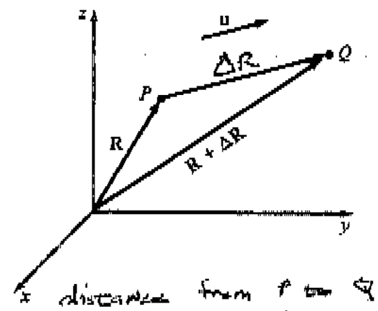
19.5 DIRECTIONAL DERIVATIVES AND THE GRADIENT

$\vec{u}$ the direction we wants to travel

Let $f(x, y, z)$ be a function (of *three* variables!) defined throughout some region of three-dimensional space, and let P be a point in this region. At what rate does f change as we move away from P in a specified direction? In the directions of the positive x-, y-, and z-axes, we know that the rates of change of f are given by the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$. But how do we calculate the rate of change of f if we move away from P in a direction that is not a coordinate direction? In analyzing this problem, we will encounter the very important concept of the gradient of a function.

Suppose that the point P under consideration has coordinates x, y, and z, so that $P = (x, y, z)$; let $\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ be the position vector of P, and let the specified direction be given by a unit vector $\mathbf{u}$, as shown in Fig. 19.10. If we move away from P in this direction to a nearby point $Q = (x + \Delta x, y + \Delta y, z + \Delta z)$, then the function f will change by an amount Δf. If we now divide this change Δf by the distance $\Delta s = |\Delta\mathbf{R}|$ between P and Q, then the quotient $\Delta f/\Delta s$ is the average rate of change of f (with respect to distance) as we move from P to Q. For instance, if the value of f at P is the temperature at this point, then $\Delta f/\Delta s$ is the average rate of change of temperature along the segment PQ. The limiting value of $\Delta f/\Delta s$ as Q approaches P, namely,

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s},$$



distance from P to Q or $|\Delta R|$ is Δs. (change in location) not f value.

Figure 19.10

is called the *derivative of f at the point P in the direction* $\mathbf{u}$, or simply the *directional derivative* of f. In the case of the temperature function, df/ds represents the instantaneous rate of change of temperature with respect to distance—roughly speaking, how fast it is getting hotter—at the point P as we move away from P in the direction specified by $\mathbf{u}$.

This is all very well, but how do we actually calculate df/ds in a specific case? To discover how to do this, we assume that $f(x, y, z)$ has continuous partial derivatives with respect to x, y, and z. Indeed, to avoid the tedious repetition of hypotheses, we make this a blanket assumption for every function we discuss, unless we explicitly state otherwise. With this, the Fundamental Lemma enables us to write Δf in the form

$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z + \epsilon_1 \Delta x + \epsilon_2 \Delta y + \epsilon_3 \Delta z, \tag{1}$$

← Fundamental Lemma in $\mathbb{R}^4$, (see Stewart $\boxed{7}$ on p945)

where $\epsilon_1, \epsilon_2, \epsilon_3 \to 0$ as Δx, Δy, and $\Delta z \to 0$, that is, as $\Delta s \to 0$. Dividing (1) by Δs now gives

$$\frac{\Delta f}{\Delta s} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial f}{\partial z}\frac{\Delta z}{\Delta s} + \epsilon_1 \frac{\Delta x}{\Delta s} + \epsilon_2 \frac{\Delta y}{\Delta s} + \epsilon_3 \frac{\Delta z}{\Delta s}, \tag{2}$$

and by taking the limit as $\Delta s \to 0$, we see that the last three terms in (2) approach zero and we obtain the formula

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} + \frac{\partial f}{\partial z}\frac{dz}{ds}. \tag{3}$$

This formula should be recognized as a special kind of chain rule, in the sense that as we move along the line through P and parallel to $\mathbf{u}$, f is a function of x, y, and z, where x, y, and z are in turn functions of the distance s, and (3) shows how to differentiate f with respect to s.

1996
Calculus, Simmons, p681-3

We observe that the first factor in each product on the right of (3) depends only on the function f and the coordinates of the point P at which the partial derivatives of f are evaluated, while the second factor in each product is independent of f and depends only on the direction in which df/ds is being calculated. These facts suggest that the right side of (3) ought to be thought of—and written—as the dot product of two vectors, as follows:

$$\begin{aligned}\frac{df}{ds} &= \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right)\cdot\left(\frac{dx}{ds}\mathbf{i} + \frac{dy}{ds}\mathbf{j} + \frac{dz}{ds}\mathbf{k}\right)\\ &= \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right)\cdot\frac{d\mathbf{R}}{ds}.\end{aligned} \tag{4}$$

grad $f = \nabla f$

The first factor here is a vector called the *gradient* of f. It is denoted by the symbol grad f, so that by definition

$$\text{grad } f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}. \tag{5}$$

With this notation, (4) can be written as

$$\frac{df}{ds} = (\text{grad } f)\cdot\frac{d\mathbf{R}}{ds}. \quad = \nabla f \cdot \frac{d\vec{R}}{ds} \tag{6}$$

think $\frac{dR}{ds} \approx \frac{\Delta R}{\Delta s}$ which is a vector divided by its magnitude.

But we know that $d\mathbf{R}/ds$ is a unit vector, and since it has the same direction as $\mathbf{u}$, it equals $\mathbf{u}$. Formula (6) is therefore equivalent to

$$\frac{df}{ds} = (\text{grad } f)\cdot\mathbf{u}. \quad \text{for calculations} \tag{7}$$

This tells us how to calculate df/ds, because (5) is presumably simple to compute from the given function f, and then to evaluate at the given point P, and the dot product (7) of two known vectors is easy to find.

For a given function f and a given point P, grad f is a fixed vector which can be placed so that its tail lies at P. We also place the tail of $\mathbf{u}$ at P, as shown in Fig. 19.11. To understand the significance of grad f, we use the definition of the dot product and the fact that $\mathbf{u}$ is a unit vector to write (7) in the form

$$\frac{df}{ds} = |\text{grad } f|\cos\theta, \quad \text{for our intuitive understanding} \tag{8}$$

where θ is the angle between grad f and $\mathbf{u}$. Since the direction of $\mathbf{u}$ can be chosen to suit our convenience, (8) immediately yields the first fundamental property of the gradient:

Property 1 The directional derivative df/ds in any given direction is the scalar projection of grad f in that direction (see Fig. 19.11).

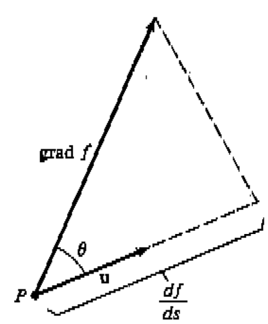


Figure 19.11 Directional derivative.

In this sense, the single vector grad f contains within itself the directional derivatives of f at P in all possible directions.

Next, if $\mathbf{u}$ is chosen to point in the same direction as grad f, so that $\theta = 0$ and $\cos 0 = 1$, then (8) shows that df/ds has its maximum value—that is, f increases most rapidly—in this direction. Also, this maximum value equals $|\text{grad } f|$. These remarks give the next two fundamental properties of the gradient:

Property 2 The vector grad f points in the direction in which f increases most rapidly.

Property 3 The length of the vector grad f is the maximum rate of increase of f.

As these remarks show, even though formulas (7) and (8) are equivalent, they play very different roles in our thinking, for we use (7) to calculate df/ds and (8) to understand the intuitive meaning of the vector grad f.

Example 1 If $f(x, y, z) = x^2 - y + z^2$, find the directional derivative df/ds at the point $(1, 2, 1)$ in the direction of the vector $4\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.

Solution At the point $(1, 2, 1)$, we have $\text{grad}\, f = 2x\mathbf{i} - \mathbf{j} + 2z\mathbf{k} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$. We obtain a unit vector $\mathbf{u}$ in the desired direction by dividing the given vector by its own length,

$$\mathbf{u} = \frac{4\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}}{\sqrt{16 + 4 + 16}} = \frac{2}{3}\mathbf{i} - \frac{1}{3}\mathbf{j} + \frac{2}{3}\mathbf{k}.$$

$\vec{u}$ must be a unit $\hat{\ }$ vector

Formula (7) now gives

$$\begin{aligned}\frac{df}{ds} &= (\text{grad}\, f) \cdot \mathbf{u} \\ &= (2\mathbf{i} - \mathbf{j} + 2\mathbf{k}) \cdot (\tfrac{2}{3}\mathbf{i} - \tfrac{1}{3}\mathbf{j} + \tfrac{2}{3}\mathbf{k}) = 3.\end{aligned}$$

Thus, the function f is increasing at the rate of 3 units per unit distance as we leave $(1, 2, 1)$ in the given direction.

Example 2 Let the temperature of the air at points in space be given by the function $f(x, y, z) = x^2 - y + z^2$. A mosquito located at $(1, 2, 1)$ wishes to get cool as soon as possible. In what direction should it fly?

Solution We saw in Example 1 that $\text{grad}\, f = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ at the point $(1, 2, 1)$. Since the direction of grad f is that in which the temperature increases most rapidly, the mosquito should fly in the opposite direction, that of $-\text{grad}\, f = -2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

The fourth fundamental property of the gradient is useful in geometry. In order to explain what it is, we denote the point under consideration by $P_0 = (x_0, y_0, z_0)$ to emphasize that it is fixed in this discussion, and we let c_0 be the value of our function f at the point P_0. Then the set of all points in space at which $f(x, y, z)$ has the same value c_0 constitutes, in general, a level surface through P_0 whose equation is $f(x, y, z) = c_0$. We wish to show that the vector grad f is normal (perpendicular) to this level surface at the point P_0, as suggested on the left in Fig. 19.12. To this end, we consider a curve that lies on the surface and passes through P_0. If we move to a nearby point Q on this curve and measure s along the curve, then $\Delta f = 0$ because f has the same value at all points on the surface, and therefore $df/ds = 0$ at P_0 in the direction of the tangent to the curve. Formula (6) remains valid and implies that

s = arclength ?

$$(\text{grad}\, f) \cdot \frac{d\mathbf{R}}{ds} = 0, \tag{9}$$