Characterizations of Differentiability in $\mathbb{R}^n$

Portland State University, Analysis Seminar

Notes by Blake Rector

February 7, 2014

Abstract. In this talk we discuss three notions of the derivative for functions defined on $\mathbb{R}^n$: the directional derivative, the Gâteaux derivative, and the the Fréchet derivative. For the directional derivative, we give a formula based on the sub-differential. For the second two, we determine a condition for existence terms of the sub-differential. In particular, we show that the second two derivatives exist at $\overline{x}$ precisely when $\partial f(\overline{x})$ is a singleton. The previous formulation of the directional derivative is used to prove this result.

Definitions and Preliminaries

Def. $\overline{\mathbb{R}}$ is the real numbers with positive ∞.

Def. For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ we define $\operatorname{dom} f := \{x \in \mathbb{R}^n | f(x) < \infty\}$.

Def. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *convex* if for all $x, y \in \mathbb{R}^n$ and all $t \in (0, 1)$ we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Fact. A convex function is locally lipshitz continuous on the interior of its domain.

Def. A *subgradient* of a convex function f at a point $\overline{x} \in \mathbb{R}^n$ is a vector $v \in \mathbb{R}^n$ such that

$\langle v, x-\overline{x}\rangle \le f(x) - f(\overline{x})$. The *subdifferential* of f at $\overline{x}$ is the set of all subgradients. We denote the subdifferential by $\partial f(\overline{x})$.

Fact. For any $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$, the subdifferential $\partial f(\overline{x})$ is nonempty and compact.

Def. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let $\overline{x} \in \operatorname{dom} f$. The *directional derivative* of the function f at the point $\overline{x}$ in the direction $d \in \mathbb{R}^n$ is the following limit (if it exists as either a real number or $\pm\infty$):

$$f'(\overline{x}, d) := \lim_{t \to 0^+} \frac{f(\overline{x}+td) - f(\overline{x})}{t}.$$

Although we do not really need it in this talk, we can also define the *left* counterpart of the directional derivative

$$f'_-(\overline{x}, d) := \lim_{t \to 0^-} \frac{f(\overline{x}+td) - f(\overline{x})}{t},$$

and we notice that the following equality holds: $f'_-(\overline{x}; d) = -f'(x; -d)$.

If we think of $\overline{x}$ as fixed in the definition of $f'(\overline{x}, d)$, then we can define a function ψ of d that we call the directional function.

Def. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let $\overline{x} \in \operatorname{dom} f$. The the *directional function* (associated with f and $\overline{x}$) is defined as $\psi(d) := f'(\overline{x}; d)$.

Fact. We have $\psi(\alpha d) = \alpha\psi(d)$ whenever $d \in \mathbb{R}^n$ and $\alpha > 0$. This is easy to prove from the definition.

Fact. If $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$, then ψ is finite and convex on $\mathbb{R}^n$. [Note this means that the directional derivative $f'(\overline{x}; d)$ exists as a real number for all directions $d \in \mathbb{R}^n$ at all points $\overline{x}$ in the interior of the domain of f.]

Before we further pursue discussion of the directional derivative, we should introduce the Fenchel conjugate, also sometimes known as the convex conjugate.

The Fenchel Conjugate

Def. Given a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, its *Fenchel conjugate* $f^* : \mathbb{R}^n \to [-\infty, \infty]$ is

$$f^*(v) = \sup\{\langle v, x\rangle - f(x) \,|\, x \in \mathbb{R}\}.$$

Note. By the nature of sup and inf we also have $-f^*(v) = \inf\{f(x) - \langle v, x\rangle \,|\, x \in \mathbb{R}\}$, since

Example. If $f(x) = x^2 + 1$, then $f^*(v) = \frac{v^2}{4} - 1$.

Proposition 2.76. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. Let $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$. Then we have the equality $f^{**}(\overline{x}) = f(\overline{x})$.

Proof. Let $\overline{x} \in \operatorname{dom} f$. First we show $f^{**}(\overline{x}) \leq f(\overline{x})$.

Notice

$$\begin{aligned}
f(\overline{x}) =& f(\overline{x}) + \langle v, \overline{x}\rangle - \langle v, \overline{x}\rangle \\
=& \langle v, \overline{x}\rangle - (\langle v, \overline{x}\rangle - f(\overline{x})) \\
\geq& \langle v, \overline{x}\rangle - \sup\{\langle v, x\rangle - f(x) | x \in \mathbb{R}^n\} \\
=& \langle v, \overline{x}\rangle - f^*(v)
\end{aligned}$$

holds for all $v \in \mathbb{R}^n$. Thus

$$\begin{aligned}
f(\overline{x}) \geq& \sup\{\langle v, \overline{x}\rangle - f^*(v) | v \in \mathbb{R}^n\} \\
=& f^{**}(\overline{x}).
\end{aligned}$$

Next we show $f^{**}(\overline{x}) \geq f(\overline{x})$.

Let $v \in \partial f(\overline{x})$ (we know $\partial f(\overline{x}) \neq 0$ since $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$). Then, by definition of $\partial f(\overline{x})$, for any $x \in \mathbb{R}^n$, we have

$$\begin{aligned}
\langle v, x - \overline{x}\rangle &\leq f(x) - f(\overline{x}) \\
\Rightarrow \langle v, x\rangle - \langle v, \overline{x}\rangle &\leq f(x) - f(\overline{x}) \\
\Rightarrow \langle v, \overline{x}\rangle - f(\overline{x}) &\geq \langle v, x\rangle - f(x) \\
\Rightarrow \langle v, \overline{x}\rangle - f(\overline{x}) &\geq \sup\{\langle v, x\rangle - f(x) | x \in \mathbb{R}^n\} \qquad \text{(Since the previous holds for all } x \in \mathbb{R}^n\text{)} \\
\Rightarrow \langle v, \overline{x}\rangle - f(\overline{x}) &\geq f^*(v) \\
\Rightarrow f(\overline{x}) &\leq \langle v, \overline{x}\rangle - f^*(v) \\
&\leq \sup\{\langle v, \overline{x}\rangle - f^*(v) | v \in \mathbb{R}^n\} \\
&= f^{**}(\overline{x})
\end{aligned}$$

Thus, $f^{**}(\overline{x}) = f(\overline{x})$. This completes the proof.

Now we continue our discussion of the directional derivative.

Directional Derivative

Lemma 2.85. Let ψ be defined as above. Then

$$\psi^*(v) = \delta_\Omega(v) = \begin{cases} 0 & v \in \Omega \\ \infty & v \notin \Omega \end{cases}$$

for all $v \in \mathbb{R}^n$ where $\Omega = \partial\psi(0)$.

Proof. To show that $\psi^*(v) = \delta_\Omega(v)$, we must show that $\psi^*(v) = 0$ when $v \in \Omega$ and $\psi^*(v) = \infty$ when $v \notin \Omega$.

First we notice that for any $v \in \mathbb{R}^n$ we have $\psi^*(v) \geq 0$ since

$$\psi^*(v) = \sup\{\langle v, d\rangle - \psi(d) | d \in \mathbb{R}^n\} \;\geq\; \langle v, 0\rangle - \psi(0) = 0 - 0 = 0.$$

Now for the first case, let $v \in \Omega = \partial\psi(0)$. Then we have by definition $\langle v, d\rangle = \langle v, d - 0\rangle \leq \psi(d) - \psi(0) = \psi(d)$ for any $d \in \mathbb{R}^n$ by the definition of $\partial\psi(0)$. Thus $\langle v, d\rangle - \psi(d) \leq 0$ for all $d \in \mathbb{R}^n$. This implies $0 \geq \sup\{\langle v, d\rangle - \psi(d) | d \in \mathbb{R}^n\} = \psi^*(v)$. So we can conclude $\psi^*(v) = 0$ for all $v \in \partial\psi(0)$.

Next we show that $\psi^*(v) = \infty$ when $v \notin \Omega = \partial\psi(0)$. Note that if $v \notin \partial\psi(0)$, then we do not have: $\langle v, d\rangle \leq \psi(d)$ for all $d \in \mathbb{R}^n$. So let $d_0 \in \mathbb{R}^n$ such that $\langle v, d_0\rangle > \psi(d_0)$. (That is, $\langle v, d_0\rangle - \psi(d_0) = m$ for some $m > 0$.) Using the fact that $\psi(\alpha d) = \alpha\psi(d)$ whenever $\alpha\rangle 0$, for any $d \in \mathbb{R}^n$ we have

$$\begin{aligned} \psi^*(v) &= \sup\{\langle v, d\rangle - \psi(d) | d \in \mathbb{R}^n\} \\ &\geq \sup\{\langle v, td_0\rangle - \psi(td_0) | t > 0\} \\ &= \sup\{t(\langle v, d_0\rangle - \psi(d_0)) | t > 0\} \\ &= \sup\{tm | t > 0\} = \infty \end{aligned}$$

Therefore $\psi^*(v) = \infty$ whenever $v \notin \partial\psi(0)$. This completes the proof.

Fact. In the setting of the previous lemma, we have $\partial\psi(0) = \partial f(\overline{x})$.

The following theorem gives us a characterization of the directional derivative in terms of the sub-differential.

Theorem 2.86. Given a convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$, we have

$$f'(\overline{x}; d) = \max\{\langle v, d\rangle | v \in \partial f(\overline{x})\}$$

for any $d \in \mathbb{R}^n$.

Proof. We know that the directional function $\psi(d) = f'(\overline{x}; d)$ is finite and convex for $\overline{x} \in \operatorname{dom} f$, so by Proposition 2.76 it has

$$\psi^{**}(d) = \psi(d)$$

as a function of d.

We also learned previously (Lemma 2.85) that $\psi^*(v) = \sup\{\langle x, v\rangle - \psi(x) \,|\, x \in \operatorname{dom}\psi\} = \delta_\Omega(v)$ where $\Omega = \partial\psi(0) = \partial f(\overline{x})$ (note that $\Omega = \partial f(\overline{x})$ is compact since f is convex and $\overline{x}$ is in the interior of its domain). Since $\psi^*(v) = \delta_\Omega(v)$, we have $\psi^{**}(d) = \delta^*_\Omega(d) = \sup\{\langle v, d\rangle \,|\, v \in \Omega\}$ (the last equality a direct consequence of the definition of the conjugate). And since Ω is compact we can replace the sup with max.

In summary, these steps are as follows:

$$\begin{aligned} f'(\overline{x}; d) = \psi(d) = \psi^{**}(d) = \delta^*_\Omega(d) &= \sup\{\langle v, d\rangle - \delta_\Omega(v) \,|\, v \in \mathbb{R}^n\} \\ &= \sup\{\langle v, d\rangle \,|\, v \in \Omega\} \\ &= \max\{\langle v, d\rangle \,|\, v \in \partial f(\overline{x})\} \end{aligned}$$

This completes the proof.

The Gâteaux and Fréchet Derivatives

Def. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$. We say that f is *Gâteaux differentiable* at $\overline{x}$ if there exists $v \in \mathbb{R}^n$ such that for all $d \in \mathbb{R}^n$ we have

$$\lim_{t\to 0} \frac{f(\overline{x} + td) - f(\overline{x}) - t\langle v, d\rangle}{t} = 0.$$

If such a v exists we denote it $f'_G(\overline{x}) = v$ and call v the Gâteaux derivative.

Note. If the Gâteaux derivative v exists, it is unique. This is easy to see since if v_1, v_2 were both Gâteaux derivatives, then subtracting the two expressions under the limit would yield $\lim_{t\to 0}\langle v_1 - v_2, d\rangle = 0$, which would imply $v_1 = v_2$.

Note. We have

$$f'_G(\overline{x}) = v \Leftrightarrow f'(\overline{x}; d) = \langle v, d\rangle \text{ for all } d.$$

We can see this is true since $f'_-(\overline{x}; d) = -f'(x; -d)$ and by using the definition of the limit:

$$\begin{aligned}
f'_G(\overline{x}) = v &\Leftrightarrow \lim_{t\to 0} \frac{f(\overline{x}+td) - f(\overline{x})}{t} = \langle v, d\rangle \\
&\Leftrightarrow \lim_{t\to 0^+} \frac{f(\overline{x}+td) - f(\overline{x})}{t} = \langle v, d\rangle \;\text{ and }\; \lim_{t\to 0^-} \frac{f(\overline{x}+td) - f(\overline{x})}{t} = \langle v, d\rangle \\
&\Leftrightarrow f'(\overline{x}; d) = \langle v, d\rangle \;\text{ and }\; f'_-(\overline{x}; d) = \langle v, d\rangle \\
&\Leftrightarrow f'(\overline{x}; d) = \langle v, d\rangle \;\text{ and }\; -f'(\overline{x}; -d) = \langle v, d\rangle \\
&\Leftrightarrow f'(\overline{x}; d) = \langle v, d\rangle \;\text{ and }\; f'(\overline{x}; -d) = \langle v, -d\rangle \\
&\Leftrightarrow f'(\overline{x}; d) = \langle v, d\rangle
\end{aligned}$$

Def. We say that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *Fréchet differentiable* at $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$ if there exists an element $v \in \mathbb{R}^n$ such that

$$\lim_{x\to\overline{x}} \frac{f(x) - f(\overline{x}) - \langle v, x - \overline{x}\rangle}{\|x - \overline{x}\|} = 0.$$

In this case, the element v is uniquely defined and denoted by $\nabla f(\overline{x}) := v$.

Proposition. In general for $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$ we have the following implications:

f is Fréchet differentiable at $\overline{x} \Rightarrow f$ is Gâteaux differentiable at $\overline{x}$

$\Rightarrow$ all directional derivatives of f exist at $\overline{x} \Rightarrow$ all partial derivatives of f exist at $\overline{x}$

Proof. To see the first implication, assume f is Fréchet differentiable at $\overline{x}$ and let $\nabla f(\overline{x}) = v$. By definition this means that

$$\lim_{x\to\overline{x}} \frac{f(x) - f(\overline{x}) - \langle v, x - \overline{x}\rangle}{\|x - \overline{x}\|} = 0.$$

Now, fix a nonzero $d \in \mathbb{R}^n$. In the limit above, consider the path $x \to \overline{x}$ by way of $(\overline{x}+td) \to \overline{x}$ as $t \to 0$. In this way, we can replace x with $\overline{x}+td$ and $x - \overline{x}$ with $(\overline{x}+td) - \overline{x} = td$ and "$x \to \overline{x}$" with

"$t \to 0$" and the [weaker] limit statement will still hold:

$$\lim_{t\to 0} \frac{f(\overline{x}+td) - f(\overline{x}) - \langle v, td\rangle}{t\|d\|} = 0.$$

Multiplying by the constant $\|d\|$ will not change the value of the limit, so we have

$$\|d\| \lim_{t\to 0} \frac{f(\overline{x}+td) - f(\overline{x}) - \langle v, td\rangle}{t\|d\|} = 0.$$

Simplifying this yields

$$\lim_{t\to 0} \frac{f(\overline{x}+td) - f(\overline{x}) - t\langle v, d\rangle}{t} = 0,$$

so f is Gâteaux differentiable at $\overline{x}$. The second implication is true by the second note following the definition of the Gâteaux derivative. The third implication is true because the partial derivatives are just the directional derivatives in the direction of each axis. This completes the proof.

The next fact together with the previous proposition tells us that Fréchet and Gâteaux differentiability are equivalent for locally lipschitz continuous functions.

Fact. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be locally lipschitz continuous around $\overline{x} \in \operatorname{dom} f$. Then the following implication holds:

$$f \text{ is Fréchet differentiable at } \overline{x} \Leftarrow f \text{ is Gâteaux differentiable at } \overline{x}$$

Finally, we arrive at our characterization. The following gives us a condition for the existence of the Gâteaux derivative (and hence the Fréchet derivative) in terms of the subdifferential.

Theorem 3.3. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function with $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$. Then the following assentations are equivalent:

(i) f is Fréchet differentiable at $\overline{x}$.

(ii) f is Gâteaux differentiable at $\overline{x}$.

(iii) $\partial f(\overline{x})$ is a singleton.

Proof. The equivalence of (i) and (ii) has already been established, so we will show (ii)$\Leftrightarrow$(iii).

First we show (ii)$\Rightarrow$(iii). Assume f is Gâteaux differentiable at $\overline{x}$ and $f'_G(\overline{x}) = v$. Then we have $f'(x; d) = \langle v, d\rangle$. We know $\partial f(\overline{x}) \neq \varnothing$ since $\overline{x} \in \operatorname{int}(\operatorname{dom} f)$, so fix $w \in \partial f(\overline{x})$. [We will show $v = w$, which will prove that $\partial f(\overline{x})$ is a singleton since v is unique.]

Since $\langle v, d\rangle = f'(\overline{x}; d) = \max\{\langle u, d\rangle \mid u \in \partial f(\overline{x})\}$, we have $\langle w, d\rangle \leq \langle v, d\rangle$. But we also have

$$\begin{aligned}\langle v, d\rangle &= -\langle v, -d\rangle \\ &= -f'(x; -d) \\ &= -\max\{\langle u, -d\rangle \mid u \in \partial f(\overline{x})\} \\ &= -\max\{-\langle u, d\rangle \mid u \in \partial f(\overline{x})\} \\ &= \min\{\langle u, d\rangle \mid u \in \partial f(\overline{x})\}, \\ &\leq \langle w, d\rangle.\end{aligned}$$

Thus $\langle v, d\rangle = \langle w, d\rangle$.

Since d was arbitrary, this means $\langle v - w, d\rangle = 0$ for all d. In particualr, if $d = w - v$, we have $\|w - v\|^2 = 0$, which implies $w = v$. Since v is unique, this proves that $\partial f(\overline{x})$ is a singleton.

Next we show (iii)$\Rightarrow$(ii). Suppose that (iii) is satisfied with $\partial f(\overline{x}) = v$. Then we have

$$f'(\overline{x}; d) = \sup\{\langle w, d\rangle \mid w \in \partial f(\overline{x})\} = \langle v, d\rangle.$$

Thus f is Gâteaux differentiable.

References

Boris S. Mordukhovich, Nguyen Mau Nam. *An Easy Path to Convex Analysis.* Morgan and Claypool Publishers. 2014.

Deotte, Chris. Examples of different types of differentiability. http://math.ucsd.edu/ cdeotte/papers/Frechet.pdf