wj32

information when you need it

MATHEMATICS

Differentiation done correctly: 3. Partial derivatives

Posted on February 23, 2013 by wj32 / 0 Comment

Navigation: 1. The derivative | 2. Higher derivatives | **3. Partial derivatives** | 4. Inverse and implicit functions | 5. Maxima and minima

While we saw that differentiable maps may be naturally split into component functions when the codomain is a product of Banach spaces, the situation for the domain is more complicated. (This is partly due to the fact that as a topological space, there is no natural injection into a product of Banach spaces.) In this post, we will look at how the existence of partial derivatives relates to differentiability, how the symmetry of higher derivatives (covered in part 2) affects mixed partial derivatives, and finally a short proof of differentiation under the integral sign.

Let $E_1, \ldots, E_m$ be Banach spaces and let $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ be an open set where each A_j is open in E_j. If $f: A_1 \times \cdots \times A_m \to F$ is any map and $x = (x_1, \ldots, x_m) \in A_1 \times \cdots \times A_m$, we can consider the map

$$t \mapsto f(x_1, \ldots, t, \ldots, x_m),$$

which can also be written as $f \circ \iota$ where $\iota: A_j \to A_1 \times \cdots \times A_m$ is given by $t \mapsto (x_1, \ldots, t, \ldots, x_m)$. If this map is differentiable at x_j, we call its derivative the **jth partial derivative of f at x** and denote it by $D_j f(x)$. Looking again at the definition of the derivative, we see that $D_j f(x): E_j \to F$ is the unique continuous linear map such that

$$\lim_{h \to 0} \frac{f(x_1, \ldots, x_j + h, \ldots, x_m) - f(x_1, \ldots, x_m) - D_j f(x)h}{|h|} = 0.$$

In practice, if we are working with functions defined on R^m then we take $E_j = \mathrm{R}$ for $j = 1, \ldots, m$ so that we have a decomposition of R^m into $\mathrm{R} \times \cdots \times \mathrm{R}$. In this situation we often see the notation

$$\frac{\partial f}{\partial x_j}(x) = D_j f(x)(1),$$

where we identify the linear map $D_j f(x): \mathrm{R} \to F$ with the value $D_j f(x)(1) \in F$.

It is not hard to see that all partial derivatives exist at x if f is differentiable at x.

Theorem 25. *Let* $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ *where each* A_j *is open in* E_j *and let* $f: A_1 \times \cdots \times A_m \to F$. *If* f *is differentiable at* $x = (x_1, \ldots, x_m) \in A_1 \times \cdots \times A_m$, *then every* $D_j f(x)$ *exists and we have* $D_j f(x) = Df(x) \circ \iota_j$ *where* $\iota_j: E_j \to E_1 \times \cdots \times E_m$ *is the canonical injection, i.e.*

$$Df(x) = \begin{bmatrix} D_1 f(x) & \cdots & D_m f(x) \end{bmatrix}.$$

Proof. Apply the chain rule to $f \circ \iota$ where $\iota: A_j \to A_1 \times \cdots \times A_m$ is given by $t \mapsto (x_1, \ldots, t, \ldots, x_m)$. Alternatively, restrict h to elements of E_j in the definition of the derivative $Df(x)$. ▢

Definition 26. *Let* $E_1, \ldots, E_m$ *and* $F_1, \ldots, F_n$ *be Banach spaces. Let* $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ *where each* A_j *is open in* E_j *and let* $f: A_1 \times \cdots \times A_m \to F_1 \times \cdots \times F_n$. *The matrix*

$$\begin{bmatrix} D_1 f_1(x) & \cdots & D_m f_1(x) \\ \vdots & \ddots & \vdots \\ D_1 f_n(x) & \cdots & D_m f_n(x) \end{bmatrix}$$

is called the **Jacobian** *matrix of* f *at* $x \in A$, *where* $f_i = \pi_i \circ f$ *and* $\pi_i: F_1 \times \cdots \times F_n \to F_i$ *is the canonical projection.*

Theorem 27. *If* $f: A_1 \times \cdots \times A_m \to F_1 \times \cdots \times F_n$ *is differentiable at* $x \in A$, *then the Jacobian matrix of* f *at* x *exists and represents* $Df(x)$.

Proof. Apply Theorem 8 followed by Theorem 25. ▢

As in R^n, it may be the case that every partial derivative $D_j f_i(x)$ exists but f is not differentiable at x. The differentiability of f implies the existence of the Jacobian matrix, but the converse is not true. Thus we do not have a true analog of Theorem 8 for partial derivatives. We do however have an analog of Theorem 18.

Theorem 28. *Let* $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ *where each* A_j *is open in* E_j *and let* $f: A_1 \times \cdots \times A_m \to F$. *Then* f *is of class* C^p *(with* $p \geq 1$*) if and only if every partial derivative*

$$D_j f : A_1 \times \cdots \times A_m \to L(E_j, F)$$

is of class C^{p-1}. In that case we have $D_j f(x) = Df(x) \circ \iota_j$ where $\iota_j : E_j \to E_1 \times \cdots \times E_m$ is the canonical injection, i.e.

$$Df(x) = \begin{bmatrix} D_1 f(x) & \cdots & D_m f(x) \end{bmatrix}.$$

Proof. It is clear from the proof of Theorem 25 that every partial derivative is of class C^{p-1} if f is of class C^p. For the converse, we only need to prove that Df exists on $A_1 \times \cdots \times A_m$ since

$$Df(x) = \sum_{j=1}^{m} D_j f(x) \circ \pi_j$$

implies that Df is of class C^{p-1} if every $D_j f$ is of class C^{p-1}, where $\pi_j : E_1 \times \cdots \times E_m \to E_j$ is the canonical projection. Let $x \in A_1 \times \cdots \times A_m$ and let $\varepsilon > 0$. Since every $D_j f$ is continuous, there exists a $\delta > 0$ such that

$$|D_j f(y) - D_j f(x)| < \frac{\varepsilon}{m}$$

for all $j = 1, \ldots, m$ and $y \in B_\delta(x) \subseteq A_1 \times \cdots \times A_m$ where $B_\delta(x)$ is the open ball of radius δ around x. Let $h = (h_1, \ldots, h_m) \in E_1 \times \cdots \times E_m$ with $|h| < \delta$. For $j = 0, \ldots, m$, let $p_j = h_1 + \cdots + h_j$ so that $p_0 = 0$, $p_m = h$, and

$$f(x + h) - f(x) = \sum_{j=1}^{m} [f(x + p_j) - f(x + p_{j-1})].$$

For each $j = 1, \ldots, m$ the line segment from $x + p_{j-1}$ to $x + p_j = x + p_{j-1} + h_j$ is contained in $B_\delta(x)$, so we have

$$f(x + p_j) - f(x + p_{j-1}) = \int_0^1 D_j f(x + p_{j-1} + th_j) h_j \, dt$$

by the <u>mean value theorem</u>. Then

$$\left| f(x + h) - f(x) - \sum_{j=1}^{m} D_j f(x) h_j \right|$$

$$\leq \left| \sum_{j=1}^{m} \left[\int_0^1 D_j f(x + p_{j-1} + th_j) h_j \, dt - \int_0^1 D_j f(x) h_j \, dt \right] \right|$$

$$\leq \sum_{j=1}^{m} |h_j| \int_0^1 |D_j f(x + p_{j-1} + th_j) - D_j f(x)| \, dt$$

$$\leq \sum_{j=1}^{m} |h_j| \frac{\varepsilon}{m}$$

$$\leq |h| \varepsilon$$

for all $|h| < \delta$, which shows that

$$Df(x) = \sum_{j=1}^{m} D_j f(x) \circ \pi_j.$$

☐ As in the case of the ordinary derivative Df, we may take higher derivatives of partial derivatives:

$$D_{j_1} \cdots D_{j_r} f : A_1 \times \cdots \times A_m \to L(E_{j_1}, L(\ldots, L(E_{j_r}, F))\ldots).$$

These are sometimes known as **mixed partial derivatives**. <u>Theorem 21</u> has an important interpretation in terms of the mixed partial derivatives of f.

Theorem 29 (Equality of mixed partial derivatives). *Let $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ where each A_j is open in E_j and let $f : A_1 \times \cdots \times A_m \to F$ be of class C^2. Then*

$$D_j D_k f(x)(u)(v) = D_k D_j f(x)(v)(u)$$

for all $1 \leq j, k \leq m$, $x \in A_1 \times \cdots \times A_m$, $u \in A_j$ and $v \in A_k$.

Proof. For $j = 1, \ldots, m$, let $\iota_j : E_j \to E_1 \times \cdots \times E_m$ be the canonical injection. We have that $D_k f(x) = Df(x) \circ \iota_k$, so $D_k f = c \circ Df$ where

$$c : L(E_1 \times \cdots \times E_m, F) \to L(E_k, F)$$
$$\lambda \mapsto \lambda \circ \iota_k.$$

Similarly, $D_j D_k f = d \circ D(D_k f)$ where

$$d : L(E_1 \times \cdots \times E_m, L(E_k, F)) \to L(E_j, L(E_k, F))$$
$$\lambda \mapsto \lambda \circ \iota_j.$$

Note that c and d are both linear maps. Therefore

$$D_j D_k f(x) = [d \circ D(D_k f)](x)$$
$$= d(D(c \circ Df)(x))$$
$$= d(Dc(Df(x)) \circ D^2 f(x))$$
$$= d(c \circ D^2 f(x))$$
$$= c \circ D^2 f(x) \circ \iota_j$$

and

$$D_j D_k f(x)(u)(v) = (c \circ D^2 f(x) \circ \iota_j)(u)(v)$$
$$= c(D^2 f(x)(\iota_j(u)))(v)$$
$$= D^2 f(x)(\iota_j(u))(\iota_k(v)).$$

A similar calculation shows that $D_k D_j f(x)(v)(u) = D^2 f(x)(\iota_k(v))(\iota_j(u))$. But

$$D^2 f(x)(\iota_k(v))(\iota_j(u)) = D^2 f(x)(\iota_j(u))(\iota_k(v))$$

by [Theorem 21](#), so the result follows. ▢

In the special case $E_j = \mathbb{R}$ for $j = 1, \ldots, m$ and $f: \mathbb{R}^m \to F$, we have

$$\frac{\partial f}{\partial x_j \partial x_k}(x) = D_j D_k f(x)(1)(1) = D_k D_j f(x)(1)(1) = \frac{\partial f}{\partial x_k \partial x_j}(x),$$

which is sometimes known as the [symmetry of second derivatives](#), or [Clairaut's theorem](#).

Often, p times continuous differentiability is defined in terms of the mixed partial derivatives of f. The following theorem shows that this definition is equivalent to ours.

Theorem 30. *Let $A_1 \times \cdots \times A_m \subseteq E_1 \times \cdots \times E_m$ where each A_j is open in E_j and let $f: A_1 \times \cdots \times A_m \to F$. Then f is of class C^p (with $p \geq 1$) if and only if the partial derivative*

$$D_{\tau(1)} \cdots D_{\tau(k)} f$$

exists on $A_1 \times \cdots \times A_m$ and is continuous, for every $k = 1, \ldots, p$ and every map τ from $\{1, \ldots, k\}$ to $\{1, \ldots, m\}$. Furthermore,

$$D_{\tau(\sigma(1))} \cdots D_{\tau(\sigma(k))} f(x)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = D_{\tau(1)} \cdots D_{\tau(k)} f(x)(v_1, \ldots, v_k)$$

for all $x \in A_1 \times \cdots \times A_m$ and any permutation σ of $\{1, \ldots, k\}$.

Proof. This follows directly from Theorem 28 and Theorem 29. ▢

Differentiation under the integral sign

Theorem 31 (Differentiation under the integral sign). *Let $A \subseteq E$ be an open set and let $[a, b]$ be a closed interval with $a < b$. Let $f: [a, b] \times A \to F$ be a continuous map such that $D_2 f$ exists on $[a, b] \times A$ and is continuous. Let $g: A \to F$ be given by*

$$g(x) = \int_a^b f(t, x)\, dt.$$

Then g is differentiable on A and

$$Dg(x) = \int_a^b D_2 f(t, x)\, dt.$$

Proof. Let $x \in A$. Let

$$\lambda = \int_a^b D_2 f(t, x)\, dt.$$

For sufficiently small h we have

$$g(x + h) - g(x) - \lambda h = \int_a^b [f(t, x + h) - f(t, x) - D_2 f(t, x)h]\, dt$$
$$= \int_a^b \left[\int_0^1 D_2 f(t, x + sh)h\, ds - D_2 f(t, x)h \right] dt$$
$$= \int_a^b \int_0^1 [D_2 f(t, x + sh) - D_2 f(t, x)]h\, ds\, dt$$

so that

$$\frac{|g(x + h) - f(x) - \lambda h|}{|h|} \leq \int_a^b \int_0^1 |D_2 f(t, x + sh) - D_2 f(t, x)|\, ds\, dt$$
$$\leq (b - a) \sup_{s, t} |D_2 f(t, x + sh) - D_2 f(t, x)|$$

where the sup is taken over all $0 \leq s \leq 1$ and $a \leq t \leq b$.

Let $\varepsilon > 0$. For each $t \in [a, b]$ there is a neighborhood $B_t \times U_t$ of (t, x) such that $|D_2 f(u, y) - D_2 f(t, x)| < \varepsilon$ whenever $(u, y) \in B_t \times U_t$, and such that B_t and U_t are open balls around t and x respectively. Since $[a, b]$ is compact, there are finitely many balls $B_{t_1}, \ldots, B_{t_n}$ that cover $[a, b]$. Then for sufficiently small h such that $x + h \in \bigcap_{k=1}^n U_{t_k}$ and all $0 \leq s \leq 1$ and $a \leq t \leq b$ we have $t \in B_{t_k}$ for some k, so

$$|D_2 f(t, x + sh) - D_2 f(t, x)| \leq |D_2 f(t, x + sh) - D_2 f(t_k, x)|$$
$$+ |D_2 f(t_k, x) - D_2 f(t, x)|$$
$$< 2\varepsilon.$$

The next post will prove exactly three things: the Banach fixed-point theorem, the inverse function theorem, and the implicit function theorem.

wj32

[View all posts by wj32 →](#)