Sidharth Ghoshal

A generalized technique for Meta-Functional Optimization

The Classical problem of operator optimization:

$$find\ y\ that\ optimizes\ \int_a^b L\big(x, y, y', y'' \dots . y^{(n)}\big)\, dx$$

Was classically solved by Euler and Lagrange as the y that satisfies

$$0 = \frac{\partial L}{\partial y} - \frac{d}{dx}\left[\frac{\partial L}{\partial y'}\right] + \frac{d^2}{dx^2}\left[\frac{\partial L}{\partial y''}\right] - \frac{d^3}{dx^3}\left[\frac{\partial L}{\partial y'''}\right] \dots \frac{d^n}{dx^n}\left[\frac{\partial L}{\partial y^{(n)}}\right]$$

An intuitive though 'loose'generalization allows us to even maximize integrals of objects other than derivatives. Observing taylor's formula which states

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2 \dots$$

We can easily show:

$$f(x+a) = f(a) + f'(a)x + \frac{1}{2} f''(a)x^2 + \dots.$$

And by commutativity of addition that allows us just as easily to state:

$$f(x+a) = f(x) + f'(x)a + \frac{1}{2} f''(x)a^2 \dots$$

Thus if we want to optimize an expression such as

$$L\big(x, y, D_{h,x}y\big) = \frac{1+x+y}{1-x-D_{h,x}y} \left(whereas\ D_{h,x}y = \frac{y(x+h) - y(x)}{h} \right)$$

The same problem can be reduced to

$$L = \frac{1+x+y}{1 - y - f'(x) - \frac{1}{2} h f''(x) - \frac{1}{3!} h^2 f'''(x) \dots}$$

And then can use the infinite order generalization of

$$0 = \frac{\partial L}{\partial y} - \frac{d}{dx}\left[\frac{\partial L}{\partial y'}\right] + \frac{d^2}{dx^2}\left[\frac{\partial L}{\partial y''}\right] - \frac{d^3}{dx^3}\left[\frac{\partial L}{\partial y'''}\right] \dots$$

To derive an optimal y. This of course leads to a natural question what if the expression we want to optimize isn't an integral at ALL? Suppose we consider an expression such as

$$Maximize \sum_{i=0}^{100} \frac{1+y(i)}{1+\left(y'(i)\right)^2}$$

The techniques mentioned earlier of course would have no application in this sum given that the sum itself cannot be cleanly transformed into an integral under all circumstances. We will show a generalization of the Euler Lagrange Equations that allow us to efficiently optimize such expressions.

Given the h-step difference

$$D_{h,x}f = \frac{f(x+h)-f(x)}{h}$$

We can define the h-step inverse difference

$$D_{h,x}^{-1}f = G(x)\ s.t. D_{h,x}G = f$$

Naturally the case of h = 1 corresponds to the sum above, and the case of h = 0 corresponds quite simply to the integral. Thus the ability to optimize

$$D_{h,x}^{-1}\left[L\left(x, y, D_{h,x}y, D_{h,x}^{2}y \dots D_{h,x}^{n}\right)\right]$$

Would effectively cover the cases we want, as well some additional theory. We thus set this up in the technique analogous to those of Lagrange. Let $y = \theta + \eta t(x)$ where θ is the optimal solution and $t(x)$ is some variational equation that is varied by the constant η. From here it follows that

$$D_{h,x}^{-1}\left[L\left(x, y, D_{h,x}y, D_{h,x}^{2}y \dots D_{h,x}^{n}\right)\right] = D_{h,x}^{-1}\left[L\left(x, \theta+\eta t(x), D_{h,x}\theta+\eta D_{h,x}t(x), \dots\right)\right] = \phi(\eta)$$

We now note that

$$\frac{d\phi(\eta)}{d\eta} = 0$$

Is of course the optimal functional. Classically this yields

$$\frac{d\phi(\eta)}{d\eta} = \frac{d}{d\eta}\left[D_{h,x}^{-1}\left[L\left(x, \theta+\eta t(x), D_{h,x}\theta+\eta D_{h,x}t(x), \dots\right)\right]\right] =$$

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial x}\frac{dx}{d\eta} + \frac{\partial L}{\partial y}\frac{dy}{d\eta} + \frac{\partial L}{\partial D_{h,x}y}\frac{dD_{h,x}y}{d\eta} \dots\right] = 0$$

Now of course that sum reduces to

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial y}t(x) + \frac{\partial L}{\partial D_{h,x}y}D_{h,x}[t(x)] + \frac{\partial L}{\partial D_{h,x}^{2}y}\ D_{h,x}^{2}[t(x)] \dots\right] = 0$$

Evaluation of this is not immediately obvious as it doesn't seem clear we can annihilate the terms involving $t(x)$. We consider the following rule

$$D_{h,x}[f(x)g(x)] = f(x)D_{h,x}[g(x)] + \ D_{h,x}[f(x)]g(x) + \ hD_{h,x}[f(x)]D_{h,x}[g(x)]$$

This can be verified by evaluating the definition of $D_{h,x}$ on both sides and cancelling terms. We note that this gives rise to the following formula

$$D_{h,x}[f(x)g(x)] - \ f(x)D_{h,x}[g(x)] - hD_{h,x}[f(x)]D_{h,x}[g(x)] = D_{h,x}[f(x)]g(x)$$

Which yields of course

$$f(x)g(x) - D_{h,x}^{-1}\left[f(x)D_{h,x}[g(x)]\right] = D_{h,x}^{-1}\left[D_{h,x}[f(x)]g(x) + hD_{h,x}[f(x)]D_{h,x}[g(x)]\ \right]$$

Now recall the original equation:

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial y}t(x) + \frac{\partial L}{\partial D_{h,x}y}D_{h,x}[t(x)] + \frac{\partial L}{\partial D_{h,x}^2 y}\ D_{h,x}^2[t(x)] \dots\right] = 0$$

Each term:

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial D_{h,x}^k y}\ D_{h,x}^k[t(x)]\right]$$

Can be recursively chunked into:

$$D_{h,x}^{-1}\left[D_{h,x}[f(x)]g(x) + hD_{h,x}[f(x)]D_{h,x}[g(x)]\ \right] = \ D_{h,x}^{-1}\left[\frac{\partial L}{\partial D_{h,x}^k y}\ D_{h,x}^k[t(x)]\right]$$

From here it follows that

$$D_{h,x}[f(x)]g(x) + hD_{h,x}[f(x)]D_{h,x}[g(x)] = \ \frac{\partial L}{\partial D_{h,x}^k y}\ D_{h,x}^k[t(x)]$$

We can of course let

$$D_{h,x}^k[t(x)] = D_{h,x}[f(x)]$$

From which it follows that

$$g(x) + \ h\,D_{h,x}[g(x)] = \frac{\partial L}{\partial D_{h,x}^k y}$$

Which has solution

$$g(x) = \left[\frac{\partial L}{\partial D_{h,x}^k y}\right]_{x\to x-h}$$

And then we can express the solution as

$$\frac{\partial L}{\partial D_{h,x}^{k} y}\ D_{h,x}^{k}[t(x)] = \left(\left[\frac{\partial L}{\partial D_{h,x}^{k} y}\right]_{x \to x-h} D_{h,x}^{k-1}[t(x)]\right) - \ D_{h,x}^{-1}\left[D_{h,x}\left[\frac{\partial L}{\partial D_{h,x}^{k} y}\right]_{x \to x-h} D_{h,x}^{k-1}[t(x)]\right]$$

Now we can repeat this process indefinitely taking every term of the form

$$\left(\left[\frac{\partial L}{\partial D_{h,x}^{k} y}\right]_{x \to x-h} D_{h,x}^{k-1}[t(x)]\right)$$

And noting that ultimately equates down to a constant of choice κ. Thus the expression

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial y} t(x) + \frac{\partial L}{\partial D_{h,x} y} D_{h,x}[t(x)] + \frac{\partial L}{\partial D_{h,x}^{2} y}\ D_{h,x}^{2}[t(x)] \ldots\right] = 0$$

Reduces to

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial y} t(x) - \ D_{h,x}\left[\frac{\partial L}{\partial D_{h,x} y}\right]_{x \to x-h} t(x) + \ D_{h,x}^{2}\left[\frac{\partial L}{\partial D_{h,x}^{2} y}\right]\ t(x) - \ \ldots\right] = 0$$

And if we insist the variational function is non zero of the interval in consideration then it follows that it must be the case that

$$D_{h,x}^{-1}\left[\frac{\partial L}{\partial y} - \ D_{h,x}\left[\frac{\partial L}{\partial D_{h,x} y}\right]_{x \to x-h} + \ D_{h,x}^{2}\left[\frac{\partial L}{\partial D_{h,x}^{2} y}\right]_{x \to x-2h} - \ \ldots\right] = 0$$

Which by the Fundamental Theorem of the Calculus of Variations gives

$$\frac{\partial L}{\partial y} - \ D_{h,x}\left[\frac{\partial L}{\partial D_{h,x} y}\right]_{x \to x-h} + \ D_{h,x}^{2}\left[\frac{\partial L}{\partial D_{h,x}^{2} y}\right]_{x \to x-2h} - \ D_{h,x}^{3}\left[\frac{\partial L}{\partial D_{h,x}^{3} y}\right]_{x \to x-3h} \ldots = 0$$

And this clearly reduces to Euler's result in the case of h=0, but yields a much larger and wider class of functional equations which we are now equipped to optimize.