

Sammanfattning LV2 MVE035/MVE600

Filip Berglund

Alma Cavallin

Atosa Daneshvar

Benjamin Hanson

Ann Heijenskjöld

Emma Larsson

Februari 2021

1 Clairaut's sats

Sats: För varje funktion $f(x) = f(x_1, \dots, x_n)$ av klass $\mathcal{C}^2$ gäller att $f''_{xy} = f''_{yx}$.

Satsen säger att ordningen man deriverar de partiella derivatorna med avseende på x och y inte spelar någon roll då de är av klassen $\mathcal{C}^2$ och de blandade andraderivatorna är kontinuerliga. Detta kan bevisas genom att använda medelvärdessatsen från envariabelanalysen och tillämpa den 2 gånger.

Exempel:

$$f(x, y) = 3x^2 - 4y^3 - 7x^2y^3$$

$$\frac{\partial f}{\partial x} = 6x - 14xy^3$$

$$\frac{\partial f}{\partial y} = -12y^2 - 21x^2y^2$$

$$\frac{\partial}{\partial x} f_x = \frac{\partial}{\partial x} (6x - 14xy^3) = 6 - 14y^3$$

$$\frac{\partial}{\partial y} f_x = \frac{\partial}{\partial y} (6x - 14xy^3) = -42xy^2$$

$$\frac{\partial}{\partial x} f_y = \frac{\partial}{\partial x} (-12y^2 - 21x^2y^2) = -42xy^2$$

$$\frac{\partial}{\partial y} f_y = \frac{\partial}{\partial y} (-12y^2 - 21x^2y^2) = -24y - 42x^2y$$

Som man kan se i exemplet innebär detta att man endast behöver derivera de partiella derivatorna 3 gånger och inte 4 för att ta reda på alla.

2 Taylors formel

Taylors formel i flera variabler bygger på den för envariabelsfallet, så en kort recap:

Låt $f : \mathbb{R} \rightarrow \mathbb{R}$ vara $k + 1$ gånger differentierbar i en omgivning D av punkten $x = a$ och låt $h > 0$ vara tillräckligt liten sådan att $[a - h, a + h] \subseteq D$. Då gäller att:

$$f(a + h) = \sum_{j=0}^k \frac{f^{(j)}(a)}{j!} h^j + \frac{f^{(k+1)}(\xi)}{(k+1)!} h^{k+1} \quad (1)$$

för något $\xi \subseteq (a, a + h)$. Om vi dessutom antar att $f \in \mathcal{C}^{k+1}$, dvs att dess $(k + 1)$:te derivata är en kontinuerlig funktion, då är den $(k + 1)$:te begränsad i varje slutet intervall.

$$f(a + h) \approx \sum_{j=0}^k \frac{f^{(j)}(a)}{j!} h^j \quad (2)$$

Denna kallas Taylorapproximationen av grad k till f i punkten $x = a$. Ofta vill man ta approximationen av grad 2 sådan att

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

Funktioner av flera variabler kan också Taylorutvecklas, för att underlätta med notationen ska vi göra alla härledningarna för funktioner $f(x, y)$ av två variabler.

Vi studerar en funktion $f(x, y)$ i närheten av en baspunkt (a, b) , och reducerar problemet till envariabelsfallet. Fixera h, k och betrakta funktionen $F(t)$ av en reell variabel t som ges av

$$F(t) = f(a + th, b + th). \quad (4)$$

Antag F kan Taylorutvecklas till vilken grad som helst. Enligt ekvation (1)

$$F(t) = \sum_{j=0}^k \frac{F^{(j)}(0)}{j!} t^j + \frac{F^{(k+1)}(\xi)}{(k+1)!} t^{k+1}, \xi \in (0, t) \quad (5)$$

Uttryck derivatorna till F i termer av f s partiella derivator. Sätt $g_1(t) = a + th, g_2(t) = b + tk$ och konstatera att $F(t) = f(g_1(t), g_2(t))$. Med kedjeregeln får vi alltså:

$$F'(t) = f_x(a + th, b + tk) \cdot g_1'(t) + f_y(a + th, b + tk) \cdot g_2'(t) = (hf_x + kf_y) \Big|_{(a+th, b+tk)}. \quad (6)$$

Notera att om $f \in \mathcal{C}^{k+1}$ så är H.L. av ekvation (6) i $\mathcal{C}^k$. Symboliskt kan vi skriva

$$F' = Df \quad (7)$$

där $D : \mathcal{C}^{k+1} \rightarrow \mathcal{C}^k$ är en så kallad differentialoperator,

$$D = h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}. \quad (8)$$

Konstatera att (7) gäller för vilket som helst par (f, F) som är relaterade enligt (6). Vi ska nu visa induktivt att för varje j så gäller

$$F^{(j)}(t) = D^j f(a + th, b + tk), D^j = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^j \quad (9)$$

så länge alla derivatorna finns. Uttrycket för D^j kan utvecklas enligt binomialsatsen så länge operatorerna $\frac{\partial}{\partial x}$ och $\frac{\partial}{\partial y}$ kommuterar m.a.o. så länge $f \in \mathcal{C}^j$ enligt Sats 2.5.9. Om $f \in \mathcal{C}^j$ kan vi skriva

$$\mathbf{a} = (a, b), \mathbf{h} = (h, k), \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \quad (10)$$

sådan att

$$D = h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} = \mathbf{h} \cdot \nabla \quad (11)$$

då kan ekvation (9) skrivas som

$$F^{(j)}(t) = (h \cdot \nabla)^j f(\mathbf{a} + t\mathbf{h}) \quad (12)$$

I synnerhet för $t = 0$ gäller

$$F^{(j)}(0) = (h \cdot \nabla)^j f(\mathbf{a}) \quad (13)$$

Antag nu att $f \in \mathcal{C}^{k+1}$. Vi kan sätta (13) i (5) för $j = 0, 1, \dots, k$, sedan sätta $t = 1$ eftersom $F(1) = f(\mathbf{a} + \mathbf{h})$ så får vi att

$$F(\mathbf{a} + \mathbf{h}) = \sum_{j=0}^k \frac{(h \cdot \nabla)^j}{j!} f(\mathbf{a}) + \frac{(h \cdot \nabla)^{k+1}}{(k+1)!} f(\xi), \quad (14)$$

där ξ är på raksträckan mellan $\mathbf{a}$ och $\mathbf{a} + \mathbf{h}$. Detta kan även skrivas som

$$F(\mathbf{a} + \mathbf{h}) = \sum_{j=0}^k \frac{(h \cdot \nabla)^j}{j!} f(\mathbf{a}) + O(\|\mathbf{h}\|^2). \quad (15)$$

Så en Taylorapproximation av grad 2, som en funktion av två variabler blir:

$$f(a + h, b + k) = f(a, b) + (hf_x + kf_y) + \frac{1}{2}(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}) \Big|_{(a,b)} + O(\|\mathbf{h}\|^3). \quad (16)$$

3 Klassificering av kritiska punkter

f sägs ha en kritisk punkt i $\mathbf{x} = \mathbf{a}$ om $\nabla f(\mathbf{a}) = \mathbf{0}$.

Det gäller att om f har en extrempunkt i punkten $\mathbf{a}$ är $\mathbf{a}$ en kritisk punkt till f (OBS! Omvändningen gäller ej).

Om vi antar att f är av klassen $\mathcal{C}^3$ (3 gånger kontinuerligt differentierbar) får vi med hjälp av Taylors formel uttrycket (17):

$$\begin{aligned} f(a+h, b+k) = & f(a, b) + hf_x(a, b) + kf_y(a, b) \\ & + 1/2(h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b) + O(\|\mathbf{h}\|^3) \end{aligned} \quad (17)$$

Om vi sedan antar att (a, b) är en kritisk punkt så följer det att $\nabla f(a, b) = (0, 0)$. Det vill säga att $f_x(a, b) = f_y(a, b) = 0$. Då kan vi skriva (17) som:

$$f(a+h, b+k) - f(a, b) = 1/2(h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b) + O(\|\mathbf{h}\|^3)$$

Här är VL differensen mellan den kritiska punkten och punkter i dess omgivning, allmänt är h och k mycket små så att omgivningen kan göras hur liten som helst. Om denna differens är positiv borde det innebära att den kritiska punkten är en minpunkt (eftersom värden i en omgivning av (a, b) är större). Likaså att en negativ differens innebär att (a, b) är en maxpunkt. Vi vill alltså avgöra tecknet för HL för att ta reda på punktens karaktär. Då konstanten $(1/2)$ framför de partiella andraderivatorna är positiv påverkar denna inte tecknet. Likaså kan vi bortse från resttermen $O(\|\mathbf{h}\|^3)$ då den oftast är försumbart liten. Det intressanta i högerledet är alltså

$$f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b)$$

vilket är på den binära kvadratiske formen:

$$Q(h, k) = Ah^2 + 2Bhk + Ck^2$$

I denna är $A = f_{xx}(a, b)$, $B = f_{xy}(a, b)$ och $C = f_{yy}(a, b)$.

3.1 Begrepp

Vi har följande möjligheter:

Om $Q(h, k) = Ah^2 + 2Bhk + Ck^2$ och (a, b) är en kritisk punkt till f så gäller:

- i) $Q(h, k)$ är positivt definit om det $\forall(h, k) \neq (0, 0)$ gäller att $Q(h, k) > 0$. Då är (a, b) en minpunkt.
- ii) $Q(h, k)$ är negativt definit om det $\forall(h, k) \neq (0, 0)$ gäller att $Q(h, k) < 0$. Då är (a, b) en maxpunkt.

- iii) $Q(h, k)$ är indefinit om det $\exists(h_1, k_1)$ s.a: $Q(h_1, k_1) > 0$ och $\exists(h_2, k_2)$ s.a: $Q(h_2, k_2) < 0$. Då är (a, b) en sadelpunkt.
- iv) $Q(h, k)$ är positivt semidefinit om det $\exists(h_1, k_1) \neq (0, 0)$ s.a: $Q(h_1, k_1) = 0$ och att det $\forall(h, k)$ gäller att $Q(h, k) \geq 0$.
- v) $Q(h, k)$ är negativt semidefinit om det $\exists(h_1, k_1) \neq (0, 0)$ s.a: $Q(h_1, k_1) = 0$ och att det $\forall(h, k)$ gäller att $Q(h, k) \leq 0$.

I fall iv) och v) kan ingen slutsats dras om den kritiska punktens karaktär. Man får använda andra metoder.

För att faktiskt bestämma vilken av de ovan nämnda som $Q(h, k)$ antar vid insättandet av den kritiska punkten (a, b) kan man antingen kvadratkomplettera uttrycket och bestämma karaktären därifrån eller också med hjälp av följande mycket användbara sats:

3.2 Sats 1.

Om $Q(h, k) = Ah^2 + 2Bhk + Ck^2$ så gäller:

- i) Om $AC - B^2 > 0$ och $A > 0 \equiv Q(h, k)$ positivt definit $\equiv (a, b)$ minpunkt.
- ii) Om $AC - B^2 > 0$ och $A < 0 \equiv Q(h, k)$ negativt definit $\equiv (a, b)$ maxpunkt.
- iii) Om $AC - B^2 < 0 \equiv Q(h, k)$ indefinit $\equiv (a, b)$ sadelpunkt.
- iv) Om $AC - B^2 = 0 \equiv Q(h, k)$ semidefinit $\equiv$ ingen slutsats kan dras om punkten (a, b) :s karaktär. När man får detta resultatet kan det vara bra att kolla på resttermen eller taylorutveckla till högre grad.

4 Vektorvärda funktioner

Genom användning av linjär algebra kan många egenskaper som gällde för skalärvärda funktioner:

$$\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n \quad (18)$$

där $n = 1$, generaliseras till funktioner där $n > 1$. Dessa kallas *vektorvärda funktioner*, och är funktioner vars värdemängd är n -dimensionella vektorer.

För att kunna göra generaliseringen är begreppet differentierbarhet centralt, då det ur differentierbarheten hos en funktion följer t ex både kontinuitet och senare att kedjeregeln kan användas. Per definition innebär det att om en given funktion $\mathbf{F}$ är differentierbar är också dess komponenter $(F_1, F_2, \dots, F_n)$ det.

4.1 Funktionalmatris

Begreppet funktionalmatris introduceras som namnet på en matris som innehåller en funktions alla partiella derivator, och betecknas med ett D . Funktionalmatrisen har det allmänna utseendet

$$D\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix} \quad (19)$$

där raderna i matrisen är gradienterna till funktionen och kolonnerna innehåller de partiella derivatorna med avseende på en viss variabel. Funktionalmatrisen är användbar då den innehåller mycket information om den givna funktionens beteende.

4.2 Kedjeregeln

I envariabelanalysen såg kedjeregeln ut på följande sätt

$$f(g(x)) = f'(g(x)) \cdot g'(x) \quad (20)$$

och har ett liknande utseende med den mer allmänna formen som används inom flervariabelanalysen, som istället har utseendet

$$D(\mathbf{F} \circ \mathbf{G})(\mathbf{t}) = D\mathbf{F}(\mathbf{G}(\mathbf{t})) \cdot D\mathbf{G}(\mathbf{t}) \quad (21)$$

där $\mathbf{F} \circ \mathbf{G}$ betecknar en sammansättning av de vektorvärda funktionerna $\mathbf{G}$ och $\mathbf{F}$. D betecknar funktionalmatrisen för dessa.

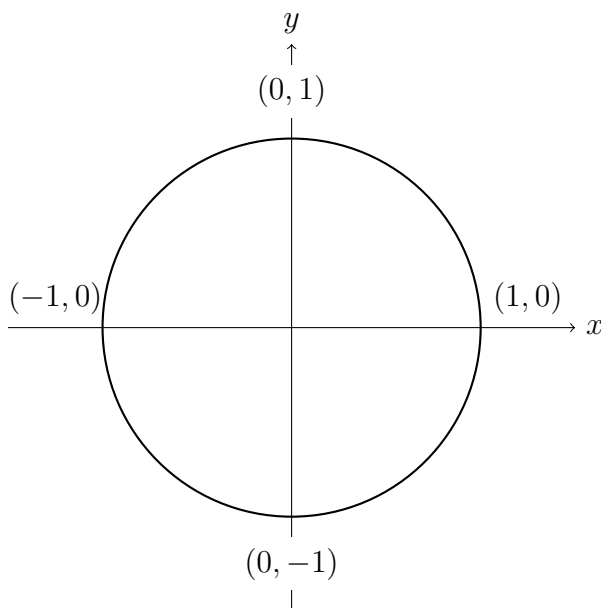
5 Implicita funktioner

Explicita funktioner är funktioner vars funktionsvärde bestäms av ett uttryck med endast invariabler, dvs de beroende och oberoende variablerna är separerade av likhetstecknet. Exempel på explicita funktioner är funktioner på formen $y = f(x)$. Explicita funktioners värde är alltid entydigt bestämt.

Implicita funktioner är funktioner vars funktionsvärde istället bestäms av en relation mellan funktionsvärdet och funktionens invariabler. En nivåkurva till en funktion av två variabler,

$$f: r(x, y) = c \quad (22)$$

där c är en konstant, är ett exempel på en implicit funktion i två variabler. Vissa implicita funktioner kan utan problem skrivas om som explicita funktioner, exempelvis $y - x = 0 \Rightarrow y = x$.



Ett mer problematiskt exempel är enhetscirkeln, uttryckt som en implicit funktion $r(x, y) = x^2 + y^2 = 1$. Om man försöker skriva om det som en explicit funktion

$$x^2 + y^2 = 1 \Rightarrow y = \pm\sqrt{1 - x^2} \quad (23)$$

ser man att funktionsvärdet ej är entydigt definierat. Om begränsningar sätts på funktionsmängden kan ekvationen dock uttryckas som en explicit funktion lokalt, dvs i en omgivning av en punkt $(a, b) \in x^2 + y^2 = 1$. Funktionen kan uttryckas explicit som $y = \sqrt{1 - x^2}$ i en omgivning av alla punkter (a, b) : $b > 0$ och som $y = -\sqrt{1 - x^2}$ i alla punkter där (a, b) : $b < 0$. De enda problematiska punkterna är de där $b = 0$, dvs punkterna $(\pm 1, 0)$, ty det finns ingen omgivning av dessa punkter där funktionsvärdet är entydigt bestämt. Notera att detta är eftersom funktionens tangent är vertikal i de två punkterna.

I fall med två variabler ger den *implicita funktionssatsen* verktyg för att undersöka om en implicit funktion $f: r(x, y) = c$ kan skrivas som en explicit funktion i en omgivning av en punkt $P = \{(a, b) : r(a, b) = c\}$. Om den sk *implicita derivatan* med avseende på b , $f'_b \neq 0$ kan funktionen skrivas explicit i en omgivning av P , givet att f är deriverbar med avseende på b . Den implicita funktionssatsen kan generaliseras ytterligare till intervall av invariabler och till godtyckligt antal variabler. I exemplet med enhetscirkeln är $f'_b = 2y$ för alla punkter på kurvan och följaktligen är $f'_a = 0$ för $y = 0$. Implicita funktionssatsen bekräftar alltså att enhetscirkeln ej kan skrivas som en explicit funktion i punkterna $(\pm 1, 0)$.

6 Funktionaldeterminant

Ett specialfall för funktionalmatriser är då $n = m$. Vi definierar ett nytt begrepp **funktionaldeterminanten**. Låt $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ vara ett differentierbart vektorfält. Då är funktionalmatrisen:

$$\det(D\mathbf{F}(\mathbf{a})) = d\mathbf{F}(\mathbf{a}). \quad (24)$$

Begreppet är väldefinierat eftersom funktionalmatrisen är en $n \times n$ matris och i punkten $\mathbf{a}$ antar varje index i matrisen ett tal vilket gör att determinanten kan beräknas.

För funktionaldeterminanten för en sammansatt funktion där

$$\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^p, \quad \mathbf{F} : \mathbb{R}^q \rightarrow \mathbb{R}^m, \quad p = q \quad \text{och} \quad \mathbf{t} \in \mathbb{R}^n$$

gäller:

$$\begin{array}{ccccccc} d\mathbf{F}(\mathbf{G}(\mathbf{t})) = & \det(D\mathbf{F}(\mathbf{G}(\mathbf{t}))) = & \det(D\mathbf{F}(\mathbf{G}(\mathbf{t})) \cdot D\mathbf{G}(\mathbf{t})) = d\mathbf{F}(\mathbf{G}(\mathbf{t})) \cdot d\mathbf{G}(\mathbf{t}) & (25) \\ \text{def} & \text{kedje} & \text{determinant linjäritet} & \end{array}$$

Anta att A är en kvadratisk matris. Då finns en sats från linjär algebra med tre ekvivalenta påståenden som säger:

- i) $\det(A) \neq 0$
- ii) Matrisen A är inverterbar
- iii) Avbildningen $\mathbf{x} \mapsto A\mathbf{x}$ är bijektiv

Detta översätter tyvärr inte helt till flervariabelanalysen. Även om $d\mathbf{F}(\mathbf{a})$ är en inverterbar matris i varje punkt $\mathbf{a}$ är det inte garanterat att avbildningen $\mathbf{x} \mapsto A\mathbf{x}$ är bijektiv i hela $\mathbb{R}^n$. Däremot kan man med hjälp av inversa funktionssatsen visa att om $\det(D\mathbf{F}(\mathbf{a})) \neq 0$ är avbildningen $\mathbf{x} \mapsto A\mathbf{x}$ bijektiv i en omgivning av punkten $\mathbf{a}$.

7 Koordinatbyte

Vi vill kunna byta mellan koordinater, till exempel $(x_1, \dots, x_n) \rightarrow (y_1, \dots, y_n)$ där varje y_i beror av de gamla variablerna

$$\begin{aligned} y_1 &= y_1(x_1, \dots, x_n) \\ &\vdots \\ y_n &= y_n(x_1, \dots, x_n). \end{aligned}$$

Hur vet man att det är ett giltigt byte? För ett giltigt byte gäller att avbildningen $\mathbf{x} \rightarrow \mathbf{y}$ är injektiv (där $\mathbf{x} = (x_1, \dots, x_n)$ och $\mathbf{y} = (y_1, \dots, y_n)$). Från inversa funktionssatsen fås att

avbildningen är bijektiv i en omgivning av punkten $\mathbf{a}$ om $d\mathbf{F}(\mathbf{a}) \neq 0$. Lite notation ger nu att

$$d\mathbf{F} = \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}, \quad (26)$$

där $\frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}$ ger en $n \times n$ matris. Med denna notation kan man även skriva kedjeregeln som:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{t}} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{t}}, \quad (27)$$

där det är vanlig matrismultiplikation mellan termerna i högerled.