

Sammanfattning av läsvecka 2 i Flervariabelanalys

Mahmut Osmanovic

William Colliander

Gustav Axelsson

Andreas Bauner

Oliver Lundblad

Mathilda Gustafsson

23 februari 2020

Innehåll

1	Bevis av satser	3
1.1	Kedjeregeln för reellvärda funktioner	3
1.2	Partiella derivator	4
2	Taylors formel för funktioner av två variabler	5
3	Kritiska punkter	6
4	Klassificering av kritiska punkter	8
5	Vektorvärda funktioner	9
5.1	Linjärisering av vektorvärda funktioner	9
5.2	Kedjeregeln för vektorvärda funktioner	10
5.2.1	Ett exempel	10
6	Variabelbyten	11
6.1	Inversa funktionssatsen	11

1 Bevis av satser

1.1 Kedjeregeln för reellvärda funktioner

Sats. Låt funktionen $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ bestå av n stycken variabler, det vill säga $f(x_1, x_2, \dots, x_n)$, och låt $g_1(t), g_2(t), \dots, g_n(t)$ vara n stycken funktioner av en variabel. Låt dessutom alla ovannämnda funktioner vara differentierbara funktioner. Differentierbarheten implicerar att var och en funktionerna enskilt är kontinuerliga, och det gäller även att deras sammansättning $f(g_1(t), g_2(t), \dots, g_n(t))$ är en kontinuerlig funktion. För derivatan gäller att

$$\frac{d}{dt}f(g_1(t), g_2(t), \dots, g_n(t)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(g_1(t), g_2(t), \dots, g_n(t)) \frac{dg_i}{dt} \quad (1)$$

Bevis. Låt $F(t) = f(g_1(t), g_2(t), \dots, g_n(t))$. Derivatan av f i en punkt a är per definition $F'(a) = \lim_{h \rightarrow 0} \frac{F(a+h) - F(a)}{h}$. Beviset går ut på att visa att gränsvärdet existerar och är lika med derivatan i satsen ovan.

Insättning av $t = (a + h)$ ger $F(a + h) = f(g_1(a + h), g_2(a + h), \dots, g_n(a + h))$. Sätt $\mathbf{a} = (g_1(a), g_2(a), \dots, g_n(a))$ och $(\mathbf{a} + \mathbf{h}) = (g_1(a + h), g_2(a + h), \dots, g_n(a + h))$, vilket ger att $F(a + h) - F(a) = f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})$.

Att f är differentierbar innebär att

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\| \cdot \rho(\mathbf{h}) \text{ där } \rho(\mathbf{h}) \rightarrow 0 \text{ då } \mathbf{h} \rightarrow 0. \quad (2)$$

Observera att

$$\mathbf{h} = (\mathbf{a} + \mathbf{h}) - \mathbf{a} = (g_1(a + h) - g_1(a), g_2(a + h) - g_2(a), \dots, g_n(a + h) - g_n(a)) \quad (3)$$

varefter applikation av linjäriseringsformeln på var och en av variablerna ger att

$$\mathbf{h} = g'_1(a)h + h\rho_1(h), \dots, g'_n(a)h + h\rho_n(h). \quad (4)$$

Täljaren i derivatans definition skrivs om med hjälp av linjärisering:

$$F(a + h) - F(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a})(g'_i(a)h + h\rho_i(h)) + \|\mathbf{h}\|\rho(\mathbf{h}). \quad (5)$$

Insättning i derivatans definition ger

$$\frac{F(a+h) - F(a)}{h} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a})(g'_i(a)) + \sum_i^n \frac{\partial f}{\partial x_i} \rho_i(h) + \frac{\|\mathbf{h}\| \rho(\mathbf{h})}{h} \quad (6)$$

Observera att den andra termen i (6) går mot noll ty $\rho_i(h) \rightarrow 0$ då $h \rightarrow 0$. Observera även att $\mathbf{h} = h(g'_1(a) + \rho_1(h), \dots, g'_n(a) + \rho_n(h))$, vilket implicerar att om $h \rightarrow 0$ så går $\mathbf{h} \rightarrow 0$ och därmed går $\rho(\mathbf{h}) \rightarrow 0$.

Observera också att: $\frac{\mathbf{h}}{h} = \|g'_1(a) + \rho_1(h), \dots, g'_n(a) + \rho_n(h)\| \rightarrow \sqrt{\sum_{i=1}^n (g'_i(a))^2}$ då $h \rightarrow 0$, varför

$$\frac{d}{dt} f(g_1(t), g_2(t), \dots, g_n(t)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} g_i(t) \frac{dg_i}{dt}. \quad (7)$$

□

1.2 Partiella derivator

Sats. Satsen visar att den ordningen efter vilken vi deriverar partiellt inte har någon betydelse för det slutgiltiga svaret, vilket förblir detsamma.

Låt $D \subseteq \mathbb{R}^2$, $f \in C^2(D)$ och $(a, b) \in D$. Då gäller att

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b). \quad (8)$$

Bevis. Vi definierar funktionen $F(x) = f(x, b+h) - f(x, b)$ och även funktionen $G(y) = f(a+h, y) - f(a, y)$. Observera att $F(a+h) - F(a) = G(b+h) - G(b)$, och att $f \in C^2$ medför att F och G är differentierbara.

Applikation av differentialkalkylens medelvärdessats på observationen ovan ger att

$$hF'(a + \theta_1 h) = hG'(b + \theta_2 h), \quad \exists \theta_1, \theta_2 \in (0, 1), \quad (9)$$

vilket ger att

$$\frac{\partial f}{\partial x}(a + \theta_1 h, b + h) - \frac{\partial f}{\partial x}(a + \theta_1 h, b) = \frac{\partial f}{\partial y}(a + h, b + \theta_2 h) - \frac{\partial f}{\partial y}. \quad (10)$$

Vi applicerar medelvärdessatsen igen och får

$$\frac{\partial^2 f}{\partial x \partial y}(a + \theta_1 h, b + \theta_3 h) = \frac{\partial^2 f}{\partial y \partial x}(a + \theta_4 h, b + \theta_2 h), \exists \theta_3, \theta_r \in (0, 1). \quad (11)$$

Att $h \rightarrow 0$ ger:

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b) \quad (12)$$

Vilket bevisar sats 8.

□

2 Taylors formel för funktioner av två variabler

Definition 2.1. Låt $D \subseteq \mathbb{R}^2$ vara en öppen mängd och punkten $(a, b) \in D$. Låt $n \geq 0$, $f = f(x, y)$ och $f \in C^{n+1}(D)$. Då gäller

$$f(a + h, b + k) = \sum_{i=0}^n \frac{(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y})^i}{i!} f(a, b) + \frac{(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y})^{n+1}}{(n+1)!} f(a + \xi h, b + \xi k), \quad (13)$$

för något $\xi \in (0, 1)$.

Det viktigaste fallet av detta är andragsutveckling av en C^3 funktion av två variabler är

$$\begin{aligned} f(a + h, b + k) &= f(a, b) + [h f_x(a, b) + k f_y(a, b)] \\ &\quad + 1/2 [h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b)] + \mathcal{O}((h^2 + k^2)^{3/2}). \end{aligned} \quad (14)$$

För godtyckligt antal variabler så skrivs formeln för två variabler om med vektornotation istället

Definition 2.2. Låt $D \subseteq \mathbb{R}^m$ vara en öppen mängd och punkten $(\mathbf{a}, \mathbf{b}) \in D$ där $\|\mathbf{h}\|$ är tillräckligt liten. Låt $n \geq 0$, $f = f(\mathbf{x})$ och $f \in C^{n+1}(D)$. Då gäller

$$f(\mathbf{a} + \mathbf{h}) = \sum_{i=1}^n \frac{(\mathbf{h} \cdot \nabla)^i}{i!} f(\mathbf{a}) + \mathcal{O}(\|\mathbf{h}\|^{n+1}). \quad (15)$$

Bevisiden för Taylors formel i två variabler är sådan att vi definierar en funktion $F(t) = f(a + th, b + tk)$ och betraktar a, b, h, k som fixa tal. Sedan används Taylors formel för en variabel och kedjeregeln för att härleda formeln i definition 2.2.

3 Kritiska punkter

Till funktioner kan kritiska punkter finnas.

Definition 3.1. Punkten $\mathbf{a}$ är en kritisk punkt till funktionen f om

$$\nabla f(\mathbf{a}) = \mathbf{0}. \quad (16)$$

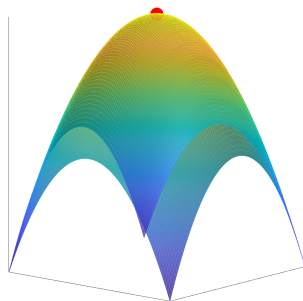
Det finns två typer av kritiska punkter. De är lokala extrempunkter och sadelpunkter.

Det finns två typer av lokala extrempunkter. De är maximipunkter och minimipunkter.

Definition 3.2. Punkten $\mathbf{a}$ är en maximipunkt till funktionen f om

$$\exists \delta > 0 : \|\mathbf{h}\| < \delta \implies f(\mathbf{a} + \mathbf{h}) < f(\mathbf{a}). \quad (17)$$

En maximipunkt är alltså en punkt som antar det största funktionsvärdet i en omgivning, se figur 1.

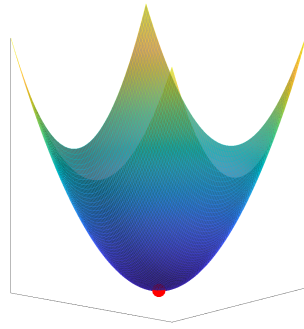


Figur 1: En maximipunkt till en funktion av två variabler.

Definition 3.3. Punkten $\mathbf{a}$ är en minimipunkt till funktionen f om

$$\exists \delta > 0 : \|\mathbf{h}\| < \delta \implies f(\mathbf{a} + \mathbf{h}) > f(\mathbf{a}). \quad (18)$$

En minimipunkt är alltså en punkt som antar det minsta funktionsvärdet i en omgivning, se figur 2.



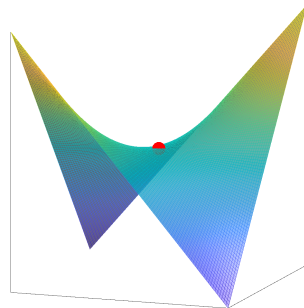
Figur 2: En minimipunkt till en funktion av två variabler.

Det finns en typ av sadelpunkter.

Definition 3.4. Den kritiska punkten $\mathbf{a}$ är en sadelpunkt till funktionen f om

$$\forall \delta > 0 \exists \|\mathbf{h}_1\|, \|\mathbf{h}_2\| < \delta : f(\mathbf{a} + \mathbf{h}_1) < f(\mathbf{a}) \wedge f(\mathbf{a} + \mathbf{h}_2) > f(\mathbf{a}). \quad (19)$$

En sadelpunkt är alltså en kritisk punkt som inte antar det största eller minsta funktionsvärdet i någon omgivning, se figur 3.



Figur 3: En sadelpunkt till en funktion av två variabler.

4 Klassificering av kritiska punkter

Analysen av kritiska punkter till funktioner kommer att begränsas till funktioner av två variabler.

Låt $D \subseteq \mathbb{R}^2$ vara en öppen mängd och $f : D \rightarrow \mathbb{R}$ en reellvärd funktion av två variabler. Antag att $\mathbf{a} \in D$ är en kritisk punkt till f . Eftersom $\mathbf{a}$ är en kritisk punkt till f så gäller ekvationen (16) i punkten $\mathbf{x} = \mathbf{a}$. Definiera $\mathbf{h} = (h, k)$. Från detta och taylorutveckling av f av grad två kring $\mathbf{x} = \mathbf{a}$ så gäller att

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \frac{1}{2}(h^2 f_{xx}(\mathbf{a}) + 2hk f_{xy}(\mathbf{a}) + k^2 f_{yy}(\mathbf{a})) + \mathcal{O}(\|\mathbf{h}\|^3). \quad (20)$$

Från ekvation 20 går det att göra följande definition.

Definition 4.1. Definiera funktionen $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2 \quad (21)$$

där $A, B, C \in \mathbb{R}$ och definieras som

$$\begin{cases} A = \frac{1}{2} \cdot f_{xx}(\mathbf{a}) \\ B = \frac{1}{2} \cdot f_{xy}(\mathbf{a}) \\ C = \frac{1}{2} \cdot f_{yy}(\mathbf{a}). \end{cases}$$

Q kallas för den *binära kvadratiska formen*.

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

Q sägs vara *negativt definit* om $Q(x, y) < 0, \forall (x, y) \neq (0, 0)$.

Q sägs vara *indefinit* om $\exists (x_1, y_1) \neq (0, 0) : Q(x_1, y_1) > 0$ och $\exists (x_2, y_2) \neq (0, 0) : Q(x_2, y_2) < 0$.

Q sägs vara *positivt semidefinit* om $Q(x, y) \geq 0, \forall (x, y) \neq (0, 0)$.

Q sägs vara *negativt semidefinit* om $Q(x, y) \leq 0, \forall (x, y) \neq (0, 0)$.

Sats. Om $AC - B^2 > 0 \wedge A > 0 \implies Q$ är positivt definit.

Om $AC - B^2 > 0 \wedge A < 0 \implies Q$ är negativt definit.

Om $AC - B^2 < 0 \implies Q$ är indefinit.

Om $AC - B^2 = 0 \implies Q$ är semidefinit.

Sats. Om Q är positivt definit för f i den kritiska punkten $\mathbf{x} = \mathbf{a} \implies \mathbf{x} = \mathbf{a}$ är ett strängt lokalt minimum till f .

Om Q är negativt definit för f i den kritiska punkten $\mathbf{x} = \mathbf{a} \implies \mathbf{x} = \mathbf{a}$ är ett strängt lokalt maximum till f .

Om Q är indefinit för f i den kritiska punkten $\mathbf{x} = \mathbf{a} \implies \mathbf{x} = \mathbf{a}$ är en sadelpunkt till f .

Anmärkning. Satsen säger inget om fallet då Q är semidefinit.

5 Vektorvärda funktioner

Definition. Om $\mathbf{F}$ är en funktion som har en $n \times 1$ vektor $(x_1, \dots, x_n)$ som input och en $m \times 1$ vektor

$$\mathbf{F}(x_1, \dots, x_n) = (F_1(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n)) = (y_1, \dots, y_m) \quad (22)$$

som output kallas funktionen för en vektorvärd funktion.

Definition. En vektorvärd funktion $\mathbf{F}$ sägs vara differentierbar om varje enskild funktion F_i som ger upphov till element y_i i outputvektorn $m \times 1$ är differentierbar.

5.1 Linjärisering av vektorvärda funktioner

Det går att linjärisera vektorvärda funktioner genom att skriva funktionen med vektorer som input och matriser som funktioner. Linjäriseringen blir då

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \mathbf{F}(\mathbf{a}) + D\mathbf{F}(\mathbf{a})\mathbf{h} + \|\mathbf{h}\| \rho(\mathbf{h}) \quad (23)$$

där $\mathbf{F}$ ($m \times 1$ vektor) är funktionen som linjäriseras, $D\mathbf{F}$ ($m \times n$ matrisen) är differentialen, $\mathbf{h}$ ($n \times 1$ vektorn) är en godtyckligt liten vektor. Mellan $D\mathbf{F}$ och $\mathbf{h}$ råder matrismultiplikation, och $\rho(\mathbf{h})$ är en funktion som går mot noll då $\mathbf{h}$ går mot nollmatrisen.

Definition. Matrisen $D\mathbf{F}(\mathbf{a})$ kallas för funktionalmatrisen till $\mathbf{F}$ i $\mathbf{a}$ där element på matrisposition (i, j) är $\frac{\partial F_i}{\partial x_j}$.

5.2 Kedjeregeln för vektorvärda funktioner

Sats. Låt $\mathbf{G}$ vara den vektorvärda funktionen

$$(x_1, \dots, x_n) \mapsto (G_1(x_1, \dots, x_n), \dots, G_m(x_1, \dots, x_n)) \quad (24)$$

och låt $\mathbf{F}$ vara den vektorvärda funktionen

$$(U_1, \dots, U_m) \mapsto (F_1(U_1, \dots, U_m), \dots, F_p(U_1, \dots, U_m)). \quad (25)$$

Låt dessa funktioner bilda den sammansatta funktionen $\mathbf{F} \circ \mathbf{G}$ där

$$(x_1, \dots, x_n) \mapsto (F_1[G_1(x_1, \dots, x_n), \dots, G_m(x_1, \dots, x_n)], \dots, F_p[G_1(x_1, \dots, x_n), \dots, G_m(x_1, \dots, x_n)]). \quad (26)$$

Om både $\mathbf{G}$ och $\mathbf{F}$ är differentierbara implicerar det att den sammansatta funktionen $\mathbf{F} \circ \mathbf{G}$ också är differentierbar och att

$$D(\mathbf{F} \circ \mathbf{G})(\mathbf{a}) = D\mathbf{F}(\mathbf{G}(\mathbf{a})) \cdot D\mathbf{G}(\mathbf{a}). \quad (27)$$

5.2.1 Ett exempel

Funktionen $\mathbf{v}(t) = \mathbf{v}(x(t), y(t))$ är en tvådimensionell vektorvärd funktion med tiden som input och hastigheten som output. Den beskriver hastighetsvektorn för ett föremål i ett plan. $\mathbf{F}(x, y)$ är en vektorvärd funktion som ändrar hastigheten och har därför hastigheten som input och hastigheten som output. För derivatan gäller att

$$D(\mathbf{F} \circ \mathbf{v})(t) = D\mathbf{F}(\mathbf{v}(t)) \cdot D\mathbf{v}(t). \quad (28)$$

Detta kan skrivas med hjälp av matriser som

$$\frac{\partial \mathbf{F}}{\partial t} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x} & \frac{\partial \mathbf{F}}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial t} \end{bmatrix}. \quad (29)$$

6 Variabelbyten

Ett av de vanligare variabelbytena är det från kartestiska till polära koordinater, men variabelbyten sker även vid integralberäkningar och lösning av differentialekvationer. Låt säga att vi vill genomföra ett variabelbyte från variablerna $(x_1, \dots, x_n)$ till $(y_1, \dots, y_n)$, och vill avgöra huruvida variabelbytet är ett godkänt byte. Vi definierar en vektorvärd funktion $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ som genomför variabelbytet, vilket innebär att

$$(x_1, \dots, x_n) \mapsto (F_1(x_1, \dots, x_n), \dots, F_n(x_1, \dots, x_n)) := (y_1, \dots, y_n). \quad (30)$$

För att variabelbytet ska vara godkänt måste funktionen $\mathbf{F}$ vara injektiv, det vill säga att flera olika inputs $(x_1, \dots, x_n)$ får inte ge samma output $(y_1, \dots, y_n)$. Huruvida en funktion är injektiv kan avgöras med hjälp av Inversa funktionssatsen.

6.1 Inversa funktionssatsen

Sats. Låt $\mathbf{y} = \mathbf{F}(\mathbf{x})$ vara en C^1 -funktion så att $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ och låt $\mathbf{a}$ vara en punkt i definitionsmängden till $\mathbf{F}$ så att $\det(D\mathbf{F}(\mathbf{a})) \neq 0$. Då gäller att det existerar omgivning u och v till $\mathbf{a}$ respektive $\mathbf{F}(\mathbf{a})$ så att funktionen $\mathbf{F} : u \rightarrow v$ är bijektiv. Dessutom gäller att $\mathbf{F}^{-1} : v \rightarrow u$ också är en C^1 -funktion.

Inversa funktionssatsen säger alltså att om funktionaldeterminanter av en funktion $\mathbf{F}$ är nollskild i en punkt $\mathbf{a}$ så är ett variabelbyte genom $\mathbf{F}$ godkänt i den punkten.

För bytet mellan kartesiska och polära koordinater, det vill säga

$$(x, y) \mapsto \left(\sqrt{x^2 + y^2}, \tan^{-1} \left(\frac{y}{x} \right) \right) := (r, \theta), \quad (31)$$

gäller att funktionaldeterminanten är nollskild då $r \neq 0$ vilket är detsamma som att $(x, y) \neq (0, 0)$. Variabelbytet är således giltigt utanför den punkten.

Referenser

- [1] Peter Hegarty, Universitetsföreläsningar i kursen Flervariabelanalys, Chalmers tekniska högskola, 2020.
- [2] Arne Persson, Lars-Christer Böiers, *Analys i Flera Variabler*, 2013.