Optimal Functions and Nanomechanics III

APP MTH 3022/7106

Barry Cox

Lecture 3

Last lecture

- Looked at extrema of functions of n variables
- Looked at some special cases for 2D
- Revised the chain rule for n variables
- Looked at Taylor's theorem in nD
- Refreshed quadratic forms and Hessian matrices
- Defined functionals and discussed integral functionals
- Tried to determine a crude brachistochrone using a one parameter family of curves

Constrained extrema

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

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

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

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

where λ_i are the undetermined Lagrange multipliers.

Why Lagrange multipliers?

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

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

So $\partial h/\partial x_i = 0$ implies that $\partial f/\partial x_i = -\lambda \partial g/\partial x_i$

Assume $\boldsymbol{x}$ is an extremal, which satisfies the constraint, consider all of the $\boldsymbol{x} + \delta\boldsymbol{x}$ in the neighborhood of $\boldsymbol{x}$ that also satisfy the constraint (i.e. $g(\boldsymbol{x} + \delta\boldsymbol{x}) = g(\boldsymbol{x}) = 0$), we also know from Taylor's theorem that

$$g(\boldsymbol{x} + \delta\boldsymbol{x}) = g(\boldsymbol{x}) + \delta\boldsymbol{x} \cdot \boldsymbol{\nabla} g + O(\delta\boldsymbol{x}^2)$$

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

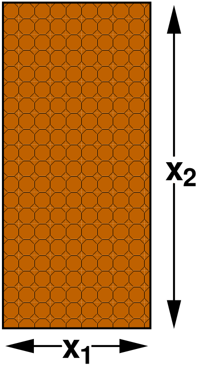
$$\delta\boldsymbol{x} \cdot \boldsymbol{\nabla} g = 0$$

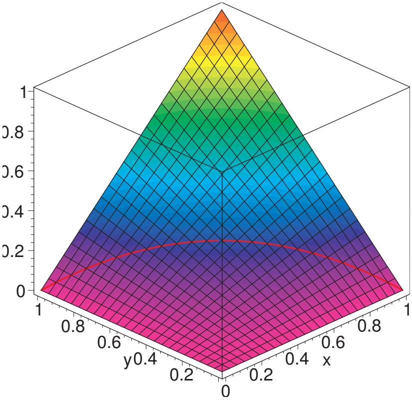
If we take $\partial f/\partial x_i = -\lambda \partial g/\partial x_i$ then

$$\delta\boldsymbol{x} \cdot \boldsymbol{\nabla} f = 0$$

Constrained: 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: Example 1 soln

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: Example 1 classification

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 $(\delta x, -\delta x)^T$.

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

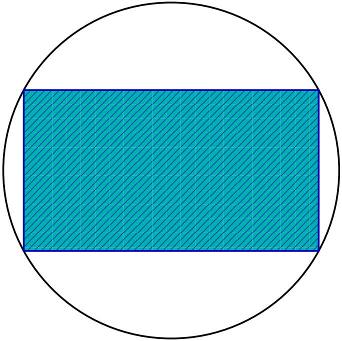
Hence, given the constraints on $\delta \boldsymbol{x}$, for all **possible** $\delta \boldsymbol{x}$, $f(\boldsymbol{x} + \delta \boldsymbol{x}) < f(\boldsymbol{x})$, and we have a local maximum.

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: Example 2, soln

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

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

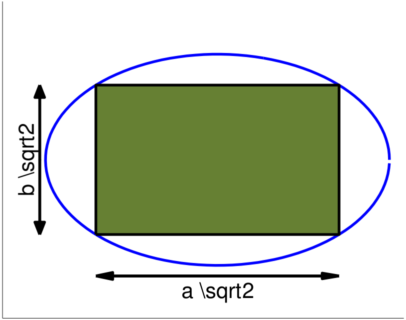
Solution is a square.

Constrained maxima: 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: Example 3, soln

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

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

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 $f(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^n$ subject to the constraints

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

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

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

where λ_i are the undetermined Lagrange multipliers.

Inequality constraints

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

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 $x^2 + y^2 \leqslant 1$.

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 $g_i(x) \geqslant 0$. 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 $\boldsymbol{x}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\lambda}$.

Slack Variables: Example

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

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

Now add a standard Lagrange multiplier, to maximize

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

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

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

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.

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

- 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 $1X = X$

Examples

- **Example 1:** the set of vectors $\boldsymbol{x} \in \mathbb{R}^n$, 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

$$\begin{aligned}
&\|x\| \geqslant 0 \\
&\|x\| = 0, \text{ iff } x = 0 \\
&\|\alpha x\| = \alpha \|x\| \\
&\|x + y\| \leqslant \|x\| + \|y\| \quad \text{(the triangle inequality)}
\end{aligned}$$

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

Normed Spaces: Examples

- **Example 1:** the vector space $\mathbb{R}^n$ can be equipped with the Euclidean norm defined by $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$. Alternatively we could use the norm defined by $\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i|$
- **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 the norm
  $\|f\|_2 = \sqrt{\int_{-\infty}^{\infty} f(x)^2\, dx}$

Normed Spaces: More Examples

- **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 $j \leqslant n$ 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 $x \in S$

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

- 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-**neighbourhood** on $(S, \|\cdot\|)$

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

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 $\alpha \in \mathbb{R}$.

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

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

- A **Functional** maps an element of a vector space to a real number, e.g. $F : S \to \mathbb{R}$.
- Typically in the Calculus of Variations 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 are not very interesting.
- Easy to find $y(x)$ which minimizes these.
- Integral functionals are more interesting.
- Example integral functionals

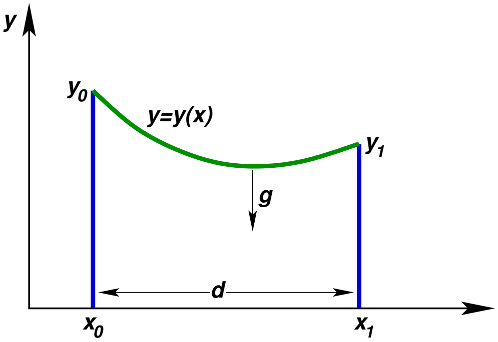
$$
\begin{aligned}
F\{y\} &= \int_a^b y(x)dx \\
F\{y\} &= \int_a^b f(x)y(x)dx \\
F\{y\} &= \int_a^b \sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx
\end{aligned}
$$

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 is the mass per unit length and g is the gravitational constant.



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

Example: the Brachistochrone

The time taken is

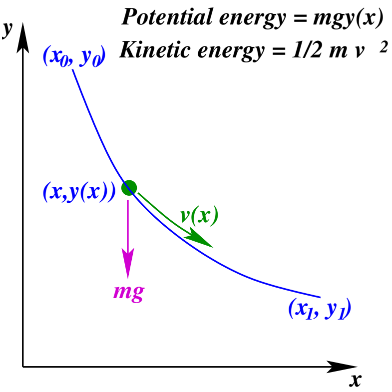
$$T\{y\} = \int_0^L \frac{ds}{v(s)}$$

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

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

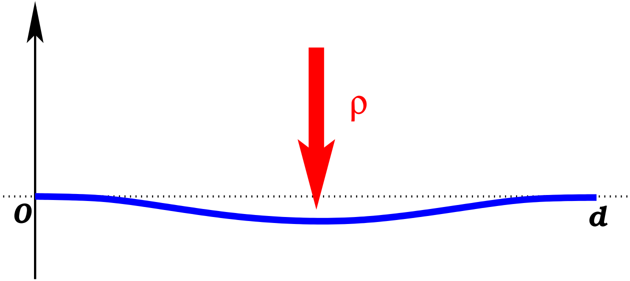
and a simple conservation law says this is constant, so

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



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

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

The potential energy is

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

Thus the total potential energy is

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