Functional Derivative

The goal of this section is to discover a suitable definition of a "functional derivative", such that we can take the derivative of a functional and still have the same rules of differentiation as normal calculus. For example, we wish to find a definition for $\frac{\delta J}{\delta y}$, where $J[y(x)]$ is a functional of $y(x)$ such that things like $\frac{\delta}{\delta y}J^2 = 2J\frac{\delta J}{\delta y}$ are still true.

Definitions

Functional

Stone's definition of *local functional* where f is a multivariable function

$$J[y] = \int_{x_1}^{x_2} f(x, y(x), y'(x), y''(x), \cdots, y^{(n)}(x))dx = \int_{x_1}^{x_2} f dx \tag{1}$$

Notice the functional J "eats" an entire function y, which is defined using its local values $y(x), y'(x)$ etc, and spits out a number through integration. In short, a functional is just a number that depends on an input function.

Variation

A *variation* of the functional is the amount the functional changes when the input function is changed by a little bit. Suppose we let $y(x) \to y(x) + \delta y(x)$, then since $\frac{d}{dx}$ is linear

$$\begin{cases} y'(x) \to y'(x) + \frac{d}{dx}\delta y(x) = y'(x) + \delta y'(x) \\ y''(x) \to y''(x) + \frac{d^2}{dx^2}\delta y(x) = y''(x) + \delta y''(x) \\ \vdots \\ y^{(n)}(x) \to y^{(n)}(x) + \frac{d^n}{dx^n}\delta y(x) = y^{(n)}(x) + \delta y^{(n)}(x) \end{cases} \tag{2}$$

thus the integrant of the new output $J[y+\delta y]$ can be expanded to first order using Taylor expansion of a multivariable function around the old input y

$$\begin{aligned} J[y+\delta y] &= \int_{x_1}^{x_2} f(x, y+\delta y, y'+\delta' y, \cdots, y^{(n)} + \delta^{(n)} y)dx \\ &= \int_{x_1}^{x_2} \left\{ f + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y' + \cdots + \frac{\partial^{(n)} f}{\partial y^{(n)}}\delta y^{(n)} \right\} dx \end{aligned} \tag{3}$$

The variation of the functional is thus, by definition, the new output minus the old output taken to first order.

$$\begin{aligned}\delta J =& J[y+\delta y] - J[y] \\ =& \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y' + \cdots + \frac{\partial^{(n)} f}{\partial y^{(n)}}\delta y^{(n)} \right\} dx \end{aligned} \tag{4}$$

we can moved all the $\frac{d}{dx}$ on δy to f using integration by parts

$$\begin{aligned}\delta J =& \frac{d}{dx}(\frac{\partial f}{\partial y'})\delta y(x)\Bigg|_{x_1}^{x_2} + \frac{d}{dx}(\frac{\partial f}{\partial y''})\delta y'(x)\Bigg|_{x_1}^{x_2} - \frac{d^2}{dx^2}(\frac{\partial f}{\partial y''})\delta y(x)\Bigg|_{x_1}^{x_2} \\ &+ \cdots + (-1)^{n-1}\frac{d^n}{dx^n}(\frac{\partial f}{\partial y^{(n)}})\delta y^{(n)}(x)\Bigg|_{x_1}^{x_2} + \int_{x_1}^{x_2} \frac{\partial f}{\partial x} dx + \\ &\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} - \cdots + (-1)^{n-1}\frac{d^n}{dx^n}\frac{\partial f}{\partial y^{(n)}} \right) \delta y(x) dx \end{aligned} \tag{5}$$

and Voila! (5) is the *variation* of the local functional defined in (1) with no additional assumption.

Functional Derivative

For a normal multi-variable function $f(x_1, x_2, \cdots, x_n)$ we have a nice form for its variation

$$df = \sum_{i=1}^{n} \left\{ \frac{\partial f}{\partial x_i} dx_i \right\} \tag{6}$$

and we know how to calculate the derivatives $\frac{\partial f}{\partial x_i}$. Here we wish to rewrite (5) such that we have a similar form for the variation of a functional

$$\delta J = \int_{x_1}^{x_2} dx \left\{ \frac{\delta J}{\delta y}(x) \delta y(x) \right\} \tag{7}$$

Unfortunately, this is only possible under special circumstances. That is, we need the variation to have **"fixed-ends"** ($\delta y^{(n)}(x_1) = \delta y^{(n)}(x_2) = 0$) and that we require **implicit f** ($\frac{\partial f}{\partial x} = 0$)[1]. Basically, we want everything before the last line of (5) to vanish. This way, comparing (5) and (7) we finally have

$$\boxed{\frac{\delta J}{\delta y} = \frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} - \cdots + (-1)^{n-1}\frac{d^n}{dx^n}\frac{\partial f}{\partial y^{(n)}}} \tag{8}$$

[1]This is slightly over kill, since we just want $\int_{x_1}^{x_2} \frac{\partial f}{\partial x} dx = 0$

As long as δJ can be written as (7), we will have our nice rules of differentiation. For example

$$\begin{aligned}\delta(J^2) \equiv&(J+\delta J)^2 - J^2 = 2J\delta J + O(\delta J^2)\\ =&\int_{x_1}^{x_2} dx \left\{2J\frac{\delta J}{\delta y}(x)\delta y(x)\right\} \Rightarrow\\ \frac{\delta J^2}{\delta y} &= 2J\frac{\delta J}{\delta y}\end{aligned} \tag{9}$$

A word of caution: This definition of functional derivative is nice, but as f involves higher derivatives of y, the fixed-end condition becomes harsher and the range of y this derivative applies to quickly diminishes. Therefore it is sometimes more useful to make variations by hand according to (5).

Lagrangian Mechanics

When the integrant of the functional only has dependence on y and y' ($f(y, y')$), (8) reduces to the popular *Fréchet derivative*

$$\frac{\delta J}{\delta y} = \frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} \tag{10}$$

this form should look familiar to all physicists, since it laid the foundation for basic Lagrangian mechanics. In a typical classical mechanics problem, we wish to minimize the *action* S, which is often a functional of a *configuration function* q, whose basic *independent variable* is time t. That is

$$S[q] = \int_{t_1}^{t_2} L(q(t), \dot{q}(t))dt \tag{11}$$

where L is the *Lagrangian.* To find extrema, we set derivative to 0

$$\begin{aligned}\frac{\delta S}{\delta q} &= 0 \Rightarrow\\ \frac{\partial L}{\partial q} &= \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\end{aligned} \tag{12}$$

Lo and behold, the *Lagrangian equation of motion.*