August 4, 2018

CHAPTER 3: PARTIAL DERIVATIVES AND DIFFERENTIATION

1. Partial Derivatives and Differentiable functions

In all this chapter, D will denote an open subset of $\mathbb{R}^n$.

Definition 1.1. Consider a function $f : D \to \mathbb{R}$ and let $p \in D$, $i = 1, \cdots, n$. We define the partial derivative of f with respect to the i-th variable at the point p as the following limit (if it exists)

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

where $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$,

defined as follows

$$e_i = (\underbrace{0, \ldots, 0}_{i-1\,\text{terms}}, i, \underbrace{0, \ldots, 0}_{n-i\,\text{terms}})$$

For example, in $\mathbb{R}^2$ the canonical basis is

$$\begin{aligned} e_1 &= (1,0) \\ e_2 &= (0,1) \end{aligned}$$

and in $\mathbb{R}^3$ the canonical basis is

$$\begin{aligned} e_1 &= (1,0,0) \\ e_2 &= (0,1,0) \\ e_3 &= (0,0,1) \end{aligned}$$

Remark 1.2. When $n = 2$, in the above definition we let

$$p = (x, y), \quad f(x, y) : \mathbb{R}^2 \to \mathbb{R}$$

we and use the notation,

$$\frac{\partial f}{\partial x}(x,y) = \lim_{t \to 0} \frac{f(x+t,y) - f(x,y)}{t}$$
$$\frac{\partial f}{\partial y}(x,y) = \lim_{t \to 0} \frac{f(x,y+t) - f(x,y)}{t}$$

Likewise, when $n = 3$, we let

$$p = (x, y, z)$$

and use the notation,

$$\frac{\partial f}{\partial x}(x,y,z) = \lim_{t \to 0} \frac{f(x+t,y,z) - f(x,y,z)}{t}$$
$$\frac{\partial f}{\partial y}(x,y,z) = \lim_{t \to 0} \frac{f(x,y+t,z) - f(x,y,z)}{t}$$
$$\frac{\partial f}{\partial y}(x,y,z) = \lim_{t \to 0} \frac{f(x,y,z+t) - f(x,y,z)}{t}$$

Example 1.3. In Economics, the partial derivatives of a utility function are called 'marginal utilities', the partial derivatives of a production function are called 'marginal products'.

Consider, for example the Cobb-Douglas production function

$$f(K, L) = 5K^{1/3}L^{2/3}$$

where f is the number of units produced, K is the capital and L is labor. That is, the above formula means that if we use K units of capital and L units of labor, then we produce $f(K, L) = 5K^{1/3}L^{2/3}$ units of a good. The constants $A = 5$, $\alpha = 1/3$ and $\beta = 2/3$ are technological parameters.

The 'marginal products' with respect to capital and labor are

$$\begin{aligned} \frac{\partial f}{\partial K} &= \frac{5}{3}K^{-2/3}L^{2/3} \\ \frac{\partial f}{\partial L} &= \frac{10}{3}K^{1/3}L^{-1/3} \end{aligned}$$

The marginal product of labor,

$$\frac{\partial f}{\partial L}(K, L)$$

is interpreted in Economics as an **approximation** to the variation in the production of the good when we are using K units of capital and L units of labor and we switch to use an additional unit $L + 1$ of labor and the same units K of capital as before.

We see that the marginal product of labor and capital is positive. That is, if we use more labor and/or more capital, production increases. On the other hand, the marginal product of labor is decreasing in labor and increasing in capital. We may interpret this as follows.

- Suppose that we keep constant the amount of capital that we are using K. If $L' > L$ then

  $$f(K, L' + 1) - f(K, L') < f(K, L + 1) - f(K, L)$$

  That is, an increase in the production when we use an additional unit of labor is decreasing in the initial labor that is being used. If we keep the capital constant, using on additional unit of labor, if we are already using a lot of labor, does not increase much the production.

  We may imagine that $f(K, L)$ is the production of a farm product in a piece of land where L is the number of the workers and the size K of the land is constant. The impact in the production when hiring and additional person is greater if few people are working in the land as compared with the case in which we already have a lot of people working in the land.
- Suppose that the amount of labor L is kept constant. If $K' > K$ then

  $$f(K', L + 1) - f(K', L) > f(K, L + 1) - f(K, L)$$

  That is, the increase in the production when we use one additional unit of labor is larger the more capital we use. Capital and labor are complementary. In the previous example, hiring and additional worker has a larger effect on the production the larger is the size of land.

Definition 1.4. Consider a function $f : D \to \mathbb{R}$. Let $p \in D$ and suppose all the partial derivatives

$$\frac{\partial f}{\partial x_1}(p),\ \frac{\partial f}{\partial x_2}(p), \cdots, \frac{\partial f}{\partial x_n}(p)$$

exist at the point p. We define the **gradient** of f at p as the following vector

$$\nabla f(p) = \left(\frac{\partial f}{\partial x_1}(p),\ \frac{\partial f}{\partial x_2}(p), \cdots, \frac{\partial f}{\partial x_n}(p) \right)$$

Definition 1.5. Consider a function $f : D \to \mathbb{R}$. Let $p \in D$ and suppose all the partial derivatives

$$\frac{\partial f}{\partial x_1}(p),\ \frac{\partial f}{\partial x_2}(p), \cdots, \frac{\partial f}{\partial x_n}(p)$$

exist at the point p. We say that f is differentiable at p if

$$\lim_{v \to 0} \frac{f(p+v) - f(p) - \nabla f(p) \cdot v}{\|v\|} = 0$$

Note that the limit is taken for $v \in \mathbb{R}^n$.

Remark 1.6. A function of **two variables** $f : D \subset \mathbb{R}^2 \to$ is differentiable at the point $p = (a, b)$ if

$$\lim_{(v_1,v_2) \to (0,0)} \frac{f(a+v_1, b+v_2) - f(a,b) - \nabla f(a,b) \cdot (v_1, v_2)}{\|(v_1, v_2)\|} = 0$$

Letting

$$x = a + v_1, \quad y = b + v_2$$

we see that $(v_1, v_2) \to (0, 0)$ is equivalent to $(x, y) \to (a, b)$, so we may write this limit as

$$\lim_{(x,y) \to (a,b)} \frac{f(x,y) - f(a,b) - \nabla f(a,b) \cdot (x-a, y-b)}{\|(x-a, y-b)\|} = 0$$

Writing this limit explicitly we wee that f is differentiable at the point $p = (a, b)$ if

$$\lim_{(x,y) \to (a,b)} \frac{f(x,y) - f(a,b) - \frac{\partial f}{\partial x}(a,b) \cdot (x-a) - \frac{\partial f}{\partial y}(a,b) \cdot (y-b)}{\sqrt{(x-a)^2 + (y-b)^2}} = 0 \tag{1.1}$$

Example 1.7. Consider the function

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

We will show that f is not differentiable at the point $p = (0, 0)$. First of all, we compute $\nabla f(0, 0)$. Note that

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \to 0} \frac{0}{t^3} = 0$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{t \to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t \to 0} \frac{0}{t^3} = 0$$

so, $\nabla f(0,0) = (0,0)$. Let us use the notation $v = (x, y)$. Then, f is differentiable at the point $p = (0,0)$ if and only if

$$\begin{aligned}
0 &= \lim_{v\to 0} \frac{f(p+v) - f(p) - \nabla f(p)\cdot v}{\|v\|} \\
&= \lim_{(x,y)\to(0,0)} \frac{f((0,0)+(x,y)) - f(0,0) - \nabla f(p)\cdot(x,y)}{\sqrt{x^2+y^2}} \\
&= \lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - (0,0)\cdot(x,y)}{\sqrt{x^2+y^2}} \\
&= \lim_{(x,y)\to(0,0)} \frac{f(x,y)}{\sqrt{x^2+y^2}} \\
&= \lim_{(x,y)\to(0,0)} \frac{xy^2}{(x^2+y^2)^{3/2}}
\end{aligned}$$

We prove that the above limit does not exist. Consider the function

$$g(x,y) = \frac{xy^2}{(x^2+y^2)^{3/2}}$$

Note that

$$\lim_{t\to 0} g(t,0) = \lim_{t\to 0} \frac{0}{(2t^2)^{3/2}} = 0$$

and note that

$$\lim_{t\to 0} g(t,t) = \lim_{t\to 0} \frac{t^3}{(2t^2)^{3/2}} = \frac{1}{(2)^{3/2}} \neq 0$$

so the limit

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{(x^2+y^2)^{3/2}}$$

does not exist and we conclude that f is not differentiable at the point $(0,0)$.

Example 1.8. Consider now the function

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

We will show that f is differentiable at the point $p = (0,0)$. First of all, we compute $\nabla f(0,0)$. Note that

$$\begin{aligned}
\frac{\partial f}{\partial x}(0,0) &= \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t\to 0} \frac{0}{t^3} = 0 \\
\frac{\partial f}{\partial y}(0,0) &= \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t\to 0} \frac{0}{t^3} = 0
\end{aligned}$$

so, $\nabla f(0,0) = (0,0)$. Let us use the notation $v = (x,y)$. Then, f is differentiable at the point $p = (0,0)$ if and only if

$$\begin{aligned} 0 &= \lim_{v\to 0} \frac{f(p+v) - f(p) - \nabla f(p)\cdot v}{\|v\|} \\ &= \lim_{(x,y)\to(0,0)} \frac{f((0,0)+(x,y)) - f(0,0) - \nabla f(p)\cdot(x,y)}{\sqrt{x^2+y^2}} \\ &= \lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - (0,0)\cdot(x,y)}{\sqrt{x^2+y^2}} \\ &= \lim_{(x,y)\to(0,0)} \frac{f(x,y)}{\sqrt{x^2+y^2}} \\ &= \lim_{(x,y)\to(0,0)} \frac{xy^3}{(x^2+y^2)^{3/2}} \end{aligned}$$

Let $\varepsilon > 0$. Take $\delta = \varepsilon$ and suppose that $0 < \sqrt{x^2+y^2} < \delta$. Then,

$$\begin{aligned} \left| \frac{xy^3}{(x^2+y^2)^{3/2}} \right| &= \frac{|x|\, y^2\, |y|}{(x^2+y^2)^{3/2}} \\ &= \frac{\sqrt{x^2}\, y^2\, |y|}{(x^2+y^2)^{3/2}} \\ &\le \frac{\sqrt{x^2+y^2}\, \left(x^2+y^2\right)\, |y|}{(x^2+y^2)^{3/2}} \\ &= \frac{\left(x^2+y^2\right)^{3/2}\, |y|}{(x^2+y^2)^{3/2}} \\ &= |y| \le \sqrt{x^2+y^2} < \delta = \varepsilon \end{aligned}$$

So,

$$\lim_{(x,y)\to(0,0)} \frac{xy^3}{(x^2+y^2)^{3/2}} = 0$$

and the function is differentiable at the point $(0,0)$.

Proposition 1.9. Let $f : D \to \mathbb{R}$. If f is differentiable at some point $p \in D$, then f is continuous at that point.

Example 1.10. Consider the function

$$f(x,y) = \begin{cases} \frac{x^3y}{x^4+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Is it continuous and/or differentiable at $(0,0)$? One computes easily the iterated limit

$$\lim_{x\to 0}\left(\lim_{y\to 0} f(x,y)\right) = 0$$

On the other hand, taking the curve

$$x(t) = t, \quad y(t) = t^2$$

we see that

$$\lim_{t\to 0} f(t,t^2) = \lim_{t\to 0} \frac{t^4}{2t^4} = \frac{1}{2} \neq 0$$

So,

$$\lim_{(x,y)\to(0,0)} f(x,y)$$

does not exist. It follows that f is not continuous at $(0,0)$. In addition, by Proposition 1.9, f is not differentiable at $(0,0)$, either.

Theorem 1.11. Let $f : D \to \mathbb{R}$ and $p \in D$. Suppose that there is some $r > 0$ such that the partial derivatives,

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}$$

exist at every point of the open ball $B(p,r)$ and are continuous functions on that ball. Then, the function f is differentiable at p.

Example 1.12. The previous Theorem applies to show that the function

$$f(x,y,z) = xe^{yz} + y\sin z$$

is differentiable at every point of $\mathbb{R}^2$.

Definition 1.13. A function $f : D \to \mathbb{R}$ is of class C^1 in D if all the partial derivatives of f exist and are continuous functions on D. In this case we write $f \in C^1(D)$.

2. Directional derivatives

Definition 2.1. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$. Fix a point $p \in D$ and a vector $v \in \mathbb{R}^n$. If the following limit exists

$$D_v f(p) = \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t}$$

it is called the <u>derivative</u> of f at p <u>along (the vector)</u> v.

If $\|v\| = 1$, then $D_v f(p)$ is called the <u>directional derivative</u> of f at p <u>in the direction of (the vector)</u> v

Remark 2.2. Let $n = 1$. Let $f : \mathbb{R} \to \mathbb{R}$, $p \in \mathbb{R}$ and $v = 1$. The above definition coincides with the derivative of a one variable function

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

Example 2.3. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = xy$ and take $p = (1,-1)$, $v = (3,4)$. Then, for $t \in \mathbb{R}$ we have that

$$p + tv = (1+3t, -1+4t)$$

so,

$$\begin{aligned} D_v f(p) &= \lim_{t\to 0} \frac{f(1+3t,-1+4t) - f(1,-1)}{t} \\ &= \lim_{t\to 0} \frac{(1+3t)(-1+4t)+1}{t} = 1 \end{aligned}$$

And, since $\|v\| = \sqrt{3^2 + 4^2} = 5$, the directional derivative of f at p in the direction of v is

$$\frac{1}{\|v\|} D_v f(p) = \frac{1}{5}$$

Remark 2.4. If we take

$$v = e_i = (0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0)$$

to be the i^{th} vector of the canonical basis, then

$$D_{e_i} f(p) = \frac{\partial f}{\partial x_i}(p)$$

is the i^{th} partial derivative of f at the point p.

Proposition 2.5. Let $f : D \to \mathbb{R}$ be differentiable at the point $p \in D$. Then,

$$D_v f(p) = \nabla f(p) \cdot v \tag{2.1}$$

Example 2.6. As in Example 2.3, let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = xy$ and take $p = (1, -1)$, $v = (3, 4)$. Then,

$$\nabla f(p) = (y, x)\Big|_{\substack{x=1\\ y=-1}} = (-1, 1)$$

and, since f is differentiable on all of $\mathbb{R}^2$, we have that

$$D_v f(p) = \nabla f(p) \cdot v = (-1, 1) \cdot (3, 4) = -3 + 4 = 1$$

as computed in Example 2.3.

Remark 2.7. One may also define the derivative of a function $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ **along (the vector)** v. To do so, we write the function f using its coordinate functions

$$f(x) = (f_1(x), f_2(x), \cdots, f_m(x))$$

with $f_i : D \to \mathbb{R}$ for each $i = 1, \cdots, m$. And define

$$D_v f(p) = (D_v f_1(p), D_v f_2(p), \cdots, D_v f_m(p))$$

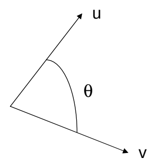
We see now that $D_v f(p)$ is a vector in $\mathbb{R}^m$. Likewise we may define the **directional derivative** of f at the point p **in the direction of a unitary vector** u.

3. Interpretation of the Gradient

The formula 2.1 may be used to give an interpretation of the gradient as follows. Recall that given two vectors u, v in $\mathbb{R}^n$, their scalar product satisfies

$$u \cdot v = \|u\| \, \|v\| \cos\theta$$

where θ is the angle between the two vectors.



Applying this observation to formula 2.1, we see that

$$D_v f(p) = \nabla f(p) \cdot v = \|\nabla f(p)\| \, \|v\| \cos\theta$$

where θ is the angle between the vectors $\nabla f(p)$ and v. Taking v to be unitary, we see that the derivative of f in the direction of v is

$$D_v f(p) = \|\nabla f(p)\| \cos\theta$$

Thus, $D_v f(p)$

- attains a maximum when $\theta = 0$, that is, when the vectors $\nabla f(p)$ and v point in the same direction.
- attains a minimum when $\theta = \pi$, that is, when the vectors $\nabla f(p)$ and v point in the opposite directions.
- is zero when $\theta = \pi/2$ or $\theta = 3\pi/2$, that is, when the vectors $\nabla f(p)$ and v are perpendicular.

It follows that,

- The function f grows the fastest in the direction of $\nabla f(p)$.
- The function f decreases the fastest in the direction opposite to $\nabla f(p)$.
- The function f remains constant in the directions perpendicular to $\nabla f(p)$.

4. The chain rule

Definition 4.1. Given a function $f(x) = (f_1(x), f_2(x), \cdots, f_m(x)) : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and a point $p \in D$, we define the Jacobian matrix of f at the point p as the following matrix of order $m \times n$

$$\mathrm{D}\, f(p) = \begin{pmatrix} \frac{\partial f_1(p)}{\partial x_1} & \frac{\partial f_1(p)}{\partial x_2} & \cdots & \frac{\partial f_1(p)}{\partial x_n} \\ \frac{\partial f_2(p)}{\partial x_1} & \frac{\partial f_2(p)}{\partial x_2} & \cdots & \frac{\partial f_2(p)}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m(p)}{\partial x_1} & \frac{\partial f_m(p)}{\partial x_2} & \cdots & \frac{\partial f_m(p)}{\partial x_n} \end{pmatrix}$$

Remark 4.2. If $f(x) = D \subset \mathbb{R}^n \to \mathbb{R}$ What is the difference between $\mathrm{D}\, f(p)$ and $\nabla\, f(p)$?

Remark 4.3. If $m = n = 1$ What is $\mathrm{D}\, f(p)$?

Definition 4.4. A function $f(x) = (f_1(x), f_2(x), \cdots, f_m(x)) : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be differentiable at a point $p \in D$ if each of the functions $f_1(x), f_2(x), \cdots, f_m(x)$ is differentiable at p.

Theorem 4.5 (The chain rule)**.** *Let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^l$. Suppose that g is differentiable at $p \in \mathbb{R}^n$ and that f is differentiable at $g(p) \in \mathbb{R}^m$. Then, the function $f \circ g$ is differentiable at p and*

$$\mathrm{D}(f \circ g)(p) = \mathrm{D}\, f(g(p))\, \mathrm{D}\, g(p)$$

Remark 4.6. The expression $\mathrm{D}(f \circ g)(p) = \mathrm{D}\, f(g(p))\, \mathrm{D}\, g(p)$ contains the product of 2 matrices.

Example 4.7 (Special case of the chain rule)**.** Let $\sigma : \mathbb{R} \to \mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable. Suppose that $\sigma(t)$ may be written as

$$\sigma(t) = (x(t), y(t))$$

Then, the chain rule says that

$$\begin{aligned} \frac{d}{dt} f(x(t), y(t)) &= \mathrm{D}(f \circ \sigma)(t) = \mathrm{D}\, f(x,y)\big|_{\substack{x=x(t)\\ y=y(t)}} \mathrm{D}\, \sigma(t) \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}\Big|_{\substack{x=x(t)\\ y=y(t)}} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \\ &= \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t) \end{aligned}$$

Example 4.8 (Special case of the chain rule)**.** Let $g(s,t) : \mathbb{R}^2 \to \mathbb{R}^2$ and $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ be differentiable. Suppose that $g(s,t)$ may be written as

$$g(s,t) = (x(s,t), y(s,t))$$

so that

$$(f \circ g)(s,t) = f(g(s,t)) = f\,(x(s,t), y(s,t))$$

hen, the chain rule says that

$$\begin{aligned} \mathrm{D}\, f\,(x(s,t), y(s,t)) &= \mathrm{D}(f \circ g)(s,t) = \mathrm{D}\, f(x,y)\big|_{\substack{x=x(s,t)\\ y=y(s,t)}} \mathrm{D}\, g(s,t) \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} & \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \end{pmatrix} \end{aligned}$$

That is,

$$\begin{aligned} \frac{\partial(f \circ g)}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial(f \circ g)}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \end{aligned}$$

Example 4.9. Consider the Cobb-Douglas production function

$$f(K,L) = 5K^{1/3}L^{2/3}$$

where f are the units produced, K is capital and L is labor. Suppose that capital and labor change with time

$$K = K(t), \quad L = L(t)$$

Then the production function

$$f(K(t), L(t))$$

is also a function of time. What is the rate of change of the production at a given time? We may answer this question using the chain rule.

$$\begin{aligned}\frac{df(K(t),L(t))}{dt} &= \frac{\partial f}{\partial K}\frac{dK}{dt}+\frac{\partial f}{\partial L}\frac{dL}{dt}\\ &= \frac{5}{3}K^{-2/3}L^{2/3}\frac{dK}{dt}+\frac{10}{3}K^{1/3}L^{-1/3}\frac{dL}{dt}\end{aligned}$$

Example 4.10. Suppose an agent has the following differentiable utility function

$$u(x,y)$$

where x is a consumption good and y is air pollution. Then, the utility of the agent is increasing in x and decreasing in y,

$$\begin{aligned}\frac{\partial u}{\partial x} &> 0\\ \frac{\partial u}{\partial y} &< 0\end{aligned}$$

Suppose that the production of x units of the good generates $y = f(x)$ units of pollution, What is the optimal level of consumption of x?

The utility of the agent when he consumes x units of the good and $y = f(x)$ units of pollution are generated is

$$u(x, f(x))$$

The agents maximizes this utility function. The first order condition is

$$\frac{du(x,f(x))}{dx} = 0$$

using the chain rule we obtain that the equation

$$0 = \frac{\partial u}{\partial x}(x,f(x)) + \frac{\partial u}{\partial y}(x,f(x))f'(x)$$

determines the optimal level of production of the good.

5. Derivative along a curve and level surfaces

Remark 5.1 (A special case of the chain rule)**.** Let $\sigma : \mathbb{R} \to \mathbb{R}^n$ be a differentiable curve and let $f : D \to \mathbb{R}$ be differentiable, where D is an open subset of $\mathbb{R}^n$. Suppose, $\sigma(t)$ can be written as

$$\sigma(t) = (\sigma_1(t), \sigma_2(t), \ldots, \sigma_n(t))$$

where each

$$\frac{d\sigma_i}{dt}(t)$$

is differentiable for every $i = 1, \ldots, n$ and every $t \in \mathbb{R}$. Hence, we may write

$$\frac{d\sigma}{dt} = \left(\frac{d\sigma_1}{dt}, \frac{d\sigma_2}{dt}, \ldots, \frac{d\sigma_n}{dt}\right)$$

Then, $f(\sigma(t))$ is differentiable and

$$\frac{d}{dt}f(\sigma(t)) = \nabla f(\sigma(t)) \cdot \frac{d\sigma}{dt}$$

Example 5.2. If $f : \mathbb{R}^2 \to \mathbb{R}$ is a function of two variables and $x = x(t)$, $y = y(t)$, the chain rule is

$$\frac{d}{dt} f(x(t), y(t)) = \left.\frac{\partial f(x,y)}{\partial x}\right|_{\substack{x=x(t)\\ y=y(t)}} \frac{dx(t)}{dt} + \left.\frac{\partial f(x,y)}{\partial y}\right|_{\substack{x=x(t)\\ y=y(t)}} \frac{dy(t)}{dt}$$

Example 5.3. If $f : \mathbb{R}^3 \to \mathbb{R}$ is a function of three variables and $x = x(t)$, $y = y(t)$, $z = z(t)$, the chain rule is

$$\frac{d}{dt} f(x(t), y(t), z(t)) = \left.\frac{\partial f(x,y,z)}{\partial x}\right|_{\substack{x=x(t)\\ y=y(t)\\ z=z(t)}} \frac{dx(t)}{dt} + \left.\frac{\partial f(x,y,z)}{\partial y}\right|_{\substack{x=x(t)\\ y=y(t)\\ z=z(t)}} \frac{dy(t)}{dt} + \left.\frac{\partial f(x,y,z)}{\partial z}\right|_{\substack{x=x(t)\\ y=y(t)\\ z=z(t)}} \frac{dz(t)}{dt}$$

Remark 5.1 provides another interpretation of the gradient. Let $p \in D$ and let $f : D \to \mathbb{R}$ be differentiable, where D is some open subset of $\mathbb{R}^n$. Let $C \in \mathbb{R}$ and suppose the level surface

$$S_C = \{x \in D : f(x) = C\}$$

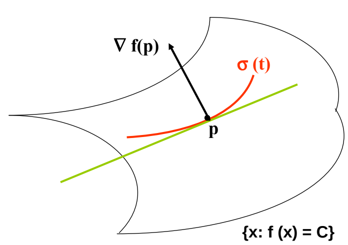
is not empty. Let $\sigma : \mathbb{R} \to \mathbb{R}^n$ be a differentiable curve and suppose that $\sigma(t) \in S_C$ for all $t \in \mathbb{R}$. That is

$$f(\sigma(t)) = c$$

for every $t \in \mathbb{R}$. Differentiating and using the above chain rule we have that

$$0 = \frac{d}{dt} f(\sigma(t)) = \nabla f(\sigma(t)) \cdot \frac{d\sigma}{dt}$$

That is $\nabla f(\sigma(t))$ and $d\sigma(t)/dt$ are perpendicular for every $t \in \mathbb{R}$.



The above argument shows that at any point $p \in S_C$, the gradient $\nabla f(p)$ is **perpendicular** to the surface level S_C.

Remark 5.4. Let us compute the plane tangent to the graph of a function of two variables. To do this, consider a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$. The graph of f is the set

$$G = \{(x, y, f(x, y)) : (x, y) \in \mathbb{R}^2\}$$

Fix a point $p = (a, b) \in \mathbb{R}^2$. Consider the following function of three variables

$$g(x, y, z) = f(x, y) - z$$

Then,the graph of f may be written as

$$G = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$$

Then, the plane T tangent to G at the point $(a, b, f(a, b))$ satisfies the two following properties

- T contains the point $(a, b, f(a, b))$.
- T is perpendicular to the gradient $\nabla g(a, b, f(a, b))$.

This information permits us to compute the equations for T as follows. First of all, it follows that an equation for T is

$$\nabla g(a, b, f(a, b)) \cdot ((x, y, z) - (a, b, f(a, b))) = 0$$

and note that

$$\nabla g(a, b, f(a, b)) = \left(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b), -1\right)$$

Hence, an equation for T is the following

$$f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b) = z \tag{5.1}$$

We may use the above to provide another interpretation of the definition of differentiability 1.5. Let

$$P_1(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b)$$

For the case of a function of two variables, Equation 1.1 says that the function f is differentiable at the point (a, b) if

$$\lim_{(x,y)\to(a,b)} \frac{|f(x, y) - P_1(x, y)|}{\|(x - a, y - b)\|} = 0$$

In view of Equation 5.1, the function f is differentiable at the point (a, b) if the tangent plane is a good approximation to the value of the function

$$f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b)$$

6. Taylor's first order approximation

Assuming that the tangent plane is a good approximation to the function, we may use the value of the variable z in equation 5.1 as an approximation to the value of $f(x, y)$ that is

$$f(x, y) \approx \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b) + f(a, b)$$

This motivates the following,

Definition 6.1. Taylor's first order polynomial for the function f around the point $p = (a, b)$ is the polynomial

$$P_1(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b)$$

The following Proposition makes precise the statement that the first order Taylor polynomial is a good approximation for the function.

Proposition 6.2. Suppose f is of class C^1 in $\mathbb{R}^2$. Let $P_1(x, y)$ be Taylor's first order polynomial for the function f around the point $p = (a, b)$ Then,

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - P_1(x, y)}{\|(x - a, y - b)\|} = 0$$

Remark 6.3. Comparing the equation in the above proposition with equation 1.1 we see that a function f of two variables is differentiable at the point p if and only if the partial derivatives of f exist at that point and the first order Taylor polynomial is a 'good' approximation to the value of f near the point p.