Variational Principles

Easter 2016

Contents

1 Introduction

Variational principles discusses how we maximise/minimise something that depends on other things that we can vary.
For many problems we have a continuous infinity of independent variables.

Example. (Dido's problem)
What curve of fixed length maximises an enclosed area?

Example. (Newton's problem, in Principia 1687)
What surface of revolution minimises resistance to motion in a fluid?

Example. (The *brachistochrone*)
What curve of wire minimises the time for the bead to fall from rest?

2 Variational principles

• Hero's principle (-100 BC): 'Light travels by the shortest path'.
• Fermat's principle (1662): Light travels on a path of *least time*. He assumed that light is slowed when it passes into a more dense material. Effectively he was assuming that the velocity is inversely proportional to the refraction index. He used this principle to derive Snell's law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2.$$

For a medium of variable refractive index $n(\mathbf{x})$, Fermat's principle is equivalent to the minimum optical path length:

$$P = \int_C n(x)\,dx$$

vary C to find the minimum of P. This led to analogies principles for mechanics.

Calculus for functions of many variables:

$$f : \mathbb{R}^n \to \mathbb{R}$$
$$\mathbf{x} = (x_1, ..., x_n) \to f(\mathbf{x}) \in \mathbb{R}$$

Assume f is smooth. Stationary points of f are points in $\mathbb{R}^n$ for which $\nabla f = \mathbf{0}$, i.e.

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right) = (0, 0, ..., 0).$$

Expanding in Taylor series about a stationary point $\mathbf{x} = \mathbf{a}$:

$$f(\mathbf{x}) = f(\mathbf{a}) + 0 + \frac{1}{2}\sum_{i,j}(x_i - a_i)(x_j - a_j) H_{ij}(\mathbf{a}) + O\left((x-a)^3\right)$$

Where H is the Hessian matrix which is symmetric,

$$H_{ij}(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_i \partial x_j}$$

Focus on one stationary point: assume wlog that $\mathbf{a} = \mathbf{0}$. Then

$$f(\mathbf{x}) - f(\mathbf{0}) = \frac{1}{2} x_i H_{ij}(\mathbf{0})\, x_j$$

Now let $x_i = R_{ij} x'_j$ for some rotation matrix R, such that the new Hessian H' in this basis is diagonal, i.e. $H' = diag(\lambda_1, ..., \lambda_n)$. Then

$$f(\mathbf{x}) - f(\mathbf{0}) = \frac{1}{2}\sum_i \lambda_i (x'_i)^2 + O\left(x'^3\right)$$

This is positive definite if $\lambda_i > 0$ for all i, i.e. a local minimum. It's negative definite if $\lambda_i < 0$ for all i, in which we have a local maximum. If some of the eigenvalues are positive while others are negative, we have a saddle point. If some eigenvalues are zero then we have a degenerate stationary point and need to investigate further (go to higher orders).

Example. Consider n=2.

$$\det H = \lambda_1 \lambda_2$$
$$\operatorname{tr} H = \lambda_1 + \lambda_2$$

If both are positive then this is a local minimum; if $\det H > 0$ and $\operatorname{tr} H < 0$ then this is a local maximum; if $\det H < 0$ this is a saddle point; if $\det H = 0$ then this case is degenerate.

Example. Find stationary points of

$$f(x, y) = x^3 + y^3 - 3xy$$