Second Year Mathematics Revision

Calculus - Part 1

Ethan Brown, Kabir Agrawal

UNSW Mathematics Society

24 Apr 2019

Table of Contents

Functions in $\mathbb{R}^n$

So far we have mostly dealt with functions with one input, and one output variable. Now we expand our definitions to include those to and from $\mathbb{R}^n$.

Definition 1

A **curve** in $\mathbb{R}^n$ is a set of vectors where the components are dependent on a single variable - that is, a function from $\mathbb{R}$ to $\mathbb{R}^n$. They are generally described parametrically, as:

$$\mathbf{f}(t) = (f_1(t), f_2(t), \cdots, f_n(t))^T .$$

Note that each function $f_1, f_2, \cdots, f_n$ is a function from $\mathbb{R}$ to $\mathbb{R}$.

Generally you will deal with curves in $\mathbb{R}^2$ or $\mathbb{R}^3$.

Functions in $\mathbb{R}^n$

In $\mathbb{R}^3$, we can talk about surfaces.

Definition 2 (MATH2111)

A **surface** is the image of a subset $D \subseteq \mathbb{R}^2$ under a function $\mathbf{f} : D \to \mathbb{R}^3$.

Effectively, this means we have a 2-D area embedded in 3-D. We can describe such surfaces three ways:

1. **Explicitly**: As a graph in the form $z = f(x, y)$;
2. **Implicitly**: As all points satisfying $g(x, y, z) = 0$;
3. **Parametrically**: As points with components $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$.

Level Curves

Level curves (also called contour lines) can help us sketch surfaces in $\mathbb{R}^3$, by combining multiple 2-D objects to form a 3-D shape.

Definition 3

Suppose S is a surface. Then a **level curve** is a curve resulting from “slicing” the surface with a plane.

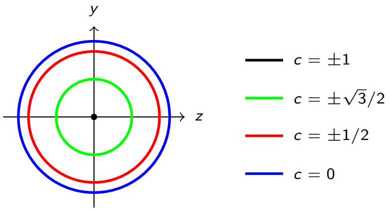
We can create a level curve by setting one component to be constant, say z, and solving for the other two.

Level Curves

Example 1

Find the level curves of $x^2 + y^2 + z^2 = 1$ for varying values of x.

Set $x = c$ to be constant, so $y^2 + z^2 = 1 - c^2$. Varying c we have:



Limits in $\mathbb{R}^n$

Limits in $\mathbb{R}^n$ are much the same as in $\mathbb{R}$, but instead of approaching from two sides, we can approach from many.

Definition 4

We write

$$\lim_{(x,y)\to(a,b)} f(x,y) = L,$$

and say the **limit of** $f(x,y)$ **as** (x,y) **approaches** (a,b) is L, if we can make $f(x,y)$ as close to L as we like, by taking (x,y) closer to (a,b), but not equal.

Limits in $\mathbb{R}^n$

Example 2

Show that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x}{x+y^2}$ does not exist.

We first approach on $x = 0$, so:

$$\lim_{(x,y)\to(0,0)} \frac{x}{x+y^2} = \lim_{y\to 0} \frac{0}{0+y^2} = 0.$$

Then, we approach on $y = 0$, so:

$$\lim_{(x,y)\to(0,0)} \frac{x}{x+y^2} = \lim_{x\to 0} \frac{x}{x+0^2} = 1.$$

Since the limits on each path do not agree, the limit does not exist.

Limits in $\mathbb{R}^n$

Regarding finding suitable paths

When trying to prove a limit does not exist, sometimes simple paths like $x = 0$ or $y = 0$ can't help you. In this case, look for other paths like $y = kx$ or $x^2 = y^3$.

Limits with mixed polynomials (like $\frac{x^2 y}{x^4 + y^2}$) can generally evaluate to non-zero limits along paths in the form $x^a = y^b$ ($y = x^2$ in this case). Simple paths like $x = 0$ tend to become zero, so being able to find these quickly is a valuable skill.

Continuity in $\mathbb{R}^n$

Continuity is much the same in $\mathbb{R}^n$ as it is in $\mathbb{R}$.

Definition 5

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **continuous at** $\mathbf{a} \in \mathbb{R}^n$ if $\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x})$ exists, and is equal to $f(\mathbf{a})$.

Definition 6

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **continuous on** $\Omega \subseteq \mathbb{R}^n$ if, for every $\mathbf{x} \in \Omega$, f is continuous at $\mathbf{x}$.

Partial Derivatives

As review of first year, we have the following definition of partial derivative.

Definition 7

The $\mathbf{k^{th}}$ **first partial derivative** of a function $f : \mathbb{R}^n \to \mathbb{R}$, denoted $\frac{\partial f}{\partial x_k}$, $\partial_k f$, $D_k f$, f_k, or f_{x_k}, is:

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_k) - f(\mathbf{x})}{h},$$

where $\mathbf{e}_k$ is the $\mathrm{k^{th}}$ standard basis vector for $\mathbb{R}^n$ (all zeroes except 1 in the $\mathrm{k^{th}}$ component).

We will use the notation $\frac{\partial f}{\partial x_k}$ or f_{x_k}.

Partial Derivatives

We write the second partial derivative as $\frac{\partial^2 f}{\partial x_i \partial x_j}$, $\partial_{ji} f$, $D_{ji} f$, f_{ji}, or $f_{x_j x_i}$.

Subscript Order

In the above, the partial derivatives are taken first with respect to x_j, then x_i. Take special care when dealing with the notation $\frac{\partial^2 f}{\partial x_i \partial x_j}$, as the order is different to the others - it is reversed. Think of it as composition of functions, where the right-most is applied first.

Partial Derivatives

Dealing with mixed partials can get annoying, so we have the following theorem.

Theorem 1 (Clairaut's Theorem)

Suppose $f : \mathbb{R}^2 \to \mathbb{R}$. If f, f_x, f_y, f_{xy} and f_{yx} all exist and are continuous on a disk around $\mathbf{a} \in \mathbb{R}^2$, then $f_{xy}(\mathbf{a}) = f_{yx}(\mathbf{a})$.

For those doing MATH2111, this can be generalised to $\mathbb{R}^n$, whereby $f_{ij}(\mathbf{a}) = f_{ji}(\mathbf{a})$ if the first and mixed second partials exist and are continuous on an open set around $\mathbf{a}$.
Since many of the functions we deal with are infinitely differentiable (e.g. sine, cosine, polynomials), the second partials will be continuous, so the mixed partials will be equal.

The Jacobian

Definition 8

If all the first partials of $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ exist at $\mathbf{a} \in \mathbb{R}^n$, then the **Jacobian matrix** of $\mathbf{f}$ at $\mathbf{a}$, written $J_{\mathbf{a}}\mathbf{f}$ or $J\mathbf{f}(\mathbf{a})$, is

$$J_{\mathbf{a}}\mathbf{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix},$$

where $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_m(\mathbf{x}))^T$.

Order of Jacobian

Remember that each column is composed of partials with respect to the same variable.

Differentiability in $\mathbb{R}^n$

Just like the derivative in $\mathbb{R}$ is a “good approximation” to $|f(x+h) - f(x)|$, the derivative in $\mathbb{R}^n$ is a “good approximation” to $||\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})||$. Specifically:

Definition 9 (MATH2011)

$\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at $\mathbf{a} \in \mathbb{R}^n$ if

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{||\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - D_{\mathbf{a}}\mathbf{f}(\mathbf{h})||}{||\mathbf{h}||} = 0,$$

where $D_{\mathbf{a}}\mathbf{f}$ is the linear map associated with $J_{\mathbf{a}}\mathbf{f}$,

$$D_{\mathbf{a}}\mathbf{f}(\mathbf{x}) = J_{\mathbf{a}}\mathbf{f} \cdot \mathbf{x}.$$

It can be shown that if all first partials exist and are continuous at $\mathbf{a}$, then the function is differentiable at $\mathbf{a}$.

Differentiability in $\mathbb{R}^n$ (MATH2111 only)

Definition 9 (MATH2111)

A function $\mathbf{f} : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at $\mathbf{a} \in \Omega$ if there exists a linear map $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

The matrix of $\mathbf{L}$ is called the **derivative** of $\mathbf{f}$ at $\mathbf{a}$, and written $D_{\mathbf{a}}\mathbf{f}$.

Theorem 2

If $\mathbf{f} : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{a}$ is an interior point of Ω, and $\mathbf{f}$ is differentiable at $\mathbf{a}$, then all first partials of the components of $\mathbf{f}$ exist at $\mathbf{a}$, and $D_{\mathbf{a}}\mathbf{f} = J_{\mathbf{a}}\mathbf{f}$.

Differentiability in $\mathbb{R}^n$ (MATH2111 only)

Example 3

Show that the function $\mathbf{f}(x, y) = (x^2 + y, x + y^2)$ is differentiable at $\mathbf{a} = (1, 0)$ with derivative $D = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$.

Substituting the values in and simplifying to numerator, we have

$$\begin{aligned}
\lim_{\mathbf{x}\to\mathbf{a}} \frac{||(x^2 - 2x + 1, y^2)||}{||(x-1, y)||} &= \lim_{\mathbf{x}\to\mathbf{a}} \sqrt{\frac{(x-1)^4 + y^4}{(x-1)^2 + y^2}} \\
&\le \lim_{\mathbf{x}\to\mathbf{a}} \sqrt{\frac{((x-1)^2 + y^2)^2}{(x-1)^2 + y^2}} \\
&= 0.
\end{aligned}$$

Since the limit cannot be negative, it is equal to zero.

Multivariable Chain Rule

Reviewing first year, we have the following theorem.

Theorem 3 (Chain Rule)

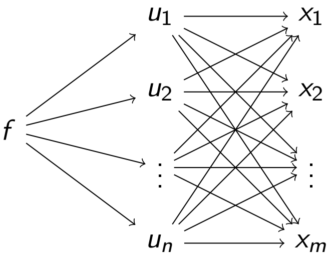
Suppose f is a function of $u_1, u_2, \cdots, u_n$, and each u_i is a function of $x_1, x_2, \cdots, x_m$. Then

$$\frac{\partial f}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial f}{\partial u_k} \frac{\partial u_k}{\partial x_i}.$$

This is the same chain rule taught in first year, but written in a more generalised form.

Multivariable Chain Rule

This general rule can be expressed as a chain diagram:



Multivariable Chain Rule

Example 4

Find $\frac{\partial f}{\partial y}$, where $f(u, v) = u^2 + v^2$, $u(x, y, z) = x^2 + y$, and $v(x, y, z) = z - y^2$.

So

$$\frac{\partial f}{\partial u} = 2u = 2(x^2 + y), \quad \frac{\partial f}{\partial v} = 2v = 2(z - y^2),$$

and

$$\frac{\partial u}{\partial y} = 1, \quad \frac{\partial v}{\partial y} = -2y.$$

Thus,

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = 2x^2 + 2y - 4yz + 4y^3.$$

Multivariable Chain Rule

Now we generalise this further, to the composition of non-scalar functions.

Theorem 4 (Chain Rule)

Suppose $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^p$. If $\mathbf{g}$ and $\mathbf{f}$ are differentiable at $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{g}(\mathbf{a})$ respectively, then so is $\mathbf{f} \circ \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^p$, and

$$J_{\mathbf{a}}(\mathbf{f} \circ \mathbf{g}) = J_{\mathbf{g}(\mathbf{a})}(\mathbf{f}) J_{\mathbf{a}}(\mathbf{g}).$$

So the derivative of a composite function is the product of each function's derivative, evaluated at the corresponding points.

Multivariable Chain Rule

Example 5 (Question 94 from MATH2011 2019)

Let $\mathbf{a} = (2,1,2)^T$ and suppose $\mathbf{f}(\mathbf{x}) = (x^2 - y^2, 2xy, z)^T$, $\mathbf{g}(\mathbf{x}) = \left(x + z^2, \frac{x}{z}\right)^T$. Find $J_\mathbf{a}(\mathbf{g} \circ \mathbf{f})$.

Note that $\mathbf{f}(\mathbf{a}) = (3,4,2)^T$. Now

$$J_\mathbf{x}(\mathbf{f}) = \begin{pmatrix} 2x & -2y & 0 \\ 2y & 2x & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad J_\mathbf{x}(\mathbf{g}) = \begin{pmatrix} 1 & 0 & 2z \\ \frac{1}{z} & 0 & -\frac{x}{z^2} \end{pmatrix}.$$

So, noting $J_\mathbf{a}(\mathbf{g} \circ \mathbf{f}) = J_{\mathbf{f}(\mathbf{a})}(\mathbf{g}) J_\mathbf{a}(\mathbf{f})$, we have

$$J_\mathbf{a}(\mathbf{g} \circ \mathbf{f}) = \begin{pmatrix} 1 & 0 & 4 \\ \frac{1}{2} & 0 & -\frac{3}{4} \end{pmatrix} \begin{pmatrix} 4 & -2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 4 \\ 2 & -1 & -\frac{3}{4} \end{pmatrix}.$$

Inverse Function Theorem

Theorem 5 (Inverse Function Theorem)

Suppose $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at $\mathbf{a} \in \mathbb{R}^n$. If $J_{\mathbf{a}}\mathbf{f}$ is an invertible matrix, then there is an inverse function $\mathbf{f}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ defined in some neighbourhood of $\mathbf{f}(\mathbf{a})$. Further,

$$J_{\mathbf{f}(\mathbf{a})}\left(\mathbf{f}^{-1}\right) = (J_{\mathbf{a}}\mathbf{f})^{-1}.$$

So, just like $\mathbb{R}$, the derivative of the inverse of a function is the inverse of the derivative of the function.

Notation Ambiguity

$J\mathbf{f}^{-1}$ is another way of writing $J\left(\mathbf{f}^{-1}\right)$. This can be confusing, since it could be interpreted as the inverse matrix of the Jacobian.

Inverse Function Theorem

Example 6

Let $\mathbf{f}(x, y) = (x^2 + y, y^2 + x)$ be differentiable everywhere. Find $J\mathbf{f}$, and hence where $\mathbf{f}$ is invertible. Find $J\mathbf{f}^{-1}$ at $(1, -1)$.

$$J\mathbf{f} = \begin{pmatrix} 2x & 1 \\ 1 & 2y \end{pmatrix}.$$

Then $|J\mathbf{f}| = 4xy - 1$. For $J\mathbf{f}$ to be invertible, we require non-zero determinant, so $4xy \neq 1$. Thus, by Inverse Function Theorem, $\mathbf{f}$ is invertible everywhere except the hyperbola $4xy = 1$. Then,

$$J_{(1,-1)}\mathbf{f}^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}^{-1} = -\frac{1}{5}\begin{pmatrix} -2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Implicit Function Theorem (MATH2111 only)

Just like Inverse Function Theorem lets us determine if an inverse function exists, we can formalise the problem of whether a system of implicit equations defines a function near a point. Denote our known variables $\mathbf{x} \in \mathbb{R}^m$ and our unknown variables $\mathbf{u} \in \mathbb{R}^n$. Then, we can expect a solution if we have n equations, $g_i(\mathbf{x}, \mathbf{u}) = 0$, where $g_i : \mathbb{R}^{m+n} \to \mathbb{R}$. Let $\mathbf{g} = (g_1, g_2, \cdots, g_n)^T$, and

$$D\mathbf{g} = \left(\begin{array}{cccc|cccc} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} & \frac{\partial g_1}{\partial u_1} & \frac{\partial g_1}{\partial u_2} & \cdots & \frac{\partial g_1}{\partial u_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m} & \frac{\partial g_2}{\partial u_1} & \frac{\partial g_2}{\partial u_2} & \cdots & \frac{\partial g_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_m} & \frac{\partial g_n}{\partial u_1} & \frac{\partial g_n}{\partial u_2} & \cdots & \frac{\partial g_n}{\partial u_n} \end{array}\right) = [A \mid B].$$

Implicit Function Theorem (MATH2111 only)

Using the previous definitions for $\mathbf{g}$, A, and B, we can state the following theorem.

Theorem 6 (Implicit Function Theorem)

Suppose that $(\mathbf{x}_0, \mathbf{u}_0)$ lies on the surface $\mathbf{g}(\mathbf{x}, \mathbf{u}) = \mathbf{0}$. If $B(\mathbf{x}_0, \mathbf{u}_0)$ is invertible, then there is a continuous function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\mathbf{g}\left(\mathbf{x}, \mathbf{f}(\mathbf{x})\right) = \mathbf{0}$ for all $\mathbf{x}$ in a neighbourhood of $\mathbf{x}_0$. Further,

$$D_{\mathbf{x}}\mathbf{f} = -\left(B\left(\mathbf{x}, \mathbf{f}(\mathbf{x})\right)\right)^{-1}\left(A\left(\mathbf{x}, \mathbf{f}(\mathbf{x})\right)\right).$$

This doesn't help us find $\mathbf{f}$ directly, but lets us find its derivative at a point and approximate it.
As long as you remember to evaluate the matrices at the right points, this can be remembered simply as

$$D\mathbf{f} = -B^{-1}A$$

.

Implicit Function Theorem (MATH2111 only)

Example 7 (Question 137 from MATH2111 2019)

Show that near $(0,1,2,1) = (x,y,u,v)$, the system

$$e^{xyu} + yuv + x - 3 = 0, \quad \ln yv + xu^3v - x^3u = 0$$

has a unique solution $(x, y, u(x,y), v(x,y))$.

So we calculate

$$D\mathbf{g} = \left(\begin{array}{cc|cc} yue^{xyu}+1 & xue^{xyu}+uv & xye^{xyu}+yv & yu \\ u^3v-3x^2u & \frac{1}{y} & -x^3 & \frac{1}{v}+xu^3 \end{array}\right).$$

At $(0,1,2,1)$, we have

$$\begin{pmatrix} xye^{xyu}+yv & yu \\ -x^3 & \frac{1}{v}+xu^3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$

which is invertible. Thus, Implicit Function Theorem holds.

Gradient

For a function $f : \mathbb{R}^n \to \mathbb{R}$, that is, a scalar function f, the Jacobian is a $1 \times n$ matrix, describing the gradient at a point of the corresponding surface.

Definition 10

The **gradient** of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\text{grad}(f) = \nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)^T,$$

where partials exist and are continuous.

This gradient vector is in the direction of maximal change. That is, moving in the direction of the gradient will change the function the most.

Directional Derivatives

Since we are no longer in one dimension, we can consider the derivative of a function from many directions, as follows.

Definition 11 (MATH2111)

The **directional derivative** of $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{a} \in \Omega$ in the direction $\hat{\mathbf{u}}$, where $\hat{\mathbf{u}}$ is a unit vector, is

$$D_{\hat{\mathbf{u}}}(f)(\mathbf{a}) = f'_{\hat{\mathbf{u}}}(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\hat{\mathbf{u}}) - f(\mathbf{a})}{h}.$$

Notation Matters

Beware the $D_{\hat{\mathbf{u}}}(f)(\mathbf{a})$ notation for directional derivatives. It is very similar to the standard derivative, and you may interpret $D_{\hat{\mathbf{u}}}(f)$ as the derivative of f evaluated at $\hat{\mathbf{u}}$. It is standard to use a hat, so $\hat{\mathbf{u}}$ instead of $\mathbf{u}$, to denote a unit vector. It is best to use $f'_{\hat{\mathbf{u}}}(\mathbf{a})$.

Directional Derivatives

We can relate the gradient and directional derivative as follows.

Theorem 7 for MATH2111 (Definition 11 for MATH2011)

The directional derivative of $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{a} \in \Omega$ in the direction $\hat{\mathbf{u}}$, where $\hat{\mathbf{u}}$ is a unit vector, is

$$f'_{\hat{\mathbf{u}}}(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \hat{\mathbf{u}}.$$

The directional derivative is a vector tangent to the corresponding surface in the direction specified. In this context, it is also called the **slope**.

Directional Derivatives

Example 8 (Question 90i from MATH2111 2018)

A skiier is at $(100, 1)$ on a mountain whose height is described by

$$h(x, y) = 2000 - \frac{x^4}{10^8} - \frac{y^2}{10^2}.$$

If they travel in the direction of steepest descent, what direction do they start moving in? What is the slope in this direction?

The gradient is $\nabla h(100, 1) = \left(-\frac{4}{100}, -\frac{2}{100}\right)$, which is in the direction $(-2, -1)$. Since we want descent, we note that increasing x and y will decrease $h(x, y)$, so the direction they move in is actually $(2, 1)$. The normalised vector is $\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$, so

$$\nabla h(100, 1) \cdot \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)^T = -\frac{1}{10\sqrt{5}}.$$

Tangent Planes

Gradients let us find tangents to surfaces, in the form of planes.

Definition 12

Suppose S is a surface defined implicitly as $f(x, y, z) = 0$, and $\mathbf{p} \in S$. Then the **tangent plane** of S at $\mathbf{p}$ is the plane through $\mathbf{p}$ with normal $\nabla f(\mathbf{p})$. That is, $\mathbf{x}$ satisfying

$$\nabla f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) = 0.$$

Note that the gradient is **normal** to the surface.

Tangent Planes

Example 9

Find the tangent plane to the graph of $f(x, y) = x^2 - y^2$ at $(1, 2)$.

Since we're dealing with a graph, we have $z = f(x, y)$, so $g(x, y, z) = x^2 - y^2 - z = 0$. Then at the point, $z = -3$, so we're considering $(1, 2, -3)$.

$$\nabla g(x, y, z) = (2x, -2y, -1)^T,$$
$$\nabla g(1, 2, -3) = (2, -4, -1)^T.$$

Thus, our plane is given by

$$(2, -4, -1)^T \cdot (x - 1, y - 2, z + 3)^T = 0.$$

In Cartesian form, this is $2x - 4y - z = -3$.

Normal Lines

Just as the gradient is normal to surfaces, it is normal to curves, too.

Example 10

Find the line normal to $x^2 + y^3 = 5$ at $(2, 1)$.

So we have $f(x, y) = x^2 + y^3 - 5 = 0$. Thus,

$$\nabla f(x, y) = (2x, 3y^2)^T,$$
$$\nabla f(2, 1) = (4, 3)^T.$$

So we want the line through $(2, 1)$ in the direction $(4, 3)$. Writing parametrically and eliminating the parameter gives

$$4y - 3x + 2 = 0.$$

Linear Approximation

Using tangent planes, we can develop an approximation method similar to that of one variable. Just as the tangent of a real function helps us approximate the function value, we can approximate multivariable functions using tangent planes.

Theorem 8

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has a tangent plane at $\mathbf{a} \in \mathbb{R}^n$, say P. Then P is the best approximation to f near $\mathbf{a}$.

In $\mathbb{R}^2$ this is

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

These approximations are first degree Taylor polynomials, as we will see soon.

Linear Approximation

Example 11

Approximate $f(x,y) = x^2 - y^2$ at $(1.05, 1.9)$. What is the absolute and relative error?

From a previous example, the tangent plane is $2x - 4y - z = -3$. Simply substituting, we find

$$f(1.05, 1.9) \approx 3 + 2(1.05) - 4(1.9) = -2.5.$$

The actual value is $f(1.05, 1.9) = -2.5075$, so this approximation has absolute error 0.0075, and relative error of

$$\left| \frac{0.0075}{-2.5075} \right| \approx 0.3\%.$$

Differentials

To find maximum error in an approximation, we calculate differentials.

Definition 13

Suppose $z = f(x, y)$. Then the **differential** $\mathrm{d}z$ is

$$\mathrm{d}z = \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y.$$

This can be generalised to $\mathbb{R}^n$ by adding more terms.
The differential is a measure of the absolute error of an approximation, so $\mathrm{d}x$ and $\mathrm{d}y$ are errors in x and y, and the partials are evaluated at the measured values.

Differentials

Example 12 (Question 75 from MATH2011 2019)

Let $T = x\cosh y$. If x is measured to be 2, and y to be $\ln 2$, with maximum error ± 0.04 and ± 0.02 respectively, find the maximum error in T.

So we have $\mathrm{d}x = 0.04$ and $\mathrm{d}y = 0.02$. Then $T_x(2, \ln 2) = 1.25$ and $T_y(2, \ln 2) = 1.5$. Thus,

$$\mathrm{d}T = 1.25 \times 0.04 + 1.5 \times 0.02 = \frac{8}{100}.$$

That is, the maximum error in the measured value of T is 0.08.

Taylor Series

Just like Taylor series from first year, we can approximate multivariable functions with polynomials.

Theorem 9 (Taylor's Theorem)

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be sufficiently differentiable near $\mathbf{a} \in \mathbb{R}^2$. Then the n^{th} **degree Taylor polynomial** of f around (a, b) is

$$\begin{aligned} P(x,y) &= f(a,b) + \frac{1}{1!}\left(f_x \cdot (x-a) + f_y \cdot (y-b)\right) \\ &+ \frac{1}{2!}\left(f_{xx} \cdot (x-a)^2 + 2 f_{xy} \cdot (x-a)(y-b) + f_{yy} \cdot (y-b)^2\right) \\ &+ \cdots + \frac{1}{n!}(\text{terms involving } n^{\text{th}} \text{ partials}), \end{aligned}$$

where all partials are evaluated at (a, b).

Taylor Series

Example 13 (Question 77b from MATH2011 2019)

Find the 5[th] degree Taylor polynomial of $\sin(x+y)$ at $(0,0)$.

We must first find up to fifth partials at $(0,0)$:

$$f = 0, f_{x_i} = 1, f_{x_i x_j} = 0, f_{x_i x_j x_k} = -1, f_{x_i x_j x_k x_l} = 0, f_{x_i x_j x_k x_l x_m} = 1.$$

Then the polynomial is:

$$P(x,y) = x + y + \frac{1}{3!}\left(-x^3 - 3x^2y - 3xy^2 - y^3\right) + \frac{1}{5!}\left(x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5\right).$$

Taylor Series

Coefficients in Taylor Series

The red highlighted numbers are very important. They correspond to binomial coefficients, so for a fourth order term, they would be 1, 4, 6, 4, 1. This is because, for almost every function you will deal with, the partial derivatives commute, so the separate terms for f_{xy} and f_{yx} can be combined. For higher order terms, this pattern persists.

Taylor Series

The **remainder** of a Taylor series is the next term (so for the previous example, the 6$^{\text{th}}$ order term), but instead of evaluating the partials at the point, we evaluate them at a point $\mathbf{z}$ on the line segment between (a, b) and (x, y).

Example 14

What is the form of the remainder term of a 2$^{\text{nd}}$ degree Taylor polynomial of two variables around $(0, 0)$?

So the remainder term is

$$\frac{1}{3!}\left(f_{xxx}x^3 + 3f_{xxy}x^2y + 3f_{xyy}xy^2 + f_{yyy}y^3\right),$$

where the partial derivatives are evaluated at some $\mathbf{z}$ between $(0, 0)$ and (x, y).

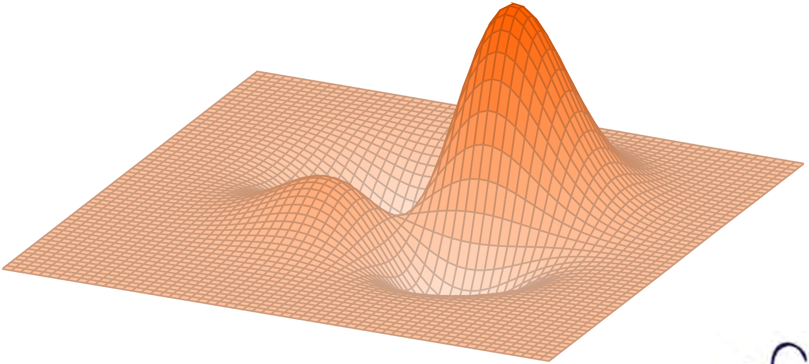
Classification of Extrema

Definition 14

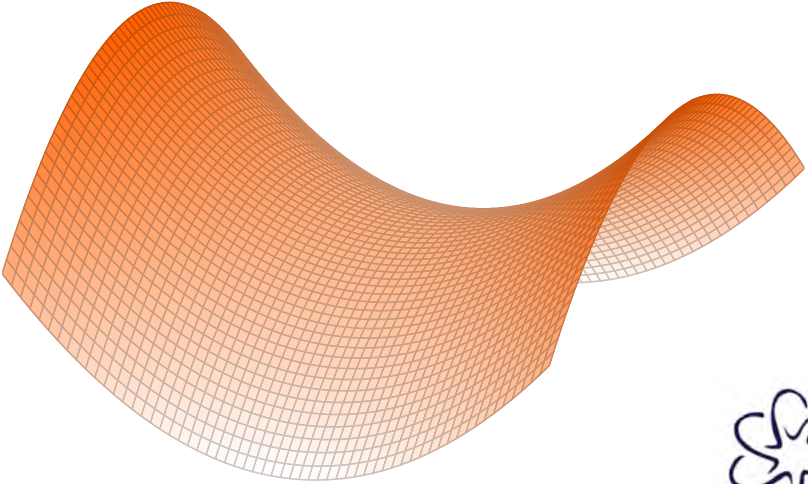
Suppose $f : \mathbb{R}^n \to \mathbb{R}$. Then $\mathbf{x} \in \mathbb{R}^n$ is:

- A **critical point** of f if $\nabla f(\mathbf{x}) = \mathbf{0}$, or does not exist.
- A **global maximum** of f if $f(\mathbf{x}) \geq f(\mathbf{a})$ for all $\mathbf{a} \in \mathbb{R}^n$.
- A **global minimum** of f if $f(\mathbf{x}) \leq f(\mathbf{a})$ for all $\mathbf{a} \in \mathbb{R}^n$.
- A **local maximum** of f if $f(\mathbf{x}) \geq f(\mathbf{a})$ for all $\mathbf{a}$ in some neighbourhood of $\mathbf{x}$.
- A **local minimum** of f if $f(\mathbf{x}) \leq f(\mathbf{a})$ for all $\mathbf{a}$ in some neighbourhood of $\mathbf{x}$.
- A **stationary point** of f if f is differentiable at $\mathbf{x}$ and $\nabla f(\mathbf{x}) = \mathbf{0}$.
- A **saddle point** of f if $\mathbf{x}$ is a stationary point, but not a maximum, nor a minimum.

Local and Global Extrema



Saddle Points



Classification of Extrema

If we consider a bounded set, then extrema may also occur on the boundary, so we also consider them critical points.

Theorem 10

Any extrema occur at critical points. That is, when the gradient is zero or doesn't exist, or on the boundary if the set is bounded.

So when we find extrema, we can ignore much of the region and focus on critical points.
MATH2011 and MATH2111 treat this area in quite different ways, so we will look at each method separately.

Classification of Extrema (MATH2011)

Theorem 11 (Second Derivative Test)

Suppose $f : \mathbb{R}^2 \to \mathbb{R}$, $\nabla f(\mathbf{a}) = \mathbf{0}$, and both f_x and f_y are continuous on a disk around $\mathbf{a}$. Let

$$D = (f_{xy}(\mathbf{a}))^2 - f_{xx}(\mathbf{a})f_{yy}(\mathbf{a}).$$

Then if

- $D < 0$ and $f_{xx}(\mathbf{a}) > 0$, $f(\mathbf{a})$ is a local minimum.
- $D < 0$ and $f_{xx}(\mathbf{a}) < 0$, $f(\mathbf{a})$ is a local maximum.
- $D > 0$, $\mathbf{a}$ is a saddle point.

D is called the **discriminant** of the critical point.
If $D = 0$ then classification requires more work; not covered in MATH2011.

Classification of Extrema (MATH2011)

Example 15 (Question 81e from MATH2011 2019)

Classify the critical points of $f(x,y) = (x^2+y^2)^2 - (x^2-y^2)$.

Solving $\nabla f(x,y) = \mathbf{0}$ we get $(0,0)$ and $\left(\pm\frac{1}{\sqrt{2}}, 0\right)$ as critical points. We also find

$$f_{xy}(0,0) = 0,\ \ f_{xx}(0,0) = -2,\ \ f_{yy}(0,0) = 2,\ \ D = 4.$$
$$f_{xy}\left(\pm\frac{1}{\sqrt{2}}, 0\right) = 0,\ f_{xx}\left(\pm\frac{1}{\sqrt{2}}, 0\right) = f_{yy}\left(\pm\frac{1}{\sqrt{2}}, 0\right) = 4,\ D = -16.$$

So by the second derivative test, $(0,0)$ is a saddle point, and $\left(\pm\frac{1}{\sqrt{2}}, 0\right)$ are local minima.

Classification of Extrema (MATH2011)

Example 16

Find maximum value of $f(x,y) = x^2 + xy - x$ on the set $S = \{(x,y) \in \mathbb{R}^2 : |x| \leq 1, |y| \leq 1\}$.

First, we find $\nabla f(x,y) = (2x+y, x) = \mathbf{0}$, giving us a critical point $(0,0)$. The function value is $f(0,0) = 0$. Now we consider the boundary of the set, in four parts:

1. If $y = 1$, then $f(x,y) = x^2$ so max 1 at $(1,1)$ and $(-1,1)$.
2. If $y = -1$, then $f(x,y) = x^2 - 2x$ so max 3 at $(-1,-1)$.
3. If $x = 1$, then $f(x,y) = y$ so max 1 at $(1,1)$.
4. If $x = -1$, then $f(x,y) = 2 - y$ so max 3 at $(-1,-1)$.

So, the maximum value of f on S is 3 at $(-1,-1)$.

Classification of Extrema (MATH2111)

To classify points in MATH2111 we use the Hessian.

Definition 15

Suppose $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$. The **Hessian** of f at $\mathbf{a} \in \Omega$ is

$$H(f, \mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix}$$

Order of Hessian

Just like the Jacobian, each column is differentiated with respect to the same variable **last**, so column one's partial derivatives all begin with ∂x_1 in the denominator, and so on.

Classification of Extrema (MATH2111)

The eigenvalues of the Hessian give us the stationary point's nature. All positive, all negative, and a mix are minimum, maximum, and saddle respectively. If the point is not a saddle and there is a zero eigenvalue, we need methods outside the scope of MATH2111. To simplify this, we use the following theorem.

Theorem 12 (Sylvester's Criterion)

If H_k is the upper left $k \times k$ submatrix of H and $\Delta_k = \det H_k$, then H is

$$\begin{aligned}
\text{Positive Definite} &\iff \Delta_k > 0 \text{ for all } k \\
\text{Positive Semidefinite} &\implies \Delta_k \geq 0 \text{ for all } k \\
\text{Negative Definite} &\iff \begin{array}{l} \Delta_k < 0 \text{ for odd } k \\ \Delta_k > 0 \text{ for even } k \end{array} \\
\text{Negative Semidefinite} &\implies \begin{array}{l} \Delta_k \leq 0 \text{ for odd } k \\ \Delta_k \geq 0 \text{ for even } k \end{array}
\end{aligned}$$

Classification of Extrema (MATH2111)

By calculating the submatrices of the Hessian we can determine if the eigenvalues are all positive or all negative (usually), and then apply the following.

Theorem 13

Suppose $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is C^2 and $\nabla f(\mathbf{a}) = \mathbf{0}$ at an interior point $\mathbf{a} \in \Omega$. Then

- If $H(f, \mathbf{a})$ is positive definite, f has a local minimum at $\mathbf{a}$.
- If $H(f, \mathbf{a})$ is negative definite, f has a local maximum at $\mathbf{a}$.
- If f has a local minimum at $\mathbf{a}$, $H(f, \mathbf{a})$ is positive semidefinite.
- If f has a local maximum at $\mathbf{a}$, $H(f, \mathbf{a})$ is negative semidefinite.

Generally we deal with $\mathbb{R}^2$.

Classification of Extrema (MATH2111)

Example 17

Classify the stationary point $(1,1,1)$ of $f(x,y,z) = x^4 + y^4 + z^4 + 4xyz$ using Sylvester's Criterion.

Through a cumbersome calculation, we find

$$H(f,(1,1,1)) = \begin{pmatrix} 12 & -4 & -4 \\ -4 & 12 & -4 \\ -4 & -4 & 12 \end{pmatrix}.$$

So, finding the determinants of submatrices, we have

$$\Delta_1 = 12, \quad \Delta_2 = 128, \quad \Delta_3 = 1024,$$

and so the Hessian is positive definite, and so all eigenvalues are positive. This means the stationary point is a minimum.

Classification of Extrema (MATH2111)

The general method for these questions is

1. Calculate the Hessian H.
2. Calculate the submatrix determinants $\Delta_1, \Delta_2, \cdots, \Delta_n$.
3. Consider the cases
   - All $\Delta_k > 0$ then we have a minimum.
   - Δ_k swaps sign, starting at negative, then we have a maximum.
   - Δ_n is non-zero, and the Hessian isn't semidefinite, then we have a saddle point.

Intuition behind submatrix determinants

Think of each matrix determinant as being a product of eigenvalues, so if $\Delta_2 > 0$, then we either have two positive, or two negative eigenvalues, or if $\Delta_1 < 0$, then we must have a negative eigenvalue. This is **not** actually what each submatrix represents, but helps remember the conditions.

Constrained optimization

Definition

A **constrained optimization** problem involves finding stationary points of $f(x, y)$ subject to some restrictions between x, y, given in the form $g_1(x, y), g_2(x, y), ..., g_n(x, y)$. In your examples, typically you will only have the 1 constraint.

Method of Lagrange Multipliers

Suppose we want to optimize the function $z = f(x, y)$ according to the constraint $c(x, y, z) = 0$, the Method of Lagrange multipliers states that we much solve $\nabla F(x, y, z) = \lambda \nabla c(x, y, z)$, where λ is a constant that solves the equation.

Lagrange multipliers

The intuition is to build some level xcurves and investigate when the graph of the constraint is tangential to the graph of the function to be optimised. Since tangents are based on the normal, we are really looking at when the normals are parallel to each other at the same point.

Questions

Lagrange Multipliers questions

1. Optimise the function $x + y + 2z$ along the surface $x^2 + y^2 + z^2 = 3$.
2. Find the optimum points of $f(x, y) = xy$ on the curve $3x^2 + y^2 = 6$.

Solutions

Since this is a function in $\mathbb{R}^{\nvDash}$, we rearrange the equation to obtain $f(x,y,z) = x + y + 2z$, with the constraint being $x^2 + y^2 + z^2 - 3 = c(x,y,z)$.
Therefore, $\nabla f = \lambda \nabla c \implies (1,1,2)^T = \lambda(2x, 2y, 2z)^T$. This means, upon equating the components, $2x\lambda = 1, 2y\lambda = 1, 2z\lambda = 2$.
So substituting $x = y = \frac{1}{2\lambda}, z = \frac{1}{\lambda}$ into the constraint, we can solve for λ:

$$\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 = 3 \implies \lambda = \sqrt{\frac{2}{3}}$$

The constrained stationary point is $\left(\frac{\sqrt{3}}{2\sqrt{2}}, \frac{\sqrt{3}}{2\sqrt{2}}, \sqrt{\frac{3}{2}}\right)$

Solution Q2

The constraint becomes $c(x,y) = 3x^2 + y^2 - 6$. So we equate $\nabla f = \lambda \nabla c$. This yiels:

$$(y,x) = \lambda(6x, 2y) \implies y = 6x\lambda, x = 2y\lambda$$

Solving these equations simultaneously, we get $y = 6(2y\lambda)\lambda \implies y = 12y\lambda^2$ Therefore, either $y = 0, \lambda^2 = \frac{1}{12}$. $y = 0 \implies x = 0$, and $\lambda = \pm\frac{1}{2\sqrt{3}} \implies y = \pm\sqrt{3}x$. Obviously, $(0,0)$ does not work because it doesn't lie on the constraint. So we are looking for the intersection point of $y = \pm\sqrt{3}x$ and $3x^2 + y^2 = 6$. This easily solves to $x = \pm 1, y = \pm\sqrt{3}$

Table of Contents

Integration

Definition

$$\int_{x=a}^{x=b} f(x)dx = \lim_{\Delta x \to 0} \sum_{x=a}^{x=b} f(x_i^*)\Delta x$$

By definition, this means that to compute an area under a curve, we construct rectangles of height $f(x_i^*)$ and width Δx over the interval $a \leq x \leq b$, and add the area of the rectangles and take the limit as the rectangles become super tiny.

Double Integration

Double integration is no different, except we integrate over a 2 dimensional region.

Definition

$$\int_{y=c}^{y=d}\int_{x=a}^{x=b} f(x,y)dxdy = \lim_{\Delta y\to 0}\lim_{\Delta x\to 0}\sum_{y=c}^{y=d}\sum_{x=a}^{x=b} f(x_i^*, y_i^*)\Delta x\Delta y$$

When integrating over a more complicated region, the process is precisely the same. The notation is more generally described as:

$$V = \iint_R f(x,y)dxdy = \iint_R f(x,y)dA$$

Evaluating integrals

Fubini's Theorem

Suppose $f(x,y)$ is a continuous function on $R = [a,b] \times [c,d]$, then

$$\int_{y=c}^{y=d} \int_{x=a}^{x=b} f(x,y)dxdy = \int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x,y)dydx$$

Thus, it is possible to reorder the integral signs to make life easier.

Double Integration Examples

Examples

Evaluate the following integrals:

1. $\iint_R (2x - 4y^3)dA$, where $R = [-5, 4] \times [0, 3]$
2. $\iint_R xe^{xy}\,dA$, where $R = [-1, 2] \times [0, 1]$
3. $\iint_D e^{\frac{x}{y}}\,dA$, where $D = \{(x, y)|1 \leq y \leq 2, y \leq x \leq y^3\}$

Solution Q1

$$\int_{x=-5}^{x=4}\int_{y=0}^{y=3}(2x-4y^3)dydx = \int_{x=-5}^{y=4}\left[2xy-y^4\right]_{y=0}^{y=3} dx$$

$$= \int_{x=-5}^{x=4}\big((6x-81)-(0-0)\big)dx$$

$$= \int_{x=-5}^{x=4}(6x-81)dx$$

$$= \left[3x^2-81x\right]_{x=-5}^{x=4}$$

$$= -756$$

Solution Q2

$$\int_{y=0}^{y=1}\int_{x=-1}^{x=2} xe^{xy}\,dxdy = \int_{x=-1}^{x=2}\int_{y=0}^{y=1} xe^{xy}\,dydx$$

$$= \int_{x=-1}^{x=2}\left[e^{xy}\right]_{y=0}^{y=1}dx$$

$$= \int_{x=-1}^{x=2}(e^{x}-e^{0})dx$$

$$= \left[e^{x}-x\right]_{x=-1}^{x=2}$$

$$= e^{2}-2-(e^{-1}-(-1))$$

$$= e^{2}-e^{-1}-3$$

Solution Q3

$$
\begin{aligned}
\int_{y=1}^{y=2}\int_{x=y}^{x=y^3} e^{\frac{x}{y}}\,dxdy &= \int_{y=1}^{y=2}\left[ye^{\frac{x}{y}}\right]_{x=y}^{x=y^3} dy \\
&= \int_{y=1}^{y=2} ye^{y^2} - ye^1 dy \\
&= \int_{y=1}^{y=2} (ye^{y^2} - ey)dy \\
&= \left[\frac{1}{2}e^{y^2} - \frac{e}{2}y^2\right]_{y=1}^{y=2} \\
&= \frac{1}{2}e^4 - \frac{e}{2}\cdot 4 - \left(\frac{1}{2}e - \frac{e}{2}\right) \\
&= \frac{1}{2}e^4 - 2e
\end{aligned}
$$

Generalised Fubini's Theorem

To change the order of integration, first draw R and then reconstruct the limit.

General rules

1. **Draw the region before working out the new region**
2. The outer most integral should be bounded by constants, while the inner integral will typically be dependent on the other variable.

Examples

Changing order of integration [MATH2011 Q102]

Evaluate the following integrals by changing the order of integration:

1. $\int_0^1 \int_{y^2}^1 2\sqrt{x}e^{x^2}\,dx\,dy$
2. $\int_0^1 \int_y^1 \sin(x^2)\,dx\,dy$

Solution Q1

$$\begin{aligned}
\int_0^1 \int_{y^2}^1 2\sqrt{x}e^{x^2}\,dx\,dy &= \int_{x=0}^{x=1} \int_{y=0}^{y=\sqrt{x}} 2\sqrt{x}e^{x^2}\,dydx \\
&= \int_{x=0}^{x=1} 2\sqrt{x}e^{x^2}\left[y\right]_{y=0}^{y=\sqrt{x}} dx \\
&= \int_0^1 2xe^{x^2}\,dx \\
&= \left[e^{x^2}\right]_0^1 \\
&= e-1
\end{aligned}$$

Solution Q2

$$\int_0^1 \int_y^1 \sin(x^2)\,dx\,dy = \int_{x=0}^{x=1} \int_{y=0}^{y=x} \sin(x^2)\,dydx$$
$$= \int_{x=0}^{x=1} \sin(x^2)\left[y\right]_0^x dx$$
$$= \int_0^1 x\sin(x^2)\,dx$$
$$= \frac{1}{2}\sin 1$$

Triple Integration

The intuition for triple integrals is precisely the same, except there needs to be 3 sets of limits, one for the x direction, one for the y and one for the z directions.

Examples

Examples

1. Find the volume of a tetrahedron bounded by $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ in the first octant.
2. Find the column common to the 2 cylinders $x^2 + y^2 \leq a^2$ and $y^2 + z^2 \leq a^2$
3. Evaluate $\iiint_S x^2 dxdydz$ where S is the region bounded by $4x^2 + y^2 = 4, z + x = 2, z = 0$

Solution Q1 [MATH2011 Q115]

$$\iiint_S dV = \int_{x=0}^{x=a} \int_{y=0}^{y=b\left(1-\frac{x}{a}\right)} \int_{z=0}^{z=c\left(1-\frac{x}{a}-\frac{y}{b}\right)} dz\, dy\, dx$$

$$= c\int_{x=0}^{x=a} \int_{y=0}^{y=b\left(1-\frac{x}{a}\right)} 1-\frac{x}{a}-\frac{y}{b} dy\, dx$$

$$= c\int_{x=0}^{x=a} \left[\left(1-\frac{x}{a}\right)y - \frac{1}{2b}y^2\right]_{y=0}^{y=b\left(1-\frac{x}{a}\right)} dx$$

$$= c\int_{x=0}^{x=a} b\left(1-\frac{x}{a}\right)^2 - \frac{b}{2}\left(1-\frac{x}{a}\right)^2 dx$$

$$= \frac{1}{2}bc\int_{x=0}^{x=a} \left(1-\frac{x}{a}\right)^2 dx$$

$$= \frac{abc}{6}$$

Solution Q2 [MATH2011 Q119]

The region in consideration is the circle $x^2 + y^2 = a^2$ in the xy-plane. So we are measuring the integral:

$$\begin{aligned}
\iiint_S dV &= \int_{x=-a}^{x=a} \int_{y=-\sqrt{a^2-x^2}}^{y=\sqrt{a^2-x^2}} \int_{z=-\sqrt{a^2-y^2}}^{z=\sqrt{a^2-y^2}} dz\,dy\,dx \\
&= \int_{y=-a}^{y=a} \int_{x=-\sqrt{a^2-y^2}}^{x=\sqrt{a^2-y^2}} 2\sqrt{a^2-y^2}\,dx\,dy \\
&\qquad\qquad\qquad \text{(changing order of integration)} \\
&= \int_{y=-a}^{y=a} 2\sqrt{a^2-y^2}\left[x\right]_{x=-\sqrt{a^2-y^2}}^{x=\sqrt{a^2-y^2}} dy \\
&= \int_{y=-a}^{y=a} 4(a^2-y^2)\,dy \\
&= \frac{16}{3}a^3
\end{aligned}$$

Solution Q3 [MATH2011 Q117]

$$\iiint_S x^2 dV = \int_{x=-1}^{x=1}\int_{y=-\sqrt{4-4x^2}}^{y=\sqrt{4-4x^2}}\int_{z=0}^{z=2-x} x^2 dz\, dy\, dx$$
$$= \int_{x=-1}^{x=1}\int_{y=-\sqrt{4-4x^2}}^{y=\sqrt{4-4x^2}} x^2(2-x) dy\, dx$$
$$= \int_{-1}^{1} x^2(2-x)\cdot 2\sqrt{4-4x^2} dx$$
$$= 4\int_{-1}^{1} x^2(2-x)\sqrt{1-x^2} dx$$
$$= 4\int_{-1}^{1} 2x^2\sqrt{1-x^2} dx - 4\int_{-1}^{1} x^3\sqrt{1-x^2} dx$$

The first integral can be solved using $u = \sin(x)$, the second integral is just 0 because the function is odd.

Integration by substitution

There are 3 main types of substitutions:

1. Polar co-ordinates
2. Cylindrical co-ordinates
3. Spherical co-ordinates

Each are basic changes of variable, the more general case will be covered in time.

Jacobians

In single variable integration:

$$\int_{x=a}^{x=b} f(x)dx = \int_{u=a'}^{u=b'} f(u)\frac{du}{dx}dx$$

So there is an extra factor that influences the rate at which the function changes. Similarly, when employing any change of variable, we use the determinant of the Jacobian to adjust the infinitesimal quantities.

Jacobians for polar co-ordinates

Polar co-ordinates

In polar system, we set $x = r\cos\theta, y = r\sin\theta$ based on the geometry of the space.

$$\begin{aligned} J &= \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} \\ &= \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \\ &= r \end{aligned}$$

Jacobian for cylindrical co-ordinates

Cylindrical co-ordinates

The system is exactly the same, except we add a 3rd dimension:
$x = r\cos\theta, y = r\sin\theta, z = z$

$$
\begin{aligned}
J &= \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{pmatrix} \\
&= \det\begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= r
\end{aligned}
$$

Jacobian for Spherical co-ordinates

Spherical co-ordinates

We let the following change of variable occur:
$x = \rho \sin\varphi \cos\theta, y = \rho \sin\varphi \sin\theta, z = \rho \cos\varphi$

$$
\begin{aligned}
J &= \det \begin{pmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \theta} \end{pmatrix} \\
&= \det \begin{pmatrix} \sin\varphi\cos\theta & \sin\varphi\sin\theta & \cos\varphi \\ \rho\cos\varphi\cos\theta & \rho\cos\varphi\sin\theta & -\rho\sin\varphi \\ -\rho\sin\varphi\sin\theta & \rho\sin\varphi\cos\theta & 0 \end{pmatrix} \\
&= \rho^2 \sin\varphi
\end{aligned}
$$

General Change of Variables

Integration by Substitution

Consider the substitutions $x = g(u, v), y = h(u, v)$. Then the following holds:

$$\iint_{\Omega} f(x, y) dx\, dy = \iint_{\Omega'} f(g(u, v), h(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\, dv$$

Examples for polar co-ordinates

Examples

1. Use polar co-ordinates to find $\iint_\Omega x^2 y^3 dR$ where $\Omega = \{(x, y) | x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}$
2. Find the area of the shape defined by the inequalities $y \geq 0, y \geq -x, x^2 + y^2 \leq 3\sqrt{x^2 + y^2} - 3x$

Solution Q1 [Paul's Online Maths Notes]

$$
\begin{aligned}
\iint_R dA &= \int_{\theta=0}^{\theta=\frac{\pi}{2}} \int_{r=0}^{r=1} (r\cos(\theta))^2 (r\sin(\theta))^3 r dr\, d\theta \\
&= \int_{\theta=0}^{\theta=\frac{\pi}{2}} r^6 \cos^2(\theta)\sin^3(\theta) dr\, d\theta \\
&= \int_{\theta=0}^{\theta=\frac{\pi}{2}} \frac{1}{7}\cos^2(\theta)\sin^2\theta\sin(\theta) d\theta \\
&= \frac{1}{7}\int_1^0 u^2(1-u^2)\cdot -du \\
&= \frac{1}{7}\int_0^1 u^2 - u^4 du \\
&= \frac{2}{105}
\end{aligned}
$$

Solution Q2 [MATH2011 Q109/MATH2111 Q149]

$$\iint_R dA = \int_{\theta=0}^{\theta=\frac{3\pi}{4}} \int_{r=0}^{r=3-3\cos\theta} r dr\, d\theta$$
$$= \int_0^{\frac{3\pi}{4}} \frac{1}{2}(3-3\cos\theta)^2 d\theta$$
$$= \text{la-da-daa-da-da-de-di-da-di-day}$$
$$= \frac{81\pi}{16} - \frac{9}{8}(4\sqrt{2}+1)$$

Examples on Cylindrical co-ordinates

Examples from homework

1. Find the volume of the solid enclosed between the spheres $x^2 + y^2 + z^2 = 4$ and $x^2 + y^2 + z^2 = 4z$
2. Find the volume inside the cone $z + 2 = \sqrt{x^2 + y^2}$ between the planes $z = 0, z = 1$

Solution Q1 [MATH2111 Q152]

$$\iiint_S dV = \int_{r=0}^{r=\sqrt{3}} \int_{\theta=0}^{\theta=2\pi} \int_{z=2-\sqrt{4-x^2-y^2}}^{z=2+\sqrt{4-r^2}} rdz\, d\theta\, dr$$

$$= \int_{r=0}^{r=\sqrt{3}} \int_{\theta=0}^{\theta=2\pi} \int_{z=2-\sqrt{4-r^2}}^{z=2+\sqrt{4-r^2}} rdz\, d\theta\, dr$$

The outer integral for the bounds on r are because we only consider how far away from the original in the xy plane we travel, θ ranges from 0 to 2π because we can go around in a full circle in the valid region for r, and z ranges from the lower sphere to the upper sphere.

Solution Q2 [MATH2111 Q153]

$$\iiint_S dV = \int_{r=2}^{r=3} \int_{\theta=0}^{\theta=2\pi} \int_{z=0}^{z=r-2} r dz\, d\theta\, dr$$

Examples of Spherical co-ordinates

Examples from homework

1. Find the volume of the region above the cone $z = \sqrt{x^2 + y^2}$ and inside the sphere $x^2 + y^2 + z^2 = 2az$
2. Use spherical coordinates to find the volume enclosed by the surface $(\sqrt{x^2 + y^2} - 1)^2 + z^2 = 1$

Solution Q1 [MATH2111 Q158]

$$\iiint_S dV = \int_{\theta=0}^{\theta=2\pi} \int_{\varphi=0}^{\varphi=\frac{\pi}{4}} \int_{\rho=0}^{\rho=a} \rho^2 \sin(\varphi) d\rho \, d\varphi \, d\theta$$

Solution Q2 [MATH2011 Q128/MATH2111 Q156 iii)]

$$\iiint_S dV = \int_{\theta=0}^{\theta=2\pi} \int_{\varphi=0}^{\varphi=\pi} \int_{\rho=0}^{\rho=\frac{1}{\sin(\varphi)}} \rho^2 \sin(\varphi) d\rho \, d\varphi \, d\theta$$

Examples on change of variables

Examples from homework

1. Let Ω be the region in the first quadrant bounded by the hyperbolas $x^2 - 2y^2 = 1, x^2 - 2y^2 = 3, xy = 1, xy = 2$. Let $u = x^2 - 2y^2, v = xy$. Sketch the region Ω in the $x - y$ plane and the region Ω' in the u, v plane that corresponds to Ω. Hence evaluate $\iint_\Omega (x^2 - 2y^2)x^2y^2(2x^2 + 4y^2)dx\,dy$
2. Integrate the function $\frac{1}{xy}$ over the region Ω' bounded by the 4 circles
$x^2 + y^2 = ax, x^2 + y^2 = a'x, x^2 + y^2 = by, x^2 + y^2 = b'y$.

Solution Q1 [MATH2011 Q137/MATH2111 Q160]

We shall consider the substitution $u = x^2 - 2y^2, v = xy$. Then the Jacobian of this substitution is:

$$\det \frac{\partial(u,v)}{\partial(x,y)} = \det \begin{bmatrix} 2x & y \\ -4y & x \end{bmatrix} = 2x \cdot x - (-4y)\cdot y = 2x^2 + 4y^2$$

$$\iint_\Omega (x^2 - 2y^2)x^2y^2(2x^2+4y^2)dx\,dy = \iint_{\Omega'} uv^2(2x^2+4y^2)\cdot \frac{1}{2x^2+4y^2}du\,dv$$

$$= \int_{u=1}^{u=3}\int_{v=1}^{v=2} uv^2\,du\,dv$$

Solution Q2 [MATH2111 Q161] I

Suppose we re-arrange the given equations: $\frac{x^2+y^2}{x} = a, a'$ and $\frac{x^2+y^2}{y} = b, b'$. Then we use the subtitutions $u = \frac{x^2+y^2}{x}, v = \frac{x^2+y^2}{y}$. Thus, we have:

$$\begin{aligned}
\det\begin{bmatrix} 1-\frac{y^2}{x^2} & \frac{2x}{y} \\ \frac{2y}{x} & -\frac{x^2}{y^2}+1 \end{bmatrix} &= \left(1-\frac{y^2}{x^2}\right)\left(1-\frac{x^2}{y^2}\right) - \left(\frac{2y}{x}\right)\left(\frac{2x}{y}\right) \\
&= 2 - \frac{x^2}{y^2} - \frac{y^2}{x^2} - (4) \\
&= -\frac{x^2}{y^2} - 2 - \frac{y^2}{x^2} \\
&= -\left(\frac{x}{y}+\frac{y}{x}\right)^2 \\
&= -\left(\frac{x^2+y^2}{xy}\right)^2
\end{aligned}$$

Solution Q2 [MATH2111 Q161] II

Hence we have: $-\left(\frac{x^2+y^2}{xy}\right)^2 = \frac{uv}{xy}$

Thus the integral becomes:

$$\iint_{\Omega'} \frac{1}{xy} dx\, dy = \iint_{\Omega} \frac{1}{xy} \frac{xy}{uv} du\, dv = \int_{v=b}^{v=b'} \int_{u=a}^{u=a'} \frac{1}{uv} du\, dv$$

Centre of Mass

Consider a solid (either a plane or 3D solid) where every infinitesimally small part of the solid has a specific **mass density** denoted $\rho(\mathbf{x})$. This means that the overall mass of the object will be given by $\iint_\Omega \rho(\mathbf{x})dA$ or $\iiint_\Omega \rho(\mathbf{x})dV$ depending on how many dimensions the solid has.

Mass of a solid

The mass of a solid is given by the formula:

$$M(S) = \int \dots \int_\Omega \rho(\mathbf{x})dR$$

Centre of Mass (MATH2011 ONLY)

Centre of Mass

The i^{th} coordinate for the centre of mass is given according to the formula:

$$c_i = \frac{\iint_\Omega x_i \rho(\mathbf{x})}{M(S)} dR$$

Here, $M(S)$ describes the mass of the solid, and $\rho(\mathbf{x})$ describes the density of the solid as a function of $\mathbf{x}$

To compute the centre of mass can be done using whatever rules you want to, provided that are valid under double and triple integration.

Leibniz Rule (MATH2111 ONLY)

Commonly known as Differentiation under the integral sign, Leibniz Rule helps evaluate weird looking integrals based on related functions.

Leibniz Rule

$$\frac{d}{dx}\int_{t=a}^{t=b} f(x,t)dt = \int_{t=a}^{t=b} \frac{\partial}{\partial x} f(x,t)dt$$

For a more general integral, we have:

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,t)dt = \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)dt + f(x,b(x))\cdot b'(x) - f(x,a(x))\cdot a'(x)$$

Examples

2018 Q1iii)

You are given that:

$$\int_{-\infty}^{\infty} \frac{\cos(3x)}{x^2 + a^2} dx = \frac{\pi}{a} e^{-3a}$$

Evaluate the following integral:

$$\int_{-\infty}^{\infty} \frac{\cos(3x)}{(x^2 + 4)^2} dx$$

Solution 2018 Q1iii)

Beginning with the given integral, we differentiate both sides with respect to a

$$\frac{d}{da}\int_{-\infty}^{\infty}\frac{\cos(3x)}{x^2+a^2}dx = \frac{d}{da}\frac{\pi}{a}e^{-3a}$$

$$\int_{-\infty}^{\infty}\frac{\partial}{\partial a}\frac{\cos(3x)}{x^2+a^2}dx = -\frac{-\pi e^{-3x}(3x+1)}{x^2}$$

$$\int_{-\infty}^{\infty}-\frac{2a\cos(3x)}{(x^2+a^2)^2}dx = -\frac{\pi e^{-3a}(3a+1)}{a^2}$$

Substituting $a = 2$

$$\int_{-\infty}^{\infty}-4\frac{\cos(3x)}{(x^2+4)^2}dx = -\frac{7\pi e^{-6}}{4}$$

$$\int_{-\infty}^{\infty}\frac{\cos(3x)}{(x^2+4)^2}dx = \frac{7\pi e^{-6}}{16}$$

Examples

2014 Q2i)

Suppose that $f : [0,1] \times [0,1] \to \mathbb{R}$ has continuous partial derivatives. Define $F : [0,1] \times [0,1] \to \mathbb{R}$ by:

$$F(u,v) = \int_0^u f(v,y)dy$$

1. If $u = u(x), v = v(x)$, find an expression for $\frac{dF}{dx}$
2. Hence, or otherwise, compute:

$$\frac{d}{dx}\int_0^x \frac{\sin(xy)}{y}dy$$

Solution 2014 Q2i) a)

By Leibniz Rule:

$$\begin{aligned}
\frac{dF}{dx} &= \frac{d}{dx}\int_0^{u(x)} f(v(x),y)dy \\
&= \int_0^{u(x)} \frac{\partial}{\partial x} f(v(x),y)dy + f(v(x),u(x))\frac{d}{dx}u(x) - f(v(x),0)\frac{d}{dx}(0) \\
&= \int_0^{u(x)} v'(x)\frac{\partial f(v(x),y)}{\partial x}dy + f(v(x),u(x))u'(x)
\end{aligned}$$

Solution 2014 Q2i) b)

Let $u(x) = x, v(x) = x$. Then upon substitution into the expression from a), we have:

$$\begin{aligned}\frac{d}{dx}\int_0^x \frac{\sin(xy)}{y}dy &= \int_0^x 1\cdot y\cdot\frac{\cos(xy)}{y} + \frac{\sin(x^2)}{x}\cdot 1\\ &= \int_0^x \cos(xy)dy + \frac{\sin(x^2)}{x}\\ &= \left[\frac{\sin(xy)}{x}\right]_0^x + \frac{\sin(x^2)}{x}\\ &= 2\frac{\sin(x^2)}{x}\end{aligned}$$