Functions of several variables:

A function f of n real variables is a rule that assigns a unique number $f(x_1, \cdots, x_n)$ to each point $(x_1, \cdots, x_n) \in D \subseteq \mathbb{R}^n$. The set D is called the domain of f.

Here, we will usually let $D = \mathbb{R}^n$.

Ex 1: $f(x,y) = x^2 \cdot \sin(y)$ $\quad f: \mathbb{R}^2 \to \mathbb{R}$

Ex 2: $f(x,y,z) = x^2 + y^2 \cdot z^4$ $\quad f: \mathbb{R}^3 \to \mathbb{R}$

Ex 3: $f(x,y) = x \cdot \sqrt{y}$ $\quad f: \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$, $\mathbb{R}^+ = [0, \infty)$

Partial derivatives

For simplicity, let $f = f(x,y)$.

The first partial derivative of f with respect to x and y are the functions given by

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

$$\frac{\partial f}{\partial y}(x,y) = \lim_{h \to 0} \frac{f(x,y+h) - f(x,y)}{h}$$

In general:
The partial derivative of f with respect to one variable is the derivative of the one-variable function obtained by keeping all other constant.

Ex 1: $f(x,y) = x^2 \cdot \sin(y)$, $\quad \frac{\partial f}{\partial x} = 2x \sin(y)$, $\quad \frac{\partial f}{\partial y} = x^2 \cdot \cos(y)$

Ex 2: $f(x,y,z) = x^2 + y^2 \cdot z^4$

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2yz^4, \quad \frac{\partial f}{\partial z} = 4y^2z^4$$

Notation: Unfortunately, there are many different notation for the partial derivatives, and they are not always used in a consistent way. So, given $f(x,y)$, ②

$$\frac{\partial f}{\partial x}\ ,\quad f_x\ ,\quad \partial_x f \qquad \text{are all the same.}$$

and similar for y, etc.

The gradient:

Given $f: \mathbb{R}^n \to \mathbb{R}$. The gradient of f is the vector function

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots \frac{\partial f}{\partial x_n}\right), \quad \text{so} \quad \nabla f : \mathbb{R}^n \to \mathbb{R}^n$$

Ex 1: $f(x,y) = x^2 \sin(y)$, $\vec{\nabla} f = (2x\sin(y),\ x^2\cos(y))$

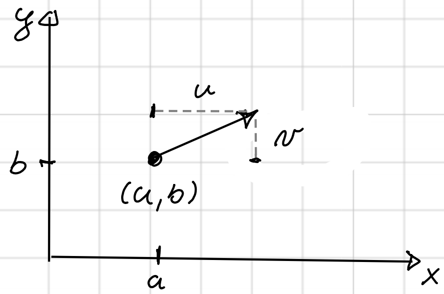
Ex 2: $f(x,y,z) = x^2 + y^2 z^4$, $\vec{\nabla} f = (2x,\ 2yz^4, 3y^2z^3)$

Directional derivative:

For simplicity, let $n = 2$.
Given a function $f(x,y)$, a point $\vec{a} = (a,b)$ and a unit vector $\vec{u} = (u,v)$ $(u^2 + v^2) = 1$.

$$\begin{aligned}\vec{x}(t) &= \vec{a} + t\cdot\vec{u}\\ &= (a + tu,\ b + tv), \quad t \in \mathbb{R}\\ &= (x(t),\ y(t))\end{aligned}$$

represent a straight line through $\vec{a}$ in the direction $\vec{u}$.



$f(\vec{x}(t))$ is then the cross section of the surface $f(x,y)$ along $\vec{x}(t)$. It is a function of t alone, so, by the chain rule we obtain:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt} = \frac{\partial f}{\partial x}\cdot u + \frac{\partial f}{\partial y}\cdot v = \vec{\nabla} f \cdot \vec{u}$$

which is called the directional derivative of f, in the direction $\vec{u}$, denoted as $D_{\vec{u}} f$.

Ex 1: $f(x,y) = x^2y + y^2$, $\vec{a} = (2,1)$, $\vec{u} = 1/\sqrt{2}\,(1,1)$

$\vec{\nabla} f = (2xy,\ x^2 + 2y)$

$\vec{\nabla} f(\vec{a}) = (2 \cdot 2 \cdot 1,\ 2^2 + 2 \cdot 1) = (4,6)$

$D_{\vec{u}} f(\vec{a}) = (4,6) \bullet \frac{1}{\sqrt{2}}(1,1) = 10/\sqrt{2}$

Ex 2: $f(x,y,z) = xe^z + y$, $\vec{a} = (1,1,0)$, $\vec{u} = \frac{1}{3}(2,2,1)$

$\vec{\nabla} f = (e^z,\ 1,\ xe^z)$

$\vec{\nabla} f(\vec{a}) = (e^0,\ 1,\ 1 \cdot e^0) = (1,1,1)$

$D_{\vec{u}} f(\vec{a}) = (1,1,1) \bullet \frac{1}{3}(2,2,1) = 5/3.$

Interpretation of the gradient:

At a point $\vec{a}$, the function $f(\vec{x})$ increase most rapidly in the direction of the vector $\vec{\nabla} f(\vec{a}) / \| \vec{\nabla} f(\vec{a}) \|$, and the maximum rate of increase is $\| \vec{\nabla} f(\vec{a}) \|$.

NB! $\| \vec{x} \| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$.

Jacobi-matrix (Jacobian).

Given $\vec{f} : \mathbb{R}^n \to \mathbb{R}^n$, thus

$$\vec{f}(\vec{x}) = \begin{pmatrix} f_1(x_1, x_2, \cdots x_n) \\ f_2(x_1, x_2, \cdots x_n) \\ \vdots \\ f_n(x_1, x_2, \cdots x_n) \end{pmatrix}$$

The Jacobian of f is the matrix function $J : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ with elements

$$J_{ij}(x) = \frac{\partial f_i}{\partial x_j}.$$

So, for $n = 2$, and

$$\vec{f}(\vec{x}) = \begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix}$$

the Jacobian is

$$J(\vec{x}) = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}$$

Ex: $\vec{f}(\vec{x}) = \begin{pmatrix} x^2 \cdot \sin(y) + z^3 \\ x \cdot z \\ x^2 + y^2 + e^{zy} \end{pmatrix}$

$$J(\vec{x}) = \begin{pmatrix} 2x \sin(y) & x^2 \cos(y) & 3z^2 \\ z & 0 & x \\ 2x & 2y + z e^{zy} & y e^{zy} \end{pmatrix}$$

Higher order partial derivatives.

This is again best described by example.
Let $n=2$. Given $f(x,y)$ $\quad (f: \mathbb{R}^2 \to \mathbb{R})$

Then
$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y^2}\right) \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

NB! $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$ (there are exceptions, those will not be covered in this course)

Notation: $\dfrac{\partial^2 f}{\partial x^2} = f_{xx} = \partial_x^2 f$, $\dfrac{\partial^2 f}{\partial y^2} = f_{yy} = \partial_y^2 f$
$$\frac{\partial^2 f}{\partial x \partial y} = f_{xy} = \partial_x \partial_y f$$

The Hessian (Hess matrix) is the matrix function
$$H_f(\vec{x}) \text{ with elements } \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

So if $f = f(x,y)$, then
$$H_f(\vec{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

- The Hessian is the Jacobian of the gradient.

Ex: $f(x,y) = x^2 y^3$
$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2xy^3, 3x^2y^2)$$
$$\frac{\partial^2 f}{\partial x^2} = 2y^3, \quad \frac{\partial^2 f}{\partial x \partial y} = 6xy^2, \quad \frac{\partial^2 f}{\partial y^2} = 6x^2 y$$

and
$$H_f(\vec{x}) = \begin{pmatrix} 2y^3 & 6xy^2 \\ 6xy^2 & 6x^2y \end{pmatrix}$$

The chain rule:

Let $f: \mathbb{R}^n \to \mathbb{R}$ and $\vec{x}: \mathbb{R} \to \mathbb{R}^n$, e.g.

$$g(t) = f(x_1(t), x_2(t), \cdots, x_n(t)).$$

Then $g(t)$ is a function of one variable, and the total derivative of g (with respect to t) is

$$\frac{dg}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n} \cdot \frac{dx_n}{dt} = \vec{\nabla} f \cdot \frac{d\vec{x}}{dt}$$

NB! We use $\frac{d}{dt}$ if there is only one free variable.

Ex: $\vec{x}(t) = \big(x(t), y(t)\big) = (\cos(t), \sin(t))$

$f(x, y) = x^2 \cdot y^3$ and $g(t) = f(x(t), y(t))$, then

$$\frac{dg}{dt} = 2 \cdot x(t) \cdot y(t)^3 \cdot \frac{dx}{dt} + 3x(t)^2 \cdot y(t)^2 \cdot \frac{dy}{dt}$$

$$= 2\cos(t) \cdot \sin^3(t) \cdot (-\sin(t)) + 3\cos^2(t) \cdot \sin(t)^2 \cdot \cos(t)$$

$$= \cos(t) \cdot \sin^2(t) \cdot \big(3\cos^2(t) - 2\sin^2(t)\big)$$

Taylor-series:

Given $f: \mathbb{R}^n \to \mathbb{R}$, $\vec{a} \in \mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$.
Then

$$f(\vec{a} + \vec{v}) = f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot \vec{v} + \vec{v}^T \cdot H_f(\vec{a}) \cdot \vec{v} + \cdots .$$