Engineering Mathematics 2B

Module 1: Differentiation

Nick Polydorides

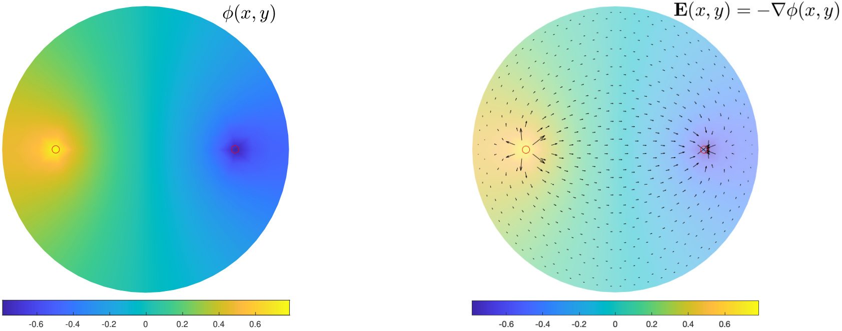
School of Engineering

THE UNIVERSITY of EDINBURGH

Module 1 contents

Motivation: Electrostatics

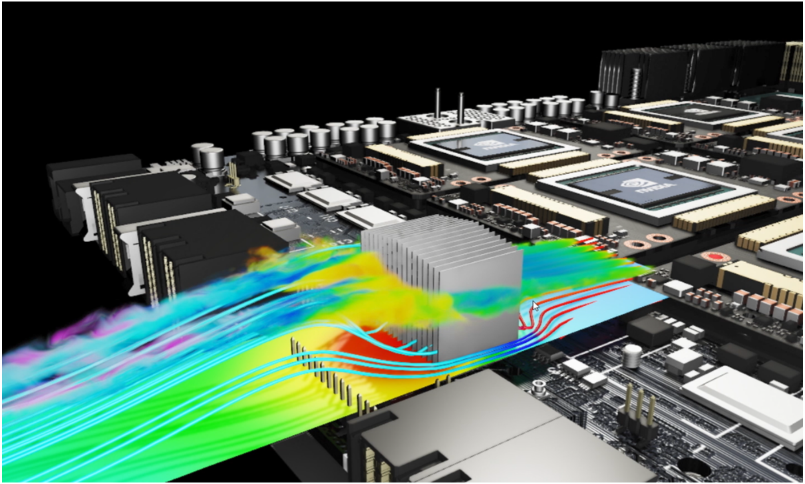
Electrostatic potential ϕ and its electric field $\mathbf{E}$ on a homogeneous metallic disk.



The +ve and −ve poles of a battery are connected to the disk at the red circles.

Motivation: Heat transfer

Temperature field and heat transfer near an electronics heat sink.



Red: high temperature, blue: low temperature. Source: Nvidia.

Fields on the xy plane

A **scalar** field on the xy plane is function $f(x, y)$ which evaluates to a scalar value or infinity at any point (x, y).

Let

$$f(x, y) = 2xy - 3y^2, \quad \text{and} \quad g(x, y) = \frac{4}{|x - y|} - xy.$$

f and g are 2D scalar fields on the xy plane. f is continuous (= attains a finite value) everywhere but g is discontinuous on the line $x = y$.

2D scalar fields needn't be defined on xy alone. They could be at xz or yz or $r\theta$, e.g.

$$f(x, z), \quad f(y, z), \quad f(r, \theta)$$

as long as they take two coordinates to yield one value.

Fields on the xy plane

Scalar fields can be added, subtracted, multiplied or divided just like scalar numbers. For example if

$$f(x,y) = 2xy - 3y^2, \quad \text{and} \quad g(x,y) = \frac{4}{|x-y|} - xy.$$

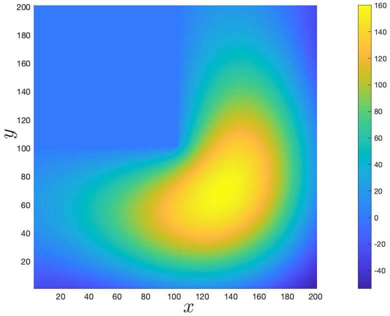
then

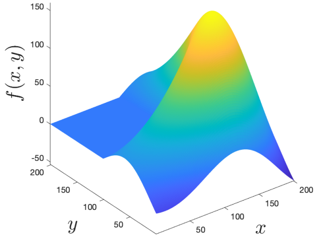
$$h(x,y) = f(x,y) + 3g(x,y) = xy - 3y^2 + \frac{12}{|x-y|},$$

and

$$f(x,y)^2 = 4x^2y^2 - 12xy^3 + 9y^4.$$

Function graphing conventions





Left: A graph of $f(x,y)$ in a range of values over its physical (=2D) domain. Right: A graph of the same $f(x,y)$ in pseudo-3D. Notice the vertical axis stands for the value of the function.

Convex & Concave functions

For a scalar field/function $f(x, y)$ the most interesting information we can ask for is its value, in particular the variation of its value in space.

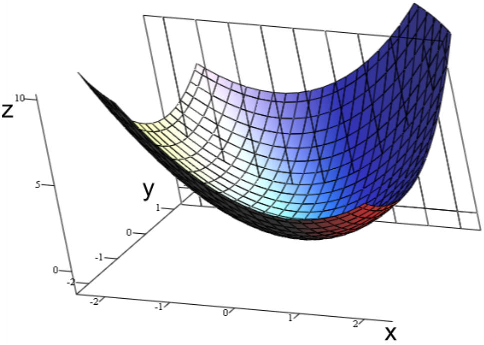
Say you are given $f(x, y)$ and asked to find one of its minima or maxima. Shouldn't we first have a way to establish that said points exist for f? (Will a graph work?)

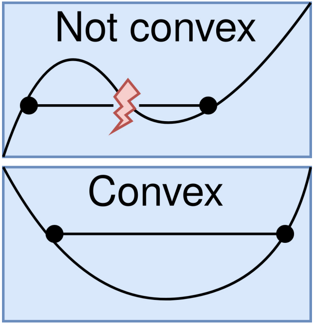
A twice-differentiable function is convex if and only if its Hessian matrix (= matrix of second partial derivatives) is positive definite (= all of its eigenvalues are positive).

Convex functions have at least one minimum.

2D convex functions

See the graph of $f(x, y) = x^2 + xy + y^2$.





f is called convex if the line segment between any two distinct points on its graph lies above the graph between the two points.

Notice: Convex functions do not have maxima!

Source: Wikipedia

2D concave functions

If $f(x, y)$ is a convex function, then $-f(x, y)$ is a concave function.

A twice-differentiable function is concave if and only if its Hessian matrix (= matrix of second partial derivatives) is negative definite (= all of its eigenvalues are negative).

Concave functions have at least one maximum. They don't have any minima.

Notice: A convex, (respectively concave) function can have more than 2 coordinates.

Vector fields

A **vector** field on the xy plane is a vector function $\mathbf{a}(x, y)$ which evaluates to a vector at any point (x, y). This vector has a finite or infinite magnitude and direction.

For vector fields we can compute addition, subtraction, dot product, cross product and magnitude, e.g. for $\mathbf{a}$ and $\mathbf{b}$ on the xy plane

$$\mathbf{a}(x,y) + 4\mathbf{b}(x,y), \quad \mathbf{a}\cdot\mathbf{b}, \quad \mathbf{a}\times\mathbf{b}, \quad |\mathbf{a}|.$$

Let

$$\mathbf{a}(x,y) = (3x^2y - 4x)\hat{\mathbf{i}} + xy\hat{\mathbf{j}} \equiv \big(3x^2y - 4x, xy\big),$$

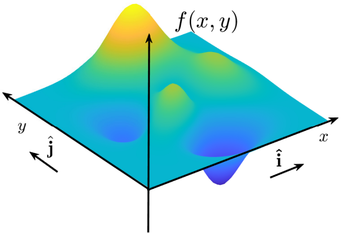
$$\mathbf{b}(x,y) = 4x\hat{\mathbf{i}} + 3\hat{\mathbf{j}} \equiv \big(4x, 3\big),$$

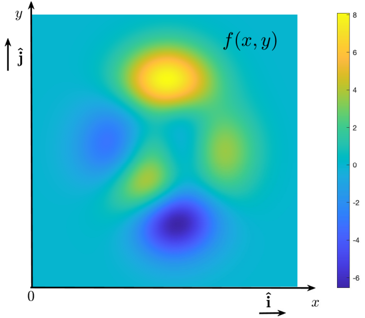
then

$$\mathbf{a}+\mathbf{b} = 3x^2y\hat{\mathbf{i}} + (xy+3)\hat{\mathbf{j}}, \quad \mathbf{a}\cdot\mathbf{b} = 12x^3y - 16x^2 + 3xy.$$

Scalar field variation in space

For a field $f(x, y)$ we may be interested in its minima or maxima or values in certain areas.





These figures show the function over some range of x and y values.

The vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ indicate the 'forward' directions in x and y respectively.

Partial derivatives

The partial derivative of a scalar field f with respect to a given direction, tells us how does the value of f varies in that direction.

On the xy plane (respectively xyz frame) we have

$$\frac{\partial f}{\partial x}(x,y), \quad \frac{\partial f}{\partial y}(x,y) \qquad \left(\text{resp. } \frac{\partial f}{\partial x}(x,y,z), \quad \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial z}\right),$$

If $\frac{\partial f}{\partial x}(x,y) > 0$ then $f(x+\delta x, y) > f(x,y)$ if δx is in the $\hat{\mathbf{i}}$ direction.

If $\frac{\partial f}{\partial x}(x,y) < 0$ then $f(x+\delta x, y) < f(x,y)$ if δx is in the $\hat{\mathbf{i}}$ direction.

Similarly, $\frac{\partial f}{\partial y}(x,y) > 0$ then $f(x, y+\delta y) > f(x,y)$ if δy is in the $\hat{\mathbf{j}}$ direction.

First and second partial derivatives

To compute the partial derivatives of a **scalar** field f with respect to a given variable we treat all other variables in f as constants.

For $f(x, y)$ continuous and twice differentiable everywhere on the xy plane, we have two first partial derivatives

$$f_x(x, y) = \frac{\partial f}{\partial x}, \quad f_y(x, y) = \frac{\partial f}{\partial y}$$

and four second partial derivatives

$$f_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial f_x}{\partial x}, \quad f_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial f_y}{\partial y}$$

$$f_{xy}(x, y) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial f_y}{\partial x}, \quad f_{yx}(x, y) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial f_x}{\partial y}$$

Notice that when differentiation order is reversible

$$f_{xy}(x, y) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}(x, y)$$

Partial derivatives example

Compute all first and second partial derivatives of

$$f(x,y) = 2x^3 - 3y^3 + 4x^2y - 2xy^2 + 5x - 1$$

Observation: f is differentiable everywhere in $-\infty < x, y < +\infty$ hence its derivatives are finite.

The first derivatives:

$$f_x = \frac{\partial f}{\partial x} = 6x^2 + 8xy - 2y^2 + 5$$

$$f_y = \frac{\partial f}{\partial y} = -9y^2 + 4x^2 - 4xy$$

Partial derivatives example

Compute all first and second partial derivatives of

$$f(x,y) = 2x^3 - 3y^3 + 4x^2y - 2xy^2 + 5x - 1$$

From the first partial derivatives

$$f_x = \frac{\partial f}{\partial x} = 6x^2 + 8xy - 2y^2 + 5, \qquad f_y = \frac{\partial f}{\partial y} = -9y^2 + 4x^2 - 4xy$$

the second partial derivatives are:

$$f_{xx} = \frac{\partial f_x}{\partial x} = 12x + 8y, \quad f_{yy} = \frac{\partial f_y}{\partial y} = -18y - 4x,$$

and

$$f_{xy} = \frac{\partial f_y}{\partial x} = 8x - 4y, \quad f_{yx} = \frac{\partial f_x}{\partial y} = 8x - 4y.$$

The gradient

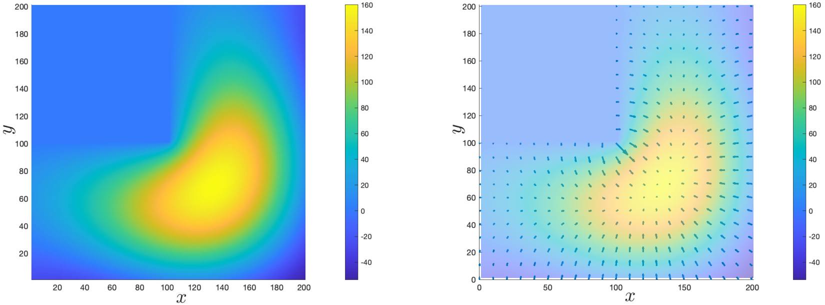
The gradient of a 2D scalar, continuous field $f(x, y)$ is a vector field

$$\nabla f(x, y) = \frac{\partial f}{\partial x}\hat{\mathbf{i}} + \frac{\partial f}{\partial y}\hat{\mathbf{j}}$$

Probed at any point (x, y) this is a vector pointing to the direction of maximum positive change in f and has a magnitude proportional to that.

If we want to find the maximum point of f from anywhere on the plane we 'follow' the gradient arrows. If we want to find the minimum of f we go opposite to the direction of the gradient.

The gradient



Left a scalar function, and right samples of its gradient on a coarser grid (for illustration clarity).

The gradient

In the 2D Cartesian frame, from the electrostatic example the gradient of the electric potential is

$$\nabla\phi(x,y) = \frac{\partial\phi}{\partial x}\hat{\mathbf{i}} + \frac{\partial\phi}{\partial y}\hat{\mathbf{j}} \equiv \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}\right)$$

and the induced Electric vector field[1] is

$$\mathbf{E}(x,y) = -\nabla\phi(x,y).$$

A scalar field f is constant (i.e. flat in 2D) if

$$\nabla f = (0,0) \Longleftrightarrow \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 \Longleftrightarrow |\nabla f| = 0$$

hence it does not vary in any direction.

[1]The minus sign is due to Maxwell's laws.

The Hessian

We saw that the two first partial derivatives of $f(x, y)$ find their way into the gradient vector

$$\nabla f(x,y) = \begin{pmatrix} f_x \\ f_y \end{pmatrix}, \quad f_x := \frac{\partial f}{\partial x}, \quad f_y := \frac{\partial f}{\partial y},$$

The four second partial derivatives can be tabulated in a 2×2 matrix called the Hessian

$$Hf(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Both the gradient vector and the Hessian matrix are functions of the coordinates, i.e. they can be evaluated at any (x, y).

Directional derivative

What's the variation in f in an arbitrary direction that's neither $\hat{\mathbf{i}}$ nor $\hat{\mathbf{j}}$?

Let $\mathbf{n}$ be a vector pointing in an arbitrary direction on the xy plane. The directional derivative is given by the dot product $\nabla f \cdot \hat{\mathbf{n}}$.

Since any vector $\mathbf{n}$ on the plane can be decomposed to two components in $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ directions respectively, then that information is contained within the gradient.

$$\nabla f \cdot \hat{\mathbf{n}} : \text{the component of } \nabla f \text{ in the } \hat{\mathbf{n}} \text{ direction}$$

Note that we must normalise $\mathbf{n}$ to get $\hat{\mathbf{n}} = \mathbf{n}/|\mathbf{n}|$.

This generalises to all vector fields: For any vector field $\mathbf{a}$, the quantity $\mathbf{a} \cdot \hat{\mathbf{n}}$ is the component of $\mathbf{a}$ in the direction of $\hat{\mathbf{n}}$.

Basic vector operations

For vector fields $\mathbf{a} = a_i(x, y)\hat{\mathbf{i}} + a_j(x, y)\hat{\mathbf{j}} \equiv (a_i, a_j)$ and $\mathbf{b} = b_i(x, y)\hat{\mathbf{i}} + b_j(x, y)\hat{\mathbf{j}} \equiv (b_i, b_j)$ recall:

The **magnitude** $|\mathbf{a}| = \sqrt{a_i^2 + a_j^2}$ is a non-negative scalar field.

The **dot product** (also known as inner or scalar product)

$$\mathbf{a} \cdot \mathbf{b} = a_i b_i + a_j b_j$$

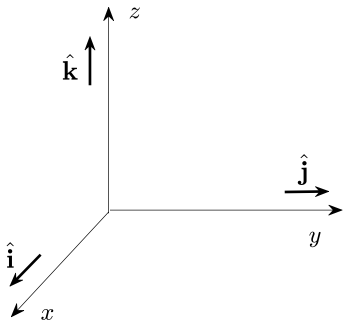
is a scalar field. When $\mathbf{a} \cdot \mathbf{b} = 0$ means $\mathbf{a}$ and $\mathbf{b}$ are orthogonal.

The **cross product** (also known as outer or vector product)

$$\mathbf{a} \times \mathbf{b} = (a_i b_j - a_j b_i)\hat{\mathbf{k}}$$

is a vector field orthogonal to the xy plane. Note $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

Going 3D



The 3D Cartesian frame xyz. Note the unit direction vectors. In this course all vectors with a 'hat' are unit magnitude vectors.

3D fields at the Cartesian frame

Extensions to the 3D space are

$$\mathbf{a} = a_i(x,y,z)\hat{\mathbf{i}} + a_j(x,y,z)\hat{\mathbf{j}} + a_k(x,y,z)\hat{\mathbf{k}} \equiv \left(a_i, a_j, a_k\right)$$

The gradient of $f(x,y,z)$ is

$$\nabla f = \frac{\partial f}{\partial x}\hat{\mathbf{i}} + \frac{\partial f}{\partial y}\hat{\mathbf{j}} + \frac{\partial f}{\partial z}\hat{\mathbf{k}}$$

The dot product of vector fields $\mathbf{a}$ and ∇f

$$\mathbf{a} \cdot \nabla f = a_i \frac{\partial f}{\partial x} + a_j \frac{\partial f}{\partial y} + a_k \frac{\partial f}{\partial z}$$

The cross product of vector fields $\mathbf{a}$ and $\mathbf{b} = \left(b_i, b_j, b_k\right)$ (best remembered as a determinant)

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_i & a_j & a_k \\ b_i & b_j & b_k \end{vmatrix}$$

Normal on a curve or a surface

For a curve embedded in 2D e.g. $f(x, y) = 0$ or a surface in 3D e.g. $f(x, y, z) = 0$ the normal vector is given by the gradient ∇f.

Let $x^2 + y^2 = r^2$ be the circle of radius r centred at the origin. The vector

$$\mathbf{n} = \nabla f = 2x\hat{\mathbf{i}} + 2y\hat{\mathbf{j}},$$

is (outward pointing) normal to the circle at a point (x, y).

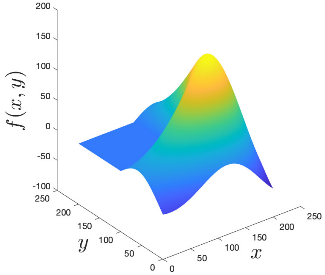
Let the xy plane whose proper definition is $z = 0$. We know that the normal on that plane is aligned to the z axes, e.g. $\pm\mathbf{k}$. To see this

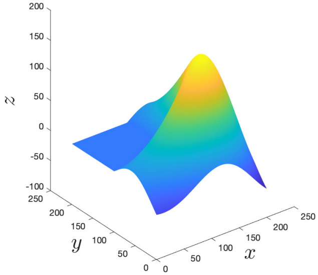
$$\mathbf{n} = \nabla f = 0\hat{\mathbf{i}} + 0\hat{\mathbf{j}} + \hat{\mathbf{k}}.$$

The next example shows that this also applies to 3D surfaces $f(x, y, z) = 0$.

Normal vector(s) on a surface

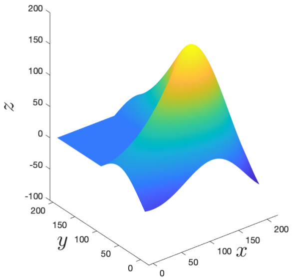
A word of caution: Read the axes labels! These two figures correspond to two very different things.

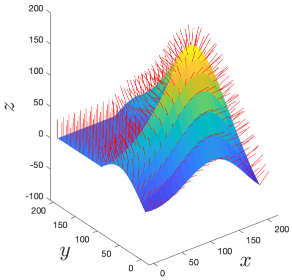




Left a graph of a 2D function in 3D space, and right, a 3D surface (function). Note the difference in the vertical axis label.

Normal vector(s) on a surface





Left a graph of the 3D surface, and right a sample of the (outward) unit normal vectors on this surface. Note that there are also inward pointing vectors.

Formulas

- The gradient

$$\nabla f(x,y,z) = \frac{\partial f}{\partial x}\hat{\mathbf{i}} + \frac{\partial f}{\partial y}\hat{\mathbf{j}} + \frac{\partial f}{\partial z}\hat{\mathbf{k}}$$

- The Hessian

$$Hf(x,y,z) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}$$

- The unit normal on $f(x,y,z) = 0$ is $\hat{\mathbf{n}} = \pm\frac{1}{|\nabla f|}\nabla f$
- The directional derivative of $f(x,y,z)$ is $\nabla f \cdot \hat{\mathbf{n}}$.

Main outcomes of module 1

You MUST know:

1. Partial derivatives, gradient and Hessian of scalar fields.
2. The component of vector field $\mathbf{a}$ in the direction of $\mathbf{n}$.
3. The normal vector on a surface.

I am assuming that you already know:

- Single variable (1D) differentiation.
- Single variable (1D) integration
- How to work out determinants for 2×2 and 3×3 matrices.