Flervariabelanalys Lv.2

Elias Nilsson, Hannes Johansson, Sofia Reiner,
Mert Roj Tekin, William Blomström, Santiago Molinos.

20 februari 2019

Sammandrag

Denna rapport är en sammanfattning av andra läsveckans föreläsningar i kursen Flervariabelanalys, MVE035.

1 Taylors formel i Flervariabelanalys

I tidigare kurser i matematisk analys har man kunnat studera och approximera envariabelfunktioner med hjälp av Taylors formel. Detta genom att skriva om funktionen av en viss punkt till ett oändligt polynom.

För att studera en differentialekvation av flera variabler kan man också använda sig av Taylorserier. Först måste dock den tidigare formeln, som bara gäller för envariabelfunktioner, generaliseras. Man anpassar Taylors formel till att inkludera fler variabler genom att använda sig av kedjeregeln.

Vi minns Taylors formel i en variabel,

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

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)\,(x-a)^j}{j!} + \frac{f^{(k+1)}(\xi)\,(x-a)^{k+1}}{(k+1)!} \tag{1}$$

för något $\xi \in (a, x)$.
Vi inför notationsbytet $x = a + h$,

$$f(a+h) = \sum_{j=0}^{k} \frac{f^{(j)}(a)\,(h)^j}{j!} + \mathcal{O}(h^{k+1}) \tag{2}$$

där $\mathcal{O}(h^{k+1})$ ersätter feltermen i Taylors formel och där $\mathcal{O}(h^{k+1}) \to 0$ när h går mot 0. Summan i (2) kallas för *Taylorapproximationen av grad k till f i punkten x=a*. Vi noterar att för $k = 1$ är taylorutvecklingen en förfining av linjariseringsformeln, som vi har definierat tidgare i kursen. Här ersätts $\rho(h)$ med $\mathcal{O}(h^{k+1})$ i formeln.

$$f(a+h) = f(a) + f'(a) \cdot h + \mathcal{O}(h^2) \tag{3}$$

Om vi istället väljer $k = 2$ får vi en bättre approximation av f i närheten av x=a. Eftersom andraderivatan inkluderas i utvecklingen kan vi dessutom använda approximationen för att klassificera kritiska punkter. Det vill säga en punkt där funktionens alla partiella derivator är noll.

För att generalisera Taylors formel till att gälla för funktioner med fler variabler använder man sig av kedjeregeln. Vi har tidigare i kursen generaliserat kedjeregeln och definierat den till att gälla flervariabelfunktioner.

Sats 2.6.10 Låt $(a, b) \in \mathbb{R}^2$ och låt $f \in C^3(D)$ där D är en omgivning av $x = (a, b)$. Då gäller att,

$$\begin{aligned} f(a+h, b+h) &= f(a,b) + h \cdot f_x(a,b) + k \cdot f_y(a,b) + \\ &+ \frac{1}{2}(h^2 \cdot f_{xx}(a,b) + 2 \cdot h \cdot k \cdot f_{xy}(a,b) + k^2 \cdot f_{yy}(a,b)) + \mathcal{O}(\| \boldsymbol{h}^3 \|) \end{aligned} \tag{4}$$

där $\boldsymbol{h} = (h, k)$

Satsen ger Tayloutvecklingen av grad 2 till f kring punkten (a,b). Man bevisar sats 2.6.10 genom att med hjälp av kedjeregeln reducera flervariabelfunktionen till en variabel så att Taylors formel kan användas. Om vi betraktar funktionen $F(t) = f(a + th, b + th)$ så ger kedjeregeln att funktionens förstaderivata blir,

$$F'(t) = h \cdot f_x(a + th, b + tk) + k \cdot f_y(a + th, b + tk) \tag{5}$$

Genom att återigen derivera $F'(t)$ och sedan sätta $t = 1$ kan vi sätta in derivatorna i Taylors formel för en variabel. Vi får då utvecklingen som visas i Sats 2.6.10. Samma procedur kan generaliseras både till fler variabler och till högre ordning. Vi kan visa det generella fallet, alltså approximationen av grad k i punkten $\boldsymbol{a}$, (en vektor).

Vi får induktivt från kedjeregeln att för $F(t) = f(\boldsymbol{a} + t\boldsymbol{h})$, där $\boldsymbol{a}$ och $\boldsymbol{h}$ är vektorer så är den j:te derivatan,

$$F^j = (\boldsymbol{h} \cdot \nabla)^j f \tag{6}$$

där ∇ är en vektor som innehåller alla f:s partiella derivator. Insättning i Taylors formel för envariabler ger oss då Taylors sats i det generella fallet.

Taylors Sats: Låt $f \in C^{k+1}(D)$ och $\boldsymbol{a} \in D$. Då gäller att,

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

där $\mathcal{O}(\| \boldsymbol{h}^{k+1} \|) \to 0$ när h går mot 0.

2 Klassificering av kritiska punkter

En vanlig sak att undersöka hos en funktion är om den har lokala extrempunkter, det vill säga om funktionen har maximipunkter eller minimipunkter.

Definition 1: Låt $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}$ och $\boldsymbol{a} \in \mathbb{R}^n$. $\boldsymbol{f}$ sägs ha en lokal minimum respektive maximum i $\boldsymbol{x} = \boldsymbol{a}$ om

(1) $\boldsymbol{f}$ är definierad i en omgivning av $\boldsymbol{a}$.

(2) $\exists\ \delta > 0 : \| \boldsymbol{x} - \boldsymbol{a} \| < \delta \Rightarrow \boldsymbol{f(x)} \geq \boldsymbol{f(a)}$ resp. $\boldsymbol{f(x)} \leq \boldsymbol{f(a)}$.

Dessutom sägs funktionen ha en ***sträng*** lokal maximipunkt resp. minimipunkt om för $\boldsymbol{x} \neq \boldsymbol{a}$ gäller $f(\boldsymbol{x}) < f(\boldsymbol{a})$ resp $f(\boldsymbol{x}) > f(\boldsymbol{a})$

Från envariabelanalys har vi att derivatan i dessa extrempunkter är lika med 0. I Definition 1 ser vi dock att det inte finns något krav på att de partiella derivatorna behöver vara 0 i punkten $\boldsymbol{x} = \boldsymbol{a}$. Dock kan vi ställa upp följande resultat,

Proposition: Låt $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{a} \in \mathbb{R}^n$ (och D är en öppen omgivning till $\boldsymbol{a}$). Om $\boldsymbol{f}$ har en lokal extrempunkt, antingen lokal max. eller lokal min. i $\mathbf{x} = \mathbf{a}$ och är differentierbar i $\mathbf{x} = \mathbf{a}$ då gäller det att $\nabla \boldsymbol{f}(\mathbf{a}) = \mathbf{0}$, dvs. alla de partiella

derivatorna är noll i $\mathbf{x} = \mathbf{a}$.
Detta kan bevisas genom att använda sig av linjariseringsformeln, anta att $\nabla \boldsymbol{f}(\mathbf{a}) \neq \mathbf{0}$ och sedan visa att detta motverkar Definition 1.

Slutsatsen kan dras att om en funktion har extremum i en punkt implicerar det att $\nabla \boldsymbol{f}(\mathbf{a}) = \mathbf{0}$ i punkten, men inte vice versa. Vi har följande definition,

Definition 2: Låt $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}$ vara differentierbar i punkten $\boldsymbol{a}$. Om $\nabla \boldsymbol{f}(\mathbf{a}) = \mathbf{0}$ sägs punkten $\boldsymbol{a}$ vara en ***kritisk/stationär punkt*** till $\boldsymbol{f}$.

Eftersom man inte kan veta om en stationär punkt också är ett lokalt extremum måste man använda sig av Taylorutvecklingar runt stationärpunkten för att kunna klassificera den. Som tidigare nämnt använder man andra ordningens Taylorutveckling som inkluderar andraderivata. Detta eftersom vi från Propositionen har att de partiella förstaderivatorna är 0.

Om vi antar att det för en funktion $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}$ gäller att $f \in C^3$ i en omgivning av en kritisk punkt (a, b), kan vi utnyttja andra ordningens Taylorutveckling och skriva,

$$
\begin{aligned}
f(a+h, b+h) - f(a,b) &= hf_x(a,b) + kf_y(a,b) + \\
&+ \frac{1}{2}(h^2 f_{xx}(a,b) + 2hkf_{xy}(a,b) + k^2 f_{yy}(a,b)) + \mathcal{O}(\| \boldsymbol{h}^3 \|)
\end{aligned}
$$

Här kommer f_x och f_y termerna bli 0 och $\mathcal{O}(\| \boldsymbol{h}^3 \|)$ kan i de flesta fall försummas. Vi ser då att om vänsterledet är mindre än 0 så kommer det finnas ett maximum i punkten (a, b) och ett minimum om differensen i vänsterledet är större än 0. Tecknet på differensen kommer självklart bero på högerledet som vi kan skriva på formen $Ah^2 + 2Bhk + Ck^2$ där $A = \frac{1}{2} f_{xx}(a,b)$, $B = \frac{1}{2} f_{xy}(a,b)$ och $C = \frac{1}{2} f_{yy}(a,b)$.

Definition 3: En funktion $Q : \mathbb{R}^n \to \mathbb{R}$ kallas för en ***binär kvadratisk form*** om $\exists\ A, B, C\ \in \mathbb{R}\ Q(h,k) = Ah^2 + 2Bhk + Ck^2$.

Definition 4: Den binära kvadratiska formen Q sägs vara

(1) ***positiv definit*** om $Q(h,k) > 0\ \forall (h,k) \neq (0,0)\ \ \Rightarrow$ Då är punkten (a, b) ett ***strängt lokalt minimum***.

(2) ***negativ definit*** om $Q(h,k) < 0\ \forall (h,k) \neq (0,0)\ \ \Rightarrow$ Då är punkten (a, b) ett ***strängt lokalt maximum***.

(3) ***indefinit*** om $Q(h_1,k_1) < 0$ och $Q(h_2,k_2) < 0\ \ \Rightarrow$ Då är punkten (a, b) varken lokalt maximum eller lokalt minimum, dvs. är ett så kallad ***sadelpunkt***.

(4) positiv (eller negativ) ***semidefinit*** om $Q(h,k) \geq 0$ (respektive $Q(h,k) \leq 0$) och $\exists\ (h_1,k_1) \neq (0,0)$ där $Q(h_1,k_1) = 0$.

Att Q är indefinit kan jämföras i envariabel med en terasspunkt. Där finns det varken ett maximum eller minimum. Om Q är semidefinit kan ingenting sägas om den kritiska punktens klassificering.
Genom att vidare analysera högerledet som är skriven på binär kvadratisk form kan slutsatser dras om vänsterledets tecken. Vi får satsen,

Sats: Låt $Q(x,y) = Ax^2 + 2Bxy + Cy^2$. 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.

Genom analys av tecken och jämförelse med Definition 4 bevisas samtliga fall i satsen.

3 Vektorvärd funktion

Definition: Låt m,n $\in \boldsymbol{N}$.
En funktion $\boldsymbol{F} : \boldsymbol{R}^n \to \boldsymbol{R}^m$ kallas för vektorvärd en funktion.
I fallet då $m = 1$, då kallas funktionen istället för en skalärvärd funktion.
Notera: $\boldsymbol{F}(x_1, .., x_n) = (\boldsymbol{F_1}(x_1, ...x_n), ..., \boldsymbol{F_m}(x_1, ...x_n))$
Definition: $\boldsymbol{F}$ sägs vara differentierbar om varje $\boldsymbol{F}_i$ är differentierbar.

4 Linjärisering

Låt varje $\boldsymbol{F}_i$ vara differentierbar i en punkt $\boldsymbol{x} = \boldsymbol{a} \in \boldsymbol{R}^n$, säg

$$\boldsymbol{F_i}(\boldsymbol{a}+\boldsymbol{h}) = \boldsymbol{F_i}(\boldsymbol{a}) + \boldsymbol{h} \cdot \nabla \boldsymbol{F_i}(\boldsymbol{a}) + \parallel \boldsymbol{h} \parallel \rho_i(\boldsymbol{h}), \tag{8}$$

där $\rho_i(\boldsymbol{h}) \to 0$, då $\boldsymbol{h} \to 0$.

Notera: $\boldsymbol{u} = (u_1, u_2, ..., u_n), \boldsymbol{v} = (v_1, v_2, ..., v_n)$,
Då är skalärprodukten $\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{i=1}^{n} \boldsymbol{u}_i \boldsymbol{v} \boldsymbol{i} = \boldsymbol{u^T} \boldsymbol{v}$

Detta kan vi applicera på (8) och få följande uttryck:

$$\boldsymbol{F_i}(\boldsymbol{a}+\boldsymbol{h}) = \boldsymbol{F_i}(\boldsymbol{a}) + \left[\frac{\partial \boldsymbol{F}_i(\boldsymbol{a})}{\partial x_i}, ..., \frac{\partial \boldsymbol{F}_i(\boldsymbol{a})}{\partial x_n}\right] \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} + \parallel \boldsymbol{h} \parallel \rho_i(\boldsymbol{h}),$$

som gäller för $i = 1, 2, 3, ..., m$.
Ställer man upp dessa m stycken ekvationer på varandra får man följande uttryck:

$$\begin{bmatrix} \boldsymbol{F_1}(\boldsymbol{a}+\boldsymbol{h}) \\ \boldsymbol{F_2}(\boldsymbol{a}+\boldsymbol{h}) \\ \vdots \\ \boldsymbol{F_m}(\boldsymbol{a}+\boldsymbol{h}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{F_1}(\boldsymbol{a}) \\ \boldsymbol{F_2}(\boldsymbol{a}) \\ \vdots \\ \boldsymbol{F_m}(\boldsymbol{a}) \end{bmatrix} + \begin{bmatrix} \text{En m} \times \text{n} \\ \text{matris vars(i,j)-} \\ \text{element är:} \\ \frac{\partial \boldsymbol{F}_i(\boldsymbol{a})}{\partial x_j} \end{bmatrix} \begin{bmatrix} \boldsymbol{F_1}(\boldsymbol{a}) \\ \boldsymbol{F_2}(\boldsymbol{a}) \\ \vdots \\ \boldsymbol{F_m}(\boldsymbol{a}) \end{bmatrix} + \parallel \boldsymbol{h} \parallel \begin{bmatrix} \rho_1(\boldsymbol{h}) \\ \rho_2(\boldsymbol{h}) \\ \vdots \\ \rho_m(\boldsymbol{h}) \end{bmatrix}$$

Vilket ger oss linjäriseringsformeln:

$\boldsymbol{F}(\boldsymbol{a}+\boldsymbol{h}) = \boldsymbol{F}(\boldsymbol{a}) + \boldsymbol{F'}(\boldsymbol{a}) \cdot \boldsymbol{h} + \parallel \boldsymbol{h} \parallel \rho(\boldsymbol{h})$, där
$\rho(\boldsymbol{h}) \to \vec{0}_m$, då $\boldsymbol{h} \to \vec{0}_n$

Terminologi: $\boldsymbol{F'(a)}$ kallas för funktionalmatriser eller Jacobimatrisen för den vektorvärda funktionen $\boldsymbol{F}$

5 Kedjeregeln för vektorvärda fuktioner

DEFINITION: Låt $\mathbf{F} : \mathbf{R^n} \to \mathbf{R^m}$ och $\mathbf{G} : \mathbf{R^p} \to \mathbf{R^q}$, som antas vara C^1. Då $n = q$ gäller att $\mathbf{F} \circ \mathbf{G} : \mathbf{R^p} \to \mathbf{R^m}$ och

(1) $\mathbf{F} \circ \mathbf{G}$ också C^1.

(2) $(\mathbf{F} \circ \mathbf{G})' = \mathbf{F}' * \mathbf{G}' \Rightarrow (\mathbf{F} \circ \mathbf{G})'(\mathbf{a}) = \mathbf{F}'\big(\mathbf{G}(\mathbf{a})\big) * \mathbf{G}'(\mathbf{a})$ för någon punkt $\mathbf{a} \in \mathbf{R^p}$. Här står $*$ för matrisprodukt.

Speciallfallet $m = n$, $\mathbf{F} : \mathbf{R^n} \to \mathbf{R^n}$, kallas för ett n-dimensionellt ***vektorfält***.

6 Formella satsen (sats 3.3.2 inversfunktionsatsen)

Låt $y = F(x)$ vara en C^1-funktion från $\mathbb{R}$ till $\mathbb{R}$ och a en punkt i definitionsmängden sådana att determinanten av $(F'(a)) \neq 0$
Då finns det öppna omgivningar U och V av a respektive $b = F(a)$ sådana att $F : V \to U$ är bijektiv och $F^{-1} : V \to U$ också är C.

7 Implicita funktionssatsen 3.4

Det är naturligt att framställa enhetscirkeln som en nivåkurva av $f(x.y) = 1$ där $f(x, y) = x^2 + y^2$

7.1 Implicit derivering

Idéen är att ekvationen $x^2 + y^2 = 1$ implicit definierar y som en funktion av x så vi kan derivera båda leden med avssende på x med hjälp av kedjeregeln och få,

$$\frac{d}{dx}(HL) = \frac{d(1)}{dx} = 0 \text{ och } \frac{d}{dx}(VL) = \frac{d}{dy}(x^2+y^2) = 2x + \frac{d}{dx}(y(x))^2 = 2x + 2y\frac{dy}{dx}$$

$$\Rightarrow 0 = 2x + 2y\frac{dy}{dx} \Rightarrow \frac{dy}{dx} = \frac{-x}{y}$$

Hela härledningen är en ogiltig beräkning då $y = 0$ för att ekvationen inte är en entydlig funktion av x i en godtycklig liten omgivning av antingen $(1, 0)$ eller $(-1, 0)$.