Lecture 12: Gradient

The **gradient** of a function $f(x, y)$ is defined as

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle \ .$$

For functions of three dimensions, we define

$$\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z) \rangle \ .$$

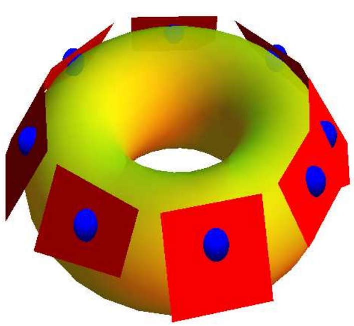
The symbol ∇ is spelled "Nabla" and named after an Egyptian harp. Here is a very important fact:

Gradients are orthogonal to level curves and level surfaces.

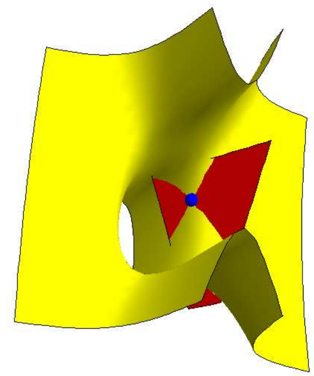
Proof. Every curve $\vec{r}(t)$ on the level curve or level surface satisfies $\frac{d}{dt} f(\vec{r}(t)) = 0$. By the chain rule, $\nabla f(\vec{r}(t))$ is perpendicular to the tangent vector $\vec{r}'(t)$.
Because $\vec{n} = \nabla f(p, q) = \langle a, b \rangle$ is perpendicular to the level curve $f(x, y) = c$ through (p, q), the equation for the tangent line is $ax + by = d$, $a = f_x(p, q)$, $b = f_y(p, q)$, $d = ap + bq$. Compactly written, this is

$$\nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0$$

and means that the gradient of f is perpendicular to any vector $(\vec{x} - \vec{x}_0)$ in the plane. It is one of the most important statements in multivariable calculus. since it provides a crucial link between calculus and geometry. The just mentioned gradient theorem is also useful. We can immediately compute tangent planes and tangent lines:



1 Compute the tangent plane to the surface $3x^2y + z^2 - 4 = 0$ at the point $(1, 1, 1)$. **Solution:** $\nabla f(x, y, z) = \langle 6xy, 3x^2, 2z \rangle$. And $\nabla f(1, 1, 1) = \langle 6, 3, 2 \rangle$. The plane is $6x + 3y + 2z = d$ where d is a constant. We can find the constant d by plugging in a point and get $6x + 3y + 2z = 11$.



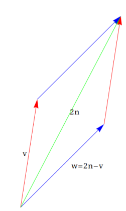
2 **Problem:** reflect the ray $\vec{r}(t) = \langle 1 - t, -t, 1 \rangle$ at the surface

$$x^4 + y^2 + z^6 = 6 \ .$$

Solution: $\vec{r}(t)$ hits the surface at the time $t = 2$ in the point $(-1, -2, 1)$. The velocity vector in that ray is $\vec{v} = \langle -1, -1, 0 \rangle$ The normal vector at this point is $\nabla f(-1, -2, 1) = \langle -4, 4, 6 \rangle = \vec{n}$. The reflected vector is

$$R(\vec{v} = 2\mathrm{Proj}_{\vec{n}}(\vec{v}) - \vec{v} \ .$$

We have $\mathrm{Proj}_{\vec{n}}(\vec{v}) = 8/68 \langle -4, -4, 6 \rangle$. Therefore, the reflected ray is $\vec{w} = (4/17)\langle -4, -4, 6 \rangle - \langle -1, -1, 0 \rangle$.



If f is a function of several variables and $\vec{v}$ is a unit vector then $D_{\vec{v}} f = \nabla f \cdot \vec{v}$ is called the **directional derivative** of f in the direction $\vec{v}$.

The name directional derivative is related to the fact that every unit vector gives a direction. If $\vec{v}$ is a unit vector, then the chain rule tells us $\frac{d}{dt} D_{\vec{v}} f = \frac{d}{dt} f(x + t\vec{v})$.

The directional derivative tells us how the function changes when we move in a given direction. Assume for example that $T(x, y, z)$ is the temperature at position (x, y, z). If we move with velocity $\vec{v}$ through space, then $D_{\vec{v}}T$ tells us at which rate the temperature changes for us. If we move with velocity $\vec{v}$ on a hilly surface of height $h(x, y)$, then $D_{\vec{v}}h(x, y)$ gives us the slope we drive on.

3 If $\vec{r}(t)$ is a curve with velocity $\vec{r}\,'(t)$ and the speed is 1, then $D_{\vec{r}'(t)}f = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$ is the temperature change, one measures at $\vec{r}(t)$. The chain rule told us that this is $d/dt f(\vec{r}(t))$.

4 For $\vec{v} = (1, 0, 0)$, then $D_{\vec{v}}f = \nabla f \cdot v = f_x$, the directional derivative is a generalization of the partial derivatives. It measures the rate of change of f, if we walk with unit speed into that direction. But as with partial derivatives, it is a **scalar**.

The directional derivative satisfies $\boxed{|D_{\vec{v}}f| \leq |\nabla f||\vec{v}|}$ because $\nabla f \cdot \vec{v} = |\nabla f||\vec{v}|\cos(\phi)| \leq |\nabla f||\vec{v}|$.

The direction $\vec{v} = \nabla f/|\nabla f|$ is the direction, where f **increases** most. It is the direction of **steepest ascent**.

If $\vec{v} = \nabla f/|\nabla f|$, then the directional derivative is $\nabla f \cdot \nabla f/|\nabla f| = |\nabla f|$. This means f **increases**, if we move into the direction of the gradient. The slope in that direction is $|\nabla f|$.

5 You are on a trip in a air-ship over Cambridge at $(1, 2)$ and you want to avoid a thunderstorm, a region of low pressure. The pressure is given by a function $p(x, y) = x^2 + 2y^2$. In which direction do you have to fly so that the pressure change is largest?
Solution: The gradient $\nabla p(x, y) = \langle 2x, 4y\rangle$ at the point $(1, 2)$ is $\langle 2, 8\rangle$. Normalize to get the direction $\langle 1, 4\rangle/\sqrt{17}$.

The directional derivative has the same properties than any derivative: $D_v(\lambda f) = \lambda D_v(f)$, $D_v(f + g) = D_v(f) + D_v(g)$ and $D_v(fg) = D_v(f)g + fD_v(g)$.

We will see later that points with $\nabla f = \vec{0}$ are candidates for **local maxima** or **minima** of f. Points (x, y), where $\nabla f(x, y) = (0, 0)$ are called **critical points** and help to understand the function f.

6 The Matterhorn is a 4'478 meter high mountain in Switzerland. It is quite easy to climb with a guide because there are ropes and ladders at difficult places. Evenso there are quite many climbing accidents at the Matterhorn, this does not stop you from trying an ascent. In suitable units on the ground, the height $f(x, y)$ of the Matterhorn is approximated by the function $f(x, y) = 4000 - x^2 - y^2$. At height $f(-10, 10) = 3800$, at the point $(-10, 10, 3800)$, you rest. The climbing route continues into the south-east direction $v = \langle 1, -1\rangle/\sqrt{2}$. Calculate the rate of change in that direction. We have $\nabla f(x, y) = \langle -2x, -2y\rangle$, so that $\langle 20, -20\rangle \cdot \langle 1, -1\rangle/\sqrt{2} = 40/\sqrt{2}$. This is a place, with a ladder, where you climb $40/\sqrt{2}$ meters up when advancing 1m forward.
The rate of change in all directions is zero if and only if $\nabla f(x, y) = 0$: if $\nabla f \neq \vec{0}$, we can choose $\vec{v} = \nabla f/|\nabla f|$ and get $D_{\nabla f}f = |\nabla f|$.

7 Assume we know $D_v f(1, 1) = 3/\sqrt{5}$ and $D_w f(1, 1) = 5/\sqrt{5}$, where $v = \langle 1, 2\rangle/\sqrt{5}$ and $w = \langle 2, 1\rangle/\sqrt{5}$. Find the gradient of f. Note that we do not know anything else about the function f.
Solution: Let $\nabla f(1, 1) = \langle a, b\rangle$. We know $a + 2b = 3$ and $2a + b = 5$. This allows us to get $a = 7/3, b = 1/3$.

Homework

1 A surface $x^2 + y^2 - z = 1$ radiates light away. It can be parametrized as $\vec{r}(x, y) = \langle x, y, x^2 + y^2 - 1\rangle$. Find the parametrization of the wave front which is distance 1 from the surface.

2 Find the directional derivative $D_{\vec{v}}f(2, 1) = \nabla f(2, 1) \cdot \vec{v}$ into the direction $\vec{v} = \langle -3, 4\rangle/5$ for the function $f(x, y) = x^5 y + y^3 + x + y$.

3 Assume $f(x, y) = 1 - x^2 + y^2$. Compute the directional derivative $D_{\vec{v}}(x, y)$ at $(0, 0)$ where $\vec{v} = \langle \cos(t), \sin(t)\rangle$ is a unit vector. Now compute

$$D_v D_v f(x, y)$$

at $(0, 0)$, for any unit vector. For which directions is this **second directional derivative** positive?

4 The **Kitchen-Rosenberg formula** gives the curvature of a level curve $f(x, y) = c$ as

$$\kappa = \frac{f_{xx}f_y^2 - 2f_{xy}f_x f_y + f_{yy}f_x^2}{(f_x^2 + f_y^2)^{3/2}}$$

Use this formula to find the curvature of the ellipsoid $f(x, y) = x^2 + 2y^2 = 1$ at the point $(1, 0)$.
P.S. This formula is known since a hundred years at least but got revived in computer vision. If you want to derive the formula, you can check that the angle

$$g(x, y) = \arctan(f_y/f_x)$$

of the gradient vector has κ as the directional derivative in the direction $\vec{v} = \langle -f_y, f_x\rangle/\sqrt{f_x^2 + f_y^2}$ tangent to the curve.

5 One numerical method to find the maximum of a function of two variables is to move in the direction of the gradient. This is called the **steepest ascent method**. You start at a point (x_0, y_0) then move in the direction of the gradient for some time c to be at $(x_1, y_1) = (x_0, y_0) + c\nabla f(x_0, y_0)$. Now you continue to get to $(x_2, y_2) = (x_1, y_1) + c\nabla f(x_1, y_1)$. This works well in many cases like the function $f(x, y) = 1 - x^2 - y^2$. It can have problems if the function has a flat ridge like in the **Rosenbrock function**

$$f(x, y) = 1 - (1 - x)^2 - 100(y - x^2)^2 \ .$$

Plot the Contour map of this function on $-0.6 \leq x \leq 1, -0.1 \leq y \leq 1.1$ and find the directional derivative at $(1/5, 0)$ in the direction $(1, 1)/\sqrt{2}$. Why is it also called the **banana function**?