

Variational Methods & Optimal Control

lecture 03

Matthew Roughan

<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

April 14, 2016

Variational Methods & Optimal Control: lecture 03 – p.1/36

Revision, part iii

Constrained extrema and Lagrange multipliers.

Variational Methods & Optimal Control: lecture 03 – p.2/36

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)

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

where λ_i are the undetermined Lagrange multipliers.

Variational Methods & Optimal Control: lecture 03 – p.3/36

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

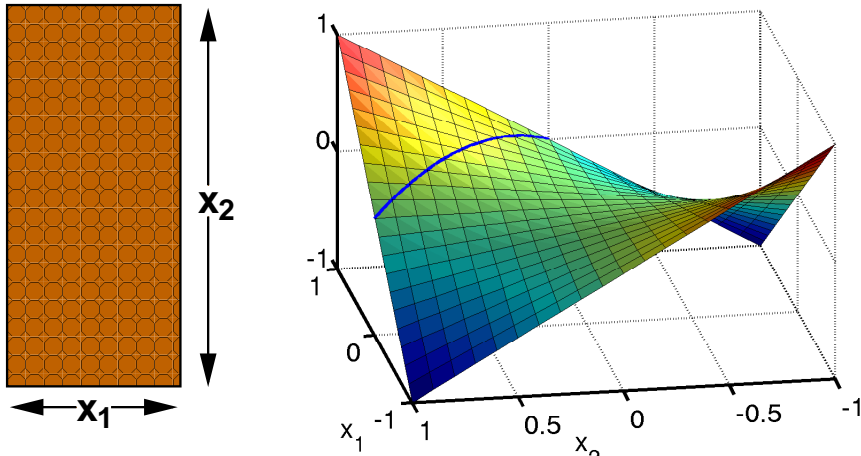
$$\delta \mathbf{x}^T \nabla f = 0$$

Variational Methods & Optimal Control: lecture 03 – p.4/36

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 $x_1 + x_2 = 1$, $x_1, x_2 > 0$.



Variational Methods & Optimal Control: lecture 03 – p.5/36

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.

Variational Methods & Optimal Control: lecture 03 – p.7/36

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

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

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

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

$$\frac{\partial h}{\partial \lambda} = x_1 + x_2 - 1 = 0$$

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

Variational Methods & Optimal Control: lecture 03 – p.6/36

Constrained maxima example 2

Largest area rectangle inscribed in a circle diameter 1.

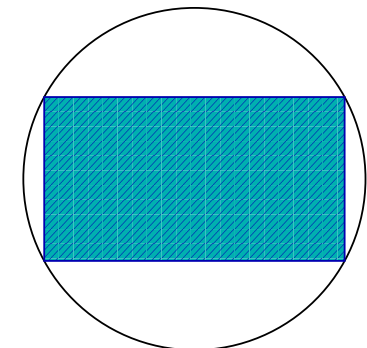
Maximize $f(x_1, x_2) = x_1 x_2$ subject to $x_1^2 + x_2^2 = 1$, $x_1, x_2 > 0$.

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

$$\frac{\partial h}{\partial x_1} = x_2 + 2\lambda x_1$$

$$\frac{\partial h}{\partial x_2} = x_1 + 2\lambda x_2$$

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



Variational Methods & Optimal Control: lecture 03 – p.8/36

Constrained maxima ex. 2, solution

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

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

So $\lambda = \pm 1/2$. 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 ex. 3, solution

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

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

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 3

Largest area rectangle inscribed in an ellipse.

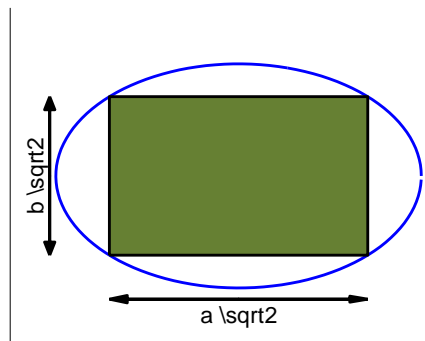
Maximize $f(x, y) = xy$ subject to $x^2/a^2 + y^2/b^2 = 1$, $x, y > 0$.

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

$$\frac{\partial h}{\partial x} = y + 2\lambda x/a^2$$

$$\frac{\partial h}{\partial y} = x + 2\lambda y/b^2$$

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



Constrained maxima example 4

Maximize $f(x_1, x_2, x_3) = x_1 x_2 x_3$ subject to $x_1 x_2 + x_1 x_3 + x_2 x_3 = 1$, and

$x_1 + x_2 + x_3 = 3$

$h = x_1 x_2 x_3 + \lambda(x_1 x_2 + x_1 x_3 + x_2 x_3 - 1) + \mu(x_1 + x_2 + x_3 - 3)$

$$\frac{\partial h}{\partial x_1} = x_2 x_3 + \lambda(x_2 + x_3) + \mu = 0$$

$$\frac{\partial h}{\partial x_2} = x_1 x_3 + \lambda(x_1 + x_3) + \mu = 0$$

$$\frac{\partial h}{\partial x_3} = x_1 x_2 + \lambda(x_1 + x_2) + \mu = 0$$

$$\frac{\partial h}{\partial \lambda} = x_1 x_2 + x_1 x_3 + x_2 x_3 - 1 = 0$$

$$\frac{\partial h}{\partial \mu} = x_1 + x_2 + x_3 - 3 = 0$$

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

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

$$\frac{\partial h}{\partial x} = 3 - \lambda = 0 \Rightarrow \lambda = 3$$

$$\frac{\partial h}{\partial \alpha} = -2\lambda\alpha = 0 \Rightarrow \alpha = 0$$

$$\frac{\partial h}{\partial \lambda} = 10 - x - \alpha^2 = 0 \Rightarrow x = 10$$

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, \dots , along with two operators (addition, and scalar multiplication) that is

- ▶ closed under addition, e.g.

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

- ▶ closed under scalar multiplication, e.g.

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

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 $C[x_0, x_1] = \{f : [x_0, x_1] \rightarrow \mathbb{R} \mid f \text{ is continuous}\}$, with addition and scalar multiplication defined by

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

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 $f : \mathbb{R} \rightarrow \mathbb{R}$ 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

$$\|x\| \geq 0$$

$$\|x\| = 0 \text{ iff } x = 0$$

$$\|\alpha x\| = |\alpha| \|x\|$$

$$\|x + y\| \leq \|x\| + \|y\| \text{ (the triangle inequality)}$$

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 $x \in S$

$$\alpha\|x\|_a \leq \|x\|_b \leq \beta\|x\|_a$$

- ▶ 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 $(S, \|\cdot\|)$

$$B(f, \varepsilon, \|\cdot\|) = \{g \in S \mid \|f - g\| \leq \varepsilon\}$$

Inner products

An **inner product** is a function $\langle \cdot, \cdot \rangle : S \times S \rightarrow \mathbb{R}$, 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 $\alpha \in \mathbb{R}$.

$$\langle f, f \rangle \geq 0$$

$$\langle f, f \rangle = 0 \text{ iff } f = 0$$

$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$$

$$\langle f, g \rangle = \langle g, f \rangle$$

$$\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$$

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. $F : S \rightarrow \mathbb{R}$.
- ▶ Typically in CoV S is a space of functions, e.g. $y(x)$
- ▶ Example Functionals

$$F\{y(x)\} = |y(0)|$$

$$F\{y(x)\} = \max_x \{y(x)\}$$

$$F\{y(x)\} = \left. \frac{dy}{dx} \right|_{x=1}$$

$$F\{y(x)\} = y(0) + y(1)$$

$$F\{y(x)\} = \sum_{n=0}^N a_n y(n)$$

Integral functionals

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

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

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

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

Variational Methods & Optimal Control: lecture 03 – p.29/36

Example: the Brachistochrone

The time taken is

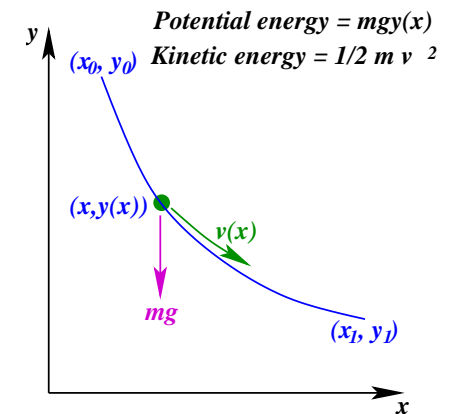
$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$

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

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



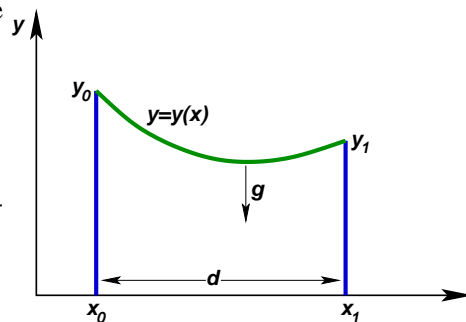
Variational Methods & Optimal Control: lecture 03 – p.31/36

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



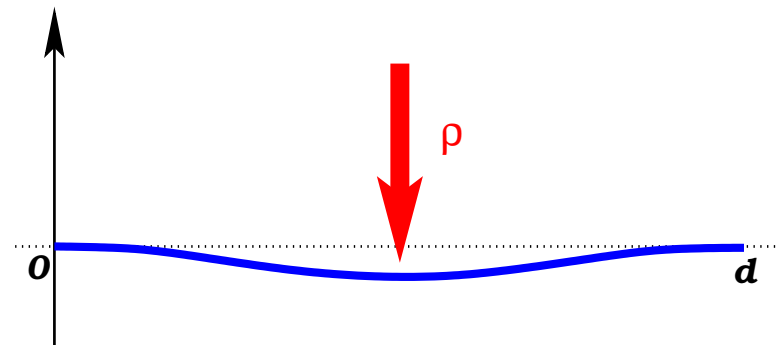
- m = mass per unit length
- g = gravitational constant

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

Variational Methods & Optimal Control: lecture 03 – p.30/36

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 $\rho(x)$.

Variational Methods & Optimal Control: lecture 03 – p.32/36

Example: a bent beam

Let $y : [0, d] \rightarrow \mathbb{R}$ describe the shape of the beam, and $\rho : [0, d] \rightarrow \mathbb{R}$ be the load per unit length on the beam.

For a bent elastic beam the potential energy from elastic forces is

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

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

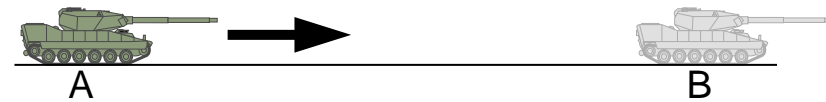
$$F\{u\} = \int_0^1 \frac{1}{2} u^2 dt$$

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.