Kjerneregelen

John Rognes

25. januar 2011

La $\mathbf{G}: A \to B$ og $\mathbf{F}: B \to \mathbb{R}^k$ være to funksjoner, der $A \subset \mathbb{R}^n$ og $B \subset \mathbb{R}^m$:

$$A \xrightarrow{G} B \xrightarrow{F} \mathbb{R}^k$$

La $\mathbf{H}: A \to \mathbb{R}^k$ være den sammensatte funksjonen gitt ved $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$.

Theorem

Anta at $\mathbf{G}$ *er deriverbar i* $\mathbf{a} \in A$, *og at* $\mathbf{F}$ *er deriverbar i* $\mathbf{b} = \mathbf{G}(\mathbf{a}) \in B$. *Da er* $\mathbf{H}$ *deriverbar i* $\mathbf{a}$, *med derivert*

$$\mathbf{H}'(\mathbf{a}) = \mathbf{F}'(\mathbf{b})\mathbf{G}'(\mathbf{a}) = \mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{G}'(\mathbf{a})\,.$$

Bevis (1/6)

At $\mathbf{G}$ er deriverbar i $\mathbf{a}$ betyr at

$$\mathbf{G}(\mathbf{a}+\mathbf{r}) = \mathbf{G}(\mathbf{a}) + \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}(\mathbf{r})$$

der restleddet $\boldsymbol{\sigma}(\mathbf{r})$ går raskere mot $\mathbf{0}$ enn $\mathbf{r}$, dvs. at $|\boldsymbol{\sigma}(\mathbf{r})|/|\mathbf{r}| \to 0$ når $\mathbf{r} \to \mathbf{0}$.

At $\mathbf{F}$ er deriverbar i $\mathbf{b} = \mathbf{G}(\mathbf{a})$ betyr at

$$\mathbf{F}(\mathbf{b}+\mathbf{s}) = \mathbf{F}(\mathbf{b}) + \mathbf{F}'(\mathbf{b})\mathbf{s} + \boldsymbol{\tau}(\mathbf{s})$$

der restleddet $\boldsymbol{\tau}(\mathbf{s})$ går raskere mot $\mathbf{0}$ enn $\mathbf{s}$, dvs. at $|\boldsymbol{\tau}(\mathbf{s})|/|\mathbf{s}| \to 0$ når $\mathbf{s} \to \mathbf{0}$.

For å vise at $\mathbf{H} = \mathbf{F} \circ \mathbf{G}$ er deriverbar i $\mathbf{a}$ må vi vise at $\mathbf{H}(\mathbf{a}+\mathbf{r}) = \mathbf{H}(\mathbf{a}) + \mathbf{H}'(\mathbf{a})\mathbf{r}$ pluss et restledd som går raskere mot $\mathbf{0}$ enn $\mathbf{r}$. Her er

$$\begin{aligned}
\mathbf{H}(\mathbf{a}+\mathbf{r}) &= \mathbf{F}(\mathbf{G}(\mathbf{a}+\mathbf{r})) = \mathbf{F}(\mathbf{G}(\mathbf{a}) + \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}(\mathbf{r})) \\
&= \mathbf{F}(\mathbf{b}+\mathbf{s}) = \mathbf{F}(\mathbf{b}) + \mathbf{F}'(\mathbf{b})\mathbf{s} + \boldsymbol{\tau}(\mathbf{s}) \\
&= \mathbf{H}(\mathbf{a}) + \mathbf{F}'(\mathbf{b})\mathbf{G}'(\mathbf{a})\mathbf{r} + \mathbf{F}'(\mathbf{b})\boldsymbol{\sigma}(\mathbf{r}) + \boldsymbol{\tau}(\mathbf{s})
\end{aligned}$$

der $\mathbf{s} = \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}(\mathbf{r})$. Restleddet er altså $\mathbf{F}'(\mathbf{b})\boldsymbol{\sigma}(\mathbf{r}) + \boldsymbol{\tau}(\mathbf{s})$.

Bevis (3/6)

Vi må vise at

$$|\mathbf{F}'(\mathbf{b})\boldsymbol{\sigma}(\mathbf{r}) + \boldsymbol{\tau}(\mathbf{s})|/|\mathbf{r}| \to 0$$

når $\mathbf{r} \to \mathbf{0}$, der $\mathbf{s} = \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}(\mathbf{r})$.
Ved trekantulikheten er det nok å vise at

1. $|\mathbf{F}'(\mathbf{b})\boldsymbol{\sigma}(\mathbf{r})|/|\mathbf{r}| \to 0$ når $\mathbf{r} \to \mathbf{0}$, og
2. $|\boldsymbol{\tau}(\mathbf{s})|/|\mathbf{r}| \to 0$ når $\mathbf{r} \to \mathbf{0}$.

Vi behandler de to tilfellene hver for seg.

Tilfelle 1:
Normen til Jacobi-matrisen $\mathbf{F}'(\mathbf{b}) = (\partial F_i/\partial u_p)_{i,p}$ er lik

$$\|\mathbf{F}'(\mathbf{b})\| = \Big(\sum_{i,p} \frac{\partial F_i}{\partial u_p}(\mathbf{b})^2\Big)^{1/2}.$$

Vi har at

$$0 \le |\mathbf{F}'(\mathbf{b})\boldsymbol{\sigma}(\mathbf{r})|/|\mathbf{r}| \le \|\mathbf{F}'(\mathbf{b})\||\boldsymbol{\sigma}(\mathbf{r})|/|\mathbf{r}| \to \|\mathbf{F}'(\mathbf{b})\| \cdot 0$$

når $\mathbf{r} \to \mathbf{0}$, så venstresiden går mot 0 når $\mathbf{r} \to \mathbf{0}$.

Bevis (5/6)

Tilfelle 2:
La $\epsilon > 0$ være et vilkårlig, positivt tall. Siden $|\boldsymbol{\tau}(\mathbf{s})|/|\mathbf{s}| \to 0$ når $\mathbf{s} \to \mathbf{0}$ vet vi at

$$|\boldsymbol{\tau}(\mathbf{s})| \leq \epsilon|\mathbf{s}|$$

for alle tilstrekkelig små $\mathbf{s}$. Siden $|\boldsymbol{\sigma}(\mathbf{r})|/|\mathbf{r}| \to 0$ når $\mathbf{r} \to \mathbf{0}$ vil også

$$|\boldsymbol{\sigma}(\mathbf{r})| \leq |\mathbf{r}|$$

for alle tilstrekkelig små $\mathbf{r}$. Da er

$$\begin{aligned} |\mathbf{s}| &= |\mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}(\mathbf{r})| \leq |\mathbf{G}'(\mathbf{a})\mathbf{r}| + |\boldsymbol{\sigma}(\mathbf{r})| \\ &\leq \|\mathbf{G}'(\mathbf{a})\||\mathbf{r}| + |\mathbf{r}| = (\|\mathbf{G}'(\mathbf{a})\| + 1)|\mathbf{r}| \end{aligned}$$

for alle tilstrekkelig små $\mathbf{r}$. Spesielt vil $\mathbf{s} \to \mathbf{0}$ når $\mathbf{r} \to \mathbf{0}$.

Bevis (6/6)

For alle tilstrekkelig små $\mathbf{r}$ er da

$$|\boldsymbol{\tau}(\mathbf{s})| \leq \epsilon|\mathbf{s}| \leq \epsilon(\|\mathbf{G}'(\mathbf{a})\| + 1)|\mathbf{r}|$$

slik at

$$0 \leq |\boldsymbol{\tau}(\mathbf{s})|/|\mathbf{r}| \leq \epsilon(\|\mathbf{G}'(\mathbf{a})\| + 1)\,.$$

Siden $\|\mathbf{G}'(\mathbf{a})\|$ er en konstant, og ϵ kan velges vilkårlig liten, impliserer dette at

$$|\boldsymbol{\tau}(\mathbf{s})|/|\mathbf{r}| \to 0$$

når $\mathbf{r} \to \mathbf{0}$.

Dette viser at $\mathbf{H}$ er deriverbar i $\mathbf{a}$, med derivert $\mathbf{H}'(\mathbf{a}) = \mathbf{G}'(\mathbf{b})\mathbf{F}'(\mathbf{a})$, som skulle vises.