Note til workshop 3:
Bestem kræfttypen med en matematisk algoritme

Ragnhild Laursen

23. august 2022

I noten her finder i en beskrivelse af Majorize-Minimization algoritmen, som også vil blive gennemgået til workshoppen. Denne note er mere detaljeret og kan derfor give jer et godt udgangspunkt til nemmere at kunne følge med til workshoppen. Vi skal differentiere en del til workshoppen, hvor vi skal bruge de differentieringsregler I kan finde i Tabel 1.

$f(x)$	$f'(x)$
k	0
x	1
$k \cdot x$	k
$k \cdot x^2$	$2 \cdot k \cdot x$

Tabel 1: Differentieringsregler vi skal bruge til workshoppen.

Da vi også skal programmere på workshoppen er der her et link der kan introducerer jer til noget af det vi skal bruge i løbet af workshoppen. Leg lidt rundt med de forskellige muligheder og se at det fungere meget som en normal lommeregner.

Majorize-Minimization algoritmen: Smart trick til minimering af funktioner med to variable

Majorize-Minimization (MM) er smart når man gerne vil finde minimum af en lidt besværlig funktion af flere variable. Idéen er at finde en nemmere funktion, som majoriserer (er større end) den oprindelige, og minimere denne i stedet. Før vi forklarer mere om denne metode, så lad os lige se på hvordan man normalt ellers kan minimere en funktion af én eller flere variable.

Minimum for en funktion af én variabel

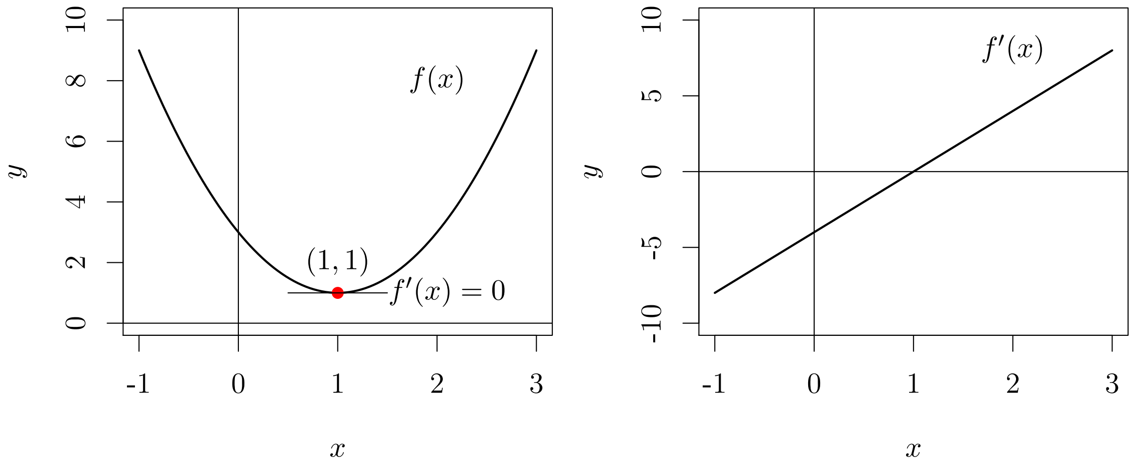
Når man normalt skal finde minimum af en funktion $f(x)$, som f.eks. den i Figur 1, så vil man finde den afledte af f og sætte den lig nul, $f'(x) = 0$. Det (De) x_0 som løser denne ligning vil være der hvor tangenten er vandret. Vi kan bagefter tjekke om løsningen er et minimum ved enten at se om den afledte er negativ før x_0 og positiv efter x_0, eller aflede funktionen igen og tjekke at den er positiv, $f''(x) > 0$. I Figur 1 har vi funktionen

$$f(x) = 2x^2 - 4x + 3$$

hvilket betyder at

$$f'(x) = 2 \cdot 2x - 4 = 4x - 4$$

som også er vist til højre i Figur 1.



Figur 1: Funktionen $f(x) = 2x^2 - 4x + 3$ og dens afledte $f'(x) = 4x - 4$.

For at finde minimum sættes den afledte funktion lig nul og vi løser for x.

$$\begin{aligned} f'(x_0) = 0 \quad &\Leftrightarrow \quad 4x_0 - 4 = 0 \\ &\Leftrightarrow \quad 4x_0 = 4 \\ &\Leftrightarrow \quad x_0 = \frac{4}{4} = \ 1. \end{aligned}$$

Vi kan samtidig se at

$$f(x_0) = 2 \cdot 1^2 - 4 \cdot 1 + 3 = 2 - 4 + 3 = 1.$$

Så vores minimum haves i $x_0 = 1$, hvor funktionens minimumsværdi er 1. Vi er sikre på at dette er et minimum fordi $f''(x) = 4 > 0$.

Minimum for en funktion af to variable

Når vi har en funktion af flere variable kan vi også nemt finde minimum, hvis vi blot kan splitte den op til funktioner af en enkelt variabel. Når vi har funktioner af en enkelt variabel har vi lige vist ovenfor hvordan vi kan finde minimum. Hvis vi f.eks. havde funktion $f(x, y)$, vil vi gerne kunne splitte den op som

$$f(x, y) = g(x) + h(y),$$

da vi således har at

$$\text{minimum af } f(x, y) = \text{minimum af } g(x) + \text{minimum af } h(y).$$

Et eksempel kunne være at

$$\begin{aligned} f(x, y) &= 3x^2 + 2y^2 - 2x - y + 6 \\ &= \underbrace{3x^2 - 2x}_{g(x)} + \underbrace{2y^2 - y + 6}_{h(y)} \\ &= g(x) + h(y). \end{aligned}$$

For at finde minimum af $f(x, y)$ finder vi derfor minimum af

$$g(x) = 3x^2 - 2x \quad \text{og} \quad h(y) = 2y^2 - y + 6$$

hver for sig. Minimum findes på samme måde som vi har set ovenfor ved at sætte den afledte lig nul. Bemærk at

$$g'(x) = 3 \cdot 2x - 2 = 6x - 2$$
$$h'(y) = 2 \cdot 2x - 1 = 4x - 1,$$

og når vi sætter de afledte lig nul så får vi hhv.

$$\begin{aligned} g'(x_0) = 0 \quad &\Leftrightarrow \quad 6x_0 - 2 = 0 \\ &\Leftrightarrow \quad 6x_0 = 2 \\ &\Leftrightarrow \quad x_0 = \frac{2}{6} = \frac{1}{3} \end{aligned}$$

og

$$\begin{aligned} h'(y_0) = 0 \quad &\Leftrightarrow \quad 4y_0 - 1 = 0 \\ &\Leftrightarrow \quad 4y_0 = 1 \\ &\Leftrightarrow \quad y_0 = \frac{1}{4}. \end{aligned}$$

Minimum er således i punktet $(x_0, y_0) = (\frac{1}{3}, \frac{1}{4})$, og funktionens minimumsværdi er

$$\begin{aligned} f(x_0, y_0) = f(\frac{1}{3}, \frac{1}{4}) &= 3 \cdot \left(\frac{1}{3}\right)^2 + 2 \cdot \left(\frac{1}{4}\right)^2 - 2 \cdot \frac{1}{3} - \frac{1}{4} + 6 \\ &= 3 \cdot \frac{1}{6} + 2 \cdot \frac{1}{16} - \frac{2}{3} - \frac{1}{4} + 6 \\ &= \frac{1}{2} + \frac{1}{8} - \frac{2}{3} - \frac{1}{4} + 6 \\ &= \frac{12}{24} + \frac{3}{24} - \frac{16}{24} - \frac{6}{24} + 6 \\ &= 6 - \frac{7}{24} \end{aligned}$$

Grunden til at vi ved, at dette er et minimum er igen fordi $g''(x) = 6 > 0$ og $h''(y) = 4 > 0$.

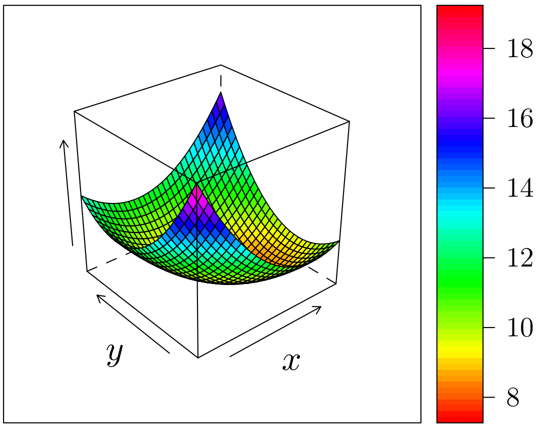
Majorize-Minimization algoritme

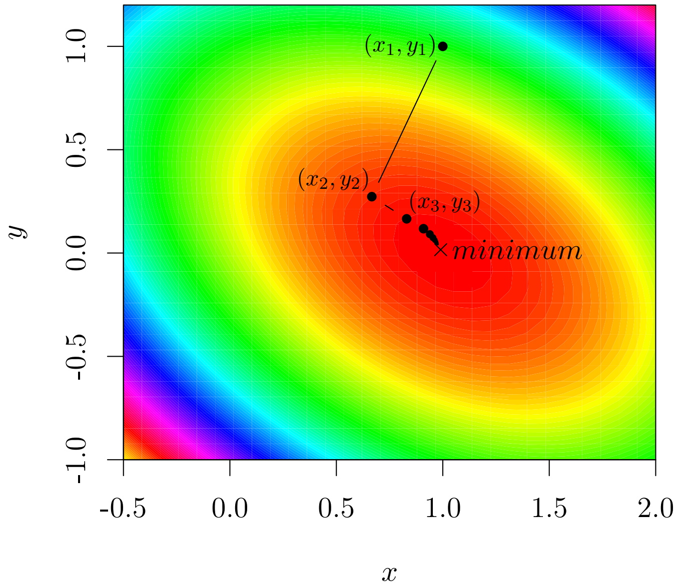
Det er dog ikke altid muligt at splitte en funktion af flere variable pænt op til funktioner af en enkelt variabel. Det kunne f.eks. være at vores funktionen så således ud

$$f(x, y) = 2x^2 - 4x + 2xy + 3y^2 - 2y + 10. \tag{1}$$

Her er leddet $2xy$ et problem, da det forhindrer os i at dele funktionen op. For at finde minimum for denne funktion kan vi bruge Majorize-Minimization algoritmen.

Den to-dimensionale funktion i (1) er vist i Figur 2, hvor man også kan se de forskellige bud på minimum igennem algoritmen til højre. Bemærk at med hvert nyt gæt er man rykket tættere på det sande minimum, som også er markeret i figuren.





Figur 2: Funktionen $f(x, y) = 2x^2 - 4x + 2xy + 3y^2 - 2y + 10$ og eksempel på MM-algoritmen, hvor man har startet i $(x_1, y_1) = (1, 1)$. Til venstre ses et 3D plot af funktionen og til højre ses et kontur plot af samme funktion.

Majoriserende funktion

Til MM-algoritmen skal vi først finde en smart funktion m der majoriserer f (altid er større end f) og samtidig er meget lettere at minimere. Det smarte ved den majoriserende funktion m er at den kan splittes op til to funktioner der kun afhænger af en enkelt variabel. Vi vil dog helst gerne have at den samtidig er tæt på f, så vi kan finde en løsning tæt på minimum. Derfor vil vi gerne have at m rør f i mindst ét punkt. Til at konstruere vores majoriserende funktion anvendes følgende ulighed for en given konstant $\alpha > 0$:

$$0 \leq \left(\sqrt{\alpha}x - \frac{1}{\sqrt{\alpha}}y\right)^2 = \alpha x^2 + \frac{1}{\alpha}y^2 - 2xy$$

som ved at rykke lidt rundt giver

$$2xy \leq \alpha x^2 + \frac{1}{\alpha}y^2.$$

Bemærk at vi har en ulighed hvor x og y går fra at være ganget sammen til at være splittet op. Det smarte ved denne ulighed er også at det ændres til en lighed hvis $\alpha = \frac{y}{x}$ (Prøv at regn det efter). Ved at bruge denne ulighed kan vi definere funktionen m som majoriserer f ved at erstatte $2xy$ med $\alpha x^2 + \frac{1}{\alpha}y^2$.

$$\begin{aligned} m(x, y) &= 2x^2 - 4x + \alpha x^2 + \frac{1}{\alpha}y^2 + 3y^2 - 2y + 10 \\ &= \underbrace{(2 + \alpha)x^2 - 4x}_{\text{Afhænger kun af x}} + \underbrace{(3 + \frac{1}{\alpha})y^2 - 2y + 10}_{\text{Afhænger kun af y}} \qquad (2) \\ &= g(x) + h(y) \end{aligned}$$

hvor

$$g(x) = (2+\alpha)x^2 - 4x$$
$$h(y) = (3+\frac{1}{\alpha})y^2 - 2y + 10.$$

I alt har vi

$$m(x,y) \geq f(x,y) \quad \text{for alle } x \text{ og } y$$

med lighed hvis og kun hvis $\alpha = \frac{y}{x}$.

Algoritme med start i $(\mathbf{x_1}, \mathbf{y_1}) = (\mathbf{1}, \mathbf{1})$

Vi starter algoritmen med at vælge to tilfældige positive tal til punktet $(x_1, y_1) = (1,1)$, som er vores første bud på hvor minimum er. I dette punkt vil vi gerne have at m og f er lig hinanden og sætter derfor $\alpha_1 = \frac{y_1}{x_1} = 1$ i (2). Vores første majoriserende funktion er dermed

$$\begin{aligned} m_1(x,y) &= (2+1)x^2 - 4x + (3+\frac{1}{1})y^2 - 2y + 10 \\ &= \underbrace{3x^2 - 4x}_{g_1(x)} + \underbrace{4y^2 - 2y + 10}_{h_1(y)} \\ &= g_1(x) + h_1(y). \end{aligned}$$

I toppen af Figur 3 har vi et tværsnit af $f(x,y)$ fra Figur 2 og den majoriserende funktion $m_1(x,y)$. Til venstre holdes $y = 1$ fast og til højre holdes $x = 1$ fast. Bemærk at den majoriserende funktion for de to tværsnit er givet ved

$$\begin{aligned} m_1(x,1) &= g_1(x) + h_1(1) = g_1(x) + 12 \\ m_1(1,y) &= g_1(1) + h_1(y) = h_1(y) - 1 \end{aligned}$$

da $g_1(1) = -1$ og $h_1(1) = 12$. Minimum af den majoriserende funktion $m_1(x,y)$ findes som ovenfor ved at aflede de to funktioner $g_1(x)$ og $h_1(y)$, sætte dem lige nul og isolere variablen:

$$\begin{aligned} g_1'(x) &= 6x - 4 \quad \Rightarrow \quad x_0 = \frac{4}{6} = \frac{2}{3} \\ h_1'(y) &= 8y - 2 \quad \Rightarrow \quad y_0 = \frac{2}{8} = \frac{1}{4}. \end{aligned}$$

Det giver os et nyt bud på minimum i punktet $(x_2, y_2) = (\frac{2}{3}, \frac{1}{4}) = (0.67, 0.25)$, som både kan ses i Figur 2 og 3. Bemærk i begge figurer at dette nye bud er tættere på det sande minimum af $f(x,y)$ end (x_1, y_1).

Vi vælger nu at sætte $\alpha_2 = \frac{y_2}{x_2} = \frac{1}{4}/\frac{2}{3} = \frac{3}{8}$ i (2), så vi nu har den majoriserende funktion

$$\begin{aligned} m_2(x,y) &= \underbrace{(2+\frac{3}{8})x^2 - 4x}_{g_2(x)} + \underbrace{(3+\frac{8}{3})y^2 - 2y + 10}_{h_2(y)} \\ &= g_2(x) + h_2(y) \end{aligned}$$

De to tværsnit, hvor vi holder $x = \frac{2}{3}$ og $y = \frac{1}{4}$ fast er vist nederst i Figur 3. De to tværsnit af den majoriserende funktion nederst i Figur 3 er givet ved

$$\begin{aligned} m_2(x, \frac{1}{4}) &= g_2(x) + h_2(\frac{1}{4}) = g_2(x) + 9.85 \\ m_2(\frac{2}{3}, y) &= g_2(\frac{2}{3}) + h_2(y) = h_2(y) - 1.6. \end{aligned}$$

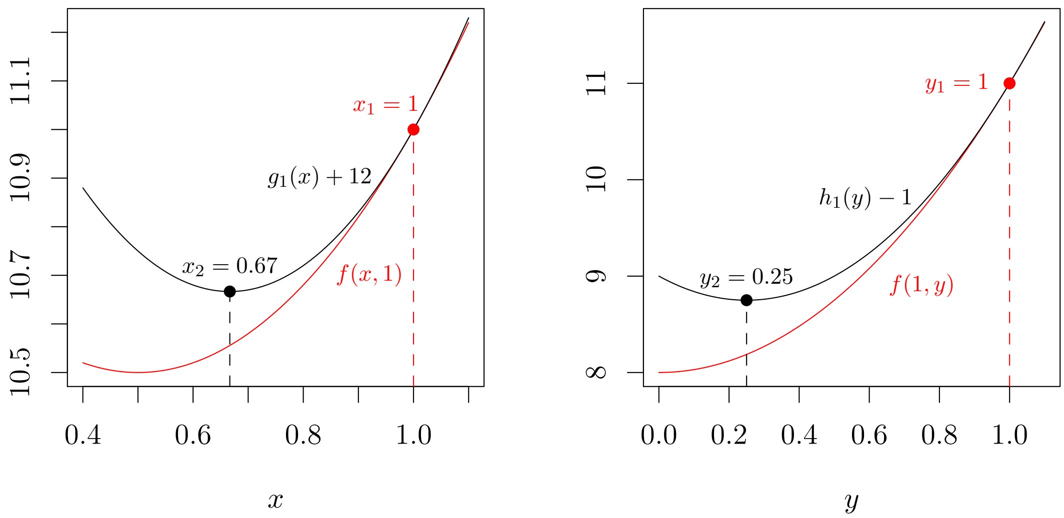
Minimum for denne nye majoriserende funktion er fundet i punktet $(x_3, y_3) = (0.84, 0.18)$, som er det nye bud på hvor minimum ligger. Igen kan vi se på Figur 2 at dette nye punkt ligger tættere på det sande minimum af $f(x, y)$. I Figur 2 har vi også vist punkterne for de næste mange bud på hvor minimum ligger. Det vigtige er at vi i hvert step rykker tættere og tættere på vores sande minimum i.e.

$$f(x_1, y_1) = m_1(x_1, y_1) \geq \min_{x,y} m_1(x, y) = m_1(x_2, y_2) \geq f(x_2, y_2),$$

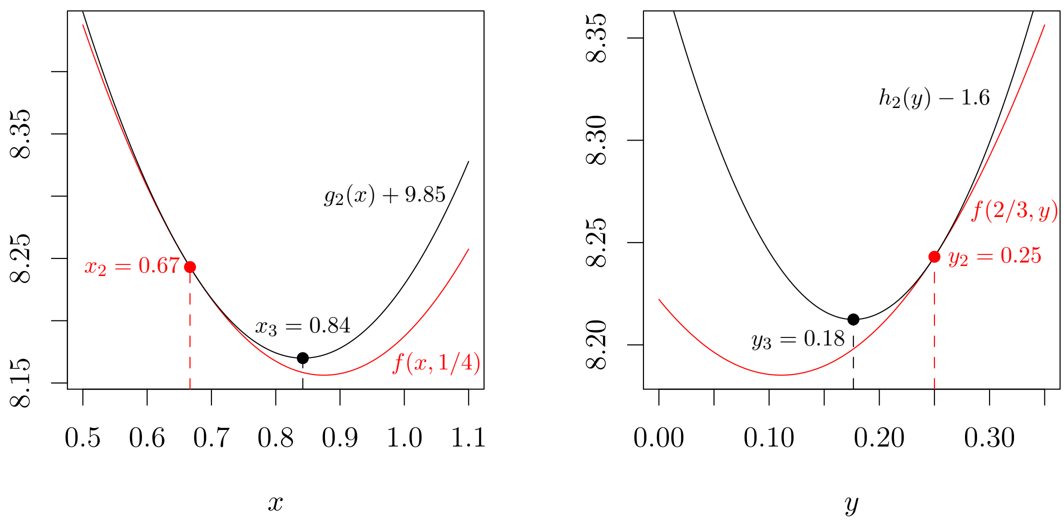
og dermed fås

$$f(x_1, y_1) \geq f(x_2, y_2) \geq f(x_3, y_3) \geq \dots$$

Vi vil fortsætte indtil funktionsværdien ikke bliver mindre længere.



(a) $(x_1, y_1) = (1, 1) \rightarrow (x_2, y_2) = (0.67, 0.25)$



(b) $(x_2, y_2) = (0.67, 0.25) \rightarrow (x_3, y_3) = (0.84, 0.18)$

Figur 3: Visualisering af MM-algoritmen i 2D for funktionen $f(x, y) = 2x^2 - 4x + 2xy + 3y^2 - 2y + 10$.

Generel algoritme

Vi kan også formulere en mere generel algoritme. Hvis vi bare tænker på α som et vilkårligt tal, så kan vi aflede $g(x)$ og $h(y)$ i (2) til

$$g'(x) = 2 \cdot (2+\alpha)x - 4 \quad \text{og} \quad h'(y) = 2 \cdot (3 + \frac{1}{\alpha})y - 2.$$

For et vilkårligt $\alpha > 0$ kan vi således finde et $x_{\min}$ som minimerer $g(x)$ og et $y_{\min}$ som minimerer $h(y)$

$$x_{\min} = \frac{2}{2+\alpha} \quad \text{og} \quad y_{\min} = \frac{1}{3+\frac{1}{\alpha}}$$

Prøv f.eks. at indsætte $\alpha_1 = 1$ og bemærk at $(x_{\min}, y_{\min}) = (\frac{2}{3}, \frac{1}{4})$, som er lig punktet (x_2, y_2) vi fandt ovenfor. For et vilkårligt positivt start sted kan vi dermed definere en generel algoritme til at finde minimum for $f(x, y)$ i (1). Dette er opsummeret i Algorithm 1.

Algorithm 1: Majorize-Minimization for $f(x, y)$

Vælg vilkårligt start sted (x_1, y_1)
for $t = 1, 2, \ldots$ **do**
  Sæt $\alpha_t = \frac{y_t}{x_t}$
  Find (x_{t+1}, y_{t+1}) som minimerer hhv. $g_t(x)$ og $h_t(y)$ ved at sætte
  $x_{t+1} = \frac{2}{2+\alpha_t}$
  $y_{t+1} = \frac{1}{3+\frac{1}{\alpha_t}}$
  stop algoritme hvis $f(x_t, y_t)$ er meget tæt på $f(x_{t+1}, y_{t+1})$.
