Saul
Youssef

Calculus ... (of variations)

... as an application of topology.

How can we interpret the differential

$$\text{“} df_x(h) = \lim_{\lambda \to 0} \frac{f(x+\lambda h) - f(x)}{\lambda} \text{”} \; ?$$

Let's assume

$$f : X \longrightarrow Y$$

Hausdorff real topological vector space.

continuous function

Connected real topological vector space, dim > 0.

<u>def</u>: A real topological vector space V is both a real vector space and a topological space where $+ : V \times V \to V$ and $\cdot : \mathbb{R} \times V \to V$ are continuous.

Let's interpret "$\lim_{x \to a} f(x) = y$" as an assertion that the function

$$x \mapsto \begin{cases} f(x) & \text{if } x \neq a \\ y & \text{if } x = a \end{cases}$$

is continuous.

The extra assumptions about X and Y are there to guarantee that limits are unique.

Thm: Such limits are unique

Proof: Suppose $\lim_{x \to a} f(x) = y$, $\lim_{x \to a} f(x) = y'$. Then

$$\begin{cases} f(x) & \text{if } x \neq a \\ y & \text{if } x = a \end{cases} - \begin{cases} f(x) & \text{if } x \neq a \\ y' & \text{if } x = a \end{cases} = \begin{cases} 0 & \text{if } x \neq a \\ y - y' & \text{if } x = a \end{cases}.$$

is continuous. Assume $y \neq y'$. Because Y is Hausdorff, we have $\mathcal{O}_{y-y'} \cap \mathcal{O}_0 = \emptyset$. Then $f^{-1}[\mathcal{O}_{y-y'}] = \{a\}$ is both open and closed. Since X is connected, $X = \{a\}$. Since $\dim(X) > 0$ $\Rightarrow\Leftarrow$.

Thus $y = y'$.

Notice how the properties of continuous functions are doing the work for us:

Thm: If $\lim_{x \to a} f(x) = y$, $\lim_{x \to a} g(x) = z$, then $\lim_{x \to a} (f+g)(x) = y + z$.

Proof:

$$\begin{cases} f(x) & \text{if } x \neq a \\ y & \text{if } x = a \end{cases} + \begin{cases} g(x) & \text{if } x \neq a \\ z & \text{if } x = a \end{cases} = \begin{cases} f(x) + g(x) & \text{if } x \neq a \\ y + z & \text{if } x = a \end{cases}.$$

Thm: If $X \xrightarrow{f} Y \xrightarrow{g} Z$ are continuous and $\lim\limits_{x \to a} f(x) = y$, then

$$\lim_{x \to a} g(f(x)) = g\left(\lim_{x \to a} f(x)\right).$$

Proof:

$$g \circ \begin{cases} f(x) & \text{if } x \neq a \\ y & \text{if } x = a \end{cases} = \begin{cases} g(f(x)) & \text{if } x \neq a \\ g(y) & \text{if } x = a \end{cases}$$

is continuous.

Compare how easy this is with your favorite calculus book!

Now we will treat the differential similarly and see what we get...

def: Given $f : X \to Y$ as before, $x \in X$. If there is a function $df_x : X \to Y$ such that

$$\lambda, h \mapsto \begin{cases} \dfrac{f(x+\lambda h) - f(x)}{\lambda} & \text{if } \lambda \neq 0 \\ df_x(h) & \text{if } \lambda = 0 \end{cases} \qquad (\equiv F_x(\lambda, h))$$

is continuous, then f is differentiable at x with differential df_x.

Thm: $df_x(a \cdot h) = a \cdot df_x(h)$.

Proof: If $a = 0$, $df_x(0) = 0$ by the same argument as for the uniqueness of limits. If $a \neq 0$,

$$\frac{1}{a} F_x(\lambda/a, a \cdot h) = \begin{cases} \dfrac{f(x+\lambda h) - f(x)}{\lambda} & \text{if } \lambda \neq 0 \\ df_x(a \cdot h)/a & \text{if } \lambda = 0 \end{cases}$$

is continuous. Thus, $df_x(a \cdot h) = a \cdot df_x(h)$.

Thm: $d(f+g)_x = df_x + dg_x$.

Proof: Let

$$G_x(\lambda, h) \equiv \begin{cases} \dfrac{g(x+\lambda h) - g(x)}{\lambda} & \text{if } \lambda \neq 0 \\ dg_x(h) & \text{if } \lambda = 0 \end{cases}.$$

$F_x(\lambda, h) + G_x(\lambda, h)$ is continuous Q.E.D.

Therem (Chain Rule): If $f : X \to Y$ is differentiable at $x \in X$ and $g : Y \to Z$ is differentiable at $f(x) \in Y$, then $d(g \circ f)_x = dg_{f(x)} \circ df_x$.

Proof: The function

$$\lambda, h \longmapsto G_{f(x)}(\lambda, F_x(\lambda, h))$$

is continuous, $\Rightarrow$

$$\lambda, h \mapsto \begin{cases} \dfrac{g(f(x) + \lambda F_x(\lambda, h)) - g(f(x))}{\lambda} & \text{if } \lambda \neq 0 \\ dg_{f(x)}(F_x(\lambda, h)) & \text{if } \lambda = 0 \end{cases}$$

is continuous, $\Rightarrow$

$$\lambda, h \mapsto \begin{cases} \dfrac{g(f(x) + f(x + \lambda h) - f(x)) - g(f(x))}{\lambda} & \text{if } \lambda \neq 0 \\ dg_{f(x)}(df_x(h)) & \text{if } \lambda = 0 \end{cases}$$

is continuous, $\Rightarrow$ $d(g \circ f)_x = dg_{f(x)} \circ df_x$.

A pair of helper theorems:

Thm: If f is differentiable at $x \in X$ and $R(h)$ is defined by $f(x+h) = f(x) + df_x(h) + R(h)$, then R is continuous and $\lim_{\lambda \to 0} R(\lambda h)/\lambda = 0$.

Proof:

$$\begin{cases} \dfrac{df_x(\lambda h) + R(\lambda h)}{\lambda} & \text{if } \lambda \neq 0 \\ df_x(h) & \text{if } \lambda = 0 \end{cases} \quad \text{is continuous}$$

$$\Rightarrow \quad \begin{cases} R(\lambda h)/\lambda & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0 \end{cases} \quad \text{is continuous} \Rightarrow \lim_{\lambda \to 0} R(\lambda h)/\lambda = 0.$$

Thm: If $f(x+h) = f(x) + D(h) + R(h)$ where R is continuous, $D(\lambda h) = \lambda D(h)$ and $\lim_{\lambda \to 0} R(\lambda h)/\lambda = 0$, then f is differentiable at x with ~~df~~ $df_x(h) = D(h)$.

Proof:

$$D(h) + \begin{cases} \dfrac{R(\lambda h)}{\lambda} & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0 \end{cases} \quad \text{is continuous}$$

$$= \begin{cases} \dfrac{D(\lambda h) + R(\lambda h)}{\lambda} & \text{if } \lambda \neq 0 \\ D(h) & \text{if } \lambda = 0 \end{cases} = \begin{cases} \dfrac{f(x+\lambda h) - f(x)}{\lambda} & \text{if } \lambda \neq 0 \\ D(h) & \text{if } \lambda = 0 \end{cases}.$$

Specialize to $f : X \to Y$ with $Y = \mathbb{R}$.

<u>Thm</u> $d(f\cdot g)_x(h) = f(x)\,dg_x(h) + g(x)\,df_x(h)$.

<u>Proof</u>: Let

$$f(x+h) = f(x) + df_x(h) + F(h)$$

$$g(x+h) = g(x) + dg_x(h) + G(h).$$

$$f(x+h)g(x+h) - f(x)g(x) =$$

$$\left.\begin{array}{l} f(x)\,dg_x(h) + g(x)\,df_x(h) \end{array}\right\} \quad A(\lambda\cdot h) = \lambda A(h)$$

$$\left.\begin{array}{l} + f(x)\,G(h) + g(x)\,F(h) \\ + df_x(h)\,G(h) + dg_x(h)\,F(h) \end{array}\right\} \quad \lim_{\lambda\to 0} B(\lambda h)/\lambda = 0$$

$$+ df_x(h)\,dg_x(h) + F(h)G(h)$$

e.g.

$$\begin{cases} F(\lambda h) & \text{if } \lambda \neq 0 \\ F(\lambda h) & \text{if } \lambda = 0 \end{cases} \cdot \begin{cases} G(\lambda h)/\lambda & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0 \end{cases}$$

$$= \begin{cases} F(\lambda h)G(\lambda h)/\lambda & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0 \end{cases} \quad \text{is continuous}$$

$$\Rightarrow \lim_{\lambda\to 0} F(\lambda h)G(\lambda h)/\lambda = 0.$$

$$\Rightarrow d(f\circ g)_x = f(x)\,dg_x(h) + g(x)\,df_x(h).$$

Thm: Suppose that $f: X \to \mathbb{R}$ has a minimum at $m \in X$ and df_m exists then $df_m = 0$.

Proof: The continuity of :

$$\left\{ \begin{array}{lll} \frac{f(m+\lambda h)-f(m)}{\lambda} & \text{if } \lambda \neq 0 & \} \geq 0 \text{ for all } h \in X \\ df_m(h) & \text{if } \lambda = 0 & \} \Rightarrow \geq 0 \text{ for all } h \end{array} \right.$$

guarantees that $df_m(h) \geq 0$ for all $h \in X$, by the same argument as in the proof that limits are unique. Suppose $df_m(h) > 0$ for some h. Then $df_m(-h)$ is negative $\Rightarrow \Leftarrow$. Thus $df_m(h) = 0$.

SUMMARY

$$\underset{\text{continuous}}{f} : \underset{\text{connected real TVS}}{X} \to Y \quad \text{Hausdorff real TVS}$$

- Limits are unique
- $\lim_{x \to a} (f + g) = \lim_{x \to a} f + \lim_{x \to a} g$
- $\lim_{x \to a} g \circ f = g\left(\lim_{x \to a} f \right)$
- $df_x(a \cdot h) = a \cdot df_x(h)$
- $d(f+g)_x = df_x + dg_x$
- $d(g \circ f)_x = dg_{f(x)} \circ df_x$ Chain Rule

$Y = \mathbb{R}$ $\left\{ \right.$

- $d(f \cdot g)_x = f(x)\, dg_x + g(x)\, df_x$ Leibniz Rule
- Differentiable f has a minimum at $m \Rightarrow df_m = 0$.

This is all very easily proved and is familiar, but what about

$$df(h + h') = df(h) + df(h') \; ?$$

This turns out to be FALSE! ...

Counter-Example:

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) \equiv \begin{cases} \dfrac{x^3}{x^2+y^2} & \text{if } (x,y) \neq 0 \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

f is differentiable at $(0,0)$ with

$$df_{(0,0)}(h_x, h_y) = \frac{h_x^3}{h_x^2 + h_y^2} \quad \text{which is not linear in } h!$$

What does this mean?

- Note that our differential works on infinite dimensional function spaces as well as on $X = \mathbb{R}^n$. This is more-n-less the "Gateaux differential" in the calculus of variations literature.
- Since we have almost everything we need in the more general setting, we can just restrict ourselves to linear differentials for ordinary diff-top applications.

- Partial derivatives

Restricting ourselves to $X = \mathbb{R}^n$ and linear differentials, the partial derivatives of a map $f : \mathbb{R}^n \to \mathbb{R}$ are just defined to be the coefficients of df_p in the standard basis $\{dx_1, dx_2, \dots, dx_n\} \in (\mathbb{R}^n)^*$.

For example, for $f : \mathbb{R}^3 \to \mathbb{R}$,

$$df_p \equiv \frac{\partial f}{\partial x}(p)\, dx + \frac{\partial f}{\partial y}(p)\, dy + \frac{\partial f}{\partial z}(p)\, dz$$

so that

$$df_p(\epsilon_x) \equiv \frac{\partial f}{\partial x}(p) = \lim_{\lambda \to 0} \frac{f(p + \lambda \epsilon_x) - f(p)}{\lambda} .$$

In the case $f : \mathbb{R} \to \mathbb{R}$,

$$df_p(1) = \frac{\partial f}{\partial x}(p) \equiv \frac{df}{dx}(p) = \lim_{\lambda \to 0} \frac{f(p + \lambda) - f(p)}{\lambda}$$

as usual.

Example: Suppose

$$\left.\begin{aligned} p &: \mathbb{R} \to \mathbb{R}^n \\ q &: \mathbb{R} \to \mathbb{R}^n \\ f &: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \end{aligned}\right\} \text{ continuous and differentiable}$$

$$F(t) \equiv f(p(t), q(t))$$

what is $\left.\frac{dF}{dt}\right|_{t=0}$?

$F = f \circ \alpha$ where $\alpha(t) \equiv (p(t), q(t))$

$$\left.\frac{dF}{dt}\right|_{t=0} = dF_0(1) = d(f \circ \alpha)_0(1) = df_{\alpha(0)} \circ d\alpha_0(1)$$

$$= df_{\alpha(0)}(\dot{p}(0), \dot{q}(0))$$

$$= \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(p(0), q(0))\, \dot{p}_i(0) + \sum_{i=1}^{m} \frac{\partial f}{\partial y_i}(p(0), q(0))\, \dot{q}_i(0).$$

Example: The Euler-Lagrange Equation

Given a continuous differentiable "Lagrangian"

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

such as

$$L(u,v) \equiv \frac{1}{2} m \langle v, v \rangle - \phi(u),$$

the action of a trajectory $u : [a,b] \to \mathbb{R}^n$ is

$$S(u) \equiv \int_a^b L(u, \dot{u})\, dt$$

Let $V \equiv \{ h : [a,b] \to \mathbb{R}^n$ such that $h(a) = h(b) = 0 \}$ be continuous and differentiable. The "principle of least action is

$$\frac{d}{d\lambda} S(u + \lambda h) \Big|_{\lambda = 0} = 0 \quad \text{for all } h \in V.$$

Thus,

$$\frac{d}{d\lambda} S(u+\lambda h)\Big|_{\lambda=0} = 0 = \int_a^b \frac{d}{d\lambda} L(u+\lambda h, \dot{u} + \lambda \dot{h})\, dt.$$

Use the previous example with $p : \lambda \mapsto u + \lambda h$

$q : \lambda \mapsto \dot{u} + \lambda \dot{h}$

$\Rightarrow$

$$0 = \int_a^b \left(\sum_{i=1}^n \frac{\partial L}{\partial u_i}(u(t),\dot u(t))\, h(t)_i + \frac{\partial L}{\partial v_i}(u(t),\dot u(t))\, \dot h(t)_i \right) dt$$

$$= \int_a^b \sum_{i=1}^n \left[\frac{\partial L}{\partial u_i}(u(t),\dot u(t)) - \frac{d}{dt}\frac{\partial L}{\partial v_i}(u(t),\dot u(t)) \right] h(t)_i \, dt$$

using integration by parts (see Taylor's theorem notes),

<u>Lemma</u>: For continuous $W : [a,b] \to \mathbb{R}^n$, if $\int_a^b \langle W(t), h(t) \rangle \, dt = 0$ for all $h \in V$, then $W(t) = 0$ for all $t \in [a,b]$.

<u>Proof</u>: Choose $\eta : [a,b] \to \mathbb{R}$ such that $\eta(a) = 0$, $\eta(b) = 0$, but, otherwise, $\eta(t) > 0$. Then $W(t) \cdot \eta(t) \in V \Rightarrow$
$\int_a^b \langle W(t), W(t) \rangle \, \eta(t) = 0 \Rightarrow W(t) = 0$.

Thus, we have that the least action $u : [a,b] \to \mathbb{R}^n$ satisfies

$$\frac{\partial L}{\partial u_i}(u(t),\dot u(t)) - \frac{d}{dt}\frac{\partial L}{\partial v_i}(u(t),\dot u(t)) = 0$$

for $i = 1, 2, \ldots, n$. The Euler-Lagrange equations.

As an example; $n = 1$, $L(u,v) = \frac{1}{2} m v^2 - \phi(u)$, $\frac{\partial L}{\partial u} = -\frac{d\phi}{du}$, $\frac{\partial L}{\partial v} = mv$, and the Euler Lagrange equation is $m\ddot u(t) = -\frac{d\phi}{du}(u(t))$, Newton's Law of motion.