Math Notes

Tyler Wilson

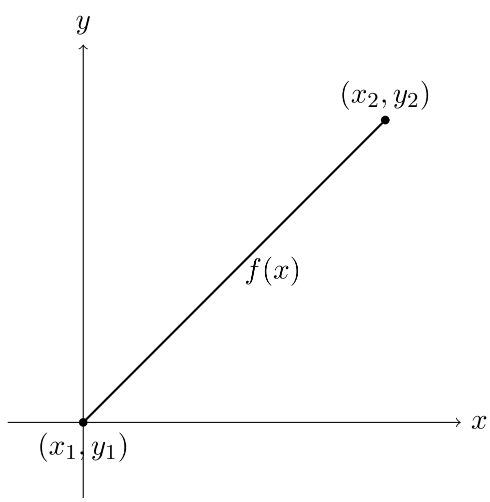
Contents

1 Variational Calculus

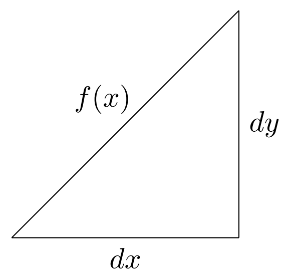
1.1 Variational Derivatives

1.1.1 Definition of the Variational Derivative

An action is a function of a function called a functional. This is best seen by example.
Using variational calculus we can prove that the shortest distance between two points is a straight line.
Let us define two points in $\mathbb{R}^2$, (x_1, y_1) and (x_2, y_2) and some function $f(x)$ that connects these points. This can be generalized to higher dimensions, but for simplicity we will stick to two dimensions for now.



We can define an infinite possible functions that connect these two points. We want to find the specific function that minimizes the distance between these two points. So we will want a function with a minimum arc length.



We can use the Pythagorean theorem to find the arc length of this function.

$$
\begin{aligned}
&\frac{dy}{dx} = f'(x) \Rightarrow dy = f'(x)dx \\
&ds^2 = dx^2 + dy^2 \\
&ds = \sqrt{dx^2 + dy^2} \\
&ds = \sqrt{dx^2 + (f'(x)dx)^2} \\
&ds = \sqrt{1 + f'(x)^2}dx \\
&S = \int_{x_1}^{x_2} dx\sqrt{1 + f'(x)^2}, \ \begin{cases} f(x_1) = y_1 \\ f(x_2) = y_2 \end{cases}
\end{aligned}
$$

We define S, the arc length, as the action that we wish to minimize. It is also considered a functional as it is defiend as $S(f(x))$.
We define $\delta f(x)$ as some small change (or wiggle) in the function $f(x)$. It is important to note that $\delta f(x)$ is not some opperation on the original function $f(x)$ but is rather some new function of x we call $\delta f(x)$ that contains very slight variations. We can then define the variation of the action as follows.
Recall that for a regular function $f(x)$, the derivative is defined as

$$
\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}
$$

We can define the variational derivative in a similar manner.

$$
\frac{\delta S}{\delta f(x)} = \lim_{\delta f(x) \to 0} \frac{S(f(x) + \delta f(x)) - S(f(x))}{\delta f(x)}
$$

Note that we usually just deal with the numerator of this expression and call it δS. We can also introduce a small parameter ϵ to make the expression easier to follow.

$$
\delta S = \lim_{\epsilon \to 0} \frac{S(f(x) + \epsilon\delta f(x)) - S(f(x))}{\epsilon}
$$

We can now use this definition to find the variation of the action for our example.

$$
\begin{aligned}
&\delta S = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_1}^{x_2} dx\sqrt{1 + (f'(x) + \epsilon\delta f'(x))^2} - \int_{x_1}^{x_2} dx\sqrt{1 + f'(x)^2} \\
&\delta S = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_1}^{x_2} dx \left(\sqrt{1 + (f'(x) + \epsilon\delta f'(x))^2} - \sqrt{1 + f'(x)^2}\right)
\end{aligned}
$$

$$\delta S = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_1}^{x_2} dx \left(\sqrt{1 + f'(x)^2 + 2f'(x)\epsilon\delta f'(x) + \epsilon^2 \delta f'(x)^2} - \sqrt{1 + f'(x)^2} \right)$$

Because ϵ is a very small we can perform a Taylor expansion on the square root.

$$\sqrt{1+x}\bigg|_{x\approx 0} = 1 + \frac{x}{2} + \mathcal{O}(x^2)$$

$$\sqrt{1 + f'(x)^2 + 2f'(x)\epsilon\delta f(x) + \epsilon^2 f'(x)^2} = \sqrt{1 + f'(x)^2}\sqrt{1 + \frac{2f'(x)\epsilon\delta f(x) + \epsilon^2 f'(x)^2}{1 + f'(x)^2}}$$

$$= \sqrt{1 + f'(x)^2}\left(1 + \frac{1}{2}\frac{2f'(x)\epsilon\delta f(x) + \epsilon^2 f'(x)^2}{1 + f'(x)^2}\right) + \mathcal{O}\left(\left(\frac{2f'(x)\epsilon\delta f(x) + \epsilon^2 f'(x)^2}{1 + f'(x)^2}\right)^2\right)$$

$$= \sqrt{1 + f'(x)^2}\left(1 + \epsilon\frac{f'(x)\delta f(x)}{1 + f'(x)^2} + \mathcal{O}(\epsilon^2)\right)$$

$$= \sqrt{1 + f'(x)^2} + \epsilon\frac{f'(x)\delta f(x)}{\sqrt{1 + f'(x)^2}} + \mathcal{O}(\epsilon^2)$$

Plugging this back into the original expression we get

$$\delta S = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_1}^{x_2} dx \left(\sqrt{1 + f'(x)^2} + \epsilon\frac{f'(x)\delta f'(x)}{\sqrt{1 + f'(x)^2}} + \mathcal{O}(\epsilon^2) - \sqrt{1 + f'(x)^2} \right)$$

$$\delta S = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_1}^{x_2} dx \left(\epsilon\frac{f'(x)\delta f'(x)}{\sqrt{1 + f'(x)^2}} + \mathcal{O}(\epsilon^2) \right)$$

$$\delta S = \lim_{\epsilon \to 0} \int_{x_1}^{x_2} dx \left(\frac{f'(x)\delta f'(x)}{\sqrt{1 + f'(x)^2}} + \mathcal{O}(\epsilon) \right)$$

$$\delta S = \lim_{\epsilon \to 0} \int_{x_1}^{x_2} dx \frac{f'(x)\delta f'(x)}{\sqrt{1 + f'(x)^2}} + \lim_{\epsilon \to 0} \int_{x_1}^{x_2} dx \mathcal{O}(\epsilon)$$

If we take the limit as $\epsilon \to 0$ then the $\mathcal{O}(\epsilon)$ term goes to zero and we are left with

$$\delta S = \int_{x_1}^{x_2} dx \frac{f'(x)\delta f'(x)}{\sqrt{1 + f'(x)^2}}$$

Using integration by parts we can rewrite this as

$$\delta S = \delta f(x)\frac{f'(x)}{\sqrt{1 + f'(x)^2}}\bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} dx \delta x \left(\frac{f'(x)}{\sqrt{1 + f'(x)^2}}\right)'$$

Note that the first term goes to zero because $\delta f(x_1) = \delta f(x_2) = 0$ as we cannot have any variation in the start and end points. We can then rewrite this as

$$\delta S = -\int_{x_1}^{x_2} dx \delta x \left(\frac{f'(x)}{\sqrt{1 + f'(x)^2}}\right)'$$

Because we are trying to find a minimum for the action, we can set $\delta S = 0$ which gives us

$$\int_{x_1}^{x_2} dx \delta x \left(\frac{f'(x)}{\sqrt{1+f'(x)^2}} \right)' = 0$$
$$\frac{d}{dx} \left(\frac{f'(x)}{\sqrt{1+f'(x)^2}} \right) = 0$$

We now have an expression only in terms of $f(x)$ which we can use to solve for the function that minimizes the action S.

$$\frac{f'(x)}{\sqrt{1+f'(x)^2}} = C$$
$$f'(x) = C\sqrt{1+f'(x)^2}$$
$$f'(x)^2 = C^2(1+f'(x)^2)$$
$$f'(x)^2 - C^2 f'(x)^2 = C^2$$
$$(1-C^2)f'(x)^2 = C^2$$
$$f'(x) = \pm\frac{C}{\sqrt{1-C^2}} = C_1$$
$$f(x) = C_1 x + C_2$$

And so we have proved that $f(x)$ is a linear function and so the shortest path between two points in $\mathbb{R}^2$ is a straight line.

Note that we can avoid the Taylor expansions for future problems by using the following identity. If we recall for functions of a single variable we have

$$df = \lim_{x \to 0} f(x + \Delta x) - f(x) = \frac{df}{dx} dx = f'(x) dx$$

We can utilize this identity for functions of many variables as well. If we have a function $f(x_1, x_2, \ldots, x_n)$ then we can write

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \vec{\nabla} f \cdot d\vec{x}$$

We can also extend this trick to functionals. If we have a functional $S[f(x)]$ then we can write

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

If we have a functional of multiple functions then we can extend this further using the multivariable rule to get

$$\delta[S(f_1, f_2, \ldots, f_n)] = \sum_{i=1}^{n} \frac{\delta S}{\delta f_i} \delta f_i$$

1.1.2 Functions of Many Variables

Intuitively the previous example should hold in higher dimensions as well (we know this to be true in $\mathbb{R}^3$). To show this we will use the same method as before but with a function of many variables. Let us define some path $\vec{p}(t)$ in $\mathbb{R}^n$ where $\vec{p}(t)$ is a vector valued function such that $\vec{p}(t) \in \mathbb{R}^n$. (In $\mathbb{R}^3$ it would be analogous to $\vec{p}(t) = \langle x(t), y(t), z(t) \rangle$). We can define the start and end points as $\vec{p}(t_1) = \vec{x}_1$ and $\vec{p}(t_2) = \vec{x}_2$. We can then define the action S once again to be the arc length of the path $\vec{p}(t)$.

$$ds^2 = dp_1^2 + dp_2^2 + \cdots + dp_n^2 = \sum_{i=1}^{n} dp_i^2$$

$$ds = \sqrt{\sum_{i=1}^{n} dp_i^2} = dt\sqrt{\sum_{i=1}^{n} \left(\frac{dp_i}{dt}\right)^2}$$

$$S = \int_{t_1}^{t_2} dt\sqrt{\sum_{i=1}^{n} \left(\frac{dp_i}{dt}\right)^2}$$

For compactness we will define the following notation

$$\frac{dp_i}{dt} = \dot{p}_i$$

We can then rewrite the action as

$$S = \int_{t_1}^{t_2} dt\sqrt{\sum_{i=1}^{n} \dot{p}_i^2}$$

Note that the derivative of a function is a different function than the original function so the action S will be a functional of $S[\dot{p}_1, \dot{p}_2, \ldots, \dot{p}_n]$. We can then define the variation of the action δS as

$$\delta S = \sum_{i=1}^{n} \frac{\delta S}{\delta \dot{p}_i} \delta \dot{p}_i$$

$$\delta S = \sum_{i=1}^{n} \int_{t_0}^{t_1} dt \frac{\delta}{\delta \dot{p}_i} \left(\sqrt{\sum_{j=1}^{n} \dot{p}_j^2}\right) \delta \dot{p}_i$$

$$\delta S = \sum_{i=1}^{n} \int_{t_0}^{t_1} dt \left(\frac{\dot{p}_i}{\sqrt{\sum_{j=1}^{n} \dot{p}_j^2}}\right) \delta \dot{p}_i$$

Now using integration by parts we can rewrite this as

$$\delta S = \sum_{i=1}^{n} \left(\frac{\dot{p}_i}{\sqrt{\sum_{j=1}^{n} \dot{p}_j^2}} \delta p_i \Bigg|_{t_0}^{t_1} - \int_{t_0}^{t_1} dt \frac{d}{dt} \left(\frac{\dot{p}_i}{\sqrt{\sum_{j=1}^{n} \dot{p}_j^2}}\right) \delta p_i \right)$$

Once again the first term will go to zero because $\delta p_i(t_0) = \delta p_i(t_1) = 0$. We will also set δS to zero again to minimize the action.

$$\sum_{i=1}^{n} \int_{t_0}^{t_1} dt \frac{d}{dt} \left(\frac{\dot{p}_i}{\sqrt{\sum\limits_{j=1}^{n} \dot{p}_j^2}} \right) \delta p_i = 0$$

$$\sum_{i=1}^{n} \frac{d}{dt} \left(\frac{\dot{p}_i}{\sqrt{\sum\limits_{j=1}^{n} \dot{p}_j^2}} \right) = 0$$

This will give us n equations of the form

$$\frac{\dot{p}_i}{\sqrt{\sum\limits_{j=1}^{n} \dot{p}_j^2}} = C_i$$

$$\dot{p}_i = C_i \sqrt{\sum_{j=1}^{n} \dot{p}_j^2}$$

$$\dot{p}_i^2 = C_i^2 \sum_{j=1}^{n} \dot{p}_j^2$$

Note that all of the functions $\dot{p}_i$ are linearly independent of one another so we can get another system of n equations from each of these equations. One of the form

$$\dot{p}_i(1 - C_i^2) = 0 \Rightarrow \dot{p}_i = 0 \Rightarrow p_i(t) = C_i t + D_i$$

All others of the form

$$C_i^2 \dot{p}_j^2 = 0 \Rightarrow \dot{p}_j = 0 \Rightarrow p_j(t) = C_j t + D_j$$

This implies that the solution if of the form

$$\vec{p}(t) = \vec{C} t + \vec{D}$$

which forms a straight line in n dimensional space.

If we have a problem such that we're trying to optimize something subject to some constraint then we can use the method of Lagrange multipliers. Let us say we have some function $f(x, y)$ that we want to optimize subject to the constraint $g(x, y) = C$. We can then define the action as $L(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - C)$.

One such example of this to find the shape that forms the maximum area for a given perimeter. We can get a generalized expression for the area by utilizing Stoke's theorem.

$$\oint_{\partial R} \vec{F} \cdot d\vec{l} = \iint_R (\nabla \times \vec{F}) \cdot d\vec{a}$$
$$\nabla \times \vec{F} = \hat{n} \Rightarrow \iint_R (\nabla \times \vec{F}) \cdot d\vec{a} = A$$
$$\Rightarrow \vec{F} = \frac{1}{2} \langle -y, x \rangle$$
$$\vec{l} = \langle x(s), y(s) \rangle \Rightarrow d\vec{l} = \langle \dot{x}, \dot{y} \rangle \, ds$$
$$A = \oint ds \frac{1}{2} \langle -y, x \rangle \cdot \langle \dot{x}, \dot{y} \rangle = \frac{1}{2} \oint ds \, (x\dot{y} - y\dot{x})$$

The length of the curve our constant and is given by

$$L = 2\pi r = \int_0^{2\pi r} dl = \int_0^{2\pi r} ds \sqrt{\dot{x}^2 + \dot{y}^2}$$

Our action that we will want to maximize is then given by

$$\mathscr{L}(x, \dot{x}, y, \dot{y}) = A - \lambda L = \int_0^{2\pi r} ds \left(\frac{1}{2}(x\dot{y} - y\dot{x}) + \lambda \sqrt{\dot{x}^2 + \dot{y}^2} \right)$$
$$\delta\mathscr{L} = \int_0^{2\pi r} ds \left(\frac{1}{2} \left(\dot{y}\delta x - \dot{x}\delta y + x\dot{\delta y} - y\dot{\delta x} \right) + \lambda \frac{\dot{x}\dot{\delta x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} + \lambda \frac{\dot{y}\dot{\delta y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \right)$$
$$\delta\mathscr{L} = \int_0^{2\pi r} ds \left(\frac{1}{2} (\dot{y}\delta x - \dot{x}\delta y - \dot{x}\delta y + \dot{y}\delta x) - \lambda \frac{d}{dt} \left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \right) \delta x - \lambda \frac{d}{dt} \left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \right) \delta y \right)$$
$$\left(\dot{y} - \lambda \frac{d}{dt} \left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \right) \right) \delta x = 0 \Rightarrow y + C_1 = \frac{\lambda\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$
$$\left(-\dot{x} - \lambda \frac{d}{dt} \left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} \right) \right) \delta y = 0 \Rightarrow -x + C_2 = \frac{\lambda\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$
$$(y + C_1)^2 + (x + C_2)^2 = \frac{\lambda^2 \dot{x}^2}{\dot{x}^2 + \dot{y}^2} + \frac{\lambda^2 \dot{y}^2}{\dot{x}^2 + \dot{y}^2} = \lambda^2$$

The constants C_1 and C_2 are just translations in x and y so if we center our origin at that point we can take $C_1 = C_2 = 0$ for simplicity. This gives the equation for a circle of radius λ

$$x^2 + y^2 = \lambda^2$$

We can get λ from our constraint equation. The path length of a circle with radius λ works out to

$$s = 2\pi\lambda = 2\pi r$$

So we see that $\lambda = r$ and we get the equation

$$x^2 + y^2 = r^2$$

The area of this will be the area of a circle which is

$$A = \pi r^2$$

So far we have used the integration by parts trick quite a few times. There is a more general form of this trick called the Euler-Lagrange equation.
If we have some action that is a functional of the form $S[f_1(t), f_2(t), \dots, \dot{f}_1(t), \dot{f}_2(t), t]$ then we can express the variation of the action as

$$\delta S = \sum_i \int_a^b dt \left(\frac{\partial S}{\partial f_i} \delta f_i + \frac{\partial S}{\partial \dot{f}_i} \delta \dot{f}_i \right)$$
$$\delta S = \sum_i \int_a^b dt \left(\frac{\partial S}{\partial f_i} - \frac{d}{dt}\left(\frac{\partial S}{\partial \dot{f}_i} \right) \right) \delta f_i + \sum_i \frac{\partial S}{\partial \dot{f}_i} \delta f_i(t) \bigg|_a^b$$
$$\delta S = \sum_i \int_a^b dt \left(\frac{\partial S}{\partial f_i} - \frac{d}{dt}\left(\frac{\partial S}{\partial \dot{f}_i} \right) \right) \delta f_i$$

This equation must equal 0 for all i so we will have n equations of the form

$$\int_a^b dt \left(\frac{\partial S}{\partial f_i} - \frac{d}{dt}\left(\frac{\partial S}{\partial \dot{f}_i} \right) \right) = 0$$
$$\frac{\partial S}{\partial f_i} - \frac{d}{dt}\left(\frac{\partial S}{\partial \dot{f}_i} \right) = 0$$

Rearranging gives us the Euler-Lagrange equation

$$\frac{\partial S}{\partial f_i} = \frac{d}{dt}\left(\frac{\partial S}{\partial \dot{f}_i} \right)$$

In many cases we can avoid many of the steps outlined above and simply plug in our action to the Euler-Lagrange equation and solve for the function we are interested in.