$ be an admissible control vector that transfers $(t_0, \mathbf{x}_0)$ to a target $(t_1, \mathbf{x}(t_1))$. Let $\mathbf{x}(t)$ be the trajectory corresponding to $\mathbf{u}(t)$. In order that $\mathbf{u}(t)$ be optimal, it is necessary that there exists $\mathbf{p}(t) = (p_1(t), p_2(t), \ldots, p_n(t))$ and a constant scalar p_0 such that

- $\mathbf{p}$ and $\mathbf{x}$ are the solution to the canonical system

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{p}} \quad \text{and} \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{x}}$$

- where the Hamiltonian is $H = \sum_{i=0}^{n} p_i f_i$ with $p_0 = -1$
- $H(\mathbf{x}, \mathbf{u}, \mathbf{p}, t) \geq H(\mathbf{x}, \hat{\mathbf{u}}, \mathbf{p}, t)$ for all alternate controls $\hat{\mathbf{u}}$
- all boundary conditions are satisfied

PMP proof sketch

Consider the general problem: minimize functional

$$F\{\mathbf{x},\mathbf{u}\} = \int_{t_0}^{t_1} f_0(t,\mathbf{x},\mathbf{u})\,dt$$

subject to constraints

$$\dot{x}_i = f_i(t,\mathbf{x},\mathbf{u})$$

We can incorporate the constraints into the functional using the Lagrange multipliers λ_i, e.g.

$$\tilde{F} = \int_{t_0}^{t_1} L(t,\mathbf{x},\dot{\mathbf{x}},\mathbf{u})\,dt = \int_{t_0}^{t_1} f_0(t,\mathbf{x},\mathbf{u}) + \sum_{i=1}^{n} \lambda_i(t)\left[\dot{x}_i - f_i(t,\mathbf{x},\mathbf{u})\right] dt$$

PMP proof sketch

Given such a function we get (by definition)

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = \lambda_i$$

So we can identify the Lagrange multipliers λ_i with the **generalized momentum** terms p_i

- the p_i are known in economics literature as **marginal valuation** of x_i or the **shadow prices**
- shows how much a unit increment in x at time t contributes to the optimal objective functional $\tilde{F}$
- the p_i are known in control as **co-state variables** (sometimes written as z_i)

PMP proof sketch

By definition (in previous lectures) the Hamiltonian is

$$\begin{aligned} H(t,\mathbf{x},\mathbf{p},\mathbf{u}) &= \sum_{i=1}^{n} p_i\dot{x}_i - L(t,\mathbf{x},\dot{\mathbf{x}},\mathbf{u}) \\ &= \sum_{i=1}^{n} p_i\dot{x}_i - f_0(t,\mathbf{x},\mathbf{u}) - \sum_{i=1}^{n} \lambda_i(t)\left[\dot{x}_i - f_i(t,\mathbf{x},\mathbf{u})\right] \\ &= -f_0(t,\mathbf{x},\mathbf{u}) + \sum_{i=1}^{n} p_i f_i(t,\mathbf{x},\mathbf{u}) \end{aligned}$$

because $\lambda_i = p_i$, so the $\dot{x}_i$ terms cancel. The final result is just the Hamiltonian as defined in the PMP.

PMP proof sketch

From previous slide the Hamiltonian can be written

$$H(t,\mathbf{x},\mathbf{p},\mathbf{u}) \;=\; -f_0(t,\mathbf{x},\mathbf{u}) + \sum_{i=1}^{n} p_i f_i(t,\mathbf{x},\mathbf{u})$$

which is the Hamiltonian defined in the PMP. Then the Canonical E-L equations (Hamilton's equations) are

$$\frac{\partial H}{\partial p_i} = \frac{dx_i}{dt} \quad \text{and} \quad \frac{\partial H}{\partial x_i} = -\frac{dp_i}{dt}$$

Note that the equations $\frac{\partial H}{\partial p_i} = \frac{dx_i}{dt}$ just revert to

$$f_i(t,\mathbf{x},\mathbf{u}) = \dot{x}_i$$

which are just the system equations.

PMP proof sketch

Finally, note that Hamilton's equations above only relate x_i and its conjugate momentum p_i. What about equations for u_i? Take the conjugate variable to be z_i, and we get (by definition) that

$$z_i = \frac{\partial L}{\partial \dot{u}_i} = 0$$

and the second of Hamilton's equations is therefore

$$\frac{\partial H}{\partial u_i} = -\frac{dz_i}{dt} = 0$$

which suggests a stationary point of H WRT u_i. In fact we look for a maximum (and note this may happen on the bounds of u_i)

PMP Example: plant growth

Plant growth problem:

- market gardener wants to plants to grow to a fixed height 2 within a fixed window of time $[0,1]$
- can supplement natural growth with lights (at night)
- growth rate dictates

$$\dot{x} = 1 + u$$

- cost of lights

$$F\{u\} = \int_0^1 \frac{1}{2} u^2 \, dt$$

PMP Example: plant growth

Minimize

$$F\{u\} = \int_0^1 \frac{1}{2} u^2 \, dt$$

Subject to $x(0) = 0$, and $x(1) = 2$ and

$$\dot{x} = f_1(t,x,u) = 1 + u$$

Hamiltonian is

$$\begin{aligned} H &= -f_0(t,x,u) + p f_1(t,x,u) \\ &= -\frac{1}{2}u^2 + p(1+u) \end{aligned}$$

PMP Example: plant growth

Hamiltonian is

$$H = -\frac{1}{2}u^2 + p(1+u)$$

Canonical equations

$$\begin{array}{ccccccc} \frac{\partial H}{\partial p} & = & \frac{dx}{dt} & \text{and} & \frac{\partial H}{\partial x} & = & -\frac{dp}{dt} \\ & \Downarrow & & & & \Downarrow & \\ 1+u & = & \dot{x} & & 0 & = & -\dot{p} \end{array}$$

LHS => system DE
RHS => $\dot{p} = 0$ means that $p = c_1$ where c_1 is a constant.

PMP Example: plant growth

Maximum principle requires H be a maximum, for which

$$\frac{\partial H}{\partial u} = -u + p = 0$$

So $u = p$, and $\dot{x} = 1 + u$ so

$$x = (1 + c_1)t + c_2$$

The solution which satisfies $x(0) = 0$ and $x(1) = 2$ is

$$x = 2t$$

So $u = c_1 = 1$, and the optimal cost is $1/2$.

PMP and Transversal conditions

Typically we fix t_0 and $\mathbf{x}(t_0)$, but often the right-hand boundary condition is not fixed, so we need transversal, or natural boundary conditions. Here, they differ from traditional CoV problems in two respects:

- The terminal cost ϕ
- The function f_0 is not explicitly dependent on $\dot{x}$

The resulting transversal conditions are

$$\sum_i \left(\frac{\partial \phi}{\partial x_i} + p_i \right) \delta x_i \bigg|_{t=t_1} + \left(\frac{\partial \phi}{\partial t} - H \right) \delta t \bigg|_{t=t_1} = 0$$

for all allowed δx_i and δt.

PMP and Transversal conditions

The resulting transversal condition is

$$\sum_i \left(\frac{\partial \phi}{\partial x_i} + p_i \right) \delta x_i \bigg|_{t=t_1} + \left(\frac{\partial \phi}{\partial t} - H \right) \delta t \bigg|_{t=t_1} = 0$$

Special cases

- when t_1 is fixed and $\mathbf{x}(t_1)$ is completely free we get

$$\left(\frac{\partial \phi}{\partial x_i} + p_i \right) \bigg|_{t=t_1} = 0, \quad \forall i$$

- when $\mathbf{x}(t_1)$ is fixed, $\delta x_i = 0$, and we get

$$\left(\frac{\partial \phi}{\partial t} - H \right) \bigg|_{t=t_1} = 0$$

Example: stimulated plant growth

Plant growth problem:

- market gardener wants to plants to grow as much as possible within a fixed window of time $[0,1]$
- supplement natural growth with lights as before
- growth rate dictates $\dot{x} = 1 + u$
- cost of lights

$$F\{u\} = \int_0^1 \frac{1}{2} u(t)^2 \, dt$$

- value of crop is proportional to the height

$$\phi(t_1, \mathbf{x}(t_1)) = x(t_1)$$

Plant growth problem statement

Write as a minimization problem

$$F\{u,x\} = -x(t_1) + \int_0^1 \frac{1}{2}u^2\,dt$$

Subject to $x(0) = 0$,

$$\dot{x} = 1 + u$$

- the terminal cost doesn't affect the shape of the solution
- but we need a natural end-point condition for t_1

Plant growth: natural boundary cond.

The problem is solved as before, but we write the natural boundary condition at $x = t_1$ as

$$\left.\left(\frac{\partial \phi}{\partial x_i} + p_i\right)\right|_{t=t_1} = 0, \quad \forall i$$

which reduces to

$$-1 + p|_{t=t_1} = 0$$

Given p is constant, this sets $p(t) = 1$, and hence the control $u = 1$ (as before).

Autonomous problems

Autonomous problems have no explicit dependence on t.

- time invariance symmetry
- hence H is constant along the optimal trajectory
- if the end-time is free (and the terminal cost is zero) then the transversality conditions ensure $H = 0$ along the optimal trajectory.

PMP Example: Gout

Optimal Treatment of Gout:

- disease characterized by excess of uric acid in blood
  - define level of uric acid to be $x(t)$
  - in absence of any control, tends to 1 according to

$$\dot{x} = 1 - x$$

- drugs are available to control disease (control u)

$$\dot{x} = 1 - x - u$$

  - aim to reduce x to zero as quickly as possible
  - drug is expensive, and unsafe (side effects)

PMP Example: Gout

Formulation: Minimize

$$F\{u\} = \int_0^{t_1} \frac{1}{2}(k^2 + u^2)\, dt$$

given constant k that measures the relative importance of the drugs cost vs the terminal time. End-conditions are $x(0) = 1$, and we wish $x(t_1) = 0$, with t_1 free. The constraint equation is

$$\dot{x} = 1 - x - u$$

Hamiltonian

$$H = -\frac{1}{2}(k^2 + u^2) + p(1 - x - u)$$

PMP Example: Gout

Canonical equations

$$\frac{\partial H}{\partial p} = \frac{dx}{dt} \Rightarrow 1 - x - u = \dot{x} \quad \text{and} \quad \frac{\partial H}{\partial x} = -\frac{dp}{dt} \Rightarrow -p = -\dot{p}$$

LHS => system DE
RHS => $\dot{p} = p$ has solution $p = c_1 e^t$
Now maximize H wrt to u, i.e., find stationary point

$$\frac{\partial H}{\partial u} = -u - p = 0$$

So $u = -p = -c_1 e^t$

PMP Example: Gout

Note

- this is an autonomous problem so $H = const$
- this is a free end-time problem so $H = 0$

Substitute values of p and u into H for $t = 0$ (i.e. $p = c_1 = -u$, and $x(0) = 1$), and we get

$$\begin{aligned} H &= -\frac{1}{2}(k^2 + u^2) + p(1 - x - u) \\ &= -\frac{k^2}{2} - \frac{c_1^2}{2} - c_1^2 \\ &= 0 \end{aligned}$$

and so $c_1 = \pm k$

PMP Example: Gout

Finally solve $\dot{x} = 1 - x - u$ where $u = -ke^t$ to get

$$x = 1 - \frac{k}{2}e^t + \frac{k}{2}e^{-t} = 1 - k \sinh t$$

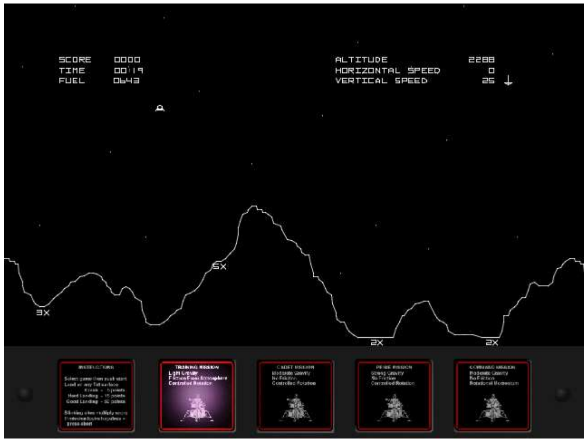
The terminal condition is $x(t_1) = 0$, and so

$$t_1 = \sinh^{-1}(1/k)$$

- when k is small the prime consideration is to use a small amount of the drug, and as $k \to 0$ then $t_1 \to \infty$
  - no optimal for $k = 0$
- when k is large, we want to get to a safe level as fast as possible, so as $k \to \infty$ we get $t_1 \sim 1/k$

PMP Example: Lunar lander

Atari game, 1979



http://www.klov.com/game_detail.php?letter=L&game_id=8465

PMP Example: Lunar lander

- need to land surface-module on the moon
  - Module mass M (ignore fuel load), uniform gravitational acceleration g (might not be $9.8m/s^2$)
  - initial height $y(0) = h$
  - initial velocity $\dot{y}(0) = v$
- controlled descent so landing is "soft"
  - height of module, and downward velocity brought to zero simultaneously
- thrust f either up or down
  - thrust is bounded, so $|f| \leq f_{\max}$
  - want to minimize fuel cost $|f|$ over time

PMP Example: Lunar lander

System defined (at any time t) by

- position y
- velocity $\dot{y}$

State equations (mass $\times$ acceleration = force)

$$M\ddot{y} = -Mg + f$$

Initial state

$$y(0) = h, \quad \text{and} \quad \dot{y}(0) = v$$

Desired final state (t_1 is free)

$$y(t_1) = 0 \quad \text{and} \quad \dot{y}(t_1) = 0$$

and we wish to minimize

$$F\{f\} = \int_0^{t_1} |f| \, dt$$

PMP Example: Lunar lander

Convert the problem to standard form by taking

$$\begin{aligned} x_1 &= y \\ x_2 &= \dot{y} \\ u &= f/M \end{aligned}$$

So the state equation becomes

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g + u \end{aligned}$$

And the initial and final conditions are

$$\begin{aligned} x_1(0) = h \quad &\text{and} \quad x_2(0) = v \\ x_1(t_1) = 0 \quad &\text{and} \quad x_2(t_1) = 0 \end{aligned}$$

PMP Example: Lunar lander

Hamiltonian

$$H = -|u| + p_1 x_2 + p_2(u - g)$$

Canonical equations

$$\frac{\partial H}{\partial p_i} = \frac{dx_i}{dt} \quad \text{and} \quad \frac{\partial H}{\partial x_i} = -\frac{dp_i}{dt}$$

Give the constraints $\dot{x}_1 = x_2$ and $\dot{x}_2 = -g + u$ and

$$\begin{aligned} \frac{\partial H}{\partial x_1} &= 0 = -\dot{p}_1 \\ \frac{\partial H}{\partial x_2} &= p_1 = -\dot{p}_2 \end{aligned}$$

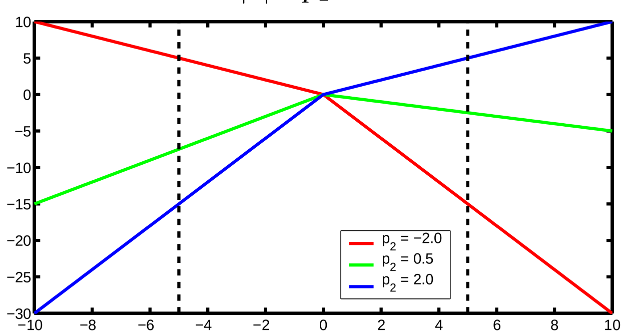
Solution $p_1 = c_1$ and $p_2 = -c_1 t + c_2$.

PMP Example: Lunar lander

Now we have to choose u to maximize H

- $|u|$ is bounded by f_{max}/M

Ignore the terms in H that are constant WRT to u and we have to maximize $-|u| + p_2 u$.



PMP Example: Lunar lander

Maximize $f(u) = -|u| + p_2 u$, with $|u| \leq 1$

- three possible locations for a maximum
  - ▷ left or right boundary, or $u = 0$
- The three values (in order from left to right) are

$$f(u) = -1 - p_2, \quad 0, \quad -1 + p_2$$

- Three cases $p_2 < -1$, $-1 < p_2 < 1$ or $p_2 > 1$
- maximum occurs at

$$u = \begin{cases} +1, & \text{if } p_2 > 1 \\ 0, & \text{if } -1 < p_2 < 1 \\ -1, & \text{if } p_2 < -1 \end{cases}$$

- If bounds are $|u| \leq f_{max}/M$, then the solution scales.

PMP Example: Lunar lander

Call p_2 a switching function, and note that we have

$$p_2 = -c_1 t + c_2$$

- during the final descent, $x_2 < 0$
  - ▷ we must be going down just before we land
- but $x_2(t_1) = 0$, so $\dot{x}_2 > 0$ near t_1
  - ▷ we must be decelerating, so that we stop at t_1
  - ▷ hence we must have positive thrust
  - ▷ optimal thrust must be at max, e.g. $u = f_{max}/M$
- so the equations for motion during final descent are

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g + f_{max}/M = k > 0 \end{aligned}$$

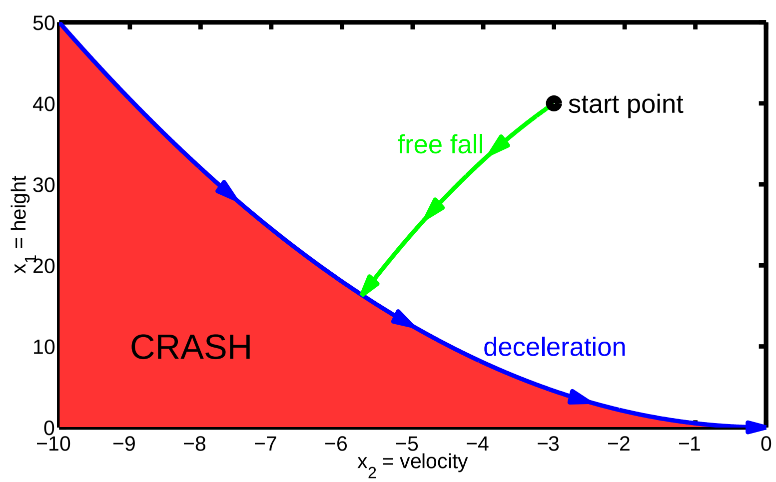
PMP Example: Lunar lander

Given final conditions the solution near landing is

$$x_1 = \frac{1}{2}k(t-t_1)^2 \quad \text{and} \quad x_2 = k(t-t_1)$$

- note $k > 0$ in final stages of landing
- note $u = f_{\max}/M$ in final stages of landing
- given $p_2 = -c_1 t + c_2$ we must have $c_1 < 0$
- hence prior stages of control include
  - a stage when $u = 0$ (free fall)
  - a stage when $u = -f_{\max}/M$ (accelerating down)
- in each stage we get an equation as above, but with different constant k, for $u = 0$ and $u = -f_{\max}/M$ the constant $k < 0$

PMP Example: Lunar lander



PMP Example: Lunar lander

Solution:

- if start above, or on the critical curve
  - if travelling upwards, max thrust down to cancel upwards velocity
  - then free-fall, until on the critical curve

$$x_1 = \frac{1}{2}k(t-t_1)^2 \quad \text{and} \quad x_2 = k(t-t_1)$$

  - max thrust up until stop on the surface
- if lie below the critical curve
  - you crash

PMP Example: Lunar lander

- What's the point of this example
  - previously, we couldn't easily deal with and objective like $|u|$
  - the function isn't "smooth"
  - PMP can work for such examples
  - it doesn't require smoothness, you just need to be able to find a maximum

Variational Methods & Optimal Control

lecture 27

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 27

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Bang-Bang controllers and other related issues

Here we consider more generally what conditions result in a bang-bang controller.

Bang-Bang controllers

A linear optimal control problem is one in which the **control variables** $\mathbf{u}$ enter the Hamiltonian linearly, e.g.

$$H = \psi(\mathbf{x}, \mathbf{p}, t) + \sigma(\mathbf{x}, \mathbf{p}, t)^T \mathbf{u}(t)$$

Examples:

- a time minimization problem, with linear state equation
$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$$
- the optimal economic growth model with $U(c) = c$, so the functional is
$$F\{c\} = \int_0^T c(t) e^{-rt}\, dt$$
subject to $\dot{k} = f(k) - \lambda k - c$ leads to the Hamiltonian
$$H = \left(e^{-rt} - p\right) c + p\left(f(k) - \lambda k\right)$$

Bang-Bang controllers

In general (for a linear problem) there will be no extremal unless the control is bounded, e.g. $m_i \leq u_i \leq M_i$, but where m_i and M_i are constant, we can re-scale the problem to consider bounded controls $|\tilde{u}_i| \leq 1$, by taking

$$\tilde{u}_i = 2\frac{u_i - m_i}{M_i - m_i} - 1$$

When the PMP is applied to this type of problem the optimal control is

$$u_i(t) = \begin{cases} 1, & \text{if } \sigma_i > 0 \\ -1 & \text{if } \sigma_i < 0 \end{cases}$$

Where $\sigma_i \neq 0$ is a **bang-bang** controller (otherwise it is singular), and σ_i is a **switching function**

Explanation

Consider a linear problem with one control u, then

$$H = \psi(\mathbf{x},\mathbf{p},t) + \sigma(\mathbf{x},\mathbf{p},t)u$$

- The PMP requires us to maximize H for all u.
- The derivative of H WRT to u is $\sigma(\mathbf{x},\mathbf{p},t)$.
- If $\sigma(\mathbf{x},\mathbf{p},t) \neq 0$ the derivative is never zero.
- Hence the maximum will occur at the bounds of u.
- If $\sigma(\mathbf{x},\mathbf{p},t) > 0$, the maximum will occur for the positive bound of u, whereas if $\sigma(\mathbf{x},\mathbf{p},t) < 0$ the maximum will occur for the negative bound.
- Hence σ is a switching function.

Example: optimal fish harvesting

- fish stock (population $x(t)$)
- grows at a fixed rate γ, so without harvesting

$$\dot{x} = \gamma x$$

- harvesting at rate u reduces the population

$$\dot{x} = \gamma x - u$$

- we wish to harvest the maximum number of fish in time T,
  - discount by rate r for future harvests
  - maximize

$$F\{u\} = \int_0^T u e^{-rt}\, dt$$

Example: optimal fish harvesting

Problem formulation: Maximize

$$F\{u\} = \int_0^T u e^{-rt}\, dt$$

subject to

$$\dot{x} = \gamma x - u$$

and $x(0) = 1$, and $x(T)$ free.

Equivalent problem: Minimize

$$F\{u\} = \int_0^T -u e^{-rt}\, dt$$

subject to

$$\dot{x} = \gamma x - u$$

and $x(0) = 1$, and $x(T)$ free.

Example: optimal fish harvesting

The Hamiltonian is

$$H = u e^{-rt} + p(\gamma x - u)$$

which is linear in the control variable.
Hamilton's equations (the canonical, or co-state equations) are

$$\frac{\partial H}{\partial p} = \frac{dx}{dt} \quad \text{and} \quad \frac{\partial H}{\partial x} = -\frac{dp}{dt}$$

The first of Hamilton's equations just gives back the growth equation $\dot{x} = \gamma x - u$, the second gives

$$\frac{\partial H}{\partial x} = \gamma p = -\frac{dp}{dt}$$

which has solution $p = c_1 e^{-\gamma t}$.

Example: optimal fish harvesting

The Hamiltonian is

$$H = ue^{-rt} + p(\gamma x - u) = p\gamma x + \left[e^{-rt} - p\right] u$$

which is linear in the control variable. The control must be bounded, and will be bang-bang with switching function

$$\sigma = e^{-rt} - p = e^{-rt} - c_1 e^{-\gamma t}$$

For $0 \le u \le 1$ we get $u = 0$ or 1.

$$u(t) = \begin{cases} 1, & \text{if } \sigma_i > 0 \\ 0, & \text{if } \sigma_i < 0 \end{cases}$$

Example: optimal fish harvesting

Given fixed end-time T, but free $x(T)$, then the natural boundary condition is $p(T) = 0$, so $c_1 = 0$, and

$$\sigma = e^{-rt} - c_1 e^{-\gamma t} = e^{-rt} > 0$$

- result is fishing at maximum rate
- if the fishing rate u is greater than the growth rate γx then the fish stock will eventually die out.

This model may be a big simplification (ignores economic factors like cost of fishing, or demand), but it does show some interesting features.

- **control is needed, or you get over-fishing!**

Time Minimization Problem

Time minimization, the functional to minimize is

$$T\{\mathbf{x},\mathbf{u}\} = \int_{t_0}^{t_1} 1 \, dt$$

Given that the starting state is $\mathbf{x}(t_0) = \mathbf{x}_0$, and the desired end state is $\mathbf{x}(t_1) = \mathbf{x}_1$, but that t_1 is not fixed, and $\mathbf{x}$ is subject to some DE

$$\dot{\mathbf{x}} = g(\mathbf{x}, \mathbf{u}, t)$$

To get a linear autonomous problem, we need that

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$$

Time Minimization Problem

Linear autonomous time minimization, the functional to minimize is

$$T\{\mathbf{x},\mathbf{u}\} = \int_{t_0}^{t_1} 1 \, dt$$

subject to

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$$

where A is a $n \times n$ constant matrix, and B is a $n \times m$ constant matrix. The controller is assumed to be bounded, e.g.

$$|u_i| \le 1, \quad \text{for } i = 1, \ldots, m$$

The Hamiltonian and generalized momentum will be

$$H = -1 + \mathbf{p}^T A\mathbf{x} + \mathbf{p}^T B\mathbf{u} \quad \text{and} \quad \dot{\mathbf{p}} = -H_{\mathbf{x}} = -A^T \mathbf{p}$$

which is linear in the controller $\mathbf{u}$.

Time Minimization Problem

We know the control will governed by the **switching function**

$$\sigma = \mathbf{p}^T B$$

so we get the control

$$u_i(t) = \begin{cases} 1, & \text{if } \mathbf{p}^T\mathbf{b}_i > 0 \\ -1 & \text{if } \mathbf{p}^T\mathbf{b}_i < 0 \\ \text{unknown} & \text{if } \mathbf{p}^T\mathbf{b}_i = 0 \end{cases}$$

where the $\mathbf{b}_i$ are the m columns of the matrix B. Given $\dot{\mathbf{p}} = -A^T\mathbf{p}$, so $\mathbf{p} = e^{-A^T(t-t_0)}\mathbf{p}_0$, it is unlikely that $\mathbf{p}^T\mathbf{b}_i = 0$, so singular control is ruled out, and the control is bang-bang.

Time Minimization Problem

In general a control may or may not exist!

- **existence:** If A is a stable matrix (i.e., all the eigenvalues of A have non-positive real parts), then for any point $\mathbf{x}_0$, there exists an optimal control which will go from $\mathbf{x}_0$ to the origin.

  This is useful because we can rewrite the problem so that the desired end-point $\mathbf{x}(t_1) = \mathbf{0}$.
- **uniqueness:** If an optimal control exists, it is unique.
- **switching:** If the eigenvalues of the $n \times n$ matrix A are all real, then there exists a unique control control, where each $u_i = \pm 1$ is piecewise constant and has no more than $n-1$ switchings.

Time Minimization Problem

Example: parking problem (from Lecture 19)

Rewrite the problem so the point B is at the origin ($x(t_1) = 0$), and the control $u =$ Force/mass is bounded by $|u| \leq 1$. The differential equation

$$\ddot{x} = u$$

can be written as two first order DEs by defining $x_1 = x$ and $x_2 = \dot{x}$, so that

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$

Time Minimization Problem

The matrix A has eigenvalues $\lambda = 0, 0$, and so satisfies the existence and uniqueness conditions. The Hamiltonian is

$$H = -1 + p_1 x_2 + p_2 u$$

So the switching function is p_2. Hamilton's equations (PMP) results in

$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = 0 \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = -p_1 \end{aligned}$$

with solution (c_1 and c_2 are constants of integration)

$$\begin{aligned} p_1 &= c_1 \\ p_2 &= -c_1 t + c_2 \end{aligned}$$

Time Minimization Problem

The switching function $p_2 = -c_1 t + c_2$ is guaranteed to change sign at most $n - 1 = 1$ times, so the possible solutions are

$$\begin{aligned} u &= 1 \text{ for all } t \in [0, T] \\ u &= -1 \text{ for all } t \in [0, T] \\ u &= \begin{cases} -1 \text{ for all } t \in [0, t_s) \\ 1 \text{ for all } t \in (t_s, T] \end{cases} \\ u &= \begin{cases} 1 \text{ for all } t \in [0, t_s) \\ -1 \text{ for all } t \in (t_s, T] \end{cases} \end{aligned}$$

Time Minimization Problem

Solving the DE for $u = \pm 1$

$$\begin{aligned} x_2 &= \pm t + c_3 \\ x_1 &= \pm \frac{1}{2} t^2 + c_3 t + c_4 \end{aligned}$$

Time can be eliminated from the above by squaring the first equation and multiplying by 1/2,

$$\begin{aligned} \frac{1}{2} x_2^2 &= \frac{1}{2} t^2 \pm c_3 t + \frac{1}{2} c_3^2 \\ x_1 &= \pm \frac{1}{2} t^2 + c_3 t + c_4 \end{aligned}$$

Time Minimization Problem

For $u = \pm 1$

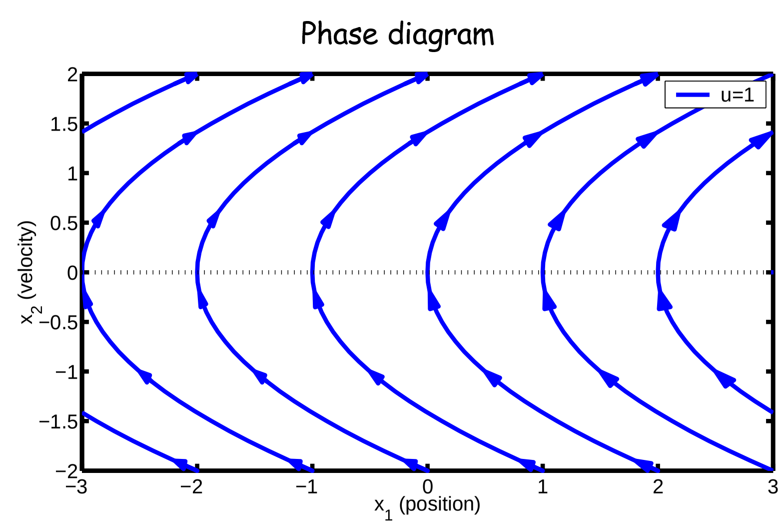
$$\begin{aligned} \frac{1}{2} x_2^2 &= \frac{1}{2} t^2 \pm c_3 t + \frac{1}{2} c_3^2 \\ x_1 &= \pm \frac{1}{2} t^2 + c_3 t + c_4 \end{aligned}$$

so we can write x_1 as a function of x_2

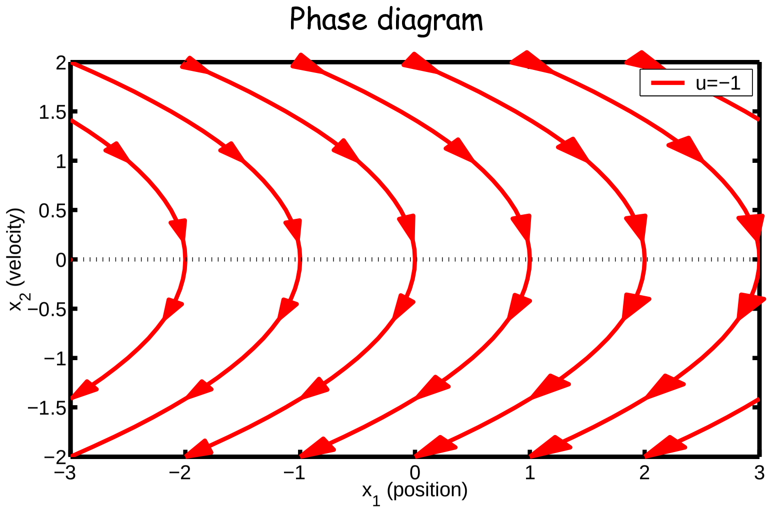
$$x_1 = \begin{cases} \frac{1}{2} x_2^2 + c_5 & \text{for } u = 1 \\ -\frac{1}{2} x_2^2 + c_6 & \text{for } u = -1 \end{cases}$$

where $c_5 = c_4 - \frac{1}{2} c_3^2$ and $c_6 = c_4 + \frac{1}{2} c_3^2$

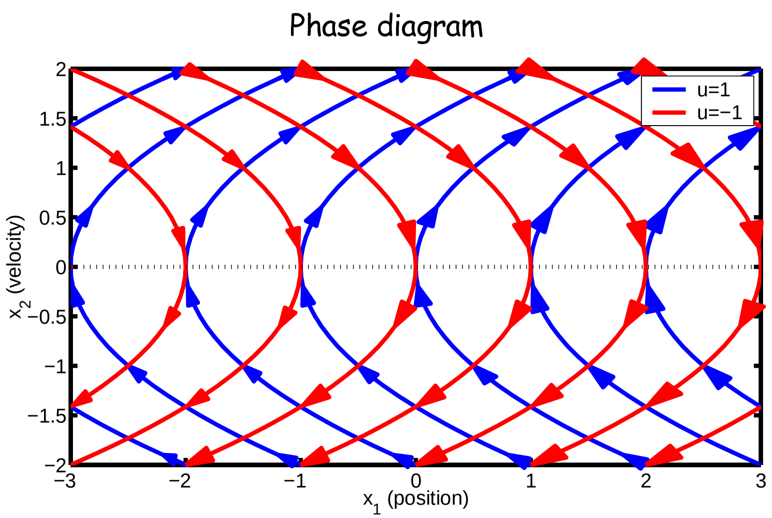
Phase diagram 1



Phase diagram 2



Combined phase diagram



Time Minimization Problem

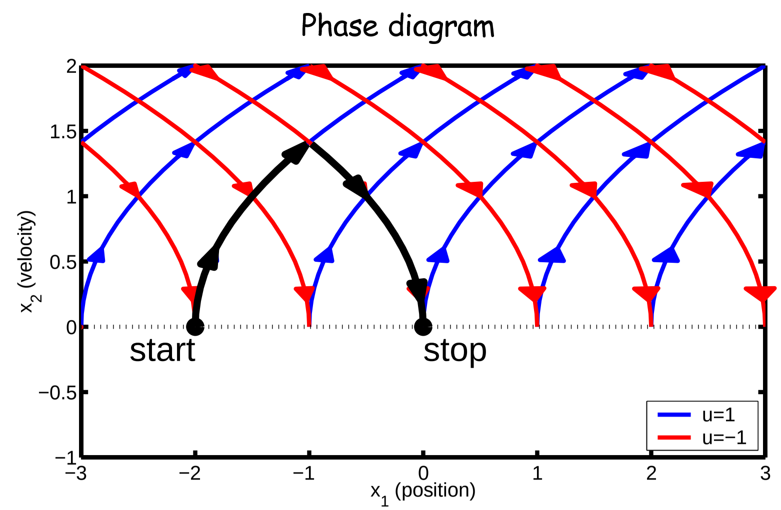
Parking problem: moving from point A (at $x = -2$) to B (at $x = 0$) and be stationary at both start and stop times. Given

$$\begin{aligned} x_1 &= \text{position} \\ x_2 &= \text{velocity} \end{aligned}$$

the end-point conditions are

$$\begin{aligned} x_1(0) &= -2 & \qquad x_1(T) &= 0 \\ x_2(0) &= 0 & \qquad x_2(T) &= 0 \end{aligned}$$

Time Minimization Problem



Time Minimization Problem

So the solution is case (4)

$$u = \begin{cases} 1 \text{ for all } t \in [0, t_s) \\ -1 \text{ for all } t \in (t_s, T] \end{cases}$$

Hence we know that the initial trajectory will be

$$\begin{aligned} x_2 &= t + c_3 \\ x_1 &= \frac{1}{2}t^2 + c_3 t + c_4 \end{aligned}$$

with $x_1(0) = -2$ and $x_2(0) = 0$, so $c_3 = 0$ and $c_4 = -2$, with result (for $t < t_s$)

$$\begin{aligned} x_2 &= t \\ x_1 &= \frac{1}{2}t^2 - 2 \end{aligned}$$

Time Minimization Problem

So the solution is case (4)

$$u = \begin{cases} 1 \text{ for all } t \in [0, t_s) \\ -1 \text{ for all } t \in (t_s, T] \end{cases}$$

Hence we know that the final trajectory will be

$$\begin{aligned} x_2 &= -t + c_3' \\ x_1 &= -\frac{1}{2}t^2 + c_3' t + c_4' \end{aligned}$$

with $x_1(T) = 0$ and $x_2(T) = 0$, so $c_3' = T$ and $c_4' = -T^2/2$, with result that for $t_s < t \leq T$

$$\begin{aligned} x_2 &= T - t \\ x_1 &= -\frac{(T-t)^2}{2} \end{aligned}$$

Time Minimization Problem

At time t_s the two paths must join, so we get the conditions

$$\begin{aligned} \lim_{t \to t_s^-} x_1(t) &= \lim_{t \to t_s^+} x_1(t) \\ \lim_{t \to t_s^-} x_2(t) &= \lim_{t \to t_s^+} x_2(t) \end{aligned}$$

When we substitute the initial and final paths, we get

$$\begin{aligned} \frac{1}{2}t_s^2 - 2 &= -\frac{(T-t_s)^2}{2} \\ t_s &= T - t_s \end{aligned}$$

The second equation requires that $t_s = T/2$, which we can observe directly from the symmetry of the phase diagram.

Time Minimization Problem

The continuity conditions are

$$\begin{aligned} \frac{1}{2}t_s^2 - 2 &= \frac{(T-t_s)^2}{2} \\ t_s &= T - t_s \end{aligned}$$

Given $t_s = T/2$ the first equation becomes

$$\frac{1}{8}T^2 - 2 = -\frac{T^2}{8}$$

which we rearrange to get

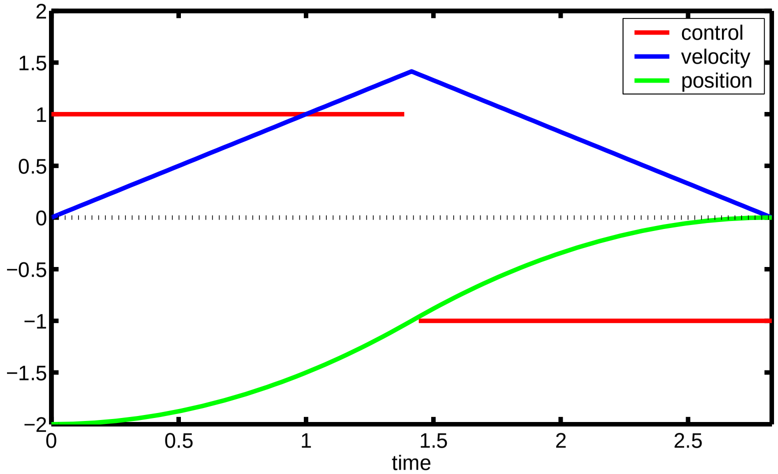
$$T^2 = 8$$

From the problem formulation $T > 0$, and so we take

$$T = 2\sqrt{2} \quad \text{and} \quad t_s = \sqrt{2}$$

Time Minimization Problem

Solution relative to time



Singular control

Linear problem,

$$H = \psi(\mathbf{x},\mathbf{p},t) + \sigma(\mathbf{x},\mathbf{p},t)^T\mathbf{u}(t)$$

Optimal control is

$$u_i(t) = \begin{cases} 1, & \text{if } \sigma_i > 0 \\ -1 & \text{if } \sigma_i < 0 \\ \text{unknown} & \text{if } \sigma_i = 0 \end{cases}$$

When $\sigma(\mathbf{x},\mathbf{p},t) = 0$ the control u has no effect on H

- the PMP fails: we have no information about the optimal control
- called singular, degenerate, irregular, or ambiguous control.

Singular control

If $\sigma(\mathbf{x},\mathbf{p},t) = 0$ only for isolated points there there is no problem. If $\sigma(\mathbf{x},\mathbf{p},t) = 0$ over an interval, then within the interval

$$\dot{\sigma}(\mathbf{x},\mathbf{p},t) = \ddot{\sigma}(\mathbf{x},\mathbf{p},t) = \ldots = 0$$

then singular control must be used.

- similar in nature to the CoV case where the functional is linear in y', and so we have a degenerate solution (see earlier lectures).

Singular control

Linear-autonomous time-minimization problem, where

$$H = \psi(\mathbf{x},\mathbf{p}) + \sigma(\mathbf{x},\mathbf{p})u(t)$$

where $\sigma(\mathbf{x},\mathbf{p}) = 0$ over some interval.

- autonomous problems implies $H = const$
- free-end time implies $H = 0$ for all $t \in [0,T]$
- So $\psi(\mathbf{x},\mathbf{p}) = 0$ over the same interval as $\sigma(\mathbf{x},\mathbf{p}) = 0$.
- Similarly for the kth order derivatives of ψ and σ
- Using the chain rule

$$\dot{\sigma}(\mathbf{x},\mathbf{p}) = \frac{\partial\sigma}{\partial\mathbf{x}}\dot{\mathbf{x}} + \frac{\partial\sigma}{\partial\mathbf{p}}\dot{\mathbf{p}} = \frac{\partial\sigma}{\partial\mathbf{x}}f(\mathbf{x},\mathbf{u}) + \frac{\partial\sigma}{\partial\mathbf{p}}\dot{\mathbf{p}} = 0$$

we may be able to solve for $\mathbf{u}$ (if not, increase k)

Singular control example

Minimize

$$F = \frac{1}{2}\int_0^T x_1^2\,dt$$

subject to

$$\begin{aligned}\dot{x}_1 &= x_2 + u\\ \dot{x}_2 &= -u\end{aligned}$$

where $|u| \leq 1$ and T is unspecified.

The Hamiltonian is

$$H = -\frac{1}{2}x_1^2 + p_1(x_2+u) - p_2 u = -\frac{1}{2}x_1^2 + p_1 x_2 + (p_1 - p_2)u$$

which is linear in u, with switching function $\sigma = p_1 - p_2$.

Singular control example

Hamilton's equations

$$\frac{\partial H}{\partial p_i} = \frac{dx_i}{dt} \quad \text{and} \quad \frac{\partial H}{\partial x_i} = -\frac{dp_i}{dt}$$

Give the state equations and

$$\begin{aligned}\frac{\partial H}{\partial x_1} &= -x_1 = -\dot{p}_1\\ \frac{\partial H}{\partial x_2} &= p_1 = -\dot{p}_2\end{aligned}$$

The solution involves three cases

1. If the switching function $\sigma = p_1 - p_2 > 0$ then $u = 1$
2. If the switching function $\sigma = p_1 - p_2 < 0$ then $u = -1$
3. If the switching function $\sigma = p_1 - p_2 = 0$ then we have singular control

Singular control example

Case 1: $\sigma = p_1 - p_2 > 0$ and $u = 1$, so

$$\begin{aligned}\dot{x}_1 &= x_2 + 1\\ \dot{x}_2 &= -1\end{aligned}$$

which has solutions

$$\begin{aligned}x_1 &= -\frac{1}{2}t^2 + (c_1+1)t + c_2\\ x_2 &= -t + c_1\end{aligned}$$

so we can write

$$x_1 = -\frac{1}{2}x_2^2 - x_2 + c_4$$

where $c_4 = c_1(c_1+1) + c_2 - c_1^2/2$

Singular control example

Case 2: $\sigma = p_1 - p_2 < 0$ and $u = -1$, so

$$\begin{aligned}\dot{x}_1 &= x_2 - 1\\ \dot{x}_2 &= 1\end{aligned}$$

which has solutions

$$\begin{aligned}x_1 &= \frac{1}{2}t^2 + (c_1-1)t + c_2\\ x_2 &= t + c_1\end{aligned}$$

so we can write

$$x_1 = \frac{1}{2}x_2^2 - x_2 + c_3$$

where $c_3 = -c_1(c_1-1) + c_2 + c_1^2/2$

Singular control example

Case 3: singular as $\sigma = p_1 - p_2 = 0$

$$\begin{aligned} \sigma &= p_1 - p_2 \\ \dot{\sigma} &= \dot{p}_1 - \dot{p}_2 \\ &= x_1 + p_1 \\ &= 0 \end{aligned}$$

Using this, and the fact that $p_1 - p_2 = 0$ in the Hamiltonian $H = -\frac{1}{2}x_1^2 + p_1 x_2 + (p_1 - p_2)u$, we get

$$H = -\frac{1}{2}x_1^2 + p_1 x_2 + (p_1 - p_2)u = -\frac{1}{2}x_1^2 - x_1 x_2$$

For autonomous problems, with free end time $H = 0$, so

$$x_1(x_2 + x_1/2) = 0$$

and hence, either $x_1 = 0$ or $x_1 + 2x_2 = 0$

Singular control example

The two solutions present two surfaces:

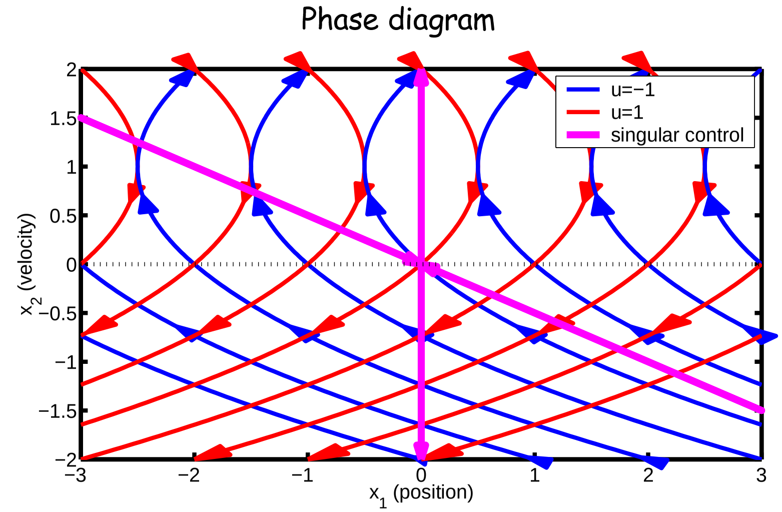
$$\begin{aligned} S_1 : &\quad x_1 = 0 \\ S_2 : &\quad x_1 + 2x_2 = 0 \end{aligned}$$

- on S_1 the derivative $\dot{x}_1 = 0$, and the state equation is $\dot{x}_1 = x_2 + u$, so $u = -x_2$.
- on S_2 the derivative $\dot{x}_2 = -\dot{x}_1/2$, and the state equations

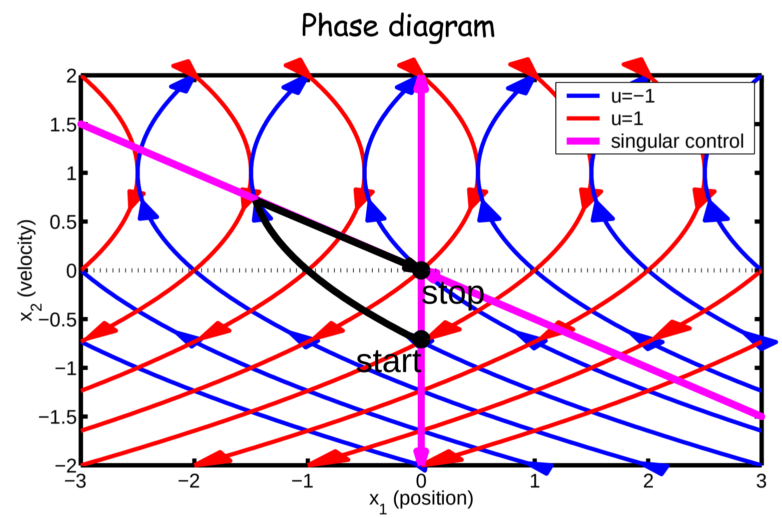
$$\begin{aligned} \dot{x}_1 &= x_2 + u \\ \dot{x}_2 &= -u \end{aligned}$$

lead to $u = x_2$

Singular control example



Singular control example



Variational Methods & Optimal Control

lecture 28

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 28

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Feedback control systems

In all of our previous examples, we solve optimization problem "all at once", i.e., we plan the shape of the curve y to optimize the functional. However, sometimes, we need a control that reacts continuously to perturbations in a system. Such controllers typically utilize feedback.

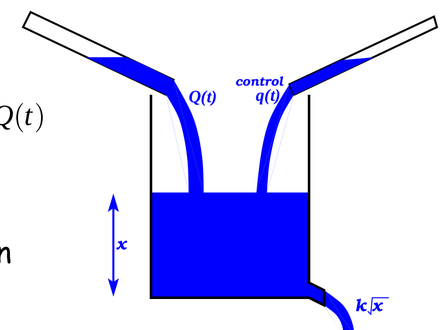
Feedback control systems

- control problem until now have been planned.
  - know the problem before hand
  - assume state is perfectly observable
  - plan the control from the start
- alternative: feedback control
  - observe the state at time t
  - make decisions on the best control at time t
  - continually update this decision

Feedback control example

Liquid level control system

- Tank of uniform cross-sectional area
- Fluid pumped in at rate $Q(t)$
- Fluid level $x(t)$
- Flow out $k\sqrt{x}$
- Given constant flow, then steady-state will be $x_s = (Q/k)^2$
- when there are fluctuations, we wish to bring the tank back to steady state, by addition of suitable input at rate $q(t)$.



Feedback control example

Liquid level optimal control problem

- we wish to operate near steady state, so part of cost is the square deviation

$$\int_0^T [x(t) - x_s]^2 \, dt$$

- Also want to minimize control expenditure, e.g.

$$\int_0^T q(t)^2 \, dt$$

- Problem is to minimize a linear combination of these two, e.g.

$$\int_0^T \left[(x(t) - x_s)^2 + \alpha^2 q(t)^2 \right] \, dt$$

Feedback control example

The dynamics of the system are

$$\dot{x} = Q + q - k\sqrt{x}$$

We can linearize the problem as follows:

- note we wish to maintain x near x_s.
- can approximate x near x_s by a Taylor series

$$\sqrt{x} = \sqrt{x_s} + \frac{1}{2\sqrt{x_s}}(x - x_s) + O\left((x - x_s)^2\right)$$

- the steady state will be when $\dot{x} = 0$, so $Q = k\sqrt{x_s}$
- change variables to $y = x - x_s$, so given $A = k/2\sqrt{x_s}$

$$\dot{y} = -Ay + q(t)$$

Feedback control example

So the optimal control problem can be written as minimize

$$F\{y\} = \int_0^T \left[y(t)^2 + \alpha^2 q(t)^2 \right] \, dt$$

subject to $\dot{y} = -Ay + q(t)$, and $y(0) = x(0) - x_s$ and $y(T) = 0$.

Substitute $q(t) = Ay + \dot{y}$ into the integral and we get

$$F\{y\} = \int_0^T \left[y(t)^2 + \alpha^2 (Ay + \dot{y})^2 \right] \, dt$$

so the E-L equations are

$$\frac{d}{dt}\frac{\partial f}{\partial \dot{y}} - \frac{\partial f}{\partial y} = 2\alpha^2 \frac{d}{dt}\left[\dot{y} + Ay\right] - 2y - 2\alpha^2\left[A\dot{y} + A^2 y\right] = 0$$

Feedback control example

$$\begin{aligned}
2\alpha^2 \frac{d}{dt}\left[\dot{y} + Ay\right] - 2y - 2\alpha^2\left[A\dot{y} + A^2 y\right] &= 0 \\
\alpha^2 \frac{d}{dt}\left[\dot{y} + Ay\right] - y - \alpha^2\left[A\dot{y} + A^2 y\right] &= 0 \\
\alpha^2\left[\ddot{y} + A\dot{y} - A\dot{y} - A^2 y\right] - y &= 0 \\
\ddot{y} - \left[\frac{1 + \alpha^2 A^2}{\alpha^2}\right] y &= 0
\end{aligned}$$

This has solution

$$y(t) = Ce^{\lambda t} + Be^{-\lambda t}$$

where $\lambda = \sqrt{\frac{1+\alpha^2 A^2}{\alpha^2}}$.

Feedback control example

The solution is

$$y(t) = Ce^{\lambda t} + Be^{-\lambda t}$$

Given the end-point constraint that $y(0) = y_0$, we get conditions $C + B = y_0$, so the solution is in the form

$$y(t) = y_0 \left(ae^{\lambda t} + (1-a)e^{-\lambda t} \right)$$

Take the case where $T \to \infty$, and we wish $y(T) = 0$, i.e. the fluctuation should go to zero in the far future, then, we require $a = 0$, and the solution will be

$$y(t) = y_0 e^{-\lambda t}$$

where $\lambda = \sqrt{\frac{1+\alpha^2 A^2}{\alpha^2}}$. We get $q(t)$ from $q(t) = Ay + \dot{y}$ which gives

$$q(t) = (A - \lambda) y_0 e^{-\lambda t} = (A - \lambda) y(t)$$

Feedback control example

- the above gives the optimal control policy $q(t)$ over the whole interval $[0, T]$
- actually, a feedback control problem would be more convenient
  - given the state $y(t)$, what should the control be
  - write the control as a function of y, e.g. $q(y)$
- example: **proportional control**

$$q(y) = My$$

where M is called the **gain**

- we choose proportional control because in previous planned solution $q(t) = (A + \lambda) y_0 e^{-\lambda t}$ which is proportional to the perturbation $y(t)$

Feedback control example

Given a proportional controller, it is easy to rewrite the dynamic equation

$$\dot{y} = -Ay + q(t)$$

using $q(y) = My$ as

$$\dot{y} + (A - M)y = 0$$

which has solution

$$y(t) = y_0 e^{-(A-M)t}$$

We wish to choose M so that it minimizes

$$F\{y\} = \int_0^T \left[y(t)^2 + \alpha^2 q(t)^2 \right] dt$$

Feedback control example

$$\begin{aligned}
F\{y\} &= \int_0^T \left[y(t)^2 + \alpha^2 q(t)^2 \right] dt \\
&= \int_0^T \left[y(t)^2 + \alpha^2 M^2 y(t)^2 \right] dt \\
&= \left(1 + \alpha^2 M^2\right) \int_0^T y(t)^2 \, dt \\
&= y_0 \left(1 + \alpha^2 M^2\right) \int_0^T e^{-2(A-M)t} \, dt \\
&= y_0 \frac{\left(1 + \alpha^2 M^2\right)}{-2(A-M)} \left[e^{-2(A-M)t} \right]_0^T
\end{aligned}$$

Feedback control example

Take the case $T \to \infty$ and we get

$$F\{y\} = y_0 \frac{\left(1+\alpha^2 M^2\right)}{2(A-M)}$$

In order to find a maximum we differentiate WRT M, to get

$$\begin{aligned} \frac{d}{dM} F\{y\} &= y_0 \frac{\alpha^2 2M 2(A-M) + 2\left(1+\alpha^2 M^2\right)}{4(A-M)^2} \\ &= y_0 \frac{2\alpha^2 M(A-M) + \left(1+\alpha^2 M^2\right)}{2(A-M)^2} \end{aligned}$$

To get a minimum, the numerator must be zero, so we set

$$2\alpha^2 M(A-M) + \left(1+\alpha^2 M^2\right) = 0$$

Feedback control example

$$\begin{aligned} 2\alpha^2 M(A-M) + \left(1+\alpha^2 M^2\right) &= 0 \\ -\alpha^2 M^2 + 2\alpha^2 AM + 1 &= 0 \\ M^2 - 2AM - 1/\alpha^2 &= 0 \end{aligned}$$

This is just a quadratic equation in M, which we solve to get

$$M = \frac{2A \pm \sqrt{4A^2 + 4/\alpha^2}}{2} = A \pm \sqrt{A^2 + 1/\alpha^2}$$

If we take the solution with the '+' sign, then $F\{y\} = y_0 \frac{\left(1+\alpha^2 M^2\right)}{2(A-M)}$ will be negative (the denominator is negative), so we are restricted to

$$M = A - \sqrt{A^2 + 1/\alpha^2}$$

Feedback control example

- $M < 0$, which makes sense, because the control has to reverse fluctuations.
- Compare the feedback and planned solutions

$$\begin{aligned} q(t) &= (A-\lambda) y_0 e^{-\lambda t} \\ q(t) &= My(t) = \left(A - \sqrt{A^2 + 1/\alpha^2}\right) y_0 e^{-(A-M)t} \end{aligned}$$

  where $\lambda = \sqrt{\frac{1+\alpha^2 A^2}{\alpha^2}}$, and notice they are the same.
- for α large, e.g. high control cost, the solution has $1/\alpha^2 \to 0$, and so $M \to 0$.
- for α small, e.g. low control cost, the fluctuation cost dominates, and $M \to -1/\alpha$

Variational Methods & Optimal Control

lecture 29

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 29

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

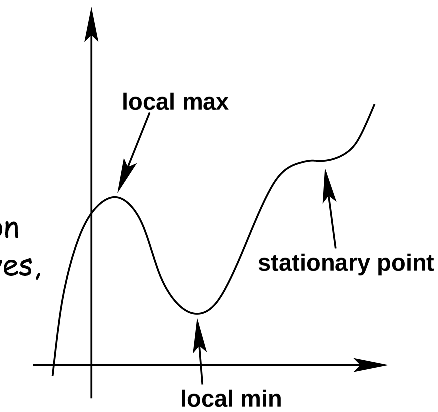
Classification of extrema

We have so far typically ignored the issue of classification of extrema, but remember that for simple stationary points we need to look at higher derivatives to see if a stationary point is a maximum, minimum or point of inflection. We need an analogous process for extremal curves as well.

Classification of extrema

Local extrema have $f'(x) = 0$

- $f''(x) > 0$ local minima
- $f''(x) < 0$ local maxima
- $f''(x) = 0$ it might be a stationary point of inflection, depending on higher order derivatives, e.g. x^4.



E-L solutions

- the E-L equations are a necessary condition
- the E-L equations are not sufficient
- along the extremal curve, the functional might have
  - ▷ a min, max, or stationary point
  - ▷ it might be global or local
- we really need to classify extremals
  - ▷ until now we have
    - ⋆ just assumed it was the minima
    - ⋆ used physical insight to understand the solution
    - ⋆ tested it by inspection
  - ▷ we could also compare to alternative curves

Examples

- **Physical intuition:** Brachystochrone (or geodesic): we look for a minimum time path. So we can see that **physically** there can't be a maximum.
- **Examine the solution:** e.g. consider the functional

$$F\{y\} = \int_0^1 y'^2 \, dx$$

conditioned on $y(0) = y(1) = 0$.

The E-L equations give straight line solutions, e.g. $y = c_1 x + c_2$, and the boundary conditions imply $c_1 = c_2 = 0$, so $y' = 0$. Clearly then $F\{y\} = 0$, which is the minimum possible value, for an integral of a non-negative function like y'^2.

Examples

- **Compare with alternative curves:** For the functional

$$F\{y\} = \int_0^1 (xy' + y'^2) \, dx$$

conditioned on $y(0) = 0$ and $y(1) = 1$.

The E-L equations give

$$y = -\frac{1}{4}x^2 + c_1 x + c2$$

and the boundary conditions give $c_1 = 5/4$, $c_2 = 0$, so the solution is

$$y = \frac{5}{4}x - \frac{1}{4}x^2$$

Examples

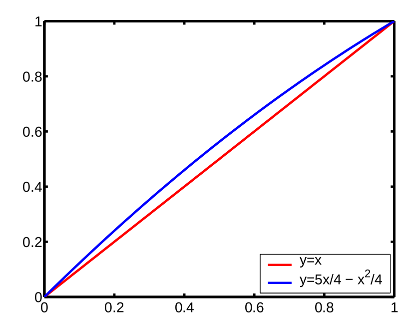
For $y = \frac{5}{4}x - \frac{1}{4}x^2$, we have $y' = \frac{5}{4} - \frac{1}{2}x$, so the function is

$$\begin{aligned} F\{y\} &= \int_0^1 \left[x\left(\frac{5}{4} - \frac{1}{2}x\right) + \left(\frac{5}{4} - \frac{1}{2}x\right)^2 \right] dx \\ &= \int_0^1 \left[\frac{25}{16} - \frac{1}{4}x^2 \right] dx \\ &= \left[\frac{25}{16}x - \frac{1}{12}x^3 \right]_0^1 \\ &= \frac{25}{16} - \frac{1}{12} \\ &= \frac{71}{48} \end{aligned}$$

Examples

For the curve $y(x) = x$, the $y' = 1$, so the functional is

$$\begin{aligned} F\{y\} &= \int_0^1 (x+1) \, dx \\ &= \left[x^2/2 + x\right]_0^1 \\ &= 3/2 \end{aligned}$$



Now $\frac{3}{2} > \frac{71}{48}$, so we should be looking at a local min.

But this isn't very formal, or rigorous!

Classification of extrema

- Above methods either
  - Aren't very formal or rigorous
  - Aren't easy to generalize
- Need to develop a means of formal classification
- The secret is by analogy to classification for functions of several variables
  - We need to look at second derivatives
  - Positive definiteness of the Hessian
- The analogy to second derivatives is called the **second variation**

Classification of extrema

Classification of extrema of functions (see Lecture 2)

Use Taylor's theorem in N-D

$$f(\mathbf{x}+\delta\mathbf{x}) = f(\mathbf{x}) + \delta\mathbf{x}^T \nabla f(\mathbf{x}) + \frac{1}{2}\delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} + O(\delta\mathbf{x}^3)$$

Where $H(\mathbf{x})$ is the Hessian matrix

$$H(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Maxima of N variables

If a smooth function $f(\mathbf{x})$ has a local extrema at $\mathbf{x}$ then $\nabla f(\mathbf{x}) = 0$, and so we can rewrite Taylor's theorem for small $\delta\mathbf{x}$ as

$$f(\mathbf{x}+\delta\mathbf{x}) - f(\mathbf{x}) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x}/2$$

A sufficient condition for the extrema $\mathbf{x}$ to be a local minimum is for the quadratic form

$$Q(\delta x_1, \ldots, \delta x_n) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\delta x_i \delta x_j$$

to be strictly positive definite.

Quadratic forms

A quadratic form

$$Q(\mathbf{x}) = \sum_{i,j} a_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x}$$

is said to be positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$.

A quadratic form is positive definite iff every eigenvalue of A is greater than zero.

A quadratic form is positive definite if all the principal minors in the top-left corner of A are positive, in other words

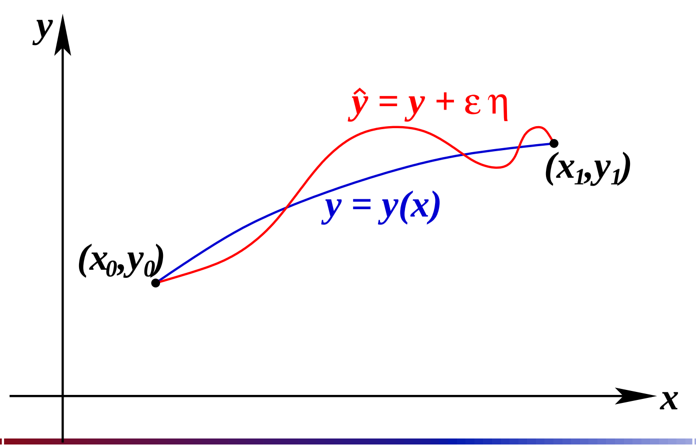
$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \cdots$$

Notes on maxima and minima

- maxima of $f(x)$ are minima of $-f(x)$.
- we need to generalize this to functionals
- we do this using the second variation
- note that even so, we only classify local min and max, the global min or max may occur at the boundary, or at one of several extrema.

The second variation

Once again consider the fixed end-point problem, with small perturbations about the extremal curve.



The second variation

Given the perturbation $\hat{y} = y + \varepsilon\eta$ we use Taylor's Theorem as when deriving the first variation, but this time we expand the $O(\varepsilon^2)$ terms as well, e.g.

$$
\begin{aligned}
f(x,\hat{y},\hat{y}') &= f(x,y,y') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right] \\
&\quad + \frac{\varepsilon^2}{2}\left[\eta^2\frac{\partial^2 f}{\partial y^2} + 2\eta\eta'\frac{\partial^2 f}{\partial y'\partial y} + \eta'^2\frac{\partial^2 f}{\partial y'^2}\right] + O(\varepsilon^3)
\end{aligned}
$$

$$
\begin{aligned}
F\{\hat{y}\} - F\{y\} &= \varepsilon\delta F(\eta,y) \\
&\quad + \frac{\varepsilon^2}{2}\int_{x_0}^{x_1}\left[\eta^2\frac{\partial^2 f}{\partial y^2} + 2\eta\eta'\frac{\partial^2 f}{\partial y\partial y'} + \eta'^2\frac{\partial^2 f}{\partial y'^2}\right]dx + O(\varepsilon^3)
\end{aligned}
$$

The second variation

$$
\begin{aligned}
F\{\hat{y}\} - F\{y\} &= \varepsilon\delta F(\eta,y) \\
&\quad + \frac{\varepsilon^2}{2}\int_{x_0}^{x_1}\left[\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'}\right]dx + O(\varepsilon^3) \\
&= \varepsilon\delta F(\eta,y) + \frac{\varepsilon^2}{2}\delta^2 F(\eta,y) + O(\varepsilon^3)
\end{aligned}
$$

Where we define the **Second Variation** by

$$
\delta^2 F(\eta,y) = \int_{x_0}^{x_1}\left[\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'}\right]dx
$$

Note for a stationary curve, we require $\delta F = 0$, so the behavior of $F\{\hat{y}\} - F\{y\}$ is captured in $\delta^2 F(\eta,y)$.

The second variation

Note that

$$2\eta\eta' = \frac{d}{dx}(\eta^2)$$

So we can write

$$\begin{aligned}\int_{x_0}^{x_1} 2\eta\eta' f_{yy'}\,dx &= \int_{x_0}^{x_1} \frac{d(\eta^2)}{dx} f_{yy'}\,dx \\ &= [\eta^2 f_{yy'}]_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta^2 \frac{d f_{yy'}}{dx}\,dx \\ &= -\int_{x_0}^{x_1} \eta^2 \frac{d f_{yy'}}{dx}\,dx\end{aligned}$$

using integration by parts and the fact that $\eta(x_0) = \eta(x_1) = 0$.

The second variation

So we can write the second variation as

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) + \eta'^2 f_{y'y'}\,dx$$

This form has the advantage that

- $\eta^2 \geq 0$
- $\eta'^2 \geq 0$
- after solving E-L equations we know f and its derivatives

Classifying extrema

For an extremal curve y to be a local minima, we require

$$\delta^2 F(\eta, y) \geq 0$$

for all valid perturbation curves η. Likewise we get a maxima if $\delta^2 F(\eta, y) \leq 0$ for all η and a stationary curve if the second variation changes sign.

- Note that we have already solved the E-L equations, and so we know y. Hence we can calculate f_{yy}, $f_{yy'}$, and $f_{y'y'}$ explicitly.
- we still need to ensure $\delta^2 F(\eta, y) \geq 0$ for all possible η.

Legendre condition

The **Legendre condition** is a **necessary** condition for a local minima.

If y is a local minima of a functional $F\{y\} = \int f(x, y, y')\,dx$, then

$$f_{y'y'}(x, y, y') \geq 0$$

along the extremal curve y.

Legendre condition

Sketch proof: Remember that f and y are known functions (now), so we know f_{yy}, $f_{yy'}$ and $f_{y'y'}$, explicitly as functions of x, and hence we can write the second variation

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 B(x) + \eta'^2 A(x)\, dx$$

where

$$\begin{aligned} A(x) &= f_{y'y'} \\ B(x) &= \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) \end{aligned}$$

Legendre condition

Sketch proof: The proof relies on the fact that we can find functions η such that $|\eta|$ is small, but $|\eta'|$ is large.

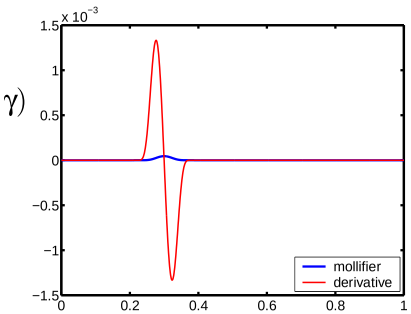
Note we cannot do the opposite, because $|\eta'|$ small, implies that η is smooth, which given the end conditions implies that $|\eta|$ will be small.

Example: **mollifier**

$$\eta(x) = \begin{cases} \exp\left(-\frac{\gamma}{\gamma^2 - (x-c)^2} \right), & \text{if } x \in [c-\gamma, c+\gamma] \\ 0, & \text{otherwise} \end{cases}$$

Mollifier

$$\eta(x) = \begin{cases} \exp\left(-\frac{\gamma}{\gamma^2 - (x-c)^2} \right), & \text{if } x \in (c-\gamma, c+\gamma) \\ 0, & \end{cases}$$



$$\eta'(x) = \begin{cases} -\frac{2\gamma(x-c)}{(\gamma^2 - (x-c)^2)^2} \exp\left(-\frac{\gamma}{\gamma^2 - (x-c)^2} \right), & \text{if } x \in (c-\gamma, c+\gamma) \\ 0, & \text{otherwise} \end{cases}$$

Ratio of derivative to function is larger for smaller γ.

Legendre condition

Sketch proof: Given $|\eta|$ small, we can essentially ignore the η^2 terms, and we get only the term

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta'^2 A(x)\, dx$$

If A changes sign, then we could choose η to be a mollifier such that it is localized in the part where A is positive, and a mollifier such that it is localized in the part of A which is negative. The two would produce integrals with different signs, and so we would get a change of sign of $\delta^2 F(\eta, y)$, which is what we are trying to avoid.

Example

Find the minimum of

$$F\{y\} = \int_0^1 \left(xy' + y'^2\right)\, dx$$

conditioned on $y(0) = 0$ and $y(1) = 1$.

The solution is

$$y = \frac{5}{4}x - \frac{1}{4}x^2$$

Then (from earlier)

$$\begin{aligned} f(x,y,y') &= xy' + y'^2 \\ &= \frac{25}{16} - \frac{1}{4}x^2 \end{aligned}$$

Example

$$\begin{aligned} f(x,y,y') &= xy' + y'^2 \\ f_{y'} &= x + 2y' \\ f_{y'y'} &= 2 \\ &> 0 \end{aligned}$$

Hence Legendre's condition is satisfied, so this **could** be a local minimum.

Sufficient conditions

- various approaches to sufficient conditions
- problem is that we have to get away from point-wise conditions
  - like the Legendre condition
  - point-wise conditions couldn't classify which of two possible arcs of a great circle is the shortest path between two points on a sphere.
- a sufficient condition is the Jacobi condition, but there are others (van Brunt, 10.4, or Cragg's p.37, or Bliss, p.37)
- still mostly only conditions for local minima, so need to do more work

All is not lost

Example: Find the minimum of

$$F\{y\} = \int_0^1 \left(xy' + y'^2\right)\, dx$$

So

$$\begin{aligned} f_{y'y'} &= 2 \\ f_{y'y'} &= 0 \\ f_{y'y'} &= 0 \end{aligned}$$

So the second variation

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) + \eta'^2 f_{y'y'}\, dx = 2\int_{x_0}^{x_1} \eta'^2\, dx \geq 0$$

for all η so we have a local minimum!

Variational Methods & Optimal Control

lecture 30

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Variational Methods & Optimal Control

lecture 30

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

July 26, 2012

Revision.

Fixed end points: lecture 4

Let $F : C^2[x_0, x_1] \to I\!R$ be a functional of the form

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

where f has continuous partial derivatives of second order with respect to x, y, and y', and $x_0 < x_1$. Let

$$S = \left\{ y \in C^2[x_0,x_1] \,\middle|\, y(x_0) = y_0 \text{ and } y(x_1) = y_1 \right\},$$

where y_0 and y_1 are real numbers. If $y \in S$ is an extremal for F, then for all $x \in [x_0, x_1]$

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0$$ $\Leftarrow$ the Euler-Lagrange equation

Special Cases: lectures 4-8

- f depends only on y'
  - e.g., geodesics in the plane
  - always results in straight lines
- f has no explicit dependence on x (autonomous case)
  - e.g., the catenary, brachystochrone, Newton's nosecone
  - use the Hamiltonian (sometimes)
- f has no explicit dependence on y
  - e.g., the geodesic on the sphere
  - $\partial f / \partial y' = const$
- $f = A(x,y)y' + B(x,y)$ (degenerate case)
  - E-L results in identity

Invariance: lecture 8

Extensions: lecture 9-11

Points to remember:

$$F\{y, z\} = \int f(z, y, z, y', z', z'')\, dx$$

dependent variables
$y(x)$ and $z(x)$
use multiple E-L equations

independent variable x
if there is more than one, the E-L equation is a partial DE

higher order derivative z''
use Euler-Poisson equation

And we can combine each of the above if more than one case applies.

Numerical methods: lecture 12-13

Constraints: lecture 14-16, 21

- Integral constraints of the form

$$\int g(x, y, y')dx = const$$

  e.g., the Isoperimetric problem.
  - use Lagrange multiplier constant λ
- Holonomic constraints, e.g., $g(x, y) = 0$
  - use Lagrange multiplier function $\lambda(x)$
- Non-holonomic constraints, e.g., $g(x, y, y') = 0$
  - use Lagrange multiplier function $\lambda(x)$
- Inequality constraints, e.g., $y(x) \geq g(x)$
  - either E-L equations, or constraint $y(x) = g(x)$
  - take care at corners, but often $y' = g'$

Free end points: lecture 17-19

- free at both end points

$$\left[p\delta y - H\delta x\right]_{x_0}^{x_1} = 0 \text{ where } p = \frac{\partial f}{\partial y'} \text{ and } H = y'\frac{\partial f}{\partial y'} - f$$

- separable end points: $p\delta y - H\delta x\Big|_{x_i} = 0$
- fixed x, free y, so $\delta x \neq 0$ and $\delta y = 0$ so $H|_{x_i} = 0$
- fixed y, free x, so $\delta x = 0$ and $\delta y \neq 0$ so $p|_{x_i} = 0$
- terminal cost $\phi(t_1, x_1(t_1))$, free $(t_1, x_1(t_1))$

$$\left[\left(\frac{\partial \phi}{\partial x} + p\right)\delta x + \left(\frac{\partial \phi}{\partial t} - H\right)\delta t\right]_{(t_1,x_1)} = 0$$

Free end points: lecture 17-19

- higher order derivatives: $f(x,y,y',y'')$

$$\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\bigg|_{x_0} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''}\bigg|_{x_1} = 0$$

$$\frac{\partial f}{\partial y''}\bigg|_{x_0} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y''}\bigg|_{x_1} = 0$$

  - ▷ first set replace $y(x_i) = y_i$ fixed
    - ⋆ e.g., a supported beam
  - ▷ second set replace $y(x_i) = y'_i$ fixed
    - ⋆ e.g., a clamped beam

Free end points: lecture 17-19

- multiple dependent variables

$$\sum_{k=1}^{n} p_k \delta q_k - H\delta t = 0 \text{ where } p_k = \frac{\partial L}{\partial \dot{q}_k} \text{ and } H = \sum_{k=1}^{n} \dot{q}_k p_k - L$$

- transversals: end points on curve (x_Γ, y_Γ)

$$\left(\frac{dx_\Gamma}{d\xi}, \frac{dy_\Gamma}{d\xi}\right) \cdot (-H, p) = p\frac{dy_\Gamma}{d\xi} - H\frac{dx_\Gamma}{d\xi} = 0$$

- special case

$$F\{y\} = \int_0^{x_1} K(x,y)\sqrt{1+y'^2}\,dx$$

transversal condition means extremal joins Γ at right angles

Corners: lecture 20

- solve E-L equations
- look for solutions for each end condition
- match up the solutions at a corner x^* so that
  - ▷ y is continuous
  - ▷ the Weierstrass-Erdman corner conditions hold

so

$$\begin{aligned} y\big|_{x^{*-}} &= y\big|_{x^{*+}} \\ p\Big|_{x^{*-}} &= p\Big|_{x^{*+}} \\ H\Big|_{x^{*-}} &= H\Big|_{x^{*+}} \end{aligned}$$

at any 'corner'

Tricks for solving problems

- Exploiting special properties: see special cases
- Hamilton's equations (Canonical Euler-Lagrange equations)
- Hamilton-Jacobi equations
- PMP

Optimal Control: lecture 17, 21-23, 26

Conservation laws: lecture 25

Classification: lecture 29