$\boxed{1}$ Et meningsløst vektorfelt er gitt ved

$$\mathbf{F}(x_1, x_2) = \begin{pmatrix} x_1 \\ (x_1 + x_2)^2 \\ x_2 \end{pmatrix} = \begin{bmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \\ F_3(x_1, x_2) \end{bmatrix}$$

$$F_2 = x_1^2 + 2x_1x_2 + x_2^2$$

og et annet ved

$$\mathbf{G}(z_1, z_2, z_3) = \begin{pmatrix} -z_1 - z_2 \\ z_1 + z_2 + z_3 \end{pmatrix} = \begin{bmatrix} G_1(z_1, z_2, z_3) \\ G_2(z_1, z_2, z_3) \end{bmatrix}$$

Finn jacobimatrisene til $\mathbf{F}$ og $\mathbf{G}$.

$$D_{\vec{x}} \vec{F} = \begin{bmatrix} \nabla F_1 \\ \nabla F_2 \\ \nabla F_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \\ \frac{\partial F_3}{\partial x_1} & \frac{\partial F_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2x_1 + 2x_2 & 2x_1 + 2x_2 \\ 0 & 1 \end{bmatrix}$$

$$D_{\vec{z}} \vec{G} = \begin{bmatrix} \nabla G_1 \\ \nabla G_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial G_1}{\partial z_1} & \frac{\partial G_1}{\partial z_2} & \frac{\partial G_1}{\partial z_3} \\ \frac{\partial G_2}{\partial z_1} & \frac{\partial G_2}{\partial z_2} & \frac{\partial G_2}{\partial z_3} \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Kjerneregelen i flere variable

La $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^p$ og $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^n$, og la $H : \mathbb{R}^m \to \mathbb{R}^p$ være gitt ved $\mathbf{H} = \mathbf{F}(\mathbf{G})$. Da er $D\mathbf{H}$ gitt ved matriseproduktet mellom $D\mathbf{F}$ og $D\mathbf{G}$:

$$D\mathbf{H} = D\mathbf{F}D\mathbf{G}$$

Dette er altfor teknisk å bevise for oss. Dersom du går med en liten matematikkstudent i dypet av deg selv et sted, kan du slå opp i Lindstrøm og se hvor mye jobb det er.

$\boxed{2}$ Finn jacobimatrisen til $\mathbf{H}(\mathbf{z}) = \mathbf{F}(\mathbf{G}(\mathbf{z}))$.
($\mathbf{F}$ og $\mathbf{G}$ som i oppgaven over.)

$$D\vec{H} = D\vec{F} \cdot D\vec{G} = \begin{bmatrix} 1 & 0 \\ 2(G_1(\vec{z}) + G_2(\vec{z})) & 2(G_1(\vec{z}) + G_2(\vec{z})) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 2(G_1(\vec{z}) + G_2(\vec{z})) \\ 1 & 1 & 1 \end{bmatrix}$$

Eksamen:

> [2] Forklar med utgangspunkt i Taylors formel hvorfor et kritisk punkt er en minimumspunkt dersom hessematrisen er positivt definitt.

Taylors formel: $f: \mathbb{R}^n \longrightarrow \mathbb{R} \Rightarrow f(\vec{x}+\vec{h}) \approx f(\vec{x}) + \nabla f(\vec{x}) \cdot \vec{h} + \frac{1}{2}\vec{h}^T H(\vec{x})\vec{h}.$

hvor hessematrisen er

$$H(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Siden $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ (når begge sider er kontinuerlige), så er $H(\vec{x})$ en symmetrisk matrise.

En symmetrisk (reell) matrise M er positivt-definitt hvis $\vec{x}^T M \vec{x} > 0$ for alle $\vec{x} \in \mathbb{R}^n \setminus \{0\}$.

Vi antar $H(\vec{x})$ er positivt definitt. Hvis $\vec{v}$ og λ er en egenvektor og -verdi for $H(\vec{x})$ da har vi

$$\vec{v}^T H(\vec{x})\, \vec{v} = \lambda > 0$$

$\Rightarrow$ alle egenverdiene til $H(\vec{x})$ er positive.

Uttrykket $f(\vec{x}+\vec{h}) \approx f(\vec{x}) + \nabla f(\vec{x}) \cdot \vec{h} + \frac{1}{2}\vec{h}^T H(\vec{x})\vec{h}$ er jo generelt sett ikke en likhet.

Men hvis de andreordens partiellderiverte til f er kontinuerlige og $\vec{h} \in \mathbb{R}^n$ er liten nok, da finnes det en funksjon $\varepsilon: \mathbb{R}^n \longrightarrow \mathbb{R}$ slik at $\lim_{\vec{h}\to 0} \varepsilon(\vec{h}) = 0$ og

$$f(\vec{x}+\vec{h}) = f(\vec{x}) + \nabla f(\vec{x}) \cdot \vec{h} + \frac{1}{2}\vec{h}^T H(\vec{x})\vec{h} + \varepsilon(\vec{h})\, |\vec{h}|^2.$$

(Dette er Setning 5.9.4 i Lindstrøm.)

Intuitivt kan vi tenke på det som følgende: taylorrekken til f har jo uendelig mange ledd, og leddene har jo en koeffisient $\frac{1}{n!}$ som blir mindre jo lengre ut i rekken vi kommer. Hvis $\vec{h}$ i tillegg er liten, så vil leddene lengre ut i rekken ha mindre å si for hvordan $f(\vec{x}+\vec{h})$ ser ut totalt sett enn de første leddene.

Så vi kan tenke at vi har samlet alt etter det tredje leddet i feilleddet $\varepsilon(\vec{h})|\vec{h}|^2$.

(Se Wikipedia: "Taylor Series # Taylor series in several variables" for mer om hele taylorrekken til f)

Lindstrøm viser også (Lemma 5.9.5(a)) at hvis egenverdiene til A er positive og $\lambda \geq 0$ er den minste egenverdien, da har vi

$$\vec{y}^T A \vec{y} \geq \lambda\, |\vec{y}|^2$$

for alle $\vec{y} \in \mathbb{R}^n$.

Merk at når vi studerer kritiske punkt har vi $\nabla f = 0$.

La nå λ være den minste egenverdien til $H(\vec{x})$. Vi har

$$f(\vec{x}+\vec{h}) - f(\vec{x}) = \nabla f(\vec{x}) \cdot \vec{h} + \frac{1}{2} \vec{h}^T H(\vec{x}) \vec{h} + \varepsilon(\vec{h}) |\vec{h}|^2$$

$$= \frac{1}{2} \vec{h}^T H(\vec{x}) \vec{h} + \varepsilon(\vec{h}) |\vec{h}|^2$$

$$\geq \frac{1}{2} \lambda |\vec{h}|^2 + \varepsilon(\vec{h}) |\vec{h}|^2$$

$$= \left(\frac{1}{2}\lambda + \varepsilon(\vec{h})\right) |\vec{h}|^2$$

Siden $\lambda > 0$, $|\vec{h}|^2 > 0$ når $\vec{h} \neq \vec{0}$, og $\varepsilon(\vec{h}) \to 0$ når $\vec{h}$ blir liten, så må dette uttrykket være ≥ 0

$$\Rightarrow f(\vec{x}+h) - f(\vec{x}) \geq 0 \quad \Rightarrow f(\vec{x}) \leq f(\vec{x}+\vec{h})$$

Altså når vi beveger oss fra det kritiske punktet $\vec{x}$ til et punkt $\vec{x}+\vec{h}$ veldig nærme (siden $\vec{h}$ er liten), så stiger verdien til f.

Dette gjelder jo uansett hvilken retning $\vec{h}$ peker i, så $\vec{x}$ må være et lokalt minimumspunkt for f.

Lindstrøm sider 461-465 forklarer mer detaljert.

Men idéen er altså å starte med Taylors formel for f i flere variabler, og så argumentere for at vi må ha $f(\vec{x}) \leq f(\vec{x}+\vec{h})$.