MATH20901 Multivariable Calculus (2015)

Dr Richard Porter

Course Information

Prerequisites

- Calculus 1 (and Linear Algebra and Geometry, Analysis 1)

Description of the course

- The course develops multivariable calculus from Calculus 1. The main focus of the course is on developing differential vector calculus, tools for changing coordinate systems and major theorems of integral calculus for functions of more than one variable.
- This unit is central to many branches of pure and applied mathematics. For example, in applied mathematics vector calculus is an integral part of describing field theories that model physical processes and dealing with the equations that arise. It is used in second year Applied Partial Differential Equations (which is a prerequisite for 3rd Fluid Dynamics and Quantum Mechanics) and 3rd year Mathematical Methods and Differential Manifolds use the material of Multivariable Calculus 2.

Resources

- Lecturer: Dr. Richard Porter, Room SM2.7
- Unit description:
  `http://www.maths.bris.ac.uk/study/undergrad/current_units/unit/?id=782`
- Web: `http://www.maths.bris.ac.uk/~marp/mvcalc`. Notes may contain extra sections for interest or additional information. Problem sheets, solutions, homework feedback forms, problems class sheets, past exam papers, links.
- Email: `richard.porter@bris.ac.uk`
- Books: Lots of books on multivariable/vector calculus. Jerrold E. Marsden & Anthony J. Tromba, "Vector Calculus", ed. 5 , W. H. Freeman and Company, 2003
- Maths Café
- Office Hours: Friday 2-3.

Problem Sheets/Problems Classes

- Homework set weekly from problems sheets.
- Timetabled problems classes/exercise classes: unseen problems/some from the problem sheets/and as many as possible from past exam papers.

Exam

- Jan 1hr30min. 2 compulsory questions.
- No calculators.

1 A review of differential calculus for functions of more than one variable

Revision of results from Calculus 1

1.1 General maps from $\mathbb{R}^m$ to $\mathbb{R}^n$

Let $\mathbf{x} \in \mathbb{R}^m = (x_1, x_2, \ldots, x_m)$.

Often in 2D $\mathbf{x} \equiv (x, y)$ or in 3D $\mathbf{x} \equiv (x, y, z)$.

We define a general mapping, or vector function, say $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ s.t. $\mathbf{x} \to \mathbf{F}(\mathbf{x})$.

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_n(\mathbf{x})).$$

The components are scalar maps denoted by $F_i : \mathbb{R}^m \to \mathbb{R}$ $(i = 1, \ldots, n)$.

When the range is $\mathbb{R}$, $\mathbf{F}$ is scalar and we often refer to these maps as scalar functions or more simply *functions* and denote them by lower-case symbols, e.g. $f : \mathbb{R}^m \to \mathbb{R}$.

Defn: A map $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ is *linear* if $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, and $\lambda, \mu \in \mathbb{R}$, $\mathbf{F}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathbf{F}(\mathbf{x}) + \mu\mathbf{F}(\mathbf{y})$.

Proposition: A map $\mathbf{F}$ is linear iff $\exists$ a matrix $A \in \mathbb{R}^{n\times m}$ s.t. $\mathbf{F} = A\mathbf{x}$.

E.g. 1) $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^2$ s.t. $(x_1, x_2, x_3) \to (x_3 - x_1, x_2 + x_1)$. Then

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A\mathbf{x}$$

(a linear map)

E.g. 2) $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ s.t. $(x_1, x_2) \to (x_2 x_1, \mathrm{e}^{x_2})$. Not a linear map.

1.2 The derivative of a map

Defn: The derivative of the map $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ is the $m \times n$ matrix $\mathbf{F}'(\mathbf{x})$ such that the i, jth element is

$$\{\mathbf{F}'(\mathbf{x})\}_{ij} = \frac{\partial F_i}{\partial x_j}$$

It is the matrix which encodes, local to a point $\mathbf{x}_0$, say, the linear map representing the tangent plane to the hypersurface at $\mathbf{x} = \mathbf{x}_0$ defined by the mapping $\mathbf{F}(\mathbf{x})$.

E.g. In 1D we have a function $f(x)$. The derivative of the function at $x = x_0$ is defined as $f'(x_0)$ and represents the gradient of the line tangential to the curve $f(x)$ at $x = x_0$.
I.e. $y = f(x_0) + (x - x_0)f'(x_0)$ is a linear map (a straight line) representing the tangent to the curve $f(x)$ at $x = x_0$ and $y' = f'(x_0)$

In general the linear map $\mathbf{y} = \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ represents the tangent plane to the surface $\mathbf{F}(\mathbf{x})$ and $\mathbf{y}' = \mathbf{F}'(\mathbf{x}_0)$.

Defn: The matrix $\mathbf{F}'(\mathbf{x})$ with elements $\partial F_i/\partial x_j$ is called the *Jacobian matrix.*

1.3 The gradient of a function

Defn: The gradient of a scalar function $f : \mathbb{R}^m \to \mathbb{R}$ (i.e. $f(\mathbf{x})$) is denoted

$$\boldsymbol{\nabla} f \equiv (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_m)$$

Note: The rows of the Jacobian matrix are formed by gradients of the components of $\mathbf{F}$, viz

$$\mathbf{F}'(\mathbf{x}) = \begin{pmatrix} \boldsymbol{\nabla} F_1 \\ \boldsymbol{\nabla} F_2 \\ \vdots \\ \boldsymbol{\nabla} F_n \end{pmatrix}$$

(More on this later)

1.4 The directional derivative

Defn: The *directional derivative* of $\mathbf{F}$ at $\mathbf{x}$ along $\mathbf{v}$ is a vector in $\mathbb{R}^n$ given by

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{x}) = \left(\frac{dF_1(\mathbf{x} + t\mathbf{v})}{dt}, \ldots, \frac{dF_n(\mathbf{x} + t\mathbf{v})}{dt}\right)_{t=0} \equiv \left.\frac{d\mathbf{F}(\mathbf{x} + t\mathbf{v})}{dt}\right|_{t=0}$$

It measures the *rate of change* of $\mathbf{F}$ in the direction of $\mathbf{v}$ and it is formulated in terms of ordinary 1D derivatives.

Note: Can be shown

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{x}) = \mathbf{F}'(\mathbf{x})\mathbf{v}.$$

Note: $\mathbf{v}$ must have unit magnitude, or $|\mathbf{v}| = 1$ where

$$|\mathbf{v}| = \sqrt{v_1^2 + v_2^2 + \ldots + v_m^2}$$

denotes the L_2−norm.

Note: If $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$ is a scalar function then

$$D_{\mathbf{v}} f = \mathbf{v} \cdot \boldsymbol{\nabla} f \tag{1}$$

1.5 Operations on maps

1. Let $\mathbf{F}, \mathbf{G}$ be maps from $\mathbb{R}^m \to \mathbb{R}^n$, then $\mathbf{F} + \mathbf{G} : \mathbb{R}^m \to \mathbb{R}^n$ is defined as

$$(\mathbf{F} + \mathbf{G})(\mathbf{x}) = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})$$

2. Let $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ and $f : \mathbb{R}^m \to \mathbb{R}$, then $f\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ is defined by

$$(f\mathbf{F})(\mathbf{x}) = f(\mathbf{x})\mathbf{F}(\mathbf{x})$$

3. If $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^p$, then $\mathbf{G} \circ \mathbf{F} : \mathbb{R}^m \to \mathbb{R}^p$ is defined by

$$(\mathbf{G} \circ \mathbf{F})(\mathbf{x}) = \mathbf{G}(\mathbf{F}(\mathbf{x}))$$

Note: A "function of a function".

1.6 Derivatives of operations on maps

The familiar rules for the derivative of a sum, product, and composition of functions (chain rule) generalise to maps.

1. With $\lambda, \mu \in \mathbb{R}$

$$(\lambda\mathbf{F} + \mu\mathbf{G})' = \lambda\mathbf{F}' + \mu\mathbf{G}',$$

or in components:

$$\frac{\partial(\lambda F_i + \mu G_i)}{\partial x_j}(\mathbf{x}) = \lambda\frac{\partial F_i}{\partial x_j}(\mathbf{x}) + \mu\frac{\partial G_i}{\partial x_j}(\mathbf{x}).$$

2. The derivative of $(f\mathbf{F})(\mathbf{x})$ is a matrix whose i, jth element is

$$\frac{\partial(fF_i)}{\partial x_j}(\mathbf{x}) = \frac{\partial f}{\partial x_j}(\mathbf{x})F_i(\mathbf{x}) + f(\mathbf{x})\frac{\partial F_i}{\partial x_j}(\mathbf{x}).$$

3.

$$(\mathbf{G} \circ \mathbf{F})'(\mathbf{x}) = \mathbf{G}'(\mathbf{F}(\mathbf{x}))\,\mathbf{F}'(\mathbf{x}), \tag{2}$$

where the right-hand side denotes the product of the $p \times n$ matrix $\mathbf{G}'(\mathbf{F}(\mathbf{x}))$ with the $n \times m$ matrix $\mathbf{F}'(\mathbf{x})$. Equivalently,

$$\frac{\partial(\mathbf{G} \circ \mathbf{F})_i}{\partial x_k}(\mathbf{x}) = \sum_{j=1}^{p} \frac{\partial G_i}{\partial x_j}(\mathbf{F}(\mathbf{x}))\frac{\partial F_j}{\partial x_k}(\mathbf{x}).$$

Note: This is the generalised *Chain rule.*

E.g. Recall the chain rule: Consider the function $g(u, v)$ and $u = u(x, y)$ and $v = v(x, y)$. Then

$$\frac{\partial g}{\partial x} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x}$$

and

$$\frac{\partial g}{\partial y} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial y}$$

This is the same as

$$\begin{pmatrix} g_x \\ g_y \end{pmatrix} = (g_u, g_v)\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$$

1.7 Inverse maps

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{G} = \mathbf{F}^{-1}$ be in *inverse* map such that

$$(\mathbf{F}^{-1} \circ \mathbf{F})(\mathbf{x}) = \mathbf{x}; \tag{3}$$

Applying (2) to (3), we find

$$(\mathbf{F}^{-1})'(\mathbf{F}(\mathbf{x}))\mathbf{F}'(\mathbf{x}) = \mathbf{I},$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. This follows since $\mathbf{x} \equiv \mathbf{I}\mathbf{x}$, a linear map with derivative $\mathbf{I}$.

Note: if $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$ is a general linear map then $\mathbf{F}'(\mathbf{x}) = A$ (easy to see/show).

Thus

$$(\mathbf{F}^{-1})'(\mathbf{F}(\mathbf{x})) = (\mathbf{F}')^{-1}(\mathbf{x})$$

or "*the derivative of the inverse is equal to the inverse of the derivative*".

1.7.1 Example: mapping Cartesian to polar coordinates

Let $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ s.t. $(r, \theta) \to (r\cos\theta, r\sin\theta) \equiv (x, y)$.

This means that

$$\mathbf{F}'(r,\theta) = \begin{pmatrix} \partial x/\partial r & \partial x/\partial \theta \\ \partial y/\partial r & \partial y/\partial \theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Taking inverses

$$(\mathbf{F}'(r,\theta))^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix}.$$

Now consider the inverse map $\mathbf{F}^{-1} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ s.t. $(x, y) \to (\sqrt{x^2+y^2}, \tan^{-1}(y/x)) \equiv (r, \theta)$. Then

$$(\mathbf{F}^{-1})'(x,y) = \begin{pmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial \theta/\partial x & \partial \theta/\partial y \end{pmatrix} = \begin{pmatrix} x/\sqrt{x^2+y^2} & y/\sqrt{x^2+y^2} \\ -y/(x^2+y^2) & x/(x^2+y^2) \end{pmatrix}$$

Finally,

$$(\mathbf{F}^{-1})'(\mathbf{F}(r,\theta)) = \begin{pmatrix} r\cos\theta/r & r\sin\theta/r \\ -r\sin\theta/r^2 & r\cos\theta/r^2 \end{pmatrix}$$

which is the same as before.

1.8 Solving equations

Question: Given a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, is there always an inverse function $\mathbf{G} \equiv \mathbf{F}^{-1}$, which satisfies

$$(\mathbf{G} \circ \mathbf{F})(\mathbf{x}) = \mathbf{x}?$$

The same question can be stated in terms of a solution to a nonlinear system of equations. Namely, let

$$\mathbf{F}(\mathbf{x}) = \mathbf{y} \tag{4}$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Or, in full,

$$\begin{aligned} F_1(x_1, \ldots, x_n) &= y_1 \\ \vdots \qquad\quad & \quad \vdots \\ F_n(x_1, \ldots, x_n) &= y_n. \end{aligned}$$

Then, given $\mathbf{y}$, is there a $\mathbf{x}$ such that (4) is solved. If so, then $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{y})$.

1.8.1 Inverse function theorem

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, with $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^n$ such that

$$\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0).$$

If the Jacobian matrix $\mathbf{F}'(\mathbf{x}_0)$ is invertible, then (4) can be solved *uniquely* as

$$\mathbf{x} = \mathbf{F}^{-1}(\mathbf{y}),$$

for $\mathbf{y}$ in the *neighbourhood* of $\mathbf{y}_0$.

Note: A matrix is invertible if and only if its determinant is non-zero. The determinant of the Jacobian matrix $\mathbf{F}'$ is often written as

$$J_{\mathbf{F}}(\mathbf{x}_0) \equiv \left.\frac{\partial(F_1, \ldots, F_n)}{\partial(x_1, \ldots, x_n)}\right|_{\mathbf{x}=\mathbf{x}_0} \tag{5}$$

and called the *Jacobian determinant*.

E.g. Consider the system of equations

$$\frac{x^2 + y^2}{x} = u, \quad \sin x + \cos y = v.$$

Q: Given (u, v), we want to solve for (x, y). Near which points does this define a unique function ?

A: We define $\mathbf{F} : \mathbb{R}^2 \backslash \{0\} \to \mathbb{R}^2$ s.t.

$$\mathbf{y} \equiv \mathbf{F}(\mathbf{x}) = \left(\frac{x^2 + y^2}{x}, \sin x + \cos y \right).$$

(so that $\mathbf{y} = (u, v)$ and $\mathbf{x} = (x, y)$.)
The Jacobian determinant is

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} (x^2 - y^2)/x^2 & 2y/x \\ \cos x & -\sin y \end{vmatrix} = \frac{y^2 - x^2}{x^2} \sin y - \frac{2y}{x} \cos x.$$

E.g. (i) near $\mathbf{x}_0 = (1, 1)$ (where $\mathbf{y} = (2, \sin(1) + \cos(1))$) we can solve for $\mathbf{x}$ in a neighborhood of $\mathbf{x}_0$; E.g. (ii) near $\mathbf{x}_0 = (\pi/2, \pi/2)$ (where $\mathbf{y} = (\pi, 1)$) where $J_{\mathbf{F}} = 0$ we can't!

1.8.2 Implicit function theorem

Similar to above. Consider an equation for $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$ in the form

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = 0 \tag{6}$$

where $\mathbf{F} : \mathbb{R}^{m+n} \to \mathbb{R}^n$.
Note: If $\mathbf{F}$ is linear in $\mathbf{y}$ then it can be written in the form $\mathbf{y} = \mathbf{G}(\mathbf{x})$ for some $\mathbf{G}$. We suppose that this is not the case.
Suppose that (6) is satisfied by the pair $\mathbf{x}_0$, $\mathbf{y}_0$. Then we can express solutions of this as $\mathbf{y} = \mathbf{y}(\mathbf{x})$ for $\mathbf{y} : \mathbb{R}^m \to \mathbb{R}^n$ in the neighbourhood of $\mathbf{x}_0$, $\mathbf{y}_0$ provided the Jacobian determinant

$$\left.\frac{\partial(F_1, \ldots, F_n)}{\partial(y_1, \ldots, y_n)}\right|_{\mathbf{x}=\mathbf{x}_0, \mathbf{y}=\mathbf{y}_0}$$

is non-zero.

E.g. Consider $f(x, y) = 0$ where $f(x, y) = x^2 + y^2 - 1$. This is satisfied by points (x_0, y_0) on the unit circle. If we try to express it as $y = y(x)$ we get into trouble since

$$y = \pm\sqrt{1 - x^2}$$

and there are two solutions. The implicit function theorem applied to this example requires the determinant of the 1×1 matrix

$$\frac{\partial f}{\partial y}$$

evaluated at (x_0, y_0) to be non-zero. This is $2y_0$ which is non-zero apart from at $y_0 = 0$. So we can express the solution $y = y(x)$ local to a point (x_0, y_0) provided $y_0 \neq 0$. Which is obvious in our case as if $y_0 > 0$ we are on the upper solution branch where $y = \sqrt{1 - x^2}$ and vice versa.

Note: We can also find values of derivatives at these points since, using the e.g. above, differentiation of $f(x, y) = x^2 + y^2 - 1 = 0$ w.r.t. x gives

$$2x + 2y\frac{dy}{dx} = 0$$

and so

$$\left.\frac{dy}{dx}\right|_{x=x_0, y=y_0} = -\frac{x_0}{y_0}.$$

This idea can be extended to vector functions.

1.9 Higher-order derivatives

Start with 2nd order

Defn: For $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$,

$$\frac{\partial^2 F_i}{\partial x_k \partial x_j}(\mathbf{x}) = \frac{\partial}{\partial x_k}\left(\frac{\partial F_i}{\partial x_j}\right)(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial F_i}{\partial x_k}\right)(\mathbf{x}) = \frac{\partial^2 F_i}{\partial x_j \partial x_k}(\mathbf{x}).$$

under normal circumstances.

E.g. $f(x,y) = x^3 - 3xy^2$. Then

$$f_{xx} \equiv \frac{\partial^2 f}{\partial x \partial x} = 6x, \qquad f_{xy} \equiv \frac{\partial^2 f}{\partial x \partial y} = -6y, \qquad f_{yx} = -6y, \qquad f_{yy} = -6x$$

Note: Extended naturally to higher orders.

1.9.1 Taylor's theorem

Higher-order derivatives are useful in Taylor's theorem in dimension ≥ 2, allowing one to approximate functions of several variables near a point.

Recall that for a scalar function of a single variable,

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x-x_0)^2}{2!} f''(x_0) + \text{higher order terms}$$

How do we generalise to higher dimensions ? Well, it gets tricky. E.g. for a scalar function $f(x,y)$,

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + (f_x, f_y) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \\ &\quad + \frac{1}{2}(x - x_0, y - y_0) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}_{x=x_0, y=y_0} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \text{higher order terms} \end{aligned}$$

with an obvious generalisation to functions of more than 2 variables.

The higher order terms are complicated and require some complex notation.

For vector functions, what we do know is that

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \text{terms of size like } |\mathbf{x} - \mathbf{x}_0|^2$$

as this is the definition of $\mathbf{F}'$.