Flervariabelanalys LV2

Linnea Hallin, Emil Ingelsten, Emma Lundqvist,
Joseph Löfving, Anton Wikström, David Winroth

22 februari 2019

1 Variabelbyten och kedjeregeln

Ett exempel på ett vanligt variabelbyte är det mellan cartesiska och polära koordinater, d.v.s. variabelsubstitutionen

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

Säg att $f : \mathbb{R}^2 \to \mathbb{R}$ är en C^1-funktion. Eftersom $f(x, y) = f(x(r, \theta), y(r, \theta))$ kan man för att beräkna $\frac{\partial f}{\partial r}$ och $\frac{\partial f}{\partial \theta}$ använda kedjeregeln för funktioner av flera variabler. Resultatet blir

$$\begin{cases} \frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}, \end{cases}$$

eller på matrisform

$$\begin{bmatrix} \partial f/\partial r \\ \partial f/\partial \theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix} \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix}. \tag{1}$$

Eftersom det för 2×2-matriser gäller att

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

så ger (1) även ett uttryck för $\frac{\partial f}{\partial x}$ och $\frac{\partial f}{\partial y}$ om $\frac{\partial f}{\partial r}$ och $\frac{\partial f}{\partial \theta}$ är kända, nämligen

$$\begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} = \frac{1}{r\cos^2\theta + r\sin^2\theta} \begin{bmatrix} r\cos\theta & -\sin\theta \\ r\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \partial f/\partial r \\ \partial f/\partial \theta \end{bmatrix},$$

som kan förenklas något till

$$\begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\frac{\sin\theta}{r} \\ \sin\theta & \frac{\cos\theta}{r} \end{bmatrix} \begin{bmatrix} \partial f/\partial r \\ \partial f/\partial \theta \end{bmatrix} = \begin{bmatrix} \partial r/\partial x & \partial \theta/\partial x \\ \partial r/\partial y & \partial \theta/\partial y \end{bmatrix} \begin{bmatrix} \partial f/\partial r \\ \partial f/\partial \theta \end{bmatrix}. \tag{2}$$

Att det andra av likhetstecknena ovan gäller kan verifieras genom direkt beräkning av de olika partiella derivatorna $\frac{\partial r}{\partial x}$, $\frac{\partial \theta}{\partial x}$, $\frac{\partial r}{\partial y}$ och $\frac{\partial \theta}{\partial y}$.

2 Taylors formel i flera variabler

För att kunna beräkna Taylors formel i flervariabelanalys är två ingredienser nödvändiga:

- Taylors formel i en variabel.
- Kedjeregeln.

Taylors formel i en variabel: Låt $k \in \mathbb{N}$, låt $a \in \mathbb{R}$, låt $f : \mathbb{R} \to \mathbb{R}$ tillhöra klassen $C^{k+1}(E)$, där E är någon omgivning av a. Då gäller att:

$$\forall x \in E : f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-a)^{k+1}, \text{ för något } \xi \in (a,x) \tag{3}$$

Där det är värt att notera att för $x = a + h$ gäller:

$$f(a+h) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}h^j + \frac{f^{(k+1)}(\xi)}{(k+1)!}h^{k+1}, \text{ för något } \xi \in (a,x) \tag{4}$$

Vilken vi kan uttrycka som:

$$f(a+h) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}h^j + \mathcal{O}(h^{k+1}) \tag{5}$$

För två variabler gäller följande för Taylorutvecklingar av ordning 2:

Låt $(a,b) \in \mathbb{R}^2$ i $f \in C^3(E)$ där E är en omgivning av (a,b). Då gäller:

$$\begin{aligned} f(a+h,b+k) = f(a,b) + h \cdot f_x(a,b) + k \cdot f_y(a,b) + \frac{1}{2}(h^2 f_{xx}(a,b) + 2hk f_{xy}(a,b) + k^2 f_{yy}(a,b)) + \\ + \mathcal{O}(\left\|\vec{h}\right\|^3), \text{ där } \vec{h} = (h,k). \end{aligned} \tag{6}$$

Bevisideén är att reducera uttrycket till en variabel med hjälp av kedjeregeln. För att göra detta, tänk dig att h, k är fixt och betrakta funktionen

$$F : [-1,1] \to \mathbb{R}, \text{där } F(t) = f(a+th, b+tk) = f(g_1(t), g_2(t)). \tag{7}$$

Faktumet att $f \in C^3(E)$ ger genom kedjeregeln att $F \in C^3((-1,1))$ Enligt Taylors sats för en variabel får vi då:

$$F(t) = F(0) + F'(0) \cdot t + F''(0) \cdot \frac{t^2}{2} + F'''(\xi)\frac{t^3}{6}, \text{för något } \xi \in (0,t). \tag{8}$$

Kedjeregeln ger då:

$$
\begin{aligned}
&F'(t) = \frac{d}{dt}(f(g_1(t), g_2(t)) = f_x(g_1(t), g_2(t) \cdot g_1'(t) + \\
&+ f_y(g_1(t), g_2(t) \cdot g_2'(t) = h \cdot f_x(a+th, b+tk) + k \cdot f_y(a+th+b+tk)
\end{aligned}
\tag{9}
$$

Om vi här definierar differentialoperatorn D som $(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})$ får vi att $F' = Df$. Det gäller då att

$$
\begin{aligned}
&F''(t) = (F')' = D(Df) = (D \bullet D)f = D^2 f = (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^2 f = \\
&= ((h\frac{\partial}{\partial x})^2 + 2hk\frac{\partial}{\partial x} \cdot \frac{\partial}{\partial y} + (k\frac{\partial}{\partial y})^2)f = \\
&= (h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy})
\end{aligned}
\tag{10}
$$

För feltermen krävs att vi vet F'''.

$$
\begin{aligned}
&F''' = D^3 f = (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^3 f = (h^3 f_{xxx} + 3h^2 k f_{xxy} + 3hk^2 f_{xyy} + k^3 f_{yyy}) \\
&\Rightarrow f(a+th, b+tk) = f(a,b) + t[hf_x(a,b) + k(f_y(a,b)] + \\
&+ \frac{t^2}{2}[h^2 f_{xx}(a,b) + 2hk f_{xy}(a,b) + k^2 f_{yy}(a,b)] + \\
&+ \frac{t^3}{6}[h^3 f_{xxx}(a+\xi h, b+\xi k) + 3h^2 k f_{xxy}(a+\xi h, b+\xi k) + \\
&+ 3hk^2 f_{xyy}(a+\xi h, b+\xi k) + k^3 f_{yyy}(a+\xi h, b+\xi k)] \\
&\text{sätt nu } t = 1 \\
&\Rightarrow f(a+h, b+k) = f(a,b) + hf_x(a,b) + kf_y(a,b) + \\
&+ \frac{1}{2}[h^2 f_{xx}(a,b) + 2hk f_{xy} + k^2 f_{yy}(a,b)] + \mathcal{O}(\left\|\vec{h}\right\|^3) \\
&\text{där } \vec{h} = (h,k).
\end{aligned}
\tag{11}
$$

$\square$

Det är tydligt att samma metod kan användas för att bevisa Taylors sats för både fler variabler och för Taylorutvecklingar av högre ordning.

Tag ett $\vec{a} \in \mathbb{R}^n$ och $F := f(\vec{a} + t\vec{h})$. Låt D vara en omgivning av $\vec{a}$. Om $f \in C^{k+1}(D)$

gäller det att,

$$\Rightarrow F(t) = \sum_{i=0}^{k} \frac{F^{(i)}(0)}{i!} \cdot t^i + \frac{F^{(k+1)}(\xi)}{(k+1)!} \cdot t^{k+1}, \text{ där } \xi \in (0,t).$$

Kedjeregeln ger

$$F' = Df, \text{ där } D = \sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} = \vec{h} \bullet \nabla, \text{ där } \nabla = (\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}).$$

Induktivt leder detta nu till att

$$F^{(i)} = D^i f = (\vec{h} \cdot \nabla)^i f \tag{12}$$

Insättning i Fs Taylorutveckling ger då att:

$$f(\vec{a} + t\vec{h}) = \sum_{i=0}^{k} \frac{(\vec{h} \cdot \nabla)^i}{i!} f(\vec{a}) \cdot t^i + \frac{(\vec{h} \cdot \nabla)^{k+1} f(\vec{a} + \xi\vec{h})}{(k+1)!} t^{k+1}, \ \xi \in (0,t)$$

Slutligen sätter vi $t = 1$

$$\Rightarrow f(\vec{a} + \vec{h}) = \sum_{i=0}^{k} \frac{(\vec{h} \cdot \nabla)^i}{i!} f(\vec{a}) + \mathcal{O}(\left\|\vec{h}\right\|^{k+1})$$

3 Lokala extrempunkter

Definition 1. Låt $f : \mathbb{R}^n \to \mathbb{R}$ vara en differentierbar funktion och låt $\vec{a} \in \mathbb{R}^n$.

Funktionen f sägs ha ett lokalt minimum resp. maximum i $\vec{x} = \vec{a}$ om:

- f är definierad i en omgivning av $\vec{a}$
- $\exists \delta > 0 : |\vec{x} - \vec{a}| < \delta \Rightarrow f(\vec{x}) \geq f(\vec{a})$ resp. $f(\vec{x}) \leq f(\vec{a})$

Proposition. Låt $f : \mathbb{R}^n \to \mathbb{R}$ vara en differentierbar funktion och låt $\vec{a} \in \mathbb{R}^n$.

Om f har en lokal extrempunkt i $\vec{x} = \vec{a}$ så gäller att $\nabla f(\vec{a}) = \vec{0}$, d.v.s. att alla f:s partiella derivator är 0 i $\vec{x} = \vec{a}$.

Bevis.
f differentierbar $\Rightarrow f(\vec{a} + \vec{h}) = f(\vec{a}) + \vec{h} \cdot \nabla f(\vec{a}) + |\vec{h}|\rho(\vec{h})$, där $\rho(\vec{h}) \to 0$ då $\vec{h} \to \vec{0}$.

Antag att $\nabla f(\vec{a}) \neq \vec{0}$. Sätt $\vec{v} = \nabla f(\vec{a})$ och välj $\vec{h} = h\vec{v}$
$\Rightarrow \vec{h} \cdot \nabla f(\vec{a}) = h|\vec{v}|^2$.

$\Rightarrow f(\vec{a} + \vec{h}) - f(\vec{a}) = h|\vec{v}|^2 + h|\vec{v}|\rho(h)$, där $\rho(h) \to 0$ då $h \to 0$.

I högerledet kommer här $h|\vec{v}|^2$ att dominera för tillräckligt små h, d.v.s.

$$\begin{cases} \text{HL} > 0 \text{ om } h > 0, \text{ där } h \text{ är tillräckligt litet} \\ \text{HL} < 0 \text{ om } h < 0, \text{ där } h \text{ är tillräckligt litet} \end{cases}$$

$\Rightarrow f$ kan ej ha en lokal extrempunkt i $\vec{x} = \vec{a}$.

$\square$

3.1 Kritiska punkter

Definition 2. Låt $f : \mathbb{R}^n \to \mathbb{R}$ vara differentierbar i $\vec{a} \in \mathbb{R}^n$.

Punkten $\vec{a}$ kallas *stationär*, eller *kritisk*, om $\nabla f(\vec{a}) = \vec{0}$.

En variabel: I en variabel gäller att om $f''(a) \neq 0$, där a är en kritisk punkt, så har f en lokal extrempunkt i $x = a$. Mer specifikt gäller att om $f''(a) > 0$ så är $x = a$ en lokal minimipunkt för f, och om $f''(a) < 0$ så är $x = a$ en lokal maximipunkt för f. Vi vill hitta en liknande regel för att klassificera kritiska punkter för funktioner av flera variabler.

Två variabler: Låt $f : \mathbb{R}^2 \to \mathbb{R}$. Antag att $f \in C^3$ i en omgivning av den kritiska punkten (a, b). Taylorutveckling av f i denna omgivning ger att

$$f(a+h, b+k) = f(a,b) + hf_x + kf_y\Big|_{(a,b)} + \frac{1}{2}(h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy})\Big|_{(a,b)} + \mathcal{O}\big((h^2+k^2)^{\frac{3}{2}}\big).$$

Eftersom (a, b) är en kritisk punkt så är $f_x(a, b) = f_y(a, b) = 0$ enligt definitionen av kritisk punkt. Därför är ekvationen ovan ekvivalent med att

$$f(a+h, b+k) - f(a,b) = \frac{1}{2}(h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy})\Big|_{(a,b)} + \mathcal{O}\big((h^2+k^2)^{\frac{3}{2}}\big). \tag{13}$$

Definition 3. En funktion $Q : \mathbb{R}^2 \to \mathbb{R}$ kallas *binär kvadratisk form* om det existerar tal $A, B, C \in \mathbb{R}$ sådana att

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2$$

Detta innebär att ekvation (13) ovan kan skrivas

$$f(a+h, b+k) - f(a,b) = Q(h,k) + \mathcal{O}\big((h^2+k^2)^{\frac{3}{2}}\big),$$

för $A = \frac{1}{2}f_{xx}(a, b)$, $B = \frac{1}{2}f_{xy}(a, b)$, $C = \frac{1}{2}f_{yy}(a, b)$.

3.2 Klassificering av kritiska punkter

Definition 4. Låt $Q : \mathbb{R}^2 \to \mathbb{R}$ vara en binär kvadratisk form. I så fall gäller följande:

- Q sägs vara *positiv definit* om $Q(x, y) > 0 \ \forall (x, y) \neq (0, 0)$.

- Q sägs vara *negativ definit* om $Q(x,y) < 0 \ \forall (x,y) \neq (0,0)$.
- Q sägs vara *indefinit* om $\exists (x_1,y_1),(x_2,y_2) \neq (0,0) : Q(x_1,y_1) > 0 \wedge Q(x_2,y_2) < 0$.
- Q sägs vara *positiv semidefinit* om $Q(x,y) \geq 0 \ \forall (x,y) \neq (0,0)$, men Q ej är positiv definit.
- Q sägs vara *negativ semidefinit* om $Q(x,y) \leq 0 \ \forall (x,y) \neq (0,0)$, men Q ej är negativ definit.

Sats (Sats 2.6.12, 1:a formuleringen)**.** Låt Q vara som ovan, $f : \mathbb{R}^2 \to \mathbb{R}$, $f \in C^3$ och (a,b) en kritisk punkt till f. Då gäller:

1. $Q(h,k)$ är positiv definit $\Rightarrow (a,b)$ är ett (strängt) lokalt minimum till f.
2. $Q(h,k)$ är negativ definit $\Rightarrow (a,b)$ är ett (strängt) lokalt maximum till f.
3. $Q(h,k)$ är indefinit $\Rightarrow (a,b)$ är en sadelpunkt till f.

Sats (Sats 1)**.** Låt Q vara som ovan. Då gäller:

1. $AC - B^2 > 0$ och $A > 0 \Rightarrow Q$ är positiv definit.
2. $AC - B^2 > 0$ och $A < 0 \Rightarrow Q$ är negativ definit.
3. $AC - B^2 < 0 \Rightarrow Q$ är indefinit.
4. $AC - B^2 = 0 \Rightarrow Q$ är semidefinit.

Sats (Sats 2, 2:a formuleringen)**.** Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f \in C^3$ och låt (a,b) vara en kritisk punkt till f. Då gäller:

1. $f_{xx}f_{yy} - f_{xy}^2 > 0$ i (a,b) och $f_{xx} > 0$ i $(a,b) \Rightarrow (a,b)$ är ett (strängt) lokalt minimum till f.
2. $f_{xx}f_{yy} - f_{xy}^2 > 0$ i (a,b) och $f_{xx} < 0$ i $(a,b) \Rightarrow (a,b)$ är ett (strängt) lokalt maximum till f.
3. $f_{xx}f_{yy} - f_{xy}^2 < 0$ i $(a,b) \Rightarrow (a,b)$ är en sadelpunkt till f.

4 Vektorvärda funktioner

Låt m, $n \in \mathbb{N}$. En funktion $\vec{F} : \mathbb{R}^n \to \mathbb{R}^m$ kallas vektorvärd då dess input och output båda kan beskrivas som vektorer. Speciellt gäller att en vektorvärd funktion med $m = 1$ kallas skalärvärd.

En vektorvärd funktion $\vec{F} : \mathbb{R}^n \to \mathbb{R}^m$ består av m olika underfunktioner vilka alla beror på n variabler. Med andra ord gäller att

$$\vec{F}(x_1, x_1, ..., x_n) = (F_1(x_1, x_2, ..., x_n), F_2(x_1, x_2, ..., x_n), ..., F_m(x_1, x_2, ..., x_n)). \tag{14}$$

Funktionen $\vec{F}$ sägs vara differentierbar om varje enskild underfunktion F_i är det.

4.1 Linjärisering av vektorvärda funktioner

På matrisform kan linjäriseringen av en vektorvärd funktion $\vec{F}:\mathbb{R}^n \to \mathbb{R}^m$ kring punkten $\vec{a}$ skrivas

$$\begin{bmatrix} F_1(\vec{a}+\vec{h}) \\ F_2(\vec{a}+\vec{h}) \\ ... \\ F_m(\vec{a}+\vec{h}) \end{bmatrix} = \begin{bmatrix} F_1(\vec{a}) \\ F_2(\vec{a}) \\ ... \\ F_m(\vec{a}) \end{bmatrix} + \begin{bmatrix} \partial F_1/\partial x_1 & \partial F_1/\partial x_2 & ... & \partial F_1/\partial x_n \\ \partial F_2/\partial x_1 & \partial F_2/\partial x_2 & ... & \partial F_2/\partial x_n \\ ... & ... & ... & ... \\ \partial F_m/\partial x_1 & \partial F_m/\partial x_2 & ... & \partial F_m/\partial x_n \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ ... \\ h_n \end{bmatrix} + |\vec{h}| \begin{bmatrix} \rho_1(\vec{h}) \\ \rho_2(\vec{h}) \\ ... \\ \rho_m(\vec{h}) \end{bmatrix} \tag{15}$$

där F_i är underfunktionerna till $\vec{F}$, $\vec{h}$ förskjutningsvektorn, x_i de variabler $\vec{F}$ beror på, h_i elementen i vektorn $\vec{h}$ och ρ_i funktioner för vilka det gäller att $\rho_i \to 0$ då $\vec{h} \to \vec{0}$. Mer kompakt kan ekvation (15) skrivas

$$\vec{F}(\vec{a}+\vec{h}) = \vec{F}(\vec{a}) + \vec{F}'(\vec{a})\vec{h} + |\vec{h}|\vec{\rho}(\vec{h}). \tag{16}$$

Matrisen $\vec{F}'(\vec{a})$ kallas *funktionalmatrisen* eller *Jacobimatrisen*.

4.2 Kedjeregeln för vektorvärda funktioner

För två vektorvärda funktioner, $\vec{F}:\mathbb{R}^n \to \mathbb{R}^m$ och $\vec{G}:\mathbb{R}^p \to \mathbb{R}^n$ där $n,m,p \in \mathbb{N}$ gäller att $\vec{F}\circ\vec{G}:\mathbb{R}^p \to \mathbb{R}^m$. Vidare gäller att om $\vec{G}$ är differentierbar i punkten $\vec{a}$ och $\vec{F}$ är differentierbar i punkten $\vec{G}(\vec{a})$ så är $\vec{F}\circ\vec{G}$ differentierbar i punkten $\vec{a}$, och

$$(\vec{F}\circ\vec{G})'(\vec{a}) = \vec{F}'(\vec{G}(\vec{a}))\vec{G}'(\vec{a}). \tag{17}$$

4.3 Specialfallet $\vec{F}:\mathbb{R}^n \to \mathbb{R}^n$

En vektorvärd funktion $\vec{F}:\mathbb{R}^n \to \mathbb{R}^n$ kallas för ett *n-dimensionellt vektorfält*. För en sådan vektorvärd funktion $\vec{F}(x_1,x_2,...,x_n) = (y_1,y_2,...,y_n)$, där $y_i = y_i(x_1,x_2,...,x_n)$ används notationen

$$\vec{F}' = \frac{\partial(y_1,y_2,...,y_n)}{\partial(x_1,x_2,...,x_n)}. \tag{18}$$

5 Inversa och implicita funktionssatserna

5.1 Inversa funktionssatsen

Låt $\vec{y} = \vec{F}(\vec{x})$ vara en C^1-funktion från $\mathbb{R}^n$ till $\mathbb{R}^n$ och $\vec{a}$ en punkt i definitionsmängden sådan att

$$\det(\vec{F}'(\vec{a})) \neq 0.$$

Då finns det öppna omgivningar U och V av $\vec{a}$ respektive $\vec{b} = \vec{F}(\vec{a})$ sådana att avbildningen

$$\vec{F}:\vec{U} \to \vec{V}$$

är bijektiv och inversfunktionen $\vec{F}^{-1}:\vec{V} \to \vec{U}$ en funktion av klassen C^1.

Ett exempel på användningen av denna sats är vid variabelbyten mellan polära och cartesiska koordinater, där $\vec{F}^{-1}(r,\theta) = (r\cos(\theta), r\sin(\theta)) = (x,y)$ och $\det(\vec{F}^{-1})' = r \neq 0$ utanför origo.

5.2 Implicita funktionssatsen för två variabler

Låt $F(x, y)$ vara en C^1-funktion av två variabler och (a, b) en punkt på nivåkurvan $F(x, y) = C$, dvs. $F(a, b) = C$.

Om $\frac{\partial F}{\partial y}(a, b) \neq 0$, då finns det en öppen omgivning U av (a, b) sådan att restriktionen av nivåkurvan till U implicit definierar en C^1-funktion $y = f(x)$ och

$$f'(x) = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))} \tag{19}$$

där förutsättningen att $\frac{\partial F}{\partial y}(a, b) = F'_y(a, b) \neq 0$ medför att $F'_y(x, y) \neq 0$ i en hel omgivning av (a, b) eftersom F'_y är kontinuerlig enligt förutsättningen.

Denna sats används vid fastställning av huruvida en ekvation implicit definierar en kontinuerligt deriverbar funktion eller ej, exempelvis vid framställning av en nivåkurva $f(x, y) = 1$ för enhetscirkeln, där $f(x, y) = x^2 + y^2$.