Functions of Several Variables.

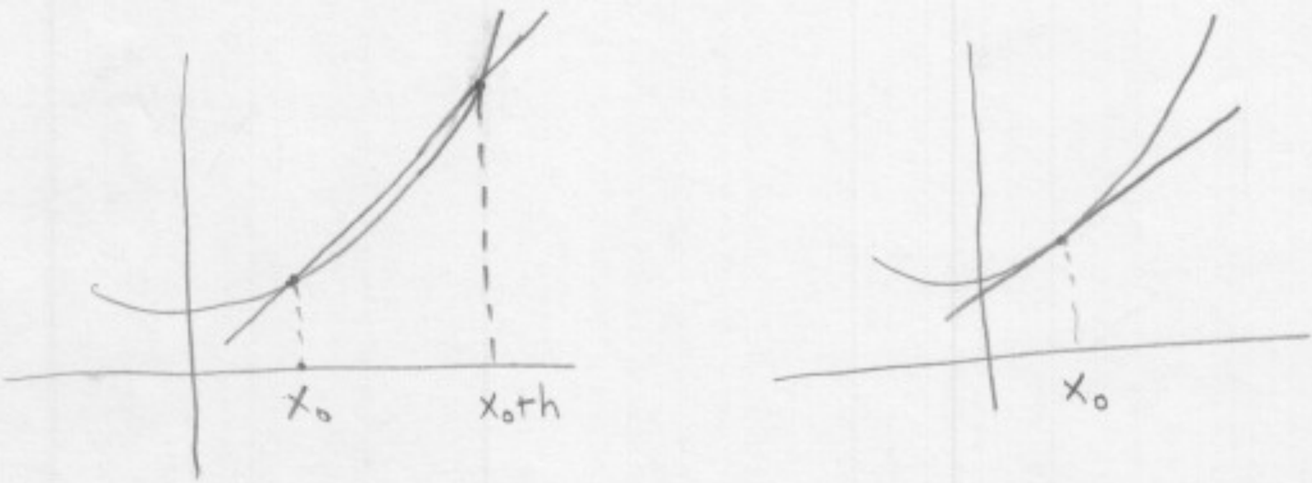
We first recall, for functions of one variable, we have:

$$f:(a,b)\to \mathbb{R}, \quad x_0\in(a,b)$$

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$$

or, equivalently,

$$f'(x_0)=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}, \quad x=x_0+h$$



We can rewrite the above limit as follows,

$$\lim_{h\to 0}\left[\frac{f(x_0+h)-f(x_0)}{h}-f'(x_0)\right]=0$$

or $$\lim_{h\to 0}\left|\frac{f(x_0+h)-f(x_0)-f'(x_0)h}{h}\right|=0$$

Note that:

$$\ell(x) = f(x_0) + f'(x_0)h, \quad h = x - x_0$$
$$= f(x_0) + f'(x_0)(x - x_0)$$

is the equation of the tangent line to the graph at x_0.

We also have, with

$$r(h) := f(x_0+h) - \ell(x_0+h)$$

that:

$$\lim_{h \to 0} \frac{r(h)}{h} = 0$$

Remark: We can regard the derivative of f at x_0, not as a real number, but as a linear operator:

$$\left.\begin{array}{l} A : \mathbb{R} \to \mathbb{R} \\ A(h) = f'(x_0)h \end{array}\right\} (*)$$

From linear algebra, if X and Y are vector spaces, we define:

$$L(X,Y) = \{ A : X \to Y : A \text{ is linear} \}$$

Recall that A is linear if:

$$A(x_1 + x_2) = Ax_1 + Ax_2 \quad \forall x_1, x_2 \in X$$
$$A(cx) = cA(x), \quad \forall c \in \mathbb{R}, x \in X$$

Therefore, the linear operator in (*) belongs to $L(\mathbb{R}, \mathbb{R})$.

Conversely, every linear function that carries $\mathbb{R}^1$ to $\mathbb{R}^1$ is multiplication by some real number.

For function of several variables we have the following:

<u>Def 9.11</u> : Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, E open, $\vec{x}_0 \in E$.

If there exists a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$\lim_{\vec{h} \to 0} \frac{|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - A\vec{h}|}{|\vec{h}|} = 0 \qquad (1)$$

then we say that f is differentiable at $\vec{x}_0$, and we write

$$f'(\vec{x}_0) = A$$

If f is differentiable at every $\vec{x} \in E$, we say that f is differentiable in E.

<u>Theorem 9.12</u> : Suppose E and f are as in Definition 9.11, $\vec{x} \in E$, and (1) holds with $A = A_1$ and with $A = A_2$. Then $A_1 = A_2$

Remarks :

(a) The relation in (1) can be rewritten in the form
$$f(\vec{x}_0+\vec{h}) - f(\vec{x}_0) = f'(\vec{x}_0)\vec{h} + r(\vec{h}), \qquad (2)$$
where the remainder $r(\vec{h})$ satisfies:
$$\lim_{\vec{h}\to 0} \frac{|r(\vec{h})|}{|\vec{h}|} = 0.$$

Hence:
$$f(\vec{x}_0+\vec{h}) = f(\vec{x}_0) + f'(\vec{x}_0)\vec{h} + \vec{r}(\vec{h})$$
$$\Rightarrow \boxed{f(\vec{x}_0+\vec{h}) \sim f(\vec{x}_0) + f'(\vec{x}_0)\vec{h}} \qquad (**)$$

So we can approximate $f(\vec{x}_0+\vec{h})$ with the function $f(\vec{x}_0)+f'(\vec{x}_0)\vec{h}$, where $f'(\vec{x}_0): \mathbb{R}^n \to \mathbb{R}^m$ is a linear function; i.e., $f'(\vec{x}_0) \in L(\mathbb{R}^n, \mathbb{R}^m)$

(b) Suppose $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable in E. For every $x \in E$, $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$. Hence, f' can be considered as a function;
$$f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$$

(c) From (2) clearly follows that if f is differentiable at $\vec{x}$ then f is continuous at $\vec{x}$.

Definition 9.16. Partial derivatives

Let $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$, i.e;

$$f(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x}))$$

For $\vec{x} \in E$, $1 \le i \le m$, $1 \le j \le n$, we define:

$$D_j f_i(\vec{x}) = \lim_{t \to 0} \frac{f_i(\vec{x} + te_j) - f_i(\vec{x})}{t}$$

$$\left(\text{or } \frac{\partial f_i}{\partial x_j}\right)$$

where $\{e_1, \dots, e_n\}$ is the standard bases of $\mathbb{R}^n$, i.e., $e_j = (0, 0, \dots, \underset{\uparrow j \text{ component}}{1}, 0, \dots 0)$

Theorem 9.17: Suppose $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{x} \in E$. Then, the partial derivaties exist; i.e:

$$\frac{\partial f_i}{\partial x_j} \text{ exist}, \quad \forall i \in \{1, \dots, m\} \quad \forall j \in \{1, \dots, n\}$$

Moreover, the linear transformation $f'(\vec{x}): \mathbb{R}^n \to \mathbb{R}^m$ is the $m \times n$ matrix:

$$f'(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{x}) \end{pmatrix}$$

<u>Definition</u>: A differentiable mapping:

$$f : E \subset \mathbb{R}^n \to \mathbb{R}^m$$

is said to be continuously differentiable in E if the function:

$$f' : E \to L(\mathbb{R}^n, \mathbb{R}^m),$$

is a continuous mapping. Recall that $L(\mathbb{R}^n, \mathbb{R}^m)$ is the set of all linear maps from $\mathbb{R}^n$ into $\mathbb{R}^m$. If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, the norm of A is defined as:

$$\|A\| := \sup_{|\vec{x}| \leq 1} |A\vec{x}|$$

We define $\mathscr{C}'(E)$ to be the set of all differentiable functions $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ such that $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous.

More precisely, if $f \in \mathscr{C}'(E)$ we have that to every $\vec{x} \in E$ and to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that

if $\vec{y} \in E$, $|\vec{x} - \vec{y}| < \delta$ then $\| f'(\vec{y}) - f'(\vec{x}) \| < \varepsilon$.

<u>Theorem 9.21</u>: Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$. Then:
$f \in \mathscr{C}'(E)$ if and only if the partial derivatives $D_j f_i$ exist and are continuous on E, for $1 \leq i \leq m$, $1 \leq j \leq n$.

Ex: Let $f: \mathbb{R}^2 \to \mathbb{R}^3$

$$f(x,y) = (xe^{2y}, yx^2).$$

Compute $f(0.9, 0.2)$ without using calculator.

We have

$$f(x,y) = (f_1(x,y), f_2(x,y))$$

$$f_1(x,y) = xe^{2y}, \quad f_2(x,y) = yx^2.$$

Since $\frac{\partial f_1}{\partial x} = e^{2y}$, $\frac{\partial f_1}{\partial y} = 2xe^{2y}$, $\frac{\partial f_2}{\partial x} = 2xy$, $\frac{\partial f_2}{\partial y} = x^2$ are all continuous functions in $\mathbb{R}^2$, Theorem 9.21 gives that $f \in \mathscr{C}'(\mathbb{R}^2)$. In particular, f is differentiable at any point of $\mathbb{R}^2$. We can then use (**) to estimate:

$$f(\vec{x}_0 + \vec{h}) \sim f(\vec{x}_0) + f'(\vec{x}_0)h$$

In this example we can take $\vec{x}_0 = (1,0)$ and $\vec{h} = (0.9, 0.2) - (1, 0) = (-0.1, 0.2)$.

$$f'(\vec{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(\vec{x}_0) & \frac{\partial f_1}{\partial y}(\vec{x}_0) \\ \frac{\partial f_2}{\partial x}(\vec{x}_0) & \frac{\partial f_2}{\partial y}(\vec{x}_0) \end{pmatrix} = \begin{pmatrix} e^{2(0)} & 2(1)e^{2(0)} \\ 2(1)(0) & (1)^2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$f(1,0) = (1,0)$$

Therefore:

$$f(0.9, 0.2) \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0.1 \\ 0.2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -0.1 + 0.4 \\ 0 + 0.2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix}$$

$$= \begin{pmatrix} 1.3 \\ 0.2 \end{pmatrix}$$

$\Rightarrow$

$$f(0.9, 0.2) \cong (1.3, 0.2)$$

The true value is:

$$f(0.9, 0.2) = (0.9\, e^{0.4}, \; 0.2\,(0.9)^2)$$
$$= (1.34\ldots, \; 0.162\ldots)$$

- Remark: If $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable, we have seen that $f'(\vec{x})$ is a linear map, i.e., $f'(\vec{x}) \in L(\mathbb{R}^n, \mathbb{R}^m)$. Moreover, this linear map is unique and given by the $m \times n$ matrix:

$$f'(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}) \end{pmatrix}$$

$\vec{x} = (x_1, \ldots, x_n)$ $\quad f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$.

For the particular case:

$$m = 1, \quad n > 1,$$

We have $f: E \subset \mathbb{R}^n \to \mathbb{R}$. In this case, $f'(\vec{x})$ is a $1 \times n$ matrix

$$f'(\vec{x}) = \left(\frac{\partial f}{\partial x_1}(\vec{x}) \quad \ldots \quad \frac{\partial f}{\partial x_n}(\vec{x}) \right)$$

We can also consider this matrix as a vector, which is denoted as the gradient vector $\nabla f(\vec{x})$:

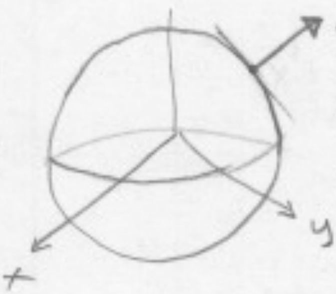
$$\nabla f(\vec{x}) = \left(\frac{\partial f}{\partial x_1}(\vec{x}), \frac{\partial f}{\partial x_2}(\vec{x}), \ldots, \frac{\partial f}{\partial x_n}(\vec{x}) \right).$$

Besides the approximation property (**), which in this particular case means that we are approximating the function with the tangent plane, the gradient vector has two fundamental geometrical applications:

1.- The gradient vector is perperdicular to the level surfaces (or level curves) of the function:

Ex: $f(x,y,z) = x^2+y^2+z^2$

$x^2+y^2+z^2=1$



$\nabla f\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ is perpendicular to the tangent plane at $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$

2.- The gradient vector points in the direction of maximum increase of the function.

Indeed, if $\vec{u}$ is a unit vector in $\mathbb{R}^n$, recall that the directional derivative of f at $\vec{x}$ is defined as:

$$D_{\vec{u}} f(\vec{x}) = \lim_{t \to 0} \frac{f(\vec{x}+t\vec{u}) - f(\vec{x})}{t}$$

other notation: $\frac{\partial f}{\partial \vec{u}}(\vec{x})$.

Theorem 9.15: If $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ and $g: \mathbb{R}^m \to \mathbb{R}^p$ are both differentiable at $\vec{x}_0$ an $f(\vec{x}_0)$ respectively, then (chain rule in several variables) the composition function:

$$h(\vec{x}) = g(f(\vec{x})); \quad h: E \subset \mathbb{R}^n \to \mathbb{R}^p$$

is differentiable at $\vec{x}_0$ and:

$$\underbrace{h'(x_0)}_{p\times n \text{ matrix}} = \underbrace{g'\,(f(\vec{x}_0))}_{p\times m \text{ matrix}} \cdot \underbrace{f'(\vec{x}_0)}_{m\times n \text{ matrix}}$$

Using the chain rule, if $\vec{f}$ is differentiable at $\vec{x}$, it is easy to show that:

$$D_{\vec{u}} f(\vec{x}) = \nabla f(\vec{x}) \cdot \vec{u}, \quad \vec{f}: E \subset \mathbb{R}^n \to \mathbb{R}$$

Since $\nabla f(\vec{x}) \cdot \vec{u} = |\nabla f(\vec{x})|\,|\vec{u}| \cos\theta$, where θ is the angle between the two vectors, it follows that $D_{\vec{u}} f$ is maximized when the unit vector $\vec{u}$ is given by:

$$\vec{u} = \frac{\nabla f(\vec{x})}{|\nabla f(\vec{x})|}, \quad |\vec{u}| = 1.$$

Ex: $n=2$, $m=1$

$f: \mathbb{R}^2 \to \mathbb{R}$ differentiable

From the definition of differentiability in (1) we obtained that f is approximated by (see (**)):

$$f(\vec{x}_0+\vec{h}) \cong f(\vec{x}_0) + f'(\vec{x}_0)\vec{h}, \qquad \vec{x} = \vec{x}_0 + \vec{h}$$

In this case, with $n=2$, let us write:

$$\vec{x}_0 = (x_0, y_0), \quad \vec{x} = (x, y)$$

Hence:

$$f'(\vec{x}_0) = \left(\frac{\partial f}{\partial x}(x_0,y_0) \quad \frac{\partial f}{\partial y}(x_0,y_0) \right)$$

Since $\vec{h} = \vec{x} - \vec{x}_0 = (x,y) - (x_0,y_0)$ we have:

$$f'(\vec{x}_0)\vec{h} = \left(\frac{\partial f}{\partial x}(x_0,y_0) \quad \frac{\partial f}{\partial y}(x_0,y_0) \right) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

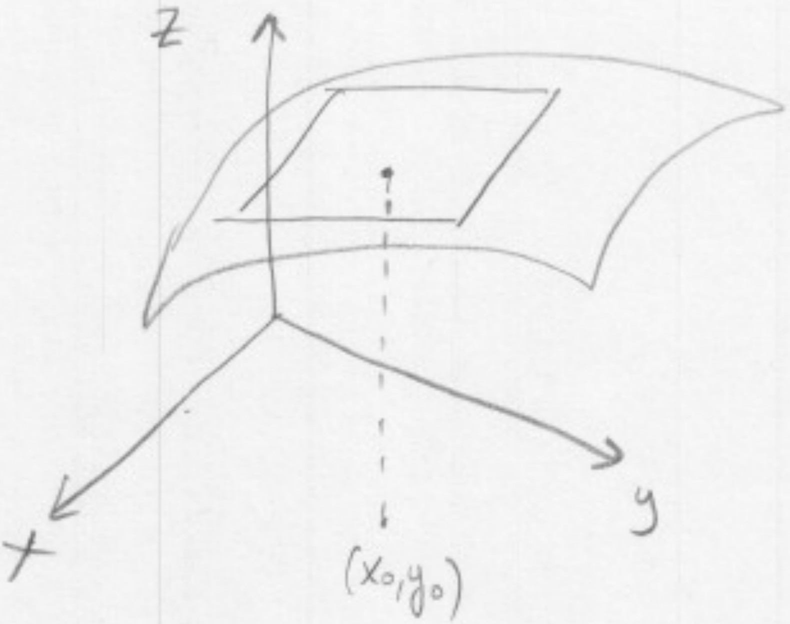
$$= \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$

Hence:

$$f(x,y) \cong f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$

which is the equation of the tangent plane.

<u>Ex</u>: If $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable at (x_0, y_0) then there exists a unique tangent plane that approximates f in a neighborhood of (x_0, y_0).



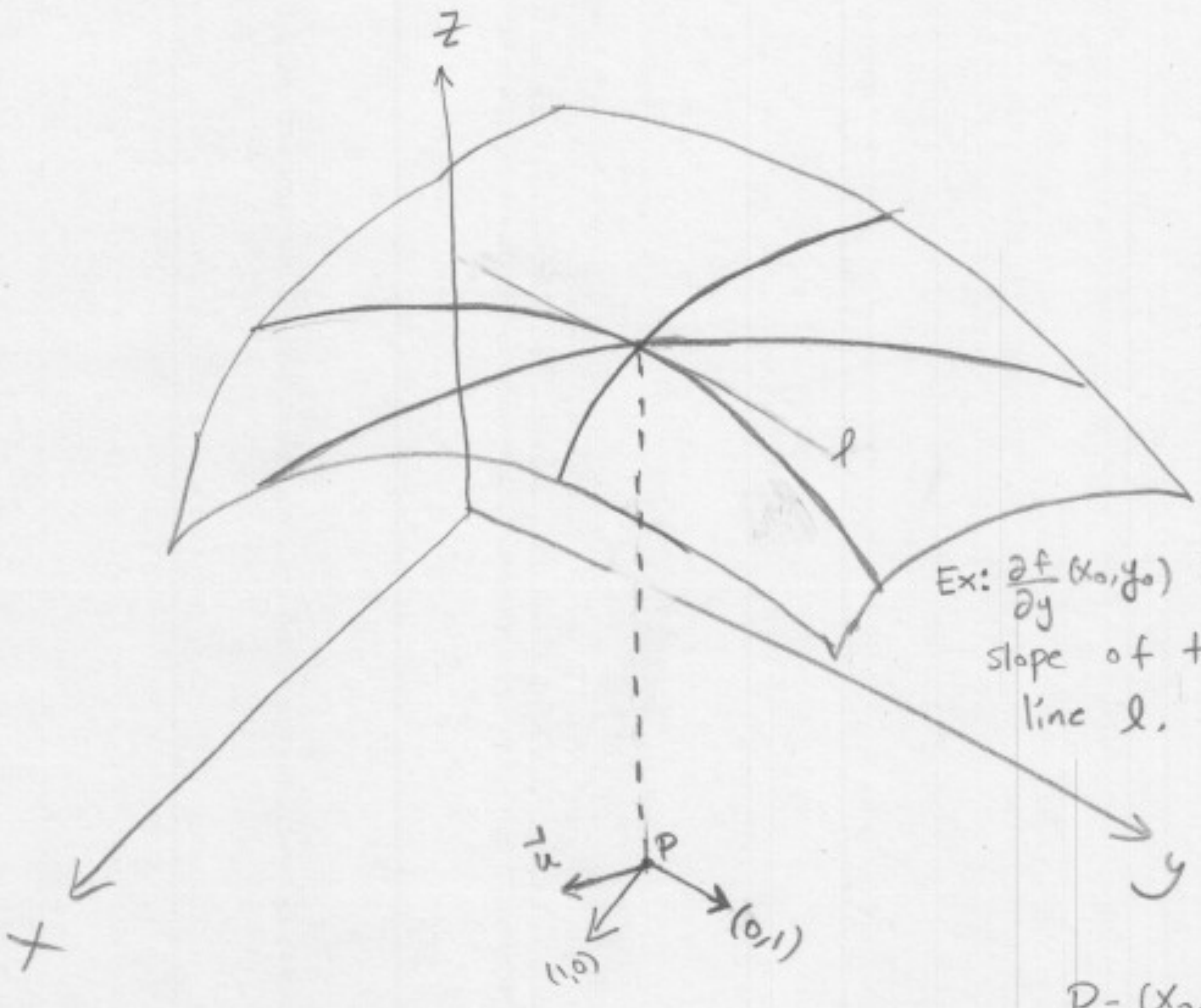
However, even if <u>all</u> the directional derivatives $D_{\vec{u}}(x_0,y_0)$ exist (recall that $\vec{u}=(1,0)$ and $\vec{u}=(0,1)$ correspond to $\frac{\partial f}{\partial x}(x_0,y_0)$ and $\frac{\partial f}{\partial y}(x_0,y_0)$ respectively), the function might not be differentiable at (x_0,y_0) and hence there is no tangent plane that approximates f.

<u>Ex</u>: Recall the geometrical meaning of $D_{\vec{u}} f(x_0, y_0)$, $\vec{u} = (u_1, u_2)$ a unit vector.

$$D_{\vec{u}} f(x_0,y_0) = \lim_{t \to 0} \frac{f(x_0 + tu_1, y_0 + tu_2) - f(x_0,y_0)}{t}$$

$\vec{u} = (1,0)$ $\quad D_{\vec{u}} f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0) = \lim_{t \to 0} \frac{f(x_0+t, y_0) - f(x_0,y_0)}{t}$

$\vec{u} = (0,1)$ $\quad D_{\vec{u}} f(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = \lim_{t \to 0} \frac{f(x_0, y_0+t) - f(x_0,y_0)}{t}$



Ex: $\frac{\partial f}{\partial y}(x_0,y_0)$ is the slope of tangent line l.

$P = (x_0, y_0)$