TMA 4180 Optimeringsteori
Comments About Derivatives

H. E. Krogstad

April 2004/2005/2007/2008

This note should *not* be considered as a part of the curriculum. However, it deals with topics encountered again and again when computing functional derivatives, and is provided to those of you that feel uneasy about doing a mathematical step without a valid proof.

For the rest of us: "*Mathematicians that insist everything they use should be proved and fully understood are welcome to apply this principle to their own PCs*"*!*

1 The Gâteaux Derivative

The directional derivative (*Gâteaux derivative*) is defined in Troutman and elsewhere as

$$\delta J(y;v) = \lim_{\varepsilon\to 0} \frac{J(y+\varepsilon v) - J(y)}{\varepsilon}. \tag{1}$$

Similar to the ordinary directional derivative and its use in optimization theory, it could seem more reasonable to use a *one-sided* limit, that is,

$$\delta^+ J(y;v) = \lim_{\varepsilon\to 0+} \frac{J(y+\varepsilon v) - J(y)}{\varepsilon}. \tag{2}$$

This is called a *one-sided Gâteaux derivative* (the notation δ^+ is not quite standard). An obvious case where this applies is when J is defined only for non-negative functions, including the 0-function. Only a one-sided derivative can then exist for the 0-function.

However, the functionals we meet are usually defined on some linear space of functions (which is larger than the domain $\mathcal{D}$ we consider), and then the existence of $\delta J(y;v_0)$ for one particular v_0 implies that $\delta J(y;\alpha v_0)$ exists for all $\alpha\in\mathbb{R}$. This is probably the reason why Troutman does not bother do discuss it.

In general, for an ordinary function ϕ, if the one-sided derivatives exist and are equal, say

$$\lim_{\varepsilon\to 0+} \frac{\phi(\varepsilon)-\phi(0)}{\varepsilon} = \lim_{\varepsilon\to 0-} \frac{\phi(\varepsilon)-\phi(0)}{\varepsilon}, \tag{3}$$

then also the ordinary derivative,

$$\phi'(0) = \lim_{\varepsilon\to 0} \frac{\phi(\varepsilon)-\phi(0)}{\varepsilon} \tag{4}$$

exists and is equal to the common one-sided limits. A similar result exists for the Gâteaux derivative.

Proposition: *Assume that the one-sided derivatives $\delta^+ J(y;v)$ and $\delta^+ J(y;-v)$ exist. Then $\delta J(y;v)$ exists if and only of*

$$\delta^+ J(y;v) = -\delta^+ J(y;-v), \tag{5}$$

and then

$$\delta J(y;v) = \delta^+ J(y;v) = -\delta^+ J(y;-v) = -\delta J(y;-v). \tag{6}$$

Proof: First of all, if $\delta J(y;v)$ exists, then

$$\begin{aligned}\delta J(y;v) &= \lim_{\varepsilon\to 0}\frac{J(y+(\varepsilon)(v))-J(y)}{\varepsilon}\\ &= \lim_{\varepsilon\to 0}\frac{J(y+(-\varepsilon)(v))-J(y)}{-\varepsilon}\\ &= -\lim_{\varepsilon\to 0}\frac{J(y+\varepsilon(-v))-J(y)}{\varepsilon} = -\delta J(y;-v).\end{aligned}$$

Eqn. 6 follows immediately since the double-sided limits imply the exsistence of the one-sided limits.

For the converse, observe that

$$\delta^{+}J(y;v) = \lim_{\varepsilon\to 0+}\frac{J(y+\varepsilon v)-J(y)}{\varepsilon}, \tag{7}$$

and

$$\begin{aligned}-\delta^{+}J(y;-v) &= -\lim_{\varepsilon\to 0+}\frac{J(y+\varepsilon(-v))-J(y)}{\varepsilon}\\ &= \lim_{\varepsilon\to 0+}\frac{J(y+(-\varepsilon)v)-J(y)}{-\varepsilon}\\ &= \lim_{\varepsilon\to 0-}\frac{J(y+\varepsilon v)-J(y)}{\varepsilon}.\end{aligned} \tag{8}$$

If both one-sided limits exist and are equal, the usual limit for $\delta J(y;v)$ also exists.

In our applications, it seems that we'll never really need one-sided derivatives. If we assume that $y+\varepsilon v$ for all ε-s in a neighborhood $\mathcal{N}$ around 0, the function

$$\phi(\varepsilon) = J(y+\varepsilon v) \tag{9}$$

will be defined on $\mathcal{N}$, and the existence of the Gâteaux derivative for the J functional at y in the direction v is equivalent to the existence of the usual derivative of ϕ at zero,

$$\delta J(y;v) = \left.\frac{dJ(y+\varepsilon v)}{d\varepsilon}\right|_{\varepsilon=0} = \phi'(0). \tag{10}$$

This formula is what we use in practice.

2 Moving the derivative under the integral sign

When computing functional derivatives we often need to interchange derivatives and integrals, and usually this works fine, but in some cases the result will not be what we expect. It is not acceptable in a mathematical text just to write "if we can change the order of derivation and integration, then ...". A simple but important counter-example is discussed at the end of the note.

Let us consider the equation

$$\frac{d}{dt}\int_a^b f(x,t)\,dx = \int_a^b \frac{\partial f(x,t)}{\partial t}dx. \tag{11}$$

The left hand side should be understood as the derivative of the function

$$\Phi(t) = \int_a^b f(x,t)\,dx, \tag{12}$$

that is, $\frac{d\Phi}{dt}(t)$.

On the right hand side the function $\frac{\partial f(x,t)}{\partial t}$ has its usual meaning, and the right hand side is equal to another function of t, say $\Psi(t)$. Eqn. 11 then states that

$$\Phi'(t) = \Psi(t). \tag{13}$$

So what is the problem? Let us first consider the derivative of $\Phi(t)$:

$$\begin{aligned}\Phi'(t) &= \lim_{\tau\to 0}\frac{\Phi(t+\tau)-\Phi(t)}{\tau}\\ &= \lim_{\tau\to 0}\frac{\int_a^b f(x,t+\tau)\,dx - \int_a^b f(x,t)\,dx}{\tau}\\ &= \lim_{\tau\to 0}\int_a^b \frac{f(x,t+\tau)-f(x,t)}{\tau}dx.\end{aligned} \tag{14}$$

On the other hand,

$$\int_a^b \frac{\partial f(x,t)}{\partial t}dx = \int_a^b \lim_{\tau\to 0}\left(\frac{f(x,t+\tau)-f(x,t)}{\tau}\right)dx. \tag{15}$$

What is needed is therefore that

$$\lim_{\tau\to 0}\int_a^b \frac{f(x,t+\tau)-f(x,t)}{\tau}dx = \int_a^b \lim_{\tau\to 0}\left(\frac{f(x,t+\tau)-f(x,t)}{\tau}\right)dx, \tag{16}$$

or,

$$\text{"}\lim_{\tau\to 0}\int_a^b = \int_a^b \lim_{\tau\to 0}\text{"}. \tag{17}$$

Unfortunately, this is not always the case. Calculus has some theorems for this to be valid, but the simplest, easiest to apply, and much more general and powerful result is provided by the Lebesgue integration theory and what is called *Lebesgue's Dominated Convergence Theorem* (The result may be proved by plain calculus, but this is surprisingly tricky).

The proof goes as follows. Assume that both sides of Eqn. 11 exist separately. This requires, in particular, that the partial derivatives $\frac{\partial f(x,t)}{\partial t}$ exist, that is,

$$\frac{\partial f(x,t)}{\partial t} = \lim_{\tau\to 0}\frac{f(x,t+\tau)-f(x,t)}{\tau}. \tag{18}$$

If there is a non-negative function H such that

$$\int_a^b H(x)\,dx < \infty, \tag{19}$$

and

$$\left|\frac{f(x,t+\tau)-f(x,t)}{\tau}\right| \le H(x), \tag{20}$$

for all τ in a fixed neighborhood $\mathcal{N}$ of 0, then Eqns. 16 and 11 hold.

The crucial condition is the existence of H, and the result is valid for both finite and infinite integration limits (Note that $\mathcal{N}$ should be independent of t and x).

When dealing with finite intervals, say $x \in [a,b]$, $t \in [t_1, t_2]$, and smooth functions, finding a suitable H is trivial

$$\begin{aligned} \left| \frac{f(x,t+\tau) - f(x,t)}{\tau} \right| &= \left| \frac{\partial f(x,t+\theta\tau)}{\partial t} \right| \leq \max_x \left| \frac{\partial f(x,t+\theta\tau)}{\partial t} \right| \\ &\leq \max_{x,t} \left| \frac{\partial f(x,t)}{\partial t} \right| = M < \infty. \end{aligned} \tag{21}$$

(The maximum is finite since it is a maximum of a continuous function on a closed and bounded set). Here $H(x) = M$ will do, since $\int_a^b M dx < \infty$.

A simpler version of this result is found as Theorem A.13 in Troutman.

Example

$$\begin{aligned} J(y) &= \int_0^1 \sin(y(x))\, dx, \\ & y \in C[0,1]. \end{aligned} \tag{22}$$

Then

$$J(y + \varepsilon v) = \int_0^1 \sin(y(x) + \varepsilon v(x))\, dx, \tag{23}$$

and it would be nice to write

$$\begin{aligned} \frac{d}{d\varepsilon} \left(\int_0^1 \sin(y(x) + \varepsilon v(x))\, dx \right) \Bigg|_{\varepsilon=0} &= \int_0^1 \frac{\partial \sin(y(x) + \varepsilon v(x))\, dx}{\partial \varepsilon} \Bigg|_{\varepsilon=0} dx \\ &= \int_0^1 \cos(y(x))\, v(x)\, dx. \end{aligned} \tag{24}$$

But this follows immediately from an argument similar to Eqn. 21:

$$\left| \frac{\partial \sin(y(x) + \varepsilon v(x))}{\partial \varepsilon} \right| = |\cos(y(x) + \varepsilon v(x))\, v(x)| \leq \max_x |v(x)| \; < \infty, \tag{25}$$

since v is continuous on [0,1].

Counter-example

A somewhat subtile counter-example, which has some significance in the theory of shock-waves, is the function

$$h(x,t) = \begin{cases} 1 & x \leq t \\ 0 & x > t \end{cases} \tag{26}$$

For $x, t \in (-1, 1)$,

$$\Phi(t) = \int_{-1}^1 h(x,t)\, dx = \int_{-1}^t 1 \cdot dx = t + 1, \tag{27}$$

and therefore

$$\frac{d\Phi(t)}{dt} = \frac{d}{dt} \int_0^1 h(x,t)\, dx = 1. \tag{28}$$

On the other hand,

$$\frac{\partial h\left(x,t\right)}{\partial t}=0 \tag{29}$$

for all $t\neq x$, and hence,

$$\Psi\left(t\right)=\int_{-1}^{1}\frac{\partial h\left(x,t\right)}{\partial t}dx=\int_{-1}^{t}\frac{\partial h\left(x,t\right)}{\partial t}dx+\int_{t}^{1}\frac{\partial h\left(x,t\right)}{\partial t}dx=0. \tag{30}$$

Thus, we need to be careful when $h\left(x,t\right)$ is not smooth.

So what about $H\left(x\right)$ in this case? Let us find the *smallest* possible H for $t=0$ and $\tau\in\mathcal{N}=\left(-1,1\right)$. If that fails to be integrable, we would expect problems (In some exceptional cases it may be OK anyway).

The smallest function H is

$$H_{m}\left(x\right)=\sup_{\tau\in\mathcal{N}}\left|\frac{f\left(x,0+\tau\right)-f\left(x,0\right)}{\tau}\right|, \tag{31}$$

where "sup" means *supremum* = *least upper bound*. From Eqn. 26 we obtain for positive τ-s

$$\left|\frac{f\left(x,\tau\right)-f\left(x,0\right)}{\tau}\right|=\begin{cases}0, & x\leq0,\\ \frac{1}{\tau}, & 0<x\leq\tau,\\ 0, & \tau<x,\end{cases} \tag{32}$$

and a similar result for negative τ-s. This function reaches the height $\left|x\right|^{-1}$ for $\left|x\right|=\left|\tau\right|$, and by taking the supremum of all functions in Eqn. 32, we find

$$H_{m}\left(x\right)=\frac{1}{\left|x\right|}, \tag{33}$$

which is *not integrable* on $\left(-1,1\right)$.