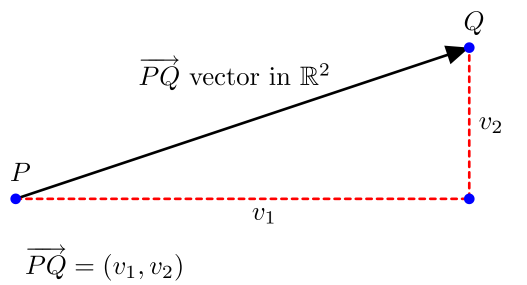
Updated: February 26, 2013

0 Day 0: Prerequisites

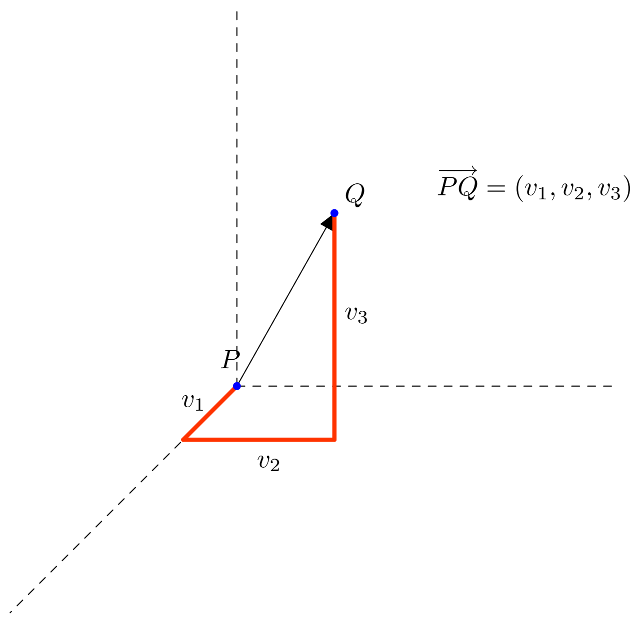
These are a few facts about vectors you should be aware of. Revise them carefully and complete information by checking any high school algebra book or college manual.

0.1 Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$

Consider two points (either in $\mathbb{R}^2$ or $\mathbb{R}^3$): P and Q. This pair of points is called a **vector**, $\overrightarrow{PQ}$. Point P is called the **origin** and point Q the **endpoint**. The easy way to think of vectors is considering them arrows starting at the origin and ending at the endpoint. For example, if $\vec{v} = \overrightarrow{PQ}$ is a vector[1] in $\mathbb{R}^2$:



In $\mathbb{R}^3$ we would have three components (v_1, v_2, v_3) describing the way of going from P to Q following each of the coordinate axes:



[1]Vectors are usually represented as $\vec{v}$ (blackboard notation) or using boldface $\mathbf{v}$. We will use indistinctly both notations.

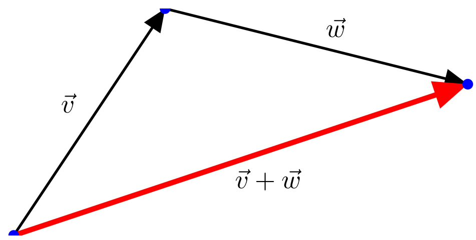
The components of a vector can be found as the difference between the coordinates of the endpoint and the origin. Notice that the components of a a vector depend only on the relative position of its origin and endpoint. In this way, a vector can be "moved" as long as the "arrow" does not change its length, its direction (moves to a parallel line) or sense (does not point at the contrary direction).

Vectors can be manipulated using their components (analytically) or using their graphical representation as arrows with an origin and an endpoint (geometrically). So, vectors can be added:

- Analytically:

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2) \qquad \text{or} \qquad (v_1, v_2, v_3) + (w_1, w_2, w_3) = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$$

- or geometrically (both in $\mathbb{R}^2$ and $\mathbb{R}^3$):

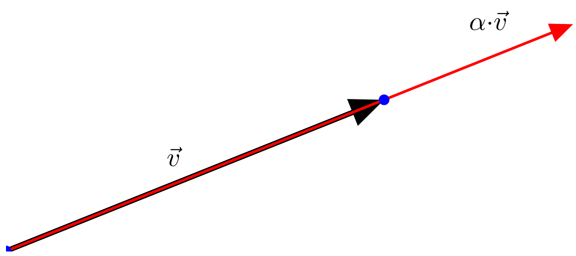


Vectors can also be multiplied by numbers:

- Analytically:

$$\alpha{\cdot}(v_1, v_2) = (\alpha{\cdot}v_1, \alpha{\cdot}v_2) \qquad \text{or} \qquad \alpha{\cdot}(v_1, v_2, v_3) = (\alpha{\cdot}v_1, \alpha{\cdot}v_2, \alpha{\cdot}v_3)$$

- or geometrically (both in $\mathbb{R}^2$ and $\mathbb{R}^3$):



If the origin coincides with the endpoint, the vector is the vector $\vec{0} = (0, 0)$ or $\vec{0} = (0, 0, 0)$.

0.2 Linear independence

Two vectors are called **linearly dependent** if one can be obtained as the other multiplied by a number. This means that both vectors have the same direction.

If two vectors are not dependent, they are **independent**. Their directions are different.

Linearly Dependent (LD) Linearly Independent (LI)

Using components, two LD vectors have proportional components:

In $\mathbb{R}^2$ $\quad \dfrac{v_1}{w_1} = \dfrac{v_2}{w_2} \quad \Leftrightarrow \quad \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} = 0;$

In $\mathbb{R}^3$ $\quad \dfrac{v_1}{w_1} = \dfrac{v_2}{w_2} = \dfrac{v_3}{w_3} \quad \Leftrightarrow \quad$ rank of matrix $\begin{pmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$ is not 2.

Two LI vectors have non-proportional components

In $\mathbb{R}^2$ $\quad \dfrac{v_1}{w_1} \neq \dfrac{v_2}{w_2} \quad \Leftrightarrow \quad \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \neq 0;$

In $\mathbb{R}^3$ $\quad \left\{ \dfrac{v_1}{w_1} = \dfrac{v_2}{w_2} = \dfrac{v_3}{w_3} \right\}$ is **not** true $\quad \Leftrightarrow \quad$ rank of matrix $\begin{pmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$ is 2.

0.3 Linear independence for more than two vectors

A vector $\vec{u}$ is a **linear combination** (LC) of two other vectors $\vec{v}, \vec{w}$ when we can find real numbers s and t such that

$$\vec{u} = s{\cdot}\vec{v} + t{\cdot}\vec{w}.$$

Now, three vectors are LD when one of them is a LC of the other two. Consequently, three vectors are LI when they are not LD, that is, when none is a LC of the other two.

In $\mathbb{R}^2$, it is impossible to have three LI vectors. In $\mathbb{R}^3$, three (non-zero and different) vectors are LI if and only if they are not on one plane. This means also that if we have three vectors on the same plane, they are LD.

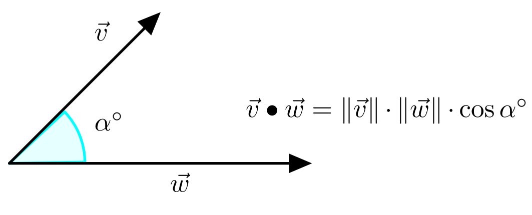
In general, the independence of a set of vectors can be determined by using matrices. We write the vectors as lines (or columns) of a matrix and then the rank of this matrix is the maximum number of independent vectors that you can find in the set. If you use Gauss' method to determine the rank of the matrix, the non-zero files left at the end of the process correspond to LI vectors.

0.4 Scalar product

The "length" of a vector is called its **module** (or **norm**). It can be found simply by using Pythagoras Theorem:

$$\|(v_1, v_2)\| = \sqrt{v_1^2 + v_2^2}\ ; \qquad \text{or} \qquad \|(v_1, v_2, v_3)\| = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

There is an important external operation called the **scalar product** or **dot product** of two vectors:



As it is geometrically obvious, if $\vec{v}$ and $\vec{w}$ are perpendicular (orthogonal), $\alpha° = 90°$ and as $\cos 90° = 0$, $\vec{v} \bullet \vec{w} = 0$. This is true in $\mathbb{R}^2$ and $\mathbb{R}^3$:

$$\vec{v} \perp \vec{w} \quad \Leftrightarrow \quad \vec{v} \bullet \vec{w} = 0.$$

This can be translated into components. It is not difficult to see (we need some trigonometry for that) that

$$\vec{v} \bullet \vec{w} = v_1{\cdot}w_1 + v_2{\cdot}w_2 + v_3{\cdot}w_3.$$

Consequently

$$\vec{v} \perp \vec{w} \quad \Leftrightarrow \quad \vec{v} \bullet \vec{w} = 0 \quad \Leftrightarrow \quad v_1{\cdot}w_1 + v_2{\cdot}w_2 + v_3{\cdot}w_3 = 0.$$

These notions will be very important in the following.

1 Day 1: Lines and planes in $\mathbb{R}^3$

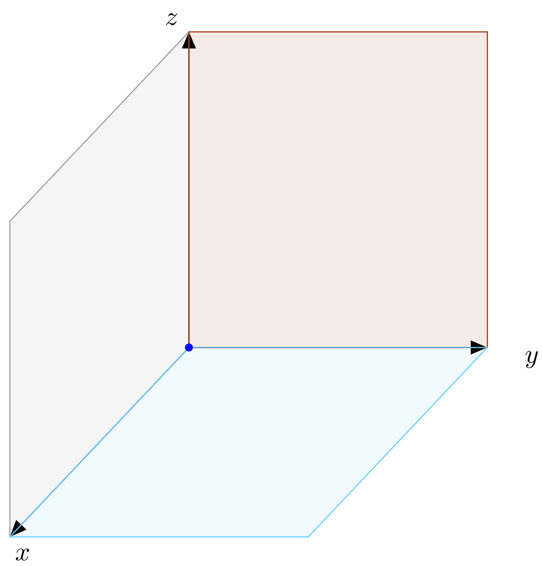
1.1 Planes

Close your eyes and imagine the classroom we are in bare of furniture and the floor completely level. The left-down corner is the origin of coordinates $(0, 0, 0)$. The x axis runs where the floor meets the left glass wall. The y axis is the meeting of the floor and the blackboard wall, and lastly, the z axis is the vertical corner going up from the origin to the ceiling. We are all in the $(+, +, +)$ area of the space $\mathbb{R}^3$.

Question 1. *What have in common all the points on the blackboard wall?* Answer to question 1.

Question 2. *And those on the floor (remember there are no seats and the floor is completely level)?* Answer to question 2.

Question 3. *And the points on the glass wall?* Answer to question 3.



We have seen three examples of equations of planes.

Now, in the same way that $3x + 2y - 5 = 0$ is the equation of a straight line in $\mathbb{R}^2$, the equation $2x - y + 3z - 4 = 0$ is the *general equation* of a plane in $\mathbb{R}^3$.

Any equation $ax + by + cz + d = 0$ is the equation of a plane in $\mathbb{R}^3$ (a, b, c not all 0).

Later, we will see why this is so.

Question 4. *What happens with the plane if $a = b = 0$?* Answer to question 4.

Question 5. *Find the equation of the plane that contains points $(0, 0, 1)$, $(0, 2, 0)$ and $(-1, 1, 0)$.* Answer to question 5.

1.2 Lines.

Let us go back to our imaginary empty classroom. Where are all the points $(0, 0, \cdot)$? These points satisfy the system of equations $x = 0$; $y = 0$ where the variable z does not appear since it can take **any** value. Close your eyes. Can you picture in your mind points like $(0, 0, 1)$, $(0, 0, 2)$, etc? What do they share?

You must have realized by now that we are talking of a **straight line**: the z-axis. In fact, $x = 0$ is a plane (the blackboard wall), $y = 0$ is another plane (the glass wall) and the z-axis is their intersection. Any straight line can be imagined as the intersection of two planes.

A straight line in $\mathbb{R}^3$ does not have "an equation" but a **system of equations**.[2]

[2] Actually, the system must have one degree of freedom, that is, the rank of the system matrix must be 2. That means that in the box, vectors (a_1, b_1, c_1) and (a_2, b_2, c_2) have to be linearly independent.

A system of equations $\left\{\begin{array}{rcl} a_1x+b_1y+c_1z+d_1 & = & 0 \\ a_2x+b_2y+c_2z+d_2 & = & 0 \end{array}\right.$ represents a line in $\mathbb{R}^3$ ((a_1,b_1,c_1) and (a_2,b_2,c_2) linearly independent).

It happens that we refer to the system as the *general equation* of the line (forcing language). Each one of the equations of the system is the equation of a plane. The line is given as the intersection of two planes.

But there are an infinity of couples of planes that intersect in the same line. Consequently, there are an infinity of "equations" of a line as the intersection of two planes. A little reflection, though, convinces us that any other plane that goes trough the same line must have an equation which is a **linear combination** of the two initial planes: given the line r with the equation $\left\{\begin{array}{rcl} a_1x+b_1y+c_1z+d_1 & = & 0 \\ a_2x+b_2y+c_2z+d_2 & = & 0 \end{array}\right.$, the plane

$$s{\cdot}(a_1x+b_1y+c_1z+d_1)+t{\cdot}(a_2x+b_2y+c_2z+d_2)=0 \qquad s,t\in\mathbb{R} \text{ (not both 0)}.$$

contains line r.

Question 6. *Find a plane that contains line* $\left\{\begin{array}{rcl} 2x-3y+z+1 & = & 0 \\ -x+y-z-3 & = & 0 \end{array}\right.$ *and passes through point* $(0,0,0)$.

Answer to question 6.

Answers to questions:

A1. All points have a first coordinate equal to zero, i.e. $x=0$. **Back**

A2. Now $z=0$. **Back**

A3. $y=0$. **Back**

A4. If $a=b=0$ then $c\neq 0$ and the equation is $cz+d=0$ or, equivalently, $z=-d/c$. The plane is parallel to the floor plane, $z=0$. **Back**

A5. If $ax+by+cz+d=0$ is the equation of the plane we seek, each point must satisfy it:

$$\left\{\begin{array}{rcl} a{\cdot}0+b{\cdot}0+c{\cdot}1+d & = & 0 \\ a{\cdot}0+b{\cdot}2+c{\cdot}0+d & = & 0 \\ a{\cdot}(-1)+b{\cdot}1+c{\cdot}0+d & = & 0 \end{array}\right.$$

This is an dependent system (3 equations, 4 unknowns). Its solution, in terms of d, is

$$\left\{\begin{array}{rcl} c & = & -d \\ b & = & -d/2 \\ a & = & b+d=d/2 \end{array}\right.$$

If we take any value of $d\neq 0$, for instance $d=2$ (which is convenient to avoid fractions) we have as equation of our plane $x-y-2z+2=0$. **Back**

A6. The plane we seek must have as equation a linear combination of the two equations that define the given line

$$s{\cdot}(2x-3y+z+1)+t{\cdot}(-x+y-z-3)=0, \quad \text{that is} \quad (2s-t)x+(-3s+t)y+(s-t)z+s-3t=0$$

If the plane has to pass through point $(0,0,0)$, then $s-3t=0$. That means that $s=3t$. Using a value ($\neq 0$) for t, for example, $t=1$, we have $s=3$ and the equation of the plane is

$$5x-8y+2z=0.$$

(See further down for a different solution to the same problem.) **Back**

2 Day 2: Vector equations of lines and planes.

The textbook, Sydsaeter & Hammond, presents a new way of seeing points, lines and planes in space: the **vector** point of view. This is very useful in many occasions.

Any point (a_1,a_2,a_3) in $\mathbb{R}^3$ can be seen as a **vector** with three components, $\mathbf{a}=(a_1,a_2,a_3)$. This vector is also called a **position vector**. We have a total one-to-one correspondence between points and their position vectors.

Now, if we have two points on a line,

Any point $\mathbf{x} = (x, y, z)$ on the line that goes through two given points in space, $\mathbf{a}$ and $\mathbf{b}$ can be obtained as

$$\mathbf{x} = \mathbf{a} + t\,(\mathbf{b} - \mathbf{a}) \qquad \text{where } t \text{ is a real number.}$$

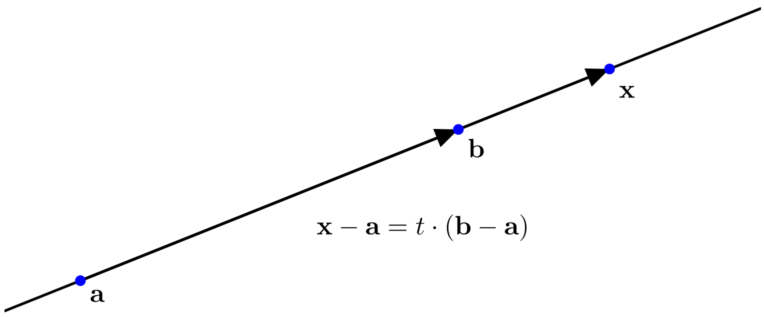
This is so because $\mathbf{x} - \mathbf{a}$ and $\mathbf{b} - \mathbf{a}$ are LD:

For $t = 0$ we obtain $\mathbf{a}$. For $t = 1$, we get $\mathbf{b}$. And so on. The points in between $\mathbf{a}$ and $\mathbf{b}$ are obtained for values of t between 0 and 1. Thus, the midpoint between $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{m} = \mathbf{a} + 0.5(\mathbf{b} - \mathbf{a}) = 0.5\mathbf{b} + 0.5\mathbf{a} = \frac{1}{2}(\mathbf{a} + \mathbf{b}).$$

The vector equation of the line that goes through points $\mathbf{a}$ and $\mathbf{b}$ is:

$$\mathbf{x} = \mathbf{a} + t\ (\mathbf{b} - \mathbf{a})\,, \text{ where } t \in \mathbb{R}. \tag{1}$$



Vector $\mathbf{b} - \mathbf{a}$ is called a **direction vector** for our line. Actually, any vector joining two different points of the line is a direction vector (proportional to each other).

Remember that $\mathbf{x}$ represents **any** point on the line and that to each point there corresponds a unique value of t and viceversa. If we break this vector equation in components, we have

$$(x, y, z) = (a_1, a_2, a_3) + t\,(b_1 - a_1, b_2 - a_2, b_3 - a_3)$$

that is

$$\begin{cases} x &= (1-t)a_1 + tb_1 \\ y &= (1-t)a_2 + tb_2 \\ z &= (1-t)a_3 + tb_3, \qquad t \in \mathbb{R} \end{cases}$$

which is another form of "equation" for our line: the *parametric equation*.

Similarly, a plane has also a *vector equation*:

Any point $\mathbf{x} = (x, y, z)$ of the plane that goes through three given points in space (non-collinear), $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ may be obtained as

$$\mathbf{x} = \mathbf{a} + s\,(\mathbf{b} - \mathbf{a}) + t\,(\mathbf{c} - \mathbf{a}) \qquad \text{where } s \text{ and } t \text{ are real numbers.} \tag{2}$$

The vectors $\mathbf{b} - \mathbf{a}$ and $\mathbf{c} - \mathbf{a}$ are called *direction vectors* of the plane. Actually, any pair of linearly independent vectors contained in the plane are direction vectors of the plane.

Equation (2) is easily obtained as vectors $\mathbf{x} - \mathbf{a}$, $\mathbf{b} - \mathbf{a}$ and $\mathbf{c} - \mathbf{a}$ must be LD if they are all in a plane. As $\mathbf{b} - \mathbf{a}$ and $\mathbf{c} - \mathbf{a}$ are LI, we must have

$$\mathbf{x} - \mathbf{a} = s\,(\mathbf{b} - \mathbf{a}) + t\,(\mathbf{c} - \mathbf{a}) \qquad \text{where } s \text{ and } t \text{ are real numbers.}$$

Question 7. *Rework question 6 at the end of the previous section.* *Answer to question 7.*

2.1 The general equation of a plane.

A plane can be defined

- by **any three** of its points (non collinear) or

- by **any one** of its points, **a**, and **any orthogonal vector** (perpendicular) to the plane (which is called a **normal** vector), **p**.

In the first case, three non-collinear points provide us with a position vector (any of the points) and two direction vectors (any two of the vectors defined by any pair of the given points). The vector equation of the plane, (2), is very easy to write.

In the second case, any other point on the plane $\mathbf{x}$ satisfies the general equation of the plane:

$$\mathbf{p} \cdot (\mathbf{x} - \mathbf{a}) = 0. \tag{3}$$

This translates immediately into components as

$$p_1(x - a_1) + p_2(y - a_2) + p_3(z - a_3) = 0.$$

Expanding this last equation, we end up with an equation like

$$p_1\, x + p_2\, y + p_3\, z + d = 0,$$

where $d = -a_1\, p_1 - a_2\, p_2 - a_3\, p_3$.

This is again the equation we met at the beginning of the lecture. Notice that the coefficients of x, y, and z are exactly the components of a normal vector to our plane.

2.2 Some examples.

Work out a few exercises as examples. Try to solve each one before looking up the solution.

1. *Find the equation of the line through the points* $(2, 3, 4)$ *and* $(-3, 0, 5)$.
   We can get different equations as the solution. The vector equation would be (I choose $\mathbf{a} = (2, 3, 4)$; could be the other point, no problem):

   $$\mathbf{x} = (2, 3, 4) + t{\cdot}(-3 - 2, 0 - 3, 5 - 4), \quad \text{that is} \quad \mathbf{x} = (2 - 5t, 3 - 3t, 4 + t),\ t \in \mathbb{R}$$

   Separating each component we get the equation in a different way (**parametric equation**):

   $$\begin{cases} x &= 2 - 5t \\ y &= 3 - 3t \\ z &= 4 + t, \qquad t \in \mathbb{R} \end{cases}$$

   As t has to be the same value in each of the three equations above, we have

   $$\frac{x - 2}{-5} = \frac{y - 3}{-3} = \frac{z - 4}{1}$$

   which is called the **continuous (or symmetric) equation** of a line. In fact, we can obtain two independent equations from the two equalities above:

   $$\begin{cases} \dfrac{x - 2}{-5} &= \dfrac{y - 3}{-3} \\ \dfrac{y - 3}{-3} &= \dfrac{z - 4}{1} \end{cases} \qquad \text{that is} \qquad \begin{cases} 3x - 5y + 9 &= 0 \\ y + 3z - 15 &= 0 \end{cases}$$

   which is the **general equation** of the line as intersection of two planes.

2. *Find the equation of the plane which is parallel to plane* $3x - 2y + z - 7 = 0$ *and passes through the origin of coordinates.*
   Two parallel planes have the **same normal** vector. Thus the plane we are looking for must have as general equation

   $$3x - 2y + z + d = 0.$$

   The value of d can be determined imposing that $(0, 0, 0)$ belong to the plane:

   $$0 + d = 0 \quad \Rightarrow \quad d = 0.$$

   The solution is $3x - 2y + z = 0$.

3. *Find the equation of the plane that passes through the three points* $(-1,1,1)$, $(2,0,-1)$, *and* $(3,-2,-1)$.
If our plane has $ax+by+cz+d=0$ as general equation, the three points must satisfy it. This provides us with a system of equations:

$$\begin{cases} -a+b+c+d &= 0 \\ 2a-c+d &= 0 \\ 3a-2b-c+d &= 0 \end{cases}$$

The system is dependent. Solving for three of the variables (let us use d as a parameter), we have

$$\begin{cases} a &= -(4/3)d \\ b &= -(2/3)d \\ c &= -(5/3)d. \end{cases}$$

Let us choose a value for d. Let us say $d=-3$ (in this way we will get positive whole numbers for a, b and c). The solution is

$$4x+2y+5z-3=0.$$

Any other value of d would lead to a multiple of the equation of the plane (that is, the same equation).

Another way of solving this problem is to think about the general equation of the plane as given by

$$\mathbf{p}\cdot(\mathbf{x}-\mathbf{a})=0.$$

Point $\mathbf{a}$ can be any of the three points: let us say $\mathbf{a}=(-1,1,1)$. We want to find the components of the normal vector (p_1,p_2,p_3). Again, replacing $\mathbf{x}$ with the coordinates of the other two points, the equation has to be satisfied:

$$\begin{cases} (p_1,p_2,p_3)\cdot(2+1,0-1,-1-1) &= 0 \\ (p_1,p_2,p_3)\cdot(3+1,-2-1,-1-1) &= 0 \end{cases} \Rightarrow \begin{cases} 3p_1-p_2-2p_3 &= 0 \\ 4p_1-3p_2-2p_3 &= 0 \end{cases}$$

Again, this is a dependent system. Solving for p_1, p_2:

$$\begin{cases} p_1 &= 4p_3/5 \\ p_2 &= 2p_3/5 \end{cases}$$

Let us choose a value for p_3. Let us say $p_3=5$ (in this way we will get whole numbers for p_1,p_2). The solution is $p_1=4$ and $p_2=2$ and the equation is

$$(4,2,5)\cdot(x+1,y-1,z-1)=0 \qquad \Rightarrow \qquad 4x+2y+5z-3=0.$$

Answers to questions:

A7. Another way of solving question 6 is finding two points of the given line and, with their help, establish the vector equation of the plane wanted. How do we find points on line r? Solving the system of equations:

$$\begin{cases} 2x-3y+z+1 &= 0 \\ -x+y-z-3 &= 0 \end{cases}$$

This is a dependent system. We may use variable z as a parameter:

$$\begin{cases} 2x-3y &= -z-1 \\ -x+y &= z+3 \end{cases} \Rightarrow \begin{cases} x &= -2z-8 \\ y &= -z-5 \end{cases}$$

When $z=0$ we get point $(-8,-5,0)$ and when $z=1$, we get $(-10,-6,1)$. We can use point $(0,0,0)$ as a base point and the vectors with these endpoints and $(0,0,0)$ as origin as direction vectors for our plane. Its vector equation will be

$$(x,y,z)=t\cdot(-8,-5,0)+s\cdot(-10,-6,1).$$

We can see that it coincides with the answer obtained before. The parametric equation will be

$$\begin{cases} x &= -8t-10s \\ y &= -5t-6s \\ z &= s \end{cases}$$

Now, this system **must** have a solution for t and s. Consequently, the determinant

$$\begin{vmatrix} -8 & -10 & x \\ -5 & -6 & y \\ 0 & 1 & z \end{vmatrix} = 0.$$

This provides us with the general equation of our plane:

$$-5x + 8y - 2z = 0.$$

Back

Incidentally, this example provides us with a new way of getting directly the:

General equation of a plane given one of its points, say (a_1, a_2, a_3), and two direction vectors, say (u_1, u_2, u_3) and (v_1, v_2, v_3):

$$\begin{vmatrix} u_1 & v_1 & x - a_1 \\ u_2 & v_2 & y - a_2 \\ u_3 & v_3 & z - a_3 \end{vmatrix} = 0.$$

3 Day 3: Functions of two variables

This course deals mainly with two-variable functions, that is, functions whose real value depends on two independent variables x and y (or x_1 and x_2): $f(x, y)$. The value of the function is usually called z. The following are examples of functions of two variables:

$$z = x^2 + y^2; \quad z = \frac{x^2}{x + y}; \quad z = \ln(x - y); \quad z = \sqrt{x^2 + y^2 - 4}. \tag{4}$$

As in the case of one-variable functions, the image of $f(x, y)$ can be any real value with the only condition that it is unique: to a pair of values (x, y) there corresponds, at the most, **only one** value of $z = f(x, y)$.

3.1 The graph of $z = f(x, y)$

We will adopt the convention of representing (x, y) as a point on the plane $z = 0$ of $\mathbb{R}^3$, the xy-plane, and the corresponding image $f(x, y)$ on the z-axis. Thus, we obtain a point in space $(x, y, f(x, y))$ whenever $f(x, y)$ exists. Joining all these points we get the **graph** of our function $z = f(x, y)$ which is a **surface** in space. In fact, a plane of equation $z = mx + ny + p$ is an easy example of a two-variable function with a recognizable graph. All planes, except those parallel to the z-axis (**Why?**), are the graphs of functions.

In class we will see a few of those graphs obtained with computer software. There are many freeware programmes that plot 3D graphs. We will use mainly two

`www.wolframalpha.com`
`www.livephysics.com/ptools/online-3d-function-grapher.php`

Question 8. *Use these programmes to plot a few two-variable functions as $z = x^2 - y^2$; $z = e^{x^2+y^2}$, $z = \sin(x + y)$; $z = 1/(x + y)$; etc.*

Answer to question 8.

3.2 The domain of $z = f(x, y)$

The set of points in $\mathbb{R}^2$ for which $f(x, y)$ is defined is called the **domain** of f. The set of values in $\mathbb{R}$ that are images $f(x, y)$ for some values (x, y) is called the **range** of f. Notice that the domain is a subset of $\mathbb{R}^2$ and is part of the plane xy-plane and the range is a subset of $\mathbb{R}$ and is part of the z-axis.

When a function is defined as an algebraic formula in the variables x and y (see the examples above), the domain is the greatest subset of $\mathbb{R}^2$ for which the formula makes sense and produces an image.

Question 9. *What is the domain of the functions in example (4) above?*

Answer to question 9.

Question 10. *Draw the domains in question 9. Remember that, for a two-variable function, the domain is a set in the xy-plane!*

Answer to question 10.

3.3 Limits and continuity

We are not going into the details of limits of two-variable functions and the formal definition of continuity. You have already worked these ideas for one-variable functions and those are not too different in the case of two variables (though limits can be a little tricky). Continuity of a function $f(x, y)$ over a domain S is interpreted as the fact that the surface of its graph does **not break**. If it does, the lines over which this happens are called **lines of discontinuity**. In the function $z = x^2/(x + y)$, the line $x + y = 0$ is a line of discontinuity.

Functions defined through formulae that involve polynomials, rational expressions, powers, exponentials, logarithms and trigonometric functions are continuous on their domains.

3.4 Level curves

A nice way of looking at a two-variable graph is to use **level curves** as in a topographical map. We plot, in the same diagram, $f(x, y) = c$ for different values of c and these curves give you an idea of how the "terrain" (the surface of $f(x, y)$) looks like. Watch the video `www.tv3.cat/3alacarta/#/videos/3315331`, minutes 6–10.

Question 11. *For the function $z = 2x + 3y - 6$ draw in the same diagram the level curves*

$$z = -4, \ z = -2, \ z = 0, \ z = 2, \ z = 4.$$

Answer to question 11.

Question 12. *For the function $z = x^2 + y^2$ draw in the same diagram the level curves*

$$z = 1, \ z = 4, \ z = 9.$$

Answer to question 12.

Question 13. *For the function $z = \ln(y - x)$ draw in the same diagram the level curves*

$$z = -1, \ z = 0, \ z = \ln 2, \ z = 1.$$

Answer to question 13.

Answers:

A8. Use the web pages indicated.

Back

A9. (a) For $z = x^2 + y^2$ the domain is $\mathbb{R}^2$ (the whole plane) because $x^2 + y^2$ can be always be calculated for any (x, y). The range is quite obvious in this case: $[0, \infty)$.

(b) For $z = \dfrac{x^2}{x + y}$ the domain is the set of points:

$$S = \{(x, y) : x + y \neq 0\}.$$

This is the whole plane except for those points on the line $x + y = 0$.

(c) For $z = \ln(x - y)$ the domain is the set of points:

$$S = \{(x, y) : x - y > 0\}.$$

This is an open half-plane: the set of points on one side of the line $y = x$.

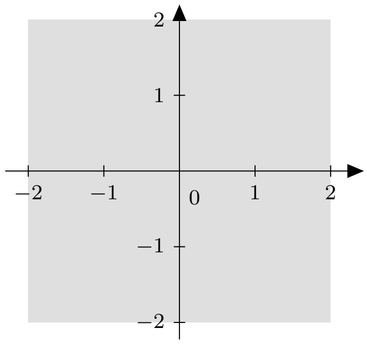
(d) Lastly, for $z = \sqrt{x^2 + y^2 - 4}$ the domain is the set of points:

$$S = \left\{(x, y) : x^2 + y^2 \geq 4\right\}.$$

This is the exterior of the circle of center $(0, 0)$ and radius 2, including the circumference itself.

Back

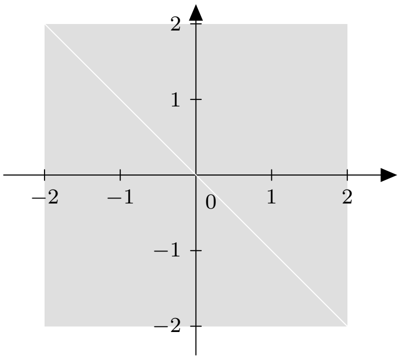
A10. (a) For $z = x^2 + y^2$ the domain is $\mathbb{R}^2$ (the whole plane). Graphically there is not much to see:



(b) For $z = \dfrac{x^2}{x+y}$ the domain is the set of points:

$$S = \{(x, y) : x + y \neq 0\}.$$

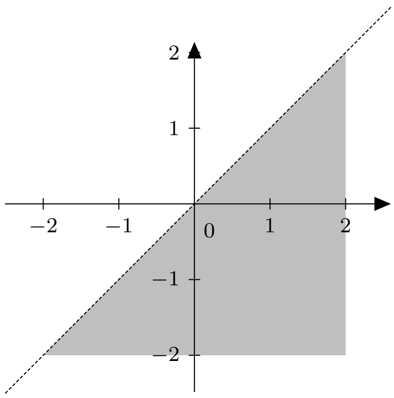
This is the whole plane except for those points on the line $x + y = 0$:



(c) For $z = \ln(x - y)$ the domain is the set of points:

$$S = \{(x, y) : x - y > 0\}.$$

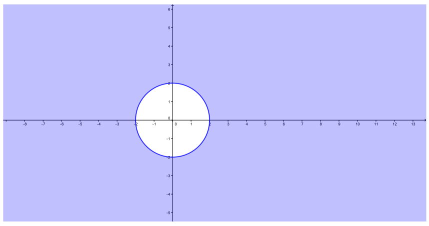
This is an open half-plane: the set of points below line $y = x$.



(d) Lastly, for $z = \sqrt{x^2 + y^2 - 4}$ the domain is the set of points:

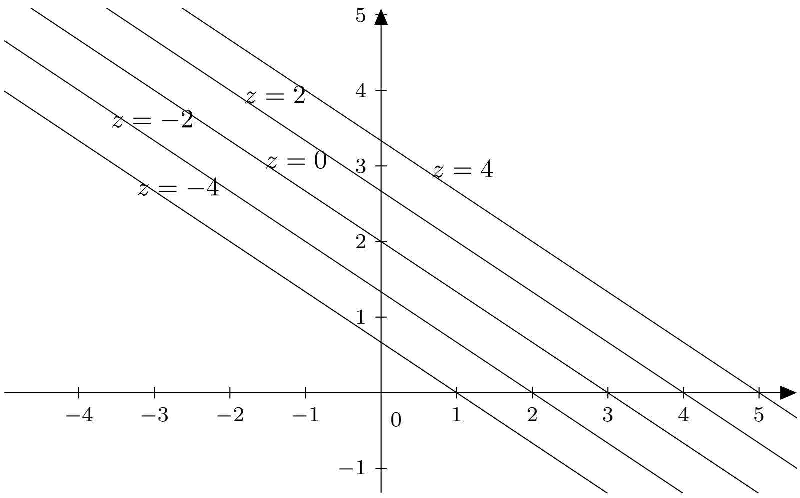
$$S = \left\{(x, y) : x^2 + y^2 \geq 4\right\}.$$

This is the exterior of the circle of center $(0, 0)$ and radius 2, including the circumference itself:



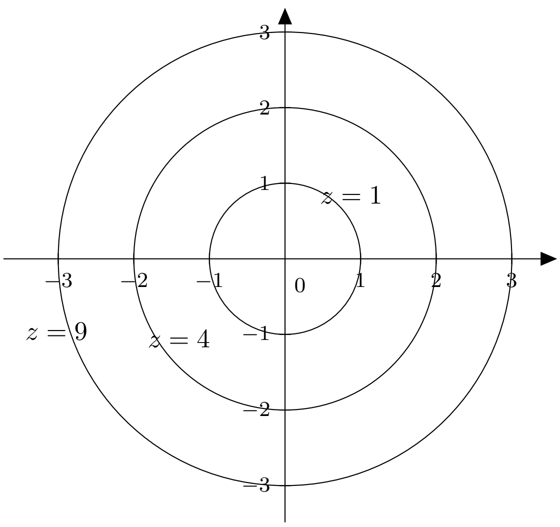
Back

A11.



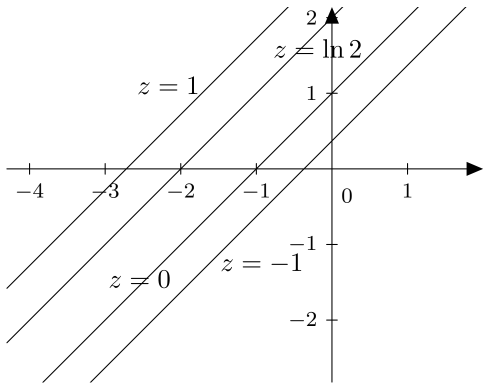
Back

A12.



Back

A13.



Back

4 Day 4: Partial derivatives

4.1 Review: one-variable derivative concept

Let us recall the definition of the derivative of a one-variable function, f, at an interior[3] point a of its domain:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}. \tag{5}$$

The numerator of the fraction in the limit is the increase (or decrease) of the value of function f as x moves from a to nearby $a+h$: $\triangle f = f(a+h) - f(a)$. The fraction in the limit reflects the "slope" of the graph of f as if it were a straight line joining points $(a, f(a))$ and $(a+h, f(a+h))$. In economics, the derivative is called the "marginal" of the function at point a. As a rough approximation, the limit in (5) is sometimes taken as if $h = 1$

$$f'(a) \approx f(a+1) - f(a)$$

and the marginal value is so the "value of the last unit". As an example, think of the profit function of a firm, $\pi(q)$, as a function of the production q. The marginal profit when $q = 30$ would be $\pi(31) - \pi(30)$, that is, the increase in profit as a consequence of moving production from 30 to 31 units.

4.2 Definition of partial derivative

Let us now think of a function of two variables: $z = f(x, y)$ and consider an interior point in its domain (a, b). How can we define the concept of derivative of f at (a, b)? Moving from (a, b) can now be done in many directions! What if we try to move to a nearby point by just changing one variable at a time? We can pass from (a, b) to $(a+h, b)$ and so just change variable x or we can move from (a, b) to $(a, b+k)$ and so change only variable y. Mimicking definition (5), we can now define **two** derivatives at (a, b):

$$f'_x(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}; \tag{6}$$

$$f'_y(a,b) = \lim_{k \to 0} \frac{f(a,b+k) - f(a,b)}{k} \tag{7}$$

which we will call **partial derivatives** with respect to (w.r.t.) x and y respectively. These derivatives have different notations (that we will use indistinctly):

$$f'_x \text{ and } f'_y \quad \text{or} \quad f'_1 \text{ and } f'_2 \quad \text{or} \quad \frac{\partial f}{\partial x} \text{ and } \frac{\partial f}{\partial y} \quad \text{or} \quad D_1 f \text{ and } D_2 f.$$

When we are given the point (a, b) at which the derivative is taken we write

$$f'_x(a,b) \text{ and } f'_y(a,b) \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{(a,b)} \text{ and } \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \quad \text{or} \quad D_1 f(a,b) \text{ and } D_2 f(a,b).$$

[3] The point has to be interior to the domain to allow h to take positive and negative values while keeping $a+h$ inside the domain.

It is not difficult to see that the same rules that are valid for one-variable derivatives are also valid for partial derivatives and so, when a function is defined by a formula, the partial derivatives —as functions of (x, y)— can be obtained using the usual calculus rules. We only have to fix the value of the other variable (consider it a constant) and proceed as if the function were a one-variable function. Take, for example,

$$f(x,y) = x^3 + 3x^2y - 5xy^2 + xy - y^4 \qquad \text{we have} \qquad \begin{cases} f'_x &= 3x^2 + 6xy - 5y^2 + y; \\ f'_y &= 3x^2 - 10xy + x - 4y^3. \end{cases}$$

4.3 Linear approximations of the derivatives

As in the one-variable case, we can also interpret the partial derivative of a two-variable function as a marginal value w.r.t. one of the variables (keeping the other one fixed):

$$f'_x(a,b) \approx f(a+1,b) - f(a,b);$$

$$f'_y(a,b) \approx f(a,b+1) - f(a,b).$$

This approach allows us to study the effect of altering one variable without changing the other. For example, if $P(K, L)$ is a production function depending on the variables K (capital) and L (labour), $P'_x(300, 70)$ would be, approximately, the increase in production when K changes from 300 to 301 (keeping labour at 70). Using symbols:

$$P'_x(300,70) \approx P(301,70) - P(300,70).$$

In fact, the approximation also works even if we use a value of h and k different from 1:

$$f'_x(a,b) \approx \frac{f(a+h,b) - f(a,b)}{h} \quad \Rightarrow \quad \boxed{f(a+h,b) \approx f(a,b) + h \cdot f'_x(a,b)}; \tag{8}$$

$$f'_y(a,b) \approx \frac{f(a,b+k) - f(a,b)}{k} \quad \Rightarrow \quad \boxed{f(a,b+k) \approx f(a,b) + k \cdot f'_y(a,b)}. \tag{9}$$

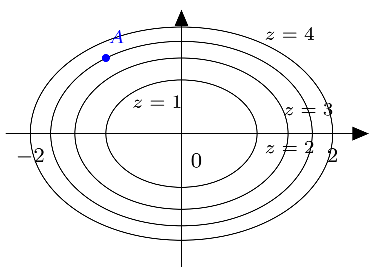
The closer h and k are from 0, the better the approximation.

As it is seen from these last formulae, the value of f at a point (a, b) and the value of its derivatives there can be used to approximate linearly the value of f at a nearby point. This is particularly useful when both, $f(a, b)$ and $f'_x(a, b)$ (or $f'_y(a, b)$) are known (or easy to find). The next question illustrates this point.

Question 14. *Find the approximate value of* $f(x, y) = \ln(\frac{x}{2} + y)$ *at point* $(0.1, 1)$ *using the linear approximation given by the derivative w.r.t.* x *at point* $(0, 1)$. *(*Hint*: what are the values of* $f(0, 1)$ *and* $f'_x(0, 1)$*?).* *Answer to question 14.*

4.4 Graphic approach

These variations on the (infinitesimal) value of functions can also be easily read on a level curves map (contour plot) of the function. Consider $f(x, y) = x^2 + 2y^2$. The level curves $z = 1$, $z = 2$, $z = 3$, etc. are (inside out) ellipses centered on $(0, 0)$:



At a point like A in the figure, lying on $z = 3$, $f'_x(A) < 0$ because moving from A horizontally to the right (that is, increasing the x coordinate by $h > 0$ while the y coordinate stays the same) causes the value of the image to decrease as it moves from $z = 3$ towards $z = 2$. With a similar reasoning, $f'_y(A) > 0$ because moving from A vertically upwards (that is, increasing the y coordinate by $k > 0$ while the x coordinate stays the same) causes the value of the image to increase as it moves from $z = 3$ towards $z = 4$.

4.5 Second-order derivatives.

Each first-order derivative of $f(x,y)$ is a two-variable function. As such we can find for it its first order derivatives. These will be the second-order derivatives of f. There are 4 of these:

$$f''_{xx},\ f''_{xy},\ f''_{yx},\ f''_{yy} \quad \text{or} \quad \frac{\partial^2 f}{\partial x^2},\ \frac{\partial^2 f}{\partial y\,\partial x},\ \frac{\partial^2 f}{\partial x\,\partial y},\ \frac{\partial^2 f}{\partial y^2}.$$

Fortunately, for the most usual functions we have that the **mixed derivatives**, f''_{xy} and f''_{yx} are equal. This fact is known as Young's Theorem:

Young's Theorem. *If f''_{xy} and f''_{yx} exist at a point and are continuous, they coincide.*

Young's Theorem extend to third-order derivatives, etc. For example, it says that

$$f'''_{xyx} = f'''_{xxy} = f'''_{yxx}$$

whenever these three derivatives exist and are continuous.

Question 15. *Check Young's Theorem with the function of the example above: $f(x,y) = x^3 + 3x^2y - 5xy^2 + xy - y^4$ for which we had*

$$f'_x = 3x^2 + 6xy - 5y^2 + y$$
$$f'_y = 3x^2 - 10xy + x - 4y^3.$$

Answer to question 15.

4.6 The Hessian matrix.

The Hessian matrix of a function $z = f(x,y)$ is the following matrix formed by the 4 second-order partial derivatives:

$$H(x,y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial y \partial x} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

The Hessian matrix plays a fundamental role at the time of classifying maxima and minima of two-variable functions.

As an easy corollary of Young's Theorem, the Hessian matrix of a function is a **symmetric** matrix.

Question 16. *Find the Hessian matrix of $z = e^{3x+2y}$. What is its value at the point $(1,1)$?* *Answer to question 16.*

Question 17. *What is the value of the determinant of the Hessian matrix we have just found?* *Answer to question 17.*

Answers to questions:

A14. In this case, $f(0,1) = \ln 1 = 0$ and $f'_x = \dfrac{1}{2\left(\frac{x}{2} + y\right)}$ gives $f'_x(0,1) = 0.5$. We use (8) with $h = 0.1$:

$$f(0.1,1) = \ln(1.05) \approx f(0,1) + 0.1 \cdot f'_x(0,1) = 0 + 0.1 \cdot 0.5 = 0.05.$$

This is a good approximation of $\ln(1.05)$ whose real value is $0.04879\ldots$ **Back**

A15. The second order mixed derivatives are

$$f''_{xy} = 6x - 10y + 1 \quad \text{and} \quad f''_{yx} = 6x - 10y + 1$$

which are exactly the same. As for the third-order mixed derivatives,

$$f'''_{xyx} = 6;\ f'''_{xxy} = (6x + 6y)'_y = 6;\ f'''_{yxx} = 6.$$

They also coincide. **Back**

A16. The first-order derivatives are

$$\frac{\partial z}{\partial x} = 3e^{3x+2y};\ \frac{\partial z}{\partial y} = 2e^{3x+2y}.$$

The second order partial derivatives are:

$$\frac{\partial^2 z}{\partial x^2} = 9e^{3x+2y}; \quad \frac{\partial^2 z}{\partial y \partial x} = 6e^{3x+2y}$$
$$\frac{\partial^2 z}{\partial y^2} = 4e^{3x+2y}; \quad \frac{\partial^2 z}{\partial x \partial y} = 6e^{3x+2y}.$$

The Hessian is

$$H(x,y) = \begin{pmatrix} \frac{\partial^2 z}{\partial x^2} & \frac{\partial^2 z}{\partial y \partial x} \\ \frac{\partial^2 z}{\partial x \partial y} & \frac{\partial^2 z}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 9e^{3x+2y} & 6e^{3x+2y} \\ 6e^{3x+2y} & 4e^{3x+2y} \end{pmatrix}.$$

Notice that the Hessian is a "function" of (x,y), that is, it changes with each point (x,y):

$$H(0,0) = \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}; \quad H(1,1) = \begin{pmatrix} 9e^5 & 6e^5 \\ 6e^5 & 4e^5 \end{pmatrix} = e^5 \cdot \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}.$$

(Remember that matrices and determinants behave differently when multiplied by an external number.) Back

A17.

$$|H(1,1)| = \begin{vmatrix} 9e^5 & 6e^5 \\ 6e^5 & 4e^5 \end{vmatrix} = e^5 \cdot e^5 \begin{vmatrix} 9 & 6 \\ 6 & 4 \end{vmatrix} = e^{10} \cdot 3 \cdot 2 \begin{vmatrix} 3 & 2 \\ 3 & 2 \end{vmatrix} = 0.$$

(Remember that matrices and determinants behave differently when multiplied by an external number.) Back

5 Day 5: Tangent plane.

The simplest surface we can imagine in $\mathbb{R}^3$ is a plane. If the plane is not parallel to the z-axis, its equation may be written in the form (**Why?**):

$$z = mx + ny + p, \quad \text{where } m, n, p \in \mathbb{R}.$$

We may now think of $mx + ny + p$ as the image of a two-variable function:

$$g(x,y) = ax + by + c.$$

Thus, a non vertical plane is the graph of the two-variable (linear) function $g(x,y)$ above. Can we find its normal vector? No problem. We write the general equation plane of our plane

$$mx + ny - z + p = 0 \quad \Rightarrow \text{ normal vector: } \mathbf{v} = (m, n, -1).$$

We now notice that $m = g'_x$ and $n = g'_y$. Consequently, our normal vector can be written also as

$$\mathbf{v} = (g'_x, g'_y, -1).$$

Let us now consider a function $f(x,y)$ and one of the points on its surface: $P = (a, b, f(a,b))$. If the surface is smooth enough (we will later define the concept of "smoothness"), you will agree that if we stay close to P the surface is approximately flat.

We define the **tangent plane** to our surface at P as the plane that passes through P and has

$$(f'_x(a,b), f'_y(a,b), -1)$$

as normal vector. Its equation will be

$$f'_x(a,b) \cdot x + f'_y(a,b) \cdot y - z + d = 0.$$

The value of d can be found by imposing that point P must belong to the plane:

$$f'_x(a,b) \cdot a + f'_y(a,b) \cdot b - f(a,b) + d = 0 \quad \Rightarrow \quad d = -f'_x(a,b) \cdot a - f'_y(a,b) \cdot b + f(a,b).$$

The equation of the **tangent plane** to $f(x,y)$ at the point $(a, b, f(a,b))$ of its surface can be written as

$$z = f(a,b) + f'_x(a,b) \cdot (x - a) + f'_y(a,b) \cdot (y - b).$$

Question 18. *Find the equation of the tangent plane to $f(x,y) = x^3 + xy^2 - 3xy$ at point $(2,1)$* *Answer to question 18.*

Question 19. *Is $x - z = 0$ the equation of one of the tangent planes to $f(x,y) = \sqrt{x^2 - y^2}$? If the answer is yes, what is the point of tangency?* *Answer to question 19.*

Question 20. *Is $2x - 3y - 2z + 4 = 0$ the equation of one of the tangent planes to $f(x,y) = x^2 + y^2$? If the answer is yes, what is the point of tangency?* *Answer to question 20.*

Answers to questions:

A18. We need to find $f(2,1)$, $f'_x(2,1)$, and $f'_y(2,1)$:

$$f(x,y) = x^3 + xy^2 - 3xy \quad \Rightarrow \quad f(2,1) = 4;$$
$$f'_x(x,y) = 3x^2 + y^2 - 3y \quad \Rightarrow \quad f'_x(2,1) = 10;$$
$$f'_y(x,y) = 2xy - 3x \quad \Rightarrow \quad f'_y(2,1) = -2.$$

The equation of the tangent plane is

$$z = 4 + 10(x-2) - 2(y-1) \quad \Leftrightarrow \quad z = 10x - 2y - 14.$$

Back

A19. Let us write the equation of the plane solved for z:

$$x - z = 0 \quad \Rightarrow \quad z = x.$$

If this is the equation of one of the tangent planes, we must have a point (a,b) in the domain of f for which

$$f'_x(a,b) = 1; \quad \text{that is} \quad \frac{a}{\sqrt{a^2-b^2}} = 1;$$
$$f'_y(a,b) = 0; \quad \text{that is} \quad \frac{-b}{\sqrt{a^2-b^2}} = 0$$

The second equation implies that $b = 0$ which, inserted into the first one, gives

$$\frac{a}{\sqrt{a^2}} = \frac{a}{|a|} = 1 \qquad \Rightarrow \qquad \text{any } a > 0 \text{ is a solution.}$$

Thus, the point should be $(a,0)$ with $a > 0$. In this case $f(a,0) = a$ and the tangent plane would be

$$z = a + (x-a) \quad \Leftrightarrow \quad z = x.$$

It coincides with the plane given. The answer is yes and there are an infinity of points of tangency: $(a,0,a)$ for $a > 0$.

Back

A20. Let us write the equation of the plane solved for z:

$$2x - 3y - 2z + 4 = 0 \quad \Rightarrow \quad z = x - \frac{3}{2}y + 2.$$

If this is the equation of one of the tangent planes, we must have a point (a,b) in the domain of f for which

$$f'_x(a,b) = 1; \quad \text{that is} \quad 2a = 1;$$
$$f'_y(a,b) = -3/2; \quad \text{that is} \quad 2b = -3/2.$$

Thus, we must have $a = 1/2$ and $b = -3/4$. In that case, $f(1/2,-3/4) = (1/2)^2 + (-3/4)^2 = 13/16$ and the tangent plane would be

$$z = \frac{13}{16} + 1 \cdot \left(x - \frac{1}{2}\right) - \frac{3}{2} \cdot \left(y + \frac{3}{4}\right) \quad \Leftrightarrow \quad z = x - \frac{3}{2}y - 5/16.$$

The equation is not equivalent to the one provided. The planes are different. The answer is no. Back

6 Day 6: Quadratic forms in two variables

We mentioned some days ago that the process to find the max/min of a two-variable function is similar to the one used in the case of a function of one variable. Particularly the "second derivative test", that allowed us to classify a candidate point as a max or a min (or none) by finding the sign of f'' at the candidate, will be replaced by finding the "sign" of the Hessian matrix of $f(x,y)$. This is the purpose of this section.

A function of the form

$$f(x,y) = ax^2 + 2bxy + cy^2$$

is called a **quadratic form** in two variables.

It is quite easy to check that

$$ax^2 + 2bxy + cy^2 = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

The square matrix above is called the **matrix of the quadratic form**.

Notice that the matrix of a quadratic form is **symmetric**.

Reciprocally, any symmetric matrix 2×2 can be seen as the matrix of a quadratic form in two variables.

Question 21. *If $f(x,y) = x^3 + 2xy$, write the Hessian of f at point $(2,1)$ as a quadratic form.* *Answer to question 21.*

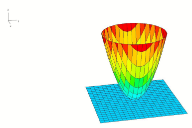
6.1 Classification of a quadratic form in two variables.

Quadratic forms are classified according to the sign of its values.

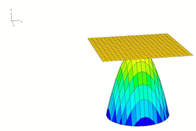
Positive Definite (PD)	$f(x,y) > 0$ for all $(x,y) \neq (0,0)$
Negative Definite (PN)	$f(x,y) < 0$ for all $(x,y) \neq (0,0)$
Positive Semidefinite (PSD)	$f(x,y) \geq 0$ for all $(x,y) \neq (0,0)$
Negative Semidefinite (NSD)	$f(x,y) \leq 0$ for all $(x,y) \neq (0,0)$
Indefinite (I)	$f(x,y) > 0$ for some (x,y) and $f(x,y) < 0$ for other (x,y)

Examples.

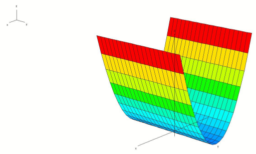
- $f(x,y) = x^2 + y^2$ is PD.
- $f(x,y) = -x^2 - y^2$ is ND.
- $f(x,y) = x^2$ is PSD.
- $f(x,y) = -x^2$ is NSD.
- $f(x,y) = x^2 - y^2$ is I.



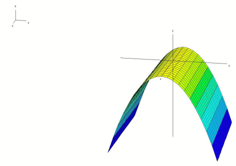
1.1 $z = x^2 + y^2$ (PD)



1.2 $z = -x^2 - y^2$ (ND)

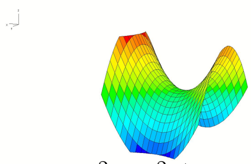


1.3 $z = x^2$ (PSD)



1.4 $z = -y^2$ (NSD)

The indefinite case has the most interesting graph, a *horse saddle mount*:



$z = x^2 - y^2$ (Indefinite)

These examples are quite obvious. In other cases, the character of $f(x,y)$ is not so easy to decide. For example

$$f(x,y) = x^2 - 6xy + 9y^2.$$

We see, for instance, that $f(1,1) = 4$, $f(1,0) = 1$, $f(0,1) = 9$, etc. It seems positive, but ... can we be sure that for all $(x,y) \neq (0,0)$, $x^2 - 6xy + 9y^2 > 0$? Fortunately there is an easy way to check. The following table tells us how to do it.

PD	$\Leftrightarrow$	$a > 0, c > 0, \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0$
ND	$\Leftrightarrow$	$a < 0, c < 0, \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0$
PSD	$\Leftrightarrow$	$a \geq 0, c \geq 0, \begin{vmatrix} a & b \\ b & c \end{vmatrix} \geq 0$
NSD	$\Leftrightarrow$	$a \leq 0, c \leq 0, \begin{vmatrix} a & b \\ b & c \end{vmatrix} \geq 0$
I	$\Leftrightarrow$	$\begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0$

The same table can be read from right to left and even simplified:

$\begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0$	$f(x,y)$ is Definite	Positive if $a > 0$ / Negative if $a < 0$
$\begin{vmatrix} a & b \\ b & c \end{vmatrix} = 0$	$f(x,y)$ is Semidefinite	Positive if $a \geq 0$ and $c \geq 0$ / Negative if $a \leq 0$ and $c \leq 0$
$\begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0$	$f(x,y)$ is indefinite	

6.2 A direct proof of the cases above

- **Case** $a = 0, c = 0$, and $b \neq 0$. In that particular case, the form reduces to $f(x,y) = 2bxy$ which is indefinite and satisfies that $\begin{vmatrix} 0 & b \\ b & 0 \end{vmatrix} < 0$.

In the following cases, we write our quadratic form in a different way by using simple algebraic manipulations.

- **Case** $a > 0$. We can change the expression of f as follows:

$$f(x,y) = ax^2 + 2bxy + cy^2 = a\cdot\left(x^2 + 2\frac{b}{a}xy + \frac{c}{a}y^2\right) =$$
$$= a\cdot\left[\left(x + \frac{b}{a}y\right)^2 - \frac{b^2}{a^2}y^2 + \frac{c}{a}y^2\right] =$$
$$= a\cdot\left[\left(x + \frac{b}{a}y\right)^2 + \left(\frac{c}{a} - \frac{b^2}{a^2}\right)y^2\right] =$$
$$= a\cdot\left(x + \frac{b}{a}y\right)^2 + \left(\frac{ac - b^2}{a^2}\right)y^2. \tag{10}$$

Now, clearly, $a\cdot(x + by/a)^2 > 0$ for $(x,y) \neq (0,0)$ and it is also obvious that

$$\text{As } a > 0, \text{ if } ac - b^2 = \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0 \quad \Rightarrow \quad f \quad \text{Positive definite;}$$
$$\text{As } a > 0, \text{ if } ac - b^2 = \begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0 \quad \Rightarrow \quad f \quad \text{Indefinite;}$$
$$\text{As } a > 0 \text{ (and } c \text{ cannot be } < 0\text{), if } ac - b^2 = \begin{vmatrix} a & b \\ b & c \end{vmatrix} = 0 \quad \Rightarrow \quad f \quad \text{Positive semidefinite.}$$

The reciprocal result [$\Leftarrow$] is also easily proved from (10).

- **Case** $a < 0$. In the previous case, change "positive" by "negative".

There is a more tedious proof (see Sydsaeter & Hammond, Section 15.8) that treats case by case.

Finally, let us classify the example given before: $f(x,y) = x^2 - 6xy + 9y^2$.
Here $a = 1, b = -3$, and $c = 9$:

$$\begin{vmatrix} 1 & -3 \\ -3 & 9 \end{vmatrix} = 0 \quad \text{and} \quad a > 0, c > 0$$

imply that $f(x,y)$ is positive semidefinite. That means that it is never negative and that it is 0 for values different from $(0,0)$. For instance, $f(3,1) = 0$.

6.3 Quadratic forms in two variables restricted to a linear constraint.

In some occasions we will be interested in classifying a quadratic form where x and y are not free but forced to fulfill a homogeneous linear constraint (this means that (x,y) lies on a line through $(0,0)$). The problem is quite easy to solve. Imagine we are dealing with $f(x,y) = ax^2 + 2bxy + cy^2$ and $px + qy = 0$. We can assume $q \neq 0$ and solve the constraint for y: $y = -(p/q)x$. Replacing in f:

$$f\left(x, -\frac{px}{q}\right) = \frac{1}{q^2}\left(aq^2 - 2bpq + cp^2\right) \cdot x^2$$

Thus, the sign of $D = aq^2 - 2bpq + cp^2$ will determine the character positive or negative of f restricted to the line. There is an easy way of remembering that value:

$$aq^2 - 2bpq + cp^2 = -\begin{vmatrix} 0 & p & q \\ p & a & b \\ q & b & c \end{vmatrix}.$$

D is exactly the opposite of what we call the **bordered hessian** of $f(x,y)$ restricted to $px + qy = 0$.

Thus, the sign (PD or ND) of $f(x,y)$ restricted to the constraint is the opposite of the sign of the bordered hessian.

Question 22. *Can f constrained to a line be indefinite?* *Answer to question 22.*

Question 23. *If f is PD, can f constrained to a line be ND?* *Answer to question 23.*

Question 24. *What happens if f is indefinite?* *Answer to question 24.*

Answers to questions:

A21. The Hessian of f is

$$H(x,y) = \begin{pmatrix} 6x & 2 \\ 2 & 0 \end{pmatrix}; \quad \text{at point } (2,1),\ H(2,1) = \begin{pmatrix} 12 & 2 \\ 2 & 0 \end{pmatrix}.$$

The Hessian is a symmetric matrix. Consequently, we can look at it as a quadratic form:

$$Q(x,y) = 12x^2 - 4xy.$$

Back

A22. Obviously not. The value $aq^2 - 2bpq + cp^2$ is either positive (then f constrained is PD), negative (then f constrained is ND) or 0 and then $f = 0$ and does not take either positive or negative values. **Back**

A23. Obviously not. If f is PD, then **for all** $(x,y) \neq (0,0)$, $f(x,y) > 0$, in particular for those $(x,y) \neq (0,0)$ situated on any line whatsoever. **Back**

A24. This is the really interesting case. If f is indefinite, then

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0.$$

But $aq^2 - 2bpq + cp^2$ is also a quadratic form on the variables (q,p). Its matrix is

$$\begin{pmatrix} a & -b \\ -b & c \end{pmatrix}$$

and its determinant is

$$\begin{vmatrix} a & -b \\ -b & c \end{vmatrix} = \begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0.$$

Consequently, $aq^2 - 2bpq + cp^2$ as a quadratic form is indefinite and will take positive or negative values depending on the values of q and p. **Back**

7 Day 7: Chain Rule. Directional derivatives. Gradient

7.1 The Chain Rule

In many occasions, the variables in a function will, in their turn, be functions of one or more variables. The situation is the following. Let $F(x,y)$ be a two-variable function and let $x = f(t)$ and $y = g(t)$. Then, in fact $F(f(t), g(t))$ can be thought of as a one-variable function of the variable t. Can we obtain dF/dt, **the total derivative**, indirectly? That is, without having an explicit algebraic expression for it?

Let us work out an example. Imagine you have $F(x,y) = x^2 + y^2$ and we know that $x = t^2$ and $y = 2t$. It is pretty obvious that, in this case, $F(x(t), y(t)) = (t^2)^2 + (2t)^2 = t^4 + 4t^2$. If we want to get dF/dt, the direct way is to use this last value:

$$F(t) = t^4 + 4t^2 \quad \Rightarrow \quad \frac{dF}{dt} = 4t^3 + 8t.$$

What about the indirect way? The indirect way is called the **Chain Rule**:

> If $F(x,y)$ is a two-variable function and $x = f(t)$ and $y = g(t)$, then
>
> $$\frac{dF}{dt} = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt}.$$

In our example we have $F'_x = 2x$ and $F'_y = 2y$. Then

$$\frac{dF}{dt} = 2x \cdot \frac{dx}{dt} + 2y \cdot \frac{dx}{dt} = 2x \cdot (2t) + 2y \cdot (2) = 2 \cdot t^2 \cdot 2t + 2 \cdot 2t \cdot 2 = 4t^3 + 8t.$$

This result is manly used in complicated cases or in theoretical reasonings. As you can imagine, the possibility of the direct substitution method is quite straightforward in most of the occasions.

7.2 Directional derivatives

Let (a,b) be a point in the domain S of a function $f(x,y)$. We know that $\partial f/\partial x|_{(a,b)}$ and $\partial f/\partial y|_{(a,b)}$ are the slopes of the function as we face the direction $(1,0)$ and $(0,1)$ respectively. That means that

$$\left.\frac{\partial f}{\partial x}\right|_{(a,b)} = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} \quad \text{and} \quad \left.\frac{\partial f}{\partial y}\right|_{(a,b)} = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}$$

If, instead of the directions above, we move from (a,b) in the direction of vector $\vec{v} = (v_1, v_2)$ and we calculate the corresponding differential quotient limit, we will obtain the **directional derivative** of f in the direction of $\vec{v}$:

$$D_{\vec{v}} f(a,b) = \lim_{h\to 0} \frac{f(a + h \cdot v_1, b + h \cdot v_2) - f(a,b)}{h}.$$

In order not to alter the units of measure, we require that vector $\vec{v} = (v_1, v_2)$ be a **unitary** vector, that is a vector of norm one: $\|\vec{v}\| = 1$. Remember that the module of a vector (v_1, v_2) is its length: $\sqrt{v_1^2 + v_2^2}$. If the given direction is not a unitary vector, the vector has to be changed by another with the same direction and sense but module 1. This is easily done by multiplying the given vector by $1/\|\vec{v}\|$.

> For any vector $\vec{v} \neq (0,0)$, the vector
>
> $$\vec{w} = \frac{1}{\|\vec{v}\|} \cdot \vec{v}$$
>
> shares the direction of $\vec{v}$, keeps it sense but has norm one.

Question 25. *Given the function $f(x,y) = x/y$ find the directional derivative in the direction of vector $\vec{v} = (-1,2)$ at point $(3,1)$.*
Answer to question 25.

7.3 An easier way to find directional derivatives.

Another way of obtaining the directional derivative is the following.

Let $g(t) = f(a + t \cdot v_1, v + t \cdot v_2)$. If $\|(v_1, v_2)\| = 1$, the directional derivative is just $dg/dt|_{t=0}$:

$$D_{\vec{w}} f = g'(t) = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} = \frac{\partial f}{\partial x} \cdot v_1 + \frac{\partial f}{\partial y} \cdot v_2 = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \bullet (v_1, v_2)$$ [4]

Now, for $t = 0$,

$$g'(0) = \left(\left.\frac{\partial f}{\partial x}\right|_{(a,b)}, \left.\frac{\partial f}{\partial y}\right|_{(a,b)}\right) \bullet (v_1, v_2)$$

In question 25 above,

$$\left(\left.\frac{\partial f}{\partial x}\right|_{(a,b)}, \left.\frac{\partial f}{\partial y}\right|_{(a,b)}\right) = \left.\left(\frac{1}{y}, -\frac{x}{y^2}\right)\right|_{(3,1)} = (1, -3).$$

Then,

$$D_{\vec{w}} f(3,1) = (1,-3) \bullet \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right) = -\frac{7}{\sqrt{5}}.$$

[4] This last operation is the dot product (scalar product) of two vectors.

7.4 The gradient.

The vector

$$\nabla f(a,b) = \left(\left.\frac{\partial f}{\partial x}\right|_{(a,b)}, \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \right)$$

is called the **gradient** of f at the point (a,b).

Our previous section shows that $D_{\vec{v}} f(a,b) = \nabla f(a,b) \bullet \vec{v} \quad$ where $\|\vec{v}\| = 1$.

7.5 Properties of the gradient.

The gradient of a function has two important properties:

i) $\nabla f(a,b)$ points in the direction of **maximum growth** from (a,b). That is, the direction of the gradient is the one with a **maximum** directional derivative.

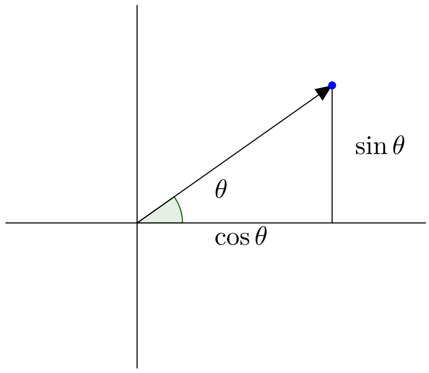
ii) $\nabla f(a,b)$ is orthogonal to the level curve at level $z = f(a,b)$.

We are going to prove these two statements.

Proof of i)

- *Short proof.* We know that the dot product of two vectors forming an angle α between them can be found as
$$\vec{u} \bullet \vec{v} = \|\vec{u}\| \cdot \|\vec{v}\| \cdot \cos\alpha$$
Then, since $-1 \leq \cos\alpha \leq 1$, the maximum value of the dot product of two vectors with a given module must be found when $\cos\alpha = 1$. That is, when $\alpha = 0°$.
Let us apply that to directional derivatives. Given a point (a,b), let us consider any unitary vector ($\|\vec{v}\| = 1$) $\vec{v}$. The directional derivative in the direction of $\vec{v}$ is
$$D_{\vec{v}} f(a,b) = \nabla f(a,b) \bullet \vec{v}.$$
It will be max when $\vec{v}$ forms a $0°$ with the gradient, that is, when it points in the same direction and sense as the gradient!

- *Not so short proof.*
Any vector of length 1 can be written as $(\cos\theta, \sin\theta)$. That is because for any angle $0° \leq \theta < 360°$, $\cos^2\theta + \sin^2\theta = 1$.



Thus, $\|(\cos\theta, \sin\theta)\| = 1$ and the directional derivative in the direction of vector $(\cos\theta, \sin\theta)$ is

$$D_{(\cos\theta,\sin\theta)} f(a,b) = \vec{\nabla} f(a,b) \bullet (\cos\theta, \sin\theta) = \left.\frac{\partial f}{\partial x}\right|_{(a,b)} \cdot \cos\theta + \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \cdot \sin\theta.$$

We can define a one-variable function

$$s(\theta) = \left.\frac{\partial f}{\partial x}\right|_{(a,b)} \cdot \cos\theta + \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \cdot \sin\theta$$

whose images are the different slopes (directional derivatives) from point (a,b) in every possible direction. What is the maximum value of $s(\theta)$? Easy:

$$s'(\theta) = -\left.\frac{\partial f}{\partial x}\right|_{(a,b)} \cdot \sin\theta + \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \cdot \cos\theta = 0$$

If the solution to this equation is θ_0, we can write

$$\tan\theta_0 = \frac{\left.\dfrac{\partial f}{\partial y}\right|_{(a,b)}}{\left.\dfrac{\partial f}{\partial x}\right|_{(a,b)}}.$$

So $\tan\theta_0$ is exactly the slope of vector

$$\left(\left.\frac{\partial f}{\partial x}\right|_{(a,b)}, \left.\frac{\partial f}{\partial y}\right|_{(a,b)}\right),$$

the direction of the gradient as we contended!!

Proof of ii)

What is the tangent line to a level curve at one of its points? Imagine we are at point (a,b). The level curve that passes through it is $f(x,y) = f(a,b)$. If, near (a,b), we consider our level curve as the graph of a one-variable function, $y = y(x)$, for which $b = y(a)$, the tangent vector is

$$\left(1, \left.\frac{dy}{dx}\right|_a\right).$$

We will say that in the neighbourhood of (a,b), y can be seen as a function of variable x: an **implicit function**.

How do we find dy/dx? Well, if $f(x,y) = f(a,b)$ we can consider $x = x; y = y(x)$. In this way, we have

$$f(x, y(x)) = f(a,b); \qquad \text{and taking derivatives on both sides w.r.t. } x,$$

we have

$$\frac{df}{dx} = 0.$$

We now use the chain rule in order to find df/dx:

$$\frac{\partial f}{\partial x}\cdot\frac{dx}{dx} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dx} = 0.$$

Obviously, $dx/dx = 1$ and we can write

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dx} = 0,$$

that is:

$$\boxed{\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}.} \tag{11}$$

We will soon give a name to this important formula: **the Implicit Function Theorem**.

For this formula to be valid is necessary that, at point (a,b), we have

$$\left.\frac{\partial f}{\partial y}\right|_{(a,b)} \neq 0.$$

The tangent vector we were looking for is

$$\left(1, -\frac{\left.\dfrac{\partial f}{\partial x}\right|_{(a,b)}}{\left.\dfrac{\partial f}{\partial y}\right|_{(a,b)}}\right),$$

or better,

$$\left(\left.\frac{\partial f}{\partial y}\right|_{(a,b)}, -\left.\frac{\partial f}{\partial x}\right|_{(a,b)}\right).$$

The equation of the tangent line (in the xy plane) is:

$$y - b = -\frac{\left.\frac{\partial f}{\partial x}\right|_{(a,b)}}{\left.\frac{\partial f}{\partial y}\right|_{(a,b)}} \cdot (x - a) \tag{12}$$

Lastly, it is seen at once that:

$$\left(\left.\frac{\partial f}{\partial x}\right|_{(a,b)}, \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \right) \bullet \left(\left.\frac{\partial f}{\partial y}\right|_{(a,b)}, -\left.\frac{\partial f}{\partial x}\right|_{(a,b)} \right) = 0$$

which proves that the gradient at a point is orthogonal to the level curve through this point.

Answers to questions:

A25. First of all, we must check the module of vector $\vec{v}$:

$$\|\vec{v}\| = \|(-1,2)\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}.$$

As the module is not 1, we change our vector by multiplying it by $1/\sqrt{5}$: $\vec{w} = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right)$.

Now,

$$D_{\vec{w}} f(3,1) = \lim_{h \to 0} \frac{f(3 + h \cdot (-1)/\sqrt{5}, 1 + h \cdot 2/\sqrt{5}) - f(3,1)}{h} = \lim_{h \to 0} \frac{\frac{3 - \frac{h}{\sqrt{5}}}{1 + \frac{2h}{\sqrt{5}}} - 3}{h} = \lim_{h \to 0} \frac{-7 \not{h}}{(\sqrt{5} + 2h) \not{h}} = -\frac{7}{\sqrt{5}}.$$

Back

8 Day 8: Implicit functions.

Notice that when we talked in (11) about implicit functions, we derived the formula (Implicit Function's Theorem)

$$\text{If } f(x,y) = c \quad \Rightarrow \quad \left.\frac{dy}{dx}\right|_{(a,b)} = -\frac{\partial f/\partial x|_{(a,b)}}{\partial f/\partial y|_{(a,b)}}, \qquad \text{valid when } \partial f/\partial y|_{(a,b)} \neq 0.$$

Question 26. *Consider function $f(x,y) = x^2 - y^2$. Find the equation of the tangent line to its level curve $z = 3$ at point $(2,1)$.* *Answer to question 26.*

Question 27. *Sydsaeter & Hammond* ***Problem 16.3.4.*** *A curve in the xy plane is given by the equation*

$$2x^2 + xy + y^2 - 8 = 0.$$

(a) Find the equation for the tangent line at the point $(2,0)$.

(b) Which points on the curve have a horizontal tangent?

Answer to question 27.

Answers to questions:

A26. The level curve $z = 3$ is the hyperbola $x^2 - y^2 = 3$. Point $(2,1)$ belongs to it because it satisfies the equation. The tangent line will have as a point-slope equation:

$$y - 1 = \left.\frac{dy}{dx}\right|_{(2,1)} \cdot (x - 2)$$

Using the Implicit Function Theorem we have:

$$\frac{dy}{dx} = -\frac{f'_x}{f'_y} = -\frac{2x}{-2y} = \frac{x}{y}. \quad \text{At point } (2,1)\text{:} \quad \left.\frac{dy}{dx}\right|_{(2,1)} = 2.$$

The equation of the tangent is (see (12)):

$$y - 1 = 2 \cdot (x - 2); \qquad \Leftrightarrow \qquad y = 2x - 3.$$

Back

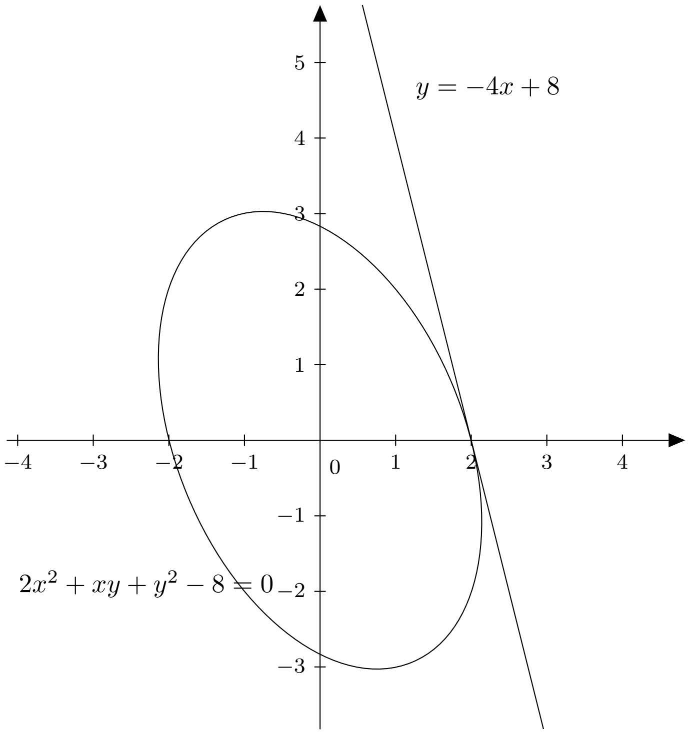
A27. (a) We need to find the first-order derivatives of $f(x,y)$ at the point $(2,0)$:

$$\frac{\partial f}{\partial x} = 4x + y; \qquad \left.\frac{\partial f}{\partial x}\right|_{(2,0)} = 8;$$
$$\frac{\partial f}{\partial y} = x + 2y; \qquad \left.\frac{\partial f}{\partial y}\right|_{(2,0)} = 2.$$

Now the equation of the tangent line is (see (12)):

$$y = -\frac{8}{2}(x-2) \qquad \Leftrightarrow \qquad y = -4x + 8.$$

The graph is:



(b) A horizontal tangent means a tangent vector like $(\cdot, 0)$, that is

$$-\left.\frac{\partial f}{\partial x}\right|_{(a,b)} = 0.$$

In our case, the points (a,b) with a horizontal tangent will be those **on the curve** satisfying $4a + b = 0$. Consequently, they will be the solution of the system

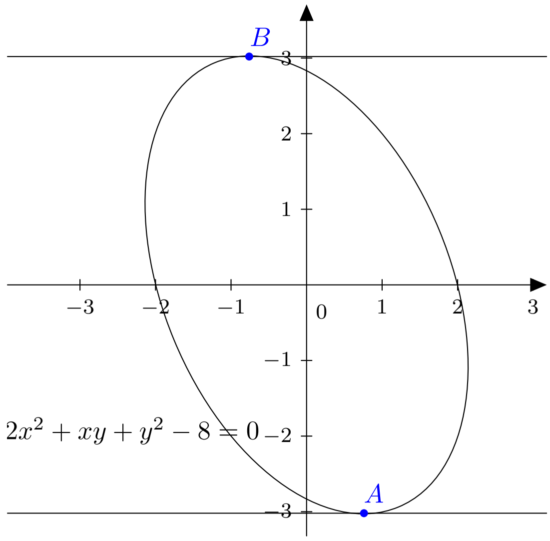
$$\begin{cases} 2a^2 + ab + b^2 - 8 &= 0 \\ 4a + b &= 0 \end{cases}$$

Replacing $b = -4a$ from the second equation in the first equation we get

$$2a^2 + a(-4a) + (-4a)^2 - 8 = 0; \quad a = \pm\sqrt{\frac{4}{7}} \approx \pm 0.76$$

and the points are

$$A(0.76, -3.02); \quad B(-0.76, 3.02).$$



Back

9 Day 9: Homogeneous functions.

A linear function of two variables, $f(x, y) = ax + by$ has a very interesting property: if we multiply our x and y by a constant, the image gets multiplied by the same constant. That is

$$f(\lambda{\cdot}x, \lambda{\cdot}y) = \lambda{\cdot}f(x, y) \qquad \text{since} \quad a(\lambda x) + b(\lambda y) = \lambda{\cdot}(ax + by).$$

Linear functions, are the only ones with this property? We are interested in functions $z = f(x, y)$ such that satisfy, for all values of (x, y) in their domain and any $\lambda \in \mathbb{R}$

$$f(\lambda{\cdot}x, \lambda{\cdot}y) = \lambda{\cdot}f(x, y). \tag{13}$$

These are called **homogeneous functions of degree one**. In a very naive way, any function where replacing x and y by λx and λy can lead to λ factored out as a common factor will do. For instance: $f(x, y) = \dfrac{x^2 + y^2}{x + y}$ is a homogeneous function of degree one:

$$f(\lambda{\cdot}x, \lambda{\cdot}y) = \frac{\lambda^2 x^2 + \lambda^2 y^2}{\lambda x + \lambda y} = \frac{\lambda^2{\cdot}(x^2 + y^2)}{\lambda(\cdot x + y)} = \lambda{\cdot}f(x, y).$$

More generally, a **homogeneous function of degree** k satisfies

$$f(\lambda{\cdot}x, \lambda{\cdot}y) = \lambda^k{\cdot}f(x, y), \tag{14}$$

where k is any real number (so we can talk of negative degrees, 0 degree, and any real value).

The easiest examples of k-degree homogeneous functions are the homogeneous polynomials. That is, polynomials where all the monomials are of degree k.

For instance, $P(x, y) = x^3 + x^2y - 3xy^2 + 5y^3$ is homogeneous of degree 3 and a quadratic form $q(x, y) = ax^2 + 2bxy + cy^2$ is homogeneous of degree 2.

Another very interesting homogeneous function is the Cobb-Douglas function: $f(x, y) = A{\cdot}x^a y^b$. Indeed

$$f(\lambda x, \lambda y) = A{\cdot}\lambda^{a+b} x^a y^b = \lambda^{a+b} f(x, y).$$

Question 28. *Check whether the following functions are homogeneous or not:*

(a) $f(x, y) = \dfrac{x + 3y}{\sqrt{5x^2 - y^2}}$

(b) $f(x, y) = \ln(xy)$

(c) $f(x, y) = e^{x+y}$

Answer to question 28.

9.1 Euler's Theorem.

$$\boxed{f(x,y) \text{ is homogeneous of degree } k \quad \Leftrightarrow \quad x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = k\cdot f(x,y).}$$

For instance, for $P(x,y) = x^3 + x^2y - 3xy^2 + 5y^3$,

$$\begin{aligned} x\frac{\partial P}{\partial x} + y\frac{\partial P}{\partial y} &= \\ &= x\cdot(3x^2 + 2xy - 3y^2) + y\cdot(x^2 - 6xy + 15y^2) = \\ &= 3x^3 + 2x^2y - 3xy^2 + x^2y - 6xy^2 + 15y^3 = \\ &= 3\cdot(x^3 + x^2y - 3xy^2 + 5y^3) = \\ &= 3\cdot P(x,y). \end{aligned}$$

Question 29. *Check whether the functions of question 28 are homogeneous or not by checking the validity of Euler's Theorem. (*Hint*: take the derivatives by hand. Later, use www.wolframalpha.com to check your results. The day of the final exam, you will not have a computer to help you!)* *Answer to question 29.*

Proof of Euler's Theorem.

[⇒]

We are assuming that f is homogeneous of degree k. Thus $f(\lambda x, \lambda y) = \lambda^k\cdot f(x,y)$. Taking derivatives w.r.t. λ on each side

$$\left.\frac{\partial f}{\partial x}\right|_{(\lambda x,\lambda y)}\cdot\frac{d(\lambda x)}{d\lambda} + \left.\frac{\partial f}{\partial y}\right|_{(\lambda x,\lambda y)}\cdot\frac{d(\lambda y)}{d\lambda} = k\lambda^{k-1}\cdot f(x,y).$$

That is

$$\left.\frac{\partial f}{\partial x}\right|_{(\lambda x,\lambda y)}\cdot x + \left.\frac{\partial f}{\partial y}\right|_{(\lambda x,\lambda y)}\cdot y = k\lambda^{k-1}\cdot f(x,y).$$

This last expression is valid for any (x,y) and for any λ. Taking $\lambda = 1$, we have

$$\left.\frac{\partial f}{\partial x}\right|_{(x,y)}\cdot x + \left.\frac{\partial f}{\partial y}\right|_{(x,y)}\cdot y = k\cdot f(x,y).$$

[⇐]

This part of the proof is more elaborate. You can find it in section 16.6 of Sydsaeter & Hammond.

9.2 Returns to scale.

Homogeneous functions are good for the comprehension of what economists call "returns to scale". When considering a production function of some description (for instance, a Cobb-Douglas function $Q(K,L) = A\cdot K^aL^b$), economists say it shows **constant returns to scale** if inputting λ times more of each factor of production (λK and λL) we get the same scale in the production (λQ). That means that our production function is a homogeneous function of degree 1. In the case of the Cobb-Douglas function above, that is true if $a + b = 1$.

9.3 Linear approximations and differential of a function.

At a point, (a,b), where the graph of $f(x,y)$ has a tangent plane:

$$z = f(a,b) + \left.\frac{\partial f}{\partial x}\right|_{(a,b)}\cdot(x-a) + \left.\frac{\partial f}{\partial y}\right|_{(a,b)}\cdot(y-b),$$

the values of the function for nearby points can be approximated using the images calculated on the tangent plane instead of the surface. That is:

$$\boxed{\text{If } (x,y) \text{ is near } (a,b),\ f(x,y) \approx f(a,b) + \left.\frac{\partial f}{\partial x}\right|_{(a,b)}\cdot(x-a) + \left.\frac{\partial f}{\partial y}\right|_{(a,b)}\cdot(y-b).}$$

Question 30. *For $f(x,y) = \ln(x+y)$ find an approximation of $f\left(\frac{e}{2} + 0.1, \frac{e}{2} + 0.2\right)$ using the tangent plane at point $\left(\frac{e}{2}, \frac{e}{2}\right)$. (No calculators).* *Answer to question 30.*

The linear part of the approximation above, as a function of variables $x-a$ and $y-b$, is called the **differential** of f:

$$df_{(a,b)}(x-a, y-b) = \left.\frac{\partial f}{\partial x}\right|_{(a,b)} \cdot (x-a) + \left.\frac{\partial f}{\partial y}\right|_{(a,b)} \cdot (y-b).$$

Instead of $x-a$ and $y-b$ it is usual to use dx and dy, read **differential** of x and **differential** of y. It is a bit confusing, but it is accepted mathematical notation. Thus, the differential of f at point (a,b) is usually written

$$\boxed{df_{(a,b)} = \left.\frac{\partial f}{\partial x}\right|_{(a,b)} dx + \left.\frac{\partial f}{\partial y}\right|_{(a,b)} dy.}$$

If the point is not specified, then we write

$$\boxed{df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.}$$

The approximation through the tangent plane now can be seen as using the differential in order to find the change in the function, $\triangle z = f(x,y) - f(a,b)$:

$$\boxed{\triangle z \approx dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.}$$

Question 31. *Find an approximate value for $V = \sqrt{(4.02)^2 + (2.99)^2}$ using the approximation $\triangle T \approx dT$ for a suitable function $T(x,y)$.*

Answer to question 31.

Answers to questions:

A28. i. $f(\lambda x, \lambda y) = \dfrac{\lambda x + 3\lambda y}{\sqrt{5(\lambda x)^2 - (\lambda y)^2}} = \dfrac{\lambda(x+3y)}{\sqrt{5\lambda^2 x^2 - \lambda^2 y^2}} = \dfrac{\lambda(x+3y)}{\lambda\sqrt{5x^2 - y^2}} = \dfrac{x+3y}{\sqrt{5x^2-y^2}}$ is homogeneous of degree 0.

ii. $f(\lambda x, \lambda y) = \ln(\lambda^2 xy) = \ln(\lambda^2) + \ln(xy)$. This is not of the form $\lambda^k \ln(xy)$. The function is not homogeneous.

iii. $f(\lambda x, \lambda y) = e^{\lambda x + \lambda y} = (e^\lambda)^{x+y}$. Not homogeneous.

Back

A29. i. For $f(x,y) = \dfrac{x+3y}{\sqrt{5x^2-y^2}}$, $f'_x = -\dfrac{y\cdot(15x+y)}{(5x^2-y^2)^{3/2}}$ and $f'_y = \dfrac{x\cdot(15x+y)}{(5x^2-y^2)^{3/2}}$. Then

$$xf'_x + yf'_y = -x\frac{y(15x+y)}{(5x^2-y^2)^{3/2}} + y\frac{x(15x+y)}{(5x^2-y^2)^{3/2}} = 0.$$

This means that our function is homogeneous of degree 0.

ii. For $f(x,y) = \ln(xy)$,

$$xf'_x + yf'_y = x\frac{1}{x} + x\frac{1}{y} = 2 \neq k\cdot\ln(xy) \quad \text{for any } k.$$

The function is not homogeneous.

iii. For $f(x,y) = e^{x+y}$,

$$xf'_x + yf'_y = xe^{x+y} + ye^{x+y} = (x+y)e^{x+y} \neq ke^{x+y} \quad \text{for any } k.$$

The function is not homogeneous.

Back

A30. The partial derivatives of f are

$$\frac{\partial f}{\partial x} = \frac{1}{x+y}; \quad \frac{\partial f}{\partial y} = \frac{1}{x+y}.$$

At point $\left(\frac{e}{2}, \frac{e}{2}\right)$ these derivatives are both $1/e$ and $f\left(\frac{e}{2}, \frac{e}{2}\right) = \ln e = 1$.
Using the tangent plane approximation,

$$f\left(\frac{e}{2} + 0.1, \frac{e}{2} + 0.2\right) \approx 1 + \frac{1}{e}\cdot 0.1 + \frac{1}{e}\cdot 0.2 = 1 + \frac{0.3}{e} = 1.11.$$

Notice that we have obtained an approximate value of $\ln(e+0.3)$ without using a calculator. The real value is $\ln(e+0.3) = 1.15$.

Back

A31. The structure of the value V suggest the use of the function $T(x,y) = \sqrt{x^2+y^2}$. For this function

$$dT = \frac{x}{\sqrt{x^2+y^2}}dx + \frac{y}{\sqrt{x+y}}dy.$$

The point $(4.02, 2.99)$ is very near $(4,3)$ and $T(4,3) = 5$. So $\triangle T = T(4.02, 2.99) - T(4,3)$ can be approximated by dT at $(4,3)$

$$dT = \left.\frac{\partial T}{\partial x}\right|_{(4,3)} dx + \left.\frac{\partial T}{\partial y}\right|_{(4,3)} dy = \frac{4}{5}{\cdot}dx + \frac{3}{5}{\cdot}dy.$$

In our case, $dx = 4.02 - 4 = 0.02$ and $dy = 2.99 - 3 = -0.01$.

$$\triangle T \approx dT = \frac{4}{5}{\cdot}0.02 + \frac{3}{5}{\cdot}(-0.01) = \frac{0.05}{5} = 0.01$$

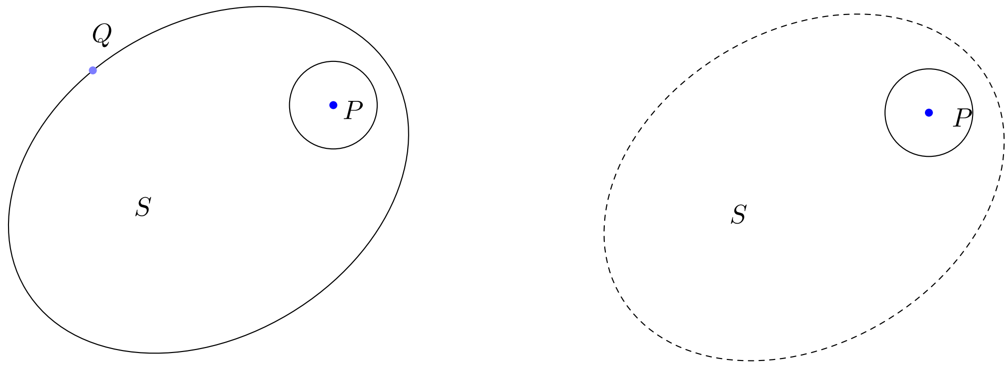
We conclude that

$$\sqrt{(4.02)^2+(2.99)^2} \approx 5 + 0.01 = 5.01.$$

The true value is 5.01004. Not a bad approximation! **Back**

10 Day 10: Local optimization. First-order conditions.

The domain of a two-variable function is a set S contained in the xy plane. An **interior** point of a set S is a point such that there exists a disk centered at the point with a positive radius entirely contained in the set. In the figure below, P is an interior point of S and Q is not.



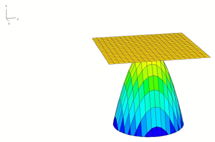
An **open** set is a set where **all its points are interior** to it. Same set S above **without** the border curve is an open set.

We will assume all functions differentiable and defined on open domains.

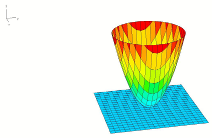
We are interested in finding points of the domain S of a function $f(x,y)$ at which f has a maximum or minimum value. The concept of maximum or minimum for a two-variable function is similar to its one-variable counterpart. We can talk of **local** or **global** max/min.

$f(x,y)$ has a **global** max at (a,b) if for all $(x,y) \in S$, $f(x,y) \le f(a,b)$. If the inequality is $<$, we say that the max is **strict**. Notice the language: the actual max is the value of f at (a,b), $f(a,b)$ but we say that f has a max at (a,b) or even that (a,b) **is** a max.

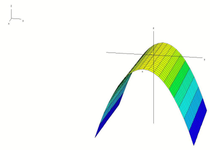
Similar definitions for min, replacing $\le$ with $\ge$.



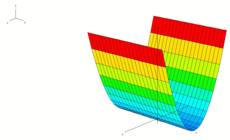
Strict Max



Strict Min



Non-strict Max



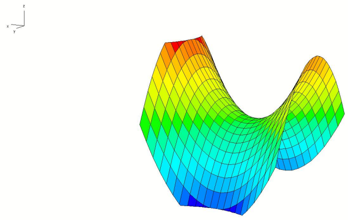
Non-strict-Min

The **local** concept is the same except that it is not required that the inequality be satisfied by all points (x,y) in the domain but only by those in a disk around (a,b).

A necessary condition for $f(x,y)$ to have a max/min in an interior point (a,b) of its domain is

$$\begin{cases} \left.\dfrac{\partial f}{\partial x}\right|_{(a,b)} & = & 0 \\ \left.\dfrac{\partial f}{\partial y}\right|_{(a,b)} & = & 0 \end{cases} \tag{15}$$

The points at which both partial derivatives are 0 are called **stationary** or **critical** points of the function. That is equivalent to saying that the tangent plane at the point is parallel to the xy plane (horizontal). We know from our study of quadratic forms that there are points where the tangent plane may be horizontal and not be a max or a min. These stationary points are called **saddle** points. They look like $(0,0,0)$ for $z = x^2 - y^2$:



In order to find the stationary points of a function $f(x,y)$, we need to solve the system of equations (15). This may be easy (if the system is a linear one) or not so easy (if the system is not linear). In this last case, some amount of ingenuity will be needed to find all the solutions.

As a general advice, these systems have to be solved by substitution: if you can solve one of the equations for x or y and replace in the other equation, you will have a one-variable equation to solve. Mind, this may be a formidable task!

If you cannot use the substitution approach, it is recommendable to try to factorize each equation as much as possible. If you succeed, and say you manage to factor the first equation into a product of two (simpler) equations $= 0$, you only have to solve the new systems formed by each one of these factors with the second equation. The following two questions illustrate this situation.

Question 32. *Find all the stationary points of* $f(x,y) = 8x^3 + \frac{1}{8}y^3 + 6xy$. *Answer to question 32.*

Question 33. *Find all the stationary points of* $f(x,y) = (x^2+y^2)^2 - (x^2+y^2)$. *Answer to question 33.*

Answers to questions:

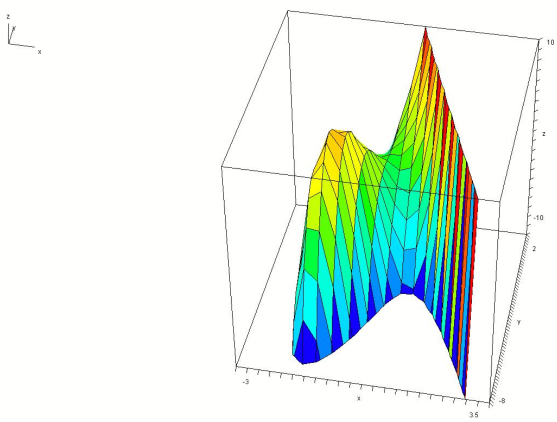
A32. We have to solve

$$\begin{cases} 24x^2 + 6y &= 0 \\ 6x + \frac{3}{8}y^2 &= 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} 4x^2 + y &= 0 \quad (E1) \\ 48x + 3y^2 &= 0 \quad (E2) \end{cases}$$

From equation (E1), $y = -4x^2$. Replacing in (E2), $48x + 48x^4 = 0$ which simplifies to $x + x^4 = 0$. This is a fourth-degree equation in x. It may have 4 solutions. Actually, factoring x out, we have only two real solutions:

$$x{\cdot}(1+x^3) = 0 \quad \Rightarrow \quad \begin{cases} x &= 0 \\ 1 + x^3 &= 0 \quad \Rightarrow \quad x = -1. \end{cases}$$

One solution, $x = 0$, replacing in E1 or E2, leads to $y = 0$ and $x = -1$ gives $y = -4$. We have only two stationary points: $(0,0)$ and $(-1,-4)$.



Back

A33. We have to solve

$$\begin{cases} 4x(x^2+y^2) - 2x &= 0 \\ 4y(x^2+y^2) - 2y &= 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} 2x(x^2+y^2) - x &= 0 \quad (E1) \\ 2y(x^2+y^2) - y &= 0 \quad (E2) \end{cases}$$

From (E1), factoring x, we deduce either $x = 0$ or $x^2 + y^2 = 1/2$.

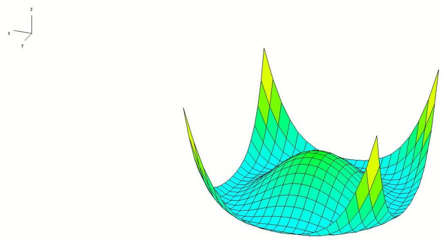
- Case $x = 0$. Replacing in (E2) we have $2y^3 - y = 0$ which leads to $y = 0$ or $2y^2 - 1 = 0$. Solutions $(0, 0), (0, 1/\sqrt{2}), (0, -1/\sqrt{2})$.
- Case $x^2+y^2 = 1/2$. Replacing in (E2) we have $0 = 0$. Consequently, all points in the circle $x^2+y^2 = 1/2$ are solutions. The two points $(0, 1/\sqrt{2}), (0, -1/\sqrt{2})$ belong to that circle.

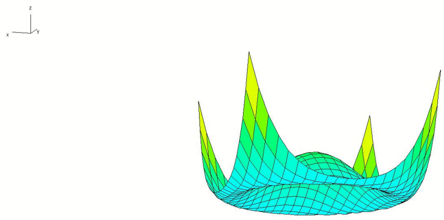
Alternatively, we could have started factoring (E2) from which we deduce either $y = 0$ or $x^2 + y^2 = 1/2$.

- Case $y = 0$. Replacing in (E1) we have $2x^3 - x = 0$ which leads to $x = 0$ or $2x^2 - 1 = 0$. Solutions $(0, 0), (1/\sqrt{2}, 0), (-1/\sqrt{2}, 0)$.
- Case $x^2 + y^2 = 1/2$. Replacing in (E1) we have $0 = 0$. Again we obtain all the points in the circle $x^2 + y^2 = 1/2$ as solutions. And again the two points $(1/\sqrt{2}, 0), (-1/\sqrt{2}, 0)$ belong to the circle.

All in all, the only stationary points are $(0, 0)$ and those (a, b) that satisfy $a^2 + b^2 = 1/2$, which is a circle of center $(0, 0)$ and radius $1/\sqrt{2}$.

Back





11 Day 11: Local optimization. Second-order conditions.

How do we tell a max from a min from a saddle point? A little reflection may help. If we are placed at a strict max, (a, b), we have that $dz(a, b) = 0$ and if we move a small distance in any direction the value of dz decreases. That means that $d(dz) = d^2_{(a,b)}z < 0$. Let us find the value of d^2z.

$$\begin{aligned}
d^2z = d(dz) &= \\
&= d\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right) = \\
&= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right)dx + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right)dy = \\
&= \frac{\partial^2 z}{\partial x^2}dx^2 + \frac{\partial^2 z}{\partial y \partial x}dy\,dx + \frac{\partial^2 z}{\partial x \partial y}dx\,dy + \frac{\partial^2 z}{\partial y^2}dy^2.
\end{aligned}$$

Using Young's Theorem we have

$$d^2z = \frac{\partial^2 z}{\partial x^2}dx^2 + 2\frac{\partial z}{\partial y \partial x}dx\,dy + \frac{\partial^2 z}{\partial y^2}dy^2.$$

which, once evaluated at (a, b), is a quadratic form in the variables dx, dy. The matrix of this quadratic form is

$$H(a,b) = \begin{pmatrix} \left.\frac{\partial^2 z}{\partial x^2}\right|_{(a,b)} & \left.\frac{\partial^2 z}{\partial y \partial x}\right|_{(a,b)} \\ \left.\frac{\partial^2 z}{\partial y \partial x}\right|_{(a,b)} & \left.\frac{\partial^2 z}{\partial y^2}\right|_{(a,b)} \end{pmatrix},$$

the hessian matrix of our function f!

Thus, the condition $d^2_{(a,b)}z < 0$ is equivalent to saying that the quadratic form above is Negative Definite, which is equivalent to

$$|H(a,b)| = \begin{vmatrix} \left.\frac{\partial^2 z}{\partial x^2}\right|_{(a,b)} & \left.\frac{\partial^2 z}{\partial y \partial x}\right|_{(a,b)} \\ \left.\frac{\partial^2 z}{\partial y \partial x}\right|_{(a,b)} & \left.\frac{\partial^2 z}{\partial y^2}\right|_{(a,b)} \end{vmatrix} > 0 \quad \text{and} \quad \left.\frac{\partial^2 z}{\partial x^2}\right|_{(a,b)} < 0.$$

If (a, b) is a strict min, then $d^2_{(a,b)}z > 0$ and the quadratic form is Positive Definite. When at (a, b) we have a saddle point, $d^2_{(a,b)}z$ will be sometimes positive and sometimes negative. The hessian will be an indefinite quadratic form.

11.1 Summary.

Maxima, minima and saddle points are amongst the stationary points of $f(x,y)$. Thus, in order to find them we need to solve the system of equations:

$$\begin{cases} \left.\dfrac{\partial f}{\partial x}\right|_{(a,b)} & = & 0 \\ \left.\dfrac{\partial f}{\partial y}\right|_{(a,b)} & = & 0 \end{cases}$$

If (a,b) is one of the solutions, in order to know whether it is a max, a min or a saddle point, we need to classify the quadratic form

$$d^2z_{(a,b)}(dx,dy) = \left.\frac{\partial^2 z}{\partial x^2}\right|_{(a,b)} dx^2 + 2\left.\frac{\partial^2 z}{\partial y \partial x}\right|_{(a,b)} dx\,dy + \left.\frac{\partial^2 z}{\partial y^2}\right|_{(a,b)} dy^2.$$

Then,

If $\lvert H(a,b)\rvert > 0$	and $\left.\dfrac{\partial^2 f}{\partial x^2}\right\|_{(a,b)} < 0$	LOCAL MAX
	and $\left.\dfrac{\partial^2 f}{\partial x^2}\right\|_{(a,b)} > 0$	LOCAL MIN
If $\lvert H(a,b)\rvert < 0$	SADDLE POINT	
If $\lvert H(a,b)\rvert = 0$	?	

The case $|H(a,b)| = 0$ could lead to a local max or min or a saddle point. It needs a special treatment for each problem and this can be sometimes difficult. Consider the three functions:

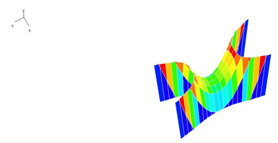
a) $f(x,y) = x^2 + y^4$; b) $g(x,y) = x^2 - y^4$; c) $h(x,y) = x^3 - 2xy^2$ (Monkey Saddle).

The three have $(0,0)$ as a stationary point and for the three we have

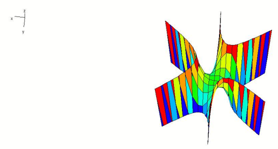
$$|H(0,0)| = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = 0.$$

At $(0,0)$ function f in a) has a min and both g in b) and h in c) have saddle points. The function in c) is called the Monkey Saddle because it has three "slopes" down from $(0,0)$, two for the legs of the monkey and one for the tail.

a) Function $f(x,y) = x^2 + y^4$;



b) Function $f(x,y) = x^2 - y^4$;



c) Function $f(x,y) = x^3 - 2xy^2$

Question 34. *Classify the stationary points of question 32.* *Answer to question 34.*

Question 35. *Find and classify the stationary points of* $f(x,y) = x^4 + x^2 - 6xy + 3y^2$. *Answer to question 35.*

Question 36. *Classify the stationary points of question 33.* *Answer to question 36.*

Answers to questions:

A34. The second-order derivatives are:

$$\frac{\partial^2 z}{\partial x^2} = 48x; \qquad \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y} = 6; \qquad \frac{\partial^2 z}{\partial y^2} = \frac{3}{4}y.$$

Thus the hessian is

$$H(x,y) = \begin{pmatrix} 48x & 6 \\ 6 & \frac{3}{4}y \end{pmatrix}$$

- Point $(0,0)$.

$$|H(0,0)| = \begin{vmatrix} 0 & 6 \\ 6 & 0 \end{vmatrix} < 0 \quad \Rightarrow \quad \text{SADDLE POINT.}$$

- Point $(-1,-4)$.

$$|H(-1,-4)| = \begin{vmatrix} -48 & 6 \\ 6 & -3 \end{vmatrix} > 0 \quad \text{and} \quad \left.\frac{\partial^2 z}{\partial x^2}\right|_{(-1,-4)} = -48 < 0 \quad \Rightarrow \quad \text{LOCAL MAX.}$$

Back

A35. For $f(x,y) = x^4 + x^2 - 6xy + 3y^2$ the first-order conditions are

$$\begin{cases} 4x^3 + 2x - 6y &= 0 \quad \text{E1} \\ -6x + 6y &= 0 \quad \text{E2} \end{cases}$$

From E2, $x = y$. Replacing in E1 we get $4x^3 - 4x = 0$ which, factored, leads to $x(x^2-1) = 0$. Three solutions for x:

$$\begin{aligned} x = 0 \quad &\Rightarrow \quad y = 0; \quad \text{candidate: } (0,0). \\ x = 1 \quad &\Rightarrow \quad y = 1; \quad \text{candidate: } (1,1). \\ x = -1 \quad &\Rightarrow \quad y = -1; \quad \text{candidate: } (-1,-1). \end{aligned}$$

The Hessian is

$$H(x,y) = \begin{pmatrix} 12x^2 + 2 & -6 \\ -6 & 6 \end{pmatrix}$$

- **Point** $(0,0)$:

$$|H(0,0)| = \begin{vmatrix} 2 & -6 \\ -6 & 6 \end{vmatrix} = -24 < 0.$$

  At $(0,0)$ we have a SADDLE POINT.
- **Point** $(1,1)$:

$$|H(1,1)| = \begin{vmatrix} 14 & -6 \\ -6 & 6 \end{vmatrix} = 48 > 0 \quad \text{and} \quad f''_{xx}(1,1) = 14 > 0$$

  At $(1,1)$ we have a MIN.
- **Point** $(-1,-1)$:

$$|H(-1,-1)| = \begin{vmatrix} 14 & -6 \\ -6 & 6 \end{vmatrix} = 48 > 0 \quad \text{and} \quad f''_{xx}(-1,-1) = 14 > 0$$

  At $(-1,-1)$ we have a MIN.

Back

A36. The second-order derivatives are:

$$\frac{\partial^2 z}{\partial x^2} = 12x^2 + 4y^2 - 2; \qquad \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y} = 8xy; \qquad \frac{\partial^2 z}{\partial y^2} = 12y^2 + 4x^2 - 2.$$

The hessian is

$$H(x,y) = \begin{pmatrix} 12x^2 + 4y^2 - 2 & 8xy \\ 8xy & 12y^2 + 4x^2 - 2 \end{pmatrix}$$

- Point $(0,0)$.

$$|H(0,0)| = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} > 0 \quad \text{and} \quad \left.\frac{\partial^2 z}{\partial x^2}\right|_{(0,0)} = -2 < 0 \quad \Rightarrow \quad \text{LOCAL MAX.}$$

- Points (a,b) such that $a^2 + b^2 = 1/2$. We can write $b = \pm\sqrt{1/2 - a^2}$ in order to replace in the hessian

$$|H(a,b)| = \begin{vmatrix} 8a^2 & \pm 4\sqrt{2} \cdot a\sqrt{1-2a^2} \\ \pm 4\sqrt{2} \cdot a\sqrt{1-2a^2} & 4(1-2a^2) \end{vmatrix} = 0 \qquad \Rightarrow \qquad ?$$[5]

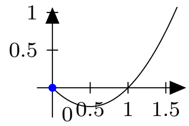
[5] Another way of getting to the same place is to write the Hessian in a slightly different way:

$$H(x,y) = \begin{pmatrix} 8x^2 + 4x^2 + 4y^2 - 2 & 8xy \\ 8xy & 8y^2 + 4y^2 + 4x^2 - 2 \end{pmatrix}$$

and now replacing $x = a, y = b$, and $a^2 + b^2 = 1/2$:

$$|H(a,b)| = \begin{vmatrix} 8a^2 & 8ab \\ 8ab & 8b^2 \end{vmatrix} = 0.$$

The graph of our function (at the end of Example 2) shows that these points are min. Can we find a way to prove it? A look at the function shows that if we replace $x^2 + y^2 = t$ in its definition and we have the one-variable function: $g(t) = t^2 - t$ for $t \geq 0$:



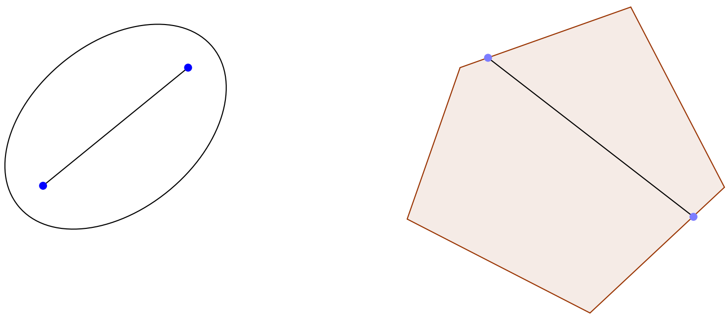
This has global min at $t = 1/2$, which corresponds to the values (a, b) such that $a^2 + b^2 = 1/2$ and a local max at $t = 0$, which correspond to $x^2 + y^2 = 0$, that is $x = 0, y = 0$. Consequently, f has a min$= -1/4$ at each (a, b) such that $a^2 + b^2 = 1/2$.

Back

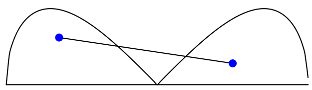
12 Day 12: Convex and concave functions.

12.1 Convex sets in $\mathbb{R}^2$.

A convex set S in $\mathbb{R}^2$ is a set that has "no dents". This means that the segment joining any two points in S is completely within S.



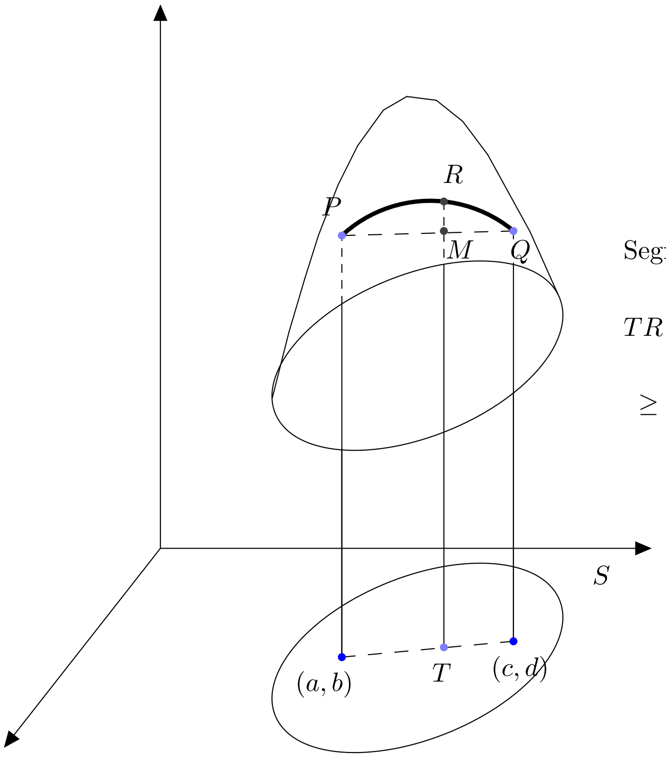
Convex sets



Non-convex set

12.2 Concave/convex functions.

Definition 1. Let S be a convex set in $\mathbb{R}^2$. We say the continuous function $f(x, y)$ defined on S is **concave** (**convex**) on S if the segment joining any two points in S is **never above** (**never below**) the graph of f.



Segment PQ not above PRQ

$TR = f((a,b) + t\cdot(c-a, d-b)) \geq$

$\geq TM = f(a,b) + t\cdot(f(c,d) - f(a,b))$

If the segment PQ is **always below** PRQ, we say the function is **strictly concave**. (Similarly for **strictly convex**)

There are many more formal ways of defining a convex/concave function, but the one above is the easiest to understand. It relies on the visual aid of the graph of the function and, consequently, it will not be useful for functions of more than 2 variables.

The easy examples are $f(x,y) = x^2 + y^2$: strictly convex on $\mathbb{R}^2$. $f(x,y) = -x^2 - y^2$: strictly concave on $\mathbb{R}^2$. $f(x,y) = x^2$; convex on $\mathbb{R}^2$. $f(x,y) = ax + by$: concave and convex at the same time. $f(x,y) = x^2 - y^2$: neither concave nor convex.

A more useful definition, valid for any number of variables, is the following:

Definition 2. Let S be a convex set in $\mathbb{R}^2$. We say that the continuous function $f(x,y)$ defined on S is **concave** [resp. **convex**] on S if

$$f((a,b) + t\cdot(c-a, d-b)) \geq f(a,b) + t\cdot(f(c,d) - f(a,b)) \qquad \text{for all } (a,b),(c,d) \in S, \text{ and } t \in [0,1]. \qquad [\text{resp. } \leq].$$

If $f(x,y)$ is differentiable (which is the case we are interested in), we can use the tangent plane to define concavity/convexity:

Definition 3. Let S be an open convex set in $\mathbb{R}^2$. We say that the differentiable function $f(x,y)$ defined on S is **concave** [resp. **convex**] on S if at any point $(a,b) \in S$, the tangent plane is above or on the graph of f. This means that for any $(a,b) \in S$

$$f(x,y) \leq \underbrace{f(a,b) + f'_x(a,b)(x-a) + f'_y(a,b)(y-b)}_{\text{image of } (x,y) \text{ using the tangent plane}} \qquad \text{for any } (x,y) \in S \qquad [\text{resp. } \geq].$$

In case the inequality above is strict for all $(x,y) \neq (a,b)$, we say that f is **strictly concave** [resp. **strictly convex**].

12.3 Second-order conditions for concavity/convexity

If our function is twice differentiable, we can use its hessian to determine its concavity and convexity on an open convex set S. The idea is that if f is concave, at **any** point $(a,b) \in S$ there will be a neighbourhood of (a,b) where the difference

$$D = f(x,y) - f(a,b) - f'_x(a,b)(x-a) - f'_y(a,b)(y-b) \leq 0$$

will have a max (= 0) at (a, b) and consequently its d^2 at (a, b) will be negative (definite or semidefinite). But $d^2_{(a,b)}$ of the function D above is exactly equal to $d^2_{(a,b)}f = H(a, b)$, the hessian of f. As this happens at **any** point in S, we can state

Let $f(x, y)$ be twice differentiable on an open convex set $S \subset \mathbb{R}^2$. Then

$$\begin{aligned} &f \text{ convex on } S \Leftrightarrow H(x,y) \text{ Positive (D or S) in } S \Leftrightarrow \begin{vmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{vmatrix} \geq 0;\ f''_{xx} \geq 0, f''_{yy} \geq 0 \text{ for all } (x,y) \in S \\ &f \text{ concave on } S \Leftrightarrow H(x,y) \text{ Negative (D or S) in } S \Leftrightarrow \begin{vmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{vmatrix} \geq 0;\ f''_{xx} \leq 0, f''_{yy} \leq 0 \text{ for all } (x,y) \in S \\ &f \text{ NCNC on } S \Leftrightarrow H(x,y) \text{ Indefinite in } S \Leftrightarrow \begin{vmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{vmatrix} < 0 \quad \text{for some } (x,y) \in S. \end{aligned} \tag{16}$$

NCNC=non-concave, non-convex.

Notice that the signs of the Hessians in the last result have to be maintained in **ALL OF THE SET** S.

As for strict concavity or convexity, we can use the above result in one sense:[6]

If $H(x, y)$ is Definite in S, the concavity/convexity will be strict.

For many functions, $H(x, y)$ will not be always ≥ 0 or ≤ 0 at all points on the domain. In some occasions one can determine the greatest convex set where f is concave (or convex) by imposing the necessary conditions on the Hessian in order to have the right signs.

Question 37. *Given $f(x, y) = x^3 + y^2 - 6xy$, determine the largest convex set S on which f is concave/convex.*

Answer to question 37.

12.4 Useful concavity/convexity conditions.

Besides the Hessian conditions, there are other ways of determining whether a function is concave or convex (or non-convex, non-concave). The following results are useful. Let f and g be defined on a convex set S:

(a) If f is convex [concave] then $-f$ is concave [convex].

(b) A linear function, $f(x, y) = ax + by$ is both convex and concave for any $a, b \in \mathbb{R}$.

(c) f and g concave [convex] $\Rightarrow f + g$ concave [convex]. **Example**: $f(x, y) = x^2 + y^2 + x + y$ is convex because it is the sum of $x^2 + y^2$ and $x + y$, both convex. **Remark**: this is true for the addition of two functions, not for the subtraction nor the product! xy is the product of two concave (and convex) functions but it is neither!

(d) f concave [convex] $\Rightarrow af, (a > 0)$ concave [convex]. **Example**: $f(x, y) = 3x^2 + 3y^2$ is convex.

(e) If $f(x, y) = F(x)$ or $F(y)$ and $F(\cdot)$ is concave [convex], then $f(x, y)$ is concave [convex]. **Example**: $f(x, y) = x^4 + y^2$ is convex because it is the sum of x^4 and y^2 which, as functions of one variable, are convex.

(f) If $f(x, y)$ is concave [convex] and $F(t)$ is *increasing* and concave [convex], then $F(f(x, y))$ is concave [convex]. **Example**: e^{x^2-y} is convex because e^t is an increasing convex function of one variable and $f(x, y) = x^2 - y$ is a convex function of two variables. **Remark**: notice that in both cases, concave and convex, $F(t)$, as a one-variable function, has to be *increasing*.

Question 38. *Use, if possible, item (f) above to determine the concavity/convexity of $g(x, y) = e^{-x^2-y^2}$.* *Answer to question 38.*

12.5 A sufficient condition for global max/min.

Imagine we know that $f(x, y)$ is concave on all of its open convex domain S and (a, b) is a stationary point in S. Can we decide right away whether (a, b) is a max, a min or a saddle point? Well if the function is concave on all of S, (a, b) **cannot** be a min nor a saddle point. Consequently it has to be a global max!

[6]Remember that there are cases for which the concavity/convexity is strict and the Hessian is **0**. For instance, $f(x, y) = x^4 + y^4$.

Theorem (Sufficient conditions for a global max/min)

Let S be an open convex domain, and let $(a, b) \in S$ be a stationary point for $f(x, y)$. Then

- If f is concave on all of S, (a, b) is a global max for f over S.
- If f is convex on all of S, (a, b) is a global min for f over S.

As the previous Theorem shows, the concavity or convexity of f on an open convex domain determines immediately the **GLOBAL** character of any stationary point in the domain.

Example. Find max/min for $f(x, y) = x^4 + y^4$.

This is one of the functions where the Hessian at the ONLY stationary point, $(0, 0)$, was equal to 0 and we could not tell whether the candidate was a max/min/Saddle Point. But now, we can determine the convexity of $x^4 + y^4$ in all of its domain $\mathbb{R}^2$. This can be done in two ways:

- Using 12.4 above: By (e) x^4 and y^4 as one-variable functions are convex, and now, by (c) $x^4 + y^4$ is convex as sum of two convex functions.
- Using (16): $H(x, y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{pmatrix}$. We have

$$|H(x, y)| = 144x^2y^2 \geq 0; \quad f''_{xx} = 12x^2 \geq 0; \quad f''_{yy} = 12y^2 \geq 0 \qquad \textbf{for all } (x, y) \in \mathbb{R}^2.$$

  Therefore, f is convex on all of its domain.

Conclusion: As $(0, 0)$ is the only stationary point of a convex function on all of its open and convex domain ($\mathbb{R}^2$), the function has a global minimum equal to 0 at $(0, 0)$. It has no max because $x^4 + y^4 \to \infty$ as x (or y or both) $\to \infty$.

Answers to questions:

A37. Let us find the Hessian: $f'_x = 3x^2 - 6y$; $f'_y = 2y - 6x$. So

$$H(x, y) = \begin{pmatrix} 6x & -6 \\ -6 & 2 \end{pmatrix} \quad \text{and} \quad |H(x, y)| = \begin{vmatrix} 6x & -6 \\ -6 & 2 \end{vmatrix} = 12x - 36.$$

Now, if $12x - 36 \geq 0$, that is $x \geq 3$ then $|H(x, y)| \geq 0$ and $6x > 0$. This implies that $H(x, y)$ is positive on the half-plane $S = \{(x, y) : x \geq 3\}$. This is the largest set where f is convex. It is never concave because $f''_{yy} = 2 > 0$ for all (x, y).
Back

A38. Let $F(t) = e^t$. It is increasing and convex. Now $g(x, y) = F(f(x, y))$ where $f(x, y) = -x^2 - y^2$. But $f(x, y) = -x^2 - y^2$ is concave. We cannot use item (f) to determine conc/conv. One could think of using $F(t) = e^{-t}$. Then $g(x, y) = F(f(x, y))$ with $f(x, y) = x^2 + y^2$, which is convex. Unfortunately, $F(t) = e^{-t}$ is decreasing and again (f) fails.
Back

13 Day 13: Global Optimization.

The aim of the rest of the course will be to find **GLOBAL max and min** of a function on a set S. In order to achieve that goal, we are going to study different strategies.

- *Concavity/convexity of f on S.* We have already seen how to use this powerful tool.
- *The Extreme Value Theorem (EVT).* A useful tool when set S is a compact set (see below).
- *Finding max/min on the frontier of the domain.* In combination of our previous tools, a way of establishing possible global max/min at points which are not stationary. We will later use a technique, *Lagrange's multipliers*, that will also help us finding max/min on a curve (which may be the frontier of a set).
- *Using SOLVER.* SOLVER is an application of EXCEL (actually a macro) that helps us finding max/min of any function on any set. We can also use other software (WolframAlpha, for instance), but SOLVER has some special features that make it very useful for big problems (many variables, many constraints).
- *Linear Programming.* Special techniques to apply when our function is linear, $f(x, y) = ax + by$ and the domain is defined by linear constraints.

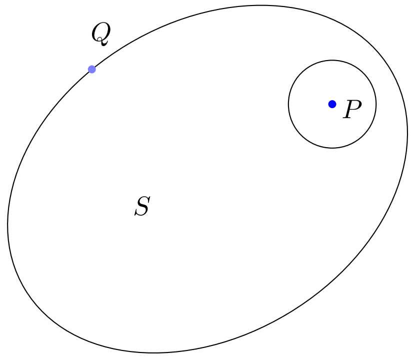
13.1 Compact sets in $\mathbb{R}^2$.

You surely remember interior points (see notes for Day 10), like point P in the figure below. Points like Q that are not interior points and may or not belong to S, are called **frontier** or **boundary** points of S. For a boundary point, any disk centered in the point has points of S and points that do not belong to S.

A set in $\mathbb{R}^2$ that contains all its **boundary** points is called a **closed** set.

A set in $\mathbb{R}^2$ that can be enclosed in a disk of center $(0,0)$ and some finite radius, r, is called a **bounded** set.

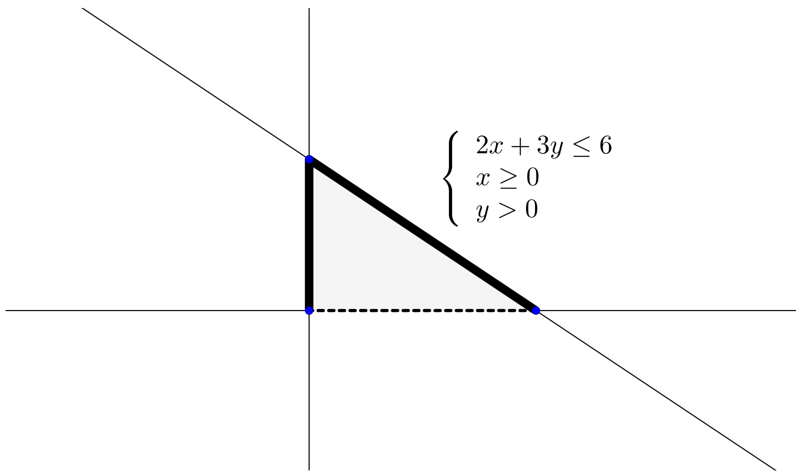
A set in $\mathbb{R}^2$ that is simultaneously **closed and bounded** is called a **compact**.



13.2 Sets defined by constraints.

Most of the sets we use are defined by conditions that have to be satisfied by their points. These conditions will usually be of the form $g(x,y) = c$ or $g(x,y) \leq (\geq, <, >)c$: equalities or inequalities involving a function. Each one these constraints will delimit a part of the plane. The combination of all, will define our set.

For example, the set $S = \{(x,y) \in \mathbb{R}^2 : 2x + 3y \leq 6, x \geq 0, y > 0\}$ is the following



If g is a continuous function, the set defined by $S = \{(x,y) : g(x,y) = 0\}$ is a curve in the xy-plane. All its points are boundary points and belong to S: S is a closed set. If the curve is bounded, S is a compact. If the curve has any branch going to infinity (like a parabola), then S is not a compact, obviously.

The set defined by $g(x,y) \geq (\leq) 0$ is a closed set (on one "side" of the curve plus the curve itself), and if the inequality is strict, an open set (not including the curve). Again, if the "side" considered is bounded, we have a compact set.

13.3 The Extreme-Value Theorem (EVT)

> If $f(x,y)$ is a continuous function defined on a compact set S of $\mathbb{R}^2$, then $f(x,y)$ has a global max and a global min at points of S.

This is an extension of the same result for one-variable functions defined on a closed interval of $\mathbb{R}$.

The result is completely clear from a naive point of view. It requires no explanation.

13.4 Global optimization on a compact set

If we are interested in max/min of a differentiable function $f(x,y)$ defined on a compact set S, we can proceed as follows:

(a) Find all stationary points of f which are interior to S.

(b) Find the max/min of f on the frontier.

(c) Evaluate f at each one of the points found before.

(d) Max/min of f are the greatest and the smallest values in the list just made above.

13.5 Problems

SH problem 17.2.1 *For the following functions find the global max/min (or prove it does not exist) by a direct argument:*

- $f(x, y) = (x-1)^2 + (y+2)^2 - 10$.
- $f(x, y) = 3 - \sqrt{2 - (x^2 + y^2)}$.

14 Day 14: Problems

SH problem 17.3.1 *Find the max/min for* $f(x, y) = 4x - 2x^2 - 2y^2$ *on* $S = \{(x, y) : x^2 + y^2 \leq 25\}$.

SH problem 17.3.2 c *Max/min* $f(x, y) = 3 + x^3 - x^2 - y^2$ *s.t.* $x^2 + y^2 \leq 1, x \geq 0$.

Problem. *Max/min of* $f(x, y) = 1 + xy - x - y$ *on* $S = \{(x, y) : y \geq x^2, y \leq 4\}$.

Problem. *If a differentiable function of one variable on an interval has only one stationary point, then a local max has to be a global max. But this is not true for functions of two variables. Show that the function* $f(x, y) = 3xe^y - x^3 - e^{3y}$ *has exactly one stationary point which is a local max that is not a global max. How do you reconcile this with Theorem 12.5 (Sufficient conditions for a global max/min)?*

Problem. *A monopolist produces two commodities, with demand functions* $p_1 = 12 - x_1; p_2 = 36 - 5x_2$, *where* x_1, x_2 *are the quantities produced of each commodity, and* p_1, p_2 *the corresponding prices per unit. If the cost function is* $C(x_1, x_2) = 2x_1x_2 + 15$, *find the production of each commodity and the prices that maximize the monopolist's profit. Consider the domain where* $x_1, x_2, p_1, p_2 \geq$.

15 Day 15: Constrained optimization. Lagrange's method

Constrained optimization refers to the problem of finding max/min of a function $f(x, y)$ where x, y are constrained by having to satisfy a given equation, $g(x, y) = c$, called **the constraint**. When we find the max/min of a function defined on a compact domain S, this is exactly what we do when we find candidates on the frontier of the domain, ∂S.

15.1 Substitution method

In order to find max/min of a function $f(x, y)$ subject to a constraint $g(x, y) = c$, we can solve $g(x, y) = c$ for x or y and replace in $f(x, y)$. The problem becomes a one-variable problem. One has to be careful about the domain of the new single variable function. This method is only possible if the solution of $g(x, y) = 0$ can be found in terms of one of the variables.

Example. Find max/min of $f(x, y) = xy$ s.t. $x + y = 6$.

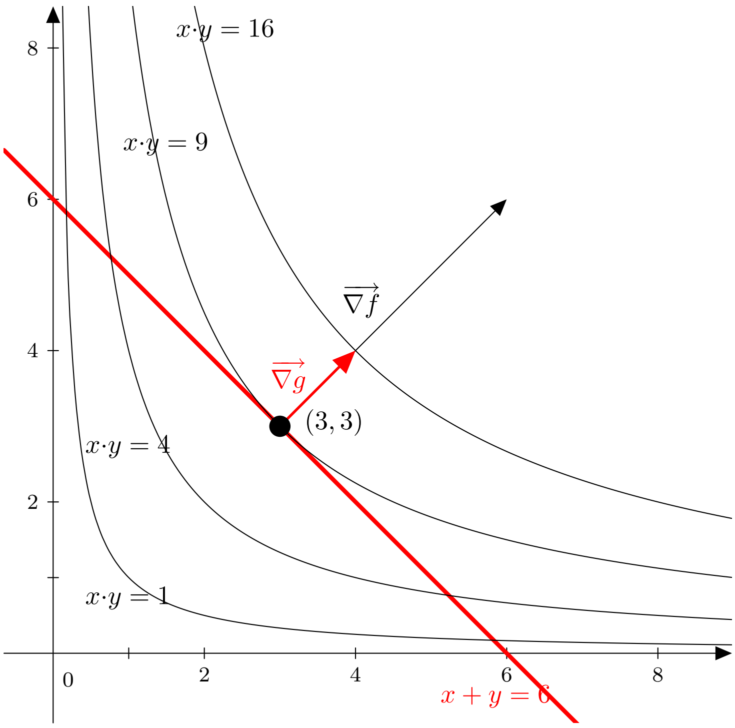
The constraint allows us to write $y = 6 - x$ and the function becomes $g(x) = f(x, 6 - x) = x(6 - x)$ with $\mathbb{R}$ as domain. Function g has $x = 3$ as a unique stationary point which is a global max since $g(x) = -x^2 + 6x$ is a concave parabola.

The solution to our constrained problem is max $f(x, y) = 9$ at $(3, 3)$.

In case that substitution is not easy, we may use:

15.2 Lagrange's method

Let us examine the previous problem from another point of view. If we graph a few of $f(x, y) = xy$ level curves together with the constraint, we will have a map of the situation:



Imagine the graph of $f(x, y)$ is a mountain and the red line is the general direction of a path on the mountain. As we move from point $(0, 6)$ towards point $(6, 0)$ **on the line** $x + y = 6$, we cross many level curves. In the figure, we only see two of those, $xy = 1$ and $xy = 4$. Clearly, there are points on the red line which are at a greater "height" on the surface than 1 or 4. But when we draw the level curve $xy = 9$ we see that our path "touches" (is tangent) the level curve $xy = 9$ at point $(3, 3)$ and, consequently, at this point we reach the max altitude on the path: 9. Continuing we descend again towards $xy = 4$, $xy = 1$, and eventually, $xy = 0$ (the axes).

What do we learn from that rationale we have just used? That at a max on the path, there must be a level curve of our function, $f(x, y) = b$ which is tangent to the given constraint, $g(x, y) = c$. Consequently, at this point both gradients have to have the same direction! (See the two vectors with origin $(3, 3)$):

$$\overrightarrow{\nabla} f(a, b) = \lambda \overrightarrow{\nabla} g(a, b) \qquad \text{where } (a, b) \text{ is the max (or min) point, and } \lambda \in \mathbb{R}.$$

This translates into

$$f'_x(a, b) = \lambda g'_x(a, b), \text{ and } f'_y(a, b) = \lambda g'_y(a, b)$$

or, even better,

$$f'_x(a, b) - \lambda g'_x(a, b) = 0 \text{ and } f'_y(a, b) - \lambda g'_y(a, b) = 0.$$

This suggests a method to solve a constrained problem: **Lagrange's method**.

(a) Write down the Lagrangian (an auxiliary function):

$$\mathcal{L}(x, y) = f(x, y) - \lambda{\cdot}(g(x, y) - c).$$

(b) Find the solutions (x, y and λ) to the system

$$\begin{cases} \mathcal{L}'_x(x, y, \lambda) &= 0 \\ \mathcal{L}'_y(x, y, \lambda) &= 0 \\ g(x, y) &= c \end{cases} \tag{17}$$

(c) These points are **candidates** to max/min (in principle, local).

15.3 Classifying Lagrange's method candidates

It is a grave error to assume that a candidate to $\max/\min f(x, y)$ s.t. $g(x, y) = c$ found by the Lagrangian can be classified directly by f's Hessian!

There are different ways of deciding whether the candidates found are max or min, as well as deciding whether they are local or global. There are examples, though, where all these ways fail. Each particular problem has to be considered.

One possibility is the Extreme-Value Theorem. If we know that $g(x, y) = c$ represents a closed and bounded curve (for example a circle or an ellipse or a polygon), then we only need to list the value of f at each candidate to find the global max and the global min. **Mind the corner points! These will be points where g is not differentiable or where g's graph has an endpoint.** They have to be included always in the list. In our problem above, there were two corner points we did not mention: $(0, 2)$ and $(2, 0)$. At these points $f(x, y) = xy$ has two local max $= 0$.

If substitution (see Section 15.1) is possible, with a careful control of the domain, we can decide on the character of our candidates.

If this cannot be done, we can try representing the level curves of f alongside with $g(x, y) = c$ and see if we can decide graphically the character (max/min, global/local) of each of the candidates.

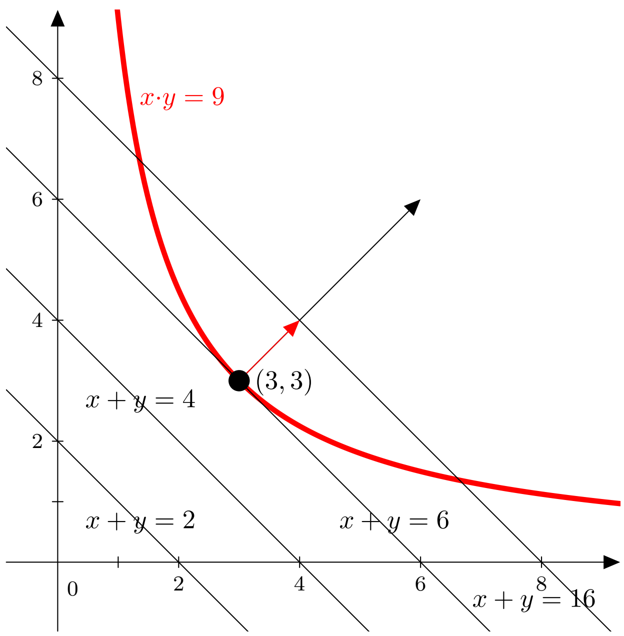
Question 39. *Find the min of $f(x, y) = x + y$ subject to $xy = 9$.* *Answer to question 39.*

Answers to questions:

A39. If we use substitution, the problem is simple enough. Let us use Lagrange. Let $\mathcal{L}(x, y) = x + y - \lambda\cdot(xy - 9)$. We need to find the solutions (x, y and λ) to the system

$$\begin{cases} 1 - \lambda y &= 0 \\ 1 - \lambda x &= 0 \\ xy &= 9 \end{cases} \quad \Rightarrow \quad x = y = 3; \lambda = 1/3.$$

We have only one candidate, $(3, 3)$ with $\lambda = 1/3$.
The domain of $f(x, y)$ is the whole of $\mathbb{R}^2$, which is not a compact. We cannot use EVT. But $f(x, y) = x + y$ is concave and convex on $\mathbb{R}^2$ (it is a linear function!) Thus $f(3, 3) = 6$ may be a min or a max. Let us graph the situation:



We can see clearly that along the curve $xy = 9$, we diminish the value of $x + y$ until we reach the tangency point $(3, 3)$, after which we increase again the value of $x + y$. We have a min at the tangency point. Back

16 Day 16: Global sufficiency in Lagrange problems

A good method to establish the global character of a max/min is to consider the concave/convex character of the Lagrangian (for each candidate) in all of its domain. This has to be done for the Lagrangian that corresponds to each candidate: $\mathcal{L}_{\lambda_0}$.

If (a, b, λ_0) is a candidate (solution to (17)) and we know that it is a global max for $\mathcal{L}_{\lambda_0}$ then it will also be a global max for f restricted on $g(x, y) = c$ since, on the constraint curve, all (x, y) satisfy

$$f(x, y) = f(x, y) - \lambda_0\cdot\underbrace{(g(x, y) - c)}_{=0} = \mathcal{L}_{\lambda_0}(x, y) \leq \mathcal{L}_{\lambda_0}(a, b) = f(a, b) - \lambda_0\cdot\underbrace{(g(a, b) - c)}_{=0} = f(a, b).$$

This allows us to state:

SH Theorem 18.2 (Global sufficiency)

If (a, b, λ_0) is a candidate of our problem (solution to (17)) and $\mathcal{L}_{\lambda_0}(x, y) = f(x, y) - \lambda_0(g(x, y) - c)$ is concave (convex) then (a, b) is a global max (min).

Notice that $\mathcal{L}_{\lambda_0}$'s concavity/convexity can be studied by f's and $-\lambda_0{\cdot}g$'s concavity/convexity.

16.1 Local sufficiency. Bordered Hessian.

We have already mentioned the local character of the candidates obtained by Lagrange's method. We have also highlighted that we cannot use f's Hessian to classify the candidates.

But we can try to use the local tools we have. We can study the sign of $d^2\mathcal{L}(a, b, \lambda_0)$ as a quadratic form constrained by $g(x, y) = c$ and, with its help, decide whether we have a max or a min. This is done in a neighbourhood of point (a, b), therefore the result is only local.

As $g(x, y) = c$, $dg(a, b) = g'_x(a, b)\cdot dx + g'_y(a, b)\cdot dy = 0$. The character (sign) of $d^2\mathcal{L}(a, b)$ on the line of equation $g'_x(a, b)\cdot dx + g'_y(a, b)\cdot dy = 0$ can be found by the sign of the bordered hessian:

$$D(a, b, \lambda_0) = \begin{vmatrix} 0 & g'_x(a, b) & g'_y(a, b) \\ g'_x(a, b) & \mathcal{L}''_{xx}(a, b, \lambda_0) & \mathcal{L}''_{xy}(a, b, \lambda_0) \\ g'_y(a, b) & \mathcal{L}''_{xy}(a, b, \lambda_0) & \mathcal{L}''_{yy}(a, b, \lambda_0) \end{vmatrix}$$

If $D(a, b, \lambda_0) > 0$ we have a local max at (a, b) and if $D(a, b, \lambda_0) < 0$ we have a local min.

Question 40. *Find max/min of $f(x, y) = x^2 + y^2$ s.t. $x^2 + 2y = 4$.* *Answer to question 40.*

Question 41. *Find max/min of $f(x, y) = x^2 + y^2$ s.t. $x + 2y = 5$.* *Answer to question 41.*

Question 42. ***SH problem 18.3.2*** *Consider the problem*

$$min \quad f(x, y) = (x - 1)^2 + y^2 \quad s.t. \quad y^2 - 8x = 0.$$

(a) *Try to solve the problem by reducing it to a minimization problem in (i) the x variable; (ii) the y variable. Comment.*

(b) *Solve the problem by using the Lagrangian method.*

(c) *Give a geometric interpretation of the problem.*

Answer to question 42.

Answers to questions:

A40. The Lagrangian is $\mathcal{L}(x, y) = x^2 + y^2 - \lambda\cdot(x^2 + 2y - 4)$. The candidates to max/min will come from solving

$$\begin{cases} \mathcal{L}'_x &= 2x - 2\lambda x = 0 \\ \mathcal{L}'_y &= 2y - 2\lambda = 0 \\ x^2 + 2y &= 4 \end{cases} \quad \Rightarrow \quad (0, 2), \lambda = 2 \quad \text{and} \quad (\pm\sqrt{2}, 1) \text{ with } \lambda = 1.$$

Now,

$\mathcal{L}_{\lambda=1}(x, y) = x^2 + y^2 - (x^2 + 2y - 4) = y^2 - 2y + 4,$ clearly a convex function. $(\pm\sqrt{2}, 1)$ GLOBAL MIN.

The value of the global min is $f(\pm\sqrt{2}, 1) = 3$. But,

$$\mathcal{L}_{\lambda=2}(x, y) = x^2 + y^2 - 2(x^2 + 2y - 4) = -x^2 + y^2 - 4y + 8, \qquad \text{not convex, not concave.}$$

In this last case, let us study the bordered Hessian. Now, $(\mathcal{L}_{\lambda=2})''_{xx} = -2$, $(\mathcal{L}_{\lambda=2})''_{xy} = 0$, and $(\mathcal{L}_{\lambda=2})''_{yy} = 2$. Besides $g'_x = 2x$ and $g'_y = 2$. At $(0, 2)$, $g'_x(0, 2) = 0$ and $g'_y(0, 2) = 2$, thus

$$D(0, 2, 2) = \begin{vmatrix} 0 & 0 & 2 \\ 0 & -2 & 0 \\ 2 & 0 & 2 \end{vmatrix} = 4 > 0.$$

This tells us that $(0, 2)$ is a local max. The value of the max is $f(0, 2) = 4$. Back

A41. The Lagrangian is $\mathcal{L}(x, y) = x^2 + y^2 - \lambda\cdot(x + 2y - 5)$. The candidates to max/min will come from solving

$$\begin{cases} \mathcal{L}'_x &= 2x - \lambda = 0 \\ \mathcal{L}'_y &= 2y - 2\lambda = 0 \\ x + 2y &= 5 \end{cases} \quad \Rightarrow \quad (1, 2) \text{ with } \lambda = 2.$$

Now, $\mathcal{L}''_{xx} = 2$, $\mathcal{L}''_{xy} = 0$, and $\mathcal{L}''_{yy} = 2$.[7] The Lagrangian is strictly convex for all (x, y). Consequently the candidate found is a global min$= 5$.[8] There is no max as x and y may tend to ∞ on the constraint and, in this case, we clearly have $x^2 + y^2 \to \infty$.

Back

[7] In this case, the Lagrangian is independent of the value of λ.

[8] In this case, the convexity of the Lagrangian has solved our problem. But if that convexity would not have been possible to determine, we can always check the bordered hessian to see if we can obtain local information. In this case, we have

$$D(a, b, \lambda_0) = \begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 2 \end{vmatrix} = -8 < 0.$$

This tells us that the point is a local min.

A42. (a) (i) Let us insert the constraint $y^2 = 8x$ in the function f:

$$k(x) = (x-1)^2 + 8x.$$

Taking the first derivative and equating to 0:

$$2(x-1) + 8 = 0 \ \Rightarrow \ x = -3.$$

$k(x)$ is a convex parabola so we have a minimum: $k(-3) = -8$ for $x = -3$.

(ii) Solving the constraint for x we have $x = y^2/8$. Replacing in f we have the one variable function

$$h(y) = f(y^2/8, y) = \left(\frac{y^2}{8} - 1\right)^2 + y^2.$$

Taking the first derivative and equating to 0

$$h'(y) = \frac{y}{2} \cdot \left(\frac{y^2}{8} - 1\right) + 2y = 0 \quad \Rightarrow \quad y\left(\frac{y^2}{16} + \frac{3}{2}\right) = 0$$

The only solution is $y = 0$. Let us check the second derivative $h''(y)$:

$$h''(y) = \frac{3y^2}{16} + \frac{3}{2} \qquad \Rightarrow \qquad h'' > 0 \text{ for all } y.$$

Thus, function h is convex throughout $\mathbb{R}$ and the minimum found at $y = 0$, $h(0) = 1$ is a global minimum.

There seems to be a contradiction between (i) and (ii) but a closer inspection shows what happens. In case (i), the domain for $k(x)$ is **not** all $\mathbb{R}$. The constraint $y^2 = 8x$ implies that $x \geq 0$. This circumstance has been overlooked in the resolution. We should have proceeded as follows.

We have to determine the minimum of $k(x) = (x-1)^2 + 8x$ in the set $[0, +\infty)$. The stationary point found, $x = -3$ does not belong to this set. Thus the minimum, if it exists, has to be found at one endpoint. In this case $x = 0$ is the desired endpoint. $k(0) = 1$ and $k'(x) > 0$ for $x > 0$ thus proving that 1 is the minimum value for k.

(b) The Lagrangian function is

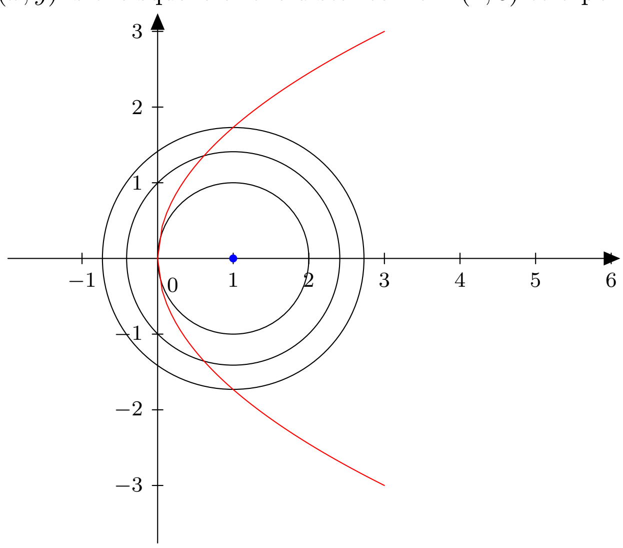
$$\mathcal{L}(x, y) = (x-1)^2 + y^2 - \lambda(y^2 - 8x).$$

The equations to consider are:

$$\begin{cases} 2(x-1) + 8\lambda &= 0 \\ 2y - 2\lambda y &= 0 \\ y^2 &= 8x \end{cases} \quad \Rightarrow \quad x = 0;\ y = 0;\ \lambda = 1/4.$$

The Lagrangian $\mathcal{L}_{\lambda=1/4}(x, y) = (x-1)^2 + (3/4)y^2 + 2x$ is the sum of three convex one-variable functions. It is then convex for all (x, y). The candidate is a global min.[9]

(c) Geometrically, the problem can be stated as finding the point on the parabola $y^2 = 8x$ which is closer to point $(1, 0)$. Function $f(x, y)$ is the square of the distance from $(1, 0)$ to a point (x, y) on the parabola:



Back

[9]The local character of the candidate can also be checked by the bordered hessian at $(0, 0, 1/4)$

$$D(0,0,1/4) = \begin{vmatrix} 0 & -8 & 0 \\ -8 & 2 & 0 \\ 0 & 0 & 3/2 \end{vmatrix} = -96 < 0 \quad \text{local min.}$$

17 Day 17: Lagrange multiplier's economic interpretation.

What happens if constant c in $g(x,y)=c$ changes a little bit? How does the solution max/min of our problem depend on the value of c?

Let $f^*=f(x^*,y^*)$ be the max/min value (solution) to a Lagrange problem. If we consider that (x^*,y^*) satisfy the equations

$$\begin{cases} f'_x(x^*,y^*)-\lambda{\cdot}g'_x(x^*,y^*) &= 0\\ f'_y(x^*,y^*)-\lambda{\cdot}g'_y(x,y) &= 0\\ g(x^*,y^*) &= c \end{cases}$$

and we consider $x^*(c)$ and $y^*(c)$ functions of c, then

$$\frac{df^*}{dc}=\lambda.$$

This is easy to prove by using the chain rule:

$$\frac{df^*}{dc}=\frac{df(x^*,y^*)}{dc}=f'_x(x^*,y^*)\frac{dx^*}{dc}+f'_y(x^*,y^*)\frac{dy^*}{dc}=\lambda g'_x(x^*,y^*)\frac{dx^*}{dc}+\lambda g'_y(x^*,y^*)\frac{dy^*}{dc}=\lambda$$

since from $g(x^*,y^*)=c$ we have (taking derivatives w.r.t. c on both sides):

$$\left.\frac{dg}{dc}\right|_{(x^*,y^*)}=g'_x(x^*,y^*)\frac{dx^*}{dc}+g'_y(x^*,y^*)\frac{dy^*}{dc}=1$$

.

Then, if $g(x,y)=c$ represents a limitation in some resource, λ is the maximum "price" we are prepared to pay for an additional unit of the resource in order to obtain a better value of f^*:

$$f^*(c+1)-f^*(c)\simeq\lambda,$$

and, from a more general point of view, if dc represents a small change in c,

$$f^*(c+dc)-f^*(c)\simeq\lambda{\cdot}dc.$$

The value of λ is called the **shadow price** of the constraint.

Sydsaeter & HammondProblem 18.2.5 Consider the problem max $10x^{1/2}y^{1/3}$ s.t. $2x+4y=m$.

(a) Solve the problem considering the solution for x,y and λ as functions of m.

(b) Check $\frac{df^*}{dm}=\lambda(m)$.

17.1 Linear programming.

Let us consider a particular max/min problem where $f(x,y)$ is a linear function and the domain is a polygon delimited by straight lines (linear equations). This kind of problem is called a Linear Programming problem (LP problem).

Actually, the basic ideas for solving a LP have already been explained. Consider the problem

$$\max/\min \quad f(x,y)=20x+30y \quad \text{s.t.} \quad \begin{cases} 3x+6y &\le 150\\ x+0.5y &\le 22\\ x+y &\le 27.5\\ x,y &\ge 0. \end{cases}$$

This is a typical situation. We will limit ourselves to problems where x and y can take no negative values. Thus, considering the constraints in the problem as the domain on which we want to optimize our function, this domain is always contained in the first quadrant of the xy-plane. As we mentioned before, as all the constraints are linear expressions, the domain will always be a closed convex polygon of the xy-plane. If the polygon is bounded, it is a compact set and we can apply the EVT to the problem. If the polygon is unbounded, the problem may have no solution.

If the solution exists, it will always be on the frontier of our polygon.

We can go somewhat further and as we are dealing with linear functions, when we restrict ourselves to one of the lines of the frontier, the max/min of our function have to be on one of the corners.[10] We can state:

[10] Eventually, a whole side may be formed of solution points if the function is constant on it. But in any case, the solution will be found at a corner point.

If the solution exists, it will always be on a vertex of our polygon.

These considerations lead to two very easy methods for solving a two-variable LP problem:

- Plot the domain (feasible region) in the xy-plane.
- **EVT method**: If you have a compact feasible region, the solution exists (by the EVT) and it is found at one of the corners. List the value of the function (objective function) at each vertex and choose the max/min among the list. Problem solved.
- The EVT method cannot be used in the case of an unbounded feasible region for obviuos reasons: the solution might not exist.
- **Sweeping method**: If you have an unbounded feasible region, the best way to proceed is by drawing on the same diagram a few level curves of the objective function (parallel lines). If there is a solution, one of the level curves will just touch the vertex solution. If there is no solution, you will always be able to get level curves inside the polygon with increasing/decreasing values.
- The sweeping method can always be used for either bounded or unbounded feasible regions.

This methods will not be explained in detail in the plenary sessions.

Read carefully, and understand completely, Section 19.1 of Sydsaeter & Hammond.

Example. Solve the LP problem

$$\max / \min \quad f(x,y) = 20x + 30y \quad \text{s.t.} \quad \begin{cases} 3x + 6y & \leq & 150 \\ x + 0.5y & \leq & 22 \\ x + y & \leq & 27.5 \\ x, y & \geq & 0. \end{cases}$$

both graphically and using SOLVER (see next sub-section).

Solving the same problem with constraint number 1 changed to $3x + 6y \leq 151$ gives us the **shadow price** of constraint 1, that is the extra value in the max/min of f obtained through an extra unit in the RHS of constraint 1:

$$\text{Shadow price of constraint 1: new max/min} - \text{old max/min.}$$

Similarly we would obtain the shadow prices of constraints 2 and 3.

Example. Use SOLVER to solve the three LP problems obtained by increasing one unit the RHS of constraints 1, 2, and 3 (one at a time). Get in this way the three shadow prices in this problem.

17.2 Use of SOLVER.

The same ideas apply to a many-variable linear function (more than 2 vars). But we do not have the help of the graphical approach. The standard method to solve this larger problems is called the SIMPLEX method which uses matrices and Gauss reduction. It is not difficult to explain but quite tedious to apply by hand. Computers do the job for us.

EXCEL's SOLVER is the ideal tool to solve large LP problems (many variables, many constraints).

Question 43. *A firm has three different production plants and four warehouses. A study on next year's production estimates the production and its assignment to the different warehouses (unit= truck load). The table shows the transport cost from each factory to each warehouse (in euros per unit)*

	Wareh. 1	*Wareh. 2*	*Wareh. 3*	*Wareh. 4*	*Production*
Plant 1	*$464*	*$513*	*$654*	*$867*	*75*
Plant 2	*$352*	*$416*	*$690*	*$791*	*125*
Plant 3	*$995*	*$682*	*$388*	*$456*	*100*
Assignment	*80*	*65*	*70*	*85*	

The last row is the minimum amount that each warehouse must store and the last column shows each plant's estimated maximum production.
Establish the distribution plan from each plant to each warehouse in order to minimize the total transportation cost.
Hint. *The variables should be the amount that each plant must send to each warehouse, x_{ij}, where i indicates Plant i and j Warehouse j.*

18 Day 18: Dual problems.

Let us consider the **primal** problem

$$\max c_1x_1 + \ldots + c_nx_n \text{ s.t. } A\cdot\vec{x} \le \vec{b}$$

where A is a $m \times n$ matrix, $A = (a_{ij})$ and $\vec{b} = (b_1, \ldots, b_m)$.

The **dual** problem consists of $\min b_1u_1 + \ldots + b_mu_m$ s.t. $A^{\mathfrak{t}}\cdot\vec{u} \ge \vec{c}$ where $A^{\mathfrak{t}}$ is a the transposed matrix of A, the $n \times m$ matrix, $A^{\mathfrak{t}} = (a_{ji})$ and $\vec{c} = (c_1, \ldots, c_n)$.

(As usual, in both problems the set of variables is non-negative.)

A max n-variable, m-constraints problem has as dual a min m-variable, n-constraints problem. It is also clear that the dual's dual is the primal.

The relationship between the solutions is not so direct:

Theorem [The Duality Theorem. Sydsaeter & Hammond 19.3] *Suppose the primal problem has a (finite) optimal solution. Then the dual problem has also a (finite) optimal solution and the corresponding values of the objective functions are equal. If the primal is unbounded and has no optimum, the dual has no feasible solution (empty feasible set).*

We see then that either both primal and dual have a solution or none has. The value of the max/min is the same but, unfortunately, there is no easy way of getting one solution point in terms of the other.

18.1 Complementary slackness

There is, though a (not easy) way of finding the solutions of the primal knowing the solutions of the dual. This is thanks to the following result.

Theorem [Complementary slackness. Sydsaeter & Hammond 19.4] *The shadow prices of the primal are the solutions to the dual and viceversa. Let us suppose that (x_i^*) is the solution point to the primal and (u_j^*) the solution point to the dual. If replacing the solution in constraint j of the primal, there is some slackness (the result is a strict inequality, that is) then $u_j^* = 0$. Reciprocally, if $u_j^* > 0$, constraint j of the primal is an equality (no slackness).*

The typical situation where we will apply the previous Theorem is the case of a primal of 4 variables (say) and 2 constraints. We cannot solve the primal graphically. But the dual has 2 variables and 4 constraints. We know how to solve that graphically. The dual's solution will be something like (u_1^*, u_2^*). Imagine $u_1^* > 0, u_2^* > 0$. This means that Constraints 1 and 2 in the primal (CP1, and CP2) are equations for the solution (=). And now suppose, in the dual, that replacing u_1^* and u_2^* leads to the slackness of Constraints 3 and. Then $x_3^* = 0$ and $x_4^* = 0$. The values of x_1^*, x_2^* can be obtained by solving the equations CP1 and CP2.

Example. Sydsaeter & Hammond Problem 19.5.2. *Solve graphically the problem:*

$$\min \quad y_1 + 2y_2 \quad s.t. \begin{cases} y_1 + 6y_2 & \ge & 15 \\ y_1 + y_2 & \ge & 5 \\ -y_1 + y_2 & \ge & -5 \\ y_1 - 2y_2 & \ge & -20 \\ y_1, y_2 & \ge & 0 \end{cases}$$

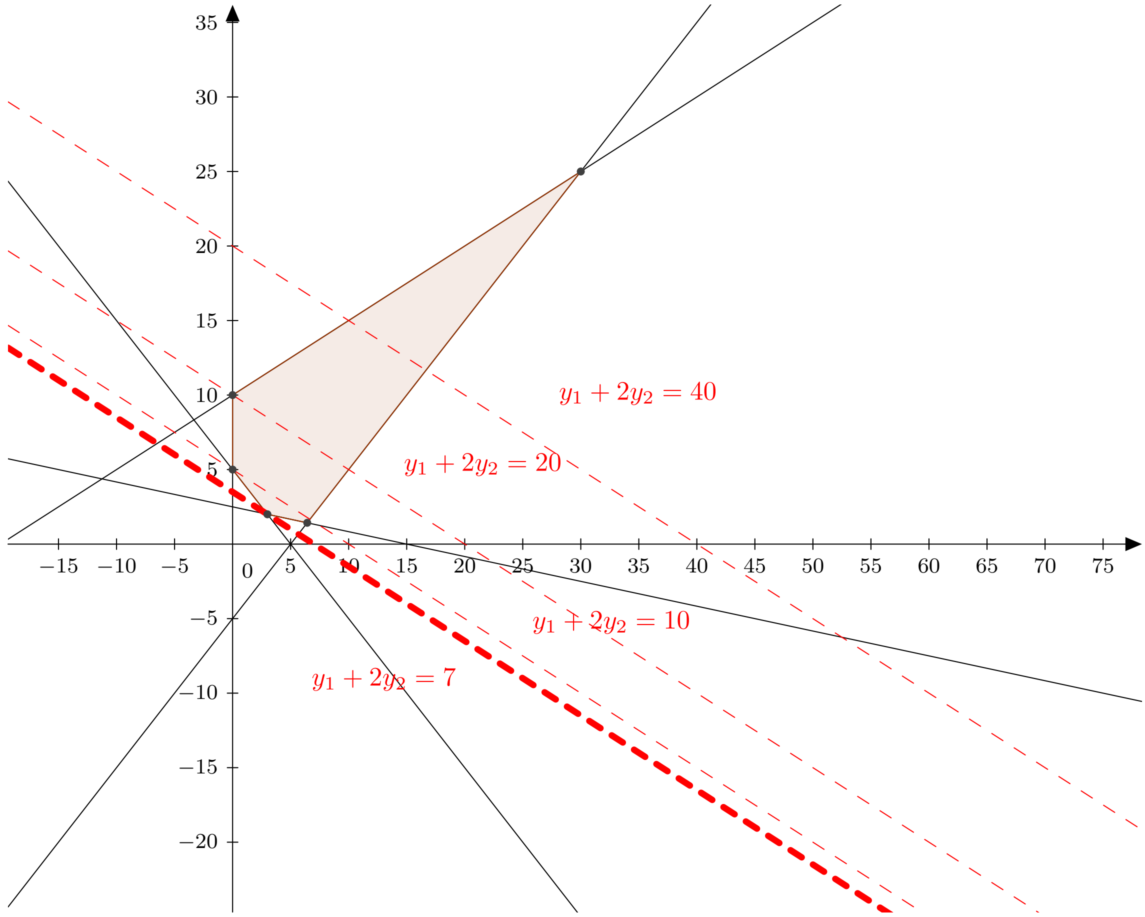
(a) Write down the dual and solve it.

(b) What happens with the optimal solution of the primal if constraint $y_1+6y_2 \ge 15$ is changed to $y_1+6y_2 \ge 15.1$?

(c) What if at the same time, $y_1 + y_2 \ge 5$ is changed into $y_1 + y_2 \ge 5.2$?

SOLUTION.

The solution is found at point $(y_1^*, y_2^*) = (3, 2)$ and the minimal value of the objective function is 7.



(a) The dual is

$$\max \quad 15x_1 + 5x_2 - 5x_3 - 20x_4 \quad \text{s.t.} \begin{cases} x_1 + x_2 - x_3 + x_4 & \leq & 1 \\ 6x_1 + x_2 + x_3 - 2x_4 & \leq & 2 \\ x_1, x_2, x_3, x_4 & \geq & 0. \end{cases}$$

We will solve the dual by using the primal solution and complementary slackness. Let us replace the optimal solution $(3, 2)$ of the primal in the constraints in order to check the existing slackness:

$$\begin{cases} 3 + 6{\cdot}2 & = 15 & \text{active (binding) constraint;} \\ 3 + 2 & = 5 & \text{active constraint;} \\ -3 + 2 & > -5 & \text{inactive (slack) constraint;} \\ 3 - 2{\cdot}2 & > -20 & \text{inactive constraint.} \end{cases}$$

The shadow prices of constraints 3 and 4 will be 0 (the constraints are slack). This means that the dual solution will be $(x_1^*, x_2^*, x_3^*, x_4^*) = (x_1^*, x_2^*, 0, 0)$.
Using the fact that $x_3^* = 0, x_4^* = 0$, the other values of the solution must be that of system

$$\begin{cases} x_1 + x_2 & = & 1 \\ 6x_1 + x_2 & = & 2 \end{cases} \quad \Rightarrow \quad \begin{cases} x_1^* & = & 1/5 \\ x_2^* & = & 4/5. \end{cases}$$

The value of the max is the same as the primal's min: $15{\cdot}(1/5) + 5{\cdot}(4/5) = 7$.

(b) Graphically we see that moving constraint 1 a little bit upwards will not alter the configuration of the solution. The feasible set gets smaller so the value of the min can only increase. We can guess confidently that, as the solutions to the dual are the shadow prices to the primal, $y_1^* = 1/5, y_2^* = 4/5$. Thus, if constraint $y_1 + 6y_2 \geq 15$ the "cost" of moving constraint 1 is $0.1{\cdot}(1/5) = 0.02$ and the new min is 7.02.

(c) Again the change in both constraints is very small . The solution will still be at the intersection of constraint 1 and 2. The new feasible region is even smaller and the new min will be:

$$7 + 0.1{\cdot}(1/5) + 0.2{\cdot}(4/5) = 7.18.$$