Kædereglen for vektorfunktioner

Reglen for differentiation af sammensat funktion lyder således for funktioner af én reel variabel
$$(f \circ g)'(x) = f' \circ g(x) g'(x).$$

Generaliseringer af denne regel kaldes ofte kædereglen, se fx 'Kompendium i calculus'. Med notation som i noten 'Differentiabilitet' kan reglen for vektorfunktioner skrives
$$D(F \circ G)(\boldsymbol{x}) = (DF) \circ G(\boldsymbol{x}) DG(\boldsymbol{x}).$$

Bemærk, at højresiden er et produkt af to Jacobimatricer. Skrevet ud i flere detaljer har vi
$$D(F \circ G)(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_m} \\ \vdots & & & \\ \frac{\partial f_p}{\partial y_1} & \frac{\partial f_p}{\partial y_2} & \cdots & \frac{\partial f_p}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}.$$
Den første Jacobimatrix er $p \times m$, og elementerne skal udregnes i punktet $G(\boldsymbol{x})$, den anden er $m \times n$, og elementerne skal udregnes i punktet $\boldsymbol{x}$. Matricen $D(F \circ G)(\boldsymbol{x})$ er altså $p \times n$ med (i, j)'te element
$$\frac{\partial (F \circ G)_i}{\partial x_j} = \frac{\partial f_i}{\partial y_1} \frac{\partial g_1}{\partial x_j} + \frac{\partial f_i}{\partial y_2} \frac{\partial g_2}{\partial x_j} + \cdots + \frac{\partial f_i}{\partial y_m} \frac{\partial g_m}{\partial x_j},$$
som ofte (lidt upræcist) skrives
$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_i}{\partial y_1} \frac{\partial y_1}{\partial x_j} + \frac{\partial f_i}{\partial y_2} \frac{\partial y_2}{\partial x_j} + \cdots + \frac{\partial f_i}{\partial y_m} \frac{\partial y_m}{\partial x_j}.$$

Lad os se på, hvordan resultatet fremkommer.

Betragt vektorfunktionen $G : D_G \to \mathbb{R}^m$, $D_G \subseteq \mathbb{R}^n$, G differentiabel i $\boldsymbol{x}_0$, og vektorfunktionenen $F : D_F \to \mathbb{R}^p$, $D_F \subseteq \mathbb{R}^m$, F differentiabel i $\boldsymbol{y}_0 = F(\boldsymbol{x}_0)$, hvor $\boldsymbol{y}_0$ er et indre punkt i $G(D_G) \cap D_F$.

Sætning. *Den sammensatte vektorfunktion $F \circ G$ er differentiabel i $\boldsymbol{x}_0$ og $D(F \circ G)(\boldsymbol{x}_0) = (DF) \circ G(\boldsymbol{x}_0) DF(\boldsymbol{x}_0)$.*

Bevis. At G er differentiabel i $\boldsymbol{x}_0$, betyder, at der eksisterer en epsilonfunktion $\mathcal{E}_1(\boldsymbol{h})$, hvor $\boldsymbol{h}$ er en tilvæksten ud fra punktet $\boldsymbol{x}_0$, således at tilvæksten $\Delta G(\boldsymbol{h}) = G(\boldsymbol{x}_0 + \boldsymbol{h}) - G(\boldsymbol{x}_0)$ kan skrives
$$\Delta G(\boldsymbol{h}) = DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h}) \|\boldsymbol{h}\|.$$

Ligeledes, at F er differentiabel i $\boldsymbol{y}_0$, betyder, at der eksisterer en epsilonfunktion $\mathcal{E}_2(\boldsymbol{k})$, hvor $\boldsymbol{k}$ er tilvæksten ud fra punktet $\boldsymbol{y}_0$, således at

$$\Delta F(\boldsymbol{k}) = DF(\boldsymbol{y}_0)\boldsymbol{k} + \mathcal{E}_2(\boldsymbol{k})\|\boldsymbol{k}\|.$$

Betragt

$$\Delta F \circ G(\boldsymbol{h}) = F \circ G(\boldsymbol{x}_0 + \boldsymbol{h}) - F \circ G(\boldsymbol{x}_0)$$

Bemærk, at $G(\boldsymbol{x}_0 + \boldsymbol{h})$ kan skrives $G(\boldsymbol{x}_0) + \Delta G(\boldsymbol{h}) = G(\boldsymbol{x}_0) + DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|$, hvorefter

$$\Delta F \circ G(\boldsymbol{h}) = F \circ \Big(G(\boldsymbol{x}_0) + DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|\Big) - F \circ G(\boldsymbol{x}_0).$$

Sæt $\boldsymbol{k} = DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|$, og bemærk, at $\boldsymbol{k} \to \boldsymbol{0}$ for $\boldsymbol{h} \to \boldsymbol{0}$. Desuden er $G(\boldsymbol{x}_0) = \boldsymbol{y}_0$, dvs.

$$\Delta F \circ G(\boldsymbol{h}) = F(\boldsymbol{y}_0 + \boldsymbol{k}) - F(\boldsymbol{y}_0)$$

Heri indsættes

$$\begin{aligned} F(\boldsymbol{y}_0 + \boldsymbol{k}) &= F(\boldsymbol{y}_0) + \Delta F(\boldsymbol{k}) \\ &= F(\boldsymbol{y}_0) + DF(\boldsymbol{y}_0)\boldsymbol{k} + \mathcal{E}_2(\boldsymbol{k})\|\boldsymbol{k}\| \\ &= F(\boldsymbol{y}_0) + (DF) \circ G(\boldsymbol{x}_0)\Big(DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|\Big) + \mathcal{E}_2(\boldsymbol{k})\|\boldsymbol{k}\|. \end{aligned}$$

Ved indsættelsen udgår $F(\boldsymbol{y}_0)$, hvorefter

$$\begin{aligned} \Delta F \circ G(\boldsymbol{h}) &= (DF) \circ G(\boldsymbol{x}_0)\Big(DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|\Big) + \mathcal{E}_2(\boldsymbol{k})\|\boldsymbol{k}\| \\ &= (DF) \circ G(\boldsymbol{x}_0)DG(\boldsymbol{x}_0)\boldsymbol{h} + (DF) \circ G(\boldsymbol{x}_0)\mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\| + \mathcal{E}_2(\boldsymbol{k})\|\boldsymbol{k}\| \\ &= (DF) \circ G(\boldsymbol{x}_0)DG(\boldsymbol{x}_0)\boldsymbol{h} + \Big((DF) \circ G(\boldsymbol{x}_0)\mathcal{E}_1(\boldsymbol{h}) + \mathcal{E}_2(\boldsymbol{k})\frac{\|\boldsymbol{k}\|}{\|\boldsymbol{h}\|}\Big)\|\boldsymbol{h}\| \end{aligned}$$

Bemærk, at $(DF) \circ G(\boldsymbol{x}_0)\mathcal{E}_1(\boldsymbol{h}) \to \boldsymbol{0}$ for $\boldsymbol{h} \to \boldsymbol{0}$. Endvidere eksisterer der en omegn om $\boldsymbol{x}_0$, hvor $\frac{\|\boldsymbol{k}\|}{\|\boldsymbol{h}\|}$ er begrænset[1)], dvs. $\mathcal{E}_2(\boldsymbol{k})\frac{\|\boldsymbol{k}\|}{\|\boldsymbol{h}\|} \to \boldsymbol{0}$ for $\boldsymbol{h} \to \boldsymbol{0}$.
Vi kan nu sammenfattende skrive

$$\Delta F \circ G(\boldsymbol{h}) = (DF) \circ G(\boldsymbol{x}_0)DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}(\boldsymbol{h})\|\boldsymbol{h}\|,$$

hvilket viser sætningens gyldighed. □

For differentialet af $F \circ G$ i $\boldsymbol{x}_0$ får vi

$$d(F \circ G)(\boldsymbol{h}) = (DF) \circ G(\boldsymbol{x}_0)DG(\boldsymbol{x}_0)\boldsymbol{h},$$

eller på en form, hvor punktet $\boldsymbol{x}_0$ er underforstået, og hvor tilvæksten $\boldsymbol{h}$ er identificeret med differentialet $d\boldsymbol{x}$ af den identiske afbildning i $\mathbb{R}^n$,

$$d(F \circ G) = (DF) \circ G\; DG\; d\boldsymbol{x}.$$

[1)] $\frac{\|\boldsymbol{k}\|}{\|\boldsymbol{h}\|} = \frac{\|DG(\boldsymbol{x}_0)\boldsymbol{h} + \mathcal{E}_1(\boldsymbol{h})\|\boldsymbol{h}\|\|}{\|\boldsymbol{h}\|} \leq \|DG(\boldsymbol{x}_0)\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\| + \|\mathcal{E}_1(\boldsymbol{h})\|$ er begrænset, idet

$$\max_{i=1,\dots,m}\Big|\Big((DG(\boldsymbol{x}_0)\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\Big)_i\Big| = \max_{i=1,\dots,m}\Big|\sum_{j=1}^{n}\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0)\frac{h_i}{\|\boldsymbol{h}\|}\Big| \leq \max_{i=1,\dots,m}\sum_{j=1}^{n}\Big|\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0)\Big| \cdot 1$$

$$\leq n \max_{\substack{i=1,\dots,m\\ j=1,\dots,n}}\Big|\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0)\Big| \;\Rightarrow\; \|DG(\boldsymbol{x}_0)\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\| \leq mn \max_{\substack{i=1,\dots,m\\ j=1,\dots,n}}\Big|\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0)\Big|.$$

Den i'te komponent af $d(F \circ G)$ kan skrives

$$\big(d(F \circ G)\big)_i = \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial y_1} \frac{\partial g_1}{\partial x_j} + \frac{\partial f_i}{\partial y_2} \frac{\partial g_2}{\partial x_j} + \cdots + \frac{\partial f_i}{\partial y_m} \frac{\partial g_m}{\partial x_j} \right) dx_j.$$

25.3.2014/BR