3 - 3 - OPTIMERING I

$\boxed{1}$ Vi setter gradienten til h lik null:

$$-2x_1 - x_2 + 4 = 0$$
$$-x_1 - 4x_2 + 6 = 0$$

som løses av $x_2 = 8/7$, $x_1 = 10/7$.

$\boxed{2}$ Hessematrisen er

$$h''(x) = \begin{pmatrix} -2 & -1 \\ -1 & -4 \end{pmatrix}$$

som har karakteristisk polynom

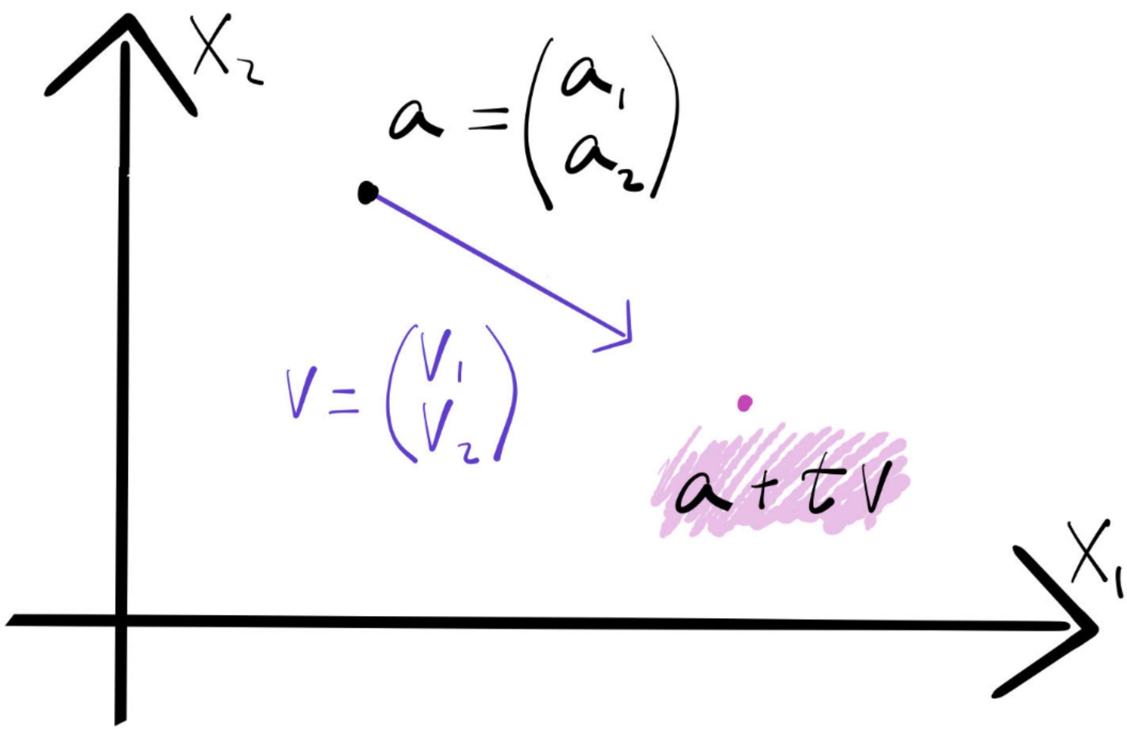
$$\det \begin{pmatrix} -2-\lambda & -1 \\ -1 & -4-\lambda \end{pmatrix} = (2+\lambda)(4+\lambda) - 1 = 7 + 6\lambda + \lambda^2$$

og egenverdier

$$\lambda = -3 \pm \sqrt{2}$$

som begge er negative, og følgelig kan vi være sikker at dette er en fjelltopp og ikke en dalbunn.

$\boxed{3}$ La oss sende en rett linje gjennom punktet a slik:



Parametriseringen for den rette linjen er

$$x(t) = a + tv = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad t \in \mathbb{R}$$

La oss definere

$$g(t) = f(x(t)) = f(a + tv)$$

og så bruke kjerneregelen på g, og få

$$\begin{aligned} g'(t) =& f'(x(t))x'(t) \\ =& f'(a+tv)v \\ =& \left(\frac{\partial f}{\partial x_1}(a+tv), \frac{\partial f}{\partial x_2}(a+tv) \right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ =& v_1 \frac{\partial f}{\partial x_1}(a+tv) + v_2 \frac{\partial f}{\partial x_2}(a+tv). \end{aligned}$$

Nå kan gjenta denne suksessen og derivere dette uttrykket med kjerneregelen, og få

$$\begin{aligned} g''(t) =& v_1 \frac{d}{dt} \left(\frac{\partial f}{\partial x_1}(a+tv) \right) + v_2 \frac{d}{dt} \left(\frac{\partial f}{\partial x_2}(a+tv) \right) \\ =& v_1 \left(v_1 \frac{\partial^2 f}{\partial x_1^2}(a+tv) + v_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(a+tv) \right) + v_2 \left(v_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(a+tv) + v_2 \frac{\partial^2 f}{\partial x_1^2}(a+tv) \right) \end{aligned}$$

Hvis du nå skriver ut

$$\begin{aligned} v^T A v =& \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ =& v_1^2 a_1 1 + v_1 v_2 a_{12} + v_1 v_2 a_{12} + v_2^2 a_{22}, \end{aligned}$$

ser du at vi kan skrive

$$g''(t) = v^T H(a+tv)v$$

der

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}$$

Vi kan nå taylorutvikle g:

$$\begin{aligned} g(t) =& g(0) + g'(0)t + g''(0)\frac{t^2}{2} + \cdots \\ =& f(a) + (f'(a)\,v)\,t + \left(v^T f''(a)\,v\right)\frac{t^2}{2} + \cdots \end{aligned}$$

Merk ellers at

$$g'(0) = f'(a)v$$

som vi kjenner igjen som den retningsderiverte til f i a.

$\boxed{4}$ La a være et av f sine kritiske punkt. Vi tar en motorsag og snitter flaten til $f(x)$ langs den parametriserte linjen $a+tv$, og så taylorutvikler vi motorsagsnittfunksjonen $g(t)=f(a+tv)$ om $t=0$, slik som i forrige oppgave. Siden $f'(a)=0$ dersom a er et kritisk punkt, blir taylorutviklingen

$$f(a+tv)=f(a)+\left(v^T f''(a)\,v\right)\frac{t^2}{2}+\cdots$$

Dersom t er bitteliten og f oppfører seg bra nok til at $\left(v^T f''(a)\,v\right)\frac{t^2}{2}$ fullstendig dominerer alle leddene som kommer etter, kan vi nå konkludere med følgende tabell. La λ_1 og λ_2 være egenverdiene til $f''(a)$.

- Dersom λ_1 og λ_2 er positive, er a et bunnpunkt.
- Dersom λ_1 og λ_2 er negative, er a et toppunkt.
- Dersom λ_1 og λ_2 har forskjellig fortegn, er a et sadelpunkt.
- Dersom en av dem er null, gir testen ingen konklusjon.

Hvordan kan vi være så freidige? Siden

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

for alle funksjoner du kommer til å plages med[1] er $f''(a)$ symmetrisk, og følgelig ortogonalt diagonaliserbar. La u_1 og u_2 være ortogonale egenvektorer til $f''(a)$. Disse utgjør en basis for $\mathbb{R}^2$, så vi kan alltid skrive

$$v=c_1u_1+c_2u_2$$

og regne ut den kvadratiske formen

$$\begin{aligned} v^T f''(a)\,v &= (c_1u_1+c_2u_2)^T f''(a)\,(c_1u_1+c_2u_2) \\ &= (c_1u_1+c_2u_2)^T(\lambda_1c_1u_1+\lambda_2c_2u_2) \\ &= \lambda_1c_1^2|u_1|^2+\lambda_2c_2^2|u_2|^2 \end{aligned}$$

Dette uttrykket er helt klart positivt for alle v i det første tilfellet, og konklusjonen følger siden vi følgelig alltid legger noe til $f(a)$ i nærheten av a når vi beveger oss bittelitt vekk fra a. Samme resonnement dersom egenverdiene er negative. Dersom egenverdiene har forskjellig fortegn, kan vi bruke samme resonnement på $v=u_1$ og $v=u_2$ og konkludere med at det finnes en retning der vi går nedover og en retning der vi går oppover, slik at a er en sadel. Hvis en av egenverdiene er null, finnes det en retning der $v^T f''(a)\,v=0$, og i så fall trenger vi flere ledd i taylorrekken for å konkludere. Hvis begge egenverdier er null, er f et lineært polynom, og da gir det ikke mening å lete etter maks- eller minimumspunkter.

Dersom alle egenverdier er positive, sier vi ellers at matrisen er positivt definitt. Tilsvarende for negative egenverdier.

[1] Det finnes patologiske unntak, slik som for eksempel $\frac{xy(x^2-y^2)}{x^2+y^2}$, der de blandede andreordens partiellderiverte ikke er like, men disse er konstruert av matematikere for å krangle med folk som er uforsiktige nok til å påstå at de blandede andreordens partiellderiverte alltid er like. Se ukens nøtter.

$\boxed{5}$ Du må lete etter maksima og minima til de tre envariable funksjonene

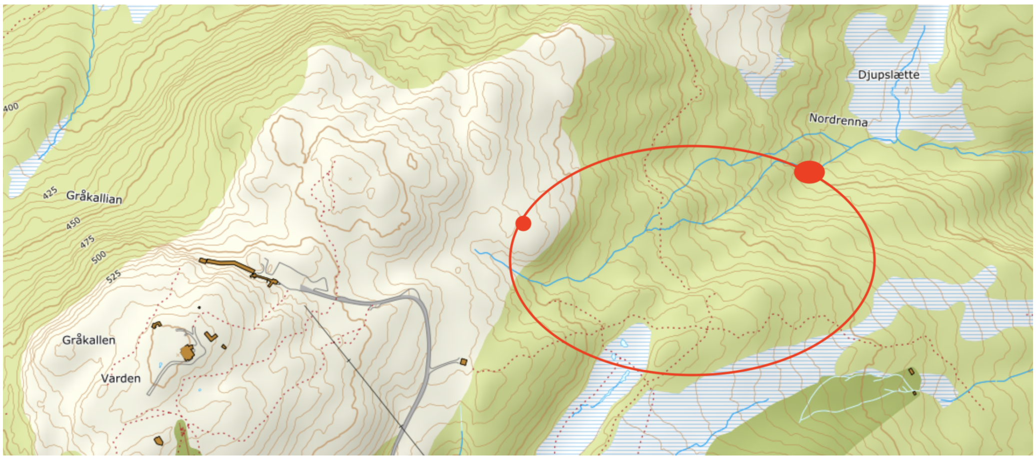
$$h(x_1, 0) = -x_1^2 + 4x_1 - 7 \qquad x_1 \in [0, 4]$$

$$h(0, x_2) = -2x_2^2 + 6x_2 - 7 \qquad x_2 \in [0, 3]$$

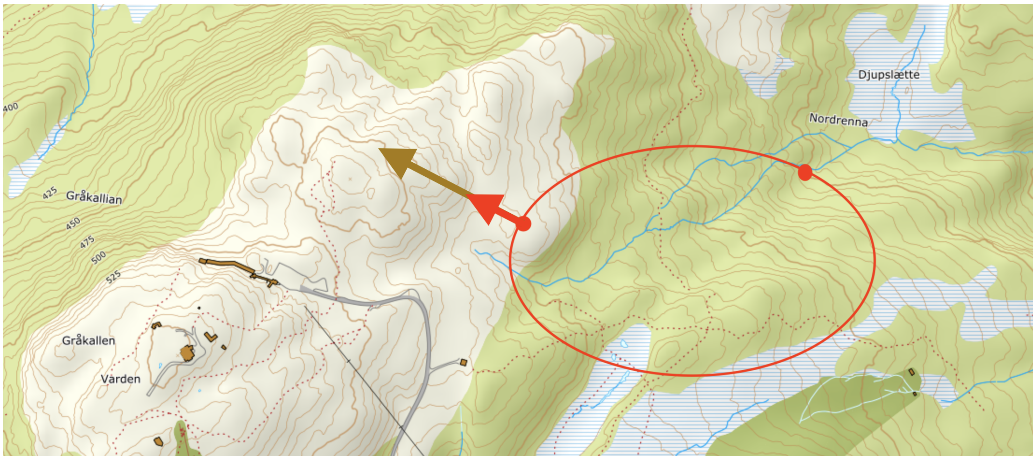
$$h(4\cos\theta, 3\sin\theta) = -16\cos^2\theta - 12\cos\theta\sin\theta - 18\sin^2\theta + 16\cos\theta + 18\sin\theta + 8 \qquad \theta \in [0, \pi/2]$$

Det er trivielt å se at de to første har globale minima i endepunktene og globale maksima i kritiske punkter. Den siste er uhåndterlig med penn og papir, men håndterbar om du har numerisk likningsløser tilgjengelig.

$\boxed{6}$ Lagranges multiplikatormetode er basert på en ide som er veldig enkel å forstå om man har det klart for seg hvordan gradienten til en funksjon skal tolkes geometrisk. Den er også til god hjelp dersom man ønsker en enkel huskeregel nettopp for å huske gradientens geometriske tolkning. Hvis jeg sier at man går en tur på den elliptiske kurven i figuren under, klarer de aller fleste uten å tenke seg om merke av turens høyeste og laveste punkt:



Nå gjelder å tenke på f (altså den funksjonen du skal maksimere eller minimere) som en funksjon som beskriver fjellet du går på. Nivåkurvene til f er det som er printa som ekvidistanselinjene på et vanlig turkart. Den snedige ideen til lagrange var å se på turtrajektorien som nivåkurven til en eller annen annen funksjon g, og så innse at i f sine topp-, bunn- eller sadelpunkter **må gradienten til f være parallell med gradienten til g, siden gradienter og nivåkurver står normalt på hverandre.**



Denne ideen gir tre likninger

$$f'(x) = \lambda g'(x)$$
$$g(x) = c$$

som kan løses for x og λ. Det er litt forskjellige betingelser ute og går; f og g må være deriverbare, og gradientene kan ikke være null, og du må ikke være i endepunktet for trajektorien og litt forskjellig.

7 Vi kan nå prøve oss på den krumme delen av $\partial\Omega$, som har likning

$$\frac{x_1^2}{16} + \frac{x_2^2}{9} = 1$$

slik at

$$g(x) = \frac{x_1^2}{16} + \frac{x_2^2}{9}.$$

Gradienten til g er

$$g'(x) = \left(\frac{x_1}{8}, \frac{2x_2}{9}\right)$$

mens gradienten til h regnet vi ut i oppgave 1:

$$f'(x) = (-2x_1 - x_2 + 4, -x_1 - 4x_2 + 6)$$

Alt i alt får vi likningssystemet

$$-2x_1 - x_2 + 4 = \lambda \frac{x_1}{8}$$

$$-x_1 - 4x_2 + 6 = \lambda \frac{2x_2}{9}$$

$$\frac{x_1^2}{16} + \frac{x_2^2}{9} = 1$$

De to første likningene kan skrives om til

$$(\lambda + 16)x_1 + 8x_2 = 32$$

$$9x_1 + 2(\lambda + 18)x_2 = 54$$

og i dette tilfellet synes jeg det er kjappest å bare invertere matrisen og skrive

$$\begin{aligned} x =& \frac{1}{2(\lambda+16)(\lambda+18)-72} \begin{pmatrix} 2(\lambda+18) & -8 \\ -9 & (\lambda+16) \end{pmatrix} \begin{pmatrix} 32 \\ 54 \end{pmatrix} \\ =& \frac{1}{(\lambda+16)(\lambda+18)-72} \begin{pmatrix} 2(\lambda+18) & -8 \\ -9 & (\lambda+16) \end{pmatrix} \begin{pmatrix} 16 \\ 27 \end{pmatrix}. \end{aligned}$$

Setter vi denne inn i ellipselikningen, får vi en fjerdegradslikning som er uhåndterlig med penn og papir, men triviell å løse med en numerisk likningsløser.

$\boxed{8}$ Gradienten til L er

$$\begin{aligned} L'(x) &= \left(\frac{\partial L}{\partial x_1}(x), \frac{\partial L}{\partial x_2}(x), \frac{\partial L}{\partial x_\lambda}(x) \right) \\ &= \left(\frac{\partial f}{\partial x_1}(x) + \lambda \frac{\partial g}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x) + \lambda \frac{\partial g}{\partial x_2}(x), g(x) \right) \end{aligned}$$

og setter vi denne lik null, får vi lagrangelikningene dersom trajektorien er gitt ved $g(x) = 0$.

$\boxed{9}$ Gradienten til fjellet er

$$h'(x) = (-2x_1 - x_2, -x_1 - 2x_2)$$

og setter vi

$$g(x) = x_1^2 + \frac{x_2^2}{4} - 1,$$

blir

$$g'(x) = (2x_1, x_2/2)$$

Lagrangelikningene blir

$$\begin{aligned} -2x_1 - x_2 &= 2\lambda x_1 \\ -x_1 - 2x_2 &= \lambda x_2/2 \end{aligned}$$

eller

$$\begin{aligned} 2(1+\lambda)x_1 + x_2 &= 0 \\ 2x_1 + (4+\lambda)x_2 &= 0 \end{aligned}$$

Dersom matrisen er inverterbar, er $x = 0$ den eneste løsningen, og dette punktet ligger ikke på ellipsen, så da får vi nesten kreve at determinanten

$$p(\lambda) = 2(1+\lambda)(4+\lambda) - 2 = 2\lambda^2 + 10\lambda + 6$$

blir null, noe som skjer når

$$\lambda = \frac{-5 \pm \sqrt{13}}{2}$$

Likningen $2(1+\lambda)x_1 + x_2 = 0$ gir i såfall at

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s \begin{pmatrix} 1 \\ -2(1+\lambda) \end{pmatrix} = s \begin{pmatrix} 1 \\ 3 \pm \sqrt{13} \end{pmatrix}.$$

Hvis vi nå setter inn disse punktene i ellipselikningen, får vi

$$s^2 \left(1 + \frac{1}{4}(3 \pm \sqrt{13})^2\right) = 1,$$

slik at

$$s = \pm \frac{1}{1 + \frac{1}{4}(3 \pm \sqrt{13})^2} = \pm \frac{4}{4 + (3 \pm \sqrt{13})^2}.$$

Alt i alt har vi nå de fire punktene

$$\pm \frac{4}{4 + (3 \pm \sqrt{13})^2} \begin{pmatrix} 1 \\ 3 \pm \sqrt{13} \end{pmatrix}.$$

Nå er det bare å putte disse inn i h og sjekke hvilken som gir størst verdi.

$\boxed{10}$ Vi setter

$$g(x) = 6(x_1 + x_2)^2 + 2(x_1 - x_2)^2,$$

og beregner

$$f'(x) = (1, 2)$$

og

$$g'(x) = (12(x_1 + x_2) + 4(x_1 - x_2), 12(x_1 + x_2) - 4(x_1 - x_2)).$$

slik at lagrangelikningene blir

$$\begin{aligned} 12(x_1 + x_2) + 4(x_1 - x_2) &= \lambda \\ 12(x_1 + x_2) - 4(x_1 - x_2) &= 2\lambda \\ 6(x_1 + x_2)^2 + 2(x_1 - x_2)^2 &= 3 \end{aligned}$$

nå er det kanskje ryddigst å sette $u = x_1 + x_2$, $v = x_1 - x_2$, slik at likningssystemet blir

$$\begin{aligned} 12u + 4v &= \lambda \\ 12u - 4v &= 2\lambda \\ 6u^2 + 2v^2 &= 3 \end{aligned}$$

Gausseliminiasjon av de to første likningene gir

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{\lambda}{8} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

og setter vi denne i ellipselikningen, får vi

$$\lambda^2 \left(\frac{6}{64} + \frac{2}{64} = 3\right)$$

som gir

$$\lambda = \pm\sqrt{24} = \pm 2\sqrt{6}$$

Siden f er plan, gir en av disse maksimum, og den andre minimum; de faktiske x-koordinatene finner vi ved å løse likningssystemene

$$\begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} = \pm \frac{\sqrt{6}}{4} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

og evaluere f for å finne hva som er maksimum og hva som er minimum.