MATH 3472

We saw last day that $D_{1,2}f(\vec{c}) = D_{2,1}f(\vec{c})$ for $f:\mathbb{R}^2 \longrightarrow \mathbb{R}$, and in general

$$D_{r,k}f(\vec{c}) = D_{k,r}f(\vec{c})$$

for $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$, in both cases assuming suitable restrictions on f.

This means that of $D_{r,r}f(\vec{c})$, $D_{r,k}f(\vec{c})$, $D_{k,r}f(\vec{c})$ and $D_{k,k}f(\vec{c})$, only three are distinct.

In general, for higher-order partials $D_{r_1,\dots,r_\ell}f$ there are fewer distinct partials than our notation would lead one to believe, as when f is "sufficiently nice" the order of the indices $r_1,\dots,r_\ell$ does not matter. Thus if $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$, and $\ell > 0$, there are in general $\dfrac{n!}{(n-k)!\,\ell!} = C(n,\ell)$ partial derivatives of order ℓ.

§12.14

We have been calling

$$f(\vec{c}+\vec{v}) = f(\vec{c}) + T_c(\vec{v}) + \|\vec{v}\|\,E(\vec{v})$$

"Taylor's formula". Recall Taylor's formula for functions $f: \mathbb{R} \longrightarrow \mathbb{R}$ is as follows:

<u>Theorem 5.19</u> Suppose $f^{(n)}$ is defined on (a,b) and $f^{(n-1)}$ is continuous on $[a,b]$. Assume $c \in [a,b]$. Then for every $x \in [a,b]$, $x \neq c$, $\exists\, x_1$ between x and c such that

$$f(x) = f(c) + \sum_{k=1}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k + \underbrace{\frac{f^{(n)}(x_1)}{n!}(x-c)^n}_{\text{"error term"}}.$$

So it is reasonable to ask: To what extent can our generalized higher-dimensional Taylor formula be extended to approximate f using higher-order derivatives?

Suppose $f: \mathbb{R}^n \longrightarrow \mathbb{R}$. Defined special symbols

$$f''(\vec{x};\vec{t}),\ f'''(\vec{x};\vec{t}),\ \ldots,\ f^{(m)}(\vec{x};\vec{t})$$

to deal with the complicated formulas that arise:

Suppose $\vec{x} \in \mathbb{R}^n$ is a point where all second-order partial derivatives exist, and $t \in \mathbb{R}^n$ is arbitrary. Define

$$f''(\vec{x};\vec{t}) = \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} f(\vec{x})\, t_j t_i$$

where $t = (t_1, \ldots, t_n)$.

Similarly set

$$f'''(\vec{x};\vec{t}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} D_{i,j,k}(\vec{x})t_i t_j t_k$$

$$\vdots$$

$$f^{(m)}(\vec{x};\vec{t}) = \sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} D_{i_1,\dots,i_m}(\vec{x})t_{i_1}t_{i_2}\cdots t_{i_m}$$

whenever all third, fourth, ..., m^{th} order partials exist.

<u>Theorem 12.14</u>: Assume f and all partials of order less than m are differentiable on $S \subseteq \mathbb{R}$ open. Suppose $\vec{a}, b \in S$ satisfy $L(\vec{a}, b) \subseteq S$. Then $\exists\, z \in L(\vec{a}, b)$ s.t.

$$f(\vec{b}) - f(\vec{a}) = \sum_{k=1}^{m-1} \frac{1}{k!} f^{(k)}(\vec{a}; \vec{b}-\vec{a}) + \frac{1}{m!} f^{(m)}(\vec{z}; \vec{b}-\vec{a}).$$

<u>Proof</u>: Choose $\delta > 0$ st $\vec{a} + t(\vec{b}-\vec{a}) \in S \;\; \forall t \in (-\delta, 1+\delta)$. This is possible since $L(\vec{a}, \vec{b}) \subseteq S$ and S is open.

Define $g : (-\delta, 1+\delta) \longrightarrow \mathbb{R}$ by

$$g(t) = f(\vec{a} + t(\vec{b}-\vec{a})),$$

so we're "making f into a function of one variable by considering f along the line $L(\vec{a}, \vec{b})$ only".

Now the one-dimensional Taylor formula applied to g is

$$f(\vec{b}) - f(\vec{a}) = g(1) - g(0) = \sum_{k=1}^{m-1} \frac{1}{k!} g^{(k)}(0) + \frac{1}{m!} g^{(m)}(\theta)$$

where $0 < \theta < 1$. This comes from $x=1$ $c=0$ in the Taylor formula previously given.

We can think of g as a composite in order to compute its derivatives in terms of the derivatives of f, namely

$g(t) = f(p(t))$, where $p: \underset{[0,1]}{\underset{\cup}{\mathbb{R}}} \longrightarrow \mathbb{R}^n$ is the function

$p(t) = \vec{a} + t(\vec{b} - \vec{a})$. Then $p = (p_1, \dots, p_n)$, where

$p_k(t) = a_k + t(b_k - a_k)$ and thus $p_k'(t) = b_k - a_k$.

Therefore by the chain rule

$$g'(t) = \sum_{j=1}^{n} D_j f(p(t))(b_j - a_j) = f'(p(t); \vec{b} - \vec{a})$$

applying the chain rule again, each term $D_j f(p(t))(b_j - a_j)$ yields

$$(b_j - a_j) \sum_{i=1}^{n} D_{i,j} f(p(t))(b_i - a_i)$$

So that overall

$$g''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} D_{i,j} f(p(t))(b_j - a_j)(b_i - a_i) = f''(p(t); \vec{b} - \vec{a})$$

Similarly we find

$$g^{(m)}(t) = f^{(m)}(p(t); \vec{b} - \vec{a}) \quad \text{for all } m.$$

Substitute these values into the 1-dimensional Taylor formula for g in order to obtain the theorem.

§ 13.1 :

Suppose $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is given by some formula, and $f(x;t) = 0$ is used to define some set in $\mathbb{R}^2$.

Q: Does the set $\{(x;t) \in \mathbb{R}^2 \mid f(x;t) = 0\}$ define a function $x = g(t)$?

Example: (Wikipedia). Suppose $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is $f(x,y) = x^2 + y^2 - 1$. Then $f(x,y) = 0 \Leftrightarrow x^2 + y^2 = 1$

So $\{(x,y) \mid f(x,y) = 0\}$ is a circle.

This does not define the graph of a function, however for each point (x,y) on the circle, if $-1 < x < 1$ and (x,y) is on the upper half, then

$$g_1(x) = \sqrt{1 - x^2}$$

provides a function whose graph locally agrees with $f(x,y) = 0$. If (x,y) is on the lower half, $g_2(x) = -\sqrt{1-x^2}$ provides a function whose graph locally agrees with $f(x,y) = 0$.

Thus for all (x,y) in $\{f(x,y) = 0\} \setminus \{(-1,0), (1,0)\}$ there is a ~~fung~~ function $g(x)$ and an open set $\mathcal{U}$ with $(x,y) \in \mathcal{U}$ such that

$$(x,y) = (x, g(x)) \text{ for all } (x,y) \text{ in } \{f(x,y) = 0\} \cap \mathcal{U}.$$

The purpose of the implicit function theorem will be to gives us functions like $g_1(x)$, $g_2(x)$ that locally define $f(x_1, \dots, x_n) = 0$ as a graph - even when we cannot compute formulas explicitly as in the previous example.

<u>Example</u>: Consider the system of linear equations

$$\sum_{j=1}^{n} a_{ij} x_j = t_i \quad (i = 1, \dots, n)$$

This ~~&~~ system has a unique solution iff $\det[a_{ij}] \neq 0$.

Each equation in the system can be rewritten as

$$f_i(\vec{x}, \vec{t}) = \sum_{j=1}^{n} a_{ij} x_j - t_j = 0$$

where $\vec{x} = (x_1, \dots, x_n)$ and $\vec{t} = (t_1, \dots, t_n)$.

So the system in its entirety is captured by setting $f = (f_1, \dots, f_n)$ and writing the vector equation

$$f(\vec{x}, \vec{t}) = \vec{0}.$$

Since $D_j f_i(\vec{x}, \vec{t}) = a_{ij}$, the Jacobian of $f(\vec{x}; \vec{t})$ is just the coefficient matrix of the corresponding system of equations. If the Jacobian is J, and is invertible with inverse J^{-1}, then (so $\det J \neq 0$)

the solution to the system is

$$\vec{x} = J^{-1}\vec{t}$$

so the solutions to $f(\vec{x},\vec{t})=0$ can be expressed as points $(J^{-1}\vec{t},\vec{t})$.

In general, nonzero Jacobian determinant ~~turns~~ at a point $\vec{c}\in\mathbb{R}^n$ turns out to be the essential ingredient in finding a set around $\vec{c}$ and a function $\vec{g}$ such that the points near $\vec{c}$ satisfying $f(\vec{x},\vec{t})=0$ can be expressed as $(g(t),t)$.
This example certainly shows that $\det J\neq 0$ is necessary, but it will turn out to be sufficient as well.

Notation: If $f = (f_1, \ldots, f_n)$ and $\vec{x} = (x_1, \ldots, x_n)$, then $Df(\vec{x}) = [D_j f_i(\vec{x})]$ is a square matrix. Its determinant is called the Jacobian determinant of f at $\vec{x}$ and is denoted $J_f(\vec{x})$.

§13.2: Functions with nonzero Jacobian determinant.

We prepare several lemmas and properties of functions with nonzero Jacobian determinant in order to use them later.

Theorem: Let $B(\vec{a}, r) \subset \mathbb{R}^n$ be a ball of radius $r > 0$ centred at $\vec{a} \in \mathbb{R}^n$. Write $\partial B = \{\vec{x} \mid \|\vec{x} - \vec{a}\| = r\}$ for the boundary of $B(\vec{a}, r)$, and let $\overline{B} = B \cup \partial B$ (the closure). Suppose $f = (f_1, \ldots, f_n)$ is continuous on $\overline{B}$, and that $D_j f_i(\vec{x})$ exists $\forall \vec{x} \in B$. Further assume that $f(\vec{x}) \neq f(\vec{a})$ $\forall \vec{x} \in \partial B$ and that $J_f(\vec{x}) \neq 0$ $\forall x \in B$.

Then $f(\vec{a})$ is in the interior of $f(B)$.

Proof: Define $g: \partial B \longrightarrow \mathbb{R}$ as follows:

$$g(\vec{x}) = \|f(\vec{x}) - f(\vec{a})\|.$$

Then $g(\vec{x}) > 0$ for all $\vec{x} \in \partial B$ since f does not map ∂B to the same point as $\vec{a}$, and g is continuous because f is continuous on ∂B.

Now since g is a continuous function with compact domain, so it attains a minimum $m>0$ somewhere on ∂B.

Set
$$T = B(f(\vec{a}); \tfrac{m}{2}).$$

We'll see $T \subset f(B)$, which proves the theorem.

So let $y \in T$, and define a function $h: \overline{B} \longrightarrow \mathbb{R}$ by
$$h(x) = \|f(x) - \vec{y}\| \quad \forall x \in \overline{B}.$$

Then h is <u>also</u> a function with compact domain and so attains a minimum on $\overline{B}$, we'll check that its minimum is attained on B (<u>not</u> on ∂B).

First observe that plugging $\vec{a}$ into h yields
$$h(\vec{a}) = \|f(\vec{a}) - \vec{y}\| < m/2$$
since $y \in T$, so whatever the minimum of h is, it must be smaller than $m/2$.

On the other hand, plugging in an arbitrary point $\vec{x}$ on ∂B gives
$$\begin{aligned} h(\vec{x}) &= \|f(\vec{x}) - \vec{y}\| \\ &= \|f(\vec{x}) - \vec{y} - f(\vec{a}) + f(\vec{a})\| \\ &\geq \|f(\vec{x}) - f(\vec{a})\| - \|f(\vec{a}) - \vec{y}\| > g(x) - \tfrac{m}{2} \geq \tfrac{m}{2}, \end{aligned}$$
so $h(\vec{x})$ does not attain its minimum on ∂B.

Let $\vec{c} \in B$ denote the point where h attains a minimum.

We will show that $f(\vec{c}) = \vec{y}$, so that $\vec{y} \in f(B)$.

Since $\vec{y}$ was an arbitrary point in T, this shows $T \subset B$ and completes the proof.

To see this, note $h(\vec{x})^2$ also has a minimum at $\vec{c}$, and

$$h(\vec{x})^2 = \|f(x) - \vec{y}\|^2 = \sum_{r=1}^{n} (f_r(\vec{x}) - y_r)^2 \text{ where } \vec{y} = (y_1, \dots, y_n).$$

The partial derivatives $D_k h^2$ must be zero at $\vec{c}$, since $\vec{c}$ is a minimum, so plugging in $\vec{c}$ above and deriving:

$$2 \sum_{r=1}^{n} (f_r(\vec{c}) - y_r) D_k f_r(\vec{c}) = \vec{0}, \text{ for } k = 1, \dots, n.$$

~~So we have a system of equations~~

Considering the system of equations

$$0 = \sum_{r=1}^{n} b_r D_k f_r(\vec{c}), \quad \text{for } k = 1, \dots, n$$

since the coefficient matrix is the Jacobian at $\vec{c}$, and we assumed $J_f(\vec{c}) \neq 0$ (as $\vec{c} \in B$), there is a unique solution: $b_k = 0 \ \forall k$.

Thus we have $f_r(\vec{c}) = y_r$ for all r, in other words $f(\vec{c}) = \vec{y}$ and $\vec{y} \in f(B)$ as claimed.

So if the Jacobian is nonzero in the nbhd of a point, the function very roughly "maps open balls to something with nonempty interior". But in fact, we get more if we ask for nonzero Jacobian on a set:

Next, we'll see that $J_f(x) \neq 0$ actually forces f to be one-to-one and open, but we will save this investigation for after the break.