Assignment 2 — MASD

Schmidt, Victor Alexander, `rqc908`
Ibsen, Helga Rykov, `mcv462`
Barchager, Thomas Haulik, `jxg170`

Wednesday 22:00, September 14th

Opgave 1

1.

(a)

$$\frac{f(t+h)-f(t)}{h}, \qquad \text{h er lille}$$

Dette udtryk skulle betyde at vi fik den afledede funktion $f'(t)$ til $f(t)$, hvis vi indsatte et lille tal på h's plads. Det gør vi ikke: jo mindre, jo tætter kommer vi på grænsen, men det giver os ikke selve grænsen, $f'(t)$. Med andre ord ville (a) give os sekantens hældning, men den vil ikke give os tangentens hældning.

(b)

$$\lim_{h\to 0}\frac{f(t+4h)-f(t)}{h}$$

Udtrykket i (b) giver os heller ikke den afledede funktion til $f(t)$. Det kan vi vise ved at gange tælleren og nævneren med en vilkårlig konstant og sætte den i tælleren foran $\lim_{h\to 0}$:

$$4\cdot\lim_{h\to 0}\frac{f(t+4h)-f(t)}{4h} \tag{1}$$

Da $h=4h$ er den samme både i tælleren og nævneren, så giver udtrykket:

$$\lim_{h\to 0}\frac{f(t+4h)-f(t)}{4h} \tag{2}$$

os grænsen for $f(t)$.

I udtrykket (1b) derimod går funktionen $f(t)$ fire gange så hurtigt mod grænsen som i (2b):

$$\begin{aligned}&\lim_{h\to 0}\frac{f(t+4h)-f(t)}{h}\\ &=4\cdot\lim_{h\to 0}\frac{f(t+4h)-f(t)}{4h}\\ &=4\cdot f'(t)\end{aligned}$$

(c) Som det blev forklaret ovenfor, giver os udtrykket

$$\lim_{h\to 0}\frac{f(t+4h)-f(t)}{4h}$$

grænsen for $f(t)$ når h går mod 0.

(d) For at vise at udtrykket

$$\lim_{h\to 0}\frac{f(t+h)-f(t-h)}{2h}$$

rent faktisk giver os grænsen for f, når $h \to 0$, skal vi simplificere udtrykket. Vi starter med at lægge $f(t)$ til og trække den fra i nævneren:

$$\lim_{h\to 0} \frac{f(t+h) - f(t) - f(t-h) + f(t)}{2h} \quad \text{sætter udtrykket på 2 brøk}$$
$$= \lim_{h\to 0} \frac{f(t+h) - f(t)}{2h} - \frac{f(t-h) - f(t)}{2h} \quad \text{sætter konstanten foran lim}$$
$$= \frac{1}{2}(f'(t) - \lim_{h\to 0} \frac{f(t-h) - f(t)}{h}) \quad \text{ganger tælleren og nævneren af anden led med -1}$$
$$= \frac{1}{2}(f'(t) + \lim_{h\to 0} \frac{f(t-h) - f(t)}{-h}) \quad \text{h er ens i tælleren og nævneren, så det giver os}$$
$$= \frac{1}{2}(f'(t) + f'(t)) = f'(t)$$

2.

Den første graf på Figur 1 ligner til forveksling den klokkeformede frekvensfunktion for normalfordeling. For at tegne dens afledede starter vi med at bestemme de punkter på grafen, hvor f's hældning er faldende, stigende og lige med 0. I disse punkter på grafen (tegnet med grønt på Figur 1(a) 1) tegner jeg tangenter, som giver os en ide om, hvor stejl hældningen er i disse punkter:

(a) f1(x)

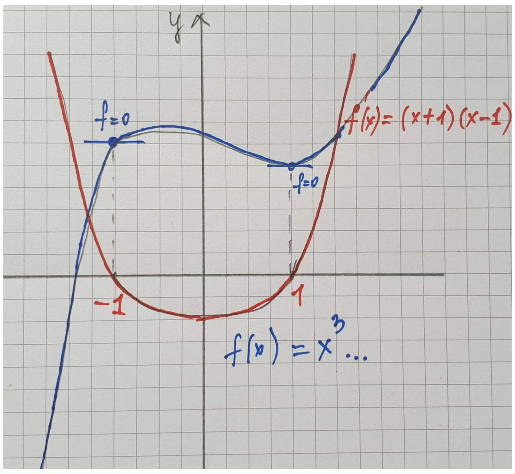
(b) f2(x)

Figur 1: De røde grafer for $f'(x)$

Hvis vi kigger på grafen f (Figur 1(a)) fra venstre til højre, så kan vi se at dens hældning i 1.punkt er større end i det næste. Jeg finder dermed to røde punkter for f' hvor x-koordinatet er den samme som i f og y-koordinatet afspejler tangenternes hældning i de to punkter. Og vi følger den samme logik på den anden side af y-aksen. Til sidst afsætter vi et punkt på f', hvor f har sit minimum, dvs. hvor f's hældning er 0.

3

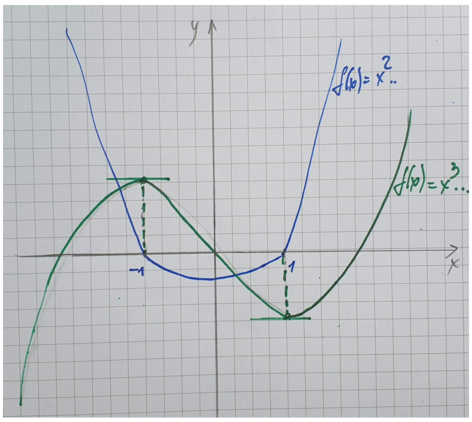
Hvis vi antager at $f'(-1) = f'(1) = 0$ og $f'(t) \neq 0$ når $t \notin$ {-1, 1}, så kan en mulig afledede funktion f' tage en form af en andengradspolynomium, som krydser x-aksen kun i to punkter $t = -1$ og $t = 1$ (se den røde graf på Figur 2 2):



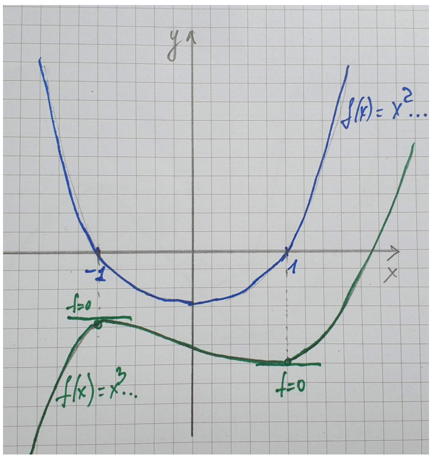
Figur 2: En 3.gradspolynomium (blå graf) som stamfunktion til f'(x) (2.gradspolynomium, den røde graf)

At funktionsværdien af f' er lige med 0 kun i $t = -1$ og $t = 1$ betyder at en mulig stamfunktion til f' har kun 2 ekstremer i de to punkter, hvor $t = -1$ og $t = 1$. I den mulig version af f' (den røde graf) som det er tegnet på Figur 2, har en mulig stamfunktion en negative hældning i intervallet {-1, 1}, og positiv hældning andetsteds og er hermed en tredjegradspolynomium.

Figur 3 3 viser to andre varianter af en mulig 3.gradspolynomium, som stamfunktion til en 2.gradspolynomium. Det betyder med andre ord at stamfunktionens to ekstremer kan ligge mange forskellige steder på y-aksen, både over og under x-aksen. Det vigtigste er at funktionen er faldene imellem sine to ekstremer.



(a) f1(x)



(b) f2(x)

Figur 3: De grønne grafer for $f(x)$

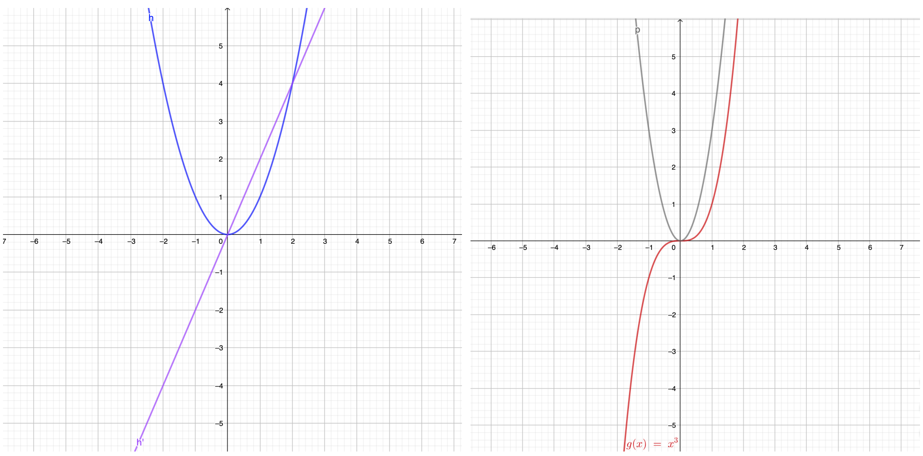
Den afledede f' kan også antage formen af en omvendt parabel, som krydser x-aksen i de samme to punkter. I dette tilfælde ville de mulige 3.gradspolymier være stigende imellem sine to ekstremer og faldene andetsteds (dvs. være spejlvendte til de grønne grafer i Figur 3 3).

4

Den anden afledede f'' af en funktion f er jo den afledede (differentialkvotienten) af f'. Vi ved fra monotonisætningen, at hvis fx den afledede er positiv, så er funktionen voksende:

1. Hvis f'' er positiv, så er f' voksende
2. Hvis f'' er negativ, så er f' aftagende

Hvis vi antager at $f'(0) = 0$, så betyder det at f har et ekstremum i 0. Og $f''(0) > 0$ betyder at f' er voksende i $f(0) = 0$. En mulig graf til f kunne derfor være en 2.gradspolynomium, som er den blå graf (h(x)) Figur 4(a) 4. Og den mulig f' til den er en ret linje som er den lilla graf i Figur 4(a) 4:



(a) h(x) og h'(x)

(b) g(x) og p(x)=g'(x)

Figur 4: Mulig grafer for $f(x)$ og $f'(x)$

Vi ved yderligere fra maks-min sætningen at hvis f'' er lige med 0, så har f' et maksimum eller et minimum der.

Hvis vi antager at $f'(0) = 0$, så betyder det at f har et ekstremum i 0. Og $f''(0) = 0$ betyder at f' har et maksimum eller et minimum i 0. En mulig funktion kunne være en 3.gradspolynomium, tegnet som den røde graf i Figur 4(b) 4 og dens afledede er en 2.gradspolynomium, tegnet som den grå graf i Figur 4(b) 4. Hvis vi kigger på den røde graf, så kan vi også sige at i $g(0)$ har den en vandret vendetangent, hvor g skifter fra at krumme nedad til at krumme opad omkring origo.

Opgave 2

Som intro til opgave 2, opstiller vi regler til differentiering af funktioner:

Kædereglen:

$$f'(g(x)) = f'(g(x)) \cdot g'(x) \tag{3}$$

Produktreglen:

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x) \tag{4}$$

Brøkreglen:

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2} \tag{5}$$

1.

Løs følgende:

$$\frac{d}{dx}\sqrt{x^2+3}$$

Vi gør brug af kædereglen:

$$\frac{d}{dx}\sqrt{x^2+3} \cdot \frac{d}{dx}(x^2+3)$$
$$\frac{1}{2\sqrt{x^2+3}} \cdot 2x$$
$$\frac{x}{\sqrt{x^2+3}}$$

2.

Løs følgende:

$$\frac{d}{dx}\sin x e^{-\cos x^2}$$

Først gør vi brug af produktreglen:

$$(\sin x)' \cdot e^{-\cos x^2} + \sin x \cdot \left(e^{-\cos x^2}\right)'$$

Nu kan vi bruge kædereglen på ledet $\left(e^{-\cos x^2}\right)'$:

$$(\sin x)' \cdot e^{-\cos x^2} + \sin x \cdot \left(e^{-\cos x^2} \cdot (-\cos x^2)'\right)$$

Igen kan kædereglen bruges, denne gang på ledet $(-\cos x^2)'$:

$$\cos x \cdot e^{-\cos x^2} + \sin x \cdot e^{-\cos x^2} \cdot -2\cos x \cdot (\cos x)'$$

$$\cos x \cdot e^{-\cos x^2} + \sin x \cdot e^{-\cos x^2} \cdot -2\cos x \cdot -\sin x$$
$$\cos x \cdot e^{-\cos x^2} + 2(\sin x)^2 \cos x e^{-\cos x^2}$$
$$\cos x \cdot e^{-\cos x^2} + (1 + 2(\sin x)^2)$$

3.

Løs følgende:

$$\frac{d}{dx}\frac{\ln x}{x^3}$$

Vi gør brug af brøkreglen:

$$\frac{(\ln x)' x^3 - \ln x (x^3)'}{(x^3)^2}$$
$$\frac{x^{-1}x^3 - \ln x 3x^2}{x^6}$$
$$\frac{1 - 3\ln x}{x^4}$$

4.

Løs følgende:

$$\frac{d}{dx}\frac{\ln 1 + e^{qx}}{q}$$

Vi gør brug af kædereglen:

$$q^{-1}(\ln 1 + e^{qx})'(1 + e^{qx})'$$
$$q^{-1}(1 + e^{qx})^{-1} q e^{qx}$$
$$(1 + e^{qx})^{-1} e^{qx} = \frac{e^{qx}}{1 + e^{qx}}$$

5.

Løs følgende:

$$\frac{d^2}{dx^2}\frac{\ln 1 + e^{qx}}{q} = \frac{d}{dx}\frac{e^{qx}}{1 + e^{qx}}$$

Vi gør brug af brøkreglen:

$$\frac{(e^{qx})'(1 + e^{qx}) - e^{qx}(1 + e^{qx})'}{(1 + e^{qx})^2}$$
$$\frac{qe^{qx}(1 + e^{qx}) - e^{qx} q e^{qx}}{(1 + e^{qx})^2}$$
$$\frac{qe^{qx} + e^{qx} q e^{qx} - e^{qx} q e^{qx}}{(1 + e^{qx})^2}$$
$$\frac{qe^{qx}}{(1 + e^{qx})^2}$$

6.

Løs følgende:

$$\frac{d}{dx} e^{(x^2+y)^3}$$

Vi gør brug af kædereglen:

$$e^{(x^2+y)^3}((x^2 + y)^3)'$$

Vi gør igen brug af kædereglen på ledet $((x^2+y)^3)'$:

$$e^{(x^2+y)^3}3(x^2+y)^2 2x = 6e^{(x^2+y)^3}(x^2+y)^2 x$$

7.

Løs følgende:

$$\frac{d}{dx}(\mathbf{x^T A x})$$

Hvis $n = 3$ vil matricerne se ud som følgende:

$$\frac{d}{dx}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Hvis vi nu ganger matricerne ud, får vi følgende udtryk:

$$x_1^2 a_{11} + x_1 x_2 a_{12} + x_1 x_3 a_{13} + x_2 x_1 a_{21} + x_2^2 a_{22} + x_2 x_3 a_{23} + x_3 x_1 a_{31} + x_3 x_2 a_{32} + x_3^2 a_{33}$$

Hvis vi differentierer ift. x_1, kan vi se at alle led hvor x_1 ikke indgår forsvinder:

$$\frac{d}{dx_1} = 2x_1 a_{11} + x_2 a_{12} + x_3 a_{13} + x_2 a_{21} + x_3 a_{31}$$

$$\frac{d}{dx_2} = 2x_2 a_{22} + x_1 a_{12} + x_3 a_{23} + x_1 a_{21} + x_3 a_{32}$$

$$\frac{d}{dx_3} = 2x_3 a_{33} + x_2 a_{32} + x_1 a_{13} + x_3 a_{23} + x_1 a_{31}$$

Vi kan se et mønster, der opfylder ovenstående ligninger for $\frac{d}{dx_1}$, $\frac{d}{dx_2}$ og $\frac{d}{dx_3}$ der ser ud som følgende:

$$\frac{d}{dx_i} = \sum_{k=1}^{m} (x_k \cdot a_{ik}) + (x_k \cdot a_{ki})$$

... hvor m er dimensionen på matricen $\mathbf{A}$. Det antages selvfølgelig at $0 < i \leq m, i \in \mathbb{N}$

8.

Løs følgende:

$$\frac{d}{dx_i}(\mathbf{Ax} - \mathbf{b})^{\mathbf{T}}(\mathbf{Ax} - \mathbf{b})$$

Igen forsøger vi os med $n = 3$, og vi prøver med $m = 3$, altså et kvadratisk matrix $\mathbf{A}$:

$$\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right)^{\mathbf{T}}\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right)$$

$$\left(\begin{bmatrix} x_1 a_{11} + x_2 a_{12} + x_3 a_{13} - b_1 \\ x_1 a_{21} + x_2 a_{22} + x_3 a_{23} - b_2 \\ x_1 a_{31} + x_2 a_{32} + x_3 a_{33} - b_3 \end{bmatrix}\right)^{\mathbf{T}}\left(\begin{bmatrix} x_1 a_{11} + x_2 a_{12} + x_3 a_{13} - b_1 \\ x_1 a_{21} + x_2 a_{22} + x_3 a_{23} - b_2 \\ x_1 a_{31} + x_2 a_{32} + x_3 a_{33} - b_3 \end{bmatrix}\right)$$

Efter transponering af venstre side, får vi følgende udtryk efter at gange sammen:

$$(x_1a_{11}+x_2a_{12}+x_3a_{13}-b_1)^2+(x_1a_{21}+x_2a_{22}+x_3a_{23}-b_2)^2+(x_1a_{31}+x_2a_{32}+x_3a_{33}-b_3)^2$$

Vi gør brug af kædereglen:

$$\begin{aligned}
\frac{d}{dx_1} =& 2a_{11}(x_1a_{11}+x_2a_{12}+x_3a_{13}-b_1)+\\
& 2a_{21}(x_1a_{21}+x_2a_{22}+x_3a_{23}-b_2)+\\
& 2a_{31}(x_1a_{31}+x_2a_{32}+x_3a_{33}-b_3)\\
\frac{d}{dx_2} =& 2a_{12}(x_1a_{11}+x_2a_{12}+x_3a_{13}-b_1)+\\
& 2a_{22}(x_1a_{21}+x_2a_{22}+x_3a_{23}-b_2)+\\
& 2a_{32}(x_1a_{31}+x_2a_{32}+x_3a_{33}-b_3)\\
\frac{d}{dx_3} =& 2a_{13}(x_1a_{11}+x_2a_{12}+x_3a_{13}-b_1)+\\
& 2a_{23}(x_1a_{21}+x_2a_{22}+x_3a_{23}-b_2)+\\
& 2a_{33}(x_1a_{31}+x_2a_{32}+x_3a_{33}-b_3)
\end{aligned}$$

For at finde et udtryk for $\frac{d}{dx_i}(\mathbf{Ax}-\mathbf{b})^{\mathbf{T}}(\mathbf{Ax}-\mathbf{b})$, har vi studeret ovenstående eksempel, og forestillet os hvordan ligningerne ville ændre sig, hvis vi ændrede på dimensionerne af $\mathbf{A}$, altså på $m \times n$. Vi har meget uformelt fundet frem til følgende:

$$\frac{d}{dx_i}(\mathbf{Ax}-\mathbf{b})^{\mathbf{T}}(\mathbf{Ax}-\mathbf{b}) = \sum_{k=1}^{m}\left(2a_{ki}\cdot\left(\left(\sum_{j=1}^{n}x_ja_{jk}\right)-b_k\right)\right)$$

... hvor dimensionerne af $\mathbf{A}$ er $m \times n$, $\mathbf{x}$ er $n \times 1$ og $\mathbf{b}$ er $m \times 1$. Dimensionerne er også givet i opgaven, de er givet her for at gøre det mere klart hvor m og n kommer fra. Dette gælder kun når $0 < i \leq n, i \in \mathbb{N}$

Opgave 3

1.

`badness` funktionen er blevet implementeret i vores kode (**.ipynb** fil). Funktionen var givet i opgaven som:

$$\frac{1}{20}\sum_{i=1}^{20}(N_i - aC_i - b)^2 \tag{6}$$

2.

For at implementere $\nabla_{(a,b)}badness$ funktionen i **.ipynb** filen, skal vi bruge den afledte af formel 6 ift. a og b, da vi skal udregne gradienten. Vi udregner derfor følgende:

$$\frac{d}{da} = \left(\frac{1}{20}\sum_{i=1}^{20}(N_i - aC_i - b)^2\right)'$$

Her kan vi gøre brug af kædereglen:

$$\frac{d}{da} = \frac{1}{20}\sum_{i=1}^{20}2(N_i - aC_i - b)\cdot -C_i$$

$$\frac{d}{da} = \frac{1}{10}\sum_{i=1}^{20}-N_iC_i + C_i^2 + bC_i$$

Vi har altså altså den ene af de afledte, til gradienten af formel 6. Nu udregner vi den anden afledte, $\frac{d}{db}$, hvor vi også starter med at gøre brug af kædereglen:

$$\frac{d}{db} = \left(\frac{1}{20}\sum_{i=1}^{20}(N_i - aC_i - b)^2\right)'$$

$$\frac{d}{db} = \frac{1}{20}\sum_{i=1}^{20}2(N_i - aC_i - b)\cdot -1$$

$$\frac{d}{db} = \frac{1}{10}\sum_{i=1}^{20}-N_i + C_i + b$$

Nu har vi altså de afledte funktioner som ∇f består af, hvis vi kalder `badness` funktionen for f

3.

I denne opgave har vi defineret en funktion, `optimal_badness` i vores **.ipynb** fil. Den tager startpunktet (a, b), `la` (λ i den givne funktion), vores datasæt for chokolade/nobel og bruger `badness_gradient` (vores $\nabla_{(a,b)}badness$) til at udregne gradienten i punktet (a, b). Derefter laver den og returnerer funktionen en tuple af nye punkter, som har bevæget sig modsat gradienten, og er udregnet ved hjælp af `gradientDescent` som var givet.

Vi har kørt `optimal_badness` i et `while-loop`, indtil $badness < 50$. Vi endte med følgende værdier for a, b og $badness$ (de eksate værdier står i **.ipynb** filen):

1. $a \approx 2,355$
2. $b \approx -1,050$
3. $badness \approx 49,99993$