Functional differentiation

March 22, 2016

1 Functions vs. functionals

What distinguishes a functional such as the action $S\left[x\left(t\right)\right]$ from a function $f\left(x\left(t\right)\right)$, is that $f\left(x\left(t\right)\right)$ is a number for each value of t, whereas the value of $S\left[x\left(t\right)\right]$ cannot be computed without knowing the entire function $x\left(t\right)$. Thus, functionals are nonlocal. If we think of functions and functionals as maps, a compound function is the composition of two maps

$$\begin{aligned} f &: R\rightarrow R \\ x &: R\rightarrow R \end{aligned}$$

giving a third map

$$f\circ x:R\rightarrow R$$

A functional, by contrast, maps an entire function space into R,

$$\begin{aligned} S &: \mathcal{F}\rightarrow R \\ \mathcal{F} &= \left\{f\left(x\right)|x:R\rightarrow R\right\} \end{aligned}$$

In this section we develop the *functional derivative*, that is, the generalization of differentiation to functionals.

2 Intuitions

2.1 Analogy with the derivative of a function

We would like the functional derivative to formalize finding the extremum of an action integral, so it makes sense to review the variation of an action. The usual argument is that we replace $x(t)$ by $x(t)+h(t)$ in the functional $S[x(t)]$, then demand that to first order in $h(t)$,

$$\delta S\equiv S[x+h]-S[x]=0$$

Now, suppose S is given by

$$S[x(t)]=\int L(x(t),\dot{x}(t))dt$$

Then replacing x by $x+h$, subtracting S and dropping all higher order terms gives

$$\begin{aligned} \delta S &\equiv \int L(x+h,\dot{x}+\dot{h})dt-\int L(x,\dot{x})dt \\ &= \int\left(L(x,\dot{x})+\frac{\partial L(x,\dot{x})}{\partial x}h+\frac{\partial L(x,\dot{x})}{\partial\dot{x}}\dot{h}\right)dt-\int L(x,\dot{x})dt \\ &= \int\left(\frac{\partial L(x,\dot{x})}{\partial x}-\frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial\dot{x}}\right)h(t)\;dt \end{aligned}$$

$$0 = \delta S = \int \left(\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} \right) h(t)\ dt$$

and since $h(t)$ is arbitrary, we conclude that the Euler-Lagrange equation must hold at each point,

$$\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} = 0$$

By analogy with ordinary differentiation, we would like to replace this statement by the demand that the extrema are given by the vanishing of the first functional derivative of $S[x]$,

$$\frac{\delta S[x(t)]}{\delta x(t)} = 0$$

This means that the functional derivative of S should be the Euler-Lagrange expression. It is tempting to define

$$\frac{\delta S[x(t)]}{\delta x(t)} \equiv \lim_{h\to 0} \frac{S[x+h] - S[x]}{h(t)}$$

but even the ratio $\frac{S[x+h]-S[x]}{h(t)}$ is not well-defined if $h(t)$ has any zeros. What does work is to convert the problem to an ordinary differentiation.

2.2 The Dirac delta

Suppose S is given by

$$S[x(t)] = \int L(x(t), \dot{x}(t))dt$$

Then replacing x by $x+h$ and subtracting S gives

$$\begin{aligned}
\delta S &\equiv \int L(x+h, \dot{x}+\dot{h})dt - \int L(x,\dot{x})dt \\
&= \int \left(L(x,\dot{x}) + \frac{\partial L(x,\dot{x})}{\partial x} h + \frac{\partial L(x,\dot{x})}{\partial \dot{x}} \dot{h} \right) dt - \int L(x,\dot{x})dt \\
&= \int \left(\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} \right) h(t)\ dt
\end{aligned}$$

Setting $\delta x = h(t)$ we may write this as

$$\delta S = \int \left(\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt'}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} \right) \delta x(t')\ dt'$$

Now write

$$\delta S = \frac{\delta S}{\delta x(t)} \delta x(t) = \left(\int \left(\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt'}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} \right) \frac{\delta x(t')}{\delta x(t)}\ dt' \right) \delta x(t)$$

or simply

$$\frac{\delta S}{\delta x(t)} = \int \left(\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt'}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} \right) \frac{\delta x(t')}{\delta x(t)}\ dt' \tag{1}$$

We might write this much by just using the chain rule. What we need is to evaluate the basic functional derivative,

$$\frac{\delta x(t')}{\delta x(t)}$$

To see what this might be, consider the analogous derivative for a countable number of degrees of freedom. Beginning with

$$\frac{\partial q^j}{\partial q^i} = \delta^j_i$$

we notice that when we sum over the i index holding j_0 fixed, we have

$$\sum_i \frac{\partial q^{j_0}}{\partial q^i} = \sum_i \delta_i^{j_0} = 1$$

since $i = j_0$ for only one value of i. We need the continuous version of this relationship. The sum over independent coordinates becomes an integral, $\sum_i \to \int dt'$, so we demand

$$\int \frac{\delta x(t')}{\delta x(t)} dt' = 1$$

This will be true provided we use a Dirac delta function for the derivative:

$$\frac{\delta x(t')}{\delta x(t)} = \delta(t' - t)$$

Substituting this expression into eq.(1) gives the desired result for $\frac{\delta S}{\delta x(t)}$:

$$\begin{aligned}
\frac{\delta S}{\delta x(t)} &= \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt'} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta(t' - t)\ dt' \\
&= \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}}
\end{aligned}$$

Notice how the Dirac delta function enters this calculation. When finding the extrema of S as before, we reach a point where we demand

$$0 = \delta S = \int \left(\frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) h(t)\ dt$$

for every function $h(t)$. To complete the argument, suppose there is a time t_0 at which $\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}}$ is nonzero. For concreteness, let $\left.\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}}\right|_{t_0} > 0$. Then by continuity there must be a neighborhood of $t_0 \in (t_1, t_2)$ on which $\frac{\partial L(x,\dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} > 0$. Choose any function $h(t)$ which is positive on (t_1, t_2) and which vanishes outside this interval. Then we have a contradiction, since the integral must be positive. An identical argument holds if the integrand is negative, so we must have

$$\left. \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right|_{t_0} = 0$$

Since the point t_0 is arbitrary, we conclude that the expression vanishes everywhere.

This argument works fine if we only wish to find the extremum of the action. However if we can develop a formalism such that we get a Dirac delta function,

$$\frac{\delta x(t')}{\delta x(t)} = \delta(t' - t)$$

then the functional derivative will exist for all curves. We now turn to a formal definition.

3 Formal definition of the functional derivative

3.1 Test functions and distributions

We begin with some definitions.

A *test function* is a smooth function which vanishes outside of a compact region.

A *distribution* is a limit of test functions. There is an alternative definition as a linear functional taking as set of test functions to the reals. I find the idea of a limit more intuitive, though the definition as functionals has calculational advantages. For example, the Dirac delta function is such a functional, for if $f(x)$ is any test function,

$$\delta(x - x_0) : f(x) \to f(x_0) \in R$$

That is, for any test function f, the delta function maps f to the real number $f(x_0)$. As a limit of smooth functions, we may define

$$\delta(x - x_0) = \lim_{\sigma \to 0} \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \right]$$

This is only one of many representations for $\delta(x - x_0)$. In the end we only use the linear functional properties. It is straightforward to show that for any test function f,

$$\lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi\sigma^2}} \int f(x) e^{-\frac{(x-x_0)^2}{2\sigma^2}} dx = f(x_0)$$

This is written as

$$\int f(x) \delta(x - x_0) dx = f(x_0)$$

which gives us an explicit calculational form for the linear functional.

3.2 The functional derivative

Given a functional of the form

$$f[x(t)] = \int g\left(x(t'), \dot{x}(t'), \ldots\right) dt'$$

we consider a sequence of 1-paramater variations of f given by replacing $x(t')$ by

$$x_n(\varepsilon, t') = x(t') + \varepsilon h_n(t, t')$$

where each h_n is a test function, and the sequence of functions h_n defines a distribution,

$$\lim_{n \to \infty} h_n(t, t') = \delta(t - t')$$

Since we may vary the path by *any* function $h(x)$, each of these functions εh_n is an allowed variation. Then the functional derivative is defined by

$$\frac{\delta f[x(t)]}{\delta x(t)} \equiv \lim_{n \to \infty} \frac{d}{d\varepsilon} f[x_n(\varepsilon, t')] \bigg|_{\varepsilon=0} \tag{2}$$

The derivative with respect to ε accomplishes the usual variation of the action by using the chain rule. Taking a derivative and setting $\varepsilon = 0$ is just a clever way to select the part of the variation linear in ε. Then we take the limit of a carefully chosen sequence of variations h_n to extract the variational coefficient from the integral with a Dirac delta function.

To see explicitly that this works, we compute:

$$\begin{aligned}
\frac{\delta f[x(t)]}{\delta x(t)} &\equiv \lim_{n \to \infty} \frac{d}{d\varepsilon} f[x_n(\varepsilon, t')] \bigg|_{\varepsilon=0} \\
&= \lim_{n \to \infty} \int \frac{dg(\varepsilon, x(t'), \dot{x}(t'), \ldots)}{d\varepsilon} \bigg|_{\varepsilon=0} \\
&= \lim_{n \to \infty} \int \frac{d}{d\varepsilon} g\left(x(t') + \varepsilon h_n(t, t'), \dot{x}(t') + \varepsilon \dot{h}_n(t, t'), \ldots\right) \bigg|_{\varepsilon=0}
\end{aligned}$$

$$\begin{aligned}
&= \lim_{n\to\infty}\int\left(\frac{\partial g}{\partial x}h_n(t,t') + \frac{\partial g}{\partial\dot{x}(t')}\frac{dh_n(t,t')}{dt'} + \ldots\right)dt' \\
&= \int\lim_{n\to\infty}h_n(t,t')\left(\frac{\partial g}{\partial x} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}} + \ldots\right)dt' \\
&= \int\left(\frac{\partial g}{\partial x} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}} + \ldots\right)\delta(t-t')\,dt' \\
&= \frac{\partial g}{\partial x} - \frac{d}{dt}\frac{\partial g}{\partial\dot{x}} + \ldots
\end{aligned}$$

A convenient shorthand notation for this procedure is

$$\begin{aligned}
\frac{\delta f[x(t)]}{\delta x(t)} &= \frac{\delta}{\delta x(t)}\int g\left(x(t'),\dot{x}(t'),\ldots\right)dt' \\
&= \int\left(\frac{\partial g}{\partial x(t')}\frac{\delta x(t')}{\delta x(t)} + \frac{\partial g}{\partial\dot{x}}\frac{d}{dt'}\frac{\delta x(t')}{\delta x(t)} + \ldots\right)dt' \\
&= \int\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}} + \ldots\right)\frac{\delta x(t')}{\delta x(t)}dt' \\
&= \int\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}} + \ldots\right)\frac{\delta x(t')}{\delta x(t)}\delta(t'-t)\,dt' \\
&= \frac{\partial g}{\partial x(t)} - \frac{d}{dt}\frac{\partial g}{\partial\dot{x}(t)} + \ldots
\end{aligned}$$

The method can be extended to more general forms of functional f.

One advantage of the formal definition is that it can be iterated,

$$\begin{aligned}
\frac{\delta^2 f[x(t)]}{\delta x(t'')\,\delta x(t')} &= \frac{\delta}{\delta x(t')}\frac{\delta f[x(t)]}{\delta x(t')} \\
&= \frac{\delta}{\delta x(t'')}\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}(t')}+\right) \\
&= \frac{\partial g}{\partial x(t')}\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}(t')}+\right)\frac{\delta x(t')}{\delta x(t'')} + \frac{\partial g}{\partial\dot{x}(t')}\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}(t')}+\right)\frac{\delta\dot{x}(t')}{\delta x(t'')} \\
&= \frac{\partial g}{\partial x(t')}\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}(t')}+\right)\delta(t'-t'') + \frac{\partial g}{\partial\dot{x}(t')}\left(\frac{\partial g}{\partial x(t')} - \frac{d}{dt'}\frac{\partial g}{\partial\dot{x}(t')}+\right)\frac{d}{dt'}\delta(t'-t'')
\end{aligned}$$

Another advantage of treating variations in this more formal way is that we can equally well apply the technique to classical field theory.

4 Field equations as functional derivatives

We can vary field actions in the same way, and the results make sense directly. Consider varying the scalar field action

$$S = \frac{1}{2}\int\left(\partial^\alpha\varphi\partial_\alpha\varphi - m^2\varphi^2\right)d^4x$$

with respect to the field φ. Setting the functional derivative of S to zero, we have

$$\begin{aligned}
0 &= \frac{\delta S[\varphi]}{\delta\varphi(x^\mu)} \\
&= \frac{1}{2}\frac{\delta}{\delta\varphi(x^\mu)}\int\left(\partial^\alpha\varphi\partial_\alpha\varphi - m^2\varphi^2\right)d^4x'
\end{aligned}$$

$$
\begin{aligned}
&= \int \left(\partial^\alpha \varphi \frac{\partial}{\partial x'^\alpha} \frac{\delta \varphi\left(x'^\mu\right)}{\delta \varphi\left(x^\nu\right)} - m^2 \varphi \frac{\delta \varphi\left(x'^\mu\right)}{\delta \varphi\left(x^\nu\right)} \right) d^4 x' \\
&= \int \left(-\partial_\alpha \partial^\alpha \varphi - m^2 \varphi \right) \frac{\delta \varphi\left(x'^\mu\right)}{\delta \varphi\left(x^\nu\right)} d^4 x' \\
&= \int \left(-\partial_\alpha \partial^\alpha \varphi - m^2 \varphi \right) \delta^4 (x'^\mu - x^\mu) d^4 x' \\
&= -\Box \varphi - m^2 \varphi
\end{aligned}
$$

and we have the field equation.

Exercise: Find the field equation for the complex scalar field by taking the functional derivative of its action.

Exercise: Find the field equation for the Dirac field by taking the functional derivative of its action.

Exercise: Find the Maxwell equations by taking the functional derivative of its action.

With this new tool at our disposal, we turn to quantization.