Tavelpresentation flervariabelanalys v. 2

Julia Björkman, Amanda Magnusson, Leo Samuelsson,
Vincent Udén, Elias Sörstrand, Markus Hjält

21 februari 2020

Civilingenjörsprogrammen i teknisk fysik och teknisk matematik

Chalmers tekniska högskola

Email:
julbjo@student.chalmers.se
hjaltm@student.chalmers.se
eliasso@student.chalmers.se
amamagn@student.chalmers.se
udenv@student.chalmers.se
leosa@student.chalmers.se

Innehåll

1 Kedjeregeln för flera variabler **3**
1.1 Allmän formel för kedjeregeln 3
1.2 Partiella derivator av högre ordning 3

2 Taylorutveckling **4**
2.1 Taylors Sats i en variabel 4
2.2 Taylors Sats för två variabler 4
2.3 Taylors Sats för godtyckligt antal variabler 5
2.4 Bevisidé för Taylors formel av ordning 2 5

3 Kritiska punkter och extrempunkter **5**
3.1 Definition av lokala extrempunkter 6
3.2 Definition av kritisk punkt i flera variabler 6
3.3 Definition av terrasspunkt 6
3.4 Sats 1 6
3.5 Sats 2 6

4 Klassificering av kritiska punkter **7**
4.1 Sats 1 8
4.2 Sats 2 8

5 Vektorvärda funktioner **8**
5.1 Linjärisering 9
5.2 Kedjeregeln för sammansättning av vektorvärda funktioner . . . 9
5.3 Variabelbyten 9

1 Kedjeregeln för flera variabler

1.1 Allmän formel för kedjeregeln

Som vi lärt oss i envariabelanalysen är en av de viktigaste derivationslagarna regeln för derivatan av en sammansatt funktion. Vi vet att

$$\frac{d}{dt}f(g(t)) = f'(g(t)) \cdot g'(t) \tag{1}$$

Ekvationen (1) visar en sammansatt funktion, där f är beroende av funktionen $g(t)$, vilken i sin tur är beroende av t. I flervariabelanalysen är det istället av intresse att studera funktioner som beror på flera antal variabler. Låt oss kolla på exemplet där vi har en funktion som ser ut på följande sätt:

$$z = f(g_1(t), g_2(t), g_3(t)) \tag{2}$$

Här har vi istället en funktion som beror på tre variabler - i detta fall g_1, g_2 och g_3, som alla i sin tur är funktioner som beror på den enda variabeln t. För att derivera denna funktion, måste vi upprepa kedjeregeln, tills dess att de inre funktionerna alla är deriverade med avseende på önskad variabel. I fallet ovan kommer derivatan att bli som följer:

$$\frac{dz}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial f}{\partial x_3}\frac{dx_3}{dt} \tag{3}$$

Det mönster som uppenbarar sig leder fram till den allmänna formeln för derivering av sammansatta funktioner i flera variabler, som formuleras på följande sätt:

$$\frac{d}{dt}f(g_1(t),...,g_n(t)) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g_1(t),...,g_n(t)) \cdot \frac{dg_i}{dt} \tag{4}$$

Vi kan även låta g vara en funktion av m stycken variabler, vilket ger:

$$\frac{d}{dt_i}f(g_1(t_1,...,t_m),...,g_n(t_1,...,t_m)) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(g_1(t_1,...,t_m),...,g_n(t_1,...,t_m)) \cdot \frac{dg_j}{dt_i} \tag{5}$$

1.2 Partiella derivator av högre ordning

Låt $D \subseteq \mathbb{R}^2$ vara en öppen mängd. Låt vidare $f \in C^2(D)$, och punkten $(a,b) \in D$. Då gäller att:

$$f_{xy}(a,b) = f_{yx}(a,b) \tag{6}$$

Ovan sats betyder att utifrån givna förutsättningar spelar det ingen roll om vi först deriverar med avseende på x och sedan y, eller tvärtom. Vi kan ta hjälp av medelvärdessatsen för att bevisa detta.

2 Taylorutveckling

Det finns elementära funktioner vars funktionsvärden ofta bara kan anges som approximationer, se t.ex. e^x och $ln(x)$. Approximationer till dessa funktioner kan göras med hjälp av polynom, eftersom de har den enklaste strukturen bland de elementära funktionerna. Detta kan göras med Taylors formel, som approximerar en funktion $f(x)$ med ett polynom runt en given punkt a. Detta formuleras i Taylors Sats.

2.1 Taylors Sats i en variabel

Låt $k \in \mathbb{N}$, $D \subseteq \mathbb{R}$ vara en öppen mängd, $f \in \boldsymbol{C}^{k+1}(D)$, $a \in D$. Då gäller

$$f(a+h) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!} \cdot h^j + O(h^{k+1}) \tag{7}$$

där $h = x - a \Leftrightarrow f(a+h) = f(x)$. $O(h^{k+1})$ är approximationens felterm, och minskar för tillräckligt små h i storlek för varje steg (ökad k) i utvecklingen. Ekvation (7) kallas för Taylorutvecklingen av grad k till f kring $x = a$. Summan, dvs huvudtermen i utvecklingen, kallas för Taylorpolynomet av grad k till f i $x = a$.

2.2 Taylors Sats för två variabler

Låt $D \subseteq \mathbb{R}^2$ vara en öppen mängd, $(a,b) \in D$ en punkt. Välj (h,k) tillräckligt liten s.a. cirkeln av radien $\sqrt{h^2+k^2}$ kring punkten (a,b) tillhör D. Låt $f \in \boldsymbol{C}^{m+1}(D)$ för något $m \in \mathbb{N}$. En approximation kan göras för $f(a+h,b+k)$, och ser då ut på följande vis:

$$f(a+h,b+k) = \sum_{j=0}^{m} \frac{(h \cdot \frac{\partial}{\partial x} + k \cdot \frac{\partial}{\partial y})^j}{j!} f(a,b) + \\ + \frac{(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^{m+1}}{(m+1)!} \cdot f(a+\xi h, b+\xi k) \tag{8}$$

med feltermen $O((h^2+k^2)^{\frac{m+1}{2}})$. Det viktigaste fallet är Taylorutvecklingen av grad och ordning 2 för en C^3-funktion $f(x,y)$ av 2 variabler:

$$f(a+h,b+k) = f(a,b) + [h \cdot f_x(a,b) + k \cdot f_y(a,b)] + \\ + \frac{1}{2}[h^2 \cdot f_{xx}(a,b) + 2hk \cdot f_{xy}(a,b) + k^2 \cdot f_{yy}(a,b)] + O((h^2+k^2)^{\frac{3}{2}}) \tag{9}$$

2.3 Taylors Sats för godtyckligt antal variabler

Låt $f \in \boldsymbol{C}^{m+1}(D)$, $m \in \mathbb{N}$, $\boldsymbol{a} = (a,b)$, $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$, och $D \subseteq \mathbb{R}^2$ en öppen mängd. Välj $\boldsymbol{h} = (h,k)$ tillräckligt liten s.a. cirkeln med radien $r = \sqrt{h^2+k^2}$ runt punkten $\boldsymbol{a}$ tillhör D. Ger följande omskrivning av (8):

$$f(\boldsymbol{a}+\boldsymbol{h}) = \sum_{j=0}^{m} \frac{(\boldsymbol{h}\cdot\nabla)^j}{j!} f(\boldsymbol{a}) + O(||\boldsymbol{h}||^{m+1}) \tag{10}$$

Denna formel kan sedan generaliseras till godtyckligt antal variabler.

2.4 Bevisidé för Taylors formel av ordning 2

Betrakta funktionen

$$F(t) = f(a+th, b+tk) \tag{11}$$

av en variabel, för $0 \leq t \leq 1$. Det är uppenbart att $F \in \boldsymbol{C}^3$, och ekvation (7) ger följande:

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \frac{F'''(\theta t)}{3!}t^3 \tag{12}$$

för något θ, $0 \leq \theta \leq 1$. Vidare ger kedjeregeln att

$$F'(t) = f'_x(a+th, b+tk)h + f'_y(a+th, b+tk)k, \tag{13}$$

samt att

$$F''(t) = f''_{xx}(a+th,b+tk)h^2 + 2f''_{xy}(a+th,b+tk)hk + f''_{yy}(a+th,b+tk)k^2 \tag{14}$$

Om $t = 1$ fås

$$f(a+h,b+k) = F(0) + F'(0) + \frac{1}{2}F''(0) + \frac{1}{6}F'''(\theta) \tag{15}$$

och då återstår endast att visa att $\frac{1}{6}F'''(\theta)$ kan skrivas som resttermen. Detta kan göras genom att derivera $F''(t)$ med hjälp av kedjeregeln och jämföra de resulterande monomen i h och k framför varje term med en konstant multiplicerad med $(h^2+k^2)^{\frac{3}{2}}$, och visa att division av resttermen med $(h^2+k^2)^{\frac{3}{2}}$ resulterar i en begränsad funktion.

3 Kritiska punkter och extrempunkter

Taylorutvecklingen kan tillämpas vid klassificiering av en funktions kritiska punkter. Vi börjar med några definitioner.

3.1 Definition av lokala extrempunkter

Om $\boldsymbol{x} = \boldsymbol{a}$ är en lokal extrempunkt till en funktion f gäller att:

1. f har ett lokalt minimum i $\boldsymbol{x} = \boldsymbol{a}$ om $\exists \delta > 0$ så att $\|\boldsymbol{h}\| < \delta \implies f(\boldsymbol{a}+\boldsymbol{h}) \geq f(\boldsymbol{a})$.
2. f har ett lokalt maximum i $\boldsymbol{x} = \boldsymbol{a}$ om $\exists \delta > 0$ så att $\|\boldsymbol{h}\| < \delta \implies f(\boldsymbol{a}+\boldsymbol{h}) \leq f(\boldsymbol{a})$

En extrempunkt är en typ av kritisk punkt, men för lokalt kritiska punkter i allmänhet finns följande definition.

3.2 Definition av kritisk punkt i flera variabler

Om $f(x_1,...,x_n)$ är en differentierbar funktion av n variabler är $\boldsymbol{a} = (a_1,...,a_n)$ en kritisk punkt till f om,

$$\frac{\partial f}{\partial x_i}(\boldsymbol{a}) = 0, \forall i = 1,...,n. \tag{16}$$

Förutom lokala extremum finns en typ av kritisk punkt som kallas terrasspunkt och denna definieras enligt följande definition.

3.3 Definition av terrasspunkt

En kritisk punkt till en funktion f i en variabel som inte är en lokal extrempunkt kallas en terrasspunkt.

Nu när kritiska punkter och extrempunkter är definierade följer två satser.

3.4 Sats 1

Om f är en differentierbar funktion och $x = a$ är ett lokalt extremum till f, då gäller att $f'(a) = 0$.

3.5 Sats 2

Låt $f \in \mathrm{C}^2$ vara en två gånger deriverbar funktion. Låt också a vara en lokalt kritisk punkt till f. Då gäller att:

1. Om $f''(a) > 0$ så är $x = a$ en strikt lokal minimipunkt till f. Det innebär att $\exists \delta > 0$ så att $h \in (-\delta,\delta), h \neq 0 \implies f(a+h) > f(a)$
2. Om $f''(a) < 0$ så är $x = a$ en strikt lokal maximumpunkt till f. Det innebär att $\exists \delta > 0$ så att $h \in (-\delta,\delta), h \neq 0 \implies f(a+h) < f(a)$

3. Om $f''(a) = 0$ kan inga slutsatser dars om punktens karaktär. Karaktären avgörs genom att Taylorutveckla till högre ordning eller undersöka om funktionen ger information om vilken karaktär punkten har.

Då kritiska punkter nu är definierade är det av intresse att kunna bestämma deras karaktär. Hur detta genomförs beskrivs i avsnitt 4.

4 Klassificering av kritiska punkter

Om vi låter $D \subseteq \mathbb{R}^2$, $(a,b) \in D$, $f \in C^3(D)$ och antag att (a,b) är en kritisk punkt till f. Då går det att taylorutveckla f till grad 3 vilket ger oss följande ekvation,

$$f(a+h,b+k) = f(a,b) + f'_x(a,b)\cdot h + f'_y(a,b)\cdot k + \\ \frac{1}{2}\left(f''_{xx}(a,b)\cdot h^2 + 2f''_{xy}(a,b)\cdot hk + f''_{yy}(a,b)\cdot k^2\right) + O((h^2+k^2)^{\frac{3}{2}}). \quad (17)$$

Vi vet att de första ordningens termer, f'_x och f'_y är lika med noll i en kritisk punkt. Detta medför att ekvationen ovan kan förenklas till,

$$f(a+h,b+k) = f(a,b) + \frac{1}{2}\left(f''_{xx}(a,b)\cdot h^2 + 2f''_{xy}(a,b)\cdot hk + f''_{yy}(a,b)\cdot k^2\right) \\ + O((h^2+k^2)^{\frac{3}{2}}). \quad (18)$$

Därefter kan vi subtrahera $f(a,b)$ på båda sidorna och sätta

$$\begin{cases} A = \frac{1}{2}f''_{xx}(a,b), \\ B = \frac{1}{2}f''_{xy}(a,b), \\ C = \frac{1}{2}f''_{yy}(a,b), \end{cases} \quad (19)$$

då A,B,C är reella tal. Görs detta i (18) går det att skriva om ekvationen som,

$$f(a+h,b+k) - f(a,b) = Ah^2 + 2Bhk + Ck^2 + O((h^2+k^2)^{\frac{3}{2}}). \quad (20)$$

Detta medför sedan att man kan skriva en ny funktion, Q, beroende av h och k då a och b är specifika tal.

$$Q(h,k) = Ah^2 + 2Bhk + Ck^2. \quad (21)$$

Där följer en definition som säger att en funktion $Q : \mathbb{R}^2 \to \mathbb{R}$ på formen (20) där A,B,$C \in \mathbb{R}$ kallas för en binär kvadratisk form.

Det går att skilja på fem olika typer av binära kvadratiska former,

1. $Q(x,y)$ sägs vara *positivt definit* om $Q(x,y) > 0 \quad \forall (x,y) \neq (0,0)$
2. $Q(x,y)$ sägs vara *negativt definit* om $Q(x,y) < 0 \quad \forall (x,y) \neq (0,0)$
3. $Q(x,y)$ sägs vara *indefinit* om $\exists (x_1,y_1) : Q(x_1,y_1) > 0$ och $\exists (x_2,y_2) : Q(x_2,y_2) < 0$
4. $Q(x,y)$ sägs vara *positivt semidefinit* om $Q(x,y) \geq 0, \forall (x,y)$ och $\exists (x_1,y_1) \neq (0,0) : Q(x_1,y_1) = 0$
5. $Q(x,y)$ sägs vara *negativt semidefinit* om $Q(x,y) \leq 0, \forall (x,y)$ och $\exists (x_1,y_1) \neq (0,0) : Q(x_1,y_1) = 0$

Från detta följer en sats, *sats 1*, som förklarar vad de olika typerna har för egenskaper.

4.1 Sats 1

Låt $Q(h,k) = f''_{xx}(a,b)h^2 + 2f''_{xy}(a,b)hk + f''_{yy}(a,b)k^2$ vara en binär kvadratisk form. Då gäller med notation från (20) och (21) att,

1. $f''_{xx}(a,b)f''_{yy}(a,b) - f''_{xy}(a,b)^2 > 0$ och $f''_{xx}(a,b) > 0 \Longrightarrow Q$ *positivt definit.*
2. $f''_{xx}(a,b)f''_{yy}(a,b) - f''_{xy}(a,b)^2 > 0$ och $f''_{xx}(a,b) < 0 \Longrightarrow Q$ *negativt definit.*
3. $f''_{xx}(a,b)f''_{yy}(a,b) - f''_{xy}(a,b)^2 < 0 \Longrightarrow Q$ *indefinit.*
4. $f''_{xx}(a,b)f''_{yy}(a,b) - f''_{xy}(a,b)^2 = 0 \Longrightarrow Q$ *semidefinit.*

Från *sats 1* följer en ny sats, *sats 2*.

4.2 Sats 2

Låt $D \subseteq \mathbb{R}^2$ och vara en öppen mängd, $f \in C^3(D)$ och antag att $(a,b) \in D$ är en stationär punkt till f. Då gäller med notation från (20) och (21) att,

1. Om Q är positivt definit $\Longrightarrow (a,b)$ är en strikt lokal minimipunkt.
2. Om Q är negativt definit $\Longrightarrow (a,b)$ är en strikt lokal maximipunkt.
3. Om Q är indefinit $\Longrightarrow (a,b)$ är en sadelpukt.

5 Vektorvärda funktioner

Tidigare har det främsta fokuset legat på reellvärda funktioner, alltså funktioner som har $\mathbb{R}$ som målmängd. Nu kommer vi även att diskutera vektorvärda funktioner. En vektorvärd funktion av n variabler är en funktion som har $\mathbb{R}^n$ som definitionsmängd och $\mathbb{R}^m$ som målmängd, där n och m tillhör de naturliga talen. Detta kan skrivas på följande sätt:

$$\mathbb{F}(x_1, ..., x_n) = (F_1(x_1, ..., x_n), ..., F_m(x_1, ..., x_n)). \tag{22}$$

$\mathbb{F}$ sägs vara differentierbar om varje F_i är differentierbar.

5.1 Linjärisering

Vektorvärda funktioner kan precis som reellvärda funktioner linjäriseras. Det är sedan tidigare känt att om varje F_i är differentierbar existerar det en funktion ρ_i för varje i så att F_i i närheten av punkten $\boldsymbol{a}$ kan skrivas enligt följande:

$$F_i(\boldsymbol{a}+\boldsymbol{h}) = F_i(\boldsymbol{a}) + \nabla F_i(\boldsymbol{a}) \cdot \boldsymbol{h} + \|\boldsymbol{h}\| \cdot \rho_i(\boldsymbol{h}), \tag{23}$$

där $\rho_i(\boldsymbol{h})$ går mot 0 då $\boldsymbol{h}$ går mot $\mathbf{0}$.

Genom att stapla samtliga av dessa ekvationer på varandra i en kolumn fås följande samband:

$$\mathbb{F}(\boldsymbol{a}+\boldsymbol{h}) = \mathbb{F}(\boldsymbol{a}) + D\mathbb{F}(\boldsymbol{a}) * \boldsymbol{h} + \|\boldsymbol{h}\| \, \boldsymbol{\rho}(\boldsymbol{h}), \tag{24}$$

där $\boldsymbol{\rho}(\boldsymbol{h})$ går mot $\mathbf{0}_m$ då $\boldsymbol{h}$ går mot $\mathbf{0}_n$.

$D\mathbb{F}(\boldsymbol{a})$ kallas för funktionalmatrisen till $\mathbb{F}$ i punkten $\boldsymbol{x} = \boldsymbol{a}$. Funktionalmatrisen är den $m \times n$ matrisen vars (i,j):te element är:

$$\frac{\partial F_i}{\partial x_j}(\boldsymbol{a}). \tag{25}$$

5.2 Kedjeregeln för sammansättning av vektorvärda funktioner

För att genomföra kedjeregeln för en sammansättning av vektorvärda funktioner används funktionalmatrisen. Om vi definerar funktionerna $\mathbb{G}$ och $\mathbb{F}$ enligt följande: $\mathbb{G} : \mathbb{R}^n \to \mathbb{R}^m$ och $\mathbb{F} : \mathbb{R}^m \to \mathbb{R}^p$ så blir sammansättningen $\mathbb{F} \circ \mathbb{G} : \mathbb{R}^n \to \mathbb{R}^p$.

Det finns en sats som säger att om $\mathbb{F}$ och $\mathbb{G}$ är differentierbara, så är även $\mathbb{F} \circ \mathbb{G}$ och i varje punkt $\boldsymbol{a} \in \mathbb{R}^n$ gäller att:

$$D(\mathbb{F} \circ \mathbb{G})(\boldsymbol{a}) = D\mathbb{F}(\mathbb{G}(\boldsymbol{a})) * D\mathbb{G}(\boldsymbol{a}). \tag{26}$$

5.3 Variabelbyten

Ett variabelbyte för i flervariabelanalys syftar till att lösa samma problem som variabelbyten i en variabel gör. I envariabelanalys utförs ett variabelbyte till exempel för att skriva om en ekvation man ej kan lösa till en form man känner igen. Sedan löses den familjära ekvationen och därefter byter man tillbaka till sin ursprungliga variabel.

För att utföra variabelbytet i en dimension används en substitutionsfunktion $\varphi(x) = y$. För att variabelbytet ska vara användbart måste man kunna byta tillbaka från y till x. Detta görs med $\varphi^{-1}(y) = x$ ty $\varphi^{-1}(\varphi(x)) = x$ per definition av invers funktion. Man förutsätter alltså att φ är injektiv och därmed inverterbar. Om φ dessutom skulle vara deriverbar är även dess invers det och inversens derivata är $\varphi^{-1\prime}(x) = \frac{1}{\varphi'(\varphi^{-1}(x))}$.

Tanken är exakt samma när man övergår till variabelbyte i högre dimensioner. Det krävs en inverterbar bytesfunktion $\mathbb{F}(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}^n, n \in \mathbb{N}$ precis som

för en variabel. För att kontrollera att $\mathbb{F}$ är en inverterbar funktion antas att den är differentierbar. Därefter linjäriseras $\mathbb{F}$ enligt (24) men feltermen anses försumbar och stryks vilket ger följande uttryck:

$$\mathbb{F}(\boldsymbol{a}+\boldsymbol{h}) = \mathbb{F}(\boldsymbol{a}) + D\mathbb{F}(\boldsymbol{a}) * \boldsymbol{h}, \tag{27}$$

där $\mathbb{F}$ har linjäriserats i punkten $\boldsymbol{x} = \boldsymbol{a}$. Notera att $\boldsymbol{a}$ är fixt vilket innebär att $\mathbb{F}$ kan ses som en funktion av en vektor, $\boldsymbol{h}$. När funktionen betraktas på det sättet är den ekvivalent med följande linjära transformation $\boldsymbol{F}(\boldsymbol{h}) = A\boldsymbol{h} + \boldsymbol{b}$. Denna transformation är inverterbar då matrisen A har en nollskilld determinant. Detta är ekvivalent med att $\det D\mathbb{F}(\boldsymbol{a}) \neq 0$. Om detta villkor är uppfyllt, är substitutionsfunktionen inverterbar i punkten $\boldsymbol{a}$ och variabelbytet går att utföra.

Ett av de vanligaste variabelbytena för vektorvärda funktioner är bytet mellan kartesiska och polära koordinater. Som exempel görs detta i två variabler

$$\mathbb{F} : (r,\theta) \to (x,y), \quad \mathbb{F}(r,\theta) = (r\cdot\cos\theta, r\cdot\sin\theta)$$
$$\mathbb{F}^{-1} : (x,y) \to (r,\theta), \quad \mathbb{F}^{-1}(x,y) = (\sqrt{x^2+y^2}, \arctan\frac{y}{x})$$

genom dessa bytesfunktioner. Funktionalmatrisen för $\mathbb{F}$ bestäms enligt (25) på följande vis

$$\begin{bmatrix} \dfrac{\partial F_1}{\partial r}(r,\theta) & \dfrac{\partial F_1}{\partial \theta}(r,\theta) \\ \dfrac{\partial F_2}{\partial r}(r,\theta) & \dfrac{\partial F_2}{\partial \theta}(r,\theta) \end{bmatrix} = \begin{bmatrix} cos(\theta) & -r\cdot\sin(\theta) \\ sin(\theta) & r\cdot\cos(\theta) \end{bmatrix}$$

och dess determinant är

$$\begin{vmatrix} cos(\theta) & -r\cdot\sin(\theta) \\ sin(\theta) & r\cdot\cos(\theta) \end{vmatrix} = r\cdot\cos^2(\theta) + r\cdot\sin^2(\theta) = r$$

som är nollskilld $\forall r,\theta$ s.a $r \neq 0$. Detta innebär att $\mathbb{F}$ är inverterbar för alla punkter förutom origo ty alla vinklar leder till origo när $r = 0$ vilket förhindrar $\mathbb{F}$ från att vara injektiv.