5.3 Higher Derivatives

Assume $f: U \to \mathbb{R}^m$ is diff'ble ~~on~~ U. To make sense of a "second derivative" recall
$$Df: U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$
Now, Df has the same domain as f but a different range. Nevertheless, $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a vector space, just like $\mathbb{R}^m$, in fact as vector spaces
$$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathbb{R}^{n \cdot m}.$$
Using this we can ask whether Df is differentiable at some p:

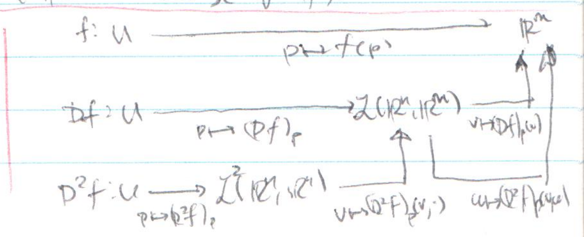
<u>Def</u>: We say $f: U \to \mathbb{R}^m$ is <u>twice-differentiable</u> at p if f is differentiable on U and Df is differentiable at p. We write
$$(D^2 f)_p = (D(Df))_p$$
and call $(D^2 f)_p$ the <u>second derivative</u> of f at p

Now, suppose f is twice-differentiable on all of U, then $p \mapsto (D^2 f)_p$ defines a map:
$$Df^2: U \to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$$
we will think of the ~~former~~ latter as: $\mathcal{L}(\mathbb{R}^{n^2}, \mathbb{R}^m)$
$$\mathcal{L}^2(\mathbb{R}^n, \mathbb{R}^m) := \cancel{\mathcal{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)} \; (=: \mathcal{L}^2) \cong M(m, n^2)$$
the space of bi-linear maps from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^m$.

Correct this review maps

Indeed, fix $p \in U$. Then
$$(D^2 f)_p \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)).$$
Therefore, for $v \in \mathbb{R}^n$,
$$(D^2 f)_p(v) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$$
(this makes sense, recall $(Df)_p(u) \in \mathbb{R}^m \leftarrow$ range of f).

Then for $w \in \mathbb{R}^n$, we have
$$((D^2 f)_p(v))(w) \in \mathbb{R}^m.$$
We write $\Downarrow = (D^2 f)_p(v, w)$.



10/29/2018

20

Ex: $f: \mathbb{R}^3 \to \mathbb{R}^2$, $f(p_1, p_2) = (p_1 + p_2, p_2^2 + p_3^3)$

Recall $(Df)_p = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2p_2 & 3p_3^2 \end{pmatrix}$ Claim $(D^2f)_p = \begin{array}{c} \begin{array}{ccccccccc} 11 & 12 & 13 & 21 & 22 & 23 & 31 & 32 & 33 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 6p_3 \end{pmatrix} \end{array}$

First compute $(Df)_{p+v} - (Df)_p$, then compute of norm & see which terms sublinear

$$(D^2f)_p(v, w) = \sum_{k=1}^{2} \sum_{i,j=1}^{3} (D^2f_k)_p(e_i, e_j)\, v_i w_j e_k$$

$(A_1 | A_2 | \cdots | A_n)$

$A_{ik} = \frac{\partial f_k(x)}{\partial x_i}$

Now, we can further ask if

$$D^2f : U \to \mathcal{L}^2 \ (\cong \mathbb{R}^{n n m})$$

is diff'ble. In this case, its derivative

$$D^3f = D(D^2f)$$

we think of as a map from U to tri-linear maps $\mathcal{L}^3(\mathbb{R}^n, \mathbb{R}^m) = \mathcal{L}(\mathbb{R}^{n^3}, \mathbb{R}^m)$, and so on.
$\cong M(m, n^3)$

In general, D^kf (if it exists) is a map from U to ~~k-linear maps~~ $\mathcal{L}(\mathbb{R}^{n^k}, \mathbb{R}^m)$

Thm If $(D^2f)_p$ exists, then $(D^2f_k)_p$ exists for each $k = 1, \ldots, m$, the second partial derivatives all exist, and

$$(D^2f_k)_p(e_i, e_j) = \frac{\partial^2 f_k(p)}{\partial x_i \partial x_j}$$

Conversely, if the second partials all exist in a nbhd of p and are cts there, then $(D^2f)_p$ exists.

1/31/2018

Pf: Assume $(D^2f)_p$ exists. Then $x \mapsto (Df)_x$ is diff'ble at $x = p$. Recall $(Df)_x = TM_x$ where

$$M_x = \left(\frac{\partial f_i(x)}{\partial x_j} \right)_{\substack{1 \le i \le n \\ 1 \le j \le n}}$$

By a previous thm, $x \mapsto (Df)_x$ is diff'ble only if each of it's components are, which implies each entry of M_x is diff'ble at p. Thus the second partial derivs exist. Furthermore:

$$(D^2 f_k)_p(e_i, e_j) = \big((D(Df_k))_p(e_i)\big)(e_j) = \lim_{t \to 0} \left(\frac{(Df_k)_{p+te_i} - (Df_k)_p}{t} \right)(e_j)$$

$= \lim_{t\to 0} \frac{(Df_k)_{p+te_i} - (Df_k)_p}{t}$

$$= \lim_{t \to 0} \lim_{s \to 0} \frac{1}{t}\big(f_k(p+te_i) - f_k \ldots$$

$$= \lim_{t \to 0} \frac{(Df_k)_{p+te_i}(e_j) - (Df_k)_p(e_j)}{t}$$

$$= \lim_{t \to 0} \frac{\frac{\partial f_k(p+te_i)}{\partial x_j} - \frac{\partial f_k(p)}{\partial x_j}}{t}$$

$$= \frac{\partial^2 f_k}{\partial x_i \partial x_j}(p)$$

exercise

Conversely, assume the second partials exist on a nbhd of p and are cts at p. Then, viewing M_x as valued in $\mathbb{R}^{nm}$, it's partial deriv exist and are cts at p. Hence its total derivative exists at p, which is exactly $(D^2 f)_p$ $\square$

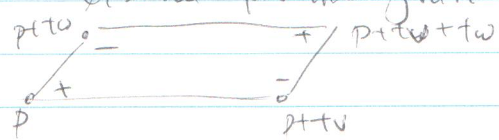
Thm If $(D^2 f)_p$ exists, then it is symmetric:

$$(D^2 f)_p(v,w) = (D^2 f)_p(w,v) \qquad \forall v,w \in \mathbb{R}^n$$

Pf: Since the symmetry concerns the domain space $\mathbb{R}^n$, we may assume $\mathbb{R}^m = \mathbb{R}$. For $t \in [0,1]$ consider the labeled parallelogram:

$D^2 f$ symm $\Leftrightarrow$ $D^2 f_u$ symm, $u = 1, \dots, m$

$v, w \in \mathbb{R}^n$



Define

$$\Delta = \Delta(t,v,w) := f(p+tv+tw) - f(p+tv) - f(p+tw) + f(p)$$

Observe that

$$\Delta(t,v,w) = \Delta(t,w,v)$$

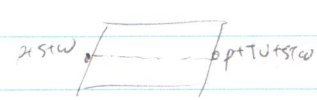
We claim

$$(D^2 f)_p(v,w) = \lim_{t \to 0} \frac{\Delta(t,v,w)}{t^2}$$

in which case symmetry of $(D^2 f)_p$ follows.

Fix t, v, w and write $\Delta = g(1) - g(0)$ for

$$g(s) = f(p+tv+stw) - f(p+stw)$$



Since f is diff'ble, so is g and by the MVT $\exists\, \theta \in (0,1)$ s.t.

$$\Delta = g(1) - g(0) = g'(\theta)\cdot(1-0) = g'(\theta)$$

By the chain-rule, we have

$$\Delta = g'(\theta) = (Df)_{p+tv+\theta tw}(tw) - (Df)_{p+\theta tw}(tw)$$
$$= t\{(Df)_{p+tv+\theta tw}(w) - (Df)_{p+\theta tw}(w)\}$$

We compute the Taylor remainder for $x \mapsto (Df)_x$:

$$(Df)_{p+x} = (Df)_p + (D^2 f)_p(x,\cdot) + (R(x))(\cdot)$$

Then $R(x,\cdot) := (R(x))(\cdot) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is sublinear wrt x. ~~This means~~ Applying this to $x = tv + \theta tw$ and $x = \theta tw$ yields:

$$\frac{\Delta}{t^2} = \frac{1}{t}\Big\{\big[(Df)_p(w) + (D^2f)_p(tv+\theta tw, w) + R(tv+\theta tw, w)\big] - \big[(Df)_p(w) + (D^2f)_p(\theta tw, w) + R(\theta tw, w)\big]\Big\}$$

$$= (D^2f)_p(v,w) + \frac{R(tv+\theta tw, w)}{t} - \frac{R(\theta tw, w)}{t}$$

Sublinearity of R completes the proof. $\square$

- Remark. The formula

$$(D^2f)_p(v,w) = \lim_{t\to 0} \frac{f(p+tv+tw) - f(p+tv) - f(p+tw) + f(p)}{t^2}$$

is the analogue of this 1-D formula:

$$f''(x) = \lim_{t\to 0} \frac{f(x+t) + f(x-t) - 2f(x)}{t^2}$$

- Cor If $f: U \to \mathbb{R}^m$ is twice-diff'ble at $p \in U$, then

$$\frac{\partial^2 f_k(p)}{\partial x_i \partial x_j} = \frac{\partial^2 f_k(p)}{\partial x_j \partial x_i} \qquad k=1,\dots,m.$$

Pf: Exercise.

- Remark: As with the first derivative, existence of the 2nd partial derivatives w/ cty at p does not imply the existence of $(D^2f)_p$. (See Homework #4)

- Cor For $f: U \to \mathbb{R}^m$, if $(D^rf)_p$ exists for some $p \in U$, then $(D^rf)_p$ is symmetric:

$$(D^rf)_p(v_1,\dots,v_r) = (D^rf)_p(v_{\sigma(1)},\dots,v_{\sigma(r)}) \quad \forall \sigma \in S_r$$

Moreover, corresponding mixed r-partial derivatives are equal.

Pf (induction) exercise. (Homework #4)

Prop: (Higher-order differentiation rules)

(a) $D^r(f+cg) = D^r f + c D^r g$

(b) $\beta: \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k} \to \mathbb{R}^m$ k-linear (i.e. $\beta \in \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}^m)$)

Define $f(p) := \beta(p, \ldots, p)$.

$$D^r f = \begin{cases} r!\, \mathrm{Symm}(\beta) & \text{if } r = k \\ 0 & \text{if } r > k \end{cases}$$

where

$$\mathrm{Symm}(\beta)(v_1, \ldots, v_r) = \frac{1}{r!} \sum_{\sigma \in S_r} \beta(v_{\sigma(1)}, \ldots, v_{\sigma(r)})$$

(c) Chain Rule:

$$(D^r(g \circ f))_p(v_1, \ldots, v_r) = \sum_{k=1}^{r} \sum_{\substack{\pi \in P(r) \\ \pi = \{B_1, \ldots, B_k\}}} (D^k g)_{f(p)}\left((D^{|B_1|} f)(v_{B_1}), \ldots, (D^{|B_k|} f)(v_{B_k})\right)$$

see?

(d) Product Rule

$$(D^r \beta(f,g))_p(v_1, \ldots, v_r) = \sum_{k=0}^{r} \sum_{\substack{\pi \in P(r) \\ \pi = \{B_1, B_2\}}} \beta\left((D^k f)_p(v_{B_1}), (D^{r-k} g)_p(v_{B_2})\right)$$

2/2/2018

Smoothness Classes:

Def: We say $f: U \to \mathbb{R}^m$ is of class C^r if it is r-times diff'ble on U and

$$D^r f: U \to \mathcal{L}^r(\mathbb{R}^n, \mathbb{R}^m) \cong \mathcal{L}(\mathbb{R}^{n^k}, \mathbb{R}^m)$$

is cts. (Note that this implies $D^k f$ is cts for $k = 0, \ldots, r-1$)

We say f is of class C^∞ or is smooth if it is of class C^r for all $r \in \mathbb{N}$.

Exercise: Determine how the ~~classes of differentiability~~ differing smoothness classes are mapped to each other via sums, composition, products etc.

Exercise: Use $\frac{\partial^r f_k}{\partial x_{i_1} \cdots \partial x_{i_r}}(p) = \langle (D^r f)_p(e_{i_1}, \ldots, e_{i_r}), e_k \rangle$ to show $D^r f$ cts iff $\frac{\partial^r f_k}{\partial x_{i_1} \cdots \partial x_{i_r}}$ cts.

Def: Let $(f_k)_{k \in \mathbb{N}}$ be a sequence of C^r functions $f_k: U \to \mathbb{R}^m$. We say the sequence is:

(a) uniformly C^r convergent if $\exists\, f: U \to \mathbb{R}^m$ of class C^r s.t.

unif. conv. $f_k \rightrightarrows f$, $Df_k \rightrightarrows f$, $\ldots$, $D^r f_k \rightrightarrows f$ as $k \to \infty$.

(b) uniformly C^r Cauchy if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ s.t. for all $k, l \geq N$ ~~and $p \in U$~~ we have

$\forall p \in U$ $|f_k(p) - f_l(p)| < \varepsilon$, $\|(Df_k)_p - (Df_l)_p\| < \varepsilon$, $\ldots$, $\|(D^r f_k)_p - (D^r f_l)_p\| < \varepsilon$

Thm: A sequence $(f_k)_{k \in \mathbb{N}}$ of C^r functions $f_k : U \to \mathbb{R}^m$ is uniformly C^r convergent if and only if it is uniformly C^r Cauchy.

Pf ($\Rightarrow$) Clear.

($\Leftarrow$) We proceed by induction. For $r = 1$, if $(f_k)_{k \in \mathbb{N}}$ is uniformly C^1 Cauchy, then by ~~completeness~~ Math 104 ~~of C^0~~ we know

$$f_k \rightrightarrows f \quad \text{and} \quad Df_k \rightrightarrows G$$

for some cts $f : U \to \mathbb{R}^m$ and some cts $G : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. We claim, $Df = G$. Indeed, fix $p \in U$ and consider $q \in U$ s.t. $[p, q] \subset U$. The C^1-MVT and unif. conv. imply:

$$f_k(q) - f_k(p) = \int_0^1 (Df_k)_{p + t(q-p)}\, dt\, (q - p)$$
$$\downarrow \qquad\qquad\qquad \downarrow$$
$$f(q) - f(p) \qquad \int_0^1 G_{p + t(q-p)}\, dt\, (q - p)$$

For q near p, $(p, q) \mapsto \int_0^1 G_{p+t(q-p)}\, dt$ is cts and so ~~by~~ the converse of the C^1-MVT, we have that f is of class C^1 with

$$(Df)_p = \int_0^1 G_{p + t(p-p)}\, dt = G_p$$

Thus $Df = G$ and $(f_k)_{k \in \mathbb{N}}$ ~~converges~~ is unif. C^1-conv.

Now, let $r \geq 2$. The ~~seq~~ sequence $(Df_k)_{k \in \mathbb{N}}$ is a uniformly C^{r-1}-Cauchy sequence. By the induction hypothesis, they are unif. C^{r-1} conv. to some ~~G~~ $G : U \to \mathbb{R}^m$ with

$$D^s(Df_k) \rightrightarrows D^s G \quad \text{for } s \leq r - 1$$

We also have $f_k \rightrightarrows f$ and by the above argument $Df = G$. Hence $(f_k)_{k \in \mathbb{N}}$ is unif. C^r conv. $\square$

Def The C^r-norm of a C^r function $f: U \to \mathbb{R}^m$ is

$$\|f\|_r := \max\{\sup_{p\in U}|f(p)|, \dots, \sup_{p\in U}\|(D^r f)_p\|\}$$

~~We~~ denote

$$C^r(U,\mathbb{R}^m) := \{f: U \to \mathbb{R}^m \mid f \text{ is of class } C^r \text{ and } \|f\|_r < \infty\}.$$

Cor: $C^r(U,\mathbb{R}^m)$ with $\|\cdot\|_r$ is a complete normed space (i.e. a Banach space)

Pf: Exercise. $\square$

Prop (C^r M-test): For $(f_k)_{k\in\mathbb{N}} \subseteq C^r(U,\mathbb{R}^m)$, if ~~s.t.~~ $\exists$ constants $(M_k)_{k\in\mathbb{N}} \subseteq [0,+\infty)$ s.t. $\|f_k\|_r \le M_k$ for each $k\in\mathbb{N}$ and $\sum M_k$ is conv., then $(f_k)_{k\in\mathbb{N}}$ is unif. C^r-convergent to some f and

$$D^s f = \sum_{k=1}^{\infty} D^s f_k \qquad \forall\, s \le r.$$

5.4 Implicit and Inverse Functions

Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open and fix $f: U \to \mathbb{R}^m$.
We fix a point $(x_0, y_0) \in U$ and suppose $f(x_0, y_0) = z_0$.
Our goal is to show that (under certain) conditions, the equation

$$f(x,y) = z_0$$

has a solution set of points near (x_0, y_0) for which $y = g(x)$ for some function g. That is,

$$f(x, g(x)) = z_0$$

and the solution set is the graph of g

$n = m = 1$