Calculus of Variations

Namrata Vaswani, namrata@iastate.edu

These notes are still under preparation. Please email me if you find any mistakes and typos.

These notes are based on Chapter 1 of [1] and some web sources.

Consider the problem of minimizing an energy functional $E(u)$ which is an integral of a function of an unknown function $u(x)$ and its derivatives w.r.t. x. u and its derivatives are only known at the boundaries of the integration domain.

Calculus of variations is used to find the gradient of a functional (here $E(u)$) w.r.t. a function (here $u(x)$), which we denote by $\nabla_u E$. Setting $\nabla_u E = 0$ gives the **Euler-Lagrange equation** and this is a necessary condition for the minimizing function to satisfy. In some cases the Euler-Lagrange can be solved directly in closed form. For other cases one uses numerical techniques for gradient descent, which gives rise to a Partial Differential Equation (PDE). In effect, Calculus of Variations extends vector calculus to enable us to evaluate derivatives of functionals.

A. *Evaluating* $\nabla_u E$

I explain here the simplest case: how to evaluate $\nabla_u E$ when E can be written as a definite integral of u and its first partial derivative $u_x \triangleq \frac{\partial u}{\partial x}$ and x is a scalar, i.e.

$$E(u) = \int_a^b L(u, u_x) dx \tag{1}$$

and $u(a)$ and $u(b)$ are known (fixed), while $u(x),\ x \in (a, b)$ is variable. Here $L(u, u_x)$ is referred to as the **Lagrangian**.

1) Defining $\nabla_u E$*:* From vector calculus (if u were a vector $u = [u_1, u_2, ..u_n]$), then the directional derivative in a direction α is

$$\frac{\partial E}{\partial \alpha} = \lim_{\epsilon \longrightarrow 0} \frac{E(u + \epsilon\alpha) - E(u)}{\epsilon} = \nabla_u E \cdot \alpha \tag{2}$$

where the dot product expands as

$$\nabla_u E \cdot \alpha = \sum_{i=1}^{n} (\nabla_u E)_i \alpha_i \tag{3}$$

One way to evaluate $\nabla_u E$ is to write out a first order Taylor series expansion of $E(u+\epsilon\alpha)$ and define $(\nabla_u E)_i$ by comparison as the term multiplying α_i.

We use this same methodology for calculus of variations, but now u is a continuous function of a variable x and α is also a continuous function of x with unit norm ($||\alpha||^2 = \int_a^b \alpha(x)^2 dx = 1$). The dot product is now defined as

$$\nabla_u E \cdot \alpha = \int_a^b (\nabla_u E)(x)\alpha(x)dx. \tag{4}$$

The boundary conditions, $u(a)$ *and* $u(b)$ *are fixed and so* $\alpha(a) = \alpha(b) = 0$*.*

2) The solution method: Expand $E(u+\epsilon\alpha)$ using first order Taylor series as

$$E(u+\epsilon\alpha) \approx E(u) + \epsilon\nabla_u E \cdot \alpha = E(u) + \epsilon\int_a^b (\nabla_u E)(x)\alpha(x)dx \tag{5}$$

and $(\nabla_u E)(x)$ is the term multiplying $\alpha(x)$ in this expansion.

Applying this to (1), we get

$$\begin{aligned} E(u+\epsilon\alpha) &= \int_a^b L(u+\epsilon\alpha, u_x+\epsilon\alpha_x)dx \\ &\approx \int_a^b L(u,u_x)dx + \epsilon\int_a^b (\frac{\partial L}{\partial u})(x)\alpha(x)dx + \epsilon\int_a^b (\frac{\partial L}{\partial u_x})(x)\alpha_x(x)dx \\ &= E(u) + \epsilon(T_1 + T_2) \end{aligned} \tag{6}$$

Now T_1 is already in the form of a dot product with α. We need to bring $T_2 = \int_a^b \frac{\partial L}{\partial u_x}\alpha_x dx$ also in this form. We do this using **integration by parts.** Recall that

$$\int_a^b f(x)g_x(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f_x(x)g(x)dx \tag{7}$$

We apply this to T_2 with $f = \frac{\partial L}{\partial u_x}$, $g = \alpha$ and $\alpha(a) = \alpha(b) = 0$, so that the first two terms of (7) vanish.

$$\begin{aligned} T_2 = \int_a^b \frac{\partial L}{\partial u_x}\alpha_x dx &= (\frac{\partial L}{\partial u_x})(b)\alpha(b) - (\frac{\partial L}{\partial u_x}(a))\alpha(a) - \int_a^b \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}\alpha dx \\ &= 0 - 0 - \int_a^b \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}\alpha dx \end{aligned} \tag{8}$$

Thus combining (6) with (8), we get

$$E(u+\epsilon\alpha) = E(u) + \epsilon\int_a^b [(\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}]\alpha dx \tag{9}$$

and thus by comparison with (5), we have

$$(\nabla_u E)(x) = [(\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}](x) \tag{10}$$

Thus the *Euler Lagrange equation (necessary condition for a minimizer) is*

$$[(\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}] = 0 \tag{11}$$

This can be either solved directly or using gradient descent. When using gradient descent to find the minimizing u, we get a PDE with an artificial time variable t as

$$\frac{\partial u}{\partial t} = -(\nabla_u E) = -[(\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x}] \tag{12}$$

B. Extensions

The solution methodology can be easily extended to cases where (i) u is a function of multiple variables, i.e. x is a vector ($x = [x_1, x_2, ..x_k]^T$) or (ii) when u itself is a vector of functions, i.e. $u(x) = [u_1(x), u_2(x)...u_m(x)]^T$ or (iii) when E depends on higher order derivatives of u. **Please UNDERSTAND the basic idea in the above derivation carefully, some of these extensions may be Exam questions.**

Exercise: Show that if $u = u(x, y)$ i.e. u is a function of two scalar variables x and y, then

$$(\nabla_u E) = (\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial u_y} \tag{13}$$

A trivial extension of this shows that if E is a function of two functions $u(x, y)$ and $v(x, y)$, then the Euler-Lagrange equation is given by

$$\begin{aligned} (\nabla_u E) = (\frac{\partial L}{\partial u}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial u_y} = 0 \\ (\nabla_v E) = (\frac{\partial L}{\partial v}) - \frac{\partial}{\partial x}\frac{\partial L}{\partial v_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial v_y} = 0 \end{aligned} \tag{14}$$

One application of this is in estimating Optical Flow using Horn and Schunk's method [2] (see Optical flow handout). More applications will be seen in Segmentation problems, which attempt to find the object contour $C(p) = [C^x(p), C^y(p)]$ (where p is a parameter that goes from 0 to 1 over the contour and $C(0) = C(1)$ for closed contour) that minimizes an image dependent energy functional $E(C)$.

REFERENCES

[1] G.Sapiro, *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, 2001.

[2] B.K.P. Horn and B.G.Schunk, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.