MULTIVARIABLE CALCULUS

1. Continuity

We consder functions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ — which are sets of n-variable functions:

$$\mathbf{f} = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}$$

and define continuity very much the same as we define it for single-variable functions.

Definition 1. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *continuous at a point* $\mathbf{a} = (a_1, \dots, a_n)$ if

$$\lim_{\mathbf{x} \to \mathbf{a}} \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| = 0 \tag{1.1}$$

where $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$ is just Euclidean distance, and $\mathbf{x} \to \mathbf{a}$ means that $\lim \|\mathbf{x} - \mathbf{a}\| = 0$.

There are many subtleties to this definition, though, because $\mathbf{x}$ can approach $\mathbf{a}$ in many ways that affect the limit in equation 1.1:

Example 2. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^1$:

$$f(x_1, x_2) = \frac{x_1 x_2}{x_1^2 + x_2^2} \tag{1.2}$$

This is not continuous at $(0, 0)$ because the limiting value depends on the direction one uses in approaching $(0, 0)$: set $x_2 = kx_1$, where k is some constant, and plug this into equation 1.2:

$$f(x_1, kx_1) = \frac{kx_1^2}{x_1^2 + k^2 x_1^2} = \frac{k}{1 + k^2} \text{ for } x_1 \neq 0$$

so the limiting value depends on what direction one takes — and the function is *not continous* at $(0, 0)$.

Even stranger things can happen, as the following example shows:

Example 3. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^1$:

$$f(x_1, x_2) = \frac{2x_1^3 x_2}{x_1^6 + x_2^2}$$

If we approach $(0, 0)$ along the line $x_2 = kx_1$ with $k \neq 0$, we get

$$f(x_1, x_2) = \frac{2kx_1^4}{x_1^6 + k^2 x_1^2} = x_1^2 \frac{2k}{k^2 + x_1^4}$$

which approaches 0 as x_1 approaches 0. On the other hand, if we approach the origin along the *curved path*$x_2 = x_1^3$, we get

$$f(x_1, x_1^3) = \frac{2x_1^6}{x_1^6 + x_1^6} = 1 \text{ for } x_1 \neq 0$$

so the limit along this curved path is 1. So this function is not continuous at $(0, 0)$ either!

2. Differentiation

In single-variable calculus, we define

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

if this limit exists. We can rewrite it as

$$\lim_{x \to a} \left(f'(a) - \frac{f(x) - f(a)}{x - a} \right) = 0$$

or define

$$r(x) = f'(a) - \frac{f(x) - f(a)}{x - a}$$

and multiply by $x - a$ to get

$$r(x)(x - a) = f'(a)(x - a) - (f(x) - f(a))$$

and rearrange this to get

$$f(x) = f(a) + f'(a)(x - a) - r(x)(x - a)$$

or

$$f(x) = f(a) + f'(a)(x - a) + E(x) \tag{2.1}$$

where the equation now says that $f(x)$ is approximately *linear* in a neighborhood of $x = a$ — it is approximately equal to $f(a) + f'(a)(x - a)$ with an *error* $E(x)$ that approaches 0 as x approaches a. In fact, not only does the error approach 0 — we have

$$E(x) = -r(x)(x - a)$$

so

$$\lim_{x \to a} \frac{E(x)}{x - a} = 0 \tag{2.2}$$

In multivariable calculus, we consider high-dimensional analogues to equation 2.1, i.e.

Definition 4. Given a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ and a point $\mathbf{a} \in \mathbb{R}^n$, we say that $\mathbf{f}$ is *differentiable at the point* $\mathbf{a}$ f there exists a *linear function* $M : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - M(\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0 \tag{2.3}$$

If such an M exists, it will be called the *total derivative* of $\mathbf{f}$ at $\mathbf{a}$.

Remark. Note that a linear function $M : \mathbb{R}^n \to \mathbb{R}^m$ is just an $m \times n$ *matrix*. Also note that, in considering the limit $\mathbf{x} \to \mathbf{a}$, all of the problems in examples 2 and 3 can occur.

Let's suppose that

$$\mathbf{f} = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}$$

and that the total derivative exists. Then we set

$$\mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + \Delta \\ a_{i+1} \\ \vdots \\ a_n \end{bmatrix}$$

and allow Δ to approach 0. This will show that the i^{th} *column* of the M matrix is

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{bmatrix}$$

so that

$$M_{i,j} = \frac{\partial f_i}{\partial x_j}$$

— evaluated at $\mathbf{x} = \mathbf{a}$ (where i =row number and j =column).

3. Special cases

3.1. The Gradient. Define

$$\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix}$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a function, the vector

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

is called its *gradient* — also written grad f.

The gradient has a geometric significance that we can describe in terms of the directional derivative:

If $\mathbf{u}$ is a unit vector, define the *directional derivative* of f *in the direction of* $\mathbf{u}$ and at a point $\mathbf{x} = (x_1, \dots, x_n)$ to be

$$\partial_{\mathbf{u}} f = \lim_{\Delta \to 0} \frac{f(\mathbf{x} + \Delta \mathbf{u}) - f(\mathbf{x})}{\Delta}$$

This is the rate of increase of f as one travels in the direction $\mathbf{u}$.

Comparing this with equation 2.3 shows that

$$\partial_{\mathbf{u}} f = \mathbf{u} \bullet \nabla f = \|\mathbf{u}\| \cdot \|\nabla f\| \cdot \cos\theta = \|\nabla f\| \cdot \cos\theta$$

since $\|\mathbf{u}\| = 1$, where θ is the angle between $\mathbf{u}$ and ∇f. This is clearly a maximum when $\theta = 0$.

It follows that

> The gradient of f, ∇f, points in the direction in which f increases the *fastest,* and its *magnitude* is that *rate* of increase.

3.2. **Divergence.** If $\mathbf{f} = (f_1, \dots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$ is a vector field, the quantity

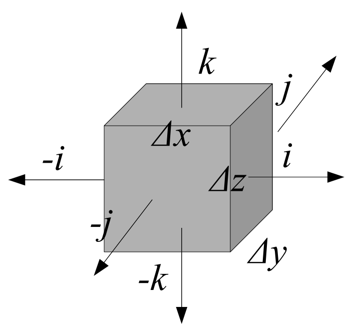
$$\nabla \bullet \mathbf{f} = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}$$

is called the *divergence* of $\mathbf{f}$ — also denoted div $\mathbf{f}$. This has a geometric significance that we will describe when $n = 3$:

Suppose we have a vector field

$$\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$$

where $\mathbf{f}(x, y, z) = f_1(x, y, z)\mathbf{i} + f_2(x, y, z)\mathbf{j} + f_3(x, y, z)\mathbf{k}$ — and we want to measure how much it "spreads out" from a tiny cube of dimensions Δx, Δy, and Δz — per the cube's *volume*:



This cube is so small we can assume the vector-field is *constant* over the faces of the cube.

To find out how much of this field *leaves* a face of the cube (say the top face) we form the dot-product $\mathbf{f}(x, y, z + \Delta z) \bullet \mathbf{k} \cdot \Delta x \Delta y = f_3(x, y, z + \Delta z)\Delta x \Delta y$ — since $\mathbf{i} \bullet \mathbf{k} = \mathbf{j} \bullet \mathbf{k} = 0$. Here, the factor $\Delta x \Delta y$ is the *area* of that face.

The same computation on the opposite face gives $\mathbf{f}(x, y, z) \bullet (-\mathbf{k}) \cdot \Delta x \Delta y = -f_3(x, y, z)\Delta x \Delta y$. The total outflow (from all 6 faces) is:

$$\begin{aligned}
f_3(x, y, z + \Delta z)\Delta x \Delta y \quad &- \quad f_3(x, y, z)\Delta x \Delta y \\
+f_2(x, y + \Delta y, z)\Delta x \Delta z \quad &- \quad f_2(x, y, z)\Delta x \Delta z \\
+f_1(x + \Delta x, y, z)\Delta y \Delta z \quad &- \quad f_1(x, y, z)\Delta y \Delta z \\
&= \quad \Delta x \Delta y (f_3(x, y, z + \Delta z) - f_3(x, y, z)) \\
&+ \quad \Delta x \Delta z (f_2(x, y + \Delta y, z) - f_2(x, y, z)) \\
&+ \quad \Delta y \Delta z (f_1(x + \Delta x, y, z) - f_1(x, y, z))
\end{aligned}$$

Now we divide by the volume of the cube, $\Delta x \Delta y \Delta z$, to get the outflow *per volume*:

$$\frac{f_3(x,y,z+\Delta z)-f_3(x,y,z)}{\Delta z}+$$
$$\frac{f_2(x,y+\Delta y,z)-f_2(x,y,z)}{\Delta y}+$$
$$\frac{f_1(x+\Delta x,y,z)-f_1(x,y,z)}{\Delta x}$$

If we take the limit as Δx, Δy, Δz all go to 0, we get

$$\frac{\partial f_3}{\partial z}+\frac{\partial f_2}{\partial y}+\frac{\partial f_1}{\partial x}=\operatorname{div}\mathbf{f}=\nabla\bullet\mathbf{f}$$

Application. This can be applied to the *Continuity Equation* for fluid flow:

$$\frac{\partial \rho}{\partial t}+\nabla\bullet(\rho\mathbf{v})=0$$

where $\rho(x,y,z,t)$ is the *density* of the fluid and $\mathbf{v}$ is the *velocity* (vector field). What does this say? If ρ (density) *increases* in some region, stuff must pour *into* that region (giving a negative value for the divergence, $\nabla\bullet(\rho\mathbf{v})$) to account for it.

3.3. **Curl.** This is only defined in the 3-dimensional case. If

$$F=\begin{bmatrix} f_1\\ f_2\\ f_3\end{bmatrix}:\mathbb{R}^3\to\mathbb{R}^3$$

is a vector-field, then

$$\nabla\times F=\text{curl } F$$

determines the vorticity of F — i.e. the extent to which F rotates around a point (this is not at all obvious!). The cross-product version gives us the formula:

$$\nabla\times F=\begin{bmatrix}\frac{\partial f_3}{\partial x_2}-\frac{\partial f_2}{\partial x_3}\\ \frac{\partial f_2}{\partial x_3}-\frac{\partial f_2}{\partial x_2}\\ \frac{\partial f_2}{\partial x_1}-\frac{\partial f_1}{\partial x_2}\end{bmatrix}$$

and we have several easy to see identities:

$$\nabla\times(\nabla\varphi)=0$$