Summary on linear approximation in 2 variables:

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

for (x,y) near (x_0,y_0).

Or

$$\boxed{f(x,y) \cong f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0).}$$

Differentiability for $f(x,y)$ at (x_0,y_0) is defined (see definition 9.4.5.1) in a way to make a more precise statement about how good the linear approximation formula is.

Theorem 9.4.5.4 makes it possible to prove that most functions we can write down by a simple formula are differentiable — if you stay away from problem points like zeroes in a denominator or $\ln u$ with $u \le 0$.

9.4.6 The Gradient vector

We define the gradient vector at (x_0,y_0) of a function $f(x,y)$ of two variables as

$$\nabla f \mid_{(x_0,y_0)} = \left(\frac{\partial f}{\partial x} \mid_{(x_0,y_0)}, \frac{\partial f}{\partial y} \mid_{(x_0,y_0)}\right),$$

or you might prefer

$$\nabla f \mid_{(x_0,y_0)} = (f_x(x_0,y_0), f_y(x_0,y_0)).$$

The symbol ∇ has a name "nabla" but we usually just pronounce it 'gradient'. We can maybe simplify the above definition by removing the evaluation at (x_0,y_0) and writing

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

where we understand that both sides are to be evaluated at the same place.

9.4.7 The gradient vector and directional derivatives

We can restate the formula

$$D_{\mathbf{u}} f(x_0,y_0) = f_x(x_0,y_0)u_1 + f_y(x_0,y_0)u_2$$

for the directional derivative as

$$D_{\mathbf{u}} f(x_0,y_0) = \left(\nabla f \mid_{(x_0,y_0)}\right) \cdot \mathbf{u}$$

Remembering how dot products work, we can write this as

$$D_{\mathbf{u}} f(x_0, y_0) = \left\| \nabla f \mid_{(x_0,y_0)} \right\| \|\mathbf{u}\| \cos\theta$$

where θ is the angle between the gradient vector and the direction $\mathbf{u}$. Since $\|\mathbf{u}\| = 1$, we have

$$D_{\mathbf{u}} f(x_0, y_0) = \left\| \nabla f \mid_{(x_0,y_0)} \right\| \cos\theta$$

Since $-1 \leq \cos\theta \leq 1$, we can then say that

$$-\left\| \nabla f \mid_{(x_0,y_0)} \right\| \leq D_{\mathbf{u}} f(x_0, y_0) \leq \left\| \nabla f \mid_{(x_0,y_0)} \right\|$$

Moreover, we can say that the largest directional derivative is the magnitude of the gradient vector and it is achieved only when the direction $\mathbf{u}$ is the unit vector in the direction of the gradient vector (when $\theta = 0$). That is the largest directional derivative is when

$$\mathbf{u} = \frac{\nabla f \mid_{(x_0,y_0)}}{\left\| \nabla f \mid_{(x_0,y_0)} \right\|}$$

The smallest possible directional derivative is minus the length of the gradient vector and that happens only when $\mathbf{u}$ is the unit vector in the opposite direction to the gradient.

(All this supposes that the gradient vector is not the zero vector, which can happen. In that case all the directional derivatives are 0.)

9.4.8 The Chain Rule for partial derivatives

We can now state one version of the chain rule that involves functions of more than one variable.

First a quick refresher on the chain rule in one variable. It says that if $y = f(x)$ and $x = x(t)$ so that in the end we can regard y as a function of t by $y = f(x(t))$, then

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}.$$

In other words it is ok to cancel the dx's in this Leibniz notation for derivatives.

For a function $z = f(x, y)$ of two variables where each of x and y is itself dependent on t by $x = (t)$ and $y = y(t)$, we can ask how the derivative dz/dt is related to the derivatives of the functions that make up the composition $z = f(x(t), y(t))$. The formula is

$$\boxed{\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}}$$

So the simple rule of "cancelling the dx's" becomes less obvious for partial derivatives. Somehow we end up with only one fraction dz/dt when we 'cancel' both the partials ∂x and ∂y.

In vector form we can express the formula as

$$\frac{dz}{dt} = (\nabla f) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt} \right)$$

So we can see we have the gradient multiplied (or dot producted with) the derivative vector (or velocity vector) of the parametric curve

$$(x, y) = (x(t), y(t)).$$

The formula can also be expressed using matrix multiplication if we treat $(x(t), y(t))$ as a 2×1 (column) matrix and ∇f as a 1×2 (row) matrix:

$$\frac{dz}{dt} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix}$$

(We are also taking the velocity vector or tangent vector $(dx/dt, dy/dt)$ to the parametric curve $(x(t), y(t))$ as a column.)

We won't prove this formula, but the way it is usually proved relies on linear approximation.

9.4.9 Level curves

We have concentrated so far on visualising function $z = f(x, y)$ by their graphs, which are surfaces (or look like landscapes) in space.

There is another way which we will now consider. It is to draw (on a two dimensional plane) the curves

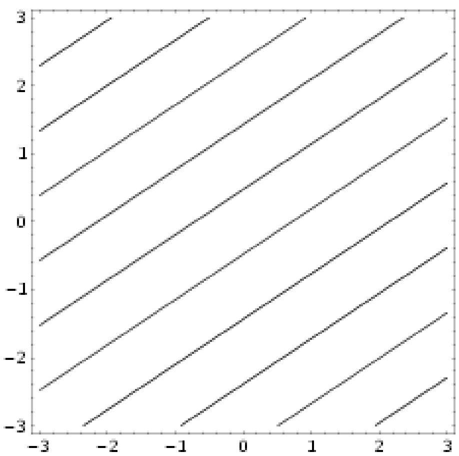
$$f(x, y) = c$$

for different values of c. These are called the *level curves* for the function $f(x, y)$.

There are everyday examples where we see these curves. On the weather chart on the TV or in the newspaper you see curves (called isobars) where the pressure is constant. So these are level curves for the function $f(x, y) =$ atmospheric pressure at point (x, y) on the earth. [Perhaps there is a small problem as the earth is not a plane.] We also sometimes see charts of the temperature today or tomorrow, usually coloured (so that red zones are hot, blue zones are cold). The dividing curves between the different colours are curves where the temperature is constant (30 degrees, 25 degrees, or whatever). So they are level curves for the temperature function.

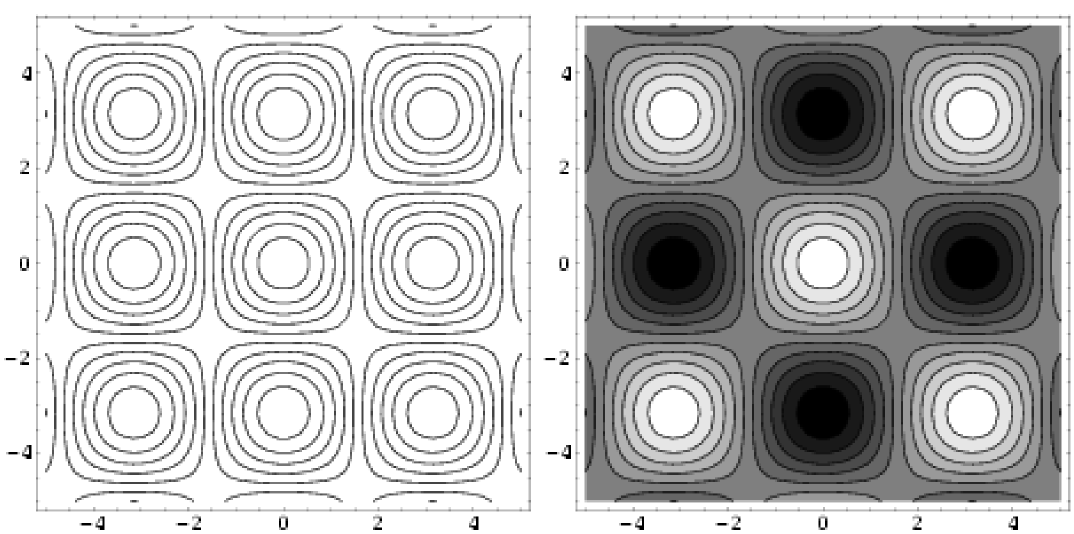
An example that is maybe better for us are the contour curves we see on certain maps. There will be many such contour curves around mountains or hills, perhaps marked with 100 metres, 200 metres and so on. They are curves on the map for tracks you could in principle follow that would be neither uphill nor downhill. They are lines of constant altitude, or level curves for the function $f(x, y) =$ altitude. Off shore they can represent negative altitudes (or depths of the sea bed). This is exactly how our level curves for functions $f(x, y)$ correspond to the graph $z = f(x, y)$ of a function of two variables.

We can maybe look to some of our examples to see how these level curves work out. Here is the example $z = 1 + 2x - 3y$ we plotted first done with contours:



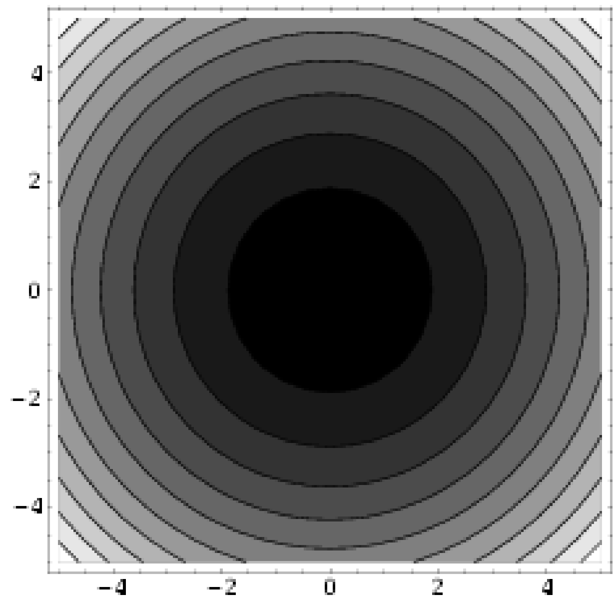
It is maybe not so useful without indicating what the values of c is on each line. We could see by hand that $1 + 2x - 3y = c$ is a line $y = (2/3)x + (1 - c)/3$ of slope $2/3$ and that is what we see in the picture — various lines of that slope.

Here is the picture for $z = \cos x \cos y$, once with just the level curves, and then with shading.



The lighter shaded regions (like near the origin $(x, y) = (0, 0)$) are higher values.

For the example $z = x^2 + y^2$, it is quite easy to see that the level curves $x^2 + y^2 = c$ are circles (of radius $\sqrt{c}$) around the origin if $c > 0$. For $c = 0$ the level 'curve' $x^2 + y^2 = 0$ is just the one point $(0, 0)$, and for $c < 0$ the level 'curve' is empty. Here is a (shaded) plot of some of the contours from the Mathematica programme:



From the fact that the level curves are circles we could immediately realise that the function is symmetrical around the origin, if we did not know that already. We can build the graph, which we saw before was a paraboloid, by lifting up the level curves $x^2 + y^2 = c$ to height c above the horizontal plane.

This may help you to see the relationship between the level curves and the graph. The map example, where they correspond to tracks you can walk along the landscape without going up or down at any stage, is also helpful.

Though a graph is maybe easier to visualise, it takes 3 dimensions. Maps with contour lines showing (say) the Mourne mountains could be harder to interpret, but they are easier to make than 3-dimensional models of the mountains.

9.4.10 Level curves and the gradient

Here we explain that the gradient vector at (x_0, y_0) of $f(x, y)$ will always be perpendicular to the level curve $f(x, y) = c$ that goes through the point (x_0, y_0). That is the one where $c = f(x_0, y_0)$.

Let us start with an example, the simple linear example $z = 1 + 2x - 3y$ and the point $(x_0, y_0) = (7, 11)$. We pictured the level curves (lines) for this earlier. So in this case we are dealing with the function $f(x, y) = 1 + 2x - 3y$ and we can work out the right value of c so that the level curve $f(x, y) = c$ passes through $(x_0, y_0) = (7, 11)$. We need $f(7, 11) = c$, or $1 + 14 - 33 = c$, $c = -18$. So the level curve is $1 + 2x - 3y = -18$ or $2x - 3y = -19$. As this is a line we can get its slope by rewriting it as $2x + 19 = 3y$, or $y = (2/3)x + 19/3$, so that the slope is $2/3$. If we find the gradient of f we get

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2, -3).$$

So the slope of the gradient vector is $-3/2$ and the product of this with the slope of the level curve (line) is $(-3/2)(2/3) = -1$. This shows they are perpendicular (in this example).

We take now another example $f(x,y) = x^2 + y^2$ and the point $(4,-1)$. We saw this example before and we know the level curves are concentric circles centered at the origin. So the tangent lines are perpendicular to the radius. In particular the tangent line to the circle through $(4,-1)$ is perpendicular to the vector $(4,-1)$. But if we work out the gradient vector we get

$$\begin{aligned} \nabla f \mid_{(4,-1)} &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \mid_{(4,-1)} \\ &= (2x, 2y) \mid_{(4,-1)} \\ &= (8, -2). \end{aligned}$$

As this gradient vector is a multiple of $(4,-1)$ we see that it is indeed perpendicular to the tangent line (to the level curve, or level circle in this example).

We've seen that the claimed fact that the gradient is always perpendicular to the tangent line to the level curve has worked out in two examples. Now we will indicate why it is always true.

The level curve $f(x,y) = c$ through a point (x_0, y_0) can be described in a different way by parametric equations

$$\begin{cases} x &= x(t) \\ y &= y(t) \end{cases}$$

in such a way that $(x(0), y(0)) = (x_0, y_0)$. We then have the equation

$$f(x(t), y(t)) = c$$

satisfied for all t. If we differentiate both sides with respect to t we get

$$\begin{aligned} \frac{d}{dt} f(x(t), y(t)) &= 0 \\ \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} &= 0 \end{aligned}$$

by using the chain rule. When we put $t = 0$ this comes down to

$$\left(\frac{\partial f}{\partial x} \mid_{(x_0,y_0)}, \frac{\partial f}{\partial y} \mid_{(x_0,y_0)}\right) \cdot \left(\frac{dx}{dt} \mid_{t=0}, \frac{dy}{dt} \mid_{t=0}\right) = 0.$$

This dot product being zero means that the gradient vector of f at the point (x_0, y_0) is perpendicular to the velocity vector to the curve.

There are some details to fill in, but this is one way to prove that the gradient is perpendicular to the level curve.

Example 9.4.10.1. Find the equation of the tangent line to

$$\cos(2x)\cos y = \frac{1}{2}$$

at the point $(\pi/8, \pi/4)$. Also do the same at $(\pi/6, 0)$.