ADVANCED CALCULUS
SPRING 2019

LESSON 0: Review

This lesson is a *very brief* review of the following topics, covered in MATH 2010 (Multivariable Calculus and Linear Algebra) and also in Chapters 1 and 2 of the Marsden-Tromba text.

- Vectors in $\mathcal{R}^3$: geometric and algebraic representations, basis vectors, parametric representation of curves, dot and cross products, projection, planes and distances.
- Cylindrical and spherical coordinates.
- Multivariate functions: partial differentiation, differentiability, linear approximation, chain rules, directional derivative, gradient.

The intent of the review is to promote a deeper understanding and retention of concepts already learned. Exercises form an integral part of the review. These are not of the routine, 'plug-and-chug' kind, but are aimed at processing information and setting up and solving problems. Some of them may well be found to be challenging.

1 Introduction

Single-variable calculus provides the language for a mathematical description of one-dimensional phenomena; to a number x we associate a number $y = f(x)$. Multivariable calculus extends the machinery to allow modelling of the multidimensional world. We shall consider functions of many variables, or multivariate functions, of form

$$y_i = f_i(x_1, x_2, \cdots, x_n), \quad i = 1, 2, \cdots, m.$$

Examples abound in science, engineering and the humanities. In meteorology, pressure, temperature, humidity and wind speed may be described as functions of position and time, four functions of four variables. In microeconomics the production is a function of input; with production consisting of as many variables as the products made by the company, and input consisting of such variables as raw materials, equipment, personnel, energy, and so on. Mathematically one is faced with functions $\boldsymbol{f} : \mathcal{R}^n \to \mathcal{R}^m$ that associate m outputs $\boldsymbol{y} = (y_1, y_2, \cdots, y_m)$ with n inputs $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, *i.e.*, $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$.

Here $\mathcal{R}^n$ is the set of ordered lists of n real numbers. An element $\boldsymbol{x} \in \mathcal{R}^n$ can be interpreted in one of two ways, either as a point representing the state or as a vector representing a change or increment in the state. For example, in $\mathcal{R}^3$, the ordered triple (x_1, x_2, x_3) may represent a point P whose coordinates are given by the ordered triple, or a vector $\vec{OP}$ from the origin O to the point P whose components in the chosen coordinate system, or basis, are given by the ordered triple. Distinction between points and vectors can be blurry, but we shall use the notation

$$(x_1, x_2, \cdots, x_n)$$

for a point and

$$< x_1, x_2, \cdots, x_n >$$

for a vector. We note further that when a list is interpreted as a vector, then the promotion of $\mathcal{R}^n$ from a space of ordered lists to a space of vectors requires endowing the space with additional properties, those of closure under addition and multiplication by a scalar, inner product, norm and distance.

Remark. When writing by hand in class, vectors will be denoted by an underbar, as in $\underline{v}$, typically using lower-case letters. In typeset notes vectors will be denoted by a boldface font, as in $\boldsymbol{v}$. *Be sure to distinguish a vector from a scalar in your written work.*

2 Vectors in $\mathcal{R}^3$

In the physical three-dimensional world, the term vector refers to a physical quantity that requires both magnitude and direction for its specification, such as displacement, velocity and acceleration. It can be represented geometrically as a directed line segment, a representation that is frame-invariant, *i.e.*, does not require specifying a coordinate frame. However, once a coordinate frame is specified, the vector can be represented by the ordered triple of its components, and is therefore an element of the vector space $\mathcal{R}^3$. As an example, consider the position vector $\boldsymbol{r}$ representing a directed line segment from the origin to the point (x_1, x_2, x_3). We write $\boldsymbol{r}$ in terms of its components as

$$\boldsymbol{r} =< x_1, x_2, x_3 > .$$

An equivalent expression is

$$\boldsymbol{r} = \sum_{i=1}^{3} x_i \boldsymbol{e}_i,$$

where the $\boldsymbol{e}_i$, $i = 1, 2, 3$, are the unit vectors along the coordinate axes, *i.e.*, $\boldsymbol{e}_1 =< 1, 0, 0 >$, $\boldsymbol{e}_2 =< 0, 1, 0 >$, $\boldsymbol{e}_3 =< 0, 0, 1 >$.

A commonly used alternative notation is $\boldsymbol{r} =< x, y, z >$ or $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, with $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ denoting the unit vector triad. On occasions we may replace $\boldsymbol{r}$ by $\boldsymbol{x}$.

Standard algebraic operations

1. Zero vector, equality, sum (parallelogram law), multiplication by a scalar.

2. Inner product or dot product,
$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$
Dot product is commutative and distributive. Also,
$$(\alpha \boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (\alpha \boldsymbol{v}) = \alpha(\boldsymbol{u} \cdot \boldsymbol{v}).$$
The dot product helps define length or norm,
$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}.$$
Length has the obvious geometric meaning. Also, geometrically,
$$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos\theta,$$
where θ is the smaller angle between the vectors when they are placed tail-to-tail. Zero dot product implies orthogonality. (Standard basis vectors are orthogonal.)
Normalization yields the unit vector[1],
$$\boldsymbol{e}_v = \frac{1}{\|\boldsymbol{v}\|} \boldsymbol{v}.$$
Distance between points $\boldsymbol{x}$ and $\boldsymbol{y}$ is the length $\|\boldsymbol{x} - \boldsymbol{y}\|$.

Some applications of the dot product

(i) Angle between two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$,
$$\cos\theta = \boldsymbol{e}_u \cdot \boldsymbol{e}_v.$$

(ii) Direction cosines and direction angles of a vector $\boldsymbol{u}$,
$$\cos\theta_1 = \boldsymbol{e}_u \cdot \boldsymbol{e}_1, \; \cos\theta_1 = \boldsymbol{e}_u \cdot \boldsymbol{e}_2, \; \cos\theta_3 = \boldsymbol{e}_u \cdot \boldsymbol{e}_3.$$

Example 2.1. Find to the nearest degree the acute angle formed by an edge and a diagonal of a cube.
Consider the unit cube in the first octant with one corner at the origin. Then the vector along the diagonal through the origin is $\boldsymbol{u} =< 1, 1, 1 >$ and the vector along one edge (the one pointing in the x_1 direction) is $\boldsymbol{e}_1 =< 1, 0, 0 >$. Therefore the desired angle is given by
$$\cos\theta = \boldsymbol{e}_1 \cdot \boldsymbol{e}_u = \boldsymbol{e}_1 \cdot \boldsymbol{u}/\|\boldsymbol{u}\| = 1/\sqrt{3}.$$

(iii) Casting a shadow: *component* of $\boldsymbol{w}$ along $\boldsymbol{v}$,
$$\text{comp } \boldsymbol{w}_v = \boldsymbol{w} \cdot \boldsymbol{e}_v.$$
The component can be positive or negative. Some texts define the component by its absolute value $|\boldsymbol{w} \cdot \boldsymbol{e}_v|$.
Projection or *vector projection* of $\boldsymbol{w}$ along $\boldsymbol{v}$,
$$\text{proj } \boldsymbol{w}_v = (\boldsymbol{w} \cdot \boldsymbol{e}_v)\boldsymbol{e}_v.$$
The projection can point in the direction of, or in a direction opposing, $\boldsymbol{v}$.

[1] We shall use $\boldsymbol{e}_v$ to denote the unit vector in the direction of vector $\boldsymbol{v}$.

(iv) Projection onto a plane: first find projection onto the unit normal vector $\boldsymbol{n}$ to the plane, then take the difference, *i.e.*,

$$\boldsymbol{w} - (\boldsymbol{w} \cdot \boldsymbol{n})\boldsymbol{n}.$$

(v) Force, displacement and work,

$$W = \boldsymbol{f} \cdot \boldsymbol{d},$$

where $\boldsymbol{f}$ is the force and $\boldsymbol{d}$ the displacement.

(vi) Flux of a (uniform) vector field $\boldsymbol{v}$ across a (planar) surface,

$$\boldsymbol{v} \cdot \boldsymbol{n},$$

where $\boldsymbol{n}$ is the unit normal to the surface.

3. Cross product,

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = (v_2w_3 - v_3w_2)\boldsymbol{e}_1 + (v_3w_1 - v_1w_3)\boldsymbol{e}_2 + (v_1w_2 - v_2w_1)\boldsymbol{e}_3.$$

The cross product satisfies

$$\boldsymbol{v} \times \boldsymbol{w} = -\boldsymbol{w} \times \boldsymbol{v}, \quad \boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w}, \quad \alpha(\boldsymbol{v} \times \boldsymbol{w}) = (\alpha\boldsymbol{v}) \times \boldsymbol{w} = \boldsymbol{v} \times (\alpha\boldsymbol{w}).$$

Geometrically,

$$\boldsymbol{v} \times \boldsymbol{w} = \|\boldsymbol{v}\|\|\boldsymbol{w}\| \sin\theta \, \boldsymbol{n},$$

Where $\boldsymbol{v}$, $\boldsymbol{w}$ and $\boldsymbol{n}$ form a right-handed triad.

Some applications of the cross product

(i) Area of parallelogram, parallelism, normal to plane containing two non-coplanar vectors.

(ii) Moment (torque) is $\boldsymbol{r} \times \boldsymbol{f}$, linear velocity of a point on a rotating body is $\boldsymbol{\omega} \times \boldsymbol{r}$, angular momentum about the origin is $\boldsymbol{r} \times m\boldsymbol{v}$.

(iii) Particle of electric charge q moves with velocity $\boldsymbol{v}$ in a field of magnetic induction $\boldsymbol{B}$. The force on the charge is $q\,\boldsymbol{v} \times \boldsymbol{B}$.

(iv) Scalar and vector triple products; volume of a parallelepiped, coplanarity of three vectors.

Example 2.2 (a). Given the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$, let $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{w}$. Does this imply that $\boldsymbol{v}$ and $\boldsymbol{w}$ must be equal? If not, in what way are $\boldsymbol{v}$ and $\boldsymbol{w}$ related? Repeat the problem for $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{w}$.

Note that $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{w}$ implies that $\boldsymbol{u} \times (\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{0}$. Therefore, either $\boldsymbol{v} = \boldsymbol{w}$ or $\boldsymbol{u}$ is parallel to $\boldsymbol{v} - \boldsymbol{w}$. Geometrically, if $\boldsymbol{v}$ and $\boldsymbol{w}$ from two sides of a triangle, then $\boldsymbol{u}$ is parallel to the third side.

Similarly, $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{w}$ implies that $\boldsymbol{u} \cdot (\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{0}$. Therefore, either $\boldsymbol{v} = \boldsymbol{w}$ or $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v} - \boldsymbol{w}$. Again, if $\boldsymbol{v}$ and $\boldsymbol{w}$ from two sides of a triangle, then $\boldsymbol{u}$ is perpendicular to the third side.

Example 2.2 (b). Discuss the solutions $\boldsymbol{u}$ of $||\boldsymbol{a} \times \boldsymbol{u}|| = b$, where $\boldsymbol{a}$ and b are known.

Let θ be the angle between $\boldsymbol{a}$ and $\boldsymbol{u}$. Then $||\boldsymbol{a} \times \boldsymbol{u}|| = b$ implies that

$$||\boldsymbol{a}||\,||\boldsymbol{u}||| \sin\theta| = b, \quad \text{or} \quad ||\boldsymbol{u}||| \sin\theta| = \frac{b}{||\boldsymbol{a}||},$$

i.e., the component of $\boldsymbol{u}$ in a direction perpendicular to $\boldsymbol{a}$ is $b/||\boldsymbol{a}||$. Therefore, if $\boldsymbol{a}$ and $\boldsymbol{u}$ have the same point of origin O, then the tip of $\boldsymbol{u}$ must lie on a cylinder of radius $b/||\boldsymbol{a}||$ whose axis is along $\boldsymbol{a}$.

At this point it is useful to digress briefly to become familiar with the index notion described in the Appendix.

Equation of a line. In the xy-plane a line may represented by, say, a point (x_0, y_0) and a direction. Direction in a plane can be defined by the slope m, leading to the point-slope form

$$y - y_0 = m(x - x_0).$$

The line through two given points in the xy-plane can be treated similarly. The line is represented by a single, static equation which indicates how the general point (x, y) must be restricted for it to lie on the line.

Two coordinates - one degree of freedom = one (constraint) equation.

- What about a line in space? Again, a point and a direction suffice. The latter is specified most conveniently by a vector, rather than by a generalization of the notion of the slope. Vector equation (parametric form):

$$\boldsymbol{r} = \boldsymbol{r}_0 + t\boldsymbol{v}.$$

  Here, t is a parameter that locates the current point. Thus the line is described *dynamically*, as if it is being traced; contrast it with the static description above.

- Equation in scalar (parametric) form:

$$\begin{aligned} x &= x_0 + t\,v_1 \\ y &= y_0 + t\,v_2 \\ z &= z_0 + t\,v_3 \end{aligned}$$

- Equation in symmetric form:

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}.$$

  Question: Why do we need two equations to represent a line in space, as opposed to a single equation in 2D?
  Answer: three coordinates - one degree of freedom = two (constraint) equations.

- Line through two points:

$$\boldsymbol{r} = \boldsymbol{r}_0 + t\,(\boldsymbol{r}_1 - \boldsymbol{r}_0).$$

Example 2.3. Find the line passing through $(1, 2, 3)$ and $(4, -2, 2)$, and its point of intersection with the plane $z = 0$.

The vector joining the two points is $< 3, -4, -1 >$. Therefore the line is $\boldsymbol{r} =< 1, 2, 3 > +t < 3, -4, -1 >=< 1 + 3t, 2 - 4t, 3 - t >$. It intersects the plane $z = 0$ at $3 - t = 0$ or $t = 3$. The corresponding point on the line is $(10, -10, 0)$.

- All manners of defining a line are equivalent to specifying $\boldsymbol{r}_0$ and $\boldsymbol{v}$.

Example 2.4. Explain why the equations

$$x = 1 + 3t,\ y = -2 + t,\ z = 2t \quad \text{and} \quad x = 4 - 6s,\ y = -1 - 2s,\ z = 2 - 4s$$

represent the same line.

If the equations represent the same line, they must point in the same direction and have a common point. The direction vectors of the two lines are $\boldsymbol{u} =< 3, 1, 2 >$ and $\boldsymbol{v} =< -6, -2, -4 >$. The vectors point in the same direction as they are multiples of each other; $\boldsymbol{v} = -2\boldsymbol{u}$. On the first line, $t = 0$ corresponds to the point $P(1, -2, 0)$. The point P also lies on the second line, and corresponds to $s = 1/2$. Hence the two lines are the same.

Equation of a plane. A plane may be defined by a point $\boldsymbol{r}_0$ and a normal vector $\boldsymbol{n}$. The equation is

$$(\boldsymbol{r}-\boldsymbol{r}_0)\cdot\boldsymbol{n}=0, \quad \text{or,} \quad a(x-x_0)+b(y-y_0)+c(z-z_0)=0, \quad \text{or,} \quad ax+by+cz=d.$$

- The above equation is linear in x, y and z. It is a static equation, a restriction on the coordinates if the point is to lie on the plane. (Three coordinates - two degrees of freedom = one (constraint) equation.)
- A dynamic equation would be $\boldsymbol{r}=\boldsymbol{r}_0+s\boldsymbol{u}+t\boldsymbol{v}$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are any two non-collinear vectors in the plane. We may also use two of the coordinates as parameters.
- The planes $x=x_1$, $y=y_1$ and $z=z_1$ are parallel to the coordinate planes.
- All manners of defining a plane are equivalent to specifying either $\boldsymbol{r}_0$, $\boldsymbol{u}$ and $\boldsymbol{v}$ or $\boldsymbol{r}_0$ and $\boldsymbol{n}$. Think about (i) how to find the equation of a plane through three given points, and (ii) how to find the perpendicular distance of a point from a plane.

Example 2.5. Show that the following lines intersect and find the point of intersection. $\underline{r}=<1+t,-2+3t,-t>$ and $\underline{r}=<2+2s,3/2+3s,-2+4s>$. Find an equation of the plane containing the two lines. Also find the distance of the plane from the origin.

If the lines intersect they must have a common point. Therefore, equating the x- and y-coordinates, we have

$$\begin{aligned} 1+t &= 2+2s, \\ -2+3t &= \frac{3}{2}+3s. \end{aligned}$$

The solution is $t=4/3$, $s=1/6$, and satisfies the equality of the z-coordinates, $-t=-2+4s$. Thus the lines do intersect. The common point can be found by setting $t=4/3$ in the equation of the first line, to get the coordinates of the point as $(7/3,2,-4/3)$.

The direction vectors of the two lines are $\boldsymbol{u}=<1,3,-1>$ and $\boldsymbol{v}=<2,3,4>$. The normal vector to the plane containing the two lines is $\boldsymbol{n}=\boldsymbol{u}\times\boldsymbol{v}=<15,-6,-3>$. One point on the plane, corresponding to $t=0$ on the first line, has the position vector $\boldsymbol{r}_0=<1,-2,0>$. Therefore the equation of the plane is $(\boldsymbol{r}-\boldsymbol{r}_0)\cdot\boldsymbol{n}=0$, or

$$15(x-1)-6(y+2)-3z=0, \quad \text{or,} \quad 5x-2y-z=9.$$

The distance d of the plane from the origin is simply the projection of $\boldsymbol{r}_0$ along the unit normal $\boldsymbol{e}_n$ to the plane, *i.e.*,

$$d=|\boldsymbol{r}_0\cdot\boldsymbol{e}_n|=|\boldsymbol{r}_0\cdot\boldsymbol{n}/||\boldsymbol{n}|||=|<1,-2,0>\cdot<15,-6,3>/\sqrt{270}|=27/\sqrt{270}=\sqrt{2.7}.$$

Think about finding the distance between (i) a point and a line, (ii) a pair of parallel lines and (iii) a pair of parallel planes.

Example 2.6. Show that the following lines are skew (nonparallel and nonintersecting) and find the distance between them.

$$x=1+7t,\ y=3+t,\ z=5-3t \quad \text{and} \quad x=4-s,\ y=6,\ z=7+2s.$$

The direction vectors of the two lines are $\boldsymbol{u}=<7,1,-3>$ and $\boldsymbol{v}=<-1,0,2>$. If they are skew they must lie in parallel planes, the common normal to which is

$$\boldsymbol{n}=\boldsymbol{u}\times\boldsymbol{v}=<2,-11,1>.$$

The shortest distance between the two lines is the distance along the common normal. Consider the vector joining two arbitrary points, one on the first line, $A(1,3,5)$, and another on the second line, $B(4,6,7)$ given by $\boldsymbol{w} = \vec{AB} =< 3,3,2 >$. The required distance is the component of $\boldsymbol{w}$ along $\boldsymbol{n}$, *i.e.*,

$$D = \left| \boldsymbol{w} \cdot \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|} \right| = \left| \frac{< 3,3,2 > \cdot < 2,-11,1 >}{\| < 2,-11,1 > \|} \right| = \frac{25}{\sqrt{126}}.$$

3 Coordinate systems in $\mathcal{R}^3$

3.1 Rectangular or Cartesian coordinates

- Origin, three mutually perpendicular lines as axes, right-handed system, eight octants.
- Three coordinates are needed for a point, $P(x,y,z)$. The full space is covered by $-\infty < x, y, z < \infty$.
- The unit vector triad $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ is fixed; does not depend upon x, y or z. Position vector $\boldsymbol{r} = \vec{OP} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$.
- An equation of type $f(x,y,z) = 0$ defines a surface. Specifically, $x = x_0$, $y = y_0$, $z = z_0$ are planes. Other examples are $2x - 3y + z - 4 = 0$ (plane), $x^2 + y^2 - 1 = 0$ (cylinder), $x^2 + y^2 + z^2 - 36 = 0$ (sphere).

  Surfaces of revolution (obtained by revolving $z = f(y)$ about the z-axis) are of the form $z = f(\sqrt{x^2+y^2})$, such as $z = \sqrt{x^2+y^2}$, $z = x^2 + y^2$.

  Subregions are defined by inequalities. Examples are $y < 3$, $-4 \leq x \leq 4$, and $x^2 + y^2 + z^2 > 4$.
- Two simultaneous equations of type $f_1(x,y,z) = 0$, $f_2(x,y,z) = 0$ define a curve, the curve of intersection of the two surfaces defined by $f_1(x,y,z) = 0$, $f_2(x,y,z) = 0$. For example the equations $x^2 + y^2 - z = 0$ and $x + y + z - 2 = 0$, whose graphs are a paraboloid and a plane respectively, taken together represent the curve of intersection of the two surfaces.

  A segment of a curve can also be defined, parametrically, by $\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}$, $a \leq t \leq b$. The arc length of the curve is given by

  $$L = \int_a^b \left\| \frac{d\boldsymbol{r}}{dt} \right\| dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt.$$
- The plane (2-d world) is a special case: 4 quadrants. In the xy-plane an equation of type $f(x,y) = 0$ defines a curve. Examples are $x + y - 1 = 0$ (straight line), $x^2 + y^2 - 1 = 0$ (circle), $y - x^2 = 0$ (parabola).

  Inequalities define planar regions. Examples are $x + y - 1 \leq 0$, $x^2 + y^2 \leq 1$, $y \geq x^2$.

3.2 Cylindrical coordinates

- Typical point is $P(r, \theta, z)$. The full space is covered by $r \geq 0$, $0 \leq \theta < 2\pi$, $-\infty < z < \infty$.
- Position vector $\boldsymbol{r} = \vec{OP} = r\boldsymbol{e}_r + z\boldsymbol{e}_z$. Unit vector triad $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z)$ depends upon θ and therefore, can change as P moves.
- An equation of the form $f(r,\theta,z) = 0$ defines a surface. Specifically, $r = r_0$ is a right-circular cylinder, $\theta = \theta_0$ is a half-plane and $z = z_0$ is a plane. Other examples are $r = z$ (cone), $r^2 = z$ (parabolic cup), $z = r^2 \cos 2\theta$ (saddle) and $r = \theta$ (a spiral cylinder).

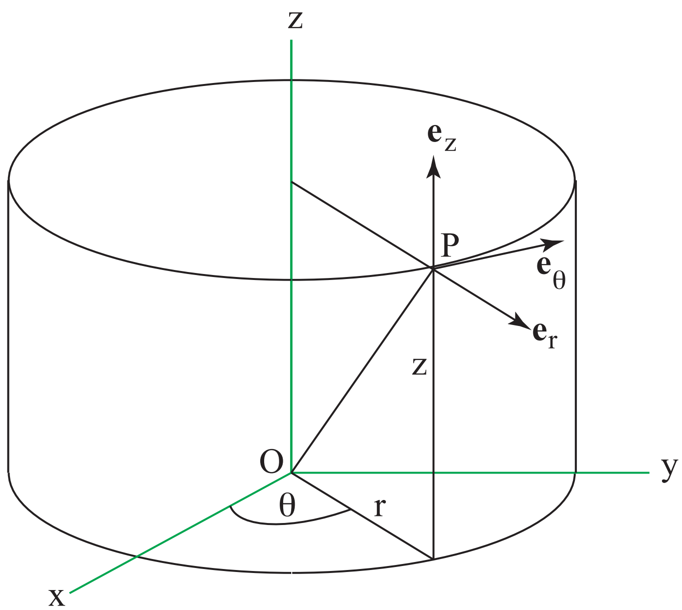

Figure 1: Cylindrical coordinates: r, θ, and z.

- Connection with cartesian coordinates: $x = r\cos\theta$, $y = r\sin\theta$, $z = z$; $r = \sqrt{x^2+y^2}$, $\theta = \tan^{-1}(y/x)$, $z = z$. Be careful about θ. For example the cartesian coordinates $(-\sqrt{3}, 1, 1)$ correspond to cylindrical coordinates $(2, 5\pi/6, 1)$ while the cartesian coordinates $(\sqrt{3}, -1, 1)$ correspond to cylindrical coordinates $(2, -\pi/6, 1)$.

- Connection with cartesian unit vector triad:

$$\begin{aligned} \boldsymbol{e}_r &= \cos\theta\,\boldsymbol{i} + \sin\theta\,\boldsymbol{j}, \\ \boldsymbol{e}_\theta &= -\sin\theta\,\boldsymbol{i} + \cos\theta\,\boldsymbol{j}, \\ \boldsymbol{e}_z &= \boldsymbol{k}. \end{aligned}$$

  We note from the above that

$$\frac{d\boldsymbol{e}_r}{d\theta} = \boldsymbol{e}_\theta, \quad \frac{d\boldsymbol{e}_\theta}{d\theta} = -\boldsymbol{e}_r. \tag{3.1}$$

- A segment of a curve is defined, parametrically, by $\boldsymbol{r}(t) = r(t)\boldsymbol{e}_r\big(\theta(t)\big) + z(t)\boldsymbol{e}_z$, $a \le t \le b$. The arc length of the curve is again given by

$$L = \int_a^b \left\| \frac{d\boldsymbol{r}}{dt} \right\| dt,$$

  but now, unlike the cartesian case, we must account for the fact that $\boldsymbol{e}_r$ is not a constant vector but depends upon θ which itself is a function of t on the curve. Thus,

$$\frac{d\boldsymbol{r}}{dt} = \frac{d}{dt}[r(t)\boldsymbol{e}_r(\theta) + z(t)\boldsymbol{e}_z] = \frac{dr}{dt}\boldsymbol{e}_r + r(t)\frac{d\boldsymbol{e}_r}{d\theta}\frac{d\theta}{dt} + \frac{dz}{dt}\boldsymbol{e}_z = \frac{dr}{dt}\boldsymbol{e}_r + r(t)\frac{d\theta}{dt}\boldsymbol{e}_\theta + \frac{dz}{dt}\boldsymbol{e}_z.$$

  Equation (3.1) has been employed in the last step above. Now the integral for the arc length becomes

$$L = \int_a^b \sqrt{\left(\frac{dr}{dt}\right)^2 + \left(r\frac{d\theta}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt.$$

3.3 Spherical coordinates

- Typical point is $P(\rho,\theta,\phi)$. The full space is covered by $\rho \geq 0,\ 0 \leq \theta < 2\pi,\ 0 \leq \phi \leq \pi$.
- Position vector $\boldsymbol{r} = \vec{OP} = \rho \boldsymbol{e}_\rho$. Unit vector triad $(\boldsymbol{e}_\rho,\ \boldsymbol{e}_\theta,\ \boldsymbol{e}_\phi)$ depends upon θ and ϕ and therefore, can change as P moves.
- An equation of the form $f(\rho,\theta,\phi)$ defines a surface. Specifically, $\rho = \rho_0$ is a sphere, $\theta = \theta_0$ is a half-plane and $\phi = \phi_0$ is a cone.

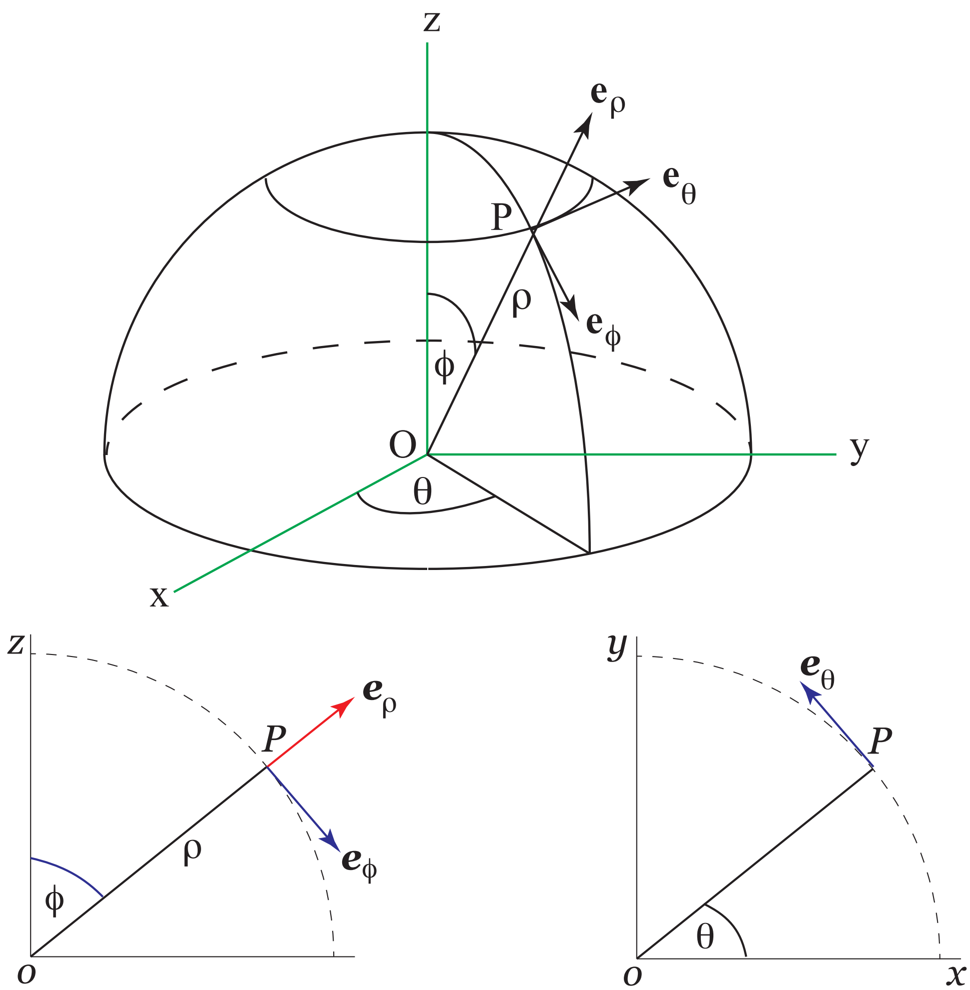

Figure 2: Spherical coordinates: radial coordinate ρ, azimuthal angle θ, and polar angle ϕ. Bottom left: side view. Bottom right: top view.

- Connection with cartesian coordinates: $x = \rho \sin\phi \cos\theta,\ y = \rho \sin\phi \sin\theta,\ z = \rho \cos\phi$.
  $\rho = \sqrt{x^2+y^2+z^2},\ \theta = \tan^{-1}(y/x),\ \phi = \cos^{-1}(z/\rho)$. Be careful about θ.
- Connection with cartesian unit vector triad:

$$\begin{aligned}
\underline{e}_\rho &= \cos\phi\ \underline{k} + \sin\phi(\cos\theta\ \underline{i} + \sin\theta\ \underline{j}), \\
\underline{e}_\phi &= -\sin\phi\ \underline{k} + \cos\phi(\cos\theta\ \underline{i} + \sin\theta\ \underline{j}), \\
\underline{e}_\theta &= -\sin\theta\ \underline{i} + \cos\theta\ \underline{j}.
\end{aligned}$$

We note that the spherical vector triad varies with the coordinates θ and ϕ, and that

$$
\begin{aligned}
\frac{\partial \underline{e}_\rho}{\partial \theta} &= \sin\phi(-\sin\theta\ \underline{i} + \cos\theta\ \underline{j}) = \boldsymbol{e}_\theta \sin\phi, \\
\frac{\partial \underline{e}_\rho}{\partial \phi} &= -\sin\phi\ \underline{k} + \cos\phi(\cos\theta\ \underline{i} + \sin\theta\ \underline{j}) = \boldsymbol{e}_\phi, \\
\frac{\partial \underline{e}_\phi}{\partial \theta} &= \cos\phi\ \boldsymbol{e}_\theta, \\
\frac{\partial \underline{e}_\phi}{\partial \phi} &= -\cos\phi\ \underline{k} - \sin\phi(\cos\theta\ \underline{i} + \sin\theta\ \underline{j}) = -\boldsymbol{e}_\rho, \\
\frac{d\boldsymbol{e}_\theta}{d\theta} &= -(\cos\theta\ \underline{i} + \sin\theta\ \underline{j}) = -\boldsymbol{e}_\rho \sin\phi - \boldsymbol{e}_\phi \cos\phi.
\end{aligned}
$$

- A segment of a curve is defined parametrically by $\boldsymbol{r} = \rho\boldsymbol{e}_\rho$, where $\rho = \rho(t)$ and $\boldsymbol{e}_\rho = \boldsymbol{e}_\rho\big(\theta(t), \phi(t)\big)$, $a \leq t \leq b$. Here we have recognized the dependence of $\boldsymbol{e}_\rho$ on θ and ϕ, which themselves are functions of t on the curve. Now the arc length is given by

$$
\begin{aligned}
L &= \int_a^b \left\| \frac{d\boldsymbol{r}}{dt} \right\| dt \\
&= \int_a^b \left\| \frac{d\rho}{dt}\boldsymbol{e}_\rho + \rho\frac{d\boldsymbol{e}_\rho}{dt} \right\| dt \\
&= \int_a^b \left\| \frac{d\rho}{dt}\boldsymbol{e}_\rho + \rho\left(\frac{\partial \boldsymbol{e}_\rho}{\partial\theta}\frac{d\theta}{dt} + \frac{\partial \boldsymbol{e}_\rho}{\partial\phi}\frac{d\phi}{dt}\right) \right\| dt \\
&= \int_a^b \left\| \frac{d\rho}{dt}\boldsymbol{e}_\rho + \rho\left(\frac{d\theta}{dt}\sin\phi\ \boldsymbol{e}_\theta + \boldsymbol{e}_\phi\ \frac{d\phi}{dt}\right) \right\| dt.
\end{aligned}
$$

Here we have used the expressions for the partial derivatives of the spherical unit vectors from the results derived above. The integral for L reduces further to

$$
L = \int_a^b \sqrt{\left(\frac{d\rho}{dt}\right)^2 + \left(\frac{d\theta}{dt}\right)^2 \rho^2 \sin^2\phi + \left(\frac{d\phi}{dt}\right)^2 \rho^2}\ dt. \tag{3.2}
$$

Example 2.7.

(a) Consider a solid sphere of radius R with its center at the origin. Describe the portion of the solid that lies below the xy-plane by using (i) spherical coordinates and (ii) cylindrical coordinates.

(b) Consider a solid object consisting of a right-circular cylinder of radius R and height H mounted on top of the hemisphere of part (a) above. Describe the solid in spherical and cartesian coordinates.

(a) In spherical coordinates the region is described by $0 \leq \rho \leq R$, $0 \leq \theta < 2\pi$, $\pi/2 \leq \phi \leq \pi$.
In cylindrical coordinates, $0 \leq r \leq R$, $0 \leq \theta < 2\pi$, $-\sqrt{R^2 - r^2} \leq z \leq 0$.

(b) In cartesian coordinates the solid is described by the inequalities

$$
-R \leq x \leq R,\ -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2},\ -\sqrt{R^2 - x^2 - y^2} \leq z \leq H.
$$

Let ϕ_0 be defined by

$$
\tan\phi_0 = \frac{R}{H}.
$$

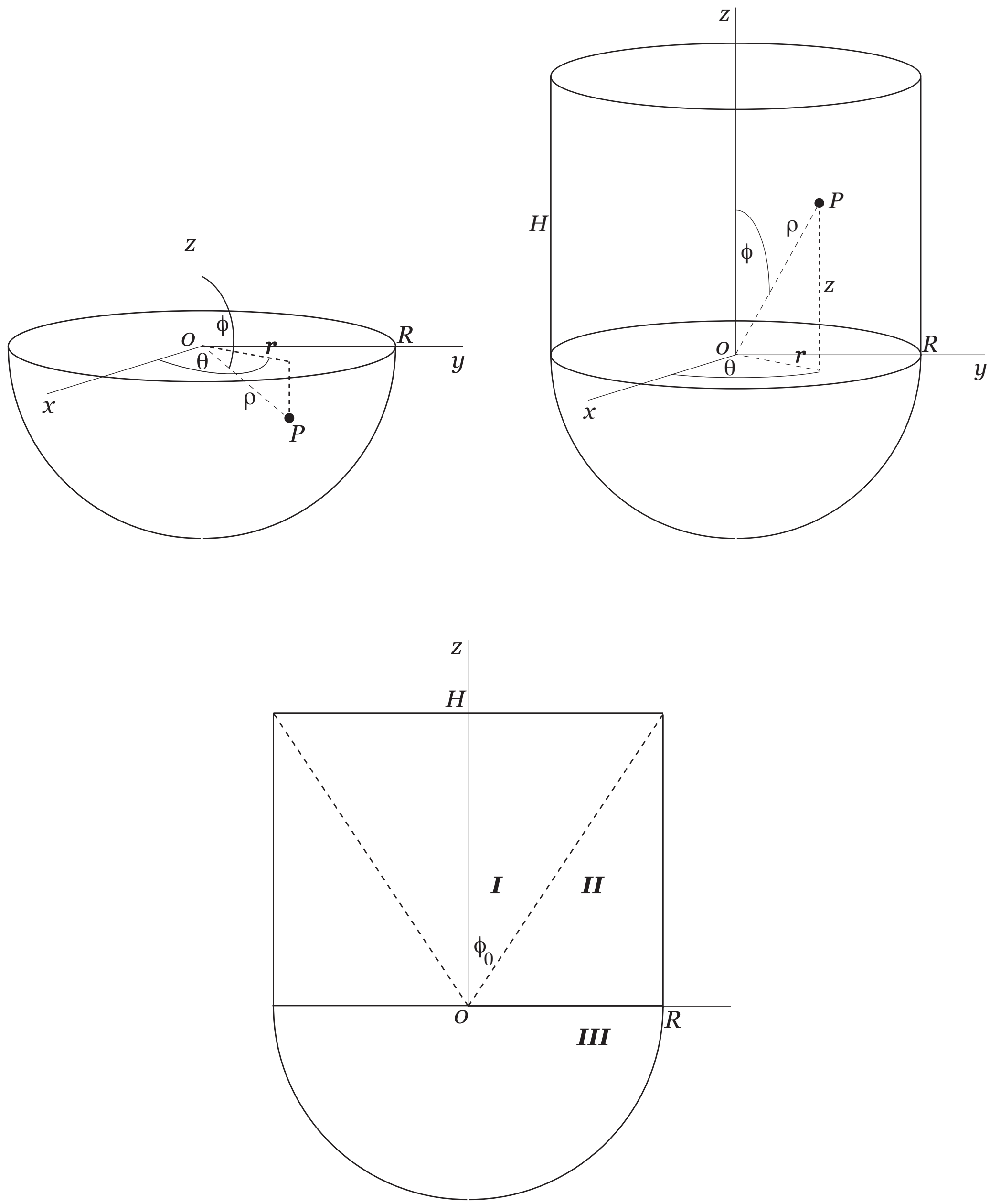

Figure 3: Objects (a), top left, and (b), top right, in Example 2.7. The bottom figure is a cross section of object (b) by a vertical plane through the z-axis, and shows the three segments of the object. Segment I is the cone of semi-vertex angle ϕ_0, segment II is the cylinder from which the cone has been carved out, and segment III is the hemisphere.

In spherical coordinates we need to subdivide the region into three segments. The first segment is the cone with its vertex at the origin and half-angle ϕ_0, described by

$$0 \le \phi \le \phi_0,\ 0 \le \rho \le \frac{H}{\cos\phi},\ 0 \le \theta < 2\pi.$$

The second segment is the cylindrical top from which the cone has been removed, and is described by

$$\phi_0 \le \phi \le \pi/2,\ 0 \le \rho \le \frac{R}{\sin\phi},\ 0 \le \theta < 2\pi.$$

The third segment is the hemisphere, described as in part (a) by

$$\pi/2 \le \phi \le \pi,\ 0 \le \rho \le R,\ 0 \le \theta < 2\pi.$$

Remark. If the cylinder is infinitely tall, *i.e.*, $H = \infty$, then $\phi_0 = 0$ and the conical segment described above disappears.

Example 2.8. Consider the sphere of radius R centered at the origin, with $A(R, \theta_0, \phi_0)$ and $B(R, \theta_1, \phi_1)$ two points on its surface. Let S be the path on the sphere connecting the two points and given by

$$S : \theta = \theta_0 + (\theta_1 - \theta_0)\frac{\phi - \phi_0}{\phi_1 - \phi_0}.$$

Show that the length of the path S is given by the integral

$$L = R\int_{\phi_0}^{\phi_1} \sqrt{1 + \left(\frac{\theta_1 - \theta_0}{\phi_1 - \phi_0}\right)^2 \sin^2\phi}\ d\phi.$$

We use ϕ as the parameter. Then the path is given by

$$\boldsymbol{r}(\phi) = R\boldsymbol{e}_\rho(\theta, \phi),\ \phi_0 \le \phi \le \phi_1,$$

where on the path S, we are given that

$$\theta = \theta_0 + (\theta_1 - \theta_0)\frac{\phi - \phi_0}{\phi_1 - \phi_0}.$$

The path length is given by

$$L = \int_{\phi_0}^{\phi_1} \left\| \frac{d\boldsymbol{r}}{d\phi} \right\| d\phi.$$

We now appeal to the integral (3.2) for arc length in spherical coordinates. Instead of t we use ϕ as the parameter. Since $t = \phi$ and $\rho = R$, a constant, the above integral reduces to

$$\begin{aligned} L &= \int_{\phi_0}^{\phi_1} \sqrt{\left(\frac{d\theta}{d\phi}\right)^2 R^2\sin^2\phi + R^2}\ d\phi \\ &= \int_{\phi_0}^{\phi_1} \sqrt{\left(\frac{\theta_1 - \theta_0}{\phi_1 - \phi_0}\right)^2 R^2\sin^2\phi + R^2}\ d\phi \\ &= R\int_{\phi_0}^{\phi_1} \sqrt{1 + \left(\frac{\theta_1 - \theta_0}{\phi_1 - \phi_0}\right)^2 \sin^2\phi}\ d\phi. \end{aligned}$$

A question we shall consider later in the course: how to find the curve with the shortest arc length (a geodesic) connecting two given points on the surface of a sphere?

4 Vectors in $\mathcal{R}^n$

Even though the direct geometrical interpretation is lost, algebraic operations on vectors extend naturally from $\mathcal{R}^3$ to $\mathcal{R}^n$, $n > 3$.

- For $\boldsymbol{x} =< x_1, x_2, \cdots, x_n >$ and $\boldsymbol{y} =< y_1, y_2, \cdots, y_n >$, the inner product (upon using the index notation, with i ranging from 1 to n) is

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_i y_i.$$

  Vanishing of the inner product implies orthogonality.

  Norm replaces length and is defined as

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{x_i x_i}.$$

- The notion of the coordinate frame is replaced by that of the *basis*. The natural basis consists of mutually orthogonal unit vectors $\boldsymbol{e}_j$, $j = 1, 2, \cdots, n$, where the j th entry of $\boldsymbol{e}_j$ is unity while all other entries are zero. We can write

$$\boldsymbol{x} = x_i \boldsymbol{e}_i.$$

- The concepts of orthogonal projection of vector $\boldsymbol{w}$ along the vector $\boldsymbol{v}$ carry over. Thus, with $\boldsymbol{e}_v$ being the unit vector along $\boldsymbol{v}$,

$$\text{comp } \boldsymbol{w}_v = \boldsymbol{w} \cdot \boldsymbol{e}_v, \quad \text{proj } \boldsymbol{w}_v = (\boldsymbol{w} \cdot \boldsymbol{e}_v)\boldsymbol{e}_v.$$

5 Functions and rates of Change

For $\boldsymbol{f} : \mathcal{R}^n \to \mathcal{R}^m$ we review the following concepts.

- Visualization
- Limit and continuity.
- Rate of change
- Smoothness
- Local approximation

5.1 Visualization

Let us focus on the case $m = 1$, *i.e.*, on scalar-valued multivariate functions. The *graph* of $f : A \subset \mathcal{R}^n \to \mathcal{R}$ is the set of points $\big(\boldsymbol{x}, f(\boldsymbol{x})\big)$ where $\boldsymbol{x} \in A$. Note that the graph is an element of $\mathcal{R}^{n+1}$.

- For $n = 1$ the graph is the set of points (x, y), with $y = f(x)$. Geometrically the set defines a curve in $\mathcal{R}^2$ (the xy-plane).
- For $n = 2$ the graph is the set of points (x, y, z), with $z = f(x, y)$. Geometrically the set defines a surface in $\mathcal{R}^3$. The function can also be visualized in $\mathcal{R}^2$ by level sets $f(x, y) = C$. The graphs of the level sets are known as contours, and can be thought of as cuts of the surface by horizontal planes at different elevations, projected to the $xy-$ plane.

  The surface can also be visualized in terms of cuts, or sections, by other (especially vertical) planes. The section of the graph of $f(x, y)$ by a plane P_1 is the set of points common to the plane and the graph. For example, the section of the graph of $f(x, y) = 100 - x^2 - y^2$ by the vertical plane $y = 8$ is the set of points common to the surface and the plane: $\{(x, y, z) \mid y = 8, z = 36 - x^2\}$. The section is a downward-opening parabola; it describes the behavior of the function when one of the independent variables is fixed.

Examples of common surfaces are $f(x,y) = -3x-4y+12$ (plane), $f(x,y) = 100-x^2-y^2$ (paraboloid, rotational symmetry), $f(x,y) = -x^2$ (cylinder). Review on your own the quadric surfaces: sphere, ellipsoid, paraboloid, hyperboloids, cone.

- For $n = 3$ the graph is the set of points $\big(x,y,z,u)\big)$ with $u = f(x,y,z)$. Visualization of this *hypersurface* is not possible as the graph lies in $\mathcal{R}^4$. The level sets (level surfaces in this case) are $f(x,y,z) = C$, which *can* be visualized in $\mathcal{R}^3$.
- The visual notions do not extend to $n > 3$.

5.2 Limit and continuity

The concept of the limit for $n = 1$ is fairly intuitive: $\lim_{x\to x_0} f(x) = L$ means that as x gets closer to x_0, $f(x)$ gets closer to L.

Formally, for any $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - x_0| < \delta$.

Extension to $n > 1$ is straightforward. The statement $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = L$ means that for any $\epsilon > 0$ there exists $\delta > 0$ such that $|f(\boldsymbol{x}) - L| < \epsilon$ whenever $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta$. Note that $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta$ is an open, deleted ball of radius δ centered at $\boldsymbol{x}_0$.

The function $f(\boldsymbol{x})$ is continuous at $\boldsymbol{x}_0$ if

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0).$$

5.3 Rates of change: differentiation

5.3.1 1-D case, $n = 1$

For $n = 1$ the change from x_0 to a neighboring point x can only occur along the x-axis. The rate of change is the derivative $f'(x_0)$ and $f(x)$ is said to be differentiable at x_0 if $f'(x_0)$ exists. The existence of the derivative implies the following.

- The graph of the linear function
$$L(x) = f(x_0) + (x - x_0)f'(x_0) \tag{5.1}$$
is tangent to the graph of f at x_0.
- Geometrically the graph of f is smooth at x_0 (no kink).
- For x close to x_0 the tangent line, or the linear function $L(x)$, provides a 'good' approximation to the graph of f. The notion of a good approximation is formalized as follows: consider the behavior of the difference $f(x) - L(x)$ as x approaches x_0.
$$\begin{aligned}\lim_{x\to x_0}\frac{f(x)-L(x)}{x-x_0} &= \lim_{x\to x_0}\frac{f(x)-\{f(x_0)+(x-x_0)f'(x_0)\}}{x-x_0}\\ &= \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} - \lim_{x\to x_0}\frac{(x-x_0)f'(x_0)}{x-x_0}\\ &= f'(x_0) - f'(x_0) = 0.\end{aligned}$$
We see that the difference $f(x) - L(x)$ approaches zero faster than $x - x_0$. It is in this sense that $L(x)$ provides a good approximation to $f(x)$ at x_0. We call $L(x)$ the *linear approximation* to $f(x)$ at x_0.

  As x_0 is incremented to $x_0 + dx$, the corresponding increment of the linear approximation, $L(x_0 + dx) - f(x_0) = f'(x_0)dx$, is denoted by df and called the *differential* of f.

- To summarize,

  existence of $f'(x_0) \Leftrightarrow$ existence of tangent line to graph of f at $x_0 \Leftrightarrow$ smoothness of the graph of f at $x_0 \Leftrightarrow$ goodness of the linear approximation to f at $x_0 \Leftrightarrow$ differentiability at x_0.

5.3.2 Multi-D case, $n > 1$

Directional derivative. The directional derivative of $f(\boldsymbol{x})$ at $\boldsymbol{x}_0$ in the direction of the *unit vector* $\boldsymbol{u}$ is defined as

$$D_u f(\boldsymbol{x}_0) = \lim_{h \to 0+} \frac{f(\boldsymbol{x}_0 + h\boldsymbol{u}) - f(\boldsymbol{x}_0)}{h}.$$

Note the one-sided nature of the directional derivative, as only positive values of h are admitted.

Partial derivatives. For $n > 1$ the change from $\boldsymbol{x}_0$ to a neighboring point $\boldsymbol{x}$ can occur along an infinity of directions, so that $f(\boldsymbol{x})$ can have an infinity of rates of change. First, consider the rates of change along the coordinate directions (or along the basis vectors). Along the basis vector $\boldsymbol{e}_1$, a neighboring point is of the form $\boldsymbol{x}_0 + h\boldsymbol{e}_1$. In other words, the first component of $\boldsymbol{x}$ changes from x_{1_0} to $x_{1_0} + h$ while the remaining components $x_{2_0}, x_{3_0}, \cdots, x_{n_0}$ remain fixed. The corresponding rate of change is the *partial derivative* of f with respect to x_1, given by

$$\frac{\partial f}{\partial x_1} = \lim_{h \to 0} \frac{f(\boldsymbol{x}_0 + h\boldsymbol{e}_1) - f(\boldsymbol{x}_0)}{h}.$$

There are n partial derivatives $\partial f/\partial x_j$, $j = 1, 2, \cdots, n$, which are defined in a similar fashion. In practice the partial derivatives are computed by differentiation with respect to the changing coordinate while the remaining coordinates are held fixed. Other notations for $\partial f/\partial x_j$ are f_{x_j} and $D_{x_j} f$ and, in index notation, $f_{,j}$.

For $n = 2$, *i.e.*, for $f(x, y)$, the partial derivative $f_x(x_0, y_0)$ is the slope of the tangent to the curve $z = f(x, y_0)$, while $f_y(x_0, y_0)$ is the slope of the tangent to the curve $z = f(x_0, y)$, both at $\big(x_0, y_0, f(x_0, y_0)\big)$.

Differentiability. For $n = 1$ we saw that the existence of the derivative, the existence of the tangent line, the smoothness of the graph and the existence of the linear approximation were all equivalent and any of them could have been used as the criterion for f to be differentiable. It turns out that for $n > 1$ the situation is more complex. To see this, consider the case $n = 2$ for which we can take advantage of the associated geometry. Assume that both partial derivatives of $f(x, y)$ exist at (x_0, y_0). If $f(x, y)$ were to have a tangent plane at (x_0, y_0), then the plane would have the equation $z = L(x, y)$ where

$$L(x, y) = f(x_0, y_0) + (x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0). \tag{5.2}$$

Note that L and f agree at (x_0, y_0), as do their partial derivatives. Note also that all quantities on the RHS of the above equation are well-defined. By analogy with the $n = 1$ case one may expect that visually, as one zooms into (x_0, y_0), the graph of $L(x, y)$, should provide an increasingly accurate approximation to that of $f(x, y)$, as shown schematically in Figure 4.

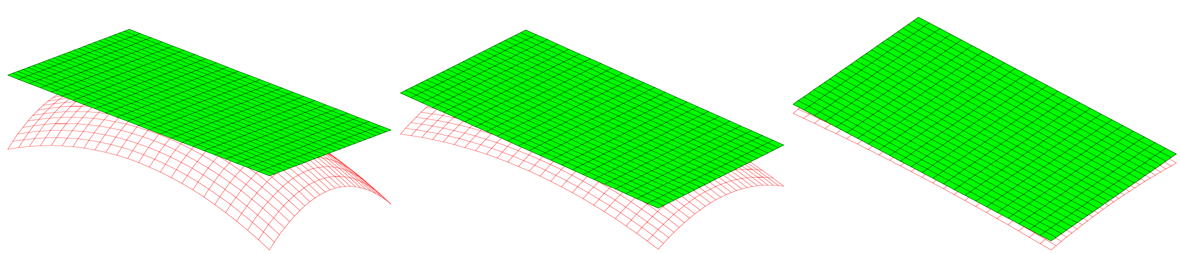

Figure 4: Graphs of $L(x, y)$ and $f(x, y)$.

The following counter-example shows that the above expectation is not always realized.

Example 5.1. Consider

$$f(x,y) = \begin{cases} \sqrt{xy}, & xy \geq 0, \\ 0, & xy < 0. \end{cases}$$

The function is identically zero in the second and fourth quadrants of the xy-plane. The graph is the surface $z = f(x,y)$, shown in Figure 5. The function vanishes on the lines $x = 0$ and $y = 0$. Therefore $f(0,0) = f_x(0,0) = f_y(0,0) = 0$. Hence the tangent plane at the origin, if it were to exist, would have the equation $z = 0$, which is just the equation for the xy-plane. However, the surface clearly has a kink at the origin, and hence a tangent plane at the origin does not exist.

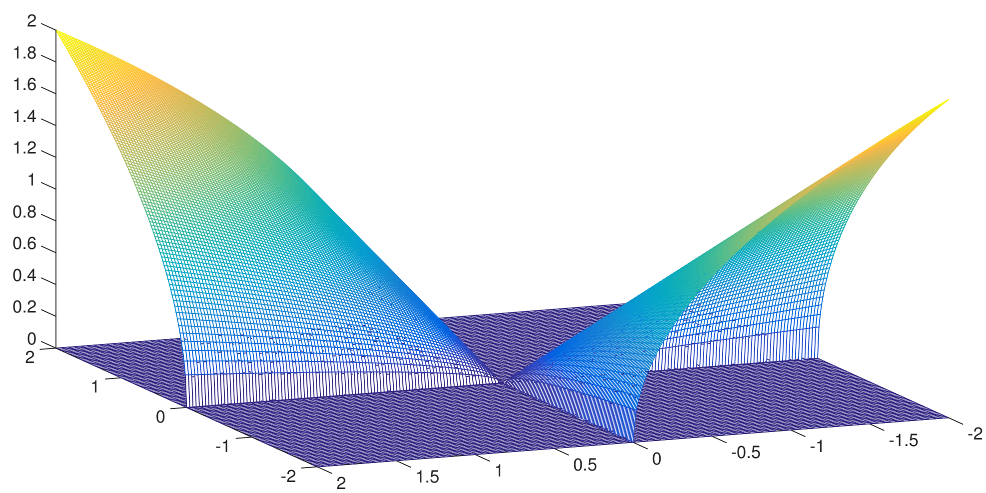

Figure 5: Graphs of $z = \sqrt{xy}$ and $z = 0$.

This example shows that the mere existence of partial derivatives does not suffice, so that we must choose a stronger criterion for differentiability. We do so by insisting that *for differentiability the linear approximation must be a good approximation,* in the sense that as $(x,y) \to (x_0, y_0)$, the difference $f(x,y) - L(x,y)$ must approach zero at a rate faster than (x,y) approaching (x_0,y_0), *i.e.*,

$$\lim_{\|(x,y)-(x_0,y_0)\| \to 0} \frac{f(x,y) - L(x,y)}{\|(x,y) - (x_0,y_0)\|} = 0.$$

Remarks.

- We now state without proof that if the above condition holds, then the partial derivatives $f_x(x_0,y_0)$ and $f_y(x_0,y_0)$ exist, and additionally, *are continuous.* In fact, *continuity of partial derivatives turns out to be necessary and sufficient for differentiability.*
- It is easy to see that the above condition is not satisfied at the origin in the example above (show it; we did so in class).

Example 5.2. For $f(x,y) = x^2y - 3xy^2$, find the tangent plane and the linear approximation at $(1,2)$.

We have $f_x = 2xy - 3y^2$, $f_y = x^2 - 6xy$. At $P(1,2)$, $f = -10$, $f_x = -8$, $f_y = -11$. Therefore the linear approximation at P, given by (5.2), is

$$L(x,y) = -10 - 8(x-1) - 11(y-2) = -8x - 11y + 20.$$

The equation of the tangent plane is $z = L(x,y)$.

The following example illustrates the utility of the linear approximation in practical computations.

Example 5.3. The radius of a right circular cylinder is measured with an error of at most 2%, and the height with an error of at most 4%. Approximate the maximum possible percentage error in the volume V calculated from these measurements.

The volume is given by $V = \pi R^2 H$, where R is the radius and H the height of the cylinder. Since changes in R and H from the exact values (R_0, H_0) are small, it makes sense to replace V by its linear approximation at (R_0, H_0), which is given by

$$\begin{aligned} V(R,H) &\approx L(R,H) = V(R_0,H_0) + (R - R_0)V_R(R_0,H_0) + (H - H_0)V_H(R_0,H_0) \\ &= V(R_0,H_0) + (R - R_0)2\pi R_0 H_0 + (H - H_0)\pi R_0^2. \end{aligned}$$

Let $R = R_0 + dR$ and $H = H_0 + dH$, and let us denote $V(R,H) - V(R_0,H_0)$ by dV. Then the above equation for V yields

$$dV \approx 2\pi R_0 H_0 dR + \pi R_0^2 dH.$$

Division by $V_0 = \pi R_0^2 H_0$ leads to

$$\frac{dV}{V_0} \approx 2\frac{dR}{R_0} + \frac{dH}{H_0}.$$

Let us assume that both the errors are positive. Then $dR/R_0 = 0.02$ and $dH/H_0 = 0.04$, so that

$$\frac{dV}{V_0} \approx 2(0.02) + 0.04 = 0.08.$$

Thus the maximum possible percentage error in the volume is 8%.

- The criterion for differentiability can now be generalized to arbitrary n. We define the linear approximation to $f(\boldsymbol{x})$ at $\boldsymbol{x}_0$ as

$$L(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \sum_{j=1}^{n} (x_j - x_{j_0}) \frac{\partial f}{\partial x_j}(\boldsymbol{x}_0),$$

and require that

$$\lim_{\|\boldsymbol{x}-\boldsymbol{x}_0\| \to 0} \frac{f(\boldsymbol{x}) - L(\boldsymbol{x}_0)}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0.$$

Note that the linear approximation can be written in terms of the gradient vector ∇f as

$$L(\boldsymbol{x}) = f(\boldsymbol{x}_0) + (\boldsymbol{x} - \boldsymbol{x}_0) \cdot \nabla f(\boldsymbol{x}_0), \tag{5.3}$$

and in terms of the index notation as

$$L(\boldsymbol{x}) = f(\boldsymbol{x}_0) + (x_i - x_{i_0}) f_{,i}(\boldsymbol{x}_0). \tag{5.4}$$

Upon contrasting (5.3) with its $n = 1$ counterpart (5.1) we note that the gradient vector ∇f for $n > 1$ plays the role of the derivative f' for $n = 1$.

- When f is differentiable, it can be shown that its directional derivative in the direction of any arbitrary unit vector $\boldsymbol{u}$ is completely determined by its partial derivatives (which are the components of the gradient) by the expression

$$D_u f = \boldsymbol{u} \cdot \nabla f = u_i f_{,i}.$$

Thus the directional derivative is the projection of ∇f in the direction of interest. The magnitude of ∇f gives the gives maximal rate of change, and its direction the direction of the maximal rate of increase. Also, ∇f is orthogonal to the level set $f = C$.

Example 5.4. Suppose $f(x,y,z) = x^2 + y^4 + x^2z^2$ gives the concentration of a chemical in a lake at the point (x,y,z), in units of grams/m^3. Suppose you are located at $P(-1,1,1)$. Assume that the unit of length is a meter.

(i) In which direction at P does the concentration (a) increase, and (b) decrease the fastest? What is the magnitude of the largest rate of increase or decrease?

(ii) Suppose you swim from P towards $Q(-2,3,5)$ at a speed of 2 meters per second. How fast does the concentration change?

(iii) Find an equation for the level surface of f through P. Also find an equation for the tangent plane to the level surface at P.

We begin by computing the gradient,

$$\nabla f =< 2x + 2xz^2, 4y^3, 2x^2z >,$$

and evaluate it at P, denoting the result by $\nabla f(P)$.

$$\nabla f(P) =< -4, 4, 2 > .$$

(i) The concentration increases fastest in the direction of $\nabla f(P)$ and decreases fastest in the direction of $-\nabla f(P)$. The magnitude of the fastest rate of increase or decrease is the magnitude of $\nabla f(P)$, given by

$$\|\nabla f(P)\| = \sqrt{16 + 16 + 4} = 6 \text{ grams/m}^4.$$

(ii) The vector from P to Q is $\vec{PQ} =< -1, 2, 4 >$, and the corresponding unit vector is

$$\boldsymbol{u}_{PQ} = \frac{1}{\sqrt{21}} < -1, 2, 4 > .$$

The rate of change of concentration with distance in the direction PQ is the directional derivative

$$D_{PQ}f(P) = \nabla f(P) \cdot \boldsymbol{u}_{PQ} =< -4, 4, 2 > \cdot \frac{1}{\sqrt{21}} < -1, 2, 4 >= \frac{20}{\sqrt{21}} \text{ grams/m}^4.$$

Multiplication with the speed of 2 m/s yields the rate of change with respect to time as

$$\frac{20}{\sqrt{21}} \times 2 = \frac{40}{\sqrt{21}} \text{ grams/(m}^3 s\text{)}.$$

(iii) At P, $f(x,y,z) = 3$. Therefore the desired level surface is $f(x,y,z) = x^2 + y^4 + x^2z^2 = 3$. The normal vector to the surface at P is $\boldsymbol{n} = \nabla f(P) =< -4, 4, 2 >$. The tangent plane at P is given by $(\boldsymbol{r} - \boldsymbol{r}_0) \cdot \boldsymbol{n} = 0$, *i.e.*,

$$-4(x+1) + 4(y-1) + 2(z-1) = 0, \quad \text{or,} \quad 2x - 2y - z + 5 = 0.$$

Remark. It is possible for f to have directional derivatives in all directions at a point but not be differentiable at that point. Think about standing at the sharp peak of a mountain.

Differentiability of $\boldsymbol{f} : \mathcal{R}^n \to \mathcal{R}^m$. The vector-valued function $\boldsymbol{f}$ is differentiable at a point if all the components f_i, $i = 1, 2, \cdots, m$, are differentiable at the point.

When referring to the partial derivatives of $\boldsymbol{f}$, it is customary to gather them all in an $m \times n$ matrix, known as the *Jacobian matrix*, denoted by $D\boldsymbol{f}$ and given below.

$$D\boldsymbol{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

Remarks.

- The entries of the i th row in the Jacobian matrix are just the components of the gradient ∇f_i, while the entries of the j th column are the components of the partial derivative $\partial \boldsymbol{f}/\partial x_j$.
- In index notation, $\{D\boldsymbol{f}\}_{ij} = f_{i,j}$.
- The linear approximation to $\boldsymbol{f}$ will now be a vector-valued function $\boldsymbol{L}(\boldsymbol{x})$, which can be written in matrix form as

$$\begin{bmatrix} L_1(\boldsymbol{x}) \\ L_2(\boldsymbol{x}) \\ \cdots \\ L_m(\boldsymbol{x}) \end{bmatrix} = \boldsymbol{f}(\boldsymbol{x}_0) + D\boldsymbol{f}(\boldsymbol{x}_0) \begin{bmatrix} x_1 - x_{1_0} \\ x_2 - x_{2_0} \\ \cdots \\ x_n - x_{n_0} \end{bmatrix}.$$

If $\boldsymbol{L}(\boldsymbol{x})$ and $\boldsymbol{x} - \boldsymbol{x}_0$ are treated as column vectors then the above equation can be written in an abbreviated form as

$$\boldsymbol{L}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_0) + D\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0). \tag{5.5}$$

This equation is of the same form as (5.1) and (5.3). Therefore, many texts refer to $D\boldsymbol{f}$ as the *derivative* of the multivariate vector-valued function $\boldsymbol{f}$. In index notation,

$$L_i(\boldsymbol{x}) = f_i(\boldsymbol{x}_0) + f_{i,j}(\boldsymbol{x}_0)(x_j - x_{j_0}). \tag{5.6}$$

5.4 The chain rule

The chain rule applies to the differentiation of a composite function and can be simply expressed as follows: *the derivative of a composition is the composition of the derivatives.*

In the simplest case, consider a function $f(x) : \mathcal{R} \to \mathcal{R}$ and a function $g(t) : \mathcal{R} \to \mathcal{R}$ such that the range of g is contained in the domain of f. Then the composite function $f \circ g : \mathcal{R} \to \mathcal{R}$ is defined as $(f \circ g)(t) = f\big(g(t)\big)$, and the chain rule becomes

$$(f \circ g)'(t) = f'\big(g(t)\big)\; g'(t).$$

It is common to write (with some abuse of notation) the above equation as

$$\frac{df}{dt} = \frac{df}{dx}\,\frac{dx}{dt}.$$

We restate it differently as

$$D(f \circ g)(t) = Df\big(g(t)\big)\; Dg(t), \tag{5.7}$$

using a notation for the derivative that was introduced earlier (see the statement following equation (5.5)).

Example 5.5. Consider the functions $\boldsymbol{g} : \mathcal{R} \to \mathcal{R}^3$ and $\boldsymbol{f} : \mathcal{R}^3 \to \mathcal{R}^2$, defined by

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x, y, z) &= \; < f_1(x, y, z), f_2(x, y, z) >, \\ \boldsymbol{g}(t) &= \; < x, y, z > = < g_1(t), g_2(t), g_3(t) > . \end{aligned}$$

Then the composite function is $\boldsymbol{f} \circ \boldsymbol{g} : \mathcal{R} \to \mathcal{R}^2$. We have

$$D\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{bmatrix},$$

and

$$Dg(t) = \begin{bmatrix} \frac{dx}{dt} \\ \\ \frac{dy}{dt} \\ \\ \frac{dz}{dt} \end{bmatrix} = \begin{bmatrix} \frac{dg_1}{dt} \\ \\ \frac{dg_2}{dt} \\ \\ \frac{dg_3}{dt} \end{bmatrix}.$$

According to the chain rule (5.7),

$$\begin{aligned} D(\boldsymbol{f}\circ\boldsymbol{g})(t) &= D\boldsymbol{f}\big(\boldsymbol{g}(t)\big)\, D\boldsymbol{g}(t) \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{dg_1}{dt} \\ \\ \frac{dg_2}{dt} \\ \\ \frac{dg_3}{dt} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x}\frac{dg_1}{dt} + \frac{\partial f_1}{\partial y}\frac{dg_2}{dt} + \frac{\partial f_1}{\partial z}\frac{dg_3}{dt} \\ \\ \frac{\partial f_2}{\partial x}\frac{dg_1}{dt} + \frac{\partial f_2}{\partial y}\frac{dg_2}{dt} + \frac{\partial f_2}{\partial z}\frac{dg_3}{dt} \end{bmatrix}. \end{aligned}$$

The above result extends in a straightforward and obvious way to the case $\boldsymbol{g}:\mathcal{R}^p \to \mathcal{R}^n$ and $\boldsymbol{f}:\mathcal{R}^n \to \mathcal{R}^m$. Now $D\boldsymbol{g}$ is an $n\times p$ matrix and $D\boldsymbol{f}$ is an $m\times n$ matrix so that the derivative $D(\boldsymbol{f}\circ\boldsymbol{g})$ is an $m\times p$ matrix.

Example 5.5, continued. Now we specialize the example above to the following functions:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x,y,z) &= < f_1(x,y,z), f_2(x,y,z) > = < x+y^2+z^3, xy^2z^3 >, \\ \boldsymbol{g}(t) &= < x,y,z > = < g_1(t), g_2(t), g_3(t) > = < t, t^2, t^3 > . \end{aligned}$$

Then

$$D\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{bmatrix} = \begin{bmatrix} 1 & 2y & 3z^2 \\ y^2z^3 & 2xyz^3 & 3xy^2z^2 \end{bmatrix},$$

$$D\boldsymbol{g}(t) = \begin{bmatrix} \frac{dg_1}{dt} \\ \\ \frac{dg_2}{dt} \\ \\ \frac{dg_3}{dt} \end{bmatrix} = \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix},$$

and

$$D\boldsymbol{f}\big(\boldsymbol{g}(t)\big) = \begin{bmatrix} 1 & 2t^2 & 3t^6 \\ t^{13} & 2t^{12} & 3t^{11} \end{bmatrix},$$

so that

$$\begin{aligned} D(\boldsymbol{f}\circ\boldsymbol{g})(t) &= D\boldsymbol{f}\big(\boldsymbol{g}(t)\big)\, D\boldsymbol{g}(t) \\ &= \begin{bmatrix} 1 & 2t^2 & 3t^6 \\ t^{13} & 2t^{12} & 3t^{11} \end{bmatrix} \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix} \\ &= \begin{bmatrix} 1+4t^3+9t^8 \\ 14t^{13} \end{bmatrix}. \end{aligned}$$

Example 5.6. Suppose that in a steady-state fluid flow,

$$\boldsymbol{v} =< -y,\ x,\ \sqrt{x^2+z^2} >$$

is the velocity of the fluid at position $< x, y, z >$. At what rate will $\boldsymbol{v}$ change as one moves from the point $(1, 0, -4)$ along the path $\boldsymbol{r} =< 2t+1, t^3, 6t-4 >$?

We have $\boldsymbol{v}(\boldsymbol{r})$ and $\boldsymbol{r}(t)$. According to the chain rule the t-derivative of the composite function $(\boldsymbol{v} \circ \boldsymbol{r})(t)$ is given by

$$(\boldsymbol{v} \circ \boldsymbol{r})'(t) = D\boldsymbol{v}(\boldsymbol{r})D\boldsymbol{r}(t).$$

We have

$$D\boldsymbol{v}(\boldsymbol{r}) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ \frac{x}{\sqrt{x^2+z^2}} & 0 & \frac{z}{\sqrt{x^2+z^2}} \end{bmatrix},$$

and

$$D\boldsymbol{r}(t) = \begin{bmatrix} 2 \\ 3t^2 \\ 6 \end{bmatrix}.$$

At $t = 0$, $\boldsymbol{r} =< 1, 0, 4 >$. Therefore,

$$D\boldsymbol{v}(\boldsymbol{r}) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ \frac{1}{17} & 0 & -\frac{4}{\sqrt{17}} \end{bmatrix},$$

and

$$D\boldsymbol{r}(t) = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix},$$

so that

$$(\boldsymbol{v} \circ \boldsymbol{r})'(0) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ \frac{1}{17} & 0 & -\frac{4}{\sqrt{17}} \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -22/\sqrt{17} \end{bmatrix}.$$

Appendix: the index notation

Component-wise algebraic computations can get quite laborious when every term has to be written out explicitly. The tedium is relieved by adopting the index notation which provides a convenient shorthand. Assume that a Cartesian reference frame has been chosen in $\mathcal{R}^3$ and that the three unit vectors along the positive coordinate axes are chosen as the basis vectors. Then any vector $\boldsymbol{a}$ can be represented as

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3,$$

or in terms of its components as

$$\boldsymbol{a} =< a_1, a_2, a_3 > .$$

In the index notation the symbol a_i is deemed sufficient to represent the vector $\boldsymbol{a}$ with the understanding that the index i can take any of the allowable values 1, 2 and 3, so that a_i in fact stands for the full set $[a_1, a_2, a_3]$. For α, β scalars, the symbol αa_i stands for the vector $\alpha\boldsymbol{a}$ and $\alpha a_i + \beta b_i$ for the linear combination $\alpha\boldsymbol{a} + \beta\boldsymbol{b}$. Symbolically we write

$$\alpha\boldsymbol{a} : \alpha a_i, \quad \alpha\boldsymbol{a} + \beta\boldsymbol{b} : \alpha a_i + \beta b_i,$$

and use the notation $\{\boldsymbol{a}\}_i = a_i$ to refer to the i th component of $\boldsymbol{a}$.

The matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is represented in the index notation by the symbol a_{ij}, with the understanding that the symbol represents the set whose nine elements are obtained by letting the indices i and j take all three allowable values, 1, 2 and 3.

Einstein introduced the convention that when an index appears twice in a term then the term represents the sum over all allowable values of the index. Thus the sum

$$S = \sum_{i=1}^{3} a_i b_i$$

is simply written as

$$S = a_i b_i.$$

Similarly we can replace $\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3$ by $\boldsymbol{a} = a_i\boldsymbol{e}_i$, and $a_{11} + a_{22} + a_{33}$ by a_{ii}.

Remarks.

- The repeating index is a dummy index and can be freely renamed. Thus,

$$a_i b_i = a_j b_j.$$

- An index cannot appear more than twice in a term. Thus $a_i b_i c_i$ does not make sense. If we wish to refer to the vector $(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}$, then the correct representation is $a_i b_i c_j$. The repeating index generates the sum $a_i b_i$ while the *free index* j indicates that we are referring to the vector whose j th component is $a_i b_i c_j$, *i.e.*,

$$\{(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}\}_j = a_i b_i c_j.$$

  Thus an index can either appear precisely once (a free index) or precisely twice (a repeating index).

- If one term in an equation contains a free index then all terms in the equation must contain the same free index. Thus the equation

$$\alpha a_i + \beta b_i = c_i$$

  makes sense but $\alpha a_i + \beta b_i = c_j$ does not.

Examples of index notation. All matrices in these examples are of order 3×3.

(i) The equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= y_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= y_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= y_3, \end{aligned}$$

have the index representation

$$a_{ij}x_j = y_i.$$

The free index i indicates that there are three equations, and the repeating index j indicates that the LHS of each equation is the sum of three terms.

(ii) The product of matrices $\boldsymbol{A} : a_{ij}$ and $\boldsymbol{B} : b_{ij}$, is represented economically as

$$\boldsymbol{AB} : a_{ik}b_{kj}.$$

The two free indices i and j indicate that the product generates nine terms, and the repeating index k indicates that each term, in turn, is itself a sum of three terms. Furthermore, the element in the i th row and j th column of the product $\boldsymbol{AB}$ is seen to be the product of the i the row of $\boldsymbol{A}$ and the j the column of $\boldsymbol{B}$. Similarly the product of three matrices is represented as $\boldsymbol{ABC} : a_{ik}b_{km}c_{mj}$. *The order of the indices is important.*

(iii) The quantity $\boldsymbol{x}^T\boldsymbol{Ax}$, where $\boldsymbol{x}$ is a column vector, has the index representation $a_{ij}x_ix_j$. This expression has no free index and is therefore a scalar, the double sum over the repeating indices i and j.

5.5 The Kronecker delta

The Kronecker delta, denoted by δ_{ij}, is defined as

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

The order of the indices can be changed, *i.e.*, $\delta_{ij} = \delta_{ji}$. Also, δ_{ij} can be viewed as the i, j component of the identity matrix $\boldsymbol{I}$, *i.e.*,

$$\{\boldsymbol{I}\}_{ij} = \delta_{ij}.$$

Remarks.

- The operation of identifying two indices and summing over them is called a *contraction*. Thus δ_{ii} is the contraction of δ_{ij}, expanded as

$$\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3.$$

- The basis vectors satisfy the relations

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

so that

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}.$$

Examples of the use of the Kronecker delta. The Kronecker delta appears in algebraic expressions frequently as a tool for selection or exchange of index, as shown in the following examples.

(i) Consider

$$\delta_{1i}a_{ij} = \delta_{11}a_{1j} + \delta_{12}a_{2j} + \delta_{13}a_{3j} = a_{1j}.$$

Note that the multiplier δ_{1i} simply replaces the index i in a_{ij} by 1 to yield a_{1j}.

(ii) Consider

$$\begin{aligned} a_ib_j\delta_{ij} &= a_1b_1\delta_{11} + a_1b_2\delta_{12} + a_1b_3\delta_{13} \\ &+ a_2b_1\delta_{21} + a_2b_2\delta_{22} + a_2b_3\delta_{23} \\ &+ a_3b_1\delta_{31} + a_3b_2\delta_{22} + a_3b_3\delta_{33} \\ &= a_1b_1 + a_2b_2 + a_3b_3. \end{aligned}$$

Thus we can write

$$a_ib_j\delta_{ij} = a_ib_i = a_jb_j,$$

where now δ_{ij} has replaced the index j in b_j by i or equivalently, the index i in a_i by j.

5.6 The alternating symbol

The alternating symbol, also known as the permutation symbol or the alternating tensor, and denoted by ϵ_{ijk}, is defined as

$$\epsilon_{ijk} = \begin{cases} 1, & ijk \text{ is an even permutation of 123,} \\ -1, & ijk \text{ is an odd permutation of 123,} \\ 0, & \text{two of the indices } ijk \text{ are the same.} \end{cases}$$

Thus,

$$\begin{aligned} \epsilon_{123} = \epsilon_{231} = \epsilon_{312} &= 1, \\ \epsilon_{132} = \epsilon_{321} = \epsilon_{213} &= -1, \\ \epsilon_{112} = \epsilon_{121} = \epsilon_{313} &= 0. \end{aligned}$$

Also,

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{kji} = -\epsilon_{ikj}.$$

Examples of the use of the alternating symbol.

(i) Recall that

$$\boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k) = \begin{cases} 1, & ijk \text{ is an even permutation of 123,} \\ -1, & ijk \text{ is an odd permutation of 123,} \\ 0, & \text{two of the indices } ijk \text{ are the same.} \end{cases}$$

Thus the alternating symbol is related to the scalar triple product of the basis vectors,

$$\boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k) = \epsilon_{ijk}.$$

(ii) The cross product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is given by

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = (a_2b_3 - a_3b_2)\boldsymbol{e}_1 + (a_3b_1 - a_1b_3)\boldsymbol{e}_2 + (a_1b_2 - a_2b_1)\boldsymbol{e}_3.$$

In index notation the above simplifies to $c_i = \epsilon_{ijk} a_j b_k$. This expression can also be derived directly by noting that if $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$, then

$$\begin{aligned} c_i &= \boldsymbol{e}_i \cdot \boldsymbol{c} \\ &= \boldsymbol{e}_i \cdot (\boldsymbol{a} \times \boldsymbol{b}) \\ &= \boldsymbol{e}_i \cdot (a_j \boldsymbol{e}_j \times b_k \boldsymbol{e}_k) \\ &= \boldsymbol{e}_i \cdot (\boldsymbol{e}_j \times \boldsymbol{e}_k) a_j b_k \\ &= \epsilon_{ijk} a_j b_k. \end{aligned}$$

Reversing the order of the vectors in the cross product,

$$\{\boldsymbol{b} \times \boldsymbol{a}\}_i = \epsilon_{ijk} a_k b_j = \epsilon_{ikj} a_j b_k = -\epsilon_{ijk} a_j b_k = -\{\boldsymbol{a} \times \boldsymbol{b}\}_i.$$

Note the interchange of j and k in the second step, allowed as j and k are dummy indices.

(iii) A further example of the economy resulting from the use of the index notation and the alternating symbol is the expression for the determinant $|\boldsymbol{A}|$ of a third-order matrix $\boldsymbol{A}$,

$$\begin{aligned} |\boldsymbol{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}$$

The above expression is succinctly represented by

$$|\boldsymbol{A}| = \epsilon_{ijk} a_{1i} a_{2j} a_{3k}, \quad \text{or equivalently,} \quad |\boldsymbol{A}| = \epsilon_{ijk} a_{i1} a_{j2} a_{k3}. \tag{5.8}$$

(iv) A useful identity connecting the Kronecker delta and the permutation symbol, known as the $\epsilon - \delta$ identity, is

$$\epsilon_{pmn} \epsilon_{pqr} = \delta_{mq} \delta_{nr} - \delta_{mr} \delta_{nq}. \tag{5.9}$$

(v) As an example of the ease of manipulation afforded by the use of the Kronecker delta, the alternating symbol and the $\epsilon - \delta$ identity, we derive the vector identity

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = (\boldsymbol{c} \cdot \boldsymbol{a}) \boldsymbol{b} - (\boldsymbol{c} \cdot \boldsymbol{b}) \boldsymbol{a}.$$

Beginning with the cross-product relation

$$\{\boldsymbol{a} \times \boldsymbol{b}\}_i = \epsilon_{ijk} a_j b_k,$$

we can write

$$\begin{aligned} \{(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}\}_i &= \epsilon_{iqr} \{\boldsymbol{a} \times \boldsymbol{b}\}_q c_r \\ &= \epsilon_{iqr} \epsilon_{qjk} a_j b_k c_r \\ &= -\epsilon_{qir} \epsilon_{qjk} a_j b_k c_r \\ &= -(\delta_{ij} \delta_{rk} - \delta_{ik} \delta_{rj}) a_j b_k c_r \\ &= -a_i b_k c_k + a_j b_i c_j \\ &= c_j a_j b_i - c_k b_k a_i \\ &= (\boldsymbol{c} \cdot \boldsymbol{a}) b_i - (\boldsymbol{c} \cdot \boldsymbol{b}) a_i. \end{aligned}$$