Monday, Sep 10

**THE Derivative.**

A function \( f \) of a single variable is differentiable at a point \( a \) if the limit

\[
\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a)
\]

exists, and the number \( f'(a) \) is the derivative of this function at this point.

Q How to generalize this notion?

We cannot simply copy the definition, because in general \( h \) will be a vector, and what does \( f'(a) \) mean then to divide by \( h \)?

Let us rewrite the above as

\[
\lim_{h \to 0} \left( \frac{f(a+h) - f(a)}{h} - f'(a) \right) = 0,
\]

or

\[
\lim_{h \to 0} \frac{1}{|h|} \left| f(a+h) - f(a) - f'(a) h \right| = 0.
\]

The function \( h \mapsto f(a+h) - f(a) \) is approximated by the linear transformation

\( df_a: \mathbb{R} \to \mathbb{R} \) defined by \( df_a(h) = f'(a) h \).

First order (the difference goes to 0 faster than \(|h|\)).
Def \( F : \mathbb{R}^n \to \mathbb{R}^m \) is differentiable at \( a \) if there is a linear transformation \( dF_a : \mathbb{R}^n \to \mathbb{R}^m \) such that
\[
\lim_{h \to 0} \frac{1}{|h|} \left| (F(a+h) - F(a)) - dF_a(h) \right| = 0.
\]

Q: What is \( dF_a \)? How can we find a matrix for it?

**Proposition** If \( F \) is differentiable at \( a \), then the \( m \times n \) matrix representing \( dF_a \) is
\[
dF_a = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\
\vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a)
\end{pmatrix}.
\]

Q: Why is this true? Let us work out the \( 2 \times 2 \) case. The general proof follows along the same lines.

**Example** \( F(x,y) = (f_1(x,y), f_2(x,y)) \)

Suppose \( F \) is diff. at \((a,b)\) with \( dF_{(a,b)} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \)

How can we find the \( c_{ij} \)?
Look at the definition for special \( \mathbf{h} \)...

First, let \( \mathbf{h} = (h, 0) \). Then

\[
\frac{1}{|h|} \left| F(a+h, b) - F(a, b) - \begin{pmatrix} c_{11} h \\ c_{21} h \end{pmatrix} \right|
\]

\[
= \frac{1}{|h|} \sqrt{\left( f_1(a+h, b) - f_1(a, b) - c_{11} h \right)^2 + \left( f_2(a+h, b) - f_2(a, b) - c_{21} h \right)^2}
\]

This goes to 0, so

\[
c_{11} = \frac{\partial f_1}{\partial x}(a, b) \quad \text{and} \quad c_{21} = \frac{\partial f_2}{\partial x}(a, b)
\]

Similarly, using \( \mathbf{h} = (0, h) \) gives

\[
c_{12} = \frac{\partial f_1}{\partial y}(a, b) \quad \text{and} \quad c_{22} = \frac{\partial f_2}{\partial y}(a, b)
\]

(Compare this with the Proposition.)
Theorem: If the partial derivs. \( \frac{\partial f}{\partial x_j} \) exist and are continuous at \( a \), then \( F \) is differentiable at \( a \).

Proof: See appendix at text for 2d version.

**Exercise:** Polar coordinates given by transformation

\[ G : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \]

\[ (r, \theta) \rightarrow (r \cos \theta, r \sin \theta) \]

\[ dG_{(r, \theta)} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \]

\[ dG_{(r, \theta)}(1) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{radial unit vector} \]

\[ dG_{(r, \theta)}(i) = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \end{pmatrix} = r \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \quad \text{vector, scaled by } r. \]
For $f : \mathbb{R}^n \to \mathbb{R}$, what is $df$?

It is a $1 \times n$ matrix

$$
    df(x_1, \ldots, x_n) = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right),
$$

which, if we view as a vector, is called the gradient of $f$:

$$
    \nabla f(x_1, \ldots, x_n) = \left< \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right>.
$$

Thus

$$
    df(x_1, \ldots, x_n)(\mathbf{v}) = \nabla f(x_1, \ldots, x_n) \cdot \mathbf{v}.
$$

The linearization or degree 1 Taylor poly of $f$ at $a = (a_1, \ldots, a_n)$ is

$$
    L(x) = f(a) + df_a(x-a) = f(a) + \sum_{\substack{i=1}}^{n} \frac{\partial f}{\partial x_i}(a)(x_i-a_i) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n-a_n).
$$
Ex \( f(x,y) \), \((a,b) \in \mathbb{R}^2 \).

Thus \( L(x, y) = f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b) \).

The graph of \( L \) is the \underline{tangent plane} to the graph \( z = f(x, y) \) at \((a, b)\).

This plane contains the tangent lines to the cross-sections of the surface through \((a, b)\).