Multivariable Calculus

Jack DeSerrano

December 13, 2020

Contents

1 Derivatives of multivariable functions

1.1 The partial derivative

Partial derivatives are derivatives for multivariable functions. Fundamentally, derivatives and partial derivatives answer the same question: how does a little nudge in one direction affect the output? Rather than using d to denote the differential, we use ∂ to represent the partial differential. Herein, to denote the partial derivative of f with respect to x, we will write $\partial_x f$.

Example 1.1. Suppose
$$f(x,\, y) = x^2\, y + \sin y.$$
To take the partial derivative of f with respect to y, we treat x as a constant and compute the derivative normally:
$$\partial_y f = x^2 + \cos y.$$

The partial derivative has a formal definition. With a two-variable function f evaluated at a and b,
$$\partial_x f(a,\, b) \equiv \lim_{h \to 0} \frac{f(a + h,\, b) - f(a,\, b)}{h}.$$
Notice the similarity between this definition and the definition of the derivative.

If we take the partial derivative of a multivariable function with respect to x, and then take the partial derivative of our result with respect to y, we will, for all intents and purposes herein, wind up with the same result as taking the partial derivative with respect to y first, then x. Notationally, we write this as, for a function f,
$$\partial_{xy} f \equiv \partial_{yx} f.$$
This is known as the *symmetry of second partial derivatives.*

1.2 Gradient and the directional derivative

Example 1.2. Suppose
$$f(x,\, y) = x^2 \sin y.$$
Taking the partial derivatives of f with respect to x and y,
$$\partial_x f = 2x\, \sin y,$$
$$\partial_y f = x^2\, \cos y.$$

The *gradient* of f outputs a vector of the partial derivatives of f:
$$\nabla f = \begin{bmatrix} 2x\, \sin y \\ x^2\, \cos y \end{bmatrix}$$
where ∇, pronounced *nabla* or *del*, is the gradient operator.

For any function f with n dimensions and x_i representing the ith input of the function,

$$\nabla f \equiv \begin{bmatrix} f_{x_1} \\ f_{x_2} \\ \vdots \\ f_{x_n} \end{bmatrix}.$$

Therefore, think of the gradient operator as the full derivative; it captures all of the partial derivatives of a multivariable function.

The *directional derivative* of a function is the rate at which the function f changes at a point in the direction of a vector $\mathbf{v}$. This directional derivative is denoted $\nabla_{\mathbf{v}} f$. To calculate the directional derivative of a function f in the direction $\mathbf{v}$, we take the dot product of the vector and the function's gradient:

$$\nabla_{\mathbf{v}} f \equiv \mathbf{v} \cdot \nabla f.$$

Normalize $\mathbf{v}$ so that its magnitude is 1 when computing slope.

The directional derivative's formal definition is as follows, where $\mathbf{x}$ is the input point.
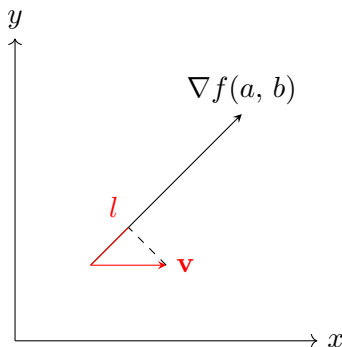
$$\nabla_{\mathbf{v}} f(\mathbf{x}) \equiv \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}.$$

From a graphical perspective, the gradient points in the direction of steepest ascent from a point.

Proof. The above proposition is equivalent to asking what unit vector $\mathbf{v}$ maximizes

$$\nabla f(a, b) \cdot \mathbf{v}$$

for the gradient of a function f at a point (a, b). Consider the figure below, which illustrates the above dot product.

Multiplying l by the magnitude of the gradient vector carries the same meaning as dotting the vector of choice with the gradient vector. If we maximize l, we maximize the dot product. The projected

length l depends on the direction of $\mathbf{v}$. If we change $\mathbf{v}$ to have the direction of the gradient, like so,
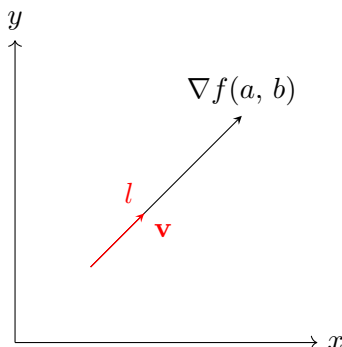
we find that $\mathbf{v}$'s direction maximizes l by the triangle inequality or simple observation. Therefore, this dot product will be maximized when $\mathbf{v}$ points in the direction of the gradient vector. Thus, the gradient always points in the direction of steepest ascent. $\qquad\square$

In general, the unit vector $\mathbf{u}$ maximizing $\mathbf{u} \cdot \mathbf{v}$ points in the same direction as $\mathbf{v}$.

1.3 The multivariable chain rule

Suppose we have a function $f(x,\,y)$ where x and y are parameterized by t. We find that

$$\frac{d}{dt}\,f(x,\,y) \equiv \partial_x f\,\frac{dx}{dt} + \partial_y f\,\frac{dy}{dt}.$$

Now suppose we define a vector-valued function $\mathbf{v}$ to contain the components of f. It is clear that the equality above is a dot product of a vector of the partial derivatives of f and a vector of ordinary derivatives with respect to t. We can, therefore, write the above equality as

$$\frac{d}{dt}\,f(\mathbf{v}) \equiv \nabla f \cdot \frac{d\mathbf{v}}{dt}.$$

This is known as the *multivariable chain rule*. It is resemblant of the single-variable chain rule.

Example 1.3. Suppose we have a two-variable function

$$f(x,\,y) = x^2\,y$$

where

$$x = \sin t,$$
$$y = \cos t.$$

4

We know that

$$\nabla f = \begin{bmatrix} 2\sin(t)\,\cos(t)\,\cos(t) \\ \sin(t)\,\sin(t) \end{bmatrix},$$

$$\frac{d\mathbf{v}}{dt} = \begin{bmatrix} \cos(t) \\ -\sin(t) \end{bmatrix}.$$

Therefore,

$$\frac{d}{dt}\,f(x,\,y) = \begin{bmatrix} 2\sin(t)\,\cos(t)\,\cos(t) \\ \sin(t)\,\sin(t) \end{bmatrix} \cdot \begin{bmatrix} \cos(t) \\ -\sin(t) \end{bmatrix}$$

$$= 2\sin(t)\,\cos(t)\,\cos(t)\,\cos(t) + \sin(t)\,\sin(t)\,\sin(t)$$

$$= \boxed{2\sin(t)\,\cos^2(t) + \sin^3(t)}$$

Look back to how we defined the multivariable chain rule. That definition is analogous to taking the directional derivative with vector $\frac{d\mathbf{v}}{dt}$ of f.

1.4   Curvature

For some point on a curve, the *radius of curvature*, denoted R, is the radius of the circle that best approximates the curve at that point. *Curvature*, denoted κ, is the reciprocal of radius of curvature. We can calculate curvature knowing that

$$\kappa \equiv \left|\frac{d\mathbf{T}}{ds}\right|,$$

where $\mathbf{T}$ is the unit tangent vector function and s is the arc length.

Example 1.4. Suppose

$$\mathbf{s}(t) = \begin{bmatrix} \cos(t)\,R \\ \sin(t)\,R \end{bmatrix}.$$

$\mathbf{s}$ represents a circle with radius R. In order to get the unit tangent vector, we must take the derivative,

$$\frac{d\mathbf{s}}{dt} = \begin{bmatrix} -\sin(t)\,R \\ \cos(t)\,R \end{bmatrix},$$

then normalize to get the unit tangent vector function:

$$\mathbf{T}(t) = \frac{d\mathbf{s}/dt}{|d\mathbf{s}/dt|}$$

$$= \frac{\begin{bmatrix} -\sin(t)\,R \\ \cos(t)\,R \end{bmatrix}}{\sqrt{\sin^2(t)\,R^2 + \cos^2(t)\,R^2}}$$

$$= \frac{\begin{bmatrix} -\sin(t)\,R \\ \cos(t)\,R \end{bmatrix}}{R}$$

$$= \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}.$$

Now that we have the unit tangent vector function, we can use the fact that

$$\left| \frac{d\mathbf{T}}{ds} \right| \equiv \frac{|d\mathbf{T}/dt|}{|d\mathbf{s}/dt|}$$

to determine curvature. Evaluating the numerator,

$$\left| \frac{d\mathbf{T}}{dt} \right| = \left| \frac{d}{dt} \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \right|$$

$$= \left| \begin{bmatrix} -\cos(t) \\ -\sin(t) \end{bmatrix} \right|$$

$$= 1.$$

we have already evaluated $\left| \frac{d\mathbf{s}}{dt} \right|$, so

$$\kappa = \frac{1}{R}.$$

This makes sense, since $\mathbf{s}$ defines a circle with radius R, and curvature is defined as the reciprocal of the radius of the circle that best fits the curve. We would follow this process for any vector-valued function of two dimensions.

Also,

$$\kappa \equiv \frac{|d\mathbf{s}/dt \times d^2\mathbf{s}/dt^2|}{|d\mathbf{s}/dt|^3},$$

where $\mathbf{s}$ is a vector valued function parameterized by t.

Example 1.5. Now suppose we want the curvature of a helix which is parameterized by

$$\mathbf{s}(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \\ t/5 \end{bmatrix}.$$

We would, again, begin by determining the unit tangent vector function of $\mathbf{s}$.

$$\mathbf{T}(t) = \frac{d\mathbf{s}/dt}{|d\mathbf{s}/dt|}$$

$$= \frac{\begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1/5 \end{bmatrix}}{\sqrt{\sin^2(t) + \cos^2(t) + 1/25}}$$

$$= \frac{\begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1/5 \end{bmatrix}}{\sqrt{26}/5}$$

$$= \begin{bmatrix} -5 \sin(t)\sqrt{26} \\ 5 \cos(t)\sqrt{26} \\ 1/\sqrt{26} \end{bmatrix}$$

We then differentiate $\mathbf{T}$ with respect to t.

$$\frac{d\mathbf{T}}{dt} = \begin{bmatrix} -5 \cos(t)/\sqrt{26} \\ -5 \sin(t)/\sqrt{26} \\ 0 \end{bmatrix}$$

Solving for the magnitude of our result,

$$\left| \begin{bmatrix} -5 \cos(t)/\sqrt{26} \\ -5 \sin(t)/\sqrt{26} \\ 0 \end{bmatrix} \right|$$

$$= \sqrt{25 \cos^2(t)/26 + 25 \sin^2(t)/26}$$

$$= \frac{5}{\sqrt{26}}.$$

We can bring everything together now.

$$\kappa = \frac{5/\sqrt{26}}{\sqrt{26}/5}$$

$$= \boxed{\frac{25}{26}}$$

1.5 Divergence

Divergence is an operator that describes how much vectors "diverge" from a certain point in a vector field. Often, we will imagine a vector field representing fluid flow. In that case, divergence

would represent whether the density of fluid particles increases or decreases at a point, and how much. We denote divergence of a vector field $\mathbf{F}$ as $\nabla \cdot \mathbf{F}$, which is pronounced *del dot F*.

The notation for divergence is resemblant of how we compute it. For a two-dimensional vector field $\mathbf{v}$ with components P and Q,

$$\nabla \cdot \mathbf{v} = \partial_x P + \partial_y Q.$$

An improper analogy of taking the divergence operator is dotting the vector of partial differential operators with the vector of which we are taking the divergence:

$$\begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} \cdot \begin{bmatrix} P(x,\, y) \\ Q(x,\, y) \end{bmatrix}.$$

To illustrate this in three dimensions, if

$$\mathbf{v}(x,\, y,\, z) = \begin{bmatrix} P(x,\, y,\, z) \\ Q(x,\, y,\, z) \\ R(x,\, y,\, z) \end{bmatrix},$$

then

$$\nabla \cdot \mathbf{v} = \partial_x P + \partial_y Q + \partial_z R.$$

1.6   Curl

Curl, in essence, is an operator that describes how vectors are rotating at a certain point in a vector field. In two dimensions, if that rotation is counterclockwise, curl is positive; if that rotation is clockwise, curl is negative. We denote curl of a vector field $\mathbf{F}$ as $\nabla \times \mathbf{F}$, which is pronounced *del cross F*.

For a two-dimensional vector field $\mathbf{v}$ with components P and Q,

$$\nabla \times \mathbf{v} = \partial_x Q - \partial_y P.$$

This is only two dimensions, though. We can describe rotation in two dimensions quite easily. We can give the rate of rotation as some radians/unit time, and give the direction of rotation by the rate being either positive or negative. That's all we need.

In three dimensions, however, there's more that we need. We need the axis around which everything is rotating and the rate at which everything is rotating. The axis and rate of rotation can be represented in a vector; however, just as in two dimensions, how do we represent the direction of rotation?

We use a straightforward right-hand rule to answer this question:

If we curl our hands in the direction of rotation, we will know the direction in which the vector representing the rotation will point. This means that we can describe any give rotation in three-dimensional space with a vector.

This means that, when taking the curl of a three-dimensional vector field, our output will be a vector field describing the rotation of each vector. Computing three-dimensional curl is where the notation—del cross—comes in handy.

Suppose

$$\mathbf{v}(x,\,y,\,z) = \begin{bmatrix} P(x,\,y,\,z) \\ Q(x,\,y,\,z) \\ R(x,\,y,\,z) \end{bmatrix}.$$

If we wanted to take the curl of $\mathbf{v}$, we would once again imagine crossing a vector of partial differential operators with $\mathbf{v}$. This is equivalent to imagining a determinant of a three-by-three matrix (which is only a notational trick, not mathematically proper):

$$\nabla \times \mathbf{v} = \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{pmatrix}$$

$$= \hat{\imath}\,(\partial_y R - \partial_z Q) - \hat{\jmath}\,(\partial_x R - \partial_z P) + \hat{k}\,(\partial_x Q - \partial_y P)$$

This is the curl of a three dimensional vector fields with components P, Q, and R. In column vector notation,

$$\nabla \times \mathbf{v} = \begin{bmatrix} \partial_y R - \partial_z Q \\ -\partial_x R + \partial_z P \\ \partial_x Q - \partial_y P \end{bmatrix}$$

Since this concept of three-dimensional curl is a little more challenging than the concepts of divergence and two-dimensional curl, it is best to go through an example.

Example 1.6. Suppose we want to take the curl of

$$\mathbf{v}(x,\,y,\,z) = \begin{bmatrix} xy \\ \cos z \\ z^2 + y \end{bmatrix}.$$

We start by taking the following determinant

$$\nabla \times \mathbf{v} = \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ xy & \cos z & z^2 + y \end{pmatrix}$$

$$= \hat{\imath}\,(\partial_y(z^2 + y) - \partial_z(\cos z)) - \hat{\jmath}\,(\partial_x(z^2 + y) - \partial_z(xy)) + \hat{k}\,(\partial_x(\cos z) - \partial_y(xy)),$$

and finish by computing partial derivatives:

$$\nabla \times \mathbf{v} = \boxed{\hat{\imath}\left(1 + \sin z\right) + \hat{k}\left(-x\right)} \text{ or}$$

$$= \boxed{\begin{bmatrix} 1 + \sin z \\ 0 \\ -x \end{bmatrix}}$$

1.7   The Laplacian

Suppose f is a two dimensional function. The *Laplacian* of f, denoted $\nabla^2 f$, outputs the divergence of the gradient of f.

$$\nabla^2 f \equiv \nabla \cdot \nabla f$$

We know that the gradient operator returns a vector field of the vectors of steepest ascent. In this gradient field, vectors tend to diverge more when the point is a minimum, and less when it is a maximum. This fact only makes sense, since all vectors will be converging to a maximum and diverging from a minimum. Therefore, you can think of the Laplacian operator as measuring how much of a minimum a point is. We can also conclude that the Laplacian is equivalent to

$$\nabla^2 f \equiv \sum_i \partial_{x_i x_i} f.$$

A function f is a *harmonic function* if and only if

$$\nabla^2 f \equiv 0.$$

1.8   The Jacobian

For a transformation L, recall that the transformation is linear if and only if it satisfies the following two properties:

$$L(a\,\mathbf{v}) = a\,L(\mathbf{v})$$
$$L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$$

Also recall that we can represent a linear transformation by multiplying a matrix by a vector, where the nth column of the matrix corresponds to the new vector representation of the nth basis vector. Multivariable calculus moves beyond the scope of linear transformations, though.

Suppose

$$\mathbf{r}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + \sin y \\ y + \sin x \end{bmatrix}.$$

This function is *locally linear*, which roughly and graphically means that if you zoom in on a certain point, the transformation will seem linear near that point. So, if we call some point undergoing this

transformation $\mathbf{x}$, then there should be a two-by-two matrix by which you multiply $\mathbf{x}$ to represent the linear transformation that seemed to have taken place. Though we will not go through the details of deriving it, the matrix with this property is called the *Jacobian matrix*. The Jacobian is also the matrix of $\mathbf{r}$'s first-order partial derivatives. Suppose we split $\mathbf{r}$ into two scalar-valued functions:

$$\left[\begin{array}{c} r_1(x,\, y) \\ r_2(x,\, y) \end{array}\right] = \left[\begin{array}{c} x + \sin y \\ y + \sin x \end{array}\right]$$

The Jacobian matrix of $\mathbf{r}$ is then

$$= \left[\begin{array}{cc} \partial_x r_1 & \partial_y r_1 \\ \partial_x r_2 & \partial_y r_2 \end{array}\right]$$

$$= \left[\begin{array}{cc} 1 & \cos y \\ \cos x & 1 \end{array}\right]$$

If you wanted to find that matrix that is analogous to the transformation near $\mathbf{x}$, then you would simply evaluate the Jacobian at $\mathbf{x}$.

We know that taking the determinant of any two-by-two matrix is akin to answering the question, "By what factor is a unit area changed after the transformation represented by that matrix?" If we take the determinant of the Jacobian, we call it, rather unsurprisingly, the *Jacobian determinant*. That tells us the factor by which a unit area changes in the neighbourhood of a certain point. So, if we wanted to determine (pun intended) the factor by which a unit area changes in the neighbourhood of $\mathbf{x}$, then we would take the determinant of $\mathbf{J_r}$ evaluated at $\mathbf{x}$.

2 Integrating multivariable functions

2.1 The line integral

Suppose $\mathbf{r}$ parameterizes a curve γ where $\mathbf{r}(a)$ and $\mathbf{r}(b)$ are the endpoints of γ. The *line integral* of γ is the area between γ and a surface described by field f. The line integral of γ along f is defined as

$$\int_\gamma f(x,\, y)\, ds \equiv \int_a^b f(\mathbf{r}(t)) \left|\frac{d\mathbf{r}}{dt}\right| dt.$$

Example 2.1. Suppose

$$f(x,\, y) = xy$$

and a curve C is parameterized by $\mathbf{r}$, where

$$\mathbf{r}(t) = \left[\begin{array}{c} \cos t \\ \sin t \end{array}\right]$$

and

$$0 \leqslant t \leqslant \frac{\pi}{2}.$$

11

If we wanted to determine the line integral of C along f, we simply use our definite integral definition and solve.

$$\int_C f(x, y) \, ds = \int_0^{\pi/2} xy \left| \frac{d\mathbf{r}}{dt} \right| dt$$

$$= \int_0^{\pi/2} xy \sqrt{(dx/dt)^2 + (dy/dt)^2} \, dt$$

$$= \int_0^{\pi/2} \cos(t) \sin(t) \sqrt{(-\sin(t))^2 + (\cos(t))^2} \, dt$$

$$= \int_0^{\pi/2} \cos(t) \sin(t) \, dt$$

If we let $u = \sin(t)$, then $du = \cos(t) \, dt$. Adjusting boundary conditions and evaluating,

$$\int_C f(\mathbf{r}) \, ds = \int_0^1 u \, du$$

$$= \frac{u^2}{2} \bigg|_0^1$$

$$= \boxed{\frac{1}{2}}$$

2.2   Line integrals and vector fields

Recall that *work* is defined as

$$W \equiv \mathbf{F} \cdot \mathbf{s}$$

where $\mathbf{F}$ is force, $\mathbf{s}$ is displacement. This is, of course, equivalent to

$$W \equiv Fs \, \cos \theta$$

where F and s are the magnitudes of the force and displacement, and θ is the angle between the direction of magnitude and force.

If $\mathbf{r}$ parameterizes some curve γ that, for instance, represents the path of a particle in a vector field $\mathbf{v}$, then the work done on the particle by the field is

$$W \equiv \int_\gamma \mathbf{v} \cdot d\mathbf{r}.$$

Example 2.2. Suppose $\mathbf{v}$ is a vector field where

$$\mathbf{v}(x, y) = y \, \hat{\imath} - x \, \hat{\jmath}$$

12

and a curve C, representing a particle's path, is parametrized by

$$x = \cos t,$$
$$y = \sin t$$

where
$$0 \leqslant t \leqslant 2\pi.$$

Let's say we want to determine the work done on the particle by the field. We start by determining $d\mathbf{r}$.

$$\frac{d\mathbf{r}}{dt} = -\sin(t)\,\hat{\imath} + \cos(t)\,\hat{\jmath}$$
$$d\mathbf{r} = -\sin(t)\,dt\,\hat{\imath} + \cos(t)\,dt\,\hat{\jmath}$$

We will then write $\mathbf{v}$ in terms of t.

$$\mathbf{v}(t) = \sin(t)\,\hat{\imath} - \cos(t)\,\hat{\jmath}$$

We now use our definite integral definition of the line integral.

$$
\begin{aligned}
\int_C \mathbf{v} \cdot d\mathbf{r} &= \int_0^{2\pi} -\sin^2(t)\,dt - \cos^2(t)\,dt \\
&= -\int_0^{2\pi} \left(\sin^2(t) + \cos^2 t \right)\,dt \\
&= -\int_0^{2\pi} dt \\
&= -t \,\Big|_0^{2\pi} \\
&= \boxed{-2\pi\ \text{J}}
\end{aligned}
$$

For a vector field $\mathbf{v}$ and two curves γ_1 and γ_2 on $\mathbf{v}$ with the same beginning and end points, the vector field is *path independent* or *conservative* if

$$\int_{\gamma_1} \mathbf{v} \cdot d\mathbf{r} \equiv \int_{\gamma_2} \mathbf{v} \cdot d\mathbf{r}.$$

Also, if $\mathbf{v}$ is the gradient of some scalar-valued function, then $\mathbf{v}$ is path independent.

A *closed line integral*, denoted $\oint$, is a line integral that integrates a closed loop. As a corollary of the definition of path independence, if a vector field is path independent, then, for any closed curve on that field,

$$\oint \mathbf{v} \cdot d\mathbf{r} \equiv 0.$$

Example 2.3. Evaluate

$$I = \oint_C \left(x^2 + y^2\right) \, dx + (2xy) \, dy$$

where C is parameterized by

$$x = \cos t,$$
$$y = \sin t$$

and

$$0 \leqslant t \leqslant 2\pi.$$

Let us define a vector field

$$\mathbf{v}(x, \, y) = \left(x^2 + y^2\right) \, \hat{\imath} + (2xy) \, \hat{\jmath}.$$

With that in mind,

$$I = \oint_C \mathbf{v} \cdot d\mathbf{r}$$

We can now determine whether or not $\mathbf{v}$ is conservative. If F is the scalar-valued function whose gradient is $\mathbf{v}$, then we must take the partial antiderivatives with respect to x and y.

$$\partial_x F = x^2 + y^2$$
$$F = \frac{x^3}{3} + xy^2 + f(y)$$

Take the antiderivative with respect to y.

$$\partial_y F = 2xy$$
$$F = xy^2 + g(x)$$

Align both results.

$$F(x, \, y) = \frac{x^3}{3} + xy^2$$

Since $\mathbf{v}$ is the gradient of some scalar-valued function, it is conservative, and therefore,

$$I = \boxed{0}$$

2.3 The double integral

The *double integral* is used to determine the area under a surface.

Example 2.4. Suppose
$$f(x, y) = xy^2$$
and we want to find the area under f for
$$0 \leqslant x \leqslant 2 \text{ and } 0 \leqslant y \leqslant 1.$$
We would use a double integral, like so,
$$\int_0^1 \int_0^2 xy^2 \, dx \, dy,$$
evaluating the inner integral first.
$$\int_0^1 \frac{x^2}{2} y^2 \bigg|_0^2 \, dy$$
$$= \int_0^1 2y^2 \, dy$$
$$= \frac{2y^3}{3} \bigg|_0^1$$
$$= \boxed{\frac{2}{3}}$$

It does not matter whether we integrate with respect to x or y first.

Example 2.5. Let
$$f(x, y) = xy^2.$$
Let's say we want to find the area under f for
$$0 \leqslant x \leqslant 1 \text{ and } y = x^2.$$
This would form some parabolic arc underneath the surface. To evaluate, we would set up our integral like so
$$\int_0^1 \int_0^{x^2} xy^2 \, dy \, dx,$$
then evaluate.
$$\int_0^1 \frac{xy^3}{3} \bigg|_0^{x^2} \, dx$$
$$= \int_0^1 \frac{x^7}{3} \, dx$$
$$= \frac{x^8}{24} \bigg|_0^1$$
$$= \boxed{\frac{1}{24}}$$

15

Always make sure your inner integral is the one with the variable boundary.

There is a more compact way to write a double integral. If the domain of the area you want to integrate is D and the area itself is A, then the double integral can be written as

$$\iint_D f(x,\, y)\, dA.$$

We can use polar coordinates to make integrations with rotational symmetry, like a disc or functions containing $x^2 + y^2$, a lot easier. A few conversions are necessary to utilize it:

$$x^2 + y^2 = r^2$$
$$dA = r\, d\theta\, dr$$

Example 2.6. Evaluate

$$I = \iint_{\mathbb{R}^2} e^{-\left(x^2 + y^2\right)}\, dA.$$

We know that

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left(x^2 + y^2\right)}\, dx\, dy.$$

Convert to polar coordinates, then evaluate the integrals.

$$I = \int_0^\infty \int_0^{2\pi} e^{-r^2} r\, d\theta\, dr$$
$$= \int_0^\infty e^{-r^2} r\, \theta \,\Big|_0^{2\pi}\, dr$$
$$= 2\pi \int_0^\infty e^{-r^2} r\, dr$$

If we let $u = -r^2$, then $-\frac{1}{2}\, du = r\, dr$.

$$I = -\frac{1}{2}\, 2\pi \int_0^\infty e^u\, du$$
$$= -\pi\, e^u \,\Big|_0^\infty$$
$$= \boxed{\pi}$$

2.4   The triple integral

The *triple integral* is the three-dimensional analogue of the double integral. It is worth asking what use the triple integral serves, since we can figure out three-dimensional areas using the double integral. The triple integral, however, can help us determine the mass of a certain three-dimensional figure if its density varies.

Example 2.7. Suppose we want to find the mass of the area bounded in the positive octant between a plane described by

$$2x + y + 3z = 6$$

and another plane described by

$$z = 2$$

that has a density function

$$\rho(x,\, y,\, z) = x^2 y z$$

in terms of kg/m^3. We need to determine boundaries for our integrals. We solve for dz's boundaries first. The bottom boundary is

$$3z = 6 - 2x - y$$
$$z = 2 - \frac{2x}{3} - \frac{y}{3}.$$

We also know the top boundary is $z = 2$, since that's where the second plane is. For dx's boundaries, we will need to see a projection of the volume in two dimensions. The lower boundary is 0 since we are bounded by the axis. The upper boundary must be in terms of y. We set $z = 0$ since dz's boundaries are the innermost, and determine the equation of the upper boundary.

$$2x + y = 6$$
$$x = 3 - \frac{y}{2}$$

dy's boundaries can only be numerical. Therefore, from the projection, its lower boundary is 0 and its upper boundary is 6. We can now set up the triple integral to determine the mass.

$$
\begin{aligned}
m &= \int_0^6 \int_0^{3-y/2} \int_{2-2x/3-y/3}^2 \rho(x,\, y,\, z)\, dz\, dx\, dy \\
&= \int_0^6 \int_0^{3-y/2} \int_{2-2x/3-y/3}^2 x^2 y z\, dz\, dx\, dy \\
&= \boxed{30.8571 \text{ kg}}
\end{aligned}
$$

2.5   The surface integral

The *surface integral* allows us to find the area of a surface; it is the double integral analogue of the line integral. To calculate a surface's surface integral, we must parameterize the surface.

Example 2.8. The parameterization of a torus is

$$\mathbf{r}(s,\, t) = ((b + a\,\cos(s))\,\sin(t))\,\hat{\imath} + ((b + a\,\cos(s))\,\cos(t))\,\hat{\jmath} + (a\,\sin(s))\,\hat{k}$$

where a is the radius of the torus' cross-sectional circles, where b is the distance from the torus' centre to the centre of the cross sections, and where

$$0 \leqslant s \leqslant 2\pi, \ 0 \leqslant t \leqslant 2\pi.$$

Determine the surface area of a torus Σ.

We will set up the surface integral as a double integral, like so

$$\iint_\Sigma d\Sigma.$$

Think of $d\Sigma$ as a patch of area on the surface. It has to move in two directions to capture all of the surface area (up/down and left/right in the case of the torus), hence the double integral. We will not derive another expression for $d\Sigma$ from the bottom up, but recall that the magnitude of the cross product of two vectors is the area of the parallelogram formed by the two vectors. We can think of $d\Sigma$ as the cross product of two vectors, specifically, the cross product of the partial derivatives of $\mathbf{r}$. We can then represent the surface integral above as

$$\iint_R |\partial_s\mathbf{r} \times \partial_t\mathbf{r}| \, ds \, dt$$

where R is the domain of $\mathbf{r}$. Let's begin with the partial derivative with respect to s.

$$\partial_s\mathbf{r} = -\left(a\,\sin(t)\sin(s)\right)\hat{\imath} - \left(a\,\cos(t)\sin(s)\right)\hat{\jmath} + \left(a\,\cos(s)\right)\hat{k}$$

Then with respect to t.

$$\partial_t\mathbf{r}\left((b + a\,\cos(s))\,\cos(t)\right)\hat{\imath} - \left((b + a\,\cos(t))\,\sin(t)\right)\hat{\jmath} + 0\,\hat{k}$$

We will now determine the cross product of these two partial derivatives.

$$
\begin{aligned}
\partial_s\mathbf{r} \times \partial_t\mathbf{r} = {} & \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ -a\sin(t)\sin(s) & -a\cos(t)\sin(s) & a\cos(s) \\ (b + a\cos(s))\cos(t) & -(b + a\cos(s))\sin(t) & 0 \end{pmatrix} \\
= {} & (a\cos(s)\cdot(b + a\cos(s))\sin(t))\,\hat{\imath} + (a\cos(s)\cdot(b + a\cos(s))\cos(t))\,\hat{\jmath} \\
& + (a\sin(t)\sin(s)\cdot(b + a\cos(s))\sin(t) + a\cos(t)\sin(s)\cdot(b + a\cos(s))\cos(t))\,\hat{k} \\
= {} & (b + a\cos(s))\left((a\cos(s)\sin(t))\,\hat{\imath} + (a\cos(s)\cos(t))\,\hat{\jmath} + \left(a\sin(s)\left(\sin^2(t) + \cos^2(t)\right)\right)\hat{k}\right) \\
= {} & (b + a\cos(s))\left((a\cos(s)\sin(t))\,\hat{\imath} + (a\cos(s)\cos(t))\,\hat{\jmath} + (a\sin(s))\,\hat{k}\right)
\end{aligned}
$$

We will now determine the magnitude of the above expression.

$$
\begin{aligned}
|\partial_s \mathbf{r} \times \partial_t \mathbf{r}| &= (b + a\cos(s))\sqrt{a^2 \cos^2(s)\sin^2(t) + a^2 \cos^2(s)\cos^2(t) + a^2 \sin^2(s)} \\
&= (b + a\cos(s))\sqrt{\left(a^2 \cos^2(s)\right)\left(\sin^2(t) + \cos^2(t)\right) + a^2 \sin^2(s)} \\
&= (b + a\cos(s))\sqrt{a^2 \left(\cos^2(s) + \sin^2(s)\right)} \\
&= (b + a\cos(s))\, a \\
&= ab + a^2 \cos(s)
\end{aligned}
$$

We can now put this into the double integral.

$$
\begin{aligned}
\iint_\Sigma d\Sigma &= \int_0^{2\pi} \int_0^{2\pi} |\partial_s \mathbf{r} \times \partial_t \mathbf{r}|\, ds\, dt \\
&= \int_0^{2\pi} \int_0^{2\pi} ab + a^2 \cos(s)\, ds\, dt \\
&= \int_0^{2\pi} abs + a^2 \sin(s) \Big|_0^{2\pi} dt \\
&= 2\pi \int_0^{2\pi} ab\, dt \\
&= 2\pi\, abt \Big|_0^{2\pi} \\
&= \boxed{4\pi^2 ab}
\end{aligned}
$$

Example 2.9. Determine

$$
I = \iint_\Sigma x^2 \, d\Sigma
$$

where Σ is the unit sphere.

The unit sphere is parameterized by

$$
\mathbf{r}(s,\, t) = (\cos(t)\cos(s))\,\hat{\imath} + (\cos(t)\sin(s))\,\hat{\jmath} + (\sin(t))\,\hat{k}
$$

where

$$
0 \leqslant s \leqslant 2\pi, \quad -\frac{\pi}{2} \leqslant t \leqslant \frac{\pi}{2}.
$$

Take the partial derivative with respect to s.

$$
\partial_s \mathbf{r} = -(\cos(t)\sin(s))\,\hat{\imath} + (\cos(t)\cos(s))\,\hat{\jmath} + 0\hat{k}
$$

Then with respect to t.

$$
\partial_t \mathbf{r} = -(\sin(t)\cos(s))\,\hat{\imath} - (\sin(t)\sin(s))\,\hat{\jmath} + (\cos(t))\,\hat{k}
$$

19

To save space, the steps in taking the cross product will not be shown. This is left as an exercise.

$$\partial_s \mathbf{r} \times \partial_t \mathbf{r} = \left(\cos^2(t)\cos(s)\right)\hat{\imath} + \left(\cos^2(t)\sin(s)\right)\hat{\jmath} + \left(\cos(t)\sin(t)\right)\hat{k}$$

Now, we take the magnitude of that result.

$$|\partial_s \mathbf{r} \times \partial_t \mathbf{r}| = \cos(t)$$

We now know that $d\Sigma = \cos(t)\,ds\,dt$. We must also express x^2 in terms of our parameters. Simple calculation shows that $x^2 = \cos^2(t)\cos^2(s)$. Let us wrap this up.

$$
\begin{aligned}
I &= \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \cos^2(t)\cos^2(s)\cos(t)\,ds\,dt \\
&= \int_{-\pi/2}^{\pi/2} \cos^3(t)\,dt \int_0^{2\pi} \cos^2(s)\,ds \\
&= \int_{-\pi/2}^{\pi/2} \cos(t) - \cos(t)\sin^2(t)\,dt \int_0^{2\pi} \frac{1}{2} + \frac{1}{2}\cos(2s)\,ds \\
&= \left(\sin(t) - \frac{\sin^3(t)}{3}\right)\Big|_{-\pi/2}^{\pi/2} \left(\frac{t}{2} + \frac{\sin(2s)}{4}\right)\Big|_0^{2\pi} \\
&= \boxed{\frac{4\pi}{3}}
\end{aligned}
$$

2.6   Flux

Suppose ρ is a three-dimensional function of mass density and $\mathbf{v}$ is a function of velocity. Let

$$\mathbf{F} = \rho\mathbf{v}$$

and suppose we have the surface integral over some surface Σ

$$\iint_\Sigma \mathbf{F} \cdot \mathbf{n}\,d\Sigma.$$

where $\mathbf{n}$ is the unit normal vector at each point. $\mathbf{F}$ is in terms of kg/m^{2}s and $d\Sigma$ is in terms of m^2, so the surface integral is in terms of kg/s; it answers the question "how much mass is travelling through Σ at a moment in time?" This is, essentially, *flux* through a two-dimensional surface.

We need to know how to construct a unit normal vector at any point on the surface. We assume that Σ can be parameterized by $\mathbf{r}(u, v)$. It is, quite simply,

$$\mathbf{n}(u,\,v) = \frac{\partial_u \mathbf{r} \times \partial_v \mathbf{r}}{|\partial_u \mathbf{r} \times \partial_v \mathbf{r}|}.$$

20

This makes sense; $\partial_u \mathbf{r}$ and $\partial_v \mathbf{r}$ are tangent to Σ, and the cross product will output a vector perpendicular to both—normal to Σ.

If we refer back to the surface integral at the beginning of this subsection, we can now determine that it is equivalent to the double integral

$$\iint_R \mathbf{F} \cdot (\partial_u \mathbf{r} \times \partial_v \mathbf{r}) \, du \, dv$$

where R is the domain of $\mathbf{r}$.

Though we have been uniform in our notation heretofore, another way to write the first surface integral is

$$\iint_\Sigma \mathbf{F} \cdot d\mathbf{\Sigma}.$$

3 Green's, Stokes', and divergence theorems

3.1 Types of regions in $\mathbb{R}^3$

A *type I region* R is defined as

$$\{(x, \, y, \, z) \,|\, (x, \, y) \in D, \, f_1(x, \, y) \leqslant z \leqslant f_2(x, \, y)\}$$

where D is the domain defined by R. This, essentially, means that if we take the cross section of R on the yz plane, each y would have at most two corresponding zs.

A *type II region* R is defined as

$$\{(x, \, y, \, z) \,|\, (y, \, z) \in D, \, f_1(y, \, z) \leqslant x \leqslant f_2(y, \, z)\}$$

where D is the domain defined by R. This means that if we take the cross section of R on the zx plane, each z would have at most two corresponding xs.

A *type III region* R is defined as

$$\{(x, \, y, \, z) \,|\, (x, \, z) \in D, \, f_1(x, \, z) \leqslant y \leqslant f_2(x, \, z)\}$$

where D is the domain defined by R. This means that if we take the cross section of R on the xy plane, each x would have at most two corresponding ys.

3.2 Green's theorem

Let C be a closed path and let $\mathbf{p}$ be a vector field such that

$$\mathbf{p}(x, \, y) = P(x, \, y)\hat{\imath}.$$

Suppose we want to determine

$$I_1 = \oint_C \mathbf{p} \cdot d\mathbf{r}.$$

We know that
$$I_1 = \oint_C P(x,\, y)\, dx$$
since there is no y component. Let a and b be the minimum and maximum x components of C respectively, and let y_1 and y_2 be functions of x, one of which describing the curve from a to b, and the other describing the curve from b to a. Therefore,

$$
\begin{aligned}
I_1 &= \int_a^b P(x,\, y_1(x))\, dx + \int_b^a P(x,\, y_2(x))\, dx \\
&= \int_a^b P(x,\, y_1(x)) - P(x,\, y_2(x))\, dx \\
&= -\int_a^b P(x,\, y_2(x)) - P(x,\, y_1(x))\, dx \\
&= -\int_a^b P(x,\, y)\,\Bigg|_{y=y_1(x)}^{y=y_2(x)}\, dx
\end{aligned}
$$

Assuming that P has a partial derivative,

$$I_1 = -\int_a^b \int_{y_1(x)}^{y_2(x)} \partial_y P\, dy\, dx.$$

We have, so far, concluded that

$$\oint_C P(x,\, y)\, dx \equiv -\iint_R \partial_y P\, dy\, dx,$$

where R is the region enclosed by C.

Let $\mathbf{q}$ be a vector field such that

$$\mathbf{q}(x,\, y) = Q(x,\, y)\hat{\jmath}.$$

Suppose we want to determine

$$I_2 = \oint_C \mathbf{q} \cdot d\mathbf{r}.$$

If we follow the same logic as we did originally, we will find that

$$I_2 = \oint_C Q(x,\, y)\, dy,$$

and eventually,

$$\oint_C Q(x,\, y)\, dy = \iint_R \partial_x Q\, dx\, dy.$$

Let
$$\mathbf{F}(x,\,y) = P(x,\,y)\hat{\imath} + Q(x,\,y)\hat{\jmath}$$
and let Ω be some closed path on $\mathbf{F}$. As a result of what we found previously,
$$\oint_{\Omega} \mathbf{F} \cdot d\mathbf{r} \equiv \iint_{R} \partial_x Q - \partial_y P \, dA$$
where $dA = dx\,dy$, where R is the region enclosed by Ω, and where the path of integration along Ω is counterclockwise. This is *Green's theorem.* As a corollary, if $\mathbf{F}$ is conservative, then
$$\partial_x Q - \partial_y P \equiv 0.$$

Example 3.1. Determine
$$I = \oint_C \left(x^2 - y^2\right)\,dx + 2xy\,dy$$
where C bounds a region R such that
$$0 \leqslant x \leqslant 1,\ 2x^2 \leqslant y \leqslant 2x.$$

By Green's theorem,

$$
\begin{aligned}
I &= \iint_R \partial_y \left(x^2 - y^2\right) - \partial_x \left(2xy\right)\,dA \\
&= \int_0^1 \int_{2x^2}^{2x} 4y\,dy\,dx \\
&= \int_0^1 2y^2 \Big|_{2x^2}^{2x}\,dx \\
&= \int_0^1 8x^2 - 8x^4\,dx \\
&= \frac{8x^3}{3} - \frac{8x^5}{5}\,\Big|_0^1 \\
&= \boxed{\frac{16}{15}}
\end{aligned}
$$

3.3 Stokes' theorem

Let Σ be a surface bounded by C and let $\mathbf{F}$ be a vector field in Σ's neighbourhood. *Stokes' theorem* says that
$$\oint_C \mathbf{F} \cdot d\mathbf{r} \equiv \iint_{\Sigma} (\nabla \times \mathbf{F}) \cdot d\mathbf{\Sigma}.$$

We can rewrite the right side of the equality to be

$$\iint_\Sigma (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, d\Sigma.$$

Suppose that

$$\mathbf{F} = P(x,\, y)\hat{\imath} + Q(x,\, y)\hat{\jmath}.$$

If we take the rewritten double integral before of $\mathbf{F}$, we find something interesting. Let C enclose a two-dimensional region R in three dimensions. Start with computing the curl (the steps of which are presented as an exercise).

$$\nabla \times \mathbf{F} = (\partial_x Q - \partial_y P)\, \hat{k}$$

Since $\hat{k}$ is the same as $\mathbf{n}$, we can say that

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \equiv \iint_R \partial_x Q - \partial_y P \, dA.$$

This equality shows how, fundamentally, Green's theorem and Stokes' theorem are related.

The normal vector of a surface can point in two directions; "inwards" or "outwards" relative to the surface. Similarly, the boundary of a surface can be made clockwise or counterclockwise. For Stokes' theorem to hold, these orientations have to be consistent. If you are familiar with the right hand rule described in the subsection on curl, that right hand rule describes the direction of the normal vector depending on the direction of integration and vice versa.

There are some conditions for Stokes' theorem. Let Σ be the surface over which we are integrating. Firstly, Σ must be *piecewise smooth*, which means that, if Σ is not continuous and differentiable everywhere, we can break Σ up into pieces that are continuous and differentiable everywhere. Let C be the boundary of Σ. C must be *simple*, meaning C does not intersect itself, closed, and piecewise smooth.

Example 3.2. Let C be the boundary formed by the intersection of the plane $y + z = 2$ and $x^2 + y^2 = 1$. Let

$$\mathbf{F} = -y^2\, \hat{\imath} + x\, \hat{\jmath} + z^2\, \hat{k}.$$

Determine

$$I = \oint_C \mathbf{F} \cdot d\mathbf{r}$$

using Stokes' theorem if C is integrated in a counterclockwise direction.

Let Σ be the region bounded by C. We know that

$$I = \iint_\Sigma (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, d\Sigma.$$

The steps of parameterizing C are an exercise. C can be parameterized by

$$\mathbf{s}(r,\,\theta) = r\cos\theta\,\hat{\imath} + r\sin\theta\,\hat{\jmath} + (2 - r\sin\theta)\,\hat{k}$$

where

$$0 \leqslant \theta \leqslant 2\pi,\ 0 \leqslant r \leqslant 1.$$

We can rewrite $\mathbf{n}\,d\Sigma$ as $(\partial_r\mathbf{s} \times \partial_\theta\mathbf{s})\,d\theta\,dr$. We can confirm that this cross product will point in the direction we need by the right hand rule for the cross product and the counterclockwise nature of C. Take the cross product.

$$\partial_r\mathbf{s} \times \partial_\theta\mathbf{s} = \det\begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \cos\theta & \sin\theta & -\sin\theta \\ -r\sin\theta & r\cos\theta & -r\cos\theta \end{pmatrix}$$

$$= r\,\hat{\jmath} + r\,\hat{k}$$

We now have to evaluate the curl of $\mathbf{F}$.

$$\nabla \times \mathbf{F} = \det\begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ -y^2 & x & z^2 \end{pmatrix}$$

$$= (1 + 2y)\,\hat{k}$$

$$= (1 + 2r\sin\theta)\,\hat{k}$$

We can now evaluate I.

$$I = \int_0^1 \int_0^{2\pi} \left((1 + 2r\sin\theta)\,\hat{k}\right) \cdot \left(r\,\hat{\jmath} + r\,\hat{k}\right)\,d\theta\,dr$$

$$= \int_0^1 \int_0^{2\pi} r + 2r^2\sin\theta\,d\theta\,dr$$

$$= \int_0^1 r\theta - 2r^2\cos\theta \,\Big|_0^{2\pi}\,dr$$

$$= \int_0^1 2\pi r\,dr$$

$$= \pi r^2 \,\Big|_0^1$$

$$= \boxed{\pi}$$

3.4 Divergence theorem

Let R be a simple three-dimensional region. Let Σ be the boundary of R. If $\mathbf{F}$ is a vector field in the neighbourhood of R, the flux through Σ is

$$\oiint_{\Sigma} (\mathbf{F} \cdot \mathbf{n}) \, d\Sigma.$$

We can think of this flux as a sum of divergences in R. By this statement, *divergence theorem* states that

$$\oiint_{\Sigma} (\mathbf{F} \cdot \mathbf{n}) \, d\Sigma \equiv \iiint_{R} (\nabla \cdot \mathbf{F}) \, dV.$$

Example 3.3. Suppose R is a simple three-dimensional region such that $-1 \leqslant x \leqslant 1$, $z = 1 - x^2$, and $y = 2 - z$. If

$$\mathbf{F} = \left(\frac{1}{2}x^2 + e^{\cos zy}\right) \hat{\imath} + (yx + \ln|z|) \hat{\jmath} + \tan(xy) \hat{k},$$

what is the flux of $\mathbf{F}$ through R?

We know from divergence theorem that the flux is nothing but

$$I = \iiint_{R} (\nabla \cdot \mathbf{F}) \, dV.$$

Calculate the divergence.

$$\nabla \cdot \mathbf{F} = 2x$$

We can now evaluate the triple integral.

$$
\begin{aligned}
I &= \int_{-1}^{1} \int_{0}^{1-x^2} \int_{0}^{2-z} 2x \, dy \, dz \, dx \\
&= \int_{-1}^{1} \int_{0}^{1-x^2} 2xy \Big|_{0}^{2-z} dz \, dx \\
&= \int_{-1}^{1} \int_{0}^{1-x^2} 2x(2 - z) \, dz \, dx \\
&= \int_{-1}^{1} 2x \left(2z - \frac{z^2}{2}\right) \Big|_{0}^{1-x^2} dx \\
&= \int_{-1}^{1} 3x - 2x^3 - x^5 \, dx \\
&= \frac{3x^2}{2} - \frac{x^4}{2} - \frac{x^6}{6} \Big|_{-1}^{1} \\
&= \boxed{0}
\end{aligned}
$$