9 Functions of several variables

9.1 Recap on planes and lines

In space points may be described by 3 coordinates (x, y, z) or (x_1, x_2, x_3) with respect to 3 perpendicular axes meeting at an origin. Thus we will usually identify $\mathbb{R}^3$, the set of ordered triples of real numbers, with space. That assumes we are fixing an origin and 3 perpendicular axes in space.

We allow ourselves to pass from points $P = (x, y, z)$ (with 3 coordinates x, y and z) to vectors $\mathbf{P} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ (with 3 components x, y and z), where $\mathbf{P}$ can be thought of as the position vector of the point. (Graphically $\mathbf{P}$ is the vector represented by the arrow from the origin to the point P. $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors in the directions of the positive x-, y- and z-axes.) It is a short step to ignore the difference between the point and the vector. That difference concerns how we think of them, rather than how we manipulate them.

The simplest objects to describe are planes and lines.

Planes are described as the points $(x, y, z) \in \mathbb{R}^3$ that satisfy a single linear equation

$$\alpha x + \beta y + \gamma z = c,$$

where (α, β, γ) is a (fixed) nonzero vector in $\mathbb{R}^3$ perpendicular to the plane and c is a constant. As a vector, (α, β, γ) (or $\alpha\mathbf{i} + \beta\mathbf{j} + \gamma\mathbf{k}$ if you prefer) is called a *normal vector* to the plane. To be a plane we must have $(\alpha, \beta, \gamma) \neq (0, 0, 0)$. We can normalise to have a unit normal vector if we want to do so (by dividing the equation across by $\|(\alpha, \beta, \gamma)\| = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$). However,

$$(-\alpha)x + (-\beta)y + (-\gamma)z = (-c)$$

is still another equation for the same plane.

Lines may be described as the intersection of two planes that are not parallel (which you can detect by checking that the normal vectors to the planes are not multiples of one another). However, given a line there are many choices for pairs of planes that intersect in that line. So it is tidier to describe lines (in space) by parametric equations, linear parametric equations.

If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ and $\mathbf{b} \neq \mathbf{0}$, then we can quite easily see that the points with position vectors

$$\mathbf{P} = \mathbf{a} + t\mathbf{b}$$

all lie on the line through the point $\mathbf{a}$ in the direction parallel to $\mathbf{b}$. As t varies through all of $\mathbb{R}$ (positive, zero and negative values) we get all the points on that line (once).

We refer to t as a *parameter* and the above equation $\mathbf{P} = \mathbf{a} + t\mathbf{b}$ as a parametric equation for the line in vector form.

If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ = the position vector of the point (a_1, a_2, a_3) on the line, if $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, and if we write $\mathbf{P} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ then what we have is that the position vectors of the points on the line are

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) + t(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})$$

This corresponds to 3 scalar equations

$$\begin{cases} x &= a_1 + b_1 t \\ y &= a_2 + b_2 t \\ z &= a_3 + b_3 t \end{cases}$$

9.1.1 In the plane

In $\mathbb{R}^2$, we are already familiar with how to deal with lines. Often we say that a line with slope m has an equation $y = mx + c$, but we have to cope with vertical lines $x = c$ separately.

We can cover all kinds of lines by saying that they have a linear equation

$$\alpha x + \beta y = c$$

with $(\alpha, \beta) \neq (0, 0)$.

We can also use the parametric equation idea to describe lines in the plane. If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ and $\mathbf{b} \neq \mathbf{0}$, then we can quite easily see that the points with position vectors

$$\mathbf{P} = \mathbf{a} + t\mathbf{b}$$

all lie on the line through the point $\mathbf{a}$ in the direction parallel to $\mathbf{b}$. We can say that $\mathbf{P} = \mathbf{a} + t\mathbf{b}$ gives the parametric equations in vector form, and this corresponds to 2 scalar equations

$$\begin{cases} x &= a_1 + b_1 t \\ y &= a_2 + b_2 t \end{cases}$$

(where $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} = (b_1, b_2)$).

9.2 Recap on products for vectors

The *dot product* (or *scalar product*) of two vectors $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ in $\mathbb{R}^3$ is defined as

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

A similar definition applies in $\mathbb{R}^2$ (leaving off the x_3 and y_3 bits).

The *norm* of a vector (point) $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ is

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

and this may also be thought of as the distance $\mathrm{dist}(\mathbf{x}, (0, 0, 0))$ of the point from the origin, or as the magnitude or length of the vector $\mathbf{x}$.

We can prove geometrically in $\mathbb{R}^3$ (or in $\mathbb{R}^2$) that

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

where θ is the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$.

Everything we have mentioned above about dot products and norms extends to $\mathbb{R}^n$ (any $n \in \mathbb{N}$), except that we cannot prove the formula about $\cos\theta$. We take the formula about $\cos\theta$ as as a definition of the angle θ. (We keep θ in the range 0 to π so as to avoid it being ambiguously defined. We also need to have nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ for $\theta = \cos^{-1}(\mathbf{x}\cdot\mathbf{y}/(\|\mathbf{x}\|\|\mathbf{y}\|))$ to make sense — to avoid having any division by zero.)

We say that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are *perpendicular* or *orthogonal* if $\mathbf{x}\cdot\mathbf{y} = 0$. This is the same as having the angle between $\mathbf{x}$ and $\mathbf{y}$ being $\theta = \pi/2$, except that we define the zero vector $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbb{R}^n$ to be perpendicular to every vector.

While dot products work in any dimension, there is a special kind of product of vectors that works only in $\mathbb{R}^3$, called the *cross product*. If $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$, then their cross product $\mathbf{x}\times\mathbf{y}$ can be defined as

$$\mathbf{x}\times\mathbf{y} = \mathbf{i}(x_2y_3 - x_3y_2) - \mathbf{j}(x_1y_3 - x_3y_1) + \mathbf{k}(x_1y_2 - x_2y_1),$$

or

$$\mathbf{x}\times\mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

While it is possible to remember the pattern to this, most people remember it by use of a kind of determinant formula.

For 2×2 matrices, the determinant is a number

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

For 3×3 matrices, the determinant is again a number and it can be described by what is called the cofactor expansion along the first row:

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Once familiar with this, we can say that

$$\mathbf{x}\times\mathbf{y} = \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = \mathbf{i}(x_2y_3 - x_3y_2) - \mathbf{j}(x_1y_3 - x_3y_1) + \mathbf{k}(x_1y_2 - x_2y_1)$$

Proposition 9.2.1. *For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ and $\lambda \in \mathbb{R}$, the cross product has the properties*

(i) $\mathbf{y}\times\mathbf{x} = -\mathbf{x}\times\mathbf{y}$*;*

(ii) $\mathbf{x}\times\mathbf{x} = \mathbf{0}$*;*

(iii) $(\mathbf{x}+\mathbf{y})\times\mathbf{z} = \mathbf{x}\times\mathbf{z} + \mathbf{y}\times\mathbf{z}$*;*
$\mathbf{x}\times(\mathbf{y}+\mathbf{z}) = \mathbf{x}\times\mathbf{y} + \mathbf{y}\times\mathbf{z}$*;*

(iv) $(\lambda \mathbf{x}) \times \mathbf{y} = \lambda(\mathbf{x} \times \mathbf{y}) = \mathbf{x} \times (\lambda \mathbf{y})$*;*

(v) $\mathbf{x} \times \mathbf{y}$ *is perpendicular to both vectors* $\mathbf{x}$ *and* $\mathbf{y}$*;*
that is $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = 0$ *and* $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y} = 0$*;*

(vi) $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\| \sin\theta$ *where* θ *is the angle between the vectors.*

Remark 9.2.2. Notice that the last two properties explain what the vector $\mathbf{x} \times \mathbf{y}$ is in a geometrical way, except that it leaves two possibilities for the cross product. These properties don't distinguish $\mathbf{x} \times \mathbf{y}$ from $\mathbf{y} \times \mathbf{x} = -\mathbf{x} \times \mathbf{y}$. There is a way to do that according to a so-called right-hand rule. It requires that we have a right-handed choice of axes. We have $\mathbf{i} \times \mathbf{j} = \mathbf{k}$. In a right handed coordinate system, a screw placed along the direction of $\mathbf{k}$ and turned from $\mathbf{i}$ towards $\mathbf{j}$ will travel in the direction of $\mathbf{k}$ (and not the direction of $-\mathbf{k}$).

9.3 Vector valued functions of a single variable

It is quite straightforward to define derivatives for vector-valued functions of a single variable. Since the case of values in $\mathbb{R}^2$ or $\mathbb{R}^3$ is not really easier than the case of $\mathbb{R}^n$, we give the definition in general.

Definition 9.3.1. For a vector valued function $\mathbf{x}\colon \mathbb{R} \to \mathbb{R}^n$ of a single variable, we write

$$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$$

where $x_1(t), x_2(t), \ldots, x_n(t)$ are n ordinary functions (scalar valued functions of a single variable). We can allow $\mathbf{x}(t)$ to be defined only for some subset of $t \in \mathbb{R}$ (say for t in some interval).

We define the *derivative* of $\mathbf{x}(t)$ to be

$$\frac{d\mathbf{x}}{dt} = \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \ldots \frac{dx_n}{dt} \right)$$

(so the vector-valued function got by differentiating the components or coordinates of $\mathbf{x}(t)$ individually, provided the components are differentiable).

Remark 9.3.2. There may be an advantage in using other notations like the prime notation $\mathbf{x}'(t)$ or the dot notation $\dot{\mathbf{x}}(t)$ (which some people like to use for a time derivative). Perhaps

$$\mathbf{x}'(t) = (x_1'(t), x_2'(t), \ldots, x_n'(t))$$

looks more tidy than the version with d/dt notation above.

Remark 9.3.3. Specialising to the case where $n = 2$ or $n = 3$, we can view a vector-valued function $\mathbf{x} = \mathbf{x}(t)$ as describing a parametric curve in $\mathbb{R}^2$ or $\mathbb{R}^3$. Or we can think of a moving point, with position $\mathbf{x}(t)$ at time t. In $\mathbb{R}^2$ we can look at the trace of the curve on a piece of paper, while in $\mathbb{R}^3$ we can at least imagine a trail followed by the point.