Reading group: Calculus of Variations and Optimal Control Theory by Daniel Liberzon

Alexandre Vieira

27th March 2017

Content

Content

Reminder

Consider

$$J : V \to \mathbb{R}$$

V : infinite-dimensional vector space (usually, a function space).
J is called a functional.

Definition: First Variation

For a function $y \in V$, we call the first variation of J at y, the linear functional $\delta J|_y : V \to \mathbb{R}$ satisfying, for all η and all α:

$$J(y + \alpha\eta) = J(y) + \delta J|_y(\eta)\alpha + o(\alpha)$$

Reminder

Suppose we want to find a local minimum of a functional J over a subset A of V, associated to a certain norm.

Definition: Admissible perturbation

We call a perturbation $\eta \in V$ *admissible* if $y^* + \alpha\eta \in A$ for all α close enough to 0.

Proposition

If y^* is a local minimum, then for all admissible perturbations η, we must have

$$\delta J|_y(\eta) = 0$$

Content

Basic Problem

Among all $\mathscr{C}^1([a,b],\mathbb{R})$ curves y, satisfying given boundary conditions

$$y(a) = y_0, y(b) = y_1$$

find (local) minima of the cost functional

$$J(y) = \int_a^b L(x, y(x), y'(x))dx$$

L called the *Lagrangian* (Analytical Mechanics Community) or the *running cost* (Optimal Control Community).

First-Order conditions

$$\min_{y \in \mathscr{C}^1} \int_a^b L(x, y(x), y'(x))dx = J(y)$$
$$\text{s.t. } y(a) = y_0, \; y(b) = y_1$$

As we did in the finite dimensional case, let us consider $\mathscr{C}^1$ perturbations η around a reference curve y:

$$y + \alpha\eta$$

In order to still comply with the *constraints* on y, we choose the perturbations such that $\eta(a) = \eta(b) = 0$.
A necessary condition for y to be optimal is, for every such perturbation η:

$$\delta J_{|y}(\eta) = 0$$

First-Order conditions

$$\delta J_{|y}(\eta) = \lim_{\alpha \to 0} \frac{J(y + \alpha\eta) - J(y)}{\alpha} = \frac{d}{d\alpha}\Big|_{\alpha=0} J(y + \alpha\eta)$$

If we assume enough smoothness for L, we can invert derivation and summation. Eventually:

$$\delta J_{|y}(\eta) = \int_a^b (L_y(x, y(x), y'(x))\eta(x) + L_z(x, y(x), y'(x))\eta'(x))dx$$

We perform an integration by part on the second term under the summation sign:

$$\int_a^b L_z(x, y(x), y'(x))\eta'(x)dx = -\int_a^b \frac{d}{dx} L_z(x, y(x), y'(x))\eta(x)dx + \underbrace{[L_z(x, y(x), y'(x))\eta(x)]_a^b}_{=0}$$

First-Order conditions

$$\delta J|_y(\eta) = \int_a^b \left(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_z(x, y(x), y'(x)) \right) \eta(x) dx$$

for all $\mathscr{C}^1$ perturbation η. We can therefore prove easily that this implies nullity of the integrand:

$$L_y(x, y(x), y'(x)) = \frac{d}{dx} L_z(x, y(x), y'(x)), \; \forall x \in [a, b]$$

Euler-Lagrange Equation

The first order condition for a weak minimum of the Basic calculus of variation problem is given by the Euler-Lagrange equation:

$$L_y = \frac{d}{dx} L_{y'}$$

Special cases

- "no y" : if $L = L(x, y')$, the EL equation reduces to:

$$L_{y'}(x, y'(x)) = cst$$

  This function is called the *momentum*.

- "no x" : if $L = L(y, y')$, the EL equation reduces to:

$$L_{y'}y' - L = cst$$

  This function is called the *Hamiltonian*.

Extension: variable-endpoint

We now complicate a bit the problem: we suppose the endpoint now free.

$$\min_{y \in \mathscr{C}^1} \int_a^b L(x, y(x), y'(x))dx = J(y)$$
$$\text{s.t. } y(a) = y_0$$

Most of the previous work still hold: the first variation now reads:

$$\delta J|_y(\eta) = \int_a^b \left(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_z(x, y(x), y'(x)) \right) \eta(x)dx + L_z(b, y(b), y'(b))\eta(b)$$

The perturbations with $\eta(b) = 0$ are still admissible, so the EL equations still holds.

Extension: variable-endpoint

It implies that:

$$L_z(b, y(b), y'(b))\eta(b) = 0$$

Since $\eta(b)$ is arbitrary, we have the following transversality condition:

$$L_z(b, y(b), y'(b)) = 0$$

Content

Hamilton's equation

We defined a few slides before the momentum:

$$p = L_{y'}(x, y, y')$$

and the Hamiltonian:

$$H(x, y, y', p) = p \cdot y' - L(x, y, y')$$

Along a curve y which is an extremal (e.g. solution of the EL equation), and considering p as a function of x, we have:

$$\frac{dy}{dx} = H_p(x, y(x), y'(x))$$

$$\frac{dp}{dx} = \frac{d}{dx} L_{y'}(x, y(x), y'(x)) = L_y(x, y(x), y'(x)) = -H_y(x, y(x), y'(x))$$

Hamilton's equation

Hamilton's equation

We call Hamilton's canonical equations:

$$y' = H_p, \; p' = -H_y$$

Several remarks:

1. Later, in the optimal control framework, p will be called the adjoint state.
2. We can see the construction of the Hamiltonian as a Legendre transform of the Lagrangian L. Recall, from Boyd's book, that the Legendre transform f^* of a function f is defined by:

$$f^*(p) = \max_{\xi}\{p\xi - f(\xi)\}$$

Content

Integral constraints

We now add constraints to the curves we define admissible. A first constraint is of the form:

$$C(y) = \int_a^b M(x, y(x), y'(x))dx = C_0$$

As it is done in the finite-dimensional case, we add this constraint to the running cost, multiplied by a multiplier λ (augmented cost). Applying directly the EL equations, there must exist a constant λ such that the optimal curve y complies with:

$$(L + \lambda^* M)_y = \frac{d}{dx}(L + \lambda^* M)_{y'}$$

Integral constraints

Given two perturbations η_1, η_2 arbitrary, consider :

$$F : (\alpha_1, \alpha_2) \mapsto (J(y + \alpha_1\eta_1 + \alpha_2\eta_2), C(y^* + \alpha_1\eta_1 + \alpha_2\eta_2))$$

The Jacobian matrix of F at $(0,0)$ is :

$$J_F(0,0) = \begin{pmatrix} \delta J|_y(\eta_1) & \delta J|_y(\eta_2) \\ \delta C|_y(\eta_1) & \delta C|_y(\eta_2) \end{pmatrix}$$

Theorem: Inverse Function Theorem

If the total derivative of a continuously differentiable function F defined from an open set of $\mathbb{R}^n$ into $\mathbb{R}^n$ is invertible at a point p (i.e., the Jacobian determinant of F at p is non-zero), then F is an invertible function near p. Moreover, the inverse function F^{-1} is also continuously differentiable.

Integral constraints

With this theorem, if y is a minimum, it implies that $J_F(0,0)$ has to be singular. It implies that there exists $(\lambda_0, \lambda^*) \in \mathbb{R}^2 \backslash (0,0)$ such that:

$$\lambda_0 \delta J_{|y}(\eta_i) + \lambda^* \delta C_{|y}(\eta_i) = 0, \; i = 1, 2$$

- If y is an extremal of C, then $\delta C_{|y} = 0$, and then either $\lambda_0 = 0$ (abnormal case), or $\delta J_{|y}(\eta_1) = 0$, whatever η_1, so y would also be an extremal of J.
- Otherwise, there exists η_1 such that $\delta C_{|y}(\eta_1) \neq 0$ and $\lambda_0 \neq 0$ (we can thus divide by λ_0 each equality, or take $\lambda_0 = 1$). Define λ^* as:

$$\lambda^* = -\frac{\delta J_{|y}(\eta_1)}{\delta C_{|y}(\eta_1)}$$

It implies that for all perturbation η_2, $\delta J_{|y}(\eta_2) + \lambda^* \delta C_{|y}(\eta_2) = 0$.

Integral constraints

We drop the 2 and write η for the perturbation.

$$\delta J|_y(\eta) = \int_a^b \left(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_z(x, y(x), y'(x)) \right) \eta(x) dx$$

$$\delta C|_y(\eta) = \int_a^b \left(M_y(x, y(x), y'(x)) - \frac{d}{dx} M_z(x, y(x), y'(x)) \right) \eta(x) dx$$

$$\delta J|_y(\eta) + \lambda^* \delta C|_y(\eta) =$$
$$\int_a^b \left((L + \lambda^* M)_y(x, y(x), y'(x)) - \frac{d}{dx} (L + \lambda^* M)_z(x, y(x), y'(x)) \right) \eta(x) dx$$

Integral constraints

$$\delta J|_y(\eta) + \lambda^* \delta C|_y(\eta) = \int_a^b \left((L + \lambda^* M)_y(x, y(x), y'(x)) - \frac{d}{dx}(L + \lambda^* M)_z(x, y(x), y'(x)) \right) \eta(x) dx$$

Since that must be true for all perturbation η, we have the following theorem:

Euler-Lagrange equation in the integral constrained case

There exist two scalars λ_0, λ^*, not simultaneously nought, such that the optimal curve y complies with:

$$(\lambda_0 L + \lambda^* M)_y = \frac{d}{dx}(\lambda_0 L + \lambda^* M)_{y'}$$

The degenerate cases appear when $\lambda_0 = 0$.

Non-Integral constraints

When the constraint is an equality constraint on the whole curve:

$$M(x, y(x), y'(x)) = 0, \ \forall x \in [a, b]$$

the same logic applies, but λ^* is not a scalar anymore. More precisely:

Euler-Lagrange equation in the non-integral constrained case

There exist a scalar λ_0 and a function λ^*, never simultaneously nought, such that the optimal curve y complies with:

$$(\lambda_0 L + \lambda^* M)_y = \frac{d}{dx}(\lambda_0 L + \lambda^* M)_{y'}$$

The degenerate cases appear when $\lambda_0 = 0$.

Content

Necessary conditions for a weak minimum

Second order conditions are here a bit more tricky. There are here just quickly presented, for future reference. Throughout this, we consider only the basic calculus of variation problem.

Legendre's necessary condition for a weak minimum

A necessary condition for the curve y to minimize the cost is: for all $x \in [a, b]$, we must have

$$L_{y'y'}(x, y(x), y'(x)) \geq 0$$

There is also a sufficient condition, necessitating the notion of conjugate point. The reader interested are referred to section 6.2 of the book.

Exercises

- Two fun (and historical!) examples: exercises 2.5 and 2.10
- A more theoretic one: exercise 2.6