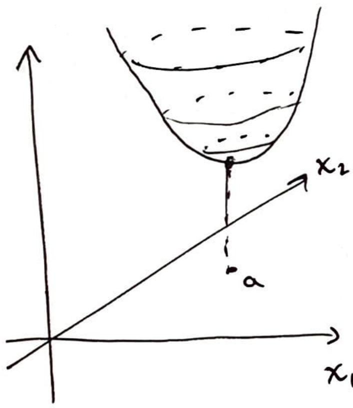
Uke 36: ~~[illegible]~~ Optimering

26. august 2024

1) Klassifisering av kritiske punkter ($\nabla f = 0$)
   → Hessematrisen og dens egenverdier
2) Finne maks/min langs en kurve
   → Parametrisering
   → Lagrange multiplikatorer

~~[illegible]~~ Klassifisering av kritiske punkter

Situasjon: Vi har funnet ~~[illegible]~~ et punkt a slik at $\nabla f(a) = 0$ ⇝ ønsker å vite om a er et toppunkt, bunnpunkt eller sadelpunkt.



Idé: hva skjer med funksjonsverdiene rundt a?
⇝ må sjekke mange retninger.

Parametriserer en vilkårlig rett linje gjennom a: $a + tv$, $t \in \mathbb{R}$, retning $v \in \mathbb{R}^2$

La $g(t) = f(a + tv)$ ← funksjonsverdiene rundt $f(a)$ for små t.

⇝ Taylorutvikler g rundt $t = 0$

$$g(t) = g(0) + g'(0) \cdot t + g''(0)\frac{t^2}{2} + \underbrace{\dots}_{\text{neglisjerbar for små } t}$$

⇝ oversette til f:

~~$f(a+tv) = f(a) + \nabla f(a) \cdot vt$~~

$g'(t) = \nabla f(a + tv) \cdot v$ ⇝ $g'(0) = \nabla f(a) \cdot v$

~~[illegible]~~

$$g''(t) = \frac{d}{dt}\left(\nabla f(a + tv) \cdot v\right)$$

$$\boxed{f(x) = \sum_{n=0}^{\infty} f^{(n)}(a) \frac{(x-a)^n}{n!} \quad \text{Taylor}}$$

f er en funksjon av to variabler som videre er en funksjon av t

$$\frac{d}{dt}\left(\nabla f(a+tv)\cdot v\right) = \frac{d}{dt}\left(\left(\frac{\partial f}{\partial x_1}\Big|_{x=a+tv}, \frac{\partial f}{\partial x_2}\Big|_{x=a+tv}\right)\cdot v\right)$$

La $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ $\quad x_1 = a_1 + tv_1$, $\; x_2 = a_2 + tv_2$

$$\frac{d}{dt}\left(v_1 \frac{\partial f}{\partial x_1} + v_2 \frac{\partial f}{\partial x_2}\right)$$

$\frac{\partial^2 f}{\partial x_2 \partial x_1}$ $\qquad$ $\frac{\partial^2 f}{\partial x_1 \partial x_2}$

$$= v_1 \frac{\partial^2 f}{\partial x_1^2}\cdot\frac{dx_1}{dt} + v_1 \cancel{\frac{\partial^2 f}{\partial x_1 \partial x_2}}\frac{dx_2}{dt} + v_2 \frac{\partial^2 f}{\partial x_2^2}\frac{dx_2}{dt} + v_2 \cancel{\frac{\partial^2 f}{\partial x_2 \partial x_1}}\frac{dx_1}{dt}$$

$$= v_1\left(\frac{\partial^2 f}{\partial x_1^2}\cdot v_1 + \frac{\partial^2 f}{\partial x_2 \partial x_1} v_2\right) + v_2\left(\frac{\partial^2 f}{\partial x_2^2}\cdot v_2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}\cdot v_1\right)$$

$$= (v_1 \;\; v_2)\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} v_1 + \frac{\partial^2 f}{\partial x_2 \partial x_1} v_2 \\ \frac{\partial^2 f}{\partial x_2^2} v_2 + \frac{\partial^2 f}{\partial x_1 \partial x_2} v_1 \end{pmatrix}$$

Kjenner igjen skalar- og matriseprodukter

(Jeg ~~...~~ rotet litt med indekser her i forelesning og, men ~~...~~ vi antar uansett at hessematrisen er symmetrisk, så det går fint)

$$= (v_1 \;\; v_2)\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

← hessematrisen til f

$$= v^T H v$$

(altså vi antar $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$ →

$$g(t) = g(0) + g'(0)\,t + g''(0)\,\frac{t^2}{2} + \dots$$

f'', kvadratisk form

$$f(a+tv) = f(a) + \underbrace{\nabla f(a)\cdot vt}_{=0,\text{ siden } a \text{ er kritisk punkt } (\nabla f(a)=0)} + v^T H v\,\frac{t^2}{2} + \dots$$

$$\approx f(a) + \underbrace{v^T H v\,\frac{t^2}{2}}$$

$> 0 \Rightarrow$ funksjonsverdier rundt a øker
⇝ a er bunnpunkt

$< 0 \Rightarrow$ —"— minker
⇝ a er toppunkt

$\Rightarrow$ hvis H er positivt definitt ($v^T H v > 0 \ \forall v \in \mathbb{R}^n, v \neq 0$)

så blir $f(a+tv) > f(a)$

$\Rightarrow$ funksjonsverdiene rundt a øker

$\Rightarrow$ a er et bunnpunkt

$\Rightarrow$ hvis H er negativt definitt ($v^T H v < 0 \ \forall v \in \mathbb{R}^n \ v \neq 0$)

så blir $f(a+tv) < f(a)$

$\Rightarrow$ funksjonsverdiene rundt a minker

$\Rightarrow$ a er et toppunkt

Så: vise at symmetrisk matrise: positivt definitt

$\Leftrightarrow$ alle egenverdier er positive.

"$\Rightarrow$" Anta A er symmetrisk og positivt definitt.

For en egenverdi λ med egenvektor u

får vi at $\underbrace{0 < u^T A u}_{\text{siden } A \text{ positivt definitt}} = u^T \lambda u = \lambda \underbrace{u^T u}_{>0} \Rightarrow \lambda > 0$

"$\Leftarrow$" Anta A symmetrisk og alle egenverdier positive

A symmetrisk $\Rightarrow$ fins ortonormal egenvektorbasis

$\Rightarrow$ alle vektorer $v \in \mathbb{R}^n$ kan skrives som

$$v = \sum_k c_k u_k \quad \text{for ortonormale egenvektorer } \{u_i\}$$

$$v^T A v = \left(\sum_k c_k u_k\right)^T A \left(\sum_k c_k u_k\right) = \left(\sum_k c_k u_k\right)^T \left(\sum_k c_k A u_k\right)$$

$$= \left(\sum_k c_k u_k\right)^T \left(\sum_k c_k \lambda_k u_k\right) = \sum_k \underbrace{|c_k|^2}_{>0} \underbrace{\lambda_k}_{>0} \underbrace{u_k^T u_k}_{=1} > 0$$

$u_i^T u_j = 0$ for $i \neq j$

av antagelse

siden ortonormale

(Tilsvarende: A er negativt definitt ($v^T A v < 0 \quad \forall \; v \in \mathbb{R}^n \setminus \{0\}$) hvis og bare hvis alle egenverdiene er negative)

KONKLUSJON:

- Alle egenverdiene til H er positive

  $\Rightarrow$ a er bunnpunkt

- Alle egenverdiene til H er negative

  $\Rightarrow$ a er toppunkt

- Både positive og negative $\rightarrow$ funksjonen øker i noen retninger, synker i andre

  $\Rightarrow$ a er et sadelpunkt

- Hvis en egenverdi er 0, har vi ingen konklusjon.

2) $f(x_1, x_2) = x_1^4 + x_2^2 + 4x_1 + 3x_2 - 4$

$x_1^4 + x_2^2 + 4x_1 + 3x_2 - 4$

Finne kritisk punkt:

$\nabla f = (4x_1^3 + 4,\ 2x_2 + 3)$

$x_1^3 = -1 \Rightarrow x_1 = -1$

$\nabla f = 0 \Rightarrow \begin{matrix} 4x_1^3 + 4 = 0 \\ 2x_2 + 3 = 0 \end{matrix} \Rightarrow x_2 = -\frac{3}{2}$

Kritisk punkt i $a = (-1, -\frac{3}{2})^T$

⇝ Finne Hessematrisen

$\frac{\partial^2 f}{\partial x_1^2} = 12x_1^2 \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right) = 0$

$\frac{\partial^2 f}{\partial x_2^2} = 2 \qquad \frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = 0$

$\Rightarrow H = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{pmatrix}$ i pkt a: $H = \begin{pmatrix} 12 & 0 \\ 0 & 2 \end{pmatrix}$

⇝ finne egenverdier:

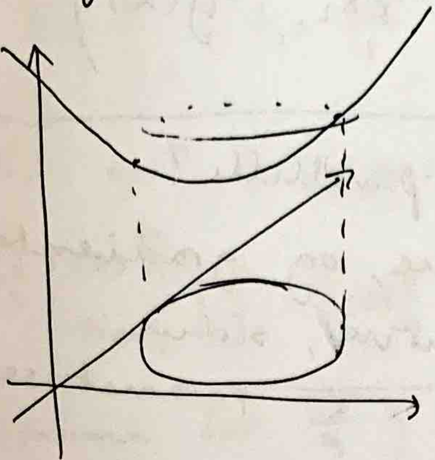
$\begin{vmatrix} 12-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = (12-\lambda)(2-\lambda) \Rightarrow \lambda = 2,\ \lambda = 12$

⇝ kun positive ⇒ H er positivt definitt

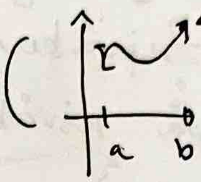
⇒ a er et bunnpunkt.

2) Finne maks/min langs kurve

Noen ganger har vi funksjoner definert på et begrenset område



⇝ kan ha ekstremalverdier langs "randen"

(

← største verdi på intervallet
⇝ må sjekke $x=a$ og $x=b$)

To metoder:

- Parametrisering
  ⇝ parametrisere randen og putte inn i f.
- Lagrange-multiplikatorer

Lagrange: ønsker å minimere/maksimere

$f(x)$ gitt $g(x) = 0$ ← likning for rand

↖ funksjon vi bryr oss om

∇f og ∇g er parallelle i maks/min langs rand

$\Rightarrow \nabla f = \lambda \nabla g$ λ konstant

⇝ løse $\begin{cases} \dfrac{\partial f}{\partial x_1} = \lambda \dfrac{\partial g}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} = \lambda \dfrac{\partial g}{\partial x_2} \\ g(x) = 0 \end{cases}$ → tre likninger med tre ukjente

dette er det samme som å lete etter kritiske punkter til Lagrangefunksjonen

$$L(x_1, x_2, \lambda) = f(x) - \lambda g(x)$$ ~~[illegible]~~ noen skriver +

Altså $\nabla L = 0$, siden $\frac{\partial L}{\partial \lambda}$

$$\nabla L = \left(\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1}, \frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2}, -g(x) \right)$$

Hvorfor er ∇f og ∇g parallelle?

$g(x) = 0$ er en nivåkurve, og gradienter er ortogonale på nivåkurver, siden

parametrisering

$$g(x_1(t), x_2(t)) = c$$

⇝ deriverer mhp t

$$\underbrace{\frac{\partial g}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial g}{\partial x_2}\frac{dx_2}{dt} = 0}$$

$$\left(\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}\right) \cdot \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}\right) = 0$$

$$\Rightarrow \nabla g \cdot \dot{x}(t) = 0$$ tangenten

∇f er ortogonal ~~på [illegible]~~ langs randen når vi har en maks/min, fordi ellers vil vi kunne bevege oss langs g og øke/minke f.

3) sjekke inni ⇝ finne kritiske punkter

$$\nabla f = (2x_2, 2x_1) \leadsto \hookrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = 0$$

⇝ har vi større verdi langs randen?

a) Randen er gitt ved $x_1^2 + x_2^2 = 1$

La $x_1 = \cos t$, $x_2 = \sin t$ $\quad t \in [0, 2\pi]$

$$f\left(\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\right) = 2\cos t \sin t = \sin 2t$$

$\left(\text{kan derivere og sette lik } 0 \text{ som vi er vant med fra vgs}\right)$

$\sin 2t$ har maksverdi når

$t = \frac{\pi}{4}$, $t = \frac{5\pi}{4}$

det gir funksjonsverdiene

$$f\left(\begin{pmatrix}\cos\frac{\pi}{4}\\ \sin\frac{\pi}{4}\end{pmatrix}\right) = 2\cos\frac{\pi}{4}\sin\frac{\pi}{4} = 2\cdot\frac{\sqrt{2}}{2}\cdot\frac{\sqrt{2}}{2} = 1$$

(samme for $\frac{5\pi}{4}$)

b) $\nabla f = (2x_2, 2x_1)$ $\nabla g = (2x_1, 2x_2)$

hvor $g(x) = x_1^2 + x_2^2 - 1$

$\nabla f = \lambda \nabla g \Rightarrow (2x_2, 2x_1) = \lambda(2x_1, 2x_2)$

$$\begin{cases} 2x_2 = \lambda 2x_1 \\ 2x_1 = \lambda 2x_2 \\ x_1^2 + x_2^2 = 1 \end{cases}$$

løser for x_1, x_2, λ

$2x_2 = \lambda \cdot \lambda 2x_2$

$\Rightarrow x_2 = 0$ eller $\lambda^2 = 1$

$x_2 = 0$ gir $x_1 = 0$, men $(0,0)$ oppfyller ikke $g(x) = 0$

så $\lambda^2 = 1$

$\lambda = 1$ gir $x_2 = x_1 \Rightarrow 2x_1^2 = 1$ $x_1 = x_2 = \pm\frac{1}{\sqrt{2}}$

$\lambda = -1$ gir $x_2 = -x_1 \Rightarrow 2x_1^2 = 1$ $x_1 = \pm\frac{1}{\sqrt{2}}$, $x_2 = \mp\frac{1}{\sqrt{2}}$

$\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$ og $\left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}\right)$ $\leadsto$ putter inn i f

~~størst~~ største funksjonsverdi blir 1.