$f : I \to \mathbb{R}$ a function defined on an interval $I \subset \mathbb{R}$

f is differentiable at $c \in I$ if the limit

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \quad \text{exists}$$

this limit is called the derivative of f at c, denoted:

$$f'(c), \quad \left.\frac{df}{dx}\right|_{x=c}, \quad \frac{df}{dx}(c), \quad \dot{f}(c)$$

If f is differentiable at every $c \in I$, then

$$\frac{df}{dx} : I \to \mathbb{R} \quad \text{is a function}$$

equivalent definition: let $h = x - c$ (change of variable)

then $x = c + h$

and $x \to c \iff h \to 0$

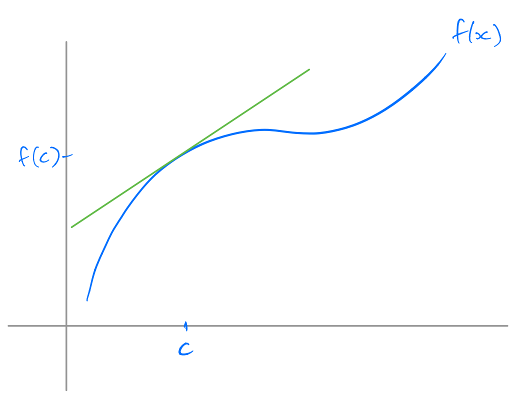
therefore $$f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}$$

Intuition: if $$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c)$$

then when x is close to c, $\frac{f(x) - f(c)}{x - c}$ is close to $f'(c)$

i.e.: $$\frac{f(x) - f(c)}{x - c} \approx f'(c)$$

rearranging: $$f(x) \approx \underbrace{f'(c)(x-c) + f(c)}$$



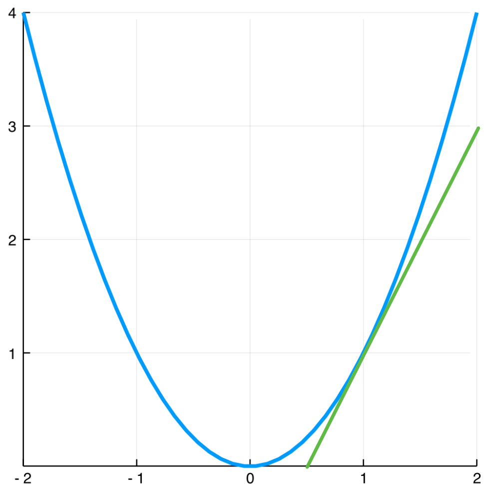
this is the equation for a line with slope $f'(c)$ passing through the point $(c, f(c))$

Example $\quad f(x) = x^2 \qquad f'(c) = 2c \qquad (f'(x) = 2x)$

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{x\to c} \frac{x^2-c^2}{x-c} = \lim_{x\to c} \frac{\cancel{(x-c)}(x+c)}{\cancel{x-c}}$$

$$= \lim_{x\to c} x+c$$

$$= 2c$$



$f'(1) = 2$

so when $x \approx 1$:

$$f(x) \approx f'(1)(x-1) + f(1)$$
$$= 2(x-1) + 1$$
$$= 2x - 1$$

i.e. $y = 2x-1$ is the line which best approximates x^2 at $x=1$. This is called the tangent line at $x=1$

Differentiation rules

If $f, u, v : I \to \mathbb{R}$ are differentiable functions then

$$\frac{d}{dx}(u+v) = \frac{du}{dx} + \frac{dv}{dx}$$

$c \in \mathbb{R}$: $\quad \frac{d}{dx}(cf) = c\,\frac{df}{dx}$

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$$ product rule

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$ quotient rule

$$\frac{d}{dx} f(u(x)) = \frac{df}{du}\frac{du}{dx}$$ chain rule

Functions take an input and assign a single output

$f \quad : \quad x \quad \longmapsto \quad y = f(x)$

They can be many-to-one, meaning two different inputs x_1, x_2, $x_1 \neq x_2$, might have the same output: $f(x_1) = f(x_2)$

e.g: $f(x) = x^2$. $\quad f(-1) = 1 = f(1)$

Or they can be one-to-one meaning that distinct inputs x_1, x_2, $x_1 \neq x_2$ always have distinct outputs:

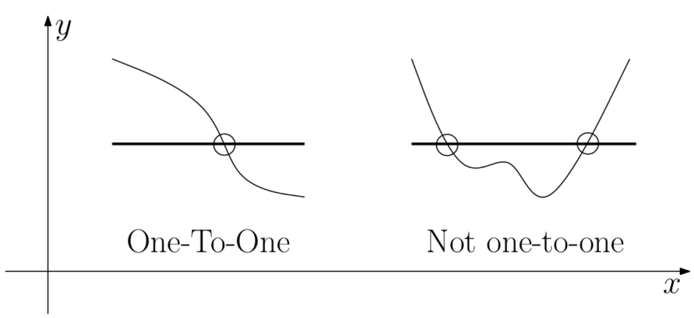
$$f(x_1) \neq f(x_2)$$

eg: $f(x) = x$.

If $x_1 \neq x_2$ then $f(x_1) = x_1 \neq x_2 = f(x_2)$

$$f(x_1) \neq f(x_2)$$

the horizontal line test:



A natural question: given an output y, what was the input?
i.e. find x such that $y = f(x)$.

- If f is many-to-one there can be multiple solutions

  eg: $f(x) = x^2$, $y = 4$. then $4 = f(x) = x^2$

  $$x = \pm 2$$

- If f is one-to-one then there is only one answer to this question and therefore the operation

$$\text{output} \longmapsto \text{corresponding input}$$
$$y \longmapsto x \text{ such that } y = f(x)$$

is a function.

It is called the inverse function and denoted f^{-1}, $f^{-1}(y)$

Note: $f^{-1}(y)$ is **not** $\dfrac{1}{f(y)}$

Example $f(x) = x + 1$

to find f^{-1}, solve $y = x + 1$ for x:

$$x = y - 1$$

then $f^{-1}(y) = y - 1$, or since y is just a name: $f^{-1}(x) = x - 1$

Properties of inverse functions

domain f^{-1} = range f (because f^{-1} takes outputs of f as its inputs)

$f^{-1}(f(x)) = x$ f^{-1} "undoes" whatever f did

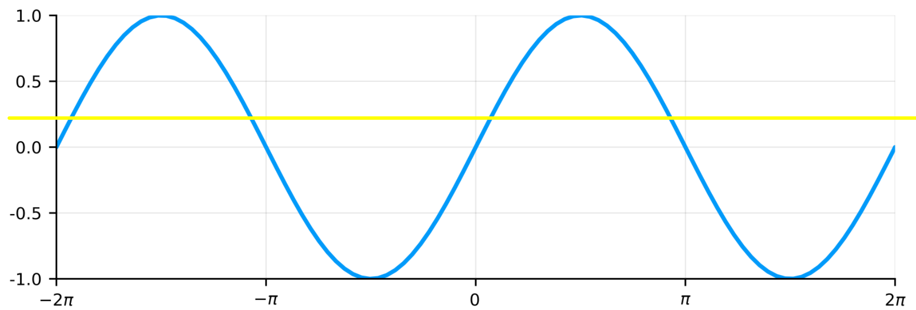
$f(f^{-1}(x)) = x$ f " " f^{-1} "

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$$ Inverse Function Theorem

when $f'(x) \neq 0$

$$f : \mathbb{R} \to \mathbb{R}$$

$$f(x) = \sin x$$

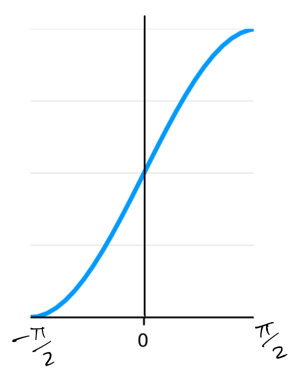


this function is not one-to-one. However, if we restrict the domain:

$$f : [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \to \mathbb{R}$$

$$x \mapsto \sin x$$

then the function is one-to-one:



so an inverse function exists.

Notation: $\sin^{-1}(x)$, $\arcsin(x)$, $\operatorname{asin}(x)$

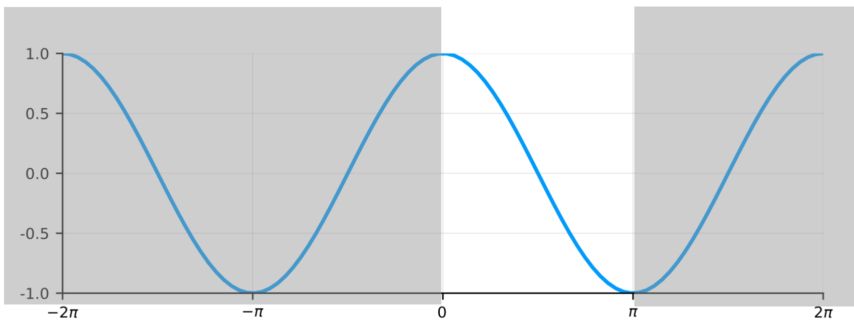
If $\sin x = y$ then $\sin^{-1}(\sin x) = x = \sin^{-1}(y)$

remember the domain of f^{-1} is the range of f, so

$$\sin^{-1} : [-1, 1] \longrightarrow [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]$$

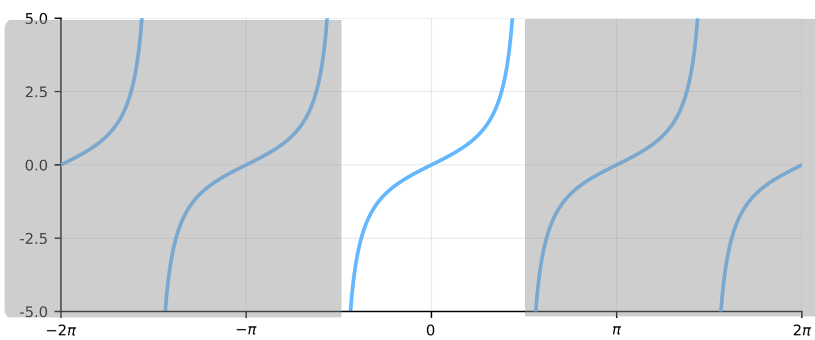
Remark: sin is also one-to-one when restricted to other intervals, eg: $[\frac{\pi}{2}, \frac{3\pi}{2}]$, but it is conventional to use the domain $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

For $\cos x$ the convention is to restrict to $[0,\pi]$



so $\cos^{-1} : [-1, 1] \longrightarrow [0, \pi]$

$\tan x$ is restricted to $(-\frac{\pi}{2}, \frac{\pi}{2})$



and $\tan^{-1} : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$

Derivatives

Inverse function theorem:

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))} \qquad \text{when } f'(x) \neq 0$$

therefore if $f(x) = \sin x$

$$\frac{d}{dx}(\sin^{-1})(x) = \frac{1}{\cos(\sin^{-1} x)}$$

This can be simplified

$$\sin^2\theta + \cos^2\theta = 1 \quad \Rightarrow \quad \cos\theta = \sqrt{1 - \sin^2\theta}$$

$$\cos(\sin^{-1} x) = \sqrt{1 - (\sin(\sin^{-1} x))^2}$$

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

therefore

$$\frac{d}{dx}(\sin^{-1})(x) = \frac{1}{\sqrt{1 - x^2}}$$

by similar methods:

$$\frac{d}{dx}(\cos^{-1})(x) = \frac{-1}{\sqrt{1 - x^2}}$$

$$\frac{d}{dx}(\tan^{-1})(x) = \frac{1}{1 + x^2}$$

you will need to memorise these derivatives.

$\mathbf{r} : I \to \mathbb{R}^n \qquad \mathbf{r}(t) = (r_1(t), r_2(t), \dots, r_n(t))$

coordinate functions.

$t_0 \in I$, define

$$\mathbf{r}'(t_0) = \lim_{t \to t_0} \frac{\mathbf{r}(t) - \mathbf{r}(t_0)}{t - t_0}$$

← this is a vector!

recalling that for any vector valued function

$$\mathbf{f}(t) = (f_1(t), \dots, f_n(t))$$

$$\lim_{t \to c} \mathbf{f}(t) = \left(\lim_{t \to c} f_1(t), \dots, \lim_{t \to c} f_n(t) \right)$$

it follows that

$$\mathbf{r}'(t) = \left(\lim_{t \to t_0} \frac{r_1(t) - r_1(t_0)}{t - t_0}, \dots, \lim_{t \to t_0} \frac{r_n(t) - r_n(t_0)}{t - t_0} \right)$$

$$\mathbf{r}'(t) = (r_1'(t), \dots, r_n'(t))$$

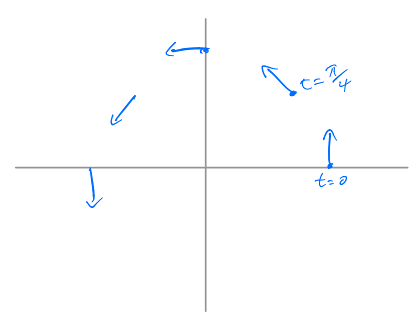
Example $\quad \mathbf{r} : [0, 2\pi) \to \mathbb{R}^2, \quad \mathbf{r}(t) = (\cos t, \sin t)$

$$\mathbf{r}'(t) = (-\sin t, \cos t)$$

$t = 0 \qquad \mathbf{r}(t) = (1, 0), \quad \mathbf{r}'(t) = (0, 1) \quad \uparrow$

$t = \frac{\pi}{4} \qquad \mathbf{r}(t) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \mathbf{r}'(t) = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \nwarrow$

$t = \frac{\pi}{2} \qquad \mathbf{r}(t) = (0, 1), \quad \mathbf{r}'(t) = (-1, 0) \quad \leftarrow$



→ tangent vector! (vector in the direction of tangent line) if $\underset{\sim}{r}'(t_0) = \lim\limits_{t \to t_0} \dfrac{\underset{\sim}{r}(t) - \underset{\sim}{r}(t_0)}{t - t_0}$

then when t is close to t_0:

$$\underset{\sim}{r}'(t_0) \approx \frac{\underset{\sim}{r}(t) - \underset{\sim}{r}(t_0)}{t - t_0}$$

rearranging:

$$\underset{\sim}{r}(t) \approx \underbrace{\underset{\sim}{r}'(t_0)}_{\text{direction vector}} (t - t_0) + \underbrace{\underset{\sim}{r}(t_0)}_{\text{starting point}} \to \text{line in } \mathbb{R}^n$$

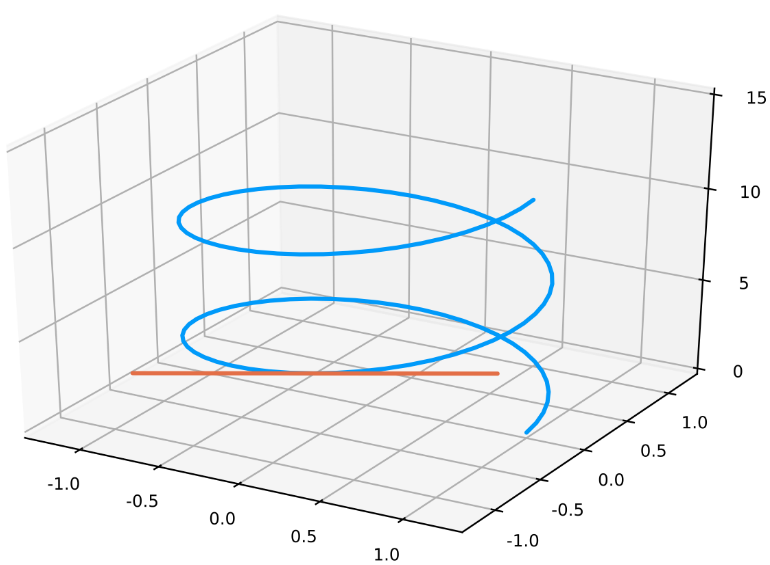
this is the line that best approximates $\underset{\sim}{r}(t)$ at $t = t_0$,
i.e. the tangent line.

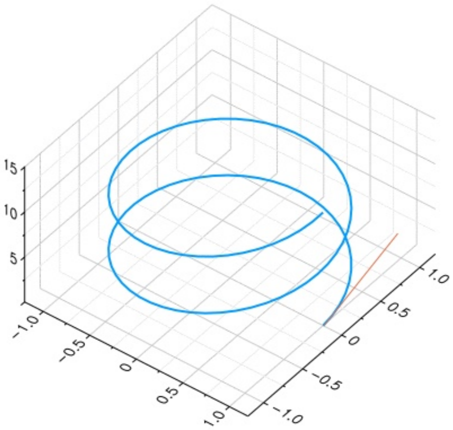
$\underset{\sim}{r}'(t_0)$ is the direction vector of the tangent line at $\underset{\sim}{r}(t_0)$.

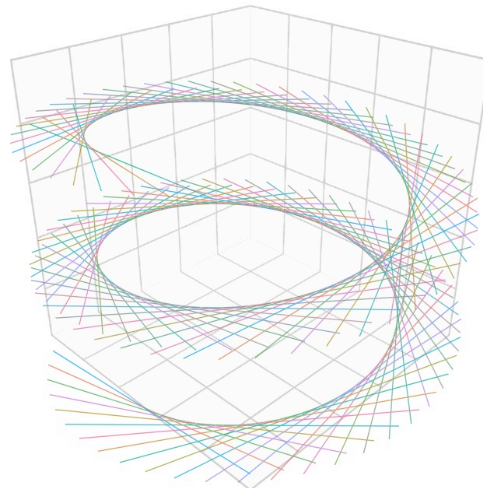
Example $\underset{\sim}{r}(t) = (\cos t, \sin t, t)$

$\underset{\sim}{r}'(t) = (-\sin t, \cos t, 1)$

tangent lines: $\underset{\sim}{r}'(t_0)(t - t_0) + \underset{\sim}{r}(t_0)$ plotted below:







Typical application: $\underset{\sim}{r}(t)$ is the position at time t of an object moving through space.

the trajectory of the object is the curve

$$C = \{ \underset{\sim}{r}(t) : t \in I \subset \mathbb{R} \}$$

the velocity at time t is the rate of change of position

$$\underset{\sim}{v}(t) = \underset{\sim}{r}'(t)$$

the speed at time t is the magnitude of velocity

$$v(t) = |\underset{\sim}{v}(t)| = |\underset{\sim}{r}'(t)|$$

no underline

the acceleration at time t is the rate of change of velocity

$$\underset{\sim}{a}(t) = \underset{\sim}{v}'(t) = \underset{\sim}{r}''(t)$$

Newton's second law of motion

$$\underset{\sim}{F} = m \underset{\sim}{a}$$

force acting on an object with mass m

Differentiation rules for vector valued functions

$\underset{\sim}{u} : I \to \mathbb{R}^n \qquad \underset{\sim}{v} : I \to \mathbb{R}^n$ differentiable vector valued functions

$$\frac{d}{dt}\left(\underset{\sim}{u}(t) + \underset{\sim}{v}(t)\right) = \underset{\sim}{u}'(t) + \underset{\sim}{v}'(t)$$

$$\frac{d}{dt}\left(c\,\underset{\sim}{u}(t)\right) = c\,\underset{\sim}{u}'(t) \qquad \text{for any } c \in \mathbb{R}$$

Product rules

$f : I \to \mathbb{R}$

$$\frac{d}{dt}\left(f(t)\,\underset{\sim}{u}(t)\right) = f'(t)\,\underset{\sim}{u}(t) + f(t)\,\underset{\sim}{u}'(t)$$

$$\frac{d}{dt}\,\underset{\sim}{u}(t) \cdot \underset{\sim}{v}(t) = \underset{\sim}{u}'(t) \cdot \underset{\sim}{v}(t) + \underset{\sim}{u}(t) \cdot \underset{\sim}{v}'(t)$$

dot product

Chain rule

If $\alpha : I \to I$ then

$$\frac{d}{dt}\,\underset{\sim}{u}(\alpha(t)) = \alpha'(t)\,\underset{\sim}{u}'(\alpha(t))$$

Examples

$$\underset{\sim}{c}(t) = (\cos t, \sin t)$$
$$= (-\sin t, \cos t)$$

Let $\alpha(t) = 2\pi t$ and $\underset{\sim}{r}(t) = \underset{\sim}{c}(\alpha(t))$

then $\underset{\sim}{r}'(t) = \alpha'(t)\,\underset{\sim}{c}'(\alpha(t))$ (chain rule)

$$= 2\pi\,(-\sin 2\pi t, \cos 2\pi t)$$

Compare speeds:

$$|\underset{\sim}{c}'(t)| = \sqrt{\sin^2 t + \cos^2 t} = 1$$

$$|\underset{\sim}{r}'(t)| = \sqrt{4\pi^2 \sin^2 2\pi t + 4\pi^2 \cos^2 2\pi t} = 2\pi\sqrt{\sin^2 2\pi t + \cos^2 2\pi t}$$
$$= 2\pi$$

$\underset{\sim}{r}(t)$ faster!

Find $\frac{d}{dt} |\underset{\sim}{c}(t)|^2 = \frac{d}{dt} \underset{\sim}{c}(t) \cdot \underset{\sim}{c}(t)$

by the (dot) product rule:

$$= \underset{\sim}{c}'(t) \cdot \underset{\sim}{c}(t) + \underset{\sim}{c}(t) \cdot \underset{\sim}{c}'(t)$$

$$= 2\underset{\sim}{c}'(t) \cdot \underset{\sim}{c}(t)$$

$$= 2(-\sin t, \cos t) \cdot (\cos t, \sin t)$$

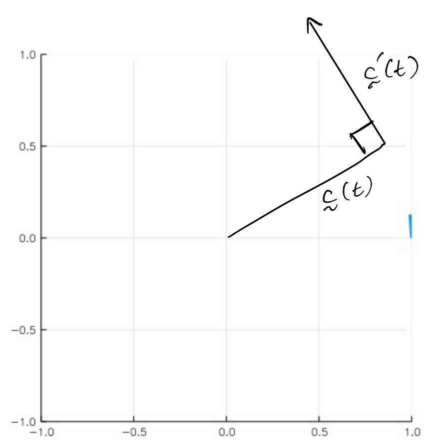
$$= 2(-\sin t \cos t + \cos t \sin t)$$

$$= 0$$

this is expected because $\underset{\sim}{c}(t)$ is a parametrization of the unit circle $\rightarrow |\underset{\sim}{c}(t)|^2 = 1$

therefore $\frac{d}{dt} |\underset{\sim}{c}(t)|^2 = 0$ ($\frac{d}{dt}$ each side)

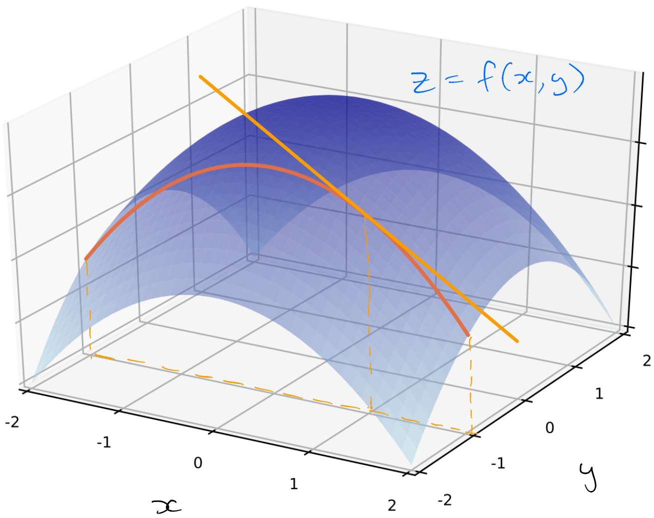
notice also that $\underset{\sim}{c}'(t) \cdot \underset{\sim}{c}(t) = 0$, i.e. the tangent vector $c'(t)$ is perpendicular to the position $\underset{\sim}{c}(t)$



Partial derivatives

how can we find tangent vectors / lines to a surface?

eg: $f(x,y) = -x^2 - y^2$



fix $y = -1$

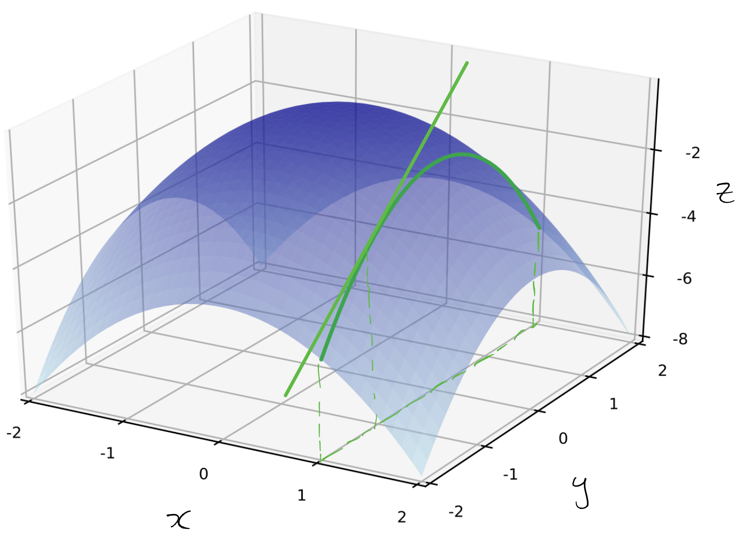
$z = f(x,-1) = -x^2 - 1$

the slope at $(x,-1)$ in the x direction is

$\frac{d}{dx}(-x^2 - 1) = -2x$

in particular at $x = 1$

slope $= -2$



fix $x = 1$

$z = f(1,y) = -1 - y^2$

the slope at $(1,y)$ in the y direction is

$\frac{d}{dy}(-1 - y^2) = -2y$

in particular at $y = -1$

slope $= 2$

These slopes are called partial derivatives

Notation $\quad \frac{\partial}{\partial x} f(1,-1) \quad$ partial derivative of f with respect to x at $(1,-1)$

$\frac{\partial}{\partial y} f(1,-1) \quad$ partial derivative of f with respect to y at $(1,-1)$

Formally:

$$\frac{\partial}{\partial x} f(a,b) = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x-a} \qquad y \text{ fixed } = b$$

$$\frac{\partial}{\partial y} f(a,b) = \lim_{y \to b} \frac{f(a,y) - f(a,b)}{y-b} \qquad x \text{ fixed } = a$$

In practice this means that to find:

$\frac{\partial f}{\partial x}(a,b)$ take the derivative of $f(x,y)$ wrt x while treating y as though it were a constant, and then substitute $x = a$, $y = b$

$\frac{\partial f}{\partial y}(a,b)$ take the derivative of $f(x,y)$ wrt y while treating x as though it were a constant, and then substitute $x = a$, $y = b$

Examples

$f(x,y) = -x^2 - y^2$, find $\frac{\partial f}{\partial x}(1,-1)$

$\frac{\partial}{\partial x}$ treating y as a constant: $\frac{\partial f}{\partial x} = -2x$

$$\frac{\partial f}{\partial x}(1,-1) = -2(1) = -2$$

$g(x,y) = x^2 + xy + y^2$, find $\frac{\partial g}{\partial y}(1,1)$

$$\frac{\partial g}{\partial y} = 0 + x + 2y \qquad \frac{\partial g}{\partial y}(1,1) = 1 + 2(1) = 3$$

Find $\frac{\partial g}{\partial x}(a,b)$:

$$\frac{\partial g}{\partial x} = 2x + y = \frac{\partial g}{\partial x}(x,y)$$

$$\frac{\partial g}{\partial x}(a,b) = 2a + b$$

All of this can be extended to functions $f : \mathbb{R}^3 \to \mathbb{R}$ or even $\mathbb{R}^n \to \mathbb{R}^m$

(we just can't visualise it as well)

eg:

$$\frac{\partial f}{\partial z}(a,b,c) = \lim_{z \to c} \frac{f(a,b,z) - f(a,b,c)}{z - c}$$

to find $\frac{\partial f}{\partial z}$ pretend <u>both</u> x and y are constants.

Examples

$f(x,y,z) = xye^z + \sin(xy)$, find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$

$$\frac{\partial f}{\partial x} = ye^z + y\cos(xy)$$

$$\frac{\partial f}{\partial y} = xe^z + x\cos(xy)$$

$$\frac{\partial f}{\partial z} = xye^z + 0$$

$g : \mathbb{R}^4 \to \mathbb{R}$ $\quad \underset{\sim}{x} = (x_1, x_2, x_3, x_4)$

$g(\underset{\sim}{x}) = x_1 x_2 - x_3 x_4$ $\quad$ Find $\frac{\partial g}{\partial x_4}$

$$\frac{\partial g}{\partial x_4} = 0 - x_3 = -x_3$$

There are several common notational alternatives for partial derivatives:

$$\underbrace{\frac{\partial f}{\partial x}, \frac{\partial}{\partial x} f, f_x}, \partial_x f, f'_x, \partial_1 f, D_1 f$$

we will mostly use the first three. Sometimes the function's arguments are written: eg: $f_x(x,y,z)$, often they are suppressed: f_x

Second partial derivatives

Suppose $f: D \subset \mathbb{R}^2 \to \mathbb{R}^2$ has partial derivative

$$\frac{\partial f}{\partial x} : D \to \mathbb{R}$$

which is differentiable with respect to x. Then

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

is called the second partial derivative with respect to x.

Alternative notation:

$$f_{xx} = \frac{\partial}{\partial x} f_x = \frac{\partial^2}{\partial x^2} f = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \quad \text{etc.}$$

Examples

$g(x,y) = x^2 + xy + y^2$

$g_x = 2x + y$

$g_{xx} = 2$

$f(x,y) = e^{2y} \sin x$

$f_x = e^{2y} \cos x$

$f_{xx} = -e^{2y} \sin x$

Similarly we define (assuming they exist)

$$f_{xy} = \frac{\partial}{\partial y} f_x = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} f\right) = \frac{\partial^2}{\partial y \partial x} f$$

$$f_{yy} = \frac{\partial}{\partial y} f_y = \frac{\partial^2}{\partial y^2} f$$

$$f_{yx} = \frac{\partial}{\partial x} f_y = \frac{\partial^2}{\partial x \partial y} f$$

f_{xy} and f_{yx} are called mixed partial derivatives

If the mixed partial derivatives are defined and continous on $\mathbb{R}^2$ then they are equal (Clairaut's / Schwarz' theorem)

Example

$$f(x,y) = e^{2y} \sin x$$

$$f_x = e^{2y} \cos x \qquad f_y = 2e^{2y} \sin x$$

$$f_{xx} = -e^{2y}\cos x, \quad f_{xy} = 2e^{2y}\cos x = f_{yx} = 2e^{2y}\cos x, \quad f_{yy} = 4e^{2y}\sin x$$

we can also calculate third-order partial derivatives

eg: $f_{xxy} = \frac{\partial}{\partial y}\frac{\partial}{\partial x}\frac{\partial}{\partial x} f = -2e^{2y}\cos x$

and fourth-order ...

and partial derivatives of functions of three or more variables

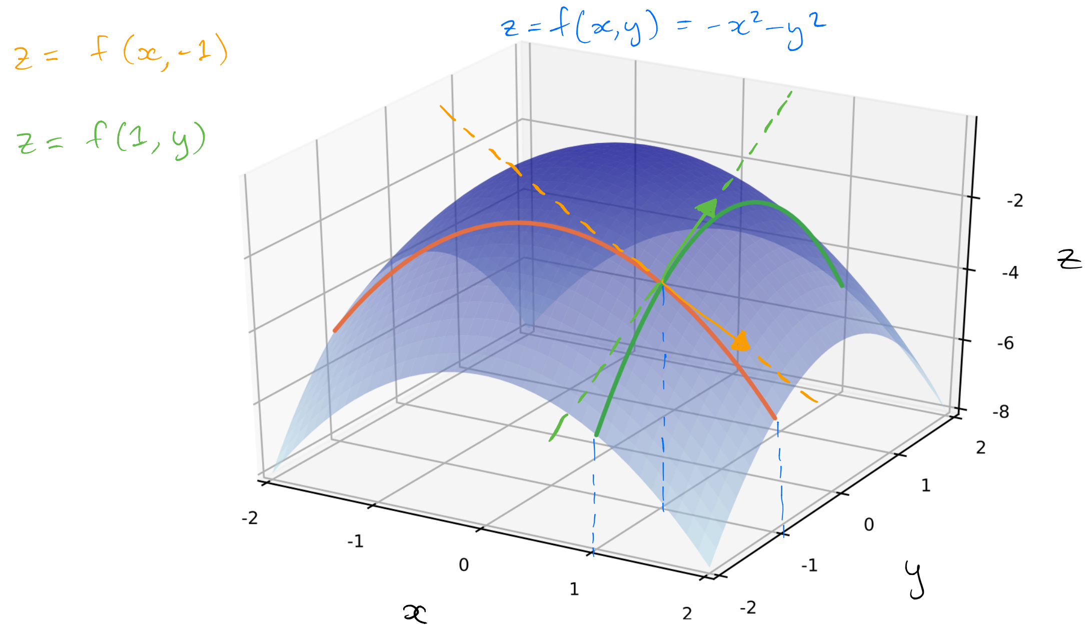
eg: $F(x,y,z) = xy - xz$

$$F_x = y - z, \quad F_{xz} = -1, \quad F_{xzy} = 0$$

$$F_{xy} = 1, \quad F_{xyz} = 0$$

Tangent vectors

Recall: The partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ give the slopes of tangent lines in the x and y directions respectively



direction vectors for the tangent lines (a.k.a tangent vectors) at $f(1,-1)$ are given by

$\left(1, 0, \frac{\partial f}{\partial x}(1,-1)\right)$ $\leftarrow$ for an increase of 1 in the x direction there is an increase/decrease of $\frac{\partial f}{\partial x}(1,-1)$ in the z direction

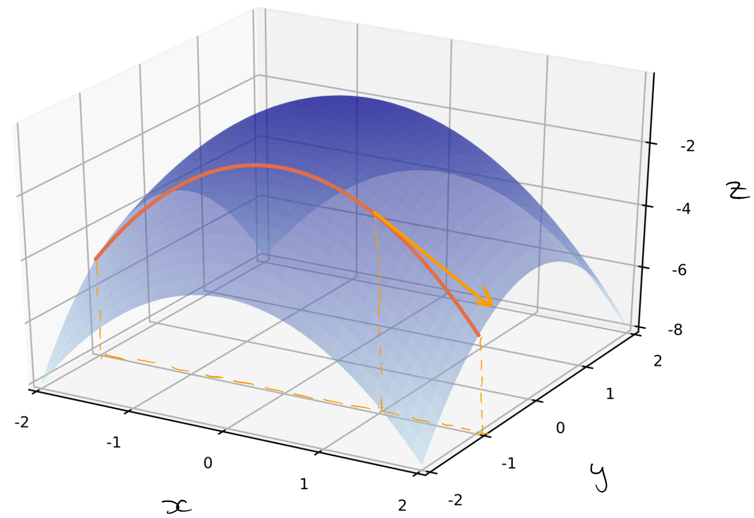
$= (1, 0, -2)$

$\left(\text{slope} = \frac{\text{rise}}{\text{run}}\right)$

$\left(0, 1, \frac{\partial f}{\partial y}(1,-1)\right)$ $\leftarrow$ for an increase of 1 in the y direction there is an increase/decrease of $\frac{\partial f}{\partial y}(1,-1)$ in the z direction

$= (0, 1, 2)$

We can also calculate these vectors as follows.



$\text{graph } f = \{(x, y, f(x,y))\}$

fix $y = -1$, this gives the orange curve $\{(x, -1, f(x,-1))\}$

which we can parametrize by $\underset{\sim}{r} : \mathbb{R} \to \mathbb{R}^3$

$$\underset{\sim}{r}(x) = (x, -1, f(x,-1))$$

we find tangent vectors to this curve by taking the derivative with respect to x

$$\underset{\sim}{r}'(x) = \left(1, 0, \frac{d}{dx}[f(x,-1)]\right) = \left(1, 0, \frac{\partial f}{\partial x}(x,-1)\right)$$

For the given example $f(x,y) = -x^2 - y^2$

$$\underset{\sim}{r}(x) = (x, -1, -x^2 - 1)$$

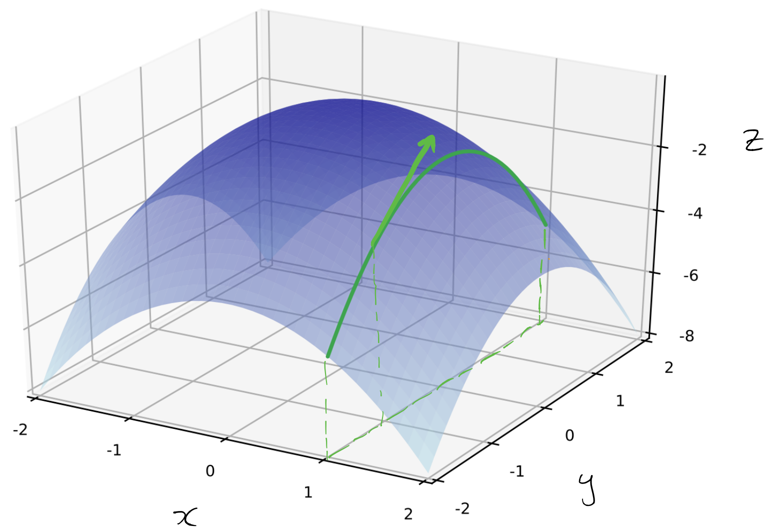
$$\underset{\sim}{r}'(x) = (1, 0, -2x) \qquad \left(\text{cf. } \frac{\partial f}{\partial x} = -2x\right)$$

$$\underset{\sim}{r}'(1) = (1, 0, -2)$$ (this is the vector in the picture).

tangent line in the x-direction at $(1,-1,-2)$:

$$\underset{\sim}{s}(t) = (1,-1,-2) + t\,(1,0,-2)$$

a tangent vector in the y direction.



fix $x = 1$ (green curve). Parametrize by

$$\underset{\sim}{q}(y) = (1, y, f(1,y)) = (1, y, -1-y^2)$$

$$\underset{\sim}{q}'(y) = \left(0, 1, \frac{\partial f}{\partial y}(1,y)\right) = (0, 1, -2y)$$

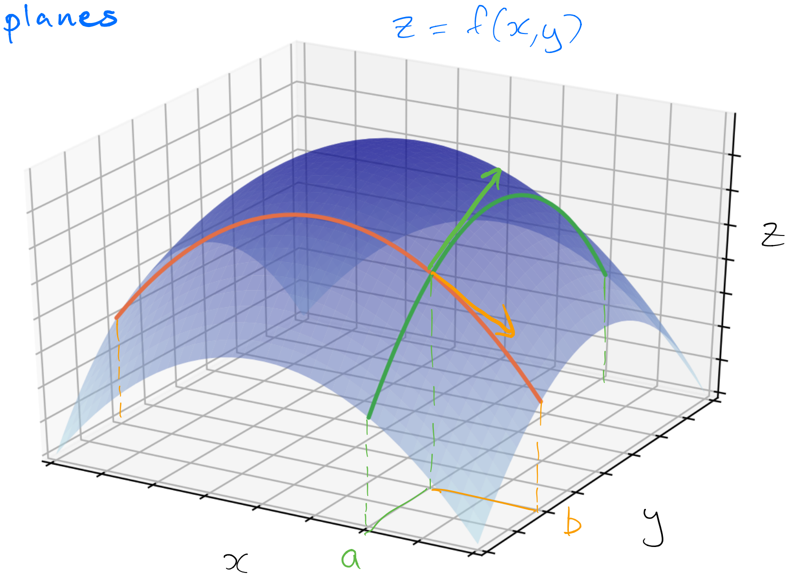
at $y = -1$:

$$\underset{\sim}{q}'(-1) = (0, 1, 2) \qquad \text{(pictured)}.$$

An equation for the tangent line in the y direction at $(1,-1,-2)$

$$\underset{\sim}{c}(t) = (1,-1,-2) + t(0,1,2)$$

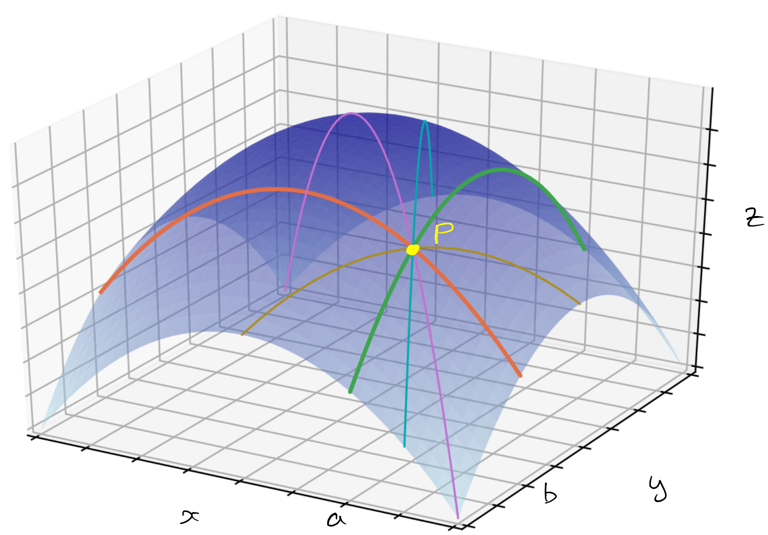
Tangent planes



The tangent vectors at $f(a,b)$ are

$$\left(1, 0, \frac{\partial f}{\partial x}(a,b)\right) \quad \text{and} \quad \left(0, 1, \frac{\partial f}{\partial y}(a,b)\right)$$

There are other curves on the surface passing through the point $P = (a, b, f(a,b))$:



$\to$ there are infinitely many tangent lines (and vectors) at this point, but they all lie in a common plane, called the tangent plane at P

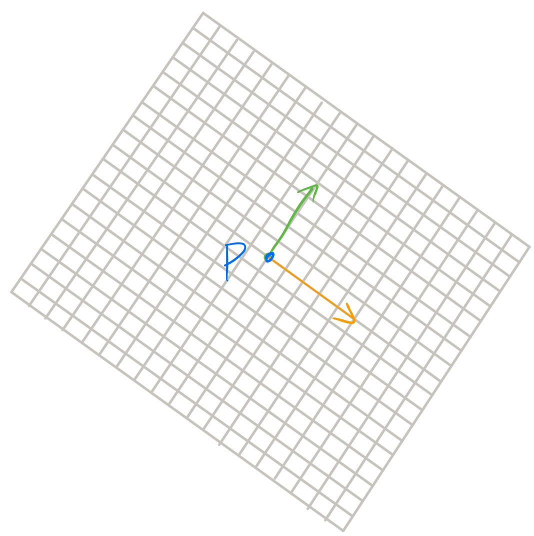
How can we describe this plane mathematically?

recall: equation of the tangent line in x-direction

$$\underset{\sim}{r}(t) = \underset{\substack{\uparrow \\ \text{starting} \\ \text{point } (P)}}{(a, b, f(a,b))} + \underset{\substack{\uparrow \\ \text{scalar} \times}}{t} \underbrace{\left(1, 0, \frac{\partial f}{\partial x}(a,b)\right)}_{\text{direction vector}}$$

starting point (P) + scalar × direction vector

observation: any point in the tangent plane can be obtained as a P + a sum of scalar multiples of the tangent vectors in the x and y directions



i.e. if $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is in the tangent plane then

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ f(a,b) \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$$

starting point P + scalar × direction vector + scalar × different direction

i.e.

$$\left.\begin{aligned} x &= a + \alpha \\ y &= b + \beta \\ z &= f(a,b) + \alpha \frac{\partial f}{\partial x}(a,b) + \beta \frac{\partial f}{\partial y}(a,b) \end{aligned}\right\} \text{parametric equations for tangent plane}$$

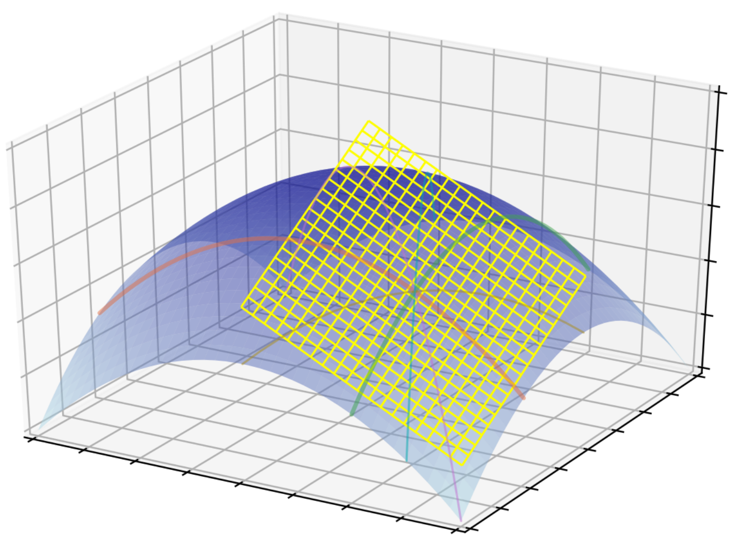
we can also characterise the tangent plane by an implicit equation, i.e. by writing z in terms of x and y:

rearranging $\quad x = a + \alpha \qquad y = b + \beta$

$$\rightarrow \quad \alpha = x - a \qquad \beta = y - b$$

substituting into the equation for z:

$$z = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$



Cross product

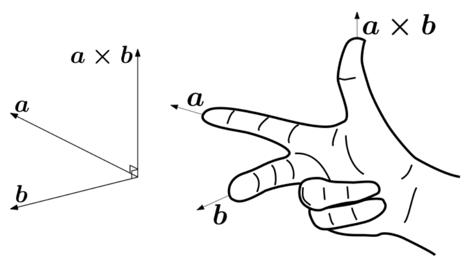
The cross product $\underset{\sim}{a}, \underset{\sim}{b} \in \mathbb{R}^3$, $\underset{\sim}{a} = (a_1, a_2, a_3)$ $\underset{\sim}{b} = (b_1, b_2, b_3)$

$$\underset{\sim}{a} \times \underset{\sim}{b} = (a_2 b_3 - a_3 b_2, \; a_3 b_1 - a_1 b_3, \; a_1 b_2 - a_2 b_1)$$

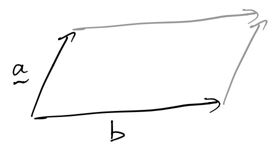
$$\begin{vmatrix} a_1 & a_2 & a_3 & a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 & b_1 & b_2 & b_3 \end{vmatrix}$$

it has some very useful properties:

- $\underset{\sim}{a} \times \underset{\sim}{b}$ is perpendicular to both $\underset{\sim}{a}$ and $\underset{\sim}{b}$ and oriented according to the right hand rule:



- the magnitude $\|\underset{\sim}{a} \times \underset{\sim}{b}\|$ is equal to the area of the parallelogram:



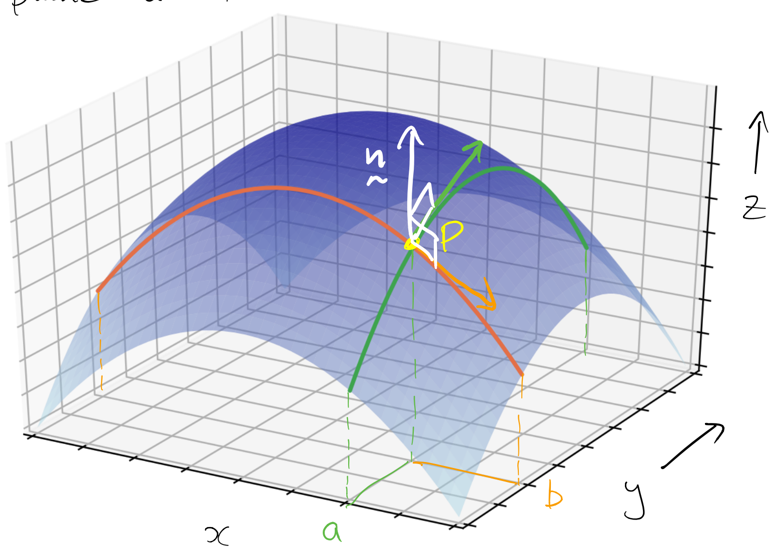
Example $\underset{\sim}{a} = (1, 3, 1)$ $\underset{\sim}{b} = (2, 1, 5)$

$$\begin{aligned} \underset{\sim}{a} \times \underset{\sim}{b} &= (15 - 1, 2 - 5, 1 - 6) \\ &= (14, -3, -5) \end{aligned}$$

$$\begin{matrix} 1 & 3 & 1 & 1 & 3 & 1 \\ 2 & 1 & 5 & 2 & 1 & 5 \end{matrix}$$

$$\begin{aligned} \underset{\sim}{a} \cdot (\underset{\sim}{a} \times \underset{\sim}{b}) &= (1, 3, 1) \cdot (14, -3, -5) \\ &= 14 - 9 - 5 \\ &= 0 \end{aligned}$$

A vector $\underset{\sim}{n}$ is called a normal vector to a given surface at the point P if it is perpendicular to every tangent vector at P, i.e. it is perpendicular (a.k.a. orthogonal) to the tangent plane at P.



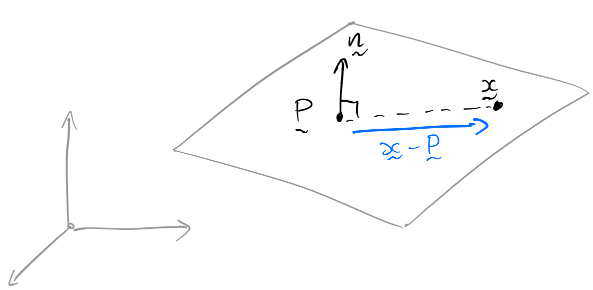
An easy way to find a normal vector is to take the cross product of the tangent vectors in the x and y-directions. Recall that at the point $\underset{\sim}{P} = (a, b, f(a,b))$ these vectors are

$$\underset{\sim}{u} = \left(1, 0, \frac{\partial f}{\partial x}(a,b)\right) \quad \text{and} \quad \underset{\sim}{v} = \left(0, 1, \frac{\partial f}{\partial y}(a,b)\right)$$

$$\underset{\sim}{n} = \underset{\sim}{u} \times \underset{\sim}{v} = (-f_x(a,b), -f_y(a,b), 1)$$

$$\begin{array}{cccccc} 1 & 0 & f_x & 1 & 0 & f_x \\ 0 & 1 & f_y & 0 & 1 & f_y \end{array}$$

A normal vector gives another way of finding an equation for the tangent plane at P:



$$\underset{\sim}{P} = (a, b, c)$$
$$\underset{\sim}{x} = (x, y, z)$$
$$\underset{\sim}{n} = (n_1, n_2, n_3)$$

$\underset{\sim}{x}$ is in the plane orthogonal to $\underset{\sim}{n}$ iff the vector from $\underset{\sim}{P}$ to $\underset{\sim}{x}$, i.e. $\underset{\sim}{x} - \underset{\sim}{P} = (x-a, y-b, z-c)$, is orthogonal to $\underset{\sim}{n}$

i.e.
$$\underset{\sim}{n} \cdot (\underset{\sim}{x} - \underset{\sim}{P}) = 0$$

$$n_1 (x-a) + n_2 (y-b) + n_3 (z-c) = 0$$

this is often written as

$$n_1 x + n_2 y + n_3 z = d$$

where $d = n_1 a + n_2 b + n_3 c$

If $\underset{\sim}{n}$ is the normal to a surface $= (-f_x(a,b), -f_y(a,b), 1)$

this equation is

$$-f_x(a,b)(x-a) - f_y(a,b)(y-b) + z - c = 0$$

Chain rule for partial derivatives

Recall if $f: \mathbb{R} \to \mathbb{R}$, $u: \mathbb{R} \to \mathbb{R}$

$f(x)$ $\quad$ $u(t)$

then we can take the composition $(f \circ u)(t) = f(u(t))$ and the chain rule gives:

$$\frac{df}{dt} = \frac{df}{du} \cdot \frac{du}{dt}$$

if we include arguments:

$$\frac{d}{dt} f(u(t)) = \frac{df}{du}(u(t)) \cdot \frac{du}{dt}(t)$$

Suppose $F: \mathbb{R}^2 \to \mathbb{R}$ and $u: \mathbb{R} \to \mathbb{R}$, $v: \mathbb{R} \to \mathbb{R}$

$F(x,y)$ $\quad$ $u(t)$ $\quad$ $v(t)$

then we can compose: $F(u(t), v(t))$

$\frac{dF}{dt}$ follows the chain rule for partial derivatives:

$$\frac{dF}{dt} = \frac{\partial F}{\partial u} \cdot \frac{du}{dt} + \frac{\partial F}{\partial v} \cdot \frac{dv}{dt}$$

<u>Example</u> $F(x,y) = x^2 + y^2$, $u(t) = t^2$, $v(t) = e^t$

find $\frac{d}{dt} F(u(t), v(t))$.

$\frac{du}{dt} = 2t$, $\frac{dv}{dt} = e^t$ $\qquad$ $F(u,v) = u^2 + v^2$ (suppressing t for now)

$$\Rightarrow \frac{\partial F}{\partial u} = 2u \qquad \frac{\partial F}{\partial v} = 2v$$

so by the chain rule

$$\frac{dF}{dt}(u,v) = 2u \cdot 2t + 2v \cdot e^t$$

expressed just in t: $= 2t^2 \cdot 2t + 2e^t \cdot e^t = 4t^3 + 2e^{2t}$

We can check this by substituting for t at the beginning and calculating $\frac{dF}{dt}$ directly:

$$F(u(t), v(t)) = u(t)^2 + v(t)^2 = (t^2)^2 + (e^t)^2$$
$$= t^4 + e^{2t}$$

$$\frac{dF}{dt} = 4t + 2e^{2t}$$

Suppose now we have a vector-valued function $\underset{\sim}{g} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\underset{\sim}{g}(s,t) = (g_1(s,t), g_2(s,t))$$

– two coordinate functions!

since g maps into $\mathbb{R}^2$, and F takes its arguments from $\mathbb{R}^2$, they can be composed: $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{F} \mathbb{R}$

$$F(\underset{\sim}{g}(s,t)) = F(g_1(s,t), g_2(s,t))$$

and we might need to find $\frac{\partial F}{\partial s}$ or $\frac{\partial F}{\partial t}$

$$\frac{\partial F}{\partial s} = \frac{\partial F}{\partial g_1} \cdot \frac{\partial g_1}{\partial s} + \frac{\partial F}{\partial g_2} \cdot \frac{\partial g_2}{\partial s}$$

$$\frac{\partial F}{\partial t} = \frac{\partial F}{\partial g_1} \cdot \frac{\partial g_1}{\partial s} + \frac{\partial F}{\partial g_2} \cdot \frac{\partial g_2}{\partial s}$$

In general if we have a multivariable function

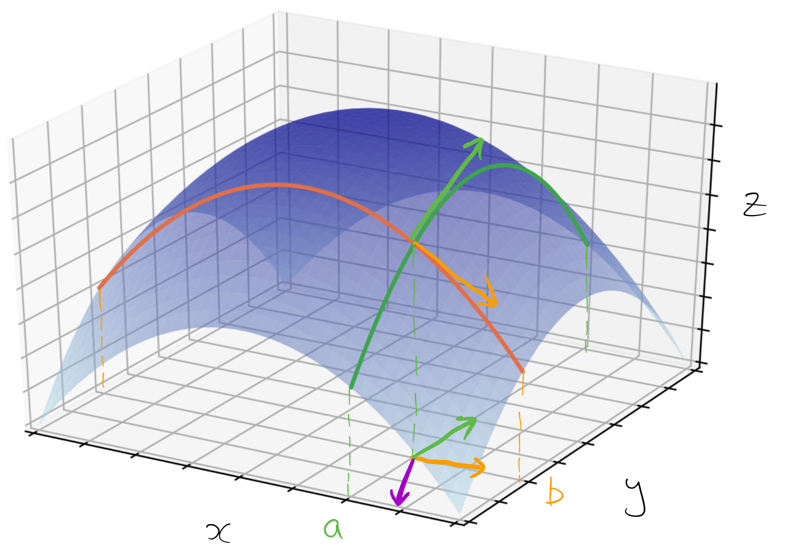
$$f(u_1, u_2, \ldots, u_n)$$

where each u_i is itself a multivariable function

$$u_i(t_1, t_2, \ldots, t_n)$$

then
$$\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial u_1} \frac{\partial u_1}{\partial t_i} + \frac{\partial f}{\partial u_2} \frac{\partial u_2}{\partial t_i} + \ldots + \frac{\partial f}{\partial u_n} \cdot \frac{\partial u_n}{\partial t_i}$$

Directional derivatives and differentiability



$z = f(x, y)$

$\frac{\partial f}{\partial x}$ rate of change of f (height) in x direction: $(1, 0)$

$\left(1, 0, \frac{\partial f}{\partial x}\right)$ tangent vector to the surface in x direction

$\frac{\partial f}{\partial y}$ rate of change of f (height) in y direction: $(0, 1)$

$\left(1, 0, \frac{\partial f}{\partial y}\right)$ tangent vector to the surface in x direction

Can we find the rate of change of f in some other direction?

(how steep is the ascent / descent in the direction $\underset{\sim}{v}$

The directional derivative of f at $\underset{\sim}{c} = (a, b)$ in the direction $\underset{\sim}{v}$ is defined by

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \lim_{h \to 0} \frac{f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) - f(\underset{\sim}{c})}{h} \qquad \text{where } \hat{\underset{\sim}{v}} = \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

if this limit exists then $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ can be expressed in terms of the partial derivatives of f:

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \hat{\underset{\sim}{v}} \qquad \text{(proof below)}$$

Defining the gradient vector

$$\nabla f(a,b) = \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b) \right)$$

the directional derivative is therefore given by

$$D_{\underset{\sim}{v}} f(a,b) = \nabla f(a,b) \cdot \hat{\underset{\sim}{v}} = \nabla f(a,b) \cdot \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

The same applies to directional derivatives of functions of more variables, eg: $\underset{\sim}{v} \in \mathbb{R}^3$, $f: \mathbb{R}^3 \to \mathbb{R}$

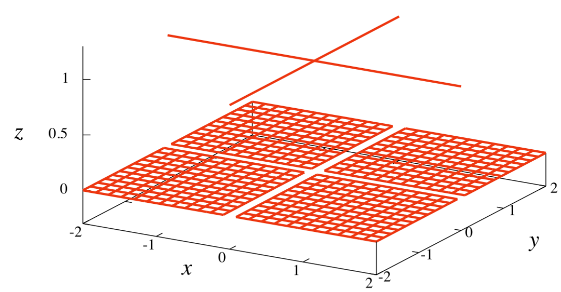
$$D_{\underset{\sim}{v}} f(a,b,c) = \nabla f(a,b,c) \cdot \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

where $$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

A function $f: \mathbb{R}^2 \to \mathbb{R}$ is called differentiable at (a,b) if the directional derivative exists in every direction.

Note that the existence of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at (a,b) doesn't guarantee that f is differentiable at (a,b),

eg: $$f(x,y) = \begin{cases} 0 & xy \neq 0 \\ 1 & xy = 0 \end{cases}$$



$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \frac{\partial f}{\partial y}(0,0) = 0$$

but no other directional derivatives exist. If we form a tangent plane using $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ at $(0,0)$ it won't be a good approximation!

It turns out that for differentiability at (a,b) we require $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ to be continuous at (a,b) (and therefore they must exist for points around (a,b)).

Proof of the formula: $D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \hat{\underset{\sim}{v}}$

First let $g(t) = f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}})$, then from the definition of $\frac{dg}{dt}$:

$$\frac{dg}{dt}(0) = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h}$$

$$= \lim_{h \to 0} \frac{f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) - f(\underset{\sim}{c})}{h} \quad \leftarrow \text{since } \begin{array}{l} g(h) = f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) \\ g(0) = f(\underset{\sim}{c}) \end{array}$$

$$= D_{\underset{\sim}{v}} f(\underset{\sim}{c}) \qquad \text{(by definition)}$$

this proves

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left. \frac{d}{dt} \right|_{t=0} f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}})$$

to which we will apply the chain rule:

let $\underset{\sim}{c} + t\hat{\underset{\sim}{v}} = \begin{pmatrix} a \\ b \end{pmatrix} + t \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \begin{pmatrix} a + t\hat{v}_1 \\ b + t\hat{v}_2 \end{pmatrix} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$

$$\frac{d}{dt} f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}}) = \frac{d}{dt} f(a + t\hat{v}_1, b + t\hat{v}_2)$$

$$= \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}$$

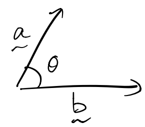
$$= \frac{\partial f}{\partial x} \cdot \hat{v}_1 + \frac{\partial f}{\partial y} \cdot \hat{v}_2$$

$$= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \hat{\underset{\sim}{v}}$$

Maximum rate of change

Directional derivatives give us a way of calculating the rate of change in any direction, so we might ask which direction gives the maximum or minimum rate of change. (at a given point on a mountain, what is the direction of steepest ascent / descent ?)

Recall: $\underset{\sim}{a} \cdot \underset{\sim}{b} = |\underset{\sim}{a}|\,|\underset{\sim}{b}| \cos\theta$ where



Suppose $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable, then

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \nabla f(\underset{\sim}{c}) \cdot \hat{\underset{\sim}{v}} = |\nabla f(\underset{\sim}{c})|\,|\hat{\underset{\sim}{v}}| \cos\theta$$

but $\hat{\underset{\sim}{v}} = \dfrac{\underset{\sim}{v}}{|\underset{\sim}{v}|}$ so $|\hat{\underset{\sim}{v}}| = 1$, $|\nabla f(\underset{\sim}{c})|$ doesn't depend on $\underset{\sim}{v}$,

"unit vector"

so the maximum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ (for $\underset{\sim}{c}$ fixed) occurs at the maximum of $\cos\theta$, which is 1 at $\theta = 0$.

$\theta = 0$ means $\nabla f(\underset{\sim}{c})$ and $\hat{\underset{\sim}{v}}$ are in the same direction, so $\hat{\underset{\sim}{v}}$ is a unit vector in the $\nabla f(\underset{\sim}{c})$ direction, i.e. $\hat{\underset{\sim}{v}} = \dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Therefore:

The maximum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ occurs when $\hat{\underset{\sim}{v}} = \dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Similarly, the minimum occurs at min $\cos\theta$, i.e. $\theta = -\pi$ which means $\hat{\underset{\sim}{v}}$ (and therefore $\underset{\sim}{v}$) is in the opposite direction to $\nabla f(\underset{\sim}{c})$

The minimum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ occurs when $\hat{\underset{\sim}{v}} = -\dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Example 3.49. *The temperature at each point of a metal plate is given by the function $T(x,y) = e^x \cos y + e^y \cos x$. In what direction does the temperature increase most rapidly at the point $(0,0)$. What is this rate of increase?*

The direction of maximum rate of change of temperature is the direction of $\nabla T(0,0)$

$$\nabla T(x,y) = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}\right) = \left(e^x \cos y - e^y \sin x, -e^x \sin y + e^y \cos x\right)$$

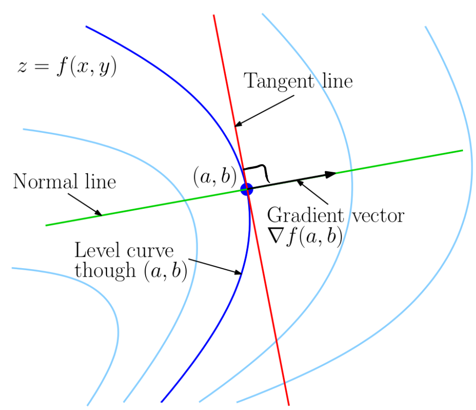
So

$$\nabla T(0,0) = (1,1)$$

The rate of change of T at $(0,0)$ in this direction is

$$D_{\nabla T(0,0)} T(0,0) = \nabla T(0,0) \cdot \frac{\nabla T(0,0)}{|\nabla T(0,0)|} = |\nabla T(0,0)| = |(1,1)|$$
$$= \sqrt{2}$$

Contrast with the level curves, $f(x,y)=k$, curves of equal height. So a tangent vector to a level curve gives a direction of no change in height. These directions turn out to be perpendicular to the directions of max/min rate of change! (steepest ascent/descent)



A similar result holds for level surfaces $f(x,y,z) = k$ $(f: \mathbb{R}^3 \to \mathbb{R})$

The gradient vector $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$ at a point (a,b,c) on the level surface $f(x,y,z) = k$ is perpendicular to every tangent vector to the level surface at (a,b,c), i.e. perpendicular to the tangent plane, i.e. a normal vector to the surface!

The Jacobian matrix

$f : \mathbb{R}^2 \to \mathbb{R}$ has a gradient $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$

which gives rates of change of f in different directions

If $\underset{\sim}{h} : \mathbb{R}^2 \to \mathbb{R}^2$, then it has two coordinate functions

$$\underset{\sim}{h}(x,y) = (h_1(x,y), h_2(x,y))$$

$\underset{\sim}{h}$ is a vector quantity changes in the value of $\underset{\sim}{h}$ can happen in two directions, so the change in $\underset{\sim}{h}$ in the x-direction (or y direction etc.) is a vector quantity made up of the change in h_1 and the change in h_2.

The Jacobian matrix, also called the (total) derivative of $\underset{\sim}{h}$, is

$$D\underset{\sim}{h} = \begin{bmatrix} \partial_x h_1 & \partial_y h_1 \\ \partial_x h_2 & \partial_y h_2 \end{bmatrix}$$

There is a lot more to be said about this derivative (extension to functions $\mathbb{R}^2 \to \mathbb{R}^3$, $\mathbb{R}^m \to \mathbb{R}^n$, matrix chain rule ...) but for this course the important thing is the Jacobian which is the determinant of the Jacobian matrix.

Example

$\underset{\sim}{h}(x,y) = (x^2, x+y)$ find the Jacobian of h

$$\frac{\partial h_1}{\partial x} = 2x \qquad \frac{\partial h_1}{\partial y} = 0$$

$$\frac{\partial h_2}{\partial x} = 1 \qquad \frac{\partial h_2}{\partial y} = 1$$

$$D\underset{\sim}{h} = \begin{bmatrix} 2x & 0 \\ 1 & 1 \end{bmatrix}$$

$$\det(D\underset{\sim}{h}) = 2x - 2y$$

Chain rule for partial derivatives

Recall if $f : \mathbb{R} \to \mathbb{R}$, $u : \mathbb{R} \to \mathbb{R}$
$f(x)$ $u(t)$

then we can take the composition $(f \circ u)(t) = f(u(t))$ and the chain rule gives:

$$\frac{df}{dt} = \frac{df}{du} \cdot \frac{du}{dt}$$

if we include arguments:

$$\frac{d}{dt} f(u(t)) = \frac{df}{du}(u(t)) \cdot \frac{du}{dt}(t)$$

Suppose $F : \mathbb{R}^2 \to \mathbb{R}$ and $u : \mathbb{R} \to \mathbb{R}$, $v : \mathbb{R} \to \mathbb{R}$
$F(x,y)$ $u(t)$ $v(t)$

then we can compose: $F(u(t), v(t))$

$\frac{dF}{dt}$ follows the chain rule for partial derivatives:

$$\frac{dF}{dt} = \frac{\partial F}{\partial u} \cdot \frac{du}{dt} + \frac{\partial F}{\partial v} \cdot \frac{dv}{dt}$$

<u>Example</u> $F(x,y) = x^2 + y^2$, $u(t) = t^2$, $v(t) = e^t$

find $\frac{d}{dt} F(u(t), v(t))$.

$\frac{du}{dt} = 2t$, $\frac{dv}{dt} = e^t$ $\qquad F(u,v) = u^2 + v^2$ (suppressing t for now)

$$\Rightarrow \frac{\partial F}{\partial u} = 2u \qquad \frac{\partial F}{\partial v} = 2v$$

so by the chain rule

$$\frac{dF}{dt}(u,v) = 2u \cdot 2t + 2v \cdot e^t$$

expressed just in t: $\quad = 2t^2 \cdot 2t + 2e^t \cdot e^t = 4t^3 + 2e^{2t}$

We can check this by substituting for t at the beginning and calculating $\frac{dF}{dt}$ directly:

$$F(u(t), v(t)) = u(t)^2 + v(t)^2 = (t^2)^2 + (e^t)^2$$
$$= t^4 + e^{2t}$$
$$\frac{dF}{dt} = 4t + 2e^{2t}$$

Suppose now we have a vector-valued function $\underset{\sim}{g} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\underset{\sim}{g}(s,t) = (g_1(s,t), g_2(s,t))$$

– two coordinate functions!

since g maps into $\mathbb{R}^2$, and F takes its arguments from $\mathbb{R}^2$, they can be composed: $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{F} \mathbb{R}$

$$F(\underset{\sim}{g}(s,t)) = F(g_1(s,t), g_2(s,t))$$

and we might need to find $\frac{\partial F}{\partial s}$ or $\frac{\partial F}{\partial t}$

$$\frac{\partial F}{\partial s} = \frac{\partial F}{\partial g_1} \cdot \frac{\partial g_1}{\partial s} + \frac{\partial F}{\partial g_2} \cdot \frac{\partial g_2}{\partial s}$$

$$\frac{\partial F}{\partial t} = \frac{\partial F}{\partial g_1} \cdot \frac{\partial g_1}{\partial s} + \frac{\partial F}{\partial g_2} \cdot \frac{\partial g_2}{\partial s}$$

In general if we have a multivariable function

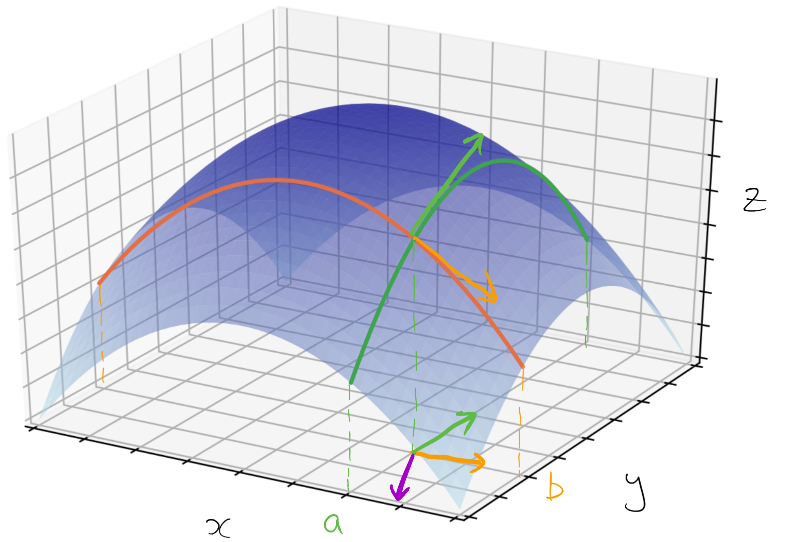
$$f(u_1, u_2, \ldots, u_n)$$

where each u_i is itself a multivariable function

$$u_i(t_1, t_2, \ldots, t_n)$$

then
$$\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial u_1} \frac{\partial u_1}{\partial t_i} + \frac{\partial f}{\partial u_2} \frac{\partial u_2}{\partial t_i} + \ldots + \frac{\partial f}{\partial u_n} \cdot \frac{\partial u_n}{\partial t_i}$$

Directional derivatives and differentiability



$z = f(x,y)$

$\dfrac{\partial f}{\partial x}$ rate of change of f (height) in x direction: $(1,0)$

$\left(1, 0, \dfrac{\partial f}{\partial x}\right)$ tangent vector to the surface in x direction

$\dfrac{\partial f}{\partial y}$ rate of change of f (height) in y direction: $(0,1)$

$\left(1, 0, \dfrac{\partial f}{\partial y}\right)$ tangent vector to the surface in x direction

Can we find the rate of change of f in some other direction?

(how steep is the ascent / descent in the direction $\underset{\sim}{v}$

The directional derivative of f at $\underset{\sim}{c} = (a,b)$ in the direction $\underset{\sim}{v}$ is defined by

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \lim_{h \to 0} \frac{f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) - f(\underset{\sim}{c})}{h} \qquad \text{where } \hat{\underset{\sim}{v}} = \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

if this limit exists then $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ can be expressed in terms of the partial derivatives of f:

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \hat{\underset{\sim}{v}} \qquad \text{(proof below)}$$

Defining the gradient vector

$$\nabla f(a,b) = \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b) \right)$$

the directional derivative is therefore given by

$$D_{\underset{\sim}{v}} f(a,b) = \nabla f(a,b) \cdot \hat{\underset{\sim}{v}} = \nabla f(a,b) \cdot \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

The same applies to directional derivatives of functions of more variables, eg: $\underset{\sim}{v} \in \mathbb{R}^3$, $f: \mathbb{R}^3 \to \mathbb{R}$

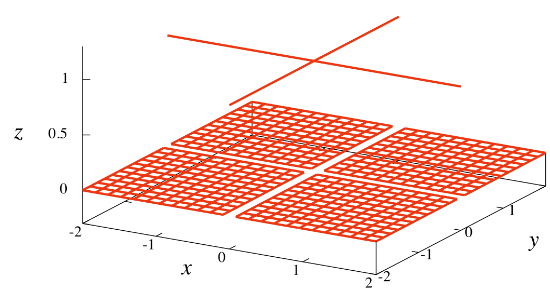
$$D_{\underset{\sim}{v}} f(a,b,c) = \nabla f(a,b,c) \cdot \frac{\underset{\sim}{v}}{\|\underset{\sim}{v}\|}$$

where $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$

A function $f: \mathbb{R}^2 \to \mathbb{R}$ is called differentiable at (a,b) if the directional derivative exists in every direction.

Note that the existence of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at (a,b) doesn't guarantee that f is differentiable at (a,b),

eg: $f(x,y) = \begin{cases} 0 & xy \neq 0 \\ 1 & xy = 0 \end{cases}$



$\frac{\partial f}{\partial x}(0,0) = 0$ $\quad\quad$ $\frac{\partial f}{\partial y}(0,0) = 0$

but no other directional derivatives exist. If we form a tangent plane using $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ at $(0,0)$ it won't be a good approximation!

It turns out that for differentiability at (a,b) we require $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ to be continuous at (a,b) (and therefore they must exist for points around (a,b)).

Proof of the formula: $D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \hat{\underset{\sim}{v}}$

First let $g(t) = f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}})$, then from the definition of $\frac{dg}{dt}$:

$$\frac{dg}{dt}(0) = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h}$$

$$= \lim_{h \to 0} \frac{f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) - f(\underset{\sim}{c})}{h} \quad \leftarrow \text{since } \begin{array}{l} g(h) = f(\underset{\sim}{c} + h\hat{\underset{\sim}{v}}) \\ g(0) = f(\underset{\sim}{c}) \end{array}$$

$$= D_{\underset{\sim}{v}} f(\underset{\sim}{c}) \qquad \text{(by definition)}$$

this proves

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \left. \frac{d}{dt} \right|_{t=0} f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}})$$

to which we will apply the chain rule:

let $\underset{\sim}{c} + t\hat{\underset{\sim}{v}} = \begin{pmatrix} a \\ b \end{pmatrix} + t \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \begin{pmatrix} a + t\hat{v}_1 \\ b + t\hat{v}_2 \end{pmatrix} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$

$$\frac{d}{dt} f(\underset{\sim}{c} + t\hat{\underset{\sim}{v}}) = \frac{d}{dt} f(a + t\hat{v}_1, b + t\hat{v}_2)$$

$$= \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}$$

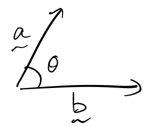
$$= \frac{\partial f}{\partial x} \cdot \hat{v}_1 + \frac{\partial f}{\partial y} \cdot \hat{v}_2$$

$$= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \hat{\underset{\sim}{v}}$$

Maximum rate of change

Directional derivatives give us a way of calculating the rate of change in <u>any</u> direction, so we might ask which direction gives the maximum or minimum rate of change. (at a given point on a mountain, what is the direction of steepest ascent / descent ?)

Recall: $\underset{\sim}{a} \cdot \underset{\sim}{b} = |\underset{\sim}{a}|\,|\underset{\sim}{b}| \cos\theta$ where



Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable, then

$$D_{\underset{\sim}{v}} f(\underset{\sim}{c}) = \nabla f(\underset{\sim}{c}) \cdot \hat{\underset{\sim}{v}} = |\nabla f(\underset{\sim}{c})|\,|\hat{\underset{\sim}{v}}| \cos\theta$$

but $\hat{\underset{\sim}{v}} = \dfrac{\underset{\sim}{v}}{|\underset{\sim}{v}|}$ so $|\hat{\underset{\sim}{v}}| = 1$, $|\nabla f(\underset{\sim}{c})|$ doesn't depend on $\underset{\sim}{v}$,

"unit vector"

so the maximum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ (for $\underset{\sim}{c}$ fixed) occurs at the maximum of $\cos\theta$, which is 1 at $\theta = 0$.

$\theta = 0$ means $\nabla f(\underset{\sim}{c})$ and $\hat{\underset{\sim}{v}}$ are in the same direction, so $\hat{\underset{\sim}{v}}$ is a unit vector in the $\nabla f(\underset{\sim}{c})$ direction, i.e. $\hat{\underset{\sim}{v}} = \dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Therefore:

The maximum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ occurs when $\hat{\underset{\sim}{v}} = \dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Similarly, the minimum occurs at min $\cos\theta$, i.e. $\theta = -\pi$ which means $\hat{\underset{\sim}{v}}$ (and therefore $\underset{\sim}{v}$) is in the opposite direction to $\nabla f(\underset{\sim}{c})$

The minimum of $D_{\underset{\sim}{v}} f(\underset{\sim}{c})$ occurs when $\hat{\underset{\sim}{v}} = -\dfrac{\nabla f(\underset{\sim}{c})}{|\nabla f(\underset{\sim}{c})|}$

Example 3.49. *The temperature at each point of a metal plate is given by the function $T(x, y) = e^x \cos y + e^y \cos x$. In what direction does the temperature increase most rapidly at the point $(0, 0)$. What is this rate of increase?*

The direction of maximum rate of change of temperature is the direction of $\nabla T(0,0)$

$$\nabla T(x,y) = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}\right) = \left(e^x \cos y - e^y \sin x, -e^x \sin y + e^y \cos x\right)$$

So

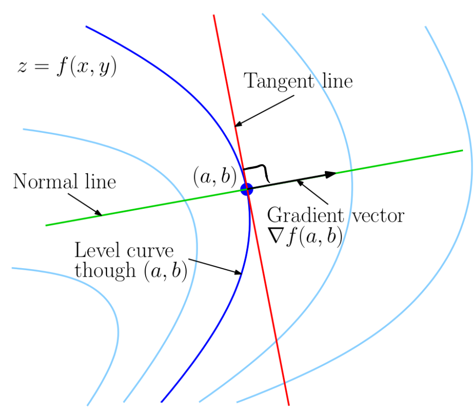
$$\nabla T(0,0) = (1,1)$$

The rate of change of T at $(0,0)$ in this direction is

$$D_{\nabla T(0,0)} T(0,0) = \nabla T(0,0) \cdot \frac{\nabla T(0,0)}{|\nabla T(0,0)|} = |\nabla T(0,0)| = |(1,1)|$$
$$= \sqrt{2}$$

Contrast with the level curves, $f(x,y)=k$, curves of equal height

So a tangent vector to a level curve gives a direction of <u>no</u> change in height. These directions turn out to be perpendicular to the directions of max/min rate of change! (steepest ascent/descent)



A similar result holds for level <u>surfaces</u> $f(x,y,z) = k$ $(f: \mathbb{R}^3 \to \mathbb{R})$

The gradient vector $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$ at a point (a,b,c) on the level surface $f(x,y,z) = k$ is perpendicular to every tangent vector to the level surface at (a,b,c), i.e. perpendicular to the tangent plane, i.e. a <u>normal</u> vector to the surface!

The Jacobian matrix

$f : \mathbb{R}^2 \to \mathbb{R}$ has a gradient $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$

which gives rates of change of f in different directions

If $\underset{\sim}{h} : \mathbb{R}^2 \to \mathbb{R}^2$, then it has two coordinate functions

$$\underset{\sim}{h}(x,y) = \left(h_1(x,y), h_2(x,y) \right)$$

$\underset{\sim}{h}$ is a vector quantity changes in the value of $\underset{\sim}{h}$ can happen in two directions, so the change in $\underset{\sim}{h}$ in the x-direction (or y direction etc.) is a vector quantity made up of the change in h_1 and the change in h_2.

The Jacobian matrix, also called the (total) derivative of $\underset{\sim}{h}$, is

$$D\underset{\sim}{h} = \begin{bmatrix} \partial_x h_1 & \partial_y h_1 \\ \partial_x h_2 & \partial_y h_2 \end{bmatrix}$$

There is a lot more to be said about this derivative (extension to functions $\mathbb{R}^2 \to \mathbb{R}^3$, $\mathbb{R}^m \to \mathbb{R}^n$, matrix chain rule ...) but for this course the important thing is the Jacobian which is the determinant of the Jacobian matrix.

Example

$\underset{\sim}{h}(x,y) = (x^2, x+y)$ find the Jacobian of h

$$\frac{\partial h_1}{\partial x} = 2x \qquad \frac{\partial h_1}{\partial y} = 0$$

$$\frac{\partial h_2}{\partial x} = 1 \qquad \frac{\partial h_2}{\partial y} = 1$$

$$D\underset{\sim}{h} = \begin{bmatrix} 2x & 0 \\ 1 & 1 \end{bmatrix}$$

$$\det(D\underset{\sim}{h}) = 2x - 2y$$