MATH11007 NOTES 18: THE DIRECTIONAL DERIVATIVE.

1. The concept of direction

Definition 1.1. A *direction* in $\mathbb{R}^d$ is a vector of unit length.

Here, the *length* of a vector $\mathbf{x} \in \mathbb{R}^d$ is defined by the familiar formula

$$\|\mathbf{x}\| := \left(\sum_{j=1}^{d} x_j^2 \right)^{\frac{1}{2}} .$$

Example 1.1. *When $d = 2$, every direction $\mathbf{u}$ may be expressed in the form*

$$\mathbf{u} = (\cos\theta,\, \sin\theta)$$

for some $\theta \in [0,\, 2\pi)$. The number θ is the angle that the vector makes with the horizontal axis.

Example 1.2. *In $\mathbb{R}^3$, it is often convenient to express a vector $\mathbf{x} = (x, y, z)$ using the* spherical coordinates *ρ, θ and φ. From Figure 1, it is readily deduced that*

$$x = \rho \sin\varphi \cos\theta, \quad y = \rho \sin\varphi \sin\theta, \quad z = \rho \cos\varphi \,.$$

It follows that, when $d = 3$, every direction $\mathbf{u}$ may be expressed in the form

$$\mathbf{u} = (\sin\varphi\cos\theta,\, \sin\varphi\sin\theta,\, \cos\varphi)$$

for some $\theta \in [0,\, 2\pi)$ and some $\varphi \in [0, \pi]$.

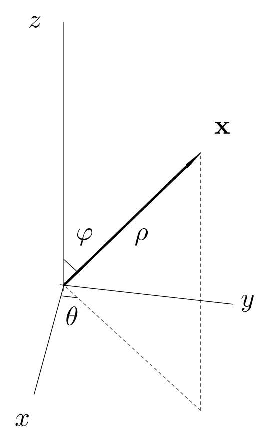

Figure 1. The spherical coordinates ρ, θ and φ of the vector $\mathbf{x}$ in three dimensions.

2. The directional derivative

For a function f of one variable x, the derivative expresses the rate of change of $f(x)$ as x varies. In higher dimension, we can ask how the function value $f(\mathbf{x})$ changes as $\mathbf{x}$ varies *along a particular direction.* Suppose, for example, that you are hillwalking and come to a point where many paths cross. If you are just out to enjoy the fresh air, you may be inclined to take it easy and choose the path (the direction) along which *height* changes least. Let your chosen path (as a function of time) be along the curve

$$\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{u}\,, \quad t \geq 0\,.$$

Here, $\mathbf{x}_0$ is your current position, and $\mathbf{u}$ is the direction in which you are heading. You height above sea level is a function of position, say $f(\mathbf{x})$. The rate of change of f in the direction $\mathbf{u}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t))\Big|_{t=0}\,.$$

Now, by the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t)) = \nabla f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) = \nabla f(\mathbf{x}(t)) \cdot \mathbf{u}\,.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t))\Big|_{t=0} = \nabla f(\mathbf{x}_0) \cdot \mathbf{u}\,.$$

Definition 2.1. The *directional derivative of f in the direction $\mathbf{u}$ at the point $\mathbf{x}_0$* is the number

$$D_{\mathbf{u}} f(\mathbf{x}_0) := \nabla f(\mathbf{x}_0) \cdot \mathbf{u}\,.$$

Example 2.1. *Let $d = 2$ and $\mathbf{u} = (1, 0)$. Then*

$$D_{(1,0)} f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0)) \cdot (1, 0) = f_x(x_0, y_0)\,.$$

So we see that f_x is the directional derivative in the direction $(1, 0)$, i.e. along the positive x axis. You will easily check that f_y is the directional derivative along the positive y axis. More generally, for

$$\mathbf{u} = (\cos\theta, \sin\theta)\,,$$

the directional derivative along $\mathbf{u}$ is

$$D_{\mathbf{u}} f(x, y) = \cos\theta \frac{\partial f}{\partial x}(x, y) + \sin\theta \frac{\partial f}{\partial y}(x, y)\,.$$

3. The direction of greatest change

We saw earlier how the gradient could be interpreted geometrically as a vector normal to a level curve. We proceed to give another useful interpretation of the gradient. For simplicity, we discuss only the case $d = 2$. As before, let

$$\mathbf{u} = (\cos\theta, \sin\theta)\,.$$

Then

$$D_{\mathbf{u}} f(x_0, y_0) = \cos\theta \frac{\partial f}{\partial x}(x_0, y_0) + \sin\theta \frac{\partial f}{\partial y}(x_0, y_0)\,.$$

The right-hand side is a function of θ. We now seek to *maximise* this function of θ; the maximiser will yield the direction of greatest increase (or decrease) of f at

(x_0, y_0). To find it, we use the familiar recipe, i.e. we differentiate the right-hand side of the above equation with respect to θ and equate to zero. This gives

$$-\sin\theta \frac{\partial f}{\partial x}(x_0, y_0) + \cos\theta \frac{\partial f}{\partial y}(x_0, y_0) = 0 \,.$$

This is to be viewed as an equation for the unknown θ. We rewrite it as

$$\nabla f(x_0, y_0) \cdot (-\sin\theta, \cos\theta) = 0 \,.$$

So the θ we are looking for satisfies

$$\nabla f(x_0, y_0) \perp (-\sin\theta, \cos\theta) \,,$$

i.e. the vector on the right must be orthogonal to the gradient. Remark however that

$$(-\sin\theta, \cos\theta) \perp (\cos\theta, \sin\theta) \,.$$

Since we are in two dimensions, we deduce that θ satisfies

$$(\cos\theta, \sin\theta) \parallel \nabla f(x_0, y_0) \,.$$

This says that the direction of greatest change is parallel to the gradient. Clearly

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

is the direction of largest increase of f (why?) and $-\mathbf{u}$ is the direction of largest decrease.

References

1. Frank Ayres, Jr. and Elliott Mendelson, *Schaum's Outline of Calculus, Fourth Edition*, McGraw–Hill, 1999.
2. E. Hairer and G. Wanner, *Analysis by its History*, Springer-Verlag, New-York, 1996.