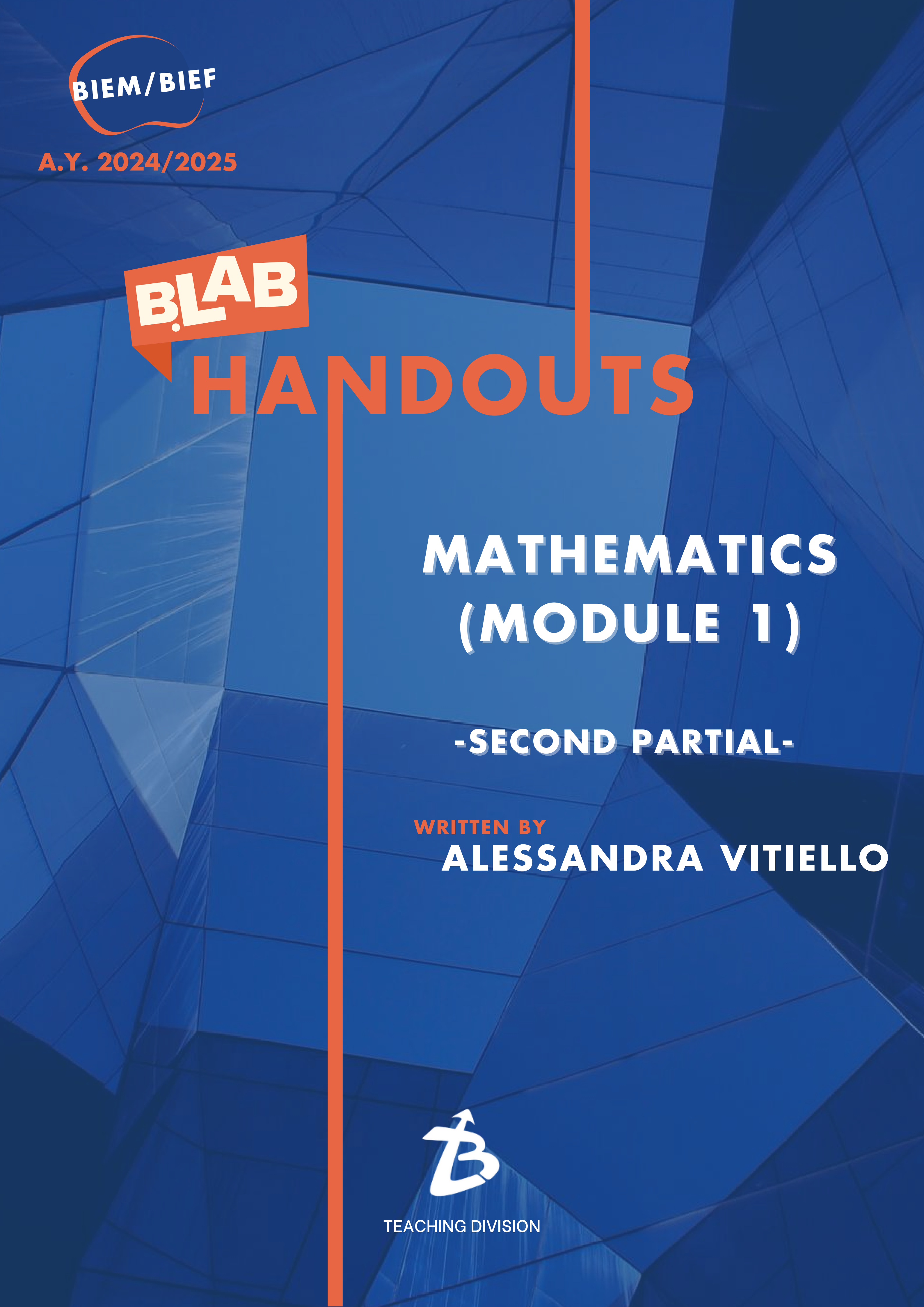


BIEM/BIEF

A.Y. 2024/2025

B.LAB

HANDOUTS

MATHEMATICS (MODULE 1)

-SECOND PARTIAL-

WRITTEN BY
ALESSANDRA VITIELLO

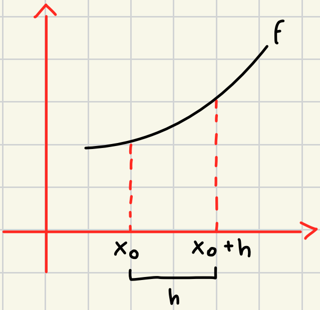
TEACHING DIVISION

VARIABLE DIFFERENCIAL CALCULS

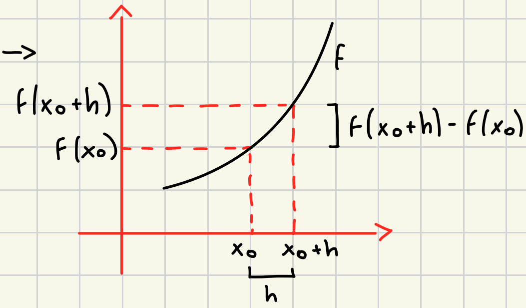
DIFFERENT QUOTIENT

given $f: A \subseteq R \to R$
x_0 interior point of A

INCREMENT / VARIATION OF X
h = increment with sign →
- $h > 0 \Rightarrow x_0 + h > x_0$
- $h < 0 \Rightarrow x_0 + h < x_0$



INCREMENT / VARIATION OF f
$f(x_0+h) - f(x_0)$ = increment with sign →
- $> 0 \Rightarrow f(x_0+h) > f(x_0)$
- $< 0 \Rightarrow f(x_0+h) < f(x_0)$



DIFFERENCE QUOTIENT OF f AT x_0

$$\frac{f(x_0+h) - f(x_0)}{h} = \frac{\Delta f}{h} = \frac{\Delta f}{\Delta x}$$

DERIVATIVE

consider $f: A \subseteq R \to R$
x_0 interior point for A
increment of $x = h$
increment of $f = f(x_0+h) - f(x_0)$
different quotient of f at $x_0 = \frac{f(x_0+h)-f(x_0)}{h}$

if the limit: $\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$

∃ finite we say f has a derivative at x_0
we call derivative of f at x_0 the value → $f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$

LEIBNIZ'S NOTION (alternative notion)

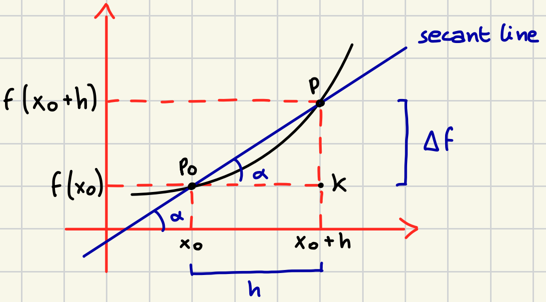
$\frac{df}{dx}(x_0)$, $\frac{dy}{dx}(x_0)$

GEOMETRICAL MEANING OF $\frac{\Delta f}{h}$

$$\frac{\Delta f}{h} = \frac{\overline{PK}}{\overline{P_0K}} = \tan\alpha = m \text{ secant}$$

$\Downarrow$

slope of the secant line



GEOMETRICAL MEANING OF $f'(x_0)$

$h \to 0 \Rightarrow x_0 + h \to x_0$

$p \to P_0$

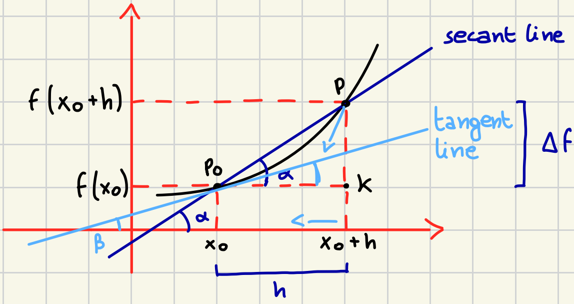
secant line → tangent line at P_0

$$f'(x_0) = \lim_{h \to 0} \frac{\Delta f}{h} = \tan\beta = m \text{ tangent}$$

$\Downarrow$

slope of the tangent line

↳ when the slope ∃ (when the tangent line is not vertical)



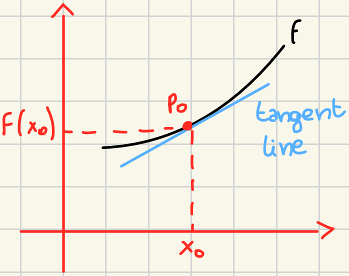
PROCEDURE TO WRITE THE TANGENT LINE TO f AT x_0

(when it's not vertical)

① tangent line to f at x_0 = streight line passing through $P_0(x_0, f(x_0))$ → $y - f(x_0) = m(x - x_0)$

② $m = f'(x_0)$ → $\underbrace{y - f(x_0) = f'(x_0)(x - x_0)}$

equation of the tangent line

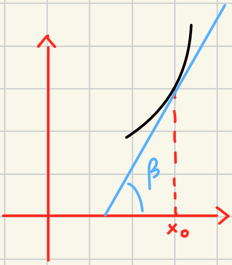


WHAT IF:

- f "quickly increasing" →

$f'(x_0) = \tan\beta$

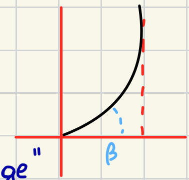
$0 < \beta < \frac{\pi}{2} \Rightarrow$ "close" to $\frac{\pi}{2}$



→ $y = \tan x$:

$\Downarrow$

$f'(x_0) > 0$ and "large"

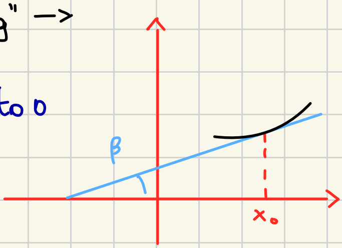


≈ 3

- f "slowly increasing" →

$f'(x_0) = \tan\beta$

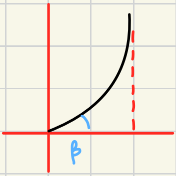
$0 < \beta < \frac{\pi}{2}$ ⇒ "close" to 0



→ $y = \tan x$:

⇓

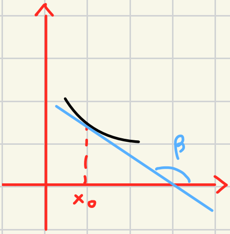
$f'(x_0) > 0$ and "small" $\approx \frac{1}{3}$



- f "slowly decreasing" →

$f'(x_0) = \tan\beta$

$\frac{\pi}{2} < \beta < \pi$ ⇒ "close" to π



→ $y = \tan x$:

⇓

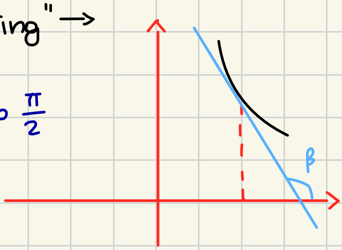
$f'(x_0) < 0$ and "small"

$\approx -\frac{1}{3}$

- f "quickly decreasing" →

$f'(x_0) = \tan\beta$

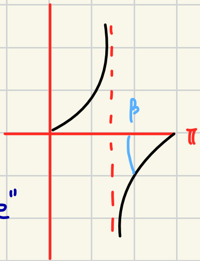
$\frac{\pi}{2} < \beta < \pi$ ⇒ "close" to $\frac{\pi}{2}$



→ $y = \tan x$:

⇓

$f'(x_0) < 0$ and "large"

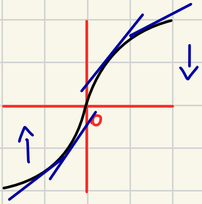


≈ -3

BUT WHAT IF :

① $\lim_{h\to 0} \frac{\Delta f}{h}$ ∃ but is not finite → EXAMPLE

$y = \sqrt[3]{x}$ ⇒



x=0 is inflection point with vertical tangent

at x=0 → $\lim_{h\to 0} \frac{f(0+h)-f(0)}{h} =$

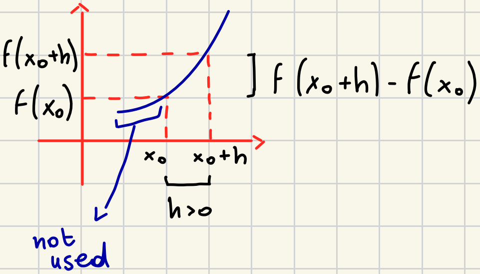
$\lim_{h\to 0} \frac{\cancel{\sqrt[3]{h}} - 0}{\underbrace{h}_{\cancel{\sqrt[3]{h}}\,\sqrt[3]{h^2}}} = \lim_{h\to 0} \frac{1}{\sqrt[3]{h^2}} = +\infty$ ⇒ $\nexists\, f'(0)$

② $\lim_{h \to 0} \frac{\Delta f}{h} \nexists \implies$

- RIGHT DERIVATIVE = $f'_+(x_0) = \lim_{h \to 0^+} \frac{f(x_0+h)-f(x_0)}{h}$ if the limit $\exists$ finite

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

  x_0 such that $\exists$ a right neighbourhood $B^+(x_0) \subseteq A$

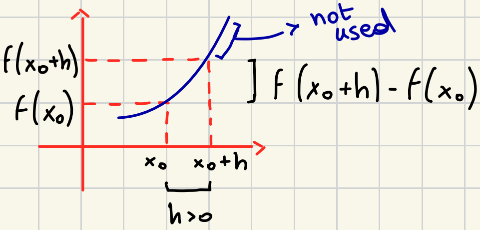


- right increment of $x = h \ (\geq 0)$
- increment of $f = f(x_0+h) - f(x_0)$ $(\gtreqless 0)$
- difference quotient $= \frac{f(x_0+h)-f(x_0)}{h}$

- LEFT DERIVATIVE = $f'_-(x_0) = \lim_{h \to 0^-} \frac{f(x_0+h)-f(x_0)}{h}$ if the limit $\exists$ finite

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

  x_0 such that $\exists$ a left neighbourhood $B^-(x_0) \subseteq A$



- left increment of $x = h \ (\geq 0)$
- increment of $f = f(x_0+h) - f(x_0)$ $(\gtreqless 0)$
- different quotient $= \frac{f(x_0+h)-f(x_0)}{h}$

THEOREM

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point

$f'(x_0) \ \exists \iff$ ① $f'_+(x_0) \ \exists$

② $f'_-(x_0) \ \exists$

③ they are equal

EXAMPE of ②

$$y = |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases} \implies f'_+(0) = \lim_{h \to 0^+} \frac{f(0+h)-f(0)}{h} = \lim_{h \to 0^+} \frac{\overset{\text{diff.}}{h-0}}{h} = 1$$

$$f'_-(0) = \lim_{h \to 0^-} \frac{f(0+h)-f(0)}{h} = \lim_{h \to 0^-} \frac{-h-0}{h} = -1$$

$\implies f'(0) \nexists$

↳ $\exists f'_+(x_0)$, $\exists f'_-(x_0)$ but they are $\neq \implies x_0 =$ corner

$$y = \sqrt{|x|} = \begin{cases} \sqrt{x} & x \geq 0 \\ \sqrt{-x} & x < 0 \end{cases} \Rightarrow$$

$$\lim_{h \to 0^+} \frac{f(0+h)-f(0)}{h} = \lim_{h \to 0^+} \frac{\cancel{\sqrt{h}} - 0}{\underbrace{h}_{\sqrt{h}\cdot\sqrt{h} = (\sqrt{h})^2}} = \lim_{h \to 0^+} \frac{1}{\sqrt{h}} = +\infty \Rightarrow f'_+(0) \nexists$$

$$\lim_{h \to 0^-} \frac{f(0+h)-f(0)}{h} = \lim_{h \to 0^-} \frac{\cancel{\sqrt{-h}} - 0}{\underbrace{h}_{-\sqrt{-h}\cdot -\sqrt{-h} = -(\sqrt{-h})^2 = -(-h) = h}} = \lim_{h \to 0^-} \frac{1}{-\sqrt{-h}} = -\infty \Rightarrow f'_-(0) \nexists$$

$\hookrightarrow$ $\exists \lim_{h \to 0^+} \frac{f(x_0+h)-f(x_0)}{h} = +\infty$,

$\exists \lim_{h \to 0^-} \frac{f(x_0+h)-f(x_0)}{h} = -\infty$, or vice versa

$\Rightarrow x_0 =$ caspial point

COMPUTATION OF DERIVATES

= given $y = f(x)$, given the general interior point x, find $f'(x)$ $\Rightarrow$ the derivative function, a function which at all points x will give the derivative of $f(x)$; if we want $f'(x_0)$, we just substitute $x = x_0$ in $f'(x_0)$

ELEMENTARY DERIVATES

① $y = k \Rightarrow y' = 0$ ($\forall$ constant $k \in \mathbb{R}$)

proof: $f'(x) = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h \to 0} \frac{k-k}{h} = \lim_{h \to 0} \frac{0}{k} = 0$

② $y = x^\alpha \Rightarrow y' = \alpha x^{\alpha-1}$

proof: $f'(x) = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h \to 0} \frac{(x+h)^\alpha - x^\alpha}{h} =$

$$\lim_{h \to 0} \frac{\left[x\left(1+\frac{h}{x}\right)\right]^\alpha - x^\alpha}{h} = \lim_{h \to 0} \frac{x^\alpha \left(1+\frac{h}{x}\right)^\alpha - x^\alpha}{h} = \lim_{h \to 0} \boxed{x^\alpha} \frac{\left(1+\frac{h}{x}\right)^\alpha - 1}{\boxed{h}} =$$

$$\frac{x^\alpha}{h} = \frac{x^{\alpha-1} \cdot h}{h} = x^{\alpha-1} \frac{1}{\frac{h}{x}}] = t$$

$$\lim_{t \to 0} x^{\alpha-1} \underbrace{\boxed{\frac{(1+t)-1}{t}}}_{\alpha} = \alpha x^{\alpha-1}$$

③ $y = e^x \Rightarrow y' = e^x \;/\; y = a^x \Rightarrow a^x \ln a$

④ $y = \ln x \Rightarrow y' = \frac{1}{x} \;/\; y = \log_a x$

⑤ $y = \sin x \Rightarrow y' = \cos x$

⑥ $y = \cos x \Rightarrow y' = -\sin x$

⑦ $y = \tan x \Rightarrow y = \frac{1}{\cos^2 x} \;/\; \tan^2 x + 1$

⑧ $y = \tan^{-1} x \Rightarrow y' =$

RULES ON DERIVATIVES

① DERIVATIVE OF A LINEAR COMBINATION

suppose f, g have a derivative $\forall x \in (a,b)$

suppose $\alpha, \beta \in \mathbb{R}$

$\alpha f + \beta g$ has a derivative $\forall x \in (a,b)$ and:

$(\alpha f + \beta g)' = \alpha f' + \beta g'$

PROOF:

$\forall x \in (a,b)$,

$$(\alpha f + \beta g)'(x) = \lim_{h \to 0} \frac{(\alpha f + \beta g)(x+h) - (\alpha f + \beta g)(x)}{h} =$$

$$\lim_{h \to 0} \frac{(\alpha f)(x+h) + (\beta g)(x+h) - (\alpha f)(x) + (\beta g)(x)}{h} =$$

$$\lim_{h \to 0} \left[\alpha \frac{f(x+h) - f(x)}{h} + \beta \frac{g(x+h) - g(x)}{h} \right] =$$

$$\alpha \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + \beta \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} = \alpha f'(x) + \beta g'(x)$$

② DERIVATIVE OF A PRODUCT

suppose f, g have a derivative $\forall x \in (a,b)$

then fg has a derivative $\forall x \in (a,b)$

$(fg)' = f'g + fg'$

③ DERIVATIVE OF A QUOTIENT

suppose f, g have a derivative $\forall x \in (a,b)$
suppose $g(x) \neq 0 \ \forall x \in (a,b)$
then f/g has a derivative $\forall x \in (a,b)$ and:

$$\left(\frac{f}{g}\right)' = \frac{gf' + fg'}{g^2}$$

④ CHAIN RULE (DERIVATIVE OF A COMPOSITION)

suppose $f: A = (a,b) \to R$, $g: B = (c,d) \to R$, with $F(A) \leq g(B)$
suppose f has a derivative at $x \in A$, g has a derivative at $F(x) \in B$
then the composition $g \circ f: A = (a,b) \to R$ has a derivative at x and:

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$$

⑤ DERIVATIVE OF THE INVERSE

suppose $f: (a,b) \to R$, injective
suppose f has a derivative at $x_0 \in (a,b)$ with $f'(x_0) \neq 0$
then $\exists$ the inverse function f^{-1}, it has a derivative at $y_0 = f(x_0)$

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$$

EXAMPLES

- f: exponential, f^{-1}: logarithm $\Rightarrow (\ln x)' = \frac{1}{(e^y)'} = \frac{1}{e^y} = \frac{1}{x}$
- f: tan, f^{-1}: $\tan^{-1}$ (arctan) $\Rightarrow (\tan^{-1} x)' = \frac{1}{(\tan y)'} = \frac{1}{1+\tan^2 y} = \frac{1}{1+x^2}$
- $y = x^{\frac{1}{x}} \Rightarrow e^{\ln(x)\frac{1}{x}} = e^{\frac{1}{x}\ln x}$

$y = e^t$

$y' = e^t$

$= e^{\frac{1}{x}\ln x}$

$= x^{\frac{1}{x}}$

$t = \frac{1}{x}\ln x = x^{-1}\ln x$

$t' = -x^{-2}\ln x \to x^{-1}\frac{1}{x}$

$= -\frac{\ln x}{x^2} + \frac{1}{x^2}$

$= \frac{1-\ln x}{x^2}$

$$y' = x^{\frac{1}{x}} \frac{1-\ln x}{x^2}$$

PIECE-WISE DEFINED FUNCTIONS

① $y = |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases} \Rightarrow y' = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$

$$y'_+(0) = \lim_{h \to 0^+} \frac{h-0}{h} = \boxed{1} \Rightarrow y'(0) \nexists$$

$$y'_-(0) = \lim_{h \to 0^-} \frac{-h-0}{h} = \boxed{-1}$$

② $y = \begin{cases} x^2 & x<0 \\ x & x \geq 0 \end{cases} \Rightarrow y' = \begin{cases} 2x & x<0 \\ 1 & x>0 \end{cases}$

$$y'_+(0) = \lim_{h \to 0^+} \frac{h-0}{h} = \boxed{1}$$

$$y'_-(0) = \lim_{h \to 0^-} \frac{h^2-0}{h} = \boxed{0}$$

$\Rightarrow y'(0) \nexists$

③ $y = \begin{cases} x^2 & x<0 \\ x^3 & x \geq 0 \end{cases} \Rightarrow y' = \begin{cases} 2x & x<0 \\ 3x^2 & x>0 \end{cases}$

$$y'_+(0) = \lim_{h \to 0^+} \frac{h^{\overset{2}{\cancel{3}}} - 0}{\cancel{h}} = \boxed{0}$$

$$y'_-(0) = \lim_{h \to 0^-} \frac{h^{\cancel{2}} - 0}{\cancel{h}} = \boxed{0}$$

$\Rightarrow y'(0) \exists = 0$

HIGHER ORDER DERIVATES

$y = x^4 \Rightarrow y' = 4x^3 \Rightarrow y'' = 12x^2 \Rightarrow y''' = 24x$ (and so on)

DERIVABILITY AND CONTINUIT

THEOREM: EXISTANCE OF DERIVATIVE -> CONTINUITY

suppose $f: (a,b) \to \mathbb{R}$ has a derivative $f'(x_0)$ at a point $x_0 \in (a,b)$ then f is continuous at the point x_0

proof

suppose $\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = f'(x_0)$

rewrite it using:

$x = x_0 + h \Rightarrow h = x - x_0$

$h \to 0 = x \to x_0$

therefore

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

then

$$\lim_{x \to x_0} [f(x) - f(x_0)] = \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} (x - x_0) \right]$$

$$= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x \to x_0} (x - x_0)$$

$= f'(x_0) \cdot 0$

$= 0$

but also

$\lim_{x \to x_0} [f(x) - f(x_0)] = \lim_{x \to x_0} f(x) - \lim_{x \to x_0} f(x_0) = \lim_{x \to x_0} f(x) - f(x_0)$

and therefore

$\lim_{x \to x_0} f(x) - f(x_0) = 0 \quad \Rightarrow \lim_{x \to x_0} f(x) = f(x_0)$

meaning that $f(x)$ is continuous at x_0

but f continuous at $x_0 \not\Rightarrow f$ has a derivative at x_0

f DOES NOT HAVE A DERIVATIVE

① f is not continuous at x_0

② corners -> $f'_+(x_0)$, $f'_-(x_0)$ finite and different

③ cusps -> $\lim_{h \to 0^+}$ and $\lim_{h \to 0^-}$ infinite and different

④ half - cusps -> one limit is finite and the other infinite

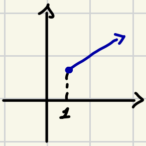
⑤ inflection points with vertical tangent

DERIVATIVE AT A BOUNDARY POINT

it is possible to define a derivative $f'(x_0)$ at a boundary point (ex. points a, b of a closed interval $[a,b]$) required that f has the one-sided derivative which is relevant

EXAMPLE

$y = x$ defined on $[1,5]$, $x_0 = 1$ =>



we say that f has a derivative at $x_0 = 1$ because $f'_+(1)$ $\exists$ finite

the same at $x_0 = 5$ because $f'_-(5)$ $\exists$ finite

DIFFENTIABILITY

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point for A

f is differentiable at x_0 = f can be approximated with a straight line in a neighbourhood $\mathcal{U}(x_0)$ of x_0

FORMAL DEFINITION = $f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point for A

f is differentiable at x_0 if: $\exists$ a neighbourhood $\mathcal{U}(x_0)$ of x_0 such that $\forall x = x_0 + h \in \mathcal{U}(x_0)$

$$f(x) = \underbrace{f(x_0) + m(x-x_0)}_{\text{non vertical straight line passing through } P_0} + \underbrace{o(x-x_0)}_{\text{negligible error}} \quad \text{as } x \to x_0$$

$\longrightarrow$ = tangent line

THEOREM: DIFFERENTIABILITY AND EXISTENCE OF A DERIVATIVE

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point for A

then

① f is differentiable at $x_0 \iff f$ has a derivative at x_0

② $m = f'(x_0)$

PROOF

since $f(x)$ is differentiable at x_0, we have:

$f(x) = f(x_0) + m(x-x_0) + o(x-x_0)$ as $x \to x_0$

substitute $x = x_0 + h \to h = x - x_0$:

$f(x_0+h) = f(x_0) + m \cdot h + o(h)$ as $h \to 0$

$f(x_0+h) - f(x_0) = m \cdot h + o(h)$ as $h \to 0$

$$\frac{f(x_0+h)-f(x_0)}{h} = m + \frac{o(h)}{h}$$

$$= m + o(1) \quad \text{as } h \to 0$$

$\left[\frac{o(h)}{h} = o(1) \text{ is true for the properties of } o \right]$

we take the limit on both sides when $h \to 0$ and obtain:

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = m$$

which means that $f(x)$ has a derivative $f'(x_0)$ at x_0 and:

$f'(x_0) = m$

$\Longleftarrow$

since $f(x)$ has a derivative at x_0, we have:

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)$$

this means that in a neighbourhood $U(x_0)$ of x_0, $\forall x = x_0+h \in U(x_0)$ we have:

$$\frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) + \theta(1) \quad \text{as } h \to 0$$

therefore:

$$\begin{aligned} f(x_0+h)-f(x_0) &= f'(x_0)h + \theta(1) \cdot h \\ &= f'(x_0)h + \theta(h) \quad \text{as } h \to 0 \end{aligned}$$

[$\theta(1) \cdot h = \theta(h)$ is true for the propositions of θ]

and

$$f(x_0+h) = f(x_0) + f'(x_0)h + \theta(h)$$

substitute $x = x_0+h \to x - x_0 = h$:

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \theta(x-x_0)$$

therefore $f(x)$ is differentiable at x_0 calling m the number $m = f'(x_0)$

DEF. OF DIFFERENTIABLE = $f(x) = f(x_0) + \underbrace{f'(x_0)(x-x_0)}_{\hookrightarrow \text{differential } df(x_0)} + \theta(x-x_0)$ as $x \to x_0$

DIFFERENTIAL

$$\left.\begin{aligned} df(x_0) &= f'(x_0)(x-x_0) \\ &= f'(x_0)\ h \\ &= f'(x_0)\ \Delta x \\ &\boxed{= f'(x_0)\ dx} \end{aligned}\right] \text{equal writings}$$

$$\Longrightarrow f'(x_0) = \frac{df(x_0)}{dx}$$

ROLE OF THE DIFFERENTIAL

$$f(x)-f(x_0) = f'(x_0)(x-x_0) + \underbrace{o(x-x_0)}_{\downarrow \atop \text{negligible error}} \Rightarrow \Delta f(x_0) = df(x_0)$$

the differential $df(x_0)$ is a good approximation of $\Delta f(x_0)$ at least when $x \to x_0$ (when the increment $x-x_0$ is "very small")

REMARK

$$\begin{tikzcd}
f \text{ has a derivative at } x_0 \arrow[rr, Leftrightarrow] \arrow[dr, Rightarrow, shift left] & & f \text{ is differentiable at } x_0 \arrow[dl, Rightarrow, shift left] \\
& f \text{ is continuous at } x_0 \arrow[ul, Rightarrow, shift left, "/" marking] \arrow[ur, Rightarrow, shift left, "/" marking] &
\end{tikzcd}$$

DISCONTINUITIES OF A DERIVATIVE FUNCTION

consider a function $g(x)$ defined on an interval $I \subseteq \mathbb{R}$ and suppose $g(x)$ is the derivative of another function $f(x)$ → not sure $g(x)$ is continuous

THEOREM

consider a function $g(x)$ defined on an interval $I \subseteq \mathbb{R}$
suppose $g(x) = f'(x) \Rightarrow g(x)$ is a derivative of the function $f(x)$
then:

① $g(x)$ has no eliminable discontinuity

② $g(x)$ has no jump

③ $g(x)$ may have essential discontinuities (at least one of the two limits $\lim_{x \to x_0^+} g(x)$ and $\lim_{x \to x_0^-} g(x)$ ∄ or is $+\infty / -\infty$)

$\Rightarrow$

- $C^{(1)}(A) = \{f : A \subseteq \mathbb{R} \to \mathbb{R},$ derivable with a continuous derivative on $A\}$
- $C^{(2)}(A) = \{f : A \subseteq \mathbb{R} \to \mathbb{R},$ twice-derivable with a continuous second derivative on $A\}$
- $C^{(n)}(A) = \{f : A \subseteq \mathbb{R} \to \mathbb{R},$ n-times derivable with a continuous n-th derivable on $A\}$
- $C^{(\infty)}(A) = \{f : A \subseteq \mathbb{R} \to \mathbb{R},$ infinitely many times derivable with all its derivatives being continuous on $A\}$

CONTINUITY

① $f: A \subseteq \mathbb{R} \to \mathbb{R}$, $x_0 \in A$

f is continuous at $x_0 \in A$ if:
$\forall \varepsilon > 0 \ \exists \sigma = \sigma(\varepsilon)$ such that
$\forall x \in A, |x - x_0| < \sigma \Rightarrow |f(x) - f(x_0)| < \varepsilon$

② $f: A \subseteq \mathbb{R} \to \mathbb{R}$

f is continuous on A if f is continuous at all points $x_0 \in A$
that is if:
$\forall x \in A, \forall \varepsilon > 0 \ \exists \sigma = \sigma(x_0, \varepsilon)$ such that
$\forall x \in A, |x - x_0| < \sigma \Rightarrow |f(x) - f(x_0)| < \varepsilon$

$\Rightarrow$

if we chang $x_0 \in A$ also σ changes
sometimes given $\varepsilon > 0$ we can choose the same $\sigma = \sigma(\varepsilon)$ for every point $x_0 \in A$ →

f is UNIFORMLY CONTINUOUS on A =

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

f is uniformly continuous on A if $\forall \varepsilon > 0 \ \exists \sigma = \sigma(\varepsilon)$ such that:
$\forall x, x_0 \in A, |x - x_0| < \sigma \Rightarrow |f(x) - f(x_0)| < \varepsilon$

f uniformly continuous on A $\underset{\not\Leftarrow}{\Rightarrow}$ f continuous on A

↳ if $f(x)$ is continuous on A but it increases/decreases too steeply then it is not uniformly continuous, however the two notions are equivalent when the domain $A \subseteq \mathbb{R}$ is compact

THEOREM

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, A compact
f is continuous on A $\Leftrightarrow$ f is uniformly continuous on A

LIPSCHITZ

a function $f: A \subseteq \mathbb{R} \to \mathbb{R}$ is Lipschitz on A if $\exists$ $k > 0$ such that:
$\forall x_1, x_2 \in A, |f(x_1) - f(x_2)| \leq k\,|x_1 - x_2|$

$\Rightarrow$

$\exists k > 0$ such that:
$\forall x_1, x_2 \in A$, with $x_1 \neq x_2$

$$\left| \frac{f(x_1)-f(x_2)}{x_1-x_2} \right| \leq k$$

this condition is linked with the behaviour of the difference quotient of f

THEOREM

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

if f is Lipschitz on A $\Rightarrow$ f is uniformly continuous on A

therefore, for a general set $A \subseteq \mathbb{R}$ and a function $f: A \subseteq \mathbb{R} \to \mathbb{R}$

Lipschitz $\underset{\not\Leftarrow}{\Rightarrow}$ uniformly continuous $\underset{\not\Leftarrow}{\Rightarrow}$ continuous

there are no general implications between derivability and the Lipschitz condition: Lipschitz $\not\Leftrightarrow$ derivability

however

THEOREM

if $f: A = [a,b] \to \mathbb{R}$, then:

$C^{(1)}(A) \Rightarrow$ Lipschitz $(\not\Leftarrow)$

USES OF DIFFERENTIAL CALCULUS: THE SEARCH FOR LOCAL MAX/MIN

$f'(x_0) = 0$ [horizontal tangent] $\not\Leftrightarrow$ x_0 local maximiser/minimiser

FERMAT'S THEOREM (necessary condition for local maximisers/minimisers)

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, if:

① x_0 is an interior point for A [excludes boundary points]

② f has a derivative at x_0 [excludes points at which $\nexists f'(x_0)$]

③ x_0 is a local maximiser/minimiser]

then:

$f'(x_0) = 0 \Rightarrow x_0 =$ stationary point

APPLICATION EXAMPLE OF FERMAT'S THEOREM

$y = x^3 - 3x$ → find local max/min

① domain $A = \mathbb{R}$ → no boundary points

f has a derivative in A → no points at which $\nexists f'(x)$

② $y' = 3x^2 - 3$ -> set $y' = 0$ => $3x^2 - 3 = 0$

$3x^2 = 3$

$x^2 = 1$ -> $\underbrace{x = 1,\ x = -1}$

we know nothing about this points

=>
we set $y' = 0$ because all interior points and such that $\exists f'(x)$ and for which we have $f'(x) \neq 0$ CANNOT be local max/min

Fermat's theorem says that all such points are not useful because if they were local max/min they would have $y' = 0$ = necessary condition for local maximisers/minimisers

ROLLE'S THEOREM

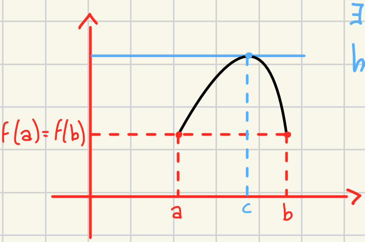
$f: A \subseteq \mathbb{R} \to \mathbb{R}$, $A = [a,b]$, if:

① f is continuous in $[a,b]$

② f has a derivative in (a,b)

③ $f(a) = f(b)$

then, $\exists c \in (a,b): f'(c) = 0$



$\exists (c): f'(c) = 0$ => horizontal tangent

LAGRANGE'S THEOREM (mean value theorem) -> generalisation of Rolle's

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, $A = [a,b]$, if:

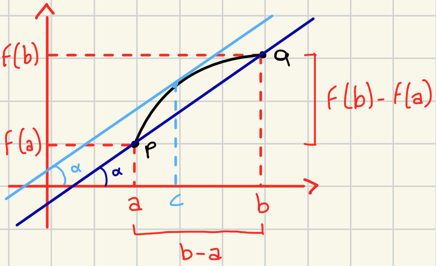
① f is continuous in $[a,b]$

② f has a derivative in (a,b)

then, $\exists c \in (a,b): f'(c) = \dfrac{f(b)-f(a)}{b-a}$

=>

$\exists$ at least one point $c \in (a,b)$: the tangent to the graph of $f(x)$ at c is parallel to the secant through $P(a, f(a))$, $Q(b, f(b))$



=>

$\dfrac{f(b)-f(a)}{b-a}$ is the average rate (mean value) of variation of $f(x)$ in $[a,b]$;

$\exists c: f'(c) = \dfrac{f(b)-f(a)}{b-a}$ means that $\exists c$ such that the instant rate of variation $f'(c)$ of $f(x)$ at c is equal to the average rate of variation $\dfrac{f(b)-f(a)}{b-a}$ of $f(x)$

INVERTIBILITY TEST (consequence of strict monotonicity test)

$f: A \subseteq R \to R$, if:

① A is an interval

② f has a derivative in A

③ $f'(x)$ always has the same sign in A [always >0 or always <0]

$\Rightarrow f(x)$ is invertible in A (since $f(x)$ is strictly monotone on interval A)

I SUFFICIENT CONDITION FOR LOCAL MAXIMISERS / MINIMISERS

$f: A \subseteq R \to R$, if:

① x_0 is an interior point of A

② f has a derivative in a neighbourhood $B(x_0)$ of x_0

③ $f'(x_0) = 0$ [$\Rightarrow x_0$ passes fermat's theorem]

* [$\lessgtr$ strict]

then: *

ⓐ if $f'(x) \leq 0$ in $B^-(x_0)$ and $f'(x) \geq 0$ in $B^+(x_0) \Rightarrow x_0$ is a local minimiser

ⓑ if $f'(x) \geq 0$ in $B^-(x_0)$ and $f'(x) \leq 0$ in $B^+(x_0) \Rightarrow x_0$ is a local maximiser

ⓒ if $f'(x) < 0$ in $B^-(x_0)$ and $f'(x) < 0$ in $B^+(x_0) \Rightarrow x_0$ is not a local max/min

EXAMPLE

Find local maximisers/minimisers of $y = x^3 - 3x$

$\Rightarrow$

① $y' = 3x^2 - 3 \to y' = 0$

$x = 1, \; x = -1$

② study the sign of y'

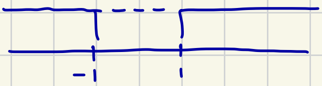
$y' > 0 \to 3x^2 - 3 > 0$

$3x^2 > 3$

$x^2 > 1$

$x < -1$ or $x > 1$ [$y' < 0$ for $-1 < x < 1$]

③ we get



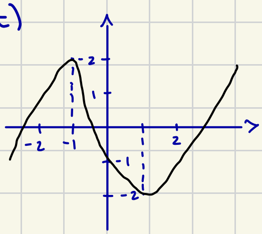
sign of y'

y increasing/decreasing

$\Rightarrow x = -1$ is a local maximiser (strict)

$x = 1$ is a local minimiser (strict)

$f(-1) = 2, \; f(1) = -2$



not global max/min

DE L'HÔPITAL'S THEOREM => computation of limits

INDETERMINATE FORM $\frac{0}{0}$

suppose f,g are differentiable on (a,b), with $a,b \in \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$

suppose $x_0 \in [a,b]$

suppose $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$

suppose $g'(x) \neq 0 \ \forall x \in (a,b)$

then:

$$\text{if } \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L \in \bar{\mathbb{R}} \Rightarrow \lim_{x \to x_0} \frac{f(x)}{g(x)} = L$$

INDETERMINATE FORM $\frac{\infty}{\infty}$

suppose f,g are differentiable on (a,b), with $a,b \in \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$

suppose $x_0 \in [a,b]$

suppose $\lim_{x \to x_0} f(x) = +\infty/-\infty$, $\lim_{x \to x_0} g(x) = +\infty/-\infty$

suppose $g'(x) \neq 0 \ \forall x \in (a,b)$

then:

$$\text{if } \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L \in \bar{\mathbb{R}} \Rightarrow \lim_{x \to x_0} \frac{f(x)}{g(x)} = L$$

EXAMPLES

notable limits $\frac{0}{0}$

① $$\lim_{x \to 0} \frac{e^x - 1}{x} \Rightarrow \lim_{x \to 0} \frac{e^x}{1} = 1 = \lim_{x \to 0} \frac{e^x - 1}{x} = 1$$

② $$\lim_{x \to 0} \frac{\ln(1+x)}{x} \Rightarrow \lim_{x \to 0} \frac{\frac{1}{1+x}}{1} = 1 = \lim_{x \to 0} \frac{\ln(1+x)}{1} = 1$$

③ $$\lim_{x \to 0} \frac{(1+x)^\alpha - 1}{x} \Rightarrow \lim_{x \to 0} \frac{\alpha(1+x)^{\alpha-1}}{1} = \alpha = \lim_{x \to 0} \frac{(1+x)^\alpha - 1}{x} = \alpha$$

④ $$\lim_{x \to 0} \frac{\sin x}{x} \Rightarrow \lim_{x \to 0} \frac{\cos x}{1} = 1 = \lim_{x \to 0} \frac{\sin x}{x} = 1$$

scales of infinities $\frac{\infty}{\infty}$

① $$\lim_{x \to +\infty} \frac{\ln x}{x} \Rightarrow \lim_{x \to +\infty} \frac{1/x}{1} = 0 = \lim_{x \to +\infty} \frac{\ln x}{x} = 0$$

② $$\lim_{x \to +\infty} \frac{e^x}{x^2} \Rightarrow \lim_{x \to +\infty} \frac{e^x}{2x} = \lim_{x \to +\infty} \frac{e^x}{2} = +\infty = \lim_{x \to +\infty} \frac{e^x}{x^2} = +\infty$$

SUFFICIENT CONDITION FOR DERIVABILITY

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point for A, suppose:

① f is continuous in a neighbourhood $\mathcal{U}(x_0)$

② f has a derivative in $\mathcal{U}(x_0) - \{x_0\}$

then:

ⓐ if $\lim_{x \to x_0^+} f'(x)$ $\exists$ finite $\Rightarrow f'_+(x_0) = \lim_{x \to x_0^+} f'(x)$

ⓑ if $\lim_{x \to x_0^-} f'(x)$ $\exists$ finite $\Rightarrow f'_-(x_0) = \lim_{x \to x_0^-} f'(x)$

ⓒ if $\lim_{x \to x_0} f'(x)$ $\exists$ finite $\Rightarrow f'(x_0) = \lim_{x \to x_0} f'(x)$

TAYLOR'S EXPANSION / MACLAURIN'S EXPANSION

TAYLOR'S THEOREM (ORDER 2)

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, then:

① x_0 interior point of A

② f has a derivative $f'(x)$ and second derivative $f''(x)$ in a neighbourhood $\mathcal{U}(x_0)$ [is twice differentiable]

then:

$\forall x = x_0 + h \in \mathcal{U}(x_0)$,

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2}_{\substack{\text{tangent parable at } x_0 \\ = \\ \text{Taylor's polynomial of degree 2}}} + \underbrace{o(x-x_0)^2}_{\substack{\text{negligible error} \\ = \\ \text{Peano's reminder}}} \Big] \xrightarrow{} \text{Taylor's expansion of order 2} \quad \text{as } x \to x_0$$

- Taylor's formula / expansion → polynomial + reminder ⇒ use order
- Taylor's polynomial → no reminder ⇒ use degree

MACLAURIN'S FORMULA / EXPANSION (special case of Taylor when $x_0 = 0$)

$\forall x \in \mathcal{U}(0)$,

$$f(x) = \underbrace{f(0) + f'(0)x + \frac{f''(0)}{2}x^2}_{\substack{\text{tangent parabola} \\ = \\ \text{Maclaurin's polynomial}}} + \underbrace{o(x^2)}_{\substack{\text{negligible error} \\ = \\ \text{Peano's reminder}}} \quad \text{as } x \to x_0$$

TAYLOR'S THEOREM (ORDER n)

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, if:

① x_0 interior point of A

② f is n times differentiable in a neighbourhood $\mathcal{U}(x_0)$

then:

$\forall x = x_0 + h \in U(x_0)$

$$\underbrace{f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3 + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n}_{\text{Taylor's polynomial of degree n}}$$

$$\underbrace{+\, o(x-x_0)^n}_{\text{Peano's reminder}} \quad \text{as } x \to x_0$$

↳ MACLAURIN'S FORMULA $\Rightarrow x_0 = 0$

MACLAURIN'S EXPANSION OF ELEMENTARY FUNCTIONS

① $y = e^x$, $x_0 = 0$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots \qquad x \to 0$$

② $y = \sin(x)$, $x_0 = 0$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots$$

③ $y = \cos(x)$, $x_0 = 0$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^{n+1} \frac{x^n}{(2n)!} + \dots \qquad x \to 0$$

④ $y = \ln(1+x)$, $x_0 = 0$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n+1} \frac{x^n}{n} + \dots \qquad x \to 0$$

⑤ $y = (1+x)^\alpha$, $x_0 = 0$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!} x^3 + \dots + \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-(n-1))}{n!} x^n + \dots$$

$$x \to 0$$

II SUFFICIENT CONDITION FOR LOCAL MAXIMISERS/MINIMISERS

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, suppose that:

① x_0 is an interior point of A

② f is n times differentiable in a neighbourhood $\mho(x_0)$

③ $f'(x_0) = 0$

④ all the higher-order derivates at x_0 are $= 0$ till $\exists\ n \geq 2$ such that $f^{(n)}(x_0) \neq 0$

then:

ⓐ if n is even and $f^{(n)}(x_0) > 0 \Rightarrow x_0$ is a local minimiser

ⓑ if n is even and $f^{(n)}(x_0) < 0 \Rightarrow x_0$ is a local maximiser

ⓒ if n is odd $\Rightarrow x_0$ is not a local maximiser/minimiser

THEOREM (weaker because cannot be applied if $f''(x_0) = 0$)

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, suppose that:

① x_0 is an interior point of A

② f is twice differentiable in a neighbourhood $\mho(x_0)$

③ $f'(x_0) = 0$

then:

ⓐ if $f''(x_0) > 0 \Rightarrow x_0$ is a local minimiser

ⓑ if $f''(x_0) < 0 \Rightarrow x_0$ is a local maximiser

GENERAL PROCEDURE FOR LOCAL MAXIMUM / MINIMUM

given $f: A \subseteq \mathbb{R} \to \mathbb{R}$

① - points $x_0 \in A$ that are boundary points $\to$ check separately
- points $x_0 \in A$ such that $\nexists f'(x_0) \longrightarrow$ check separately

② all other points $\to$ Fermat's necessary condition $\Rightarrow f'(x) = 0$

③ on those points which pass Fermat's condition $\to$
- I sufficient condition $\Rightarrow$ only $f'(x)$ but involves inequalities

or

- II sufficient condition $\Rightarrow$ also higher-order derivates but no inequalities involved

CONCAVITY / CONVEXITY

THEOREM: CONVEXITY AND TANGENT

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

if A is an open interval and f is differentiable on A, then:

f is convex in $A \iff \forall x_0, x \in A$,

$$f(x_0) + f'(x_0)(x - x_0) \leq f(x) \quad [\text{CONCAVE} \Rightarrow \geq f(x)]$$

THEOREM: CONVEXITY TEST FOR DIFFERENTIABLE FUNCTIONS

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

if A is an open interval and f is differentiable on A, then:

① f is convex in $A \iff f'$ is increasing in A [CONCAVE $\Rightarrow$ decreasing]

② f is strictly convex in $A \iff f'$ is strictly increasing in A

THEOREM: CONVEXITY TEST FOR TWICE-DIFFERENTIABLE FUNCTIONS

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

if A is an open interval and f is twice-differentiable on A, then:

① f is convex in $A \iff f'' \geq 0$ in A [CONCAVE $\Rightarrow f'' \leq 0$]

② $f'' > 0$ in $A \Rightarrow f$ is strictly convex in A [CONCAVE $\Rightarrow f'' < 0$]

GLOBAL MAXIMA/MINIMA

GENERAL METHODS TO SEE IF GLOBAL MAXIMA/MINIMA EXIST

① draw the graph

② use Weierstrass (or Tonelli) → if the function $f(x)$ is continuous on a domain $A \subseteq \mathbb{R}$ which is a compact set $A \subseteq \mathbb{R}$, Weierstrass's theorem guarantees that both a global maximum and a global minimum $\exists$ BUT it does not say how to find them

THEOREM: I SUFFICIENT CONDITION, GLOBAL VERSION

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, if:

① A is an interval

② x_0 is an interior point of A

③ f has a finite derivative in A

④ $f'(x_0) = 0$

then: *

ⓐ if $f'(x) \leq 0 \ \forall x \in A, x < x_0$ and $f'(x) \geq 0 \ \forall x \in A, x > x_0$ then x_0 = global minimiser

ⓑ if $f'(x) \geq 0 \ \forall x \in A, x < x_0$ and $f'(x) \leq 0 \ \forall x \in A, x > x_0$ then x_0 = global minimiser

* [$\gtrless$ strict]

THEOREM: II SUFFICIENT CONDITION, GLOBAL VERSION

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, if:

① A is an interval

② x_0 is an interior point of A

③ f is twice differentiable in A

④ $f'(x_0) = 0$

then: *

ⓐ if $f''(x) \geq 0$ in A then x_0 = global minimiser

ⓑ if $f''(x) \leq 0$ in A then x_0 = global maximiser

* [$\gtrless$ strict]

↱ or MINIMUM

THEOREM: NECESSARY AND SUFFICIENT MAXIMUM CONDITION FOR A CONCAVE AND DIFFERENTIABLE FUNCTION

↳ or CONVEX

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, if:

① A is an interval

② f is concave (or convex) in A

③ f has a finite derivative in A

④ x_0 interior point of A

then:

$f'(x_0) = 0 \iff x_0$ is a local maximiser (or minimiser) $\iff x_0$ is a global maximiser (or minimiser)

HOW TO DROW A GRAPH

① domain A

② limits at the boundary $\to$ horizontal / vertical asymptotes

③ sign of $f(x)$ + intersections with coordinate axes

④ study of $f'(x) \to f$ increasing/decreasing + local max / min (I sufficient condition)

⑤ study of $f''(x) \to f$ convex / concave + local max / min (II sufficient condition)

⑥ graph + study of exceptional points + global max / min

EXAMPLE $\to$ $y = \dfrac{x^2+2}{x^2-2}$

① domain

$x^2 - 2 \neq 0 \implies x \neq \sqrt{2}, -\sqrt{2}$

$A = (-\infty, -\sqrt{2}) \cup (-\sqrt{2}, \sqrt{2}) \cup (\sqrt{2}, +\infty)$

② limits / asymptotes

$$\underbrace{\lim_{x \to -\infty} \frac{x^2+2}{x^2-2} = 1^+}_{\text{horizontal asymptote}} \qquad \underbrace{\lim_{x \to +\infty} \frac{x^2+2}{x^2-2} = 1^+}_{\text{horizontal asymptote}}$$

$y = 1$ as $x \to -\infty$ $\qquad$ $y = 1$ as $x \to +\infty$

$$\lim_{x \to \sqrt{2}^-} \frac{x^2+2}{x^2-2} = -\infty \qquad \lim_{x \to \sqrt{2}^+} \frac{x^2+2}{x^2-2} = +\infty$$

vertical asymptote $x = \sqrt{2}$

$$\lim_{x \to -\sqrt{2}^-} \frac{x^2+2}{x^2-2} = +\infty \qquad \lim_{x \to -\sqrt{2}^+} \frac{x^2+2}{x^2-2} = -\infty$$

vertical asymptote $x = -\sqrt{2}$

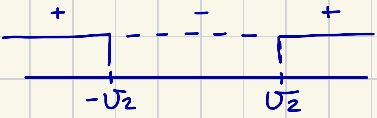
③ sign / intersection with axes

$\frac{x^2+2}{x^2-2} > 0$ N > 0 always

D > 0 if $x^2-2>0$

$\Rightarrow x^2 > 2$

$\Rightarrow x < -\sqrt{2}, \; x > \sqrt{2}$



intersection with y-axes $\begin{cases} x = 0 \\ y = -1 \end{cases}$

intersection with x-axes $\begin{cases} y = 0 \\ \text{never} \end{cases}$

④ f'(x)

$$f(x) = \frac{x^2+2}{x^2-2} \rightarrow y' = \frac{(x^2-2)2x - (x^2+2)2x}{(x^2-2)^2}$$

$$= \frac{2x^3 - 4x - 2x^3 - 4^x}{(x^2-2)^2}$$

$$= \frac{-8x}{(x^2-2)^2}$$

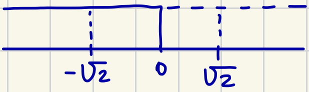
$y' = 0 \rightarrow -8x = 0 \Rightarrow x = 0$

$y' > 0 \rightarrow \frac{-8x}{(x^2-2)^2} > 0$

D > 0 (excluded $x = -\sqrt{2}, \sqrt{2}$ where $f(x) \nexists$)

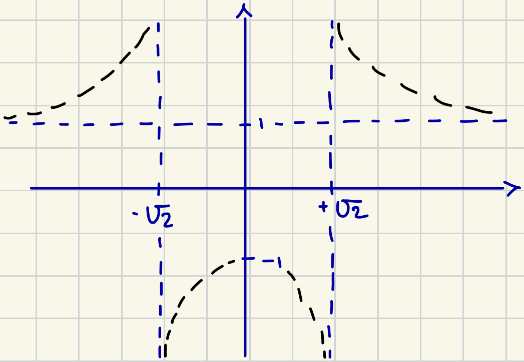
N > 0 $-8x > 0 \Rightarrow x < 0$

x = 0 is local maximiser



⑤ f''(x) not required

⑥ graph



theorems and proofs for differential calculus

DERIVATIVE OF A CONSTANT

$y = k \Rightarrow y' = 0$ ($\forall$ constant $k \in R$)

PROOF

$f: R \rightarrow R$, $f(x) = k$, $\forall x \in R$, $\forall K \neq 0$:

$$f'(x) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{h \to 0} \frac{k-k}{h} = \lim_{h \to 0} \frac{0}{k} = 0$$

DERIVATIVE OF $y = x^{\alpha}$

$$y = x^{\alpha} \Rightarrow y' = \alpha x^{\alpha-1}$$

PROOF

$$f'(x) = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h \to 0} \frac{(x+h)^{\alpha} - x^{\alpha}}{h} =$$

$$\lim_{h \to 0} \frac{\left[x\left(1+\frac{h}{x}\right)\right]^{\alpha} - x^{\alpha}}{h} = \lim_{h \to 0} \frac{x^{\alpha}\left(1+\frac{h}{x}\right)^{\alpha} - x^{\alpha}}{h} = \lim_{h \to 0} \boxed{x^{\alpha}} \frac{\left(1+\frac{h}{x}\right)^{\alpha} - 1}{\boxed{h}} =$$

$$\frac{x^{\alpha}}{h} = \frac{x^{\alpha-1} \cdot x}{h} =$$

$$x^{\alpha-1} \frac{1}{\frac{h}{x}}] = t$$

$$\lim_{t \to 0} x^{\alpha-1} \underbrace{\boxed{\frac{(1+t)-1}{t}}}_{\alpha} = \alpha x^{\alpha-1}$$

DERIVATIVE OF $y = e^x$

$f: R \to R, f(x) = a^x, a > 0$

is derivable $\forall x \in R$, $f': R \to R$ given by $f'(x) = a^x \log a$

in particular:

$y = e^x \Rightarrow y' = e^x$

PROOF

$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = \lim_{h \to 0} \frac{a^x (a^h - 1)}{h} = a^x \lim_{h \to 0} \frac{a^h - 1}{h} =$
$a^x \log a$

DERIVATIVE OF A LINEAR COMBINATION

suppose f, g have a derivative $\forall x \in (a,b)$
suppose $\alpha, \beta \in \mathbb{R}$
$\alpha f + \beta g$ has a derivative $\forall x \in (a,b)$ and:
$(\alpha f + \beta g)' = \alpha f' + \beta g'$

f, g have a der. $\forall x \in (a,b)$
$\alpha, \beta \in \mathbb{R}$
$\alpha f + \beta g$ has a der. $\forall x \in (a,b)$
$(\alpha f + \beta g)' = \alpha f' + \beta g'$

PROOF:

$\forall x \in (a,b)$,

$$(\alpha f + \beta g)'(x) = \lim_{h \to 0} \frac{(\alpha f + \beta g)(x+h) - (\alpha f + \beta g)(x)}{h} =$$

$$\lim_{h \to 0} \frac{(\alpha f)(x+h) + (\beta g)(x+h) - (\alpha f)(x) - (\beta g)(x)}{h} =$$

$$\lim_{h \to 0} \left[\alpha \frac{f(x+h) - f(x)}{h} + \beta \frac{g(x+h) - g(x)}{h} \right] =$$

$$\alpha \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + \beta \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} = \alpha f'(x) + \beta g'(x)$$

DERIVATIVE OF $y = \ln x$

$y = \ln x \Rightarrow y' = \frac{1}{x}$

PROOF

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{h \to 0} \frac{\ln(x+h) - \ln x}{h} = \lim_{h \to 0} \frac{\ln\left(\frac{x+h}{x}\right)}{h}$$

$$= \lim_{h \to 0} \frac{1}{h} \ln\left(1 + \frac{h}{x}\right) = \lim_{h \to 0} \ln\left(1 + \frac{h}{x}\right)^{\frac{1}{h}}$$

let $t = \frac{h}{x}$; $h = tx$; $\frac{1}{h} = \frac{1}{tx}$; $\frac{1}{t} \cdot \frac{1}{x}$

$$= \lim_{t \to 0} \ln(1+t)^{\frac{1}{t} \cdot \frac{1}{x}} = \lim_{t \to 0} \ln\left[(1+t)^{\frac{1}{t}}\right]^{\frac{1}{x}} = \frac{1}{x} \lim_{t \to 0} \ln(1+t)^{\frac{1}{t}}$$

$$= \frac{1}{x} \ln \underbrace{\left[\lim_{t \to 0} (1+t)^{\frac{1}{t}}\right]}_{= e} = \frac{1}{x} \underbrace{\ln e}_{= 1} = \frac{1}{x}$$

DERIVATIVE OF $y = \arctan x$

$f: [-\frac{\pi}{2} ; \frac{\pi}{2}] \rightarrow R$ given by $f(y) = \tan y$

then, $f^{-1}: R \rightarrow [-\frac{\pi}{2} ; \frac{\pi}{2}]$ is given by $f^{-1}(x) = \arctan x$

PROOF

from the derivative of the inverse function

$$\frac{d \arctan x}{dx} = \frac{1}{1+\tan^2 y} = \frac{1}{1+x^2}$$

RELATIONSHIP BETWEEN DERIVABILITY AND CONTINUITY

suppose $f: (a,b) \to \mathbb{R}$ has a derivative $f'(x_0)$ at a point $x_0 \in (a,b)$ then f is continuous at the point x_0

PROOF

suppose $\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)$

rewrite it using:

$x_0+h=x \Rightarrow h = x-x_0$

$h \to 0 = x \to x_0$

therefore

$$\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$$

then

$$\lim_{x \to x_0} [f(x)-f(x_0)] = \lim_{x \to x_0} \left[\frac{f(x)-f(x_0)}{x-x_0} (x-x_0) \right]$$

$$= \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \cdot \lim_{x \to x_0} (x-x_0) = f'(x_0) \cdot 0 = 0$$

but also

$$\lim_{x \to x_0} [f(x)-f(x_0)] = \lim_{x \to x_0} f(x) - \lim_{x \to x_0} f(x_0) = \lim_{x \to x_0} f(x) - f(x_0)$$

and therefore

$$\lim_{x \to x_0} f(x) - f(x_0) = 0 \Rightarrow \lim_{x \to x_0} f(x) = f(x_0)$$

meaning that $f(x)$ is continous at x_0

but f continuous at x_0 $\not\Rightarrow$ f has a derivative at x_0

RELATIONSHIP BETWEEN DERIVABILITY AND DIFFERENTIABILITY

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, x_0 interior point for A

then

① f is differentiable at x_0 $\Leftrightarrow$ f has a derivative at x_0

② $m = f'(x_0)$

PROOF

since $f(x)$ is differentiable at x_0, we have:

$f(x) = f(x_0) + m(x-x_0) + o(x-x_0)$ as $x \to x_0$

substitute $x = x_0 + h \to h = x - x_0$:

$f(x_0+h) = f(x_0) + m \cdot h + o(h)$ as $h \to 0$

$f(x_0+h) - f(x_0) = m \cdot h + o(h)$ as $h \to 0$

$$\frac{f(x_0+h)-f(x_0)}{h} = m + \frac{o(h)}{h}$$

$$= m + o(1) \quad \text{as } h \to 0$$

[$\frac{o(h)}{h} = o(1)$ is true for the properties of o]

we take the limit on both sides when $h \to 0$ and obtain:

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = m$$

which means that $f(x)$ has a derivative $f'(x_0)$ at x_0 and:

$f'(x_0) = m$

$\Longleftarrow$

since $f(x)$ has a derivative at x_0, we have:

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)$$

this means that in a neighbourhood $U(x_0)$ of x_0, $\forall x = x_0 + h \in U(x_0)$ we have:

$$\frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) + o(1) \quad \text{as } h \to 0$$

therefore:

$f(x_0+h) - f(x_0) = f'(x_0)h + o(1) \cdot h$

$= f'(x_0)h + o(h)$ as $h \to 0$

[$o(1) \cdot h = o(h)$ is true for the propositions of o]

and

$$f(x_0+h) = f(x_0) + f'(x_0)h + o(h)$$

substitute $x = x_0 + h \rightarrow x - x_0 = h$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

therefore $f(x)$ is differentiable at x_0 calling m the number $m = f'(x_0)$

FERMAT'S THEOREM

$f : A \subseteq \mathbb{R} \to \mathbb{R}$ defined on a set A in $\mathbb{R}$

C subset of A

f differentiable at an interior point $x_0 \in C$

x_0 maximiser/minimiser of f on C

then $f'(x_0) = 0 \implies x_0 =$ stationary point

PROOF

let $x_0 \in C$ interior point and a local maximiser on C

$\exists$ a neighbourhood $B_\varepsilon(x_0)$ such that $f(x_0) \geq f(x_0)$ $\forall x \in B_\varepsilon(x_0) \cap C$

$h \in (0, \varepsilon)$, $x_0 + h \in B_\varepsilon(x_0) \implies$

$$\frac{f(x_0+h) - f(x_0)}{h} \leq 0 \quad \forall h \in (0, \varepsilon)$$

$$\lim_{h \to 0^+} \frac{f(x_0+h) - f(x_0)}{h} \leq 0$$

$h \in (-\varepsilon, 0)$, $x_0 + h \in B_\varepsilon(x_0) \implies$

$$\frac{f(x_0+h) - f(x_0)}{h} \geq 0 \quad \forall h \in (-\varepsilon, 0)$$

$$\lim_{h \to 0^-} \frac{f(x_0+h) - f(x_0)}{h} \geq 0$$

together:

$$0 \leq \lim_{h \to 0^-} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{h \to 0} \frac{f(x_0+h) + f(x_0)}{h} = \lim_{h \to 0^+} \frac{f(x_0+h) - f(x_0)}{h} \leq 0$$

since the hypothesis $f'(x_0)$ exists:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = 0$$

ROLLE'S THEOREM

$f: A \subseteq \mathbb{R} \to \mathbb{R}$, $A = [a,b]$, if:

① f is continuous in $[a,b]$

② f has a derivative in (a,b)

③ $f(a) = f(b)$

then, $\exists x_0 \in (a,b) : f'(x_0) = 0$

PROOF

according to Weierstrass $\exists\, x_1, x_2 \in [a,b]$ such that:

$f(x_1) = \min_{x \in [a,b]} f(x) = m$

$f(x_2) = \max_{x \in [a,b]} f(x) = M$

if $m = M$ the f is constant, $f(x) = m = M$ so $f'(x) = 0$ $\forall x \in (a,b)$

if $m < M$ then at least one between x_1 and x_2 must be an interior point for (a,b), in fact they can't be both boundary points since $f(a) = f(b)$

if $x_1 \in (a,b)$ for the Fermat's theorem $f'(x_1) = 0$, same in the case of x_2

LAGRANGE'S VALUE THEOREM

let $f:[a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b)
then there exists $\hat{x} \in (a,b)$ such that:

$$f'(\hat{x}) = \frac{f(b)-f(a)}{b-a} \quad *$$

PROOF

let $g:[a,b] \to \mathbb{R}$ be the auxiliary function defined by:

$$g(x) = f(x) - \left(f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right)$$

it is the difference between f and the straight line passing through the points $(a, f(a))$ and $(b, f(b))$
the function g is continuous on $[a,b]$ and differentiable on (a,b)
moreover $g(a) = g(b) = 0$
by Rolle's theorem there exists $\hat{x} \in (a,b)$ such that $g'(\hat{x}) = 0$, but
$g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ and therefore $f'(\hat{x}) - \frac{f(b)-f(a)}{b-a} = 0$

that is, $\hat{x}$ satisfies condition $*$

CHARACTERIZATION OF FUNCTIONS WITH NULL DERIVATIVE

let $f:[a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b)

then $f'(x)=0$ for every $x\in(a,b)$ if and only if f is constant, that is, if and only if there exists $k\in\mathbb{R}$ such that $f(x)=k \quad \forall x\in[a,b]$

PROOF

"only if" $\Rightarrow$

"if" part is the simple property of derivatives of derivability and continuity

let $x\in(a,b]$ and let us apply the mean value theorem on the interval $[a,x]$

it yields a point $\hat{x}\in(a,x)$ such that:

$0=f'(\hat{x})=\dfrac{f(x)-f(a)}{x-a}$ that is $f(x)=f(a)$

since x is any point in $(a,b]$ it follows that $f(x)=f(a)$ for any $x\in[a,b]$

CHARACTERIZATION OF FUNCTIONS WITH THE SAME DERIVATIVE

let $f, g : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b)
then $f'(x) = g'(x)$ $\forall x \in (a,b)$ if and only if there exists $k \in \mathbb{R}$ such that $f(x) = g(x) + k$ $\forall x \in [a,b]$

two functions that have the same first derivative are thus equal up to an additive constant k

PROOF

since "if" is obvious we prove "only if" $\Rightarrow$
let $h : [a,b] \to \mathbb{R}$ be the auxiliary function $h(x) = f(x) - g(x)$
we have $h'(x) = f'(x) - g'(x) = 0$ $\forall x \in (a,b)$
therefore by the characterization of functions with null derivative h is constant on $[a,b]$
that is there exists $k \in \mathbb{R}$ such that $h(x) = k$ $\forall x \in [a,b]$ so $f(x) = g(x) + k$ $\forall x \in [a,b]$

MONOTONICITY TEST ON INTERVAL

let $f:(a,b) \to \mathbb{R}$ be a differentiable function with $a,b \in \bar{\mathbb{R}}$

then f is (globally) increasing on (a,b) if and only if $f'(x) \geq 0 \ \forall x \in (a,b)$

because of the clause $a,b \in \bar{\mathbb{R}}$ the interval (a,b) can be unbounded for example $(a,b) = \mathbb{R}$

a dual result with negativity of the derivative on (a,b) holds for the decreasing monotonicity

PROOF

"only if" =>

suppose that f is increasing, let $x \in (a,b)$

$\forall h > 0$ we have $f(x+h) \geq f(x)$ hence $\dfrac{f(x+h)-f(x)}{h} \geq 0$

it follows that:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h \to 0^+} \frac{f(x+h)-f(x)}{h} \geq 0$$

"if" =>

let $f'(x) \geq 0 \ \forall x \in (a,b)$

let $x_1, x_2 \in (a,b)$ with $x_1 < x_2$

by the mean value theorem there exists $\hat{x} \in [x_1, x_2]$ such that $f'(\hat{x}) = \dfrac{f(x_2)-f(x_1)}{x_2 - x_1}$

since $f'(\hat{x}) \geq 0$ and $x_2 - x_1 > 0$, this shows that $f(x_2) \geq f(x_1)$

STRICT MONOTONICITY TEST ON AN INTERVAL

let $f: (a,b) \to \mathbb{R}$ be a differentiable function, with $a, b \in \bar{\mathbb{R}}$

if $f'(x) > 0 \ \forall x \in (a,b)$, then f is (globally) strictly increasing on (a,b)

PROOF

let $f'(x) > 0 \ \forall x \in (a,b)$ and let $x_1, x_2 \in (a,b)$ with $x_1 < x_2$

by the mean value theorem there exists $c \in [x_1, x_2]$ such that $\overset{*}{f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}}$

since $f'(c) > 0 \ \forall c$ and $x_2 - x_1 > 0$, from $*$ it follows that $f(x_2) > f(x_1)$

MACLAURIN FORMULA FOR $\log(1+x)$

the function $f: (-1, \infty) \to \mathbb{R}$ given by $f(x) = \log(1+x)$ is n times differentiable at each point of its domain, with

$$f^{(n)}(x) = (-1)^{n+1} \frac{(n-1)!}{(1+x)^n} \quad \forall n \geq 1$$

therefore, Taylor's expansion of order n of f at $x_0 > -1$ is:

$$\log(1+x) = \log(1+x_0) + \sum_{k=1}^{n} (-1)^{k+1} \frac{(x-x_0)^k}{k(1+x_0)^k} + o((x-x_0)^n)$$

a simple polynomial thus locally approximates the logarithmic function in particular the Maclaurin's expansion of order n of f is:

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots + (-1)^{n+1} \frac{x^n}{n} + o(x^n) = \sum_{k=1}^{n} (-1)^{k+1} \frac{x^k}{k} + o(x^n)$$

N-VARIABLE DIFFERENTIAL CALCULUS

INTRODUCTION TO PARTIAL DERIVATIVES

consider an interior point x_0 of the domain $A \subseteq \mathbb{R}^2$

$\underline{x}_0 = (x_{0_1}, x_{0_2})$ starting point

$\underline{h} = (h_1, h_2)$ increment of $\underline{x}$ (vector in one of infinitely many directions)

$\underline{x}_0 + \underline{h} = (x_{0_1} + h_1, x_{0_2} + h_2)$ final point

$f(\underline{x}_0) = f(x_{0_1}, x_{0_2})$

$f(\underline{x}_0 + \underline{h}) = f(x_{0_1} + h_1, x_{0_2} + h_2) - f(\underline{x}_{0_1}, x_{0_2}) = \Delta f$ increment of f

$\Longrightarrow$

we only accept to move in a direction which is ∥ to one of the axes

there are two such directions ⟶

① move ∥ to x_1- axis

- $\underline{x}_0 = (x_{0_1}, x_{0_2})$
- $\underline{h} = (h, 0)$ increment of $\underline{x}$ when we move ∥ to x_1- axis
  - ↳ $\underline{h}$ is still a vector but it is now only determined by one number, its non-null I component h
- $\underline{x}_0 + \underline{h} = (x_{0_1} + h_1, x_{0_2})$
- $f(\underline{x}_0) = f(x_{0_1}, x_{0_2})$
- $f(\underline{x}_0 + \underline{h}) = f(x_{0_1} + h_1, x_{0_2})$
- $f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) = f(x_{0_1} + h, x_{0_2}) - f(x_{0_1}, x_{0_2}) = \Delta f$ increment of f when we move ∥ to x_1- axis
- $\dfrac{\Delta f}{\underline{h}} = \dfrac{f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0)}{h} = \dfrac{f(x_{0_1} + h, x_{0_2}) - f(x_{0_1}, x_{0_2})}{h}$
  $\underline{h} \longrightarrow$ number
- partial derivative with respect to x_1 (if the increment $\underline{h}$ is ∥ to x_1-axis)
  $f'_{x_1}(\underline{x}_0) = \lim\limits_{h \to 0} \dfrac{f(x_{0_1} + h, x_{0_2}) - f(x_{0_1}, x_{0_2})}{h}$ $\exists$ finite

② move ∥ to x_2- axis

- $\underline{x}_0 = (x_{0_1}, x_{0_2})$
- $\underline{h} = (0, h)$
- $\underline{x}_0 + \underline{h} = (x_{0_1}, x_{0_2} + h)$
- $f(\underline{x}_0) = f(x_{0_1}, x_{0_2})$
- $f(\underline{x}_0 + \underline{h}) = f(x_{0_1}, x_{0_2} + h)$
- $f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) = f(x_{0_1}, x_{0_2} + h) - f(x_{0_1}, x_{0_2}) = \Delta f$ increment of f when we move ∥ to x_2- axis
- $\dfrac{\Delta f}{h} = \dfrac{f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0)}{h} = \dfrac{f(x_{0_1}, x_{0_2} + h) - f(x_{0_1}, x_{0_2})}{h}$

- partial derivative with respect to x_2

$$f'_{x_2}(\underline{x}_0) = \lim_{h \to 0} \frac{f(x_{01}, x_{02}+h) - f(x_{01}, x_{02})}{h} \quad \exists \text{ finite}$$

NOTIONS of partial derivatives of the function $f(x_1, x_2)$

$$f'_{x_1}(x_1, x_2),\ f'_{x_2}(x_1, x_2) \Rightarrow \frac{\partial f}{\partial x_1}(x_1, x_2),\ \frac{\partial f}{\partial x_2}(x_1, x_2)$$

f'_{x_1} is the derivative of $f(x_1, x_2)$ when only x_1 varies and x_2 remains constant; f'_{x_2} is the derivative of $f(x_1, x_2)$ when only x_2 varies and x_1 remains constant

EXAMPLE

$$f(x_1, x_2) = x_1^2 + x_2^3 \Rightarrow f'_{x_1}(x_1, x_2) = 2x_1,\ f'_{x_2}(x_1, x_2) = 3x_2^2$$

GRADIENT VECTOR

if $f(x_1, x_2)$ has finite partial derivatives f'_{x_1}, f'_{x_2} at $\underline{x}_0$ then the vector which contains the two partial derivatives is the gradient vector of f at $\underline{x}_0$ $\Rightarrow$ $\operatorname{grad} f(\underline{x}_0) = [f'_{x_1}(\underline{x}_0), f'_{x_2}(\underline{x}_0)]$ vector of numbers

with the general point $\underline{x} = (x_1, x_2) \Rightarrow \operatorname{grad} f(x_1, x_2) = [f'_{x_1}(x_1, x_2), f'_{x_2}(x_1, x_2)]$
vector of functions

another notion $\Rightarrow \nabla f(\underline{x}_0) \to \nabla f(x_1, x_2)$ [∇ = nabla]

FORMAL DEFINITION

$f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

take an interior point $\underline{x}_0$ of A: $\underline{x}_0 = (x_{01}, x_{02}, \dots, x_{0n})$

take an increment $\underline{h}$ for which only one component is $\neq 0$ (the increment only involves one variable x_j): $\underline{h} = (0, \dots, 0, \overset{\rightarrow x_j}{h}, 0 \dots, 0)$

then: $\underline{x}_0 + \underline{h} = (x_{01}, \dots, x_{0j-1}, x_{0j}+h, x_{0j+1}, \dots, x_{0n})$

then consider:

$f(\underline{x}_0) = f(x_{01}, \dots, x_{0n})$

$f(\underline{x}_0 + \underline{h}) = f(x_{01}, \dots, x_{0j-1}, x_{0j}+h, x_{0j+1}, \dots, x_{0n})$

and:

$f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) = f(x_{01}, \dots, x_{0j-1}, x_{0j}+h, x_{0j+1}, \dots, x_{0n}) - f(x_{01}, \dots, x_{0n}) = \Delta f$

then take the difference quotient:

$$\frac{\Delta f}{h} = \frac{f(x_{01}, \dots, x_{0j-1}, x_{0j}+h, x_{0j+1}, \dots, x_{0n}) - f(x_{01}, \dots, x_{0n})}{h}$$

we can now define the partial derivative of f at $\underline{x}_0$ with respect to the variable x_j:

$$f'_{x_1}(\underline{x}_0) = \lim_{h \to 0} \frac{f(x_{01},\dots,x_{0j-1},x_{0j}+h,x_{0j+1},\dots,x_{0n}) - f(x_{01},\dots,x_{0n})}{h} \quad \exists \text{ finite}$$

$\forall j = 1,2,\dots,n$

we define: $f'_{x_1}(\underline{x}_0),\dots,f'_{x_n}(\underline{x}_0)$ $\left[\frac{\partial f}{\partial x_1}(\underline{x}_0),\dots,\frac{\partial f}{\partial x_n}(\underline{x}_0)\right]$

we apply it to a general point $\underline{x} = (x_1,\dots,x_n)$ from the function $f(x_1,\dots,x_n)$ we get n functions:

$f'_{x_1}(x_1,\dots,x_n),\dots,f'_{x_n}(x_1,\dots,x_n)$ $\left[\frac{\partial f}{\partial x_1}(x_1,\dots,x_n),\dots,\frac{\partial f}{\partial x_n}(x_1,\dots,x_n)\right]$

called the n partial derivatives of the function $f(x_1,\dots,x_n)$

GRADIENT VECTOR

at $\underline{x}_0$:

$\text{grad} f(\underline{x}_0) = (f'_{x_1}(\underline{x}_0),\dots,f'_{x_n}(\underline{x}_0))$ vector of numbers

at $\underline{x} = (x_1,\dots,x_n)$:

$\text{grand} f(x_1,\dots,x_n) = (f'_{x_1}(x_1,\dots,x_n),\dots,f'_{x_n}(x_1,\dots,x_n))$ vector of functions

INTRODUCTION TO DIFFERENTIABILITY

$f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

consider an interior point $\underline{x}_0$ of A, then:

→ only strong notion

$$\begin{tikzcd} & f \text{ is differentiable at } \underline{x}_0 \arrow[dl, Rightarrow, shift right] \arrow[dr, Rightarrow, shift left] & \\ \exists \text{ partial derivatives of } f \text{ at } \underline{x}_0 \arrow[ur, Rightarrow, shift right, "/"'] \arrow[rr, Leftrightarrow, "/"] & & f \text{ is continuous at } \underline{x}_0 \arrow[ul, Rightarrow, shift left, "/"] \end{tikzcd}$$

CASE OF $f: A \subseteq \mathbb{R}^2 \to \mathbb{R}$

consider an interior point $\underline{x}_0$ of A

f is differentiable at $\underline{x}_0$ if:

$\forall \underline{x} = \underline{x}_0 + \underline{h} \in U(\underline{x}_0)$,

$$f(\underline{x}) = \underbrace{f(\underline{x}_0) + \underline{m}\cdot(\underline{x}-\underline{x}_0)}_{\text{approximating plane}} + \underbrace{o(\|\underline{x}-\underline{x}_0\|)}_{\text{negligible error}} \quad \text{as } \underline{x} \to \underline{x}_0$$

more explicitly:

$$f(x_1,x_2) = f(x_{01},x_{02}) + (m_1,m_2)\begin{pmatrix} x_1 - x_{01} \\ x_2 - x_{02} \end{pmatrix} + o(\|\underline{x}-\underline{x}_0\|)$$

$$= \underbrace{f(x_{01},x_{02}) + m_1(x_1 - x_{01}) + m_2(x_2 - x_{02})}_{\text{plane in 3D-space}} + \underbrace{o(\|\underline{x}-\underline{x}_0\|)}_{\text{negligible part}} \quad \text{as } \underline{x} \to \underline{x}_0$$

GENERAL CASE $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

consider an interior point $\underline{x}_0$ of A

f differentiable at $\underline{x}_0$ if:

$\forall \underline{x} = \underline{x}_0 + \underline{h} \in U(\underline{x}_0)$,

$$f(\underline{x}) = \underbrace{f(\underline{x}_0) + \underline{m} \cdot (\underline{x} - \underline{x}_0)}_{\text{approximating I-degree part}} + \underbrace{o(\|\underline{x} - \underline{x}_0\|)}_{\text{negligible error}} \quad \text{as } x \to x_0$$

THEOREM: DIFFERENTIABILITY → ∃ PARTIAL DERIVATIVES, CONTINUITY

$f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

consider an interior point $\underline{x}_0$ of A

if f is differentiable at $\underline{x}_0$, then:

① f has all partial derivatives

$f'_{x_1}(\underline{x}_0), \dots, f'_{x_n}(\underline{x}_0)$, and, $\underline{m} = \text{grad} f(\underline{x}_0) = (f'_{x_1}(\underline{x}_0), \dots, f'_{x_n}(\underline{x}_0))$

[$n=2 \Rightarrow$ the approximating plane is exactly the tangent plane at $\underline{x}_0$]

② f is continuous at $\underline{x}_0$

$\Rightarrow$

when f is differentiable at $\underline{x}_0$, then:

$\forall \underline{x} = \underline{x}_0 + \underline{h} \in U(\underline{x}_0)$,

$$f(\underline{x}) = \underbrace{f(\underline{x}_0) + \underbrace{\text{grad} f(\underline{x}_0) \cdot (\underline{x} - \underline{x}_0)}_{\text{differentiable } df(\underline{x}_0)}}_{\text{approximating I-degree part}} + \underbrace{o(\|\underline{x} - \underline{x}_0\|)}_{\text{negligible part}} \quad \text{as } x \to x_0$$

more explicitly:

$$df(\underline{x}_0) = \text{grad} f(\underline{x}_0) \cdot (\underline{x} - \underline{x}_0) = [f'_{x_1}(\underline{x}_0), \dots, f'_{x_n}(\underline{x}_0)] \begin{bmatrix} x_1 - x_{0_1} \\ \vdots \\ x_n - x_{0n} \end{bmatrix}$$

$$= f'_{x_1}(\underline{x}_0)(x_1 - x_{0_1}) + \dots + f'_{x_n}(\underline{x}_0)(x_n - x_{0n})$$

$$= f'_{x_1}(\underline{x}_0)\, dx_1 + \dots + f'_{x_n}(\underline{x}_0)\, dx_n$$

EXAMPLE

given $f(x_1, x_2) = 3x_1^2 + 5x_1 x_2^2$, write the differentiable df at $P(1,1)$ $\Rightarrow$

$f'_{x_1} = 6x_1 + x_2^2$, $f'_{x_2} = 10 x_1 x_2$

therefore: $f'_{x_1}(1,1) = 11$, $f'_{x_2}(1,1) = 10$

$df(1,1) = 11(x_1 - 1) + 10(x_2 - 1) = 11\,dx_1 + 10\,dx_2$

THEOREM

$f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

if A is open and f has continuous partial derivatives on A

then f is differentiable on A

$\Rightarrow$

given the set $A \subseteq \mathbb{R}^n$ we will use the notion:

$$C^{(1)}(A) = \left\{ \text{functions } f: A \subseteq \mathbb{R}^n \to \mathbb{R} \text{ that have continuous partial derivatives } \forall \underline{x} \in A \right\}$$

MAXIMA/MINIMA FOR $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$

UNCONSTRAINED OPTIMALISATION

look for max/min

$f(x_1, \dots, x_2)$

CONSTRAINED OPTIMALISATION

look for max/min $f(x_1, \dots, x_2)$

subject to some constraints

UNCONSTRAINED OPTIMALISATION

① necessary condition $\Rightarrow$ Fermat's theorem

② sufficient condition

consider $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$, if:

① $\underline{x}_0$ is an interior point of A

② f has all partial derivatives at $\underline{x}_0$

③ $\underline{x}_0$ is a local maximiser/minimiser for f

then:

$\operatorname{grad} f(\underline{x}_0) = \underline{0}$

that is:

$$\begin{cases} f'_{x_1}(\underline{x}_0) = 0 \\ \vdots \\ f'_{x_n}(\underline{x}_0) = 0 \end{cases} \Rightarrow$$

if $\underline{x}_0$ is such that $\operatorname{grad} f(\underline{x}_0) = \underline{0}$

x_0 is called a stationary point of $f(\underline{x})$

we CANNOT apply it when

① $\underline{x}_0$ is a boundary point of A

② $\underline{x}_0$ is a point at which the partial derivatives do not exist

③ Fermat's theorem is not invertible $\Rightarrow$ if at a point $\underline{x}_0$ we know that

$\text{grad} f(\underline{x}_0) = \underline{0}$ (the tangent plane is horizontal)

GENERAL PROCEDURE TO FIND UNCONSTRAINED LOCAL MAXIMISERS / MINIMISERS

① find exceptional points ->
  - boundary points
  - points at which $f'_{x_1}, \dots, f'_{x_n}$ do not all $\exists$

② Find all other points ->
  - necessary condition (Fermat's theorem)
  - sufficient condition

it may be that a point which passes the necessary condition gives neither a local maximum nor a local minimum

THEOREM: NECESSARY AND SUFFICIENT MAXIMUM (MINIMUM) CONDITION FOR A CONCAVE (CONVEX) AND DIFFERENTIABLE FUNCTION

$f: A \subseteq \mathbb{R}^n \to \mathbb{R}$, if:

① A is a convex set in $\mathbb{R}^n$

② f is concave (convex) in A

③ f is differentiable in A

④ $\underline{x}_0$ interior point of A

then:

$\text{grad} f(\underline{x}_0) = \underline{0} \iff \underline{x}_0$ is a local maximiser (minimiser) $\iff \underline{x}_0$ is a global maximiser (minimiser)

theorems and proofs for n-variable differential calculus

FERMAT'S THEOREM IN R^n

let $f: A \subseteq R^n \to R$ be defined on a set A in R^n and c a subset of A

suppose f is differentiable at an interior point $\hat{x}$ of c

if $\hat{x}$ is a local extremal point (maximizer/minimizer) of f on c then $\nabla f(\hat{x})=0$

PROOF

if $\hat{x}$ is a local extremum of f on c there exists a neighborhood $U \subseteq c$ of $\hat{x}$ such that

$f(\hat{x}) \geq f(x)$ (maximizer) or $f(\hat{x}) \leq f(x)$ (minimizer) for all $x \in U$

by the differentiability of f at $\hat{x}$, for any direction v, we have:

$f(\hat{x}+tv) = f(\hat{x}) + t\nabla f(\hat{x}) \cdot v + o(t)$ where $o(t)$ vanishes faster than t as $t \to 0$

at a local extremum $f(\hat{x}+tv) \leq f(\hat{x})$ (maximizer) or $f(\hat{x}+tv) \geq f(\hat{x})$ (minimizer) for small t

dividing and taking the limit as $t \to 0$, the directional derivative in any direction v is 0:

$\nabla f(\hat{x}) \cdot v = 0 \quad \forall v$

since the gradient is 0 in all directions, $\nabla f(\hat{x}) = 0$

LINEAR ALGEBRA

VECTORS

SUBSPACE = $V \subseteq \mathbb{R}^n$, $V \neq \emptyset$

V is a vector subspace of $\mathbb{R}^n$ if:

① $\forall \underline{x}, \underline{y} \in V \Rightarrow \underline{x} + \underline{y} \in V$

② $\forall \underline{x} \in V, \forall \alpha \in \mathbb{R} \Rightarrow \alpha \underline{x} \in V$ (V is closed with respect to $\underline{x}+\underline{y}$, $\alpha \underline{x}$)

THEOREM

$V \subseteq \mathbb{R}^n$, $V \neq \emptyset$

V is a vector subspace of $\mathbb{R}^n \Leftrightarrow \forall \underline{x}, \underline{y} \in V, \forall \alpha, \beta \in \mathbb{R}, \alpha \underline{x} + \beta \underline{y} \in V$

(V is closed with respect to linear combinations $\alpha \underline{x} + \beta \underline{y}$)

THEOREM: SUBSPACES IN $\mathbb{R}^2$

$A \subseteq \mathbb{R}^2$, $A \neq \emptyset$

A is a subspace of $\mathbb{R}^2 \Leftrightarrow$ ① $A = \mathbb{R}^2$ or

② $A = \{\underline{0}\}$ or

③ A is a straight line passing through $\underline{0}$, that is:

$A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : ax + by = 0 \right\}$

THEOREM: SUBSPACES IN $\mathbb{R}^n$

$A \subseteq \mathbb{R}^n$, $A \neq \emptyset$

A is a subspace of $\mathbb{R}^n \Leftrightarrow$ ① $A = \mathbb{R}^n$ or

② $A = \{\underline{0}\}$ or

③ $A = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \right.$: all the relations involving the components $x_1, \dots, x_n$ can be written as I·degree equalities with no constants $\neq 0$

necessary condition = if A is a subspace of $\mathbb{R}^n \to A$ must contain the origin $\underline{0}$

THEOREM: INTERSECTION OF SUBSPACES

the intersection of any collection of vector subspaces of $\mathbb{R}^n$ is a vector subspace of $\mathbb{R}^n$

PROOF

suppose $\{V_i\}$ is a collection of vector subspaces of $\mathbb{R}^n$

since $\underline{0} \in V_i \ \forall i$, $\underline{0} \in \cap V_i$; therefore $\cap V_i \neq \emptyset$

consider $\underline{x}, \underline{y} \in \cap V_i$ and $\alpha, \beta \in \mathbb{R}$
since $\underline{x}, \underline{y} \in \cap V_i$, we have that $\underline{x}, \underline{y} \in V_i \ \forall i$
since every V_i is a vector subspace of $\mathbb{R}^n$
this implies that $\alpha \underline{x} + \beta \underline{y} \in V_i \ \forall i$
therefore $\alpha \underline{x} + \beta \underline{y} \in \cap V_i$, so $\cap V_i$ is a vector subspace of $\mathbb{R}^n$

LINEARLY DEPENDENT / INDEPENDENT VECTORS

=

$\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$ are linearly dependent
if they are not linearly independent
that is if $\exists \ \alpha_1, \dots, \alpha_m \in \mathbb{R}$ such that:
$\alpha \underline{x}_1 + \dots + \alpha_m \underline{x}_m = \underline{0}$ and $\alpha_1, \dots, \alpha_m$
are not all equal to zero

=

$\underline{x}_1, \dots, x_m \in \mathbb{R}^n$ are linearly independent
if whenever $\alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m = \underline{0} \Rightarrow$
$\alpha_1 = \dots = \alpha_m = 0$

- if $\underline{x}_1, \dots, \underline{x}_m$ contains $\underline{0}$, then $\underline{x}_1, \dots, \underline{x}_m$ are linearly dependent $\Rightarrow$
  $0 \underline{x}_1 + \dots + 0 \underline{x}_{m-1} + k \underline{0} = \underline{0}$ also when $k \neq 0$
- one single vector $\underline{x}$ $(\forall m \geq 1) \to$
  - $\underline{x} \neq \underline{0}$ = linearly independent
  - $\underline{x} = \underline{0}$ = linearly dependent

THEOREM $(m \geq 2)$

$\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$, with $m \geq 2$, then:
① $\underline{x}_1, \dots, \underline{x}_m$ linearly independent = none of them can be written as a linear combination of the others
② $\underline{x}_1, \dots, \underline{x}_m$ linearly dependent = at least one of them can be written as a linear combination of the others

$(m = 2)$
① linearly independent = if none of the vectors is a multiple of the other
② linearly dependent = at least one of them is a multiple of the other

FUNDAMENTAL VECTORS OF $\mathbb{R}^3$ $\Rightarrow$ linearly independent

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{e}_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$$

FUNDAMENTAL VECTORS OF $\mathbb{R}^n$ => linearly independent

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \underline{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

remark

$\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$, then:

① if $m > n$ they are linearly dependent

② if $m \leq n$ we don't know (depends on the vectors)

SPAN S

- consider a set S of vectors in $\mathbb{R}^n$
- span S = intersection $\cap V_i$ of all vector subspaces V_i of $\mathbb{R}^n$ which contains S
- span $S = \cap V_i$ is the smallest vector subspace of $\mathbb{R}^n$ which contains S
- it is the vector subspace which is spanned generated by the set S -> smallest enlargement of S with the property of being a vector space

remark

take $\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$, then:

① if $m < n$ it is not possible that they span all $\mathbb{R}^n$

② if $m \geq n$ we don't know (depends on the vectors)

BASIS OF $\mathbb{R}^n$

take $\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$

they are a basis of $\mathbb{R}^n$ if:

① they are linearly independent (none of them is superflous)

② they span all $\mathbb{R}^n$ (enough to span $\mathbb{R}^n$)

if we call $S = \{\underline{x}_1, \dots, \underline{x}_m\}$, span $S = \mathbb{R}^n$

remark (I)

① $m > n$ => they can't be linearly independent so they are not a basis of $\mathbb{R}^n$

② $m < n$ => they can't span all $\mathbb{R}^n$, so they are not a basis of $\mathbb{R}^n$

- the number of vectors in a basis of $\mathbb{R}^n$ must be $m = n$

remark (II)

- suppose to take exactly n vectors in $\mathbb{R}^n$
- to check they are a basis of $\mathbb{R}^n$, two facts must be checked ->

① they are linearly dependent

② they span all R^n

- this property [$\underline{x}_1, \dots, \underline{x}_n \in R^n$ are linearly independent $\Leftrightarrow$ they span all R^n] holds $\rightarrow$ if the vectors are the correct number only the linear independance needs to be checked
- Ex. the fundamental basis of R^n $\rightarrow$

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \underline{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in R^n$$

the coefficients such that $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ can be written as a linear combination $\underline{x} = \alpha_1 \underline{e}_1 + \dots + \alpha_n \underline{e}_n$ are exactly the components $x_1, \dots, x_n$

DIMENSION OF R^n

- the dimension of R^n is n
- it gives the number of vectors in any basis of R^n and it corresponds to the geometrical notion of dimension

remark

- any two vectors $\underline{e}_j, \underline{e}_k \in R^n$ with $j \neq k$ are such that $\underline{e}_j \perp \underline{e}_k$ ($\underline{e}_j \cdot \underline{e}_k = 0$ $\Rightarrow$ they are orthogonal)
- all vectors $\underline{e}_j \in R^n$ are such that : $\|\underline{e}_j\| = 1$ ($\underline{e}_j \cdot \underline{e}_j = 1$ $\Rightarrow$ unitary norm)
- the two properties are collected together and conclude that any two vectors $\underline{e}_j, \underline{e}_k \in R^n$ are orthonormal
- the fundamental basis $\underline{e}_1, \dots, \underline{e}_n$ is an orthonormal basis
- in those situations in which we can't use exactly the fundamental basis we will try to use a basis contributed by orthogonal vectors or by orthonormal vectors (better)

THEOREM: ORTHOGONAL VECTORS ARE LINEARLY INDEPENDENT

consider k vectors $\underline{x}_1, \dots, \underline{x}_k \in R^n$ with $2 \leq k \leq n$

if these vectors are 2-by-2 orthogonal and they do not include the null vector (if they are orthonormal), then they are linearly independent

THEOREM

consider an orthonormal basis $\underline{x}_1, \dots, \underline{x}_n$ in R^n

then $\forall \underline{v} \in R^n$ it is:

$$v = \alpha_1 \underline{x}_1 + \dots + \alpha_n \underline{x}_n \text{ with } \alpha_1 = \underline{v} \cdot \underline{x}_1, \dots, \alpha_n = \underline{v} \cdot \underline{x}_n$$

if $\underline{x}_1, \dots, \underline{x}_n$ is an orthonormal basis of $\mathbb{R}^n$ there is a quick way to write the coefficients $\alpha_1, \dots, \alpha_n$ (exactly as it is for the basis $\underline{e}_1, \dots, \underline{e}_n$)

$S = \{\underline{x}_1, \dots, \underline{x}_m\}$, SPAN $S \subset \mathbb{R}^n$

=> span S is a subspace of $\mathbb{R}^n$ smaller than $\mathbb{R}^n$

THEOREM

suppose V is a subspace of $\mathbb{R}^n$, then $\exists$ a set $S = \{\underline{x}_1, \dots, \underline{x}_m\}$ such that:

V = span S

=> every subspace of $\mathbb{R}^n$ can be seen as the span S of some finite set S

BASIS OF A SUBSPACE

take a subspace $V \subseteq \mathbb{R}^n$

take $\underline{x}_1, \dots, \underline{x}_m \in V$

they are a basis of V if:

① they are linearly independent

② they span all V

=> if we call $S = \{\underline{x}_1, \dots, \underline{x}_m\}$, span S = V

THEOREM: PROPERTY OF UNIQUE WRITING

take a vector subspace $V \subseteq \mathbb{R}^n$

take $\underline{x}_1, \dots, \underline{x}_m \in V$

they are a basis of V <=> $\forall \underline{x} \in V$, $\underline{x}$ can be written in a unique way as a linear combination of $\underline{x}_1, \dots, \underline{x}_m$

THEOREM

take a vector subspace $V \subseteq \mathbb{R}^n$

any two bases of V will contain the same number k of vectors

it will be:

$k = n \iff V = \mathbb{R}^n$

$k < n \iff V \subset \mathbb{R}^n$

DIMENSION OF A SUBSPACE

- take a vector subspace $V \subseteq \mathbb{R}^n$
- dimension of V = the number of vectors in any basis of V
- applied to $V = \mathbb{R}^n$ => $\dim \mathbb{R}^n = n$

- applied to $V \subset R^n \Rightarrow \dim V = k < n$

remark (I)

- consider a subspace $V \subseteq R^n$ with dimension $k \leq n$
- take k vectors $\underline{x}_1, \dots, \underline{x}_k \in V$ (same number as $\dim V$)
- in order to guarantee they are a basis of V only one of the two properties needs to be checked
- $\underline{x}_1, \dots, \underline{x}_k$ are linearly independent $\Leftrightarrow$ they span all the space V

remark (II)

if $V = \{\underline{0}\}$, then:

① basis of V: empty set $\emptyset$

② dimension of V: 0

THEOREM: SUBSPACES IN R^2

take $A \subseteq R^2$, $A \neq \emptyset$

A is a subspace of $R^2 \Leftrightarrow$ ① $A = R^2$ $\dim A = 2$ or

② $A = \{\underline{0}\}$ $\dim A = 0$ or

③ A is a straight line passing through $\underline{0}$:

$A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in R^2 : ax + by = 0 \right\}$ $\dim A = 1$

THEOREM: SUBSPACES IN R^3

take $A \subseteq R^3$, $A \neq \emptyset$

A is a subspace of $R^3 \Leftrightarrow$ ① $A = R^3$ $\dim A = 3$ or

② $A = \{\underline{0}\}$ $\dim A = 0$ or

③ A is a plane passing through $\underline{0}$: $\rightarrow$ basis has 2 vectors

$A = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : ax + by + cz = 0 \right\}$ $\dim A = 2$ or

④ A is a straight line passing through $\underline{0}$:

$A = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in R^3 : a_1 x + b_1 y + c_1 z = 0, a_2 x + b_2 y + c_2 z = 0 \right\}$

$\dim A = 1$ -> basis has 1 vector

MATRICES

$$A = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & 1 \end{bmatrix} \rightarrow \text{row} \quad \Rightarrow \quad 2 \text{ rows} + 3 \text{ columns} = 2 \times 3 \text{ matrix}$$

column

DEF. table containing numbers organised in rows and columns

more generally:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

$A = [a_{ij}]$ $\quad i = 1, \dots, m$, $\; j = 1, \dots, n$ $\Rightarrow$ its elements have a double index which corresponds to double orders (by row / by column)

also:

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 4 & 5 & 0 \end{bmatrix} \Rightarrow \text{collection of column vectors}$$

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 4 & 5 & 0 \end{bmatrix} \Rightarrow \text{collection of row vectors}$$

if $m = n$:

$$A = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 2 & 0 \\ 1 & 3 & 5 \end{bmatrix} \Rightarrow \text{square matrix of order } n$$

↳ main diagonal

SETS OF MATRICES

$\mathcal{M}(m,n) = \{\text{all } m \times n \text{ matrices}\}$

$\mathcal{M}(n) = \{\text{all square matrices of order } n\}$

OPERATIONS IN $\mathcal{M}(m,n)$

① $A + B \Rightarrow$

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \Rightarrow A + B = \begin{bmatrix} 7 & -4 & 3 \\ 0 & 6 & 6 \end{bmatrix}$$

DEF. $\forall A = [a_{ij}],\ B = [b_{ij}] \in \mathcal{R}(m,n)$, $C = A + B \in \mathcal{R}(m,n)$ and is given by:
$c_{ij} = a_{ij} + b_{ij}$, $\forall i = 1, \dots, m;\ \forall j = 1, \dots, n$

PROPER.

- commutative = $\forall A, B \in \mathcal{R}(m,n)$, $A + B = B + A$
- associative = $\forall A, B, C \in \mathcal{R}(m,n)$, $(A+B)+C = A + (B+C) = A + B + C$
- $\exists$ null matrix = $\exists\, O \in \mathcal{R}(m,n)$, $O = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$, such that $\forall A \in \mathcal{R}(m,n)$, $A + O = A$
- $\forall$ matrix $\exists$ opposite = $\forall A = [a_{ij}] \in \mathcal{R}(m,n)$ $\exists\, B = -A = [-a_{ij}] \in \mathcal{R}(m,n)$ such that $A + (-A) = O$

② αA =>

$$A = \begin{bmatrix} 3 & -1 \\ 4 & 0 \\ 2 & 5 \end{bmatrix},\quad \alpha = 2 \quad \Rightarrow \quad \alpha A = \begin{bmatrix} 6 & -2 \\ 8 & 0 \\ 4 & 10 \end{bmatrix}$$

DEF. $\forall A = [a_{ij}] \in \mathcal{R}(m,n)$, $\forall \alpha \in R$, $B = \alpha A \in \mathcal{R}(m,n)$ and is given by
$b_{ij} = \alpha a_{ij}$, $\forall i = 1, \dots, m;\ \forall j = 1, \dots, n$

PROPER.

- distributive = $\forall \alpha, \beta \in R$, $\forall A \in \mathcal{R}(m,n)$, $(\alpha + \beta) A = \alpha A + \beta A$
- distributive = $\forall \alpha \in R$, $\forall A, B \in \mathcal{R}(m,n)$, $\alpha (A + B) = \alpha A + \alpha B$
- associative = $\forall \alpha, \beta \in R$, $\forall A \in \mathcal{R}(m,n)$, $(\alpha\beta) A = \alpha(\beta A) = \alpha \beta A$
- number 1 is a unit for matrices = $\forall A \in \mathcal{R}(m,n)$, $1A = A$

③ $A \cdot B$ (row-column product) =>

DEF. $\forall A \in \mathcal{R}(m,n)$, $\forall B \in \mathcal{R}(n,p)$, $C = A \cdot B \in \mathcal{R}(m,p)$ and is given by:
c_{ik} = (i-th row of A) $\cdot$ (k-th column of B)

$$= (a_{i1}\ a_{i2} \cdots a_{in}) \cdot \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}$$

$$= a_{i1} + b_{1k} + a_{i2} b_{2k} + \dots + a_{in} b_{nk}$$

$$= \sum_{j=1}^{n} a_{ij} b_{jk}, \quad \forall i = 1, \dots, m;\ \forall k = 1, \dots, n$$

DEFECTS

- $A \cdot B$ is not an internal operation =
  $A \in \mathcal{M}(m,n),\ B \in \mathcal{M}(n,p) \Rightarrow C = A \cdot B \in \mathcal{M}(m,p)$ [but it becomes internal with square matrices $\rightarrow A \in \mathcal{M}(n), B \in \mathcal{M}(n) \Rightarrow C = A \cdot B \in \mathcal{M}(n)$]
- $A \cdot B$ is not commutative =
  - you can define $A \cdot B$ but not $B \cdot A$
  - you can define both but they are matrices of different kinds
  - you can define both and they are both of the same kind (only square matrices) but still usually $AB \neq BA$

PROPER.

- associative = $\forall A \in \mathcal{M}(m,n), \forall B \in \mathcal{M}(n,p), \forall C \in \mathcal{M}(p,q), (A \cdot B) \cdot C = A \cdot (B \cdot C)$
  $= A \cdot B \cdot C$
- associative = $\forall A \in \mathcal{M}(m,n), \forall B \in \mathcal{M}(n,p), \forall \alpha \in \mathbb{R}, (\alpha A) \cdot B = \alpha (A \cdot B) = \alpha AB$
- distributive = $\forall A \in \mathcal{M}(m,n), \forall B, C \in \mathcal{M}(n,p), A(B+C) = AB + AC$
- $\exists$ 1-sided identity matrices = $\forall m,n$ $\exists$ 1-sided identity matrices :
  $I_m = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathcal{M}(m)$,
  $I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathcal{M}(n)$,
  such that $\forall A \in \mathcal{M}(m,n)$:
  $I_m A = A \in \mathcal{M}(m,n)$
  $A I_n = A \in \mathcal{M}(m,n)$
- $\exists$ identity matrix = $\forall n$ $\exists$ identity matrix of order n, $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in \mathcal{M}(n)$,
  such that $\forall A \in \mathcal{M}(n)$, $A I_n = I_n A = A$

SQUARE MATRICES

$A^2 = A \cdot A \rightarrow A^n = \underbrace{A \ldots\ldots A}_{}$]-> n times , $A^1 = A$, $A^0 = I_n$ [$a^0 = 1$]

④ A^T A transpose (defined $\forall$ matrix) $\Rightarrow$

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 4 & 7 & 0 \end{pmatrix} \rightarrow A^T = \begin{pmatrix} 3 & 4 \\ -1 & 7 \\ 2 & 0 \end{pmatrix}$$

PROPER.

- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(\alpha A)^T = \alpha A^T$
- $(AB)^T = B^T A^T$
- if A is a square matrix it is symmetric if: $a_{ij} = a_{ji}$ $\forall i,j$, or if $A^T = A$

⑤ PARTICULAR SQUARE MATRICES =>

- upper triangular = $A = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$ $\forall i > j, a_{ij} = 0$

- lower triangual = $A = \begin{bmatrix} 3 & 0 & 0 \\ -1 & 2 & 0 \\ 4 & 1 & 0 \end{bmatrix}$ $\forall i < j, a_{ij} = 0$

- diagonal = $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $\forall i \neq j, a_{ij} = 0$

- scalar = $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ diagonal + $a_{ii} = a$

VECTOR SPACES : EXAMPLES

① the set $P[x]$ of all polynomials in one variable x (operations: $p(x) + q(x), \alpha p(x)$)

② the set S of all convergent sequences $a_n : \mathbb{N} \to \mathbb{R}$ (operations: $a_n + b_n, \alpha a_n$)

③ the set $C(I) / C^{(1)}(I) / \dots / C^{(n)}(I)$ of all continuous / differentiable with continuous derivative / $\cdots$ / n-times differentiable with continuous n^{th} derivative functions on an interval $I \subseteq \mathbb{R}$ (operations: $f + g, \alpha f$)

④ the set $C(A) / C^{(1)}(A)$ of all continuous / differentiable with continuous partial derivatives functions on a convex set $A \subseteq \mathbb{R}^n$ (operations: $f + g, \alpha f$)

LINEAR FUNCTIONS

a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be a linear function if:

① $\forall \underline{x}, \underline{y} \in \mathbb{R}^n, f(\underline{x} + \underline{y}) = f(\underline{x}) + f(\underline{y})$ [additivity]

② $\forall \underline{x} \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}, f(\alpha \underline{x}) = \alpha f(\underline{x})$ [homogeneity]

THEOREM

a function $f : \mathbb{R}^n \to \mathbb{R}$ is a linear function $\iff$ $\forall \underline{x}, \underline{y} \in \mathbb{R}^n, \forall \alpha, \beta \in \mathbb{R},$

$$f(\alpha \underline{x} + \beta \underline{y}) = \alpha f(\underline{x}) + \beta f(\underline{y})$$

THEOREM

if f is a linear function $f : \mathbb{R}^n \to \mathbb{R}$, then:

① $f(\underline{0}) = 0$

② $\forall \underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n, \forall \alpha_1, \dots, \alpha_m \in \mathbb{R}, f(\alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m) = \alpha_1 f(\underline{x}_1) + \dots + \alpha_m f(\underline{x}_m)$

remark

- linear function $\iff f(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n$ of the I degree and with no constants $\neq 0$
- coefficients $a_1, \dots, a_n$ are exactly given by $f(\underline{e}_1), \dots, f(\underline{e}_n)$
- f affine (concave + convex) $\iff f(\underline{x}) = \underline{a} \cdot \underline{x} + b = a_1 x_1 + \dots + a_n x_n + b$

THEOREM: CONTINUITY OF LINEAR FUNCTIONS

every linear function $f: \mathbb{R}^n \to \mathbb{R}$ is continuous

that is: $\lim_{\underline{x} \to \underline{x}_0} f(\underline{x}) = f(\underline{x}_0) \quad \forall \underline{x}_0 \in \mathbb{R}$

numbers

INCREASING FUNCTIONS $f: \mathbb{R}^n \to \mathbb{R}$

- increasing = $\underline{x} \geq \underline{y} \Rightarrow f(\underline{x}) \geq f(\underline{y}), \forall \underline{x}, \underline{y} \in \mathbb{R}^n$
- strongly increasing = increasing + $\underline{x} \gg \underline{y} \Rightarrow f(\underline{x}) > f(\underline{y})$
- strictly increasing = $\underline{x} > \underline{y} \Rightarrow f(\underline{x}) > f(\underline{y}), \forall \underline{x}, \underline{y} \in \mathbb{R}^n$

$\Rightarrow$ strictly increasing $\underset{\not\Leftarrow}{\Rightarrow}$ strongly increasing $\underset{\not\Leftarrow}{\Rightarrow}$ increasing

POSITIVE FUNCTIONS $f: \mathbb{R}^n \to \mathbb{R}$

- positive = $\underline{x} \geq \underline{0} \Rightarrow f(\underline{x}) \geq 0, \forall \underline{x} \in \mathbb{R}^n$
- strongly positive = positive + $\underline{x} \gg \underline{0} \Rightarrow f(\underline{x}) > 0, \forall \underline{x} \in \mathbb{R}^n$
- strictly positive = $\underline{x} > \underline{0} \Rightarrow f(\underline{x}) > 0, \forall \underline{x} \in \mathbb{R}^n$

$\Rightarrow$ strictly positive $\underset{\not\Leftarrow}{\Rightarrow}$ strongly positive $\underset{\not\Leftarrow}{\Rightarrow}$ positive

THEOREM

consider a linear function $f: \mathbb{R}^n \to \mathbb{R}$, then:

① f is increasing $\iff f$ is positive

② f is strongly increasing $\iff f$ is strongly positive

③ f is strictly increasing $\iff f$ is strictly positive

THEOREM: RIESZ-MARKOV

consider a function $f: \mathbb{R}^n \to \mathbb{R}$

f is linear and increasing $\iff \exists$ a unique positive vector $\underline{a} \in \mathbb{R}^n$ such that:

$f(\underline{x}) = \underline{a} \cdot \underline{x}, \forall \underline{x} \in \mathbb{R}^n$

in particular:

① f strongly increasing (strongly positive) $\iff \underline{a}$ strictly positive

② f strictly increasing (strictly positive) <=> $\underline{a}$ strongly positive

FUNCTIONS $f: \mathbb{R}^n \to \mathbb{R}^m$

ex. $f: \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \longrightarrow f(\underline{x}) = \begin{pmatrix} x_1 + 3x_3 \\ x_1 + x_2 + x_3 \end{pmatrix} \in \mathbb{R}^2$

a function $f: \mathbb{R}^n \to \mathbb{R}^m$ is said to be a linear function (linear operator) if:

① $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$, $f(\underline{x} + \underline{y}) = f(\underline{x}) + f(\underline{y})$ => additivity

② $\forall \underline{x} \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$, $f(\alpha \underline{x}) = \alpha f(\underline{x})$ => homogeneity

THEOREM

a function $f: \mathbb{R}^n \to \mathbb{R}^m$ is a linear function <=> $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$, $\forall \alpha, \beta \in \mathbb{R}$,

$$f(\alpha \underline{x} + \beta \underline{y}) = \alpha f(\underline{x}) + \beta f(\underline{y})$$

THEOREM

if f is a linear function $f: \mathbb{R}^n \to \mathbb{R}^m$, then:

① $f(\underline{0}) = \underline{0}$

② $\forall \underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$, $\forall \alpha_1, \dots, \alpha_m \in \mathbb{R}$, $f(\alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m) = \alpha_1 f(\underline{x}_1) + \dots + \alpha_m f(\underline{x}_m)$

THEOREM: CONTINUITY OF LINEAR OPERATIONS

every linear operator $f: \mathbb{R}^n \to \mathbb{R}^m$ is continuous

that is: $\lim_{\underline{x} \to \underline{x}_0} f(\underline{x}) = f(\underline{x}_0)$, $\forall \underline{x}_0 \in \mathbb{R}^n$

IMAGE, KERNEL

take a linear function $f: \mathbb{R}^n \to \mathbb{R}^m$

then we can always associate to f two different sets:

① image / image space => $f(\mathbb{R}^n) = \mathrm{Im} f$

$= \{\underline{y} \in \mathbb{R}^m : \exists\, \underline{x} \in \mathbb{R}^n \text{ with } \underline{y} = f(\underline{x})\} \subseteq \mathbb{R}^m$

② kernel => $\ker f = \{\underline{x} \in \mathbb{R}^n : f(\underline{x}) = \underline{0}\} \subseteq \mathbb{R}^n$

DEFINITION => consider $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: rank f = dim Imf

THEOREM

Kerf is always a vector subspace of $\mathbb{R}^n$

DEFINITION => consider $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: nullity of f = dim kerf

THEOREM: RANK AND NULLITY

take $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: $\dim \ker f = n - \dim \operatorname{Im} f$

THEOREM: SURJECTIVITY CONDITIONS

take $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: f surjective $\Leftrightarrow \operatorname{Im} f = \mathbb{R}^m \Leftrightarrow \operatorname{rank} f = \dim \operatorname{Im} f = m$

THEOREM: INJECTIVITY CONDITIONS

take $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: f injective $\Leftrightarrow \ker f = \{\underline{0}\} \Leftrightarrow$ nullity of $f =$ $\dim \ker f = 0 \Leftrightarrow \operatorname{rank} f = \dim \operatorname{Im} f = n$

THEOREM: BIJECTIVITY CONDITIONS

take $f: \mathbb{R}^n \to \mathbb{R}^m$ linear, then: f bijective $\Leftrightarrow$ ① $m = n$
② $\operatorname{rank} f = m = n$

DETERMINANT OF A SQUARE MATRIX

particular cases ->

① A has order 1 $\Rightarrow A = (a) \Rightarrow \det A = a$

② A has order 2 $\Rightarrow A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

ex. ① $A = \begin{pmatrix} 2 & 4 \\ 1 & 5 \end{pmatrix} \to \det A = 2 \cdot 5 - 4 \cdot 1 = 6 \neq 0 \Rightarrow$ linearly independent

② $A = \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix} \to \det A = 2 \cdot 3 - 6 \cdot 1 = 0 \Rightarrow$ linearly dependent

general case $\Rightarrow \det A = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$

③ A has order 3 $\Rightarrow A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

ex. $A = \begin{pmatrix} 2 & 3 & 5 \\ -1 & 4 & -1 \\ 0 & 1 & -1 \end{pmatrix} \to \det A = 2 \cdot 4 \cdot (-1) + 3 \cdot (-1) \cdot 0 + 5 \cdot (-1) \cdot 1 = -14 \neq 0$

$$\hookrightarrow \left(\begin{array}{ccc} 2 & 3 & 5 \\ -1 & 4 & -1 \\ 0 & 1 & -1 \end{array}\right) \begin{array}{cc} 2 & 3 \\ -1 & 4 \\ 0 & 1 \end{array}$$

SUBMATRICES

① submatrix of a matrix = given a matrix $A \in \mathcal{M}(m,n)$ we call submatrix of A any matrix which is obtained from A canceling out entire rows and/or colums

ex. $A = \begin{pmatrix} 2 & -1 & \cancel{5} \\ \cancel{3} & \cancel{2} & \cancel{4} \end{pmatrix}$ ✓ $A = \begin{pmatrix} 2 & -1 & \cancel{5} \\ 3 & \cancel{2} & 4 \end{pmatrix}$ ×

② complement submatrix = given a square matrix: $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ we call complement submatrix of the element a_{ij} the square submatrix of A which we obtain by canceling out the row i and the column j on which the element a_{ij} stands and we denote it as A_{ij}

LAPLACE'S DEFINITION

take a square matrix A of order n, then:

① if $n=1$, that is $A = (a)$, then $\det A = a$

② if $n>1$ →

ex. $A = \begin{pmatrix} 3 & 0 & 0 & -2 \\ 5 & 1 & 2 & 4 \\ 2 & -1 & 3 & 2 \\ 1 & 2 & 0 & 5 \end{pmatrix}$ we concentrate on the first row and define:

$$\det A = a_{11}(-1)^{1+1}\det A_{11} + a_{12}(-1)^{1+2}\det A_{12} + a_{13}(-1)^{1+3}\det A_{13} + a_{14}(-1)^{1+4}\det A_{14} + \ldots + a_{1n}(-1)^{1+n}\det A_{1n}$$

DEF. MINOR = consider a matrix A; consider a square submatrix B of A; then $\det B$ is called a minor of the matrix A

DEF. COMPLEMENT MINOR = consider a square matrix A; we call complement minor of the element a_{ij} the determinant $\det A_{ij}$ of the complement submatrix A_{ij} of the element a_{ij}

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

DEF. COFACTOR (algebric complement) = $(-1)^{i+j}\det A_{ij}$ ⇒ the complement minor $\det A_{ij}$ is multiplied by $+1/-1$ according to whether the sum of the indices $i+j$ is even/odd

⇒

① $n=1$: $A = (a)$, $\det A = a$

② $n>1$: $\det A = \sum_{j=i}^{n} a_{1j}(-1)^{1+j}\det A_{1j}$ / $\det A = a_{11}$ (cofactor of a_{11}) + a_{12} (cofactor of a_{12}) +…+ a_{1n}(cofactor of a_{1n})

ex. $A = \begin{pmatrix} 3 & 0 & 0 & -2 \\ 5 & 1 & 2 & 4 \\ 2 & -1 & 3 & 2 \\ 1 & 2 & 0 & 5 \end{pmatrix} \rightarrow \det A = 3 \cdot \det \begin{pmatrix} 1 & 2 & 4 \\ -1 & 3 & 2 \\ 2 & 0 & 5 \end{pmatrix} - 2(-1) \det \begin{pmatrix} 5 & 1 & 2 \\ 2 & -1 & 3 \\ 1 & 2 & 0 \end{pmatrix} =$

$= 3 \cdot 9 + 2 \cdot (-17) = -7 \neq 0$

FIRST LAPLACE'S THEOREM

① $\forall i, \sum_{j=1}^{n} a_{ij} \det A_{ij} (-1)^{i+j} = \det A$ [we can use any row i]

② $\forall j, \sum_{i=1}^{n} a_{ij} \det A_{ij} (-1)^{i+j} = \det A$ [we can use any column j]

$\Rightarrow$ in exercises use the one with more zeros in it

THEOREM

consider $\underline{x}_1, \dots, \underline{x}_n \in \mathbb{R}^n$, build the square matrix with $\underline{x}_1, \dots, \underline{x}_n$ as columns $A = (\underline{x}_1, \dots, \underline{x}_n)$, then:

- $\underline{x}_1, \dots, \underline{x}_n$ are linearly independent $\Leftrightarrow \det A \neq 0$
- $\underline{x}_1, \dots, \underline{x}_n$ are linearly dependent $\Leftrightarrow \det A = 0$

BEHAVIOUR OF det A WITH RESPECT TO OPERATIONS

① $\det(A^T) = \det A$

② $\det(A \cdot B) = \det A \cdot \det B$

③ $\det(A+B) \neq \det A + \det B$

④ (if A has order n) $\det(\alpha A) = \alpha^n \det A$

INVERSE OF A SQUARE MATRIX

consider a square matrix of order n

we say that:

A is invertible or A has an inverse B

if $\exists$ a square matrix B of order n such that

$AB = BA = I_n$

THEOREM: EXISTANCE AND UNIQUENESS

consider a square matrix A of order n

① A has an inverse $B \Leftrightarrow \det A \neq 0$

② if A has an inverse B then B is unique

$\Rightarrow$ B inverse of A $(B = A^{-1})$

ADJOINT MATRIX

take a square matrix A

its adjoint matrix adj A is the matrix we obtain with this procedure:

$A \to A^* =$ matrix of the cofactors of A $\to$ $\text{adj}\, A = (A^*)^T$

THEOREM

suppose A is a square of order n

suppose $\det A \neq 0$, then:

$$A^{-1} = \frac{1}{\det A} \,\text{adj}\, A$$

PROPERTIES

① $\forall A$ square, with $\det A \neq 0 \to (A^{-1})^{-1} = A$

② $\forall A$ square, with $\det A \neq 0 \to (A^T)^{-1} = (A^{-1})^T$

③ $\forall A, B$ square, with $\det A \neq 0, \det B \neq 0 \to (AB)^{-1} = B^{-1}A^{-1}$

④ $\forall A, B$ square, with $\det A \neq 0, \det B \neq 0 \to (A+B)^{-1} = A^{-1} + B^{-1}$

⑤ $\forall A$ square, with $\det A \neq 0$, $\forall \alpha \neq 0 \to (\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$

⑥ $\forall A$ square, with $\det A \neq 0 \to \det(A^{-1}) = \frac{1}{\det A}$

THEOREM

consider a linear function $f: \mathbb{R}^n \to \mathbb{R}^n$

call A its representation matrix which is a square matrix of order n, then:

① f invertible $\iff$ A invertible

② when f is invertible, the representation matrix of $f^{-1}: \mathbb{R}^n \to \mathbb{R}^n$ is the square matrix A^{-1}

RANK OF A MATRIX

RANK BY COLUMNS

given a m×n matrix A

the rank by columns of the matrix A is the maximum number of linearly independent columns in A

RANK BY ROWS

give a m×n matrix A,

the rank by rows of the matrix A is the maximum number of linearly independent rows in A we have:

THEOREM

given a m×n matrix A,
the rank by columns of A is always coincident with the rank by rows of A
therefore, from now on, we will just call it the rank of A: it gives both the maximum number of linearly independent columns and the maximum number of linearly independent rows in A

THEOREM: COMPUTATION OF RANK

consider a m×n matrix A, then:
rank A = k <=> k is the maximum order for which in A we have a non-null minor

KRONECKER'S THEOREM

consider a m×n matrix A
if ∃ a non-null minor of A of order k, and all the minors of A of order k+1 which contains this non-null minor are null, then rank A = k

remark: THEOREM

consider a linear function $f: \mathbb{R}^n \to \mathbb{R}^m$
consider its representation matrix A, then:

① rank f = k <=> rank A = k

② f surjective <=> rank A = m (n. of rows)
f injective <=> rank A = n (n. of columns)
f bijective <=> m = n, rank A = n
[det A ≠ 0]

LINEAR SYSTEM

HOW TO WRITE IT

- explicit writing =
  - one equation -> $a_1x_1 + \ldots + a_nx_n = b$
  - m equations ->

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{cases}$$

[a_{ij} : i = equation index, j = variable index]

- matrix writing = $A\underline{x} = \underline{b}$ =>

$\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ vector of variables

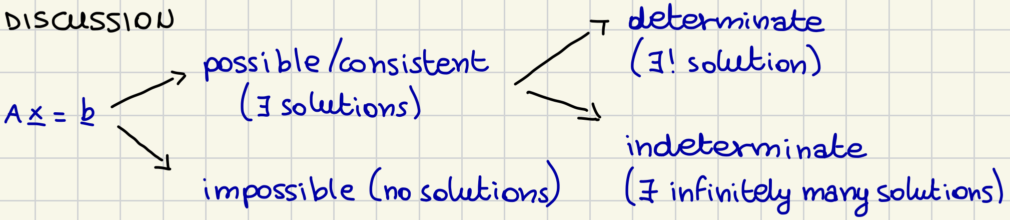
$\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ vector of constant terms $\Rightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ matrix of coefficients

- column writing = $\underline{a}_1 x_1 + \underline{a}_2 x_2 + \ldots + \underline{a}_n x_n = \underline{b}$ =>

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_2 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

DISCUSSION



THEOREM: DISTINCTION BETWEEN DETERMINATE/INDETERMINATE CASES

given a linear system $A\underline{x} = \underline{b}$ with m equations and n variables, then:

① if rank $A <$ rank $(A|\underline{b})$ => impossible

② if rank A = rank $(A|\underline{b}) = n$ => determinate (∃! solution)

③ if rank A = rank $(A|\underline{b}) = k < n$ => indeterminate, with $n-k$ degrees of freedom (∃ infinitely many solutions)

TYPES OF SYSTEMS

① $A\underline{x} = \underline{0}$ => homogeneous linear systems always possible (the solution is always $\underline{x} = \underline{0}$)

② $A\underline{x} = \underline{b}$, $\underline{b} \neq \underline{0}$ => non-homogeneous linear systems (may be possible or not possible)

remark

looking for $\underline{x}$: $A\underline{x} = \underline{b}$ $\approx$ looking for $\underline{x}$: $f(\underline{x}) = \underline{b}$

STRUCTURE OF SOLUTIONS

① homogeneous linear systems, $A\underline{x} = \underline{0}$ =>

- always possible + one solution is always $\underline{x} = \underline{0}$
- consider the set of all solutions : $V = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{0}\} = \text{Ker} f$ =>
  - the set of all solutions V is always a subspace of $\mathbb{R}^n$
  - dimension: $\dim V = n - \text{rank} A$ [by the rank nullity theorem]
  - therefore: - if $\text{rank} A = n$, one solution $\underline{x} = \underline{0}$, $V = \{\underline{0}\}$ and $\dim V = 0$
    if $\text{rank} A = k < n$, infinitely many solutions which form a subspace V of $\mathbb{R}^n$ with $\dim V = k - k$

② non-homogeneous linear systems, $A\underline{x} = \underline{b}$, $\underline{b} \neq \underline{0}$ =>

- possible or impossible -> check if $\text{rank} A = \text{rank}(A|b)$ + $\underline{x} = \underline{0}$ is never a solution
- consider the set of all solutions : $W = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{b}, \underline{b} \neq 0\}$, we have that $W = \{\underline{x} \in \mathbb{R}^n : f(\underline{x}) = \underline{b}, \underline{b} \neq 0\}$ [W is never a subspace of $\mathbb{R}^n$, $\underline{0} \notin \underline{W}$]
- THEOREM: STRUCTURE OF SOLUTIONS IN THE NON-HOMOGENEOUS CASE

  consider a linear system $A\underline{x} = \underline{b}$, $\underline{b} \neq \underline{0}$

  suppose it is possible ($\text{rank} A = \text{rank}(A|b)$), then:

  if you call $\underline{x}$ the general solution of $A\underline{x} = \underline{b}$ and you call $\underline{x}^*$ one particular solution of $A\underline{x} = \underline{b}$, you have $\underline{x = x^* + z}$

  where $\underline{z}$ is the general solution of the homogeneous linear system $A\underline{x} = \underline{b}$

  the task of solving a non-homogeneous linear system $A\underline{x} = \underline{b}$ can be reduced to finding one solution of $A\underline{x} = \underline{b}$ and solving the homogeneous linear system $A\underline{x} = \underline{0}$ instead

SOLUTION

given $A\underline{x} = \underline{b}$, suppose we checked that $A\underline{x} = \underline{b}$ is possible, we can find the solutions using Cramer ->

$A\underline{x} = \underline{b}$, with A square matrix of order n and $\det A \neq 0$

CRAMER'S RULE

consider $A\underline{x} = \underline{b}$, with A square of order n

if $\det A \neq 0$, then: $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and $x_1 = \dfrac{\det A_1}{\det A}, \dots, x_n = \dfrac{\det A_n}{\det A}$

where every matrix A_k is the matrix we obtain from matrix A, substituting the column k with the vector $\underline{b}$

A not square or A square, $\det A = 0 \Rightarrow$

CRAMER'S METHOD (example of the general procedure)

consider a linear system $A\underline{x} = \underline{b}$ where A is any $m \times n$ matrix

$$A = \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & -1 \\ 4 & 5 & -1 \\ 2 & 5 & -3 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -1 \\ -6 \\ -8 \end{pmatrix}$$

① $\operatorname{rank} A = \operatorname{rank} \begin{pmatrix} \boxed{3} & 2 & 1 \\ -1 & 0 & -1 \\ 4 & 5 & -1 \\ 2 & 5 & -3 \end{pmatrix} \quad \underbrace{\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}}_{} \rightarrow B$

$\det B = 0 + 2 = 2 \neq 0$

$\det \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & -1 \\ 4 & 5 & -1 \end{pmatrix} = 0 \qquad \text{III}c = \text{I}c - \text{II}c$

$\det \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & -1 \\ 2 & 5 & -3 \end{pmatrix} = 0 \qquad \text{III}c = \text{I}c - \text{II}c$

$\Rightarrow \operatorname{rank} A = 2$

$\operatorname{rank}(A|\underline{b}) = \operatorname{rank} \begin{pmatrix} 3 & 2 & 1 & -1 \\ -1 & 0 & -1 & -1 \\ 4 & 5 & -1 & -6 \\ 2 & 5 & -3 & -8 \end{pmatrix} \quad \rightarrow B$

proceding this way, since $\text{IV}c = \text{I}c - 2\,\text{II}c$, we get that also $\operatorname{rank}(A|\underline{b}) = 2$
therefore:
$A\underline{x} = \underline{b}$ is possible, indeterminate, with $3-2 = 1$ degree of freedom

② cancel out all the rows which have no intersection with the square submatrix B you find when computing the rank, that is:

$$A = \begin{pmatrix} 3 & 2 & 1 \\ -1 & 0 & -1 \\ \cancel{4} & \cancel{5} & \cancel{-1} \\ \cancel{2} & \cancel{5} & \cancel{-3} \end{pmatrix} \rightarrow B, \quad \underline{b} = \begin{pmatrix} -1 \\ -1 \\ \cancel{-6} \\ \cancel{-8} \end{pmatrix} \quad \Rightarrow \text{dependent} = \text{useless}$$

③ write the system explicitly: $\begin{cases} 3x_1 + 2x_2 + x_3 = -1 \\ -x_1 - x_3 = -1 \end{cases}$

we move all the columns which have no intersection with the square submatrix B to the right-hand side of the equation

$$\begin{cases} 3x_1 + 2x_2 = -1 - x_3 \\ -x_1 = -1 + x_3 \end{cases} \qquad \text{that is: } B = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}, \ \underline{b} = \begin{pmatrix} -1 - x_3 \\ -1 + x_3 \end{pmatrix}$$

④ on the left we have a square matrix B with $\det B \neq 0$ so we use Cramer's rule

$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, with:

$$x_1 = \frac{\det B_1}{\det B} = \frac{\det \begin{pmatrix} -1-x_3 & 2 \\ -1+x_3 & 0 \end{pmatrix}}{2} = \frac{2-2x_3}{2} = 1-x_3$$

$$x_2 = \frac{\det B_2}{\det B} = \frac{\det \begin{pmatrix} 3 & -1-x_3 \\ -1 & -1+x_3 \end{pmatrix}}{2} = \frac{-3+3x_3-1-x_3}{2} = -2+x_3$$

we find $k=2$ variables ($k = \operatorname{rank} A$) in terms of $n-k=1$ free parameter (degree of freedom):

$$\begin{cases} x_1 = 1-x_3 \\ x_2 = -2+x_3 \\ x_3 = x_3 \end{cases}$$

PROBLEM OF LEAST SQUARES

consider a linear system: $A\underline{x} = \underline{b}$ with $A \in \mathcal{M}(m,n)$

suppose that it has no solutions (often the case when $m>n$) that is when you have more equations than variables

if the linear system has no solutions it means that:

$\nexists\, \underline{x} \in \mathbb{R}^n$ such that $A\underline{x} = \underline{b}$ [equivalent: $A\underline{x} - \underline{b} = \underline{0}$]

? $\exists\, \underline{x}^* \in \mathbb{R}^n$ such that x^* minimises the error $\|A\underline{x} - \underline{b}\|$? = distance between the vector $\underline{b}$ of the constants and the image $T(\underline{x}) = A\underline{x}$ of the linear operator $T: \mathbb{R}^n \to \mathbb{R}^m$ that is individuated by the matrix A

this error is null where $\underline{x}^*$ solves the linear system $A\underline{x} = \underline{b}$, but in general we have: $\|A\underline{x} - \underline{b}\| \geq 0$

$\underline{x}^*$ minimises the error $\|A\underline{x} - \underline{b}\|$ if and only if it minimises its square $(\|A\underline{x} - \underline{b}\|)^2$

DEF. a vector $\underline{x}^* \in \mathbb{R}^n$ is said to be a least squares solution of the linear system: $A\underline{x} = \underline{b}$, if $\underline{x}^*$ is a solution of the optimisation problem: $\min\limits_{\underline{x} \in \mathbb{R}^n} (\|A\underline{x} - \underline{b}\|)^2$
that is if $\underline{x}^*$ is the best vector $\underline{x}$ that we can find in $\mathbb{R}^n$, in order to minimise the quantity $(\|A\underline{x} - \underline{b}\|)^2$
the name comes from the fact that $\|A\underline{x} - \underline{b}\|^2$ is a sum of squares

if the linear system $A\underline{x} = \underline{b}$ has a solution $\underline{x}$, it is also a least squares solution but a least squares solutions $\underline{x}^*$ may exist also when the linear system $A\underline{x} = \underline{b}$ has no solution $\underline{x}$

THEOREM

consider the linear system $A\underline{x}=\underline{b}$, with $A \in \mathcal{M}(m,n)$

suppose $m \geq n$

then if rank $A = n$ [if A has maximum rank], the optimalisation problem:

$\min_{\underline{x} \in \mathbb{R}^n} (\|A\underline{x}-\underline{b}\|)^2$ has a solution and this solution is unique

theorems and proofs for
linear algebra

THE INTERSECTION OF VECTOR SUBSPACES

the intersection of any collection of vector subspaces of $\mathbb{R}^n$ is a vector subspace

PROOF

let $\{V_i\}$ be any collection of vector subspaces of $\mathbb{R}^n$

since $0 \in V_i \; \forall i$, we have $\cap_i V_i \neq \emptyset$

let $x, y \in \cap_i V_i$ and $\alpha, \beta \in \mathbb{R}$

since $x, y \in V_i \; \forall i$ and therefore $\alpha x + \beta y \in V_i \; \forall i$ since each V_i is a vector subspace of $\mathbb{R}^n$

hence $\alpha x + \beta y \in \cap_i V_i$ and so $\cap_i V_i$ is a vector subspace of $\mathbb{R}^n$

CHARACTERIZATION WITH LINEAR COMBINATIONS

S subset of R^n

$\underline{x} \in$ Span S <=> $\underline{x}$ is a linear combination of vectors of s

=> $\exists\, \underline{x}_1, \dots, \underline{x}_m \in S$, $\exists\, \alpha_1, \dots, \alpha_m \in R$ such that $\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m$

PROOF

=> only if

V = set of all vectors $\underline{x} \in R^n$ that can be written as a linear combination of vectors of S

if $\underline{x}, \underline{y} \in V$ then $\underline{x}, \underline{y}$ are linear combinations of vectors of s and therefore also $\underline{x} + \underline{y}$ is a linear combination of vectors of s which also implies $\underline{x} + \underline{y} \in V$

V is a vector subspace of R^n containg S so it must be that span $S \subseteq V$, that is:

each $\underline{x} \in$ span S can be written as a linear combination of vectors of S

<= if

suppose $\underline{x} \in R^n$ linear combination of vectors of s: $\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m$ for some $\alpha_1, \dots, \alpha_m$ and $\underline{x}_1, \dots, \underline{x}_m \in S$

since $\underline{x}_1, x_2 \in S$ they belong to span s

since span S is a vector space also $\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2$ must belong to span S

since $\underline{x}_3 \in S$ it also belongs to span s

since span s is a vector space also $\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \alpha_3 \underline{x}_3$ belongs to span s

proceding in this way in the end we get that also $\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m$ belongs to span S

as a consequence of the theorem if: $S = \{ \underline{x}_1, \dots, \underline{x}_m \in R^n \}$ then,

span $S = \{ \underline{x} \in R^n : \exists\, \alpha_1, \dots, \alpha_m \in R$ such that $\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m \}$

PROPERTY OF UNIQUE WRITING FOR A BASIS

take $\underline{x}_1, \dots, \underline{x}_m \in \mathbb{R}^n$

they are a basis of $\mathbb{R}^n$ <=> $\forall \underline{x} \in \mathbb{R}^n$, $\underline{x}$ can be written in a unique way as a linear combination of $\underline{x}_1, \dots, \underline{x}_m$

PROOF

=> only if

suppose $\underline{x}_1, \dots, \underline{x}_m$ are a basis of $\mathbb{R}^n$

by definition of basis every $\underline{x} \in \mathbb{R}^n$ can be written as a linear combination of $\underline{x}_1, \dots, \underline{x}_m$

$\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m$

$\underline{x} = \beta_1 \underline{x}_1 + \dots + \beta_m \underline{x}_m$

this means that: $(\alpha_1 - \beta_1)\underline{x}_1 + \dots + (\alpha_m - \beta_m)\underline{x}_m = \underline{0}$

but since $\underline{x}_1, \dots, \underline{x}_m$ are linearly independent, this means that it must be:

$\alpha_1 - \beta_1 = 0, \dots, \alpha_m - \beta_m = 0$

that is: $\alpha_1 = \beta_1, \dots, \alpha_m = \beta_m$

which proves the unique possible writing

<= if

suppose that $\forall \underline{x} \in \mathbb{R}^n$, $\underline{x}$ can be written in a unique way as a linear combination of $\underline{x}_1, \dots, \underline{x}_m$

this implies that $\text{span } S = \mathbb{R}^n$

we have to prove that $\underline{x}_1, \dots, \underline{x}_m$ are linearly independent

but if we suppose that $\exists\, \alpha_1, \dots, \alpha_m$ such that: $\alpha_1 \underline{x}_1 + \dots + \alpha_m \underline{x}_m = \underline{0}$

it is enough to remark that also $0\underline{x}_1 + \dots + 0\underline{x}_m = \underline{0}$

since $\underline{0}$ can only be written in a unique way as a linear combination of $\underline{x}_1, \dots, \underline{x}_m$ this implies that: $\alpha_1 = 0, \dots, \alpha_m = 0$

that is: $\underline{x}_1, \dots, \underline{x}_m$ are linearly independent

DETERMINATION OF COEFFICIENTS OF A VECTOR IN R^n WITH RESPECT TO AN ORTHONORMAL BASIS

consider an orthonormal basis $\underline{x}_1, \ldots, \underline{x}_n$ in R^n

the for every $\underline{v} = \alpha_1 \underline{x}_1 + \ldots + \alpha_n \underline{x}_n$ with $\alpha_1 = \underline{v} \cdot \underline{x}_1, \ldots, \alpha_n = \underline{v} \cdot \underline{x}_n$

meaning that if $\underline{x}_1, \ldots, \underline{x}_n$ is an orthonormal basis if R^n, there is a quick way to write the coefficients $\alpha_1, \ldots, \alpha_n$

PROOF

let $\{x_1, x_2, \ldots, x_n\}$ be an orthonormal basis of R^n

for any vector $v \in R^n$ it can be expressed as a linear combination of the basis vectors:

$v = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the coefficients we want to determine

by the orthonormality of the basis:

$$x_i \cdot x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

take the dot product of both sides of $v = \sum_{i=1}^{n} a_i x_i$ with x_k where $k \in \{1, 2, \ldots, n\}$:

$v \cdot x_k = \left(\sum_{i=1}^{n} \alpha_i x_i\right) \cdot x_k$

using the linearity of the dot product:

$v \cdot x_k = \sum_{i=1}^{n} \alpha_i (x_i \cdot x_k)$

from the orthonormality property:

$$x_i \cdot x_k = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$$

therefore all terms in the sum vanish except when $i = k$:

$v \cdot x_k = \alpha_k (x_k \cdot x_k) = \alpha_k \cdot 1 = \alpha_k$

thus the coefficient α_k is given by:

$\alpha_k = v \cdot x_k$

the coefficient $\alpha_1, \alpha_2, \ldots, \alpha_n$ of the vector v with respect to the orthonormal basis $\{x_1, x_2, \ldots, x_n\}$ can be quickly determined as:

$\alpha_k = v \cdot x_k$, for $k = 1, 2, \ldots, n$

RIESZ REPRESENTATION THEOREM

a function $f: R^n \to R$ is linear if and only if there exists a unique vector $\chi \in R^n$ such that:

$f(x) = \chi \cdot x \quad \forall x \in R^n$

PROOF

$\Leftarrow$ "if"

assume $f(x) = \chi \cdot x$ for some $\chi \in R^n$

we show that f is linear

for $x, y \in R^n$, $f(x+y) = \chi \cdot (x+y) = \chi \cdot x + \chi \cdot y = f(x) + f(y)$

for $\alpha \in R$ and $x \in R^n$, $f(\alpha x) = \chi \cdot (\alpha x) = \alpha(\chi \cdot x) = \alpha f(x)$

since f satisfies the additivity and homogeneity, f is linear

$\Rightarrow$ "only if"

assume $f: R^n \to R$ is linear

we show that there exists a unique $\chi \in R^n$ such that $f(x) = \chi \cdot x$

define χ as $\chi_i = f(e_i)$ for $i = 1, \dots, n$ where e_i is the i-th standard basis vector of R^n

for any $x = (x_1, x_2, \dots, x_n) \in R^n$, $x = \sum_{i=1}^n x_i e_i$

and by linearity of f, $f(x) = f(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_i f(e_i) = \sum_{i=1}^n x_i \chi_i = \chi_i \cdot x_i$

if $f(x) = y^* \cdot x$ for another vector y^* then $\chi \cdot x = y^* \cdot x \quad \forall x \in R^n$

choosing $x = e_i$, we get $\chi_i = y^*_i$, thus $\chi = y^*$

RIESZ REPRESENTATION THEOREM FOR OPERATORS $f: R^n \to R^m$

① a function $f: R^n \to R^m$ is linear $\Leftrightarrow$ there exists a matrix $A \in \mathcal{M}(m,n)$ called the representation matrix of f such that $\forall \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n$ we have: $f(\underline{x}) = A\underline{x}$

② when it exists this matrix A is unique

PROOF

$\Leftarrow$ "if"

suppose $\exists A \in \mathcal{M}(m,n)$ such that:

$f(\underline{x}) = A\underline{x} \quad \forall x \in R^n$, then:

① $\forall \underline{x}, \underline{y} \in R^n$, $f(\underline{x}+\underline{y}) = A(\underline{x}+\underline{y}) = A\underline{x} + A\underline{y} = f(\underline{x}) + f(\underline{y})$

② $\forall \underline{x} \in R^n$, $\forall \alpha \in R$, $f(\alpha\underline{x}) = A(\alpha\underline{x}) = \alpha(A\underline{x}) = \alpha f(\underline{x})$

$\Rightarrow$ "only if"

we use the fact that:

$\forall \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n$ we can write $\underline{x} = x_1\underline{e}_1 + \dots + x_n\underline{e}_n$

suppose f is linear, $\forall \underline{x} \in R^n$ we have:

$$\begin{aligned} f(\underline{x}) &= f(x_1\underline{e}_1 + \dots + x_n\underline{e}_n) \\ &= f(x_1\underline{e}_1) + \dots + f(x_n\underline{e}_n) \\ &= x_1 f(\underline{e}_1) + \dots + x_n f(\underline{e}_n) \\ &= \underline{\alpha}_1 x_1 + \dots + \underline{\alpha}_n x_n = A\underline{x} \end{aligned}$$

if we consider the matrix $A = (\underline{a}_1, \dots, \underline{a}_n) = (f(\underline{e}_1), \dots, f(\underline{e}_n))$ whose colums are given by the vectors: $\underline{\alpha}_1 = f(\underline{e}_1), \dots, \underline{\alpha}_n = f(\underline{e}_n)$

② uniqness $\Rightarrow$

suppose there exist $A, A' \in \mathcal{M}(m,n)$ such that $\forall \underline{x} \in R^n$, $f(\underline{x}) = A\underline{x} = A'\underline{x}$

if we take $\underline{x} = \underline{e}_1$ we find:

$f(\underline{e}_1) = A\underline{e}_1 = \underline{\alpha}_1$

$f(\underline{e}_1) = A'\underline{e}_1 = \underline{\alpha}'_1$

where $\underline{\alpha}_1$ is the first column of A and $\underline{\alpha}'_1$ is the first column of A'

therefore $\underline{\alpha}_1 = \underline{\alpha}'_1$

taking $\underline{x} = \underline{e}_2$ we prove that it is $\underline{\alpha}_2 = \underline{\alpha}'_2$ and so on, therefore $A = A'$

THE IMAGE SPACE IS A SUBSPACE SPANNED BY THE IMAGE OF STANDARD BASIS

① $f(\mathbb{R}^n) = \mathrm{Im} f$ is always a vector subspace of $\mathbb{R}^m$

② a spanning set for $\mathrm{Im} f$ is always: $f(\underline{e}_1), \dots, f(\underline{e}_n)$

PROOF

① consider $\underline{y}_1, \underline{y}_2 \in \mathrm{Im} f$, $\alpha_1, \alpha_2 \in \mathbb{R}$

then $\exists\, \underline{x}_1 \in \mathbb{R}^n : \underline{y}_1 = f(\underline{x}_1)$, $\exists\, \underline{x}_2 \in \mathbb{R}_n : \underline{y}_2 = f(\underline{x}_2)$

since f is linear, $f(\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2) = \alpha_1 f(\underline{x}_1) + \alpha_2 f(\underline{x}_2) = \alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2$

this means that $\alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2 \in \mathrm{Im} f$

therefore $\mathrm{Im} f$ is a vector subspace of $\mathbb{R}^m$

② consider $\underline{y} \in \mathrm{Im} f$

then $\exists\, \underline{x} \in \mathbb{R}^n : \underline{y} = f(\underline{x})$

considering the fundamental basis $\underline{e}_1, \dots, \underline{e}_n$ of $\mathbb{R}^n$, we have that $\underline{x} = x_1 \underline{e}_1 + \dots + x_n \underline{e}_n$

since f is linear we have $\underline{y} = f(\underline{x}) = f(x_1 \underline{e}_1 + \dots + x_n \underline{e}_n) = x_1 f(\underline{e}_1) + \dots + x_n f(\underline{e}_n)$

this means that $\underline{y}$ is a linear combination of $f(\underline{e}_1), \dots, f(\underline{e}_n)$

therefore $f(\underline{e}_1), \dots, f(\underline{e}_n)$ is a spanning set for $\mathrm{Im} f$

THE KERNEL IS A SUBSPACE

if $T: \mathbb{R}^n \to \mathbb{R}^m$ is linear then $\ker T$ is a vector subspace of $\mathbb{R}^n$

PROOF

let $x, x' \in \ker T$, i.e. $T(x) = 0$ and $T(x') = 0$

for every $\alpha, \beta \in \mathbb{R}$ we have: $T(\alpha x + \beta x') = \alpha T(x) + \beta T(x') = \alpha 0 + \beta 0 = 0$

thus, $\alpha x + \beta x' \in \ker T$ and this proves that $\ker T$ is a vector subspace of $\mathbb{R}^n$

DETERMINANT OF THE INVERSE MATRIX

for any invertible square matrix A, the determinant of its inverse satisfies:

$\det(A^{-}) = \frac{1}{\det(A)}$ where $\det(A) \neq 0$

PROOF

by definition of the inverse matrix, $A \cdot A^{-1} = I$, where I is the identity matrix

taking the determinant of both sides, we have $\det(A \cdot A^{-1}) = \det(I)$

using the determinant property $\det(AB) = \det(A) \cdot \det(B)$, this becomes:

$\det(A) \cdot \det(A^{-1}) = \det(I)$

since the determinant of the identity matrix is 1 ($\det(I) = 1$), it follows that:

$\det(A) \cdot \det(A^{-1}) = 1$

rearranging we find: $\det(A^{-1}) = \frac{1}{\det(A)}$

this result holds as long as A is invertible, which implies $\det(A) \neq 0$

KRONECKER - CAPELLI'S THEOREM

given a linear system $A\underline{x} = \underline{b}$ with m equations and n variables,
$A\underline{x} = \underline{b}$ is possible $\Leftrightarrow$ $\operatorname{rank} A = \operatorname{rank}(A|\underline{b})$

PROOF

$A\underline{x} = \underline{b}$ is possible $\Leftrightarrow$
$\exists\ \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n :\ A\underline{x} = \underline{b} \Leftrightarrow$
$\exists\ x_1, \dots, x_n \in \mathbb{R} :\ \underline{b}$ can be written as a linear combination $\underline{b} = x_1 \underline{\alpha}_1 + \dots + x_n \underline{\alpha}_n$ of the columns of $A \Leftrightarrow$
$\underline{b}$ is linearly dependent on the columns of $A \Leftrightarrow$
$\operatorname{rank} A = \operatorname{rank}(A|\underline{b})$

CRAMER'S THEOREM

consider the system $A\underline{x} = \underline{b}$ where A is a square matrix of order n:

① the system $A\underline{x} = \underline{b}$ is determinable (it has a unique solution) if and only if $\det(A) \neq 0$

② if $\det(A) \neq 0$ then the solution is given by $\underline{x} = A^{-1}\underline{b}$

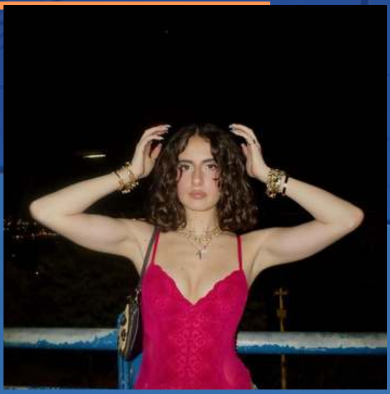
PROOF

① if $\det(A) \neq 0$ the matrix A is invertible by definition of the determinant
the invertibility of A ensures the existence of A^{-1}, the inverse of A
for the system $A\underline{x} = \underline{b}$, multiplying both sides by A^{-1}, we get: $A^{-1}(A\underline{x}) = A^{-1}\underline{b}$
using the property $A^{-1}A = I$, this simplifies to $\underline{x} = A^{-1}\underline{b}$
since A^{-1} exists this is the unique solution

② if $\det(A) = 0$, the matrix A is not invertible, which implies A is singular
for a singular matrix, the system $A\underline{x} = \underline{b}$ may have no solution (if $\underline{b}$ lies outside the column space of A) or infinitely many solutions if $\underline{b}$ lies in the column space of A but A does not have full rank
in either cases the system is not uniquely solvable (it is not determinable)

the system $A\underline{x} = \underline{b}$ is uniquely solvable if and only if $\det(A) \neq 0$ in which case the solution is $\underline{x} = A^{-1}\underline{b}$

FOR DOUBTS OR SUGGESTIONS ON THE HANDOUTS



ALESSANDRA VITIELLO

alessandra.vitiello@studbocconi.it
@alexxvitiello
+39 344 491 6912

FOR INFO ON THE TEACHING DIVISION

VITTORIA NASONTE

vittoria.nasonte@studbocconi.it
@_vittorian_
+39 3274441476

ELENA CACIOLI

elena.cacioli@studbocconi.it
@elenacaciolii_
+39 3928931605

TEACHING DIVISION

OUR PARTNERS

7oC+ CLUB

ETHAN SUSTAINABILITY

DELIVERY VALLEY
NO GENDER KITCHEN

LA PIADINERIA

LA PIADINERIA dal 1994