Integration by parts

We begin our journey by talking about a trick that could be seen as the reverse for the product rule for derivatives. We will use it to develop an entirely new way to solve problems by minimizing a function of functions. This will basically follow Wikipedia contributors (2023), but I'm copying it here for completion's sake.

Sketch for deriving the integration-by-parts formula

We will use the $'$ symbol to indicate a derivative with respect to x. If you know the basics of derivation, you recall that:

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

Now, let's integrate on both sides:

$$\begin{aligned}\int (f(x)g(x))' \, dx &= \int f'(x)g(x)dx + \int f(x)g'(x)dx \\ f(x)g(x) &= \int f'(x)g(x)dx + \int f(x)g'(x)dx\end{aligned}$$

If we rename $u = f(x)$ and $dv = g'(x)dx$, then we get the familiar expressions for indefinite and definite integrals:

$$\begin{aligned}\int u \, dv &= uv - \int v \, du \\ \int_a^b u \, dv &= uv\Big|_a^b - \int_a^b v \, du\end{aligned}$$

Finally, imagine that the functions vanish on the extremes a and b. That is, $u(b)v(b) - u(a)v(a) = 0$. In that case, you can simplify even further:

$$\int u \, dv = -\int v \, du$$

Calculus of variations

Malliavin Calculus is sometimes called an extension of calculus of variations to stochastic processes. What does that even means? Let's deal with the first part.

Derivation and application

I'll assume that you know the basics of calculus. That means you can solve problems like this: let's say we have a function f and we want to know the value of x where f attains its minimum:

$$\min f(x) = x^2 - 4x + 5$$

To solve it, we take the derivative, make it equal to zero, and obtain the arguments that make it so.

$$\begin{aligned} f'(x) = 2x - 4 &= 0 \\ 2x &= 4 \\ x &= 2 \end{aligned}$$

Which means that $\min f(x) = f(2) = 2^2 - 4 * 2 + 5 = 1$.

This approach works great when you try to minimize a function **value** over the reals, $f : \mathbb{R} \to \mathbb{R}$, or in general for multivariate functions, $f : \mathbb{R}^n \to \mathbb{R}$.

But let's say we want to find the **function** $y = f(x)$ that minimizes an expression. We don't know how that function looks, or want to impose any preconceived notions like being a polynomial or a sum of sine/cosine pairs. What do we do, then?

From now on, it's autopilot from any course (I'll be using the Wikipedia article as the baseline). The first step is to define a "function of functions" as the starting point for minimization, and that function is essentially the integral like this:

$$J[y] = \int_{x_1}^{x_2} L(x, y, y')dx$$

To keep things grounded on reality, we will use the classical example of finding the function that represents the smallest distance between two points, a and b. Starting with a "coarse grain" case with deltas, we are minimizing the sum of all the little changes in the path between those two points using Pythagoras:

$$\min \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$$

As little changes become infinitesimal, and with some extreme abuse of notation, we switch to an integral and obtain the expression for L:

$$\begin{aligned}
\int_a^b \sqrt{dx^2 + dy^2} &= \int_a^b \sqrt{(dx)^2 \left(1 + \left(\frac{dy}{dx}\right)^2\right)} \\
&= \int_a^b \sqrt{1 + (y')^2}\sqrt{(dx)^2} \\
&= \int_a^b \sqrt{1 + (y')^2}\, dx \\
&= \int_a^b L(y')\, dx
\end{aligned}$$

Now, we will introduce a generic function $\eta(x)$ called the variation. This function represents a perturbation and the only requirement that we will impose on it is that $\eta(a) = \eta(b) = 0$. We can assume that $f(x) = y$ is the solution and η is the deviation from the solution, which we can multiply by a small number ε. We can then define a function over ε, $\Phi(\varepsilon) = J[f + \varepsilon * \eta]$, so that $\Phi(0) = f$, our solution.

The trick here is that now, instead of minimizing something that depends on unknown functions, which we can't even start to comprehend, we minimize something that depends on a number, ε. Not only that, we constructed ε and η so that we know, in advance, that the function we are looking for, the solution, is at $\Phi(\varepsilon = 0) = J[f]$ and $\Phi'(0) = 0$ because it's the minimum. Absolute genius.

$$\Phi'(0) = 0 = \left.\frac{d\Phi}{d\varepsilon}\right|_{\varepsilon=0} = \int_a^b \left.\frac{dL}{d\varepsilon}\right|_{\varepsilon=0} dx$$

We take the total derivative of Φ over ε , just like before, but now $y = f + \varepsilon\eta$ and $y' = f' + \varepsilon\eta'$:

$$\begin{aligned}
\frac{dL}{d\varepsilon} &= \frac{\partial L}{\partial x} \underbrace{\frac{dx}{d\varepsilon}}_{=0} + \frac{\partial L}{\partial y} \underbrace{\frac{dy}{d\varepsilon}}_{=\eta} + \frac{\partial L}{\partial y'} \underbrace{\frac{dy'}{d\varepsilon}}_{=\eta'} \\
&= \frac{\partial L}{\partial y}\eta + \frac{\partial L}{\partial y'}\eta'
\end{aligned}$$

This looks meaningless, but now comes the magic:

$$
\begin{aligned}
\int_a^b \left.\frac{dL}{d\varepsilon}\right|_{\varepsilon=0} dx &= \int_a^b \left(\frac{\partial L}{\partial f}\eta + \frac{\partial L}{\partial f'}\eta'\right) dx \\
&= \int_a^b \frac{\partial L}{\partial f}\eta\, dx + \underbrace{\left.\frac{\partial L}{\partial f'}\eta\right|_a^b}_{=0} - \int_a^b \eta\frac{d}{dx}\frac{\partial L}{\partial f'}\, dx \\
&= \int_a^b \left(\frac{\partial L}{\partial f}\eta - \eta\frac{d}{dx}\frac{\partial L}{\partial f'}\right) dx \\
0 &= \int_a^b \eta(x)\left(\frac{\partial L}{\partial f} - \frac{d}{dx}\frac{\partial L}{\partial f'}\right) dx
\end{aligned}
$$

We start from $\Phi'(0)$ and we do an *integration by parts*. This is crucial, without that operation this whole scheme falls apart. We also know that $\eta(a) = \eta(b) = 0$, so that term vanishes. Joining terms and taking a common factor we reach our final conclusion. The final piece of magic is realizing that this integral is zero, no matter what $\eta(x)$ we pick. By the Fundamental Lemma of Calculus of Variations, that leaves only one possibility, and it's so important that it's called the Euler-Lagrange equation:

$$
\frac{\partial L}{\partial f} - \frac{d}{dx}\frac{\partial L}{\partial f'} = 0
$$

Whatever the solution function f is, it must fulfill that condition.

Shortest Path Example

Why is it helpful? Let's return to the shortest path between two points, and I'm using $f = y$ interchangeably. When we left off, we had arrived at:

$$
L(f') = \sqrt{1 + (f')^2}
$$

We apply the Euler-Lagrange equation and we see where it leads us. What makes this example super clean is that there's no explicit f in it, only f', so we end up with a very small formula:

$$
\begin{aligned}
\frac{\partial L}{\partial f} - \frac{d}{dx}\frac{\partial L}{\partial f'} &= 0 \\
-\frac{d}{dx}\frac{f'(x)}{\sqrt{1 + [f'(x)]^2}} &= 0
\end{aligned}
$$

If the derivative is zero, then we can integrate that and the result is an unknown constant $c < 1$. We do some minor algebra, plus an integration at the end, and we are done!

$$
\begin{aligned}
\frac{f'(x)}{\sqrt{1+[f'(x)]^2}} &= c \\
\frac{[f'(x)]^2}{1+[f'(x)]^2} &= c^2 \\
[f'(x)]^2 &= c^2 + c^2[f'(x)]^2 \\
[f'(x)]^2 &= \frac{c^2}{1-c^2} \\
f'(x) &= \sqrt{\frac{c^2}{1-c^2}} = m \\
f(x) &= mx + b
\end{aligned}
$$

The function with the shortest path between two points is a straight line, and the only thing we needed to know is how the distance between two points is calculated to get a closed-form, analytic solution.

Wikipedia contributors. 2023. "Calculus of Variations — Wikipedia, the Free Encyclopedia." https://en.wikipedia.org/wiki/Calculus_of_variations.