MATH2010E LECTURE 6: PROPERTIES FOR DIFFERENTIATIONS

1. Differentiations and partial derivatives

Let $f : \Omega \to \mathbb{R}$ be a differentiable function, $\Omega \subset \mathbb{R}^n$ open. We called f a C^1-function if and only if all first order partial derivatives $\frac{\partial f}{\partial x_i}$ are continuous. Similarly, we call a C^1-function f be a C^2-function if and only if all second order partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ are continuous, $i, j \in \{1, 2, ..., n\}$.

Proposition 1.1. If f is a C^1 function, then f is differentiable on Ω.

Remark 1.2. The existence of first order partial differentiation at a point, however, can not imply the differentiation of a function at that point.

Proof. Recall from the last Lecture, we showed that if f is differentiable at p, then its linear approximation will be

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)(x_i - p_i) + f(p) \tag{1.1}$$

near p. Let us denote this function by L. We use $\{\vec{e}_i\}$ to denote the standard orthonormal basis of $\mathbb{R}^n$. For any $q \neq p$, we write $q - p = \sum_{i=1}^{n} h_i \vec{e}_i$. Then

(1.2)

$$\begin{aligned} f(q) =& \Big(f(q) - f(p)\Big) + f(p) \\ =& \Big(f(p + \sum_{i=1}^{n} h_i\vec{e}_i) - f(p + \sum_{i=1}^{n-1} h_i\vec{e}_i)\Big) + \Big(f(p + \sum_{i=1}^{n-1} h_i\vec{e}_i) - f(p + \sum_{i=1}^{n-2} h_i\vec{e}_i)\Big) \\ &+ \cdots + \Big(f(p + h_1\vec{e}_1) - f(p)\Big) + f(p) \\ =& \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(p + \sum_{i=1}^{j-1} h_i\vec{e}_i + \xi_j\vec{e}_j)h_j + f(p) \end{aligned}$$

by mean value theorem (The value $\xi_j \in (0, h_j)$ if $h_j > 0$ and $\xi_j \in (h_j, 0)$ if $h_j < 0$. When $h_j = 0$, we can simply take $\xi_j = 0$).
If we take $q = x$, then the RHS of (1.2) is read as

$$\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(p + \sum_{i=1}^{j-1} h_i\vec{e}_i + \xi_j\vec{e}_j)(x_j - p_j) + f(p). \tag{1.3}$$

Combine this equality and (1.1), we have

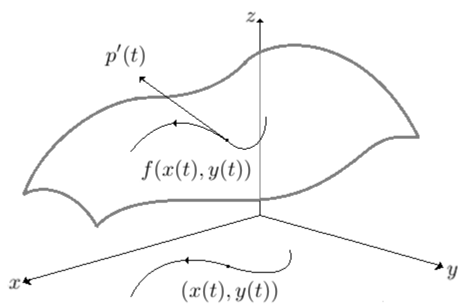
$$\frac{|f(x) - L(x)|}{|x - p|} \leq \sum_{j=1}^{n} |\frac{\partial f}{\partial x_j}(p + \sum_{i=1}^{j-1} h_i\vec{e}_i + \xi_j\vec{e}_j) - \frac{\partial f}{\partial x_i}(p)| \tag{1.4}$$

Because $f \in C^1$, so RHS of (1.4) converges to 0 as our wish. $\square$

2. Chain rule

Let f be a C^1-function defined on $\Omega \subset \mathbb{R}^2$. The graph of f gives a surface defined on $\mathbb{R}^3$. Now, suppose we have a parametrized curve $p(t) = (x(t), y(t), f(x(t), y(t)))$ on this surface, $t \in (a, b)$. Then, we can differentiate p with respect to t to obtain the tangent vector:

$$p'(t) = (x'(t), y'(t), \frac{d}{dt}(f(x(t), y(t)))). \tag{2.1}$$



Pic. 1

To obtain the third component, we have to use the chain rule:

$$\frac{d}{dt}(f(x(t), y(t))) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}. \tag{2.2}$$

It is not hard to prove that this tangent vector is on the tangent plane of f.

We should also consider problem of the changing coordinates. Let $x(r, s), y(r, s)$ be differentiable functions $(x, y) : \Omega \to \mathbb{R}^2$ with its range containing Ω. Then we can parametrize f by (r, s) and compute the partial derivatives by the chain rule:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}; \tag{2.3}$$

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}. \tag{2.4}$$

3. Applications

Implicit differentiations. An important application of the formula in (2.2) is the differentiation of implicit functions. Let $F(x, y) = 0$ be a implicit function. One may expect that $y = f(x)$ locally, or $x = f(y)$. Suppose that we have $y = f(x)$ satisfying $F(x, f(x)) = 0$ near x_0. A question we can ask is finding the derivative

$$\frac{dy}{dx}(x_0). \tag{3.1}$$

To solve this problem, we can choose a parameter t such that $x = t$ and $y = f(t)$ near x_0. By applying (2.2) directly, we have

$$0 = \frac{d}{dt}F(x(t), y(t)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dt} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx}. \tag{3.2}$$

Here we simply use the notations $\partial_x F$ to denote $\frac{\partial F}{\partial x}$ and $\partial_y F$ to denote $\frac{\partial F}{\partial y}$. Then

$$\frac{dy}{dx} = -\frac{\partial_x F}{\partial_y F}. \tag{3.3}$$

Directional derivatives. Let $f : \Omega \to \mathbb{R}$ be a C^1, n-variable function. We have all partial derivatives exist and continuous on Ω. Define

$$\nabla(f) = \Big(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\Big) \tag{3.4}$$

to be the **gradient** of f. Notice that this is a vector value function $\nabla f : \Omega \to \mathbb{R}^n$. Let $p \in \Omega$ and $\vec{v} \in \mathbb{R}^n$, we can also define the direction derivative

$$D_{\vec{v}} f(p) = \nabla(f)(p) \cdot \vec{v} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(p) v_j. \tag{3.5}$$

One can consider the line $p(t) = p + t\vec{v}$, then $D_{\vec{v}} f(p)$ is just the z component of formula (2.1) at $t = 0$. Geometrically, this is the derivative of f along the $\vec{v}$ direction, multiplied by $|\vec{v}|$.

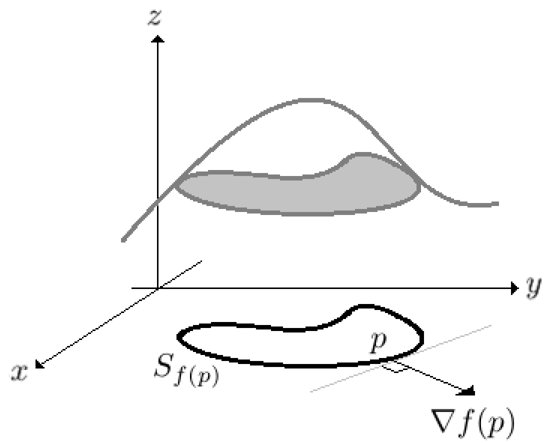
According to this observation, we consider all $\vec{v}$ with $|\vec{v}| = 1$. Then we notice that the maximum of $D_{\vec{v}} f(p)$ happens if and only if $\vec{v} = \frac{\nabla f(p)}{|\nabla f(p)|}$, unless $\nabla f(p) = 0$. Namely, $\vec{v}$ and $\nabla f(p)$ have the same direction. Meanwhile, when $D_{\vec{v}} f(p)$ attaches maximum, f increases most along $\vec{v}$.

Let $\nabla f(p) \neq 0$. Suppose $n = 2$ and $p \in S_{f(p)}$ for a level curve $S_{f(p)}$. Then we have the following theorem.

Theorem 3.1. *We have the tangent line $S_{f(p)}$ at p is perpendicular to $\nabla f(p)$.*

Proof. To prove this theorem, we should know the tangent direction of f at $f(p)$ first. Let us call this direction $\vec{v}$. Notice that along this direction, f stays a constant. So $D\vec{v} f(p) = 0$. Therefore, we have $\vec{v} \cdot \nabla f(p) = 0$, which proves this theorem. $\square$

The same result holds for $n = 3$ and $S_{f(p)}$ being a surface. $\nabla f(p)$ is perpendicular to the tangent plane of f at p. We will prove this later.



Pic. 2