IV. Variational Derivative

We have studied Gâteaux variation and Fréchet differential and the relationship between them. There is one more subtle variant of this, which is called variational derivative. It is useful in some applications and in proving some theorems. More importantly, it tells us an alternative way of looking at the concept of variation based purely on the techniques of ordinary calculus. In fact, it can be interpreted as the "partial derivative" equivalent for calculus of variations. As the history goes, Euler had apparently derived his eponymous necessary condition using this concept.

Let us begin with the notation. The variational derivative of a functional $J = \int_{x_0}^{x_f} F(x, y, y')dx$ is denoted as $\frac{\delta J}{\delta y}$ and is given by

$$\frac{\delta J}{\delta y} = F_y - \frac{d}{dx}(F_{y'}) \tag{1}$$

You can immediately see that it is nothing but the E-L expression that should be zero. When J has a more general form, the expression for $\frac{\delta J}{\delta y}$ will be the corresponding expression in the E-L equation that we equate to zero. Let us see what rationale underlies this definition.

Because we want to use only the techniques of ordinary calculus, let us "discretize" $y(x)$ and consider finitely many discrete points $x_k\left(k = 1, 2, ..., N\right)$ within the interval $\left(x_0, x_f\right)$. See Fig. 1. As can be seen in this figure, by way of discretization, we are approximating the continuous curve of $y(x)$ by a polygon.

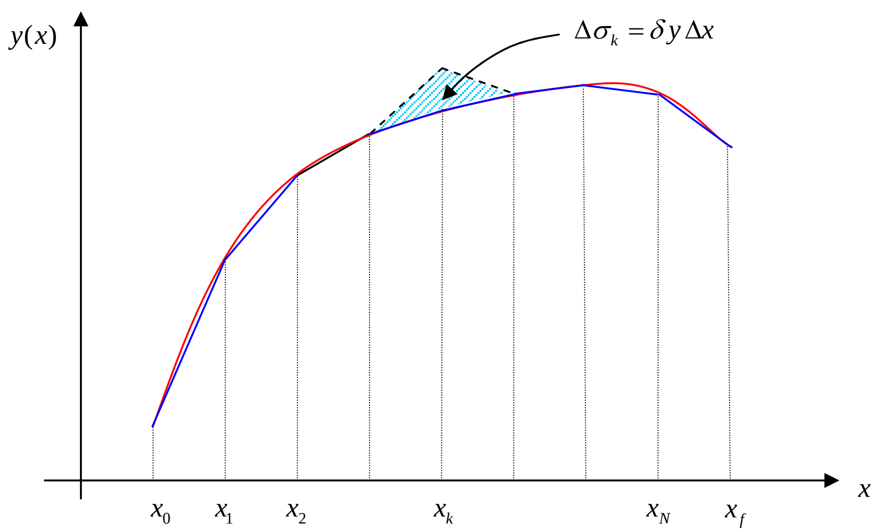

Figure 1. Discretization of a continuous curve $y(x)$ by a polygon. All subdivisions on the x-axis are equal to Δx. A local perturbation at x_k is considered and its effect is shown with the dashed lines.

Now, the functional can be approximated as follows.

$$\sum_{k=1}^{N} \quad J \approx J_N = \sum_{k=1}^{N} F\left(x_k, y_k, \frac{(y_{k+1}-y_k)}{(x_{k+1}-x_k)} \right)(x_{k+1}-x_k) = \sum_{k=1}^{N} F\left(x_k, y_k, \frac{(y_{k+1}-y_k)}{(x_{k+1}-x_k)} \right)\Delta x \quad (2)$$

where we have assumed that all subdivisions along the x-axis are equal to Δx. Our variables to minimize J_N are now $\{y_1, y_2, \cdots, y_N\}$. Consider the partial derivative of J_N with respect to y_k.

$$\frac{\partial J_N}{\partial y_k} = F_y(x_k, y_k, \frac{(y_{k+1}-y_k)}{\Delta x})\Delta x + F_{y'}(x_{k-1}, y_{k-1}, \frac{(y_k - y_{k-1})}{\Delta x}) - F_{y'}(x_k, y_k, \frac{(y_{k+1}-y_k)}{\Delta x}) \quad (3)$$

Here, we have only used the chain rule of differentiation. As $\Delta x \to 0$, the RHS of Eq. (3) goes to zero. Now, divide the LHS and RHS of Eq. (3) by Δx to get

$$\frac{\partial J_N}{\partial y_k \Delta x} = F_y(x_k, y_k, \frac{(y_{k+1}-y_k)}{\Delta x}) + \frac{F_{y'}(x_{k-1}, y_{k-1}, \frac{(y_k - y_{k-1})}{\Delta x}) - F_{y'}(x_k, y_k, \frac{(y_{k+1}-y_k)}{\Delta x})}{\Delta x} \quad (4)$$

When $\Delta x \to 0$, $\partial y_k \Delta x$, which can be interpreted as the shaded area in Fig. 1, also tends to zero. In fact, we then denote $\partial y_k \Delta x$ as $\Delta\sigma_k$ or, in general, simply as $\delta y \Delta x$ evaluated at $x = x_k$. Furthermore, as $\Delta x \to 0$, $J_N \to J$. We take the limit of Eq. (4) as $\Delta x \to 0$.

$$\lim_{\Delta x \to 0} \frac{\partial J_N}{\partial y_k \Delta x} = \frac{\delta J}{\delta y} = F_y - \frac{d}{dx}\left(F_{y'}\right) \quad (5)$$

Notice how we defined the variational derivative in Eq. (5). We can think of $\frac{\delta J}{\delta y}$ as the limiting case of $\frac{J(y+h) - J(y)}{\Delta\sigma}$ where h is the perturbation (i.e., variation) of y at some $\hat{x}$ and $\Delta\sigma$ is the extra area under $y(x)$ due to that perturbation. Therefore, we write

$$\Delta J = J(y+h) - J(y) = \left\{ \left. \frac{\delta J}{\delta y} \right|_{x=\hat{x}} + \varepsilon \right\} \Delta\sigma \quad (6)$$

where ε is a small discretization error. When the discretization error is insignificantly small, we can write

$$\Delta J \approx \left. \frac{\delta J}{\delta y} \right|_{x=\hat{x}} \Delta\sigma \quad (7)$$

Thus, the variational derivative helps us get the first order change in the value of the functional for a local perturbation of $y(x)$ at $x = \hat{x}$. Think of Taylor series of expansion of a function of many variables and try to relate this concept of first order change in the value of the function.