

Differentiable funktioner af flere variable

Resultaterne vedrørende differentiable funktioner er spredt ud over flere afsnit i [WFT]. Jeg giver derfor her en koncis præsentation af resultaterne. Vi erstatter derfor dele af Chapters 5–6 med disse noter.

Vi starter med lidt notation. De lineære afbildninger $T: \mathbf{R}^n \rightarrow \mathbf{R}^m$ betegnes med $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$. Enhver lineær afbildning $T: \mathbf{R}^n \rightarrow \mathbf{R}^m$ kan repræsenteres ved en $m \times n$ matrix. Vi skelner ofte ikke mellem den lineære afbildning og dens matrix med hensyn til de kanoniske baser i $\mathbf{R}^n$ og $\mathbf{R}^m$. Normen på både $\mathbf{R}^n$ og $\mathbf{R}^m$ betegnes med $\|\mathbf{x}\|$. De kanoniske basisvektorer i $\mathbf{R}^n$ og $\mathbf{R}^m$ betegnes med $\mathbf{e}_i$. Vektorerne i $\mathbf{R}^n$ og $\mathbf{R}^m$ er søjlevektorer.

Vi bemærker, at enhver lineær afbildning $T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ er kontinuert. Dette resultat følger af, at vi har $\|T\mathbf{x}\| \leq \|T\|\|\mathbf{x}\|$, hvor $\|T\|$ er operator normen af den lineære afbildning.

En funktion $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ kan skrives som en søjlevektor af funktioner $(\mathbf{F})_i = F_i: \mathcal{O} \rightarrow \mathbf{R}^m$, $1 \leq i \leq m$,

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} F_1(\mathbf{x}) \\ F_2(\mathbf{x}) \\ \vdots \\ F_m(\mathbf{x}) \end{bmatrix}.$$

Definition 1. Lad $\mathcal{O} \subseteq \mathbf{R}^n$ være en åben mængde og lad $\mathbf{a} \in \mathcal{O}$. En funktion $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ siges at være *differentiabel* i $\mathbf{a}$, hvis der findes en lineær afbildning $T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ og en funktion $\mathbf{E}_\mathbf{a}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ med egenskaben $\lim_{\mathbf{h} \rightarrow 0} \mathbf{E}_\mathbf{a}(\mathbf{h}) = \mathbf{0}$, således at

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \mathbf{F}(\mathbf{a}) + T\mathbf{h} + \|\mathbf{h}\|\mathbf{E}_\mathbf{a}(\mathbf{h}) \quad (1)$$

for alle $\mathbf{h} \in \mathbf{R}^n$, således at $\mathbf{a} + \mathbf{h} \in \mathcal{O}$. Den lineære afbildning T kaldes den *totale afledede* af $\mathbf{F}$ i $\mathbf{a}$ og betegnes med $\mathbf{DF}(\mathbf{a})$.

Vi bemærker, at da $\mathcal{O}$ er åben, findes der et $\delta > 0$, således at $B_\delta(\mathbf{a}) \subseteq \mathcal{O}$, og dermed at (1) gælder for alle $\mathbf{h}$ med $\|\mathbf{h}\| < \delta$.

Det følger umiddelbart fra definitionen, at vi har nedenstående resultater.

Sætning 2. Hvis $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ er differentiabel i $\mathbf{a} \in \mathcal{O}$, så er $\mathbf{F}$ kontinuert i $\mathbf{a}$.

Bevis. Overlades til læseren. □

Sætning 3. Lad $\mathcal{O} \subseteq \mathbf{R}^n$ være en åben mængde og lad $\mathbf{a} \in \mathcal{O}$. En funktion $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ er differentiabel i $\mathbf{a}$, hvis og kun hvis alle komponentfunktionerne F_i er differentiable i $\mathbf{a}$, $1 \leq i \leq m$.

Bevis. Overlades til læseren. □

Sætning 4. Lad $\mathbf{x}_0 \in \mathbf{R}^m$ og lad $T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$. Definer $\mathbf{F}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ ved $\mathbf{F}(\mathbf{x}) = \mathbf{x}_0 + T\mathbf{x}$. Så er $\mathbf{F}$ differentiabel i alle $\mathbf{a} \in \mathbf{R}^n$, og vi har $\mathbf{DF}(\mathbf{a}) = T$.

Bevis. Vi har, at

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \mathbf{x}_0 + T(\mathbf{a} + \mathbf{h}) = \mathbf{x}_0 + T\mathbf{a} + T\mathbf{h} = \mathbf{F}(\mathbf{a}) + T\mathbf{h}.$$

Dermed er (1) opfyldt med $\mathbf{E}_\mathbf{a}(\mathbf{h}) = \mathbf{0}$ for alle $\mathbf{h} \in \mathbf{R}^n$. □

Afbildningerne defineret i Sætning 4 kaldes affine afbildninger.

Det næste resultat viser, at definitionen af differentiabilitet medfører, at alle de partielle afledede eksisterer i punkter, hvor der er differentiabilitet.

Sætning 5. Lad $\mathcal{O} \subseteq \mathbf{R}^n$ være en åben mængde og lad $\mathbf{a} \in \mathcal{O}$. Antag, at $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ er differentiabel i $\mathbf{a}$. Så eksisterer de partielle afledede i $\mathbf{a}$. Vi har, at

$$\frac{\partial F_i}{\partial x_j}(\mathbf{a}) = (\mathbf{DF}(\mathbf{a})(\mathbf{e}_j))_i, \quad 1 \leq j \leq n. \quad (2)$$

Bevis. Lad j være givet, $1 \leq j \leq n$. Vi bruger lemmaet i [1]. Bestem $\delta > 0$, således at $\mathbf{a} + t\mathbf{e}_j \in \mathcal{O}$ for alle t , $|t| < \delta$. Vi har da

$$\mathbf{F}(\mathbf{a} + t\mathbf{e}_j) = \mathbf{F}(\mathbf{a}) + \mathbf{DF}(\mathbf{a})(t\mathbf{e}_j) + \|t\mathbf{e}_j\| \mathbf{E}_{\mathbf{a}}(t\mathbf{e}_j)$$

Ser vi på den i -te komponent, har vi

$$\begin{aligned} F_i(\mathbf{a} + t\mathbf{e}_j) &= F_i(\mathbf{a}) + (\mathbf{DF}(\mathbf{a})(t\mathbf{e}_j))_i + |t|(\mathbf{E}_{\mathbf{a}}(t\mathbf{e}_j))_i \\ &= F_i(\mathbf{a}) + t(\mathbf{DF}(\mathbf{a})(\mathbf{e}_j))_i + t \frac{|t|}{t} (\mathbf{E}_{\mathbf{a}}(t\mathbf{e}_j))_i. \end{aligned}$$

Det følger af lemmaet i [1], at den j -te partielle afledede af F_i eksisterer og er givet ved (2). $\square$

Dette resultat medfører følgende vigtige resultat.

Korollar 6. Matricen for den lineære afbildning $\mathbf{DF}(\mathbf{a})$ med hensyn til de kanoniske baser er givet ved

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{a}) & \frac{\partial F_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{a}) & \frac{\partial F_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{a}) & \frac{\partial F_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_m}{\partial x_n}(\mathbf{a}) \end{bmatrix}. \quad (3)$$

Matricen (3) kaldes *Jacobi matricen* i $\mathbf{a}$ for $\mathbf{F}$. Bemærk placeringen af indgangene i forhold til indices for afbildningen $\mathbf{F}$.

Vi beviser nu kædereglen for differentiation. Beviset skal sammenlignes med beviset i [1].

Sætning 7. Lad $\mathcal{O} \subseteq \mathbf{R}^n$ og $\mathcal{O}_1 \subseteq \mathbf{R}^m$ være åbne mængder. Lad $\mathbf{a} \in \mathcal{O}$ og $\mathbf{b} \in \mathcal{O}_1$. Antag, at $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$ er differentiabel i $\mathbf{a}$, at $\mathbf{G}: \mathcal{O}_1 \rightarrow \mathbf{R}^k$ er differentiabel i $\mathbf{b}$, og at $\mathbf{F}(\mathbf{a}) = \mathbf{b}$. Så er $\mathbf{H} = \mathbf{G} \circ \mathbf{F}$ differentiabel i $\mathbf{a}$, og vi har, at

$$\mathbf{DH}(\mathbf{a}) = \mathbf{DG}(\mathbf{b})\mathbf{DF}(\mathbf{a}). \quad (4)$$

Her er $\mathbf{DG}(\mathbf{b})\mathbf{DF}(\mathbf{a})$ sammensætningen af de to lineære afbildninger $\mathbf{DG}(\mathbf{b})$ og $\mathbf{DF}(\mathbf{a})$. Vi har valgt at bruge denne notation, som er den sædvanlige for lineære afbildninger, ligesom notationen $\mathbf{H} = \mathbf{G} \circ \mathbf{F}$ er den sædvanlige for sammensætning af generelle afbildninger.

Bevis. Vi antager i det følgende, at $\mathbf{h}$ er tilstrækkelig lille, så at $\mathbf{a} + \mathbf{h} \in \mathcal{O}$. Vi starter med udregningen

$$\begin{aligned} \mathbf{H}(\mathbf{a} + \mathbf{h}) - \mathbf{H}(\mathbf{a}) &= \mathbf{G}(\mathbf{F}(\mathbf{a} + \mathbf{h})) - \mathbf{G}(\mathbf{F}(\mathbf{a})) \\ &= \mathbf{G}(\mathbf{F}(\mathbf{a}) + \mathbf{v}) - \mathbf{G}(\mathbf{F}(\mathbf{a})) \\ &= \mathbf{G}(\mathbf{b} + \mathbf{v}) - \mathbf{G}(\mathbf{b}). \end{aligned}$$

Her er $\mathbf{v} = \mathbf{DF}(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\mathbf{E}_a(\mathbf{h})$. Dernæst bruger vi differentiability af $\mathbf{G}$ til at skrive

$$\mathbf{G}(\mathbf{b} + \mathbf{v}) - \mathbf{G}(\mathbf{b}) = \mathbf{DG}(\mathbf{b})(\mathbf{v}) + \|\mathbf{v}\|\tilde{\mathbf{E}}_b(\mathbf{v}).$$

Vi kombinerer nu de to udregninger og bruger, at $\mathbf{DG}(\mathbf{b})$ er lineær. Vi antager, at $\mathbf{h} \neq \mathbf{0}$.

$$\begin{aligned} \mathbf{G}(\mathbf{b} + \mathbf{v}) - \mathbf{G}(\mathbf{b}) &= \mathbf{DG}(\mathbf{b})(\mathbf{DF}(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\|\mathbf{E}_a(\mathbf{h})) + \|\mathbf{v}\|\tilde{\mathbf{E}}_b(\mathbf{v}) \\ &= \mathbf{DG}(\mathbf{b})(\mathbf{DF}(\mathbf{a})(\mathbf{h})) + \|\mathbf{h}\|\mathbf{DG}(\mathbf{b})(\mathbf{E}_a(\mathbf{h})) + \|\mathbf{v}\|\tilde{\mathbf{E}}_b(\mathbf{v}) \\ &= \mathbf{DG}(\mathbf{b})(\mathbf{DF}(\mathbf{a})(\mathbf{h})) + \|\mathbf{h}\|\left[\mathbf{DG}(\mathbf{b})(\mathbf{E}_a(\mathbf{h})) + \frac{\|\mathbf{v}\|}{\|\mathbf{h}\|}\tilde{\mathbf{E}}_b(\mathbf{v})\right]. \end{aligned}$$

Vi mangler at vise, at leddene i $[\dots]$ går mod nul, når $\mathbf{h}$ går mod nul. Vi har, at

$$\|\mathbf{DG}(\mathbf{b})(\mathbf{E}_a(\mathbf{h}))\| \leq C\|\mathbf{E}_a(\mathbf{h})\|$$

og højre side går mod nul, når $\mathbf{h}$ går mod nul. Vi har også, at

$$\|\mathbf{v}\| \leq \|\mathbf{h}\|(\|\mathbf{DF}(\mathbf{a})\| + \|\mathbf{E}_a(\mathbf{h})\|),$$

således at $\|\mathbf{v}\|/\|\mathbf{h}\| \leq C$ for $\|\mathbf{h}\| < \delta$. Da endvidere $\mathbf{v} \rightarrow \mathbf{0}$ for $\mathbf{h} \rightarrow \mathbf{0}$, er resultatet vist. $\square$

Vi har følgende hovedresultat, som etablerer en forbindelse mellem differentiability og de partielle afledede.

Sætning 8. *Lad $\mathcal{O} \subseteq \mathbf{R}^n$ være åben. Lad $\mathbf{a} \in \mathcal{O}$ og lad $\delta > 0$, så at $B_\delta(\mathbf{a}) \subseteq \mathcal{O}$. Lad $\mathbf{F}: \mathcal{O} \rightarrow \mathbf{R}^m$, således at de partielle afledede eksisterer i $B_\delta(\mathbf{a})$ og er kontinuerte i $\mathbf{a}$. Så er $\mathbf{F}$ differentiable i $\mathbf{a}$.*

Bevis. Ifølge Sætning 3 er det nok at se på en funktion $F: \mathcal{O} \rightarrow \mathbf{R}$, altså en af komponenterne af $\mathbf{F}$. Hvis F er differentiable i $\mathbf{a}$, så har vi, at

$$DF(\mathbf{a}) = \left[\frac{\partial F}{\partial x_1}(\mathbf{a}), \frac{\partial F}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial F}{\partial x_n}(\mathbf{a}) \right]. \quad (5)$$

Antag, at $\|\mathbf{h}\| < \delta$, således at $\mathbf{a} + \mathbf{h} \in B_\delta(\mathbf{a})$. Definer $\mathbf{y}_0 = \mathbf{a}$, $\mathbf{y}_1 = \mathbf{a} + h_1\mathbf{e}_1$, $\mathbf{y}_2 = \mathbf{y}_1 + h_2\mathbf{e}_2$, og generelt $\mathbf{y}_k = \mathbf{y}_{k-1} + h_k\mathbf{e}_k$, $k = 2, \dots, n$. Vi bemærker, at $\mathbf{y}_n = \mathbf{a} + \mathbf{h}$, og at $\mathbf{y}_k \in B_\delta(\mathbf{a})$ for $k = 0, 1, \dots, n$. Vi har nu, at

$$F(\mathbf{a} + \mathbf{h}) - F(\mathbf{a}) = \sum_{k=1}^n (F(\mathbf{y}_k) - F(\mathbf{y}_{k-1})).$$

Vi betegner liniestykket mellem $\mathbf{y}_{k-1}$ og $\mathbf{y}_k$ med $\ell(\mathbf{y}_{k-1}, \mathbf{y}_k)$. Da $\mathbf{y}_k = \mathbf{y}_{k-1} + h_k\mathbf{e}_k$, kan vi bruge den en-dimensionale middelværdissætning til at bestemme $\mathbf{z}_k \in \ell(\mathbf{y}_{k-1}, \mathbf{y}_k)$, således at

$$F(\mathbf{y}_k) - F(\mathbf{y}_{k-1}) = \frac{\partial F}{\partial x_k}(\mathbf{z}_k)h_k.$$

Vi har da, at

$$F(\mathbf{a} + \mathbf{h}) = F(\mathbf{a}) + \sum_{k=1}^n \frac{\partial F}{\partial x_k}(\mathbf{a})h_k + \sum_{k=1}^n \left(\frac{\partial F}{\partial x_k}(\mathbf{z}_k) - \frac{\partial F}{\partial x_k}(\mathbf{a}) \right) h_k.$$

Vi definerer nu

$$E_a(\mathbf{h}) = \sum_{k=1}^n \left(\frac{\partial F}{\partial x_k}(\mathbf{z}_k) - \frac{\partial F}{\partial x_k}(\mathbf{a}) \right) \frac{h_k}{\|\mathbf{h}\|}.$$

Vi har da, at

$$|E_{\mathbf{a}}(\mathbf{h})| \leq \sum_{k=1}^n \left| \frac{\partial F}{\partial x_k}(\mathbf{z}_k) - \frac{\partial F}{\partial x_k}(\mathbf{a}) \right|.$$

Vi skal vise, at højre side går mod nul, når $\mathbf{h}$ går mod nul. Lad $\varepsilon > 0$ være givet. Da $\partial F / \partial x_k$ er kontinuert i $\mathbf{a}$, kan vi bestemme δ_k , således at

$$\left| \frac{\partial F}{\partial x_k}(\mathbf{z}_k) - \frac{\partial F}{\partial x_k}(\mathbf{a}) \right| < \frac{\varepsilon}{n}$$

for $|h_k| < \delta_k$. Sæt $\delta = \min\{\delta_1, \delta_2, \dots, \delta_n\}$. så har vi, at

$$\sum_{k=1}^n \left| \frac{\partial F}{\partial x_k}(\mathbf{z}_k) - \frac{\partial F}{\partial x_k}(\mathbf{a}) \right| < \varepsilon$$

for $\|\mathbf{h}\| < \delta$. Vi har dermed vist, at

$$F(\mathbf{a} + \mathbf{h}) = F(\mathbf{a}) + \sum_{k=1}^n \frac{\partial F}{\partial x_k}(\mathbf{a}) h_k + \|\mathbf{h}\| E_{\mathbf{a}}(\mathbf{h}),$$

og dermed, at F er differentiabel i $\mathbf{a}$. □

Vi bruger ofte en lidt anden notation og terminologi for den totale afledede af en funktion $F: \mathcal{O} \rightarrow \mathbf{R}$ i et punkt $\mathbf{a} \in \mathcal{O}$. Den totale afledede kaldes *gradienten* af F i $\mathbf{a}$ og vi bruger notationen

$$(\nabla F)(\mathbf{a}) = \left[\frac{\partial F}{\partial x_1}(\mathbf{a}), \frac{\partial F}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial F}{\partial x_n}(\mathbf{a}) \right]. \quad (6)$$

Hvis notationen skal være helt konsistent, så skal $DF(\mathbf{a})$ betragtes som en $1 \times n$ matrix og $(\nabla F)(\mathbf{a})$ som en vektor med n komponenter. I notationen ovenfor er den skrevet som en rækkevektor, men det er faktisk bedre at betragte den som en søjlevektor. Med den konvention er da

$$\frac{\partial F}{\partial x_j}(\mathbf{a}) = (\nabla F)(\mathbf{a}) \cdot \mathbf{e}_j.$$

Her er $\mathbf{x} \cdot \mathbf{y}$ det indre produkt mellem vektorerne $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$.

Litteratur

- [1] Arne Jensen, Supplerende materiale 1, 27. oktober 2013.