

Tavelpresentation

Anton Carlsson, Anna Davoodi, Vilgot Jansson,
Maja Rhodin, Hugo Rådegård och Hugo Spencer

January 2021

Innehåll

1	Clairauts Sats	3
2	Taylors formel i flera variabler	3
3	Klassificering av kritiska punkter	4
4	Vektorvärda funktioner	5
4.1	Linjärisering av vektorvärda funktioner	5
4.2	Kedjeregeln	6
5	Vektorfält	7
5.1	Funktionaldeterminanter	7
5.2	Inversa funktionssatsen	8

1 Clairauts Sats

Definition: Om $D \subseteq \mathbb{R}^n$ är en domän och $k \geq 0$ är ett heltal så defineras $\mathcal{C}^k(D) = \{f : D \rightarrow \mathbb{R} \mid \text{både } f \text{ och samtliga av } f\text{'s partiella derivator av grad } \leq k \text{ är kontinuerliga i } D\}$.

Sats: (Specialfall av Clairauts Sats)

Låt $D \subseteq \mathbb{R}^2$ vara en domän och $f : D \rightarrow \mathbb{R}$. Om en funktion $f \in \mathcal{C}^2(D)$, så gäller $\forall (a, b) \in D$ att $f_{xy}(a, b) = f_{yx}(a, b)$.

Anmärkning: Det generella fallet av Clairauts sats säger att om $f \in \mathcal{C}^k(D)$, $D \subseteq \mathbb{R}^n$ så spelar ordningen för f 's partiella derivator ingen roll upp till grad k . Detta kan visas genom ett induktionsargument från specialfallet.

Notation: En mer effektiv notation är:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = f_{xy}.$$

2 Taylors formel i flera variabler

För att utforma Taylors formel för ett godtyckligt antal variabler påminner vi oss först om satsen för en funktion av en variabel.

Låt $f : \mathbb{R} \rightarrow \mathbb{R}$ vara en funktion av klassen $\mathcal{C}^{k+1}$, låt $D \subseteq \mathbb{R}$ vara en domän och $a \in D$. Då är

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

för något $\xi \in (a, a+h)$.

Eftersom $f^{(k+1)}(x)$ är kontinuerlig i ett slutet intervall kan den sista termen skrivas $\mathcal{O}(h^{k+1})$. Summan i ekvationen är ett polynom av grad k som kallas Taylorpolynomet. Detta ger följande som den så kallade Taylorapproximationen:

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

Vi vill nu utvidga denna formeln till funktioner av ett godtyckligt antal variabler. Detta görs smidigast genom att studera Taylorutvecklingen för två variabler. Vi låter $D \subseteq \mathbb{R}^2$ vara en domän och $f \in \mathcal{C}^{k+1}(D)$. Låt $(a, b) \in D$ och sätt $F(t) = f(a+th, b+tk)$. F betraktas här som en sammansättning $f(g_1(t), g_2(t))$ där g_1 och g_2 är två linjära funktioner. Med detta går att utnyttja kedjeregeln för att få fram hur upprepade derivering av F påverkar f genom en så kallad *differentialoperator*, D som betecknas som följande

$$\begin{cases} F' = Df \\ F'' = D^2f \\ \vdots \\ F^{(j)} = D^j f \end{cases}$$

där

$$D = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right) = \mathbf{h} \cdot \nabla.$$

Genom att nu sätta $t = 1$ i funktionen F fås att $F(1) = f(g_1(1), g_2(1)) = f(a + h, b + k)$. För ett godtyckligt antal variabler låter vi nu $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{h} = (h_1, \dots, h_n)$ samt $\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)$. Vi kan nu uttrycka den allmänna Taylorutvecklingen för en funktion av ett godtyckligt antal variabler som

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

De två mest centrala Taylorutvecklingarna för denna kursen är de av grad $k = 1, 2$ i två variabler.

För $k = 1$, $f \in \mathcal{C}^2$ fås

$$f(a + h, b + k) = f(a, b) + hf_x(a, b) + kf_y(a, b) + \mathcal{O}(\|\mathbf{h}\|^2),$$

samt för $k = 2$, $f \in \mathcal{C}^3$

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

3 Klassificering av kritiska punkter

Om $f : \mathbb{R}^n \rightarrow \mathbb{R}$ är differentierbar i $\mathbf{x} = \mathbf{a}$ sägs den ha en kritisk/stationär punkt om $\nabla f(\mathbf{a}) = \mathbf{0}$. Att hitta en funktions kritiska punkter kan vara användbart för att hitta de lokala extrempunkterna då de alltid är kritiska punkter. Ett sätt att avgöra typen av en kritisk punkt till en differentierbar funktion f av 2 variabler i punkten (a, b) är att studera funktionens Taylorapproximation av grad 2 där $f_x = f_y = 0$, som då blir:

$$\begin{aligned} f(a + h, b + k) - f(a, b) &= Q(h, k) + \mathcal{O}(\|\mathbf{h}\|^3) = \\ &= \frac{1}{2}(h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b)) + \mathcal{O}(\|\mathbf{h}\|^3). \end{aligned}$$

Analogt med i envariabelsanalysen fås en lokal maximipunkt (respektive minimipunkt) där den kvadratiske termen $Q(h, k)$ är mindre (respektive större) än 0 i en omgivning till punkten (a, b) . För att underlätta notationen sätts nu $f_{xx}(a, b) = A, f_{xy}(a, b) = B, f_{yy}(a, b) = C$. För att underlätta att avgöra om $Q(h, k)$ är större än 0 kan nu detta uttryck kvadratkompletteras, vilket förutsatt att $A \neq 0$ ger:

$$Q(h, k) = Ah^2 + 2Bhk + Ck^2 = A \left(\left(h + \frac{B}{A}k \right)^2 + (AC - B^2) \left(\frac{k}{A} \right)^2 \right).$$

Informationen som kan utläsas ur uttrycket samt definitioner för begrepp till funktionen $Q(h, k)$ finns i tabellen nedan:

$AC - B^2$	A	$Q(h, k)$	(a, b)
> 0	> 0	Positivt definit	Lokalt minimum
> 0	< 0	Negativt definit	Lokalt maximum
< 0	-	Indefinit	Sadelpunkt
$= 0$	-	Semidefinit	Oklart

4 Vektorvärda funktioner

Definition : En *vektorvärd* funktion är en funktion $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ där $n, m \in \mathbb{N}$. Vi skriver:

$$\mathbf{F}(x_1, \dots, x_n) = (F_1(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n)).$$

Definition : En funktion $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sägs vara *differentierbar* i en punkt $\mathbf{a} \in \mathbb{R}^n$ om varje komponent F_i ($i = 1, \dots, m$) är differentierbar. Då alla komponenter F_i är skalärvärda funktioner följer differentierbarhet för vektorvärda funktioner direkt från definitionen av differentierbarhet för skalärvärda funktioner.

4.1 Linjärisering av vektorvärda funktioner

Härledningen för linjärisering av vektorvärda funktioner följer direkt från linjärisering av skalärvärda funktioner. Då en godtycklig komponent F_i är differentierbar kan vi beskriva linjäriseringen för F_i som:

$$F_i(\mathbf{a} + \mathbf{h}) = F_i(\mathbf{a}) + \nabla F_i(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\| \rho_i(\mathbf{h})$$

där varje $\rho_i(\mathbf{h}) \rightarrow 0$, då $\mathbf{h} \rightarrow \mathbf{0}$. Vi noterar att $\mathbf{F}$ kan skrivas om som en kolonnvektor där varje komponent F_i ($i = 1, \dots, m$) kan staplas på följande sätt,

$$\mathbf{F} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}$$

Med hjälp av den notationen formuleras linjärisering för en vektorvärd funktion som:

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \begin{bmatrix} F_1(\mathbf{a}) \\ F_2(\mathbf{a}) \\ \vdots \\ F_m(\mathbf{a}) \end{bmatrix} + \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial F_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial F_m}{\partial x_n}(\mathbf{a}) \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} + \|\mathbf{h}\| \begin{bmatrix} \rho_1(\mathbf{h}) \\ \vdots \\ \rho_m(\mathbf{h}) \end{bmatrix}$$

där varje element $\rho_i \rightarrow 0$ då $\mathbf{h} \rightarrow \mathbf{0}$. Gradientens motsvarighet i vektorvärda funktioner blir en $m \times n$ -matris som kallas funktionalmatris.

Obs 1: Skalärmultiplikationen mellan gradienten och $\mathbf{h}$ ersätts med matrismultiplikation i formuleringen ovan.

Obs 2: Varje rad $i = 1, \dots, m$ i funktionalmatrisen består av gradienten för funktionen F_i .

Definition: Funktionalmatrisen till $\mathbf{F}$ i punkten $\mathbf{a}$ betecknas $D\mathbf{F}(\mathbf{a})$ och har på positionen (i, j) elementet $\frac{\partial F_i}{\partial x_j}(\mathbf{a})$.

Slutligen formulerar vi linjäriseringen med en mer effektiv notation:

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \mathbf{F}(\mathbf{a}) + D\mathbf{F}(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h})$$

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

4.2 Kedjeregeln

Även kedjeregeln kan generaliseras för vektorvärda funktioner. För att visa detta betraktas två godtyckliga vektorvärda funktioner $\mathbf{G}(\mathbf{t}) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ och $\mathbf{F}(\mathbf{x}) : \mathbb{R}^p \rightarrow \mathbb{R}^m$ och deras sammansättning $(\mathbf{F} \circ \mathbf{G})(\mathbf{t}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$. För att derivatan som söks med kedjeregeln ska vara definierad måste både $\mathbf{F}$ och $\mathbf{G}$ vara differentierbara. Målet med kedjeregeln är att relatera funktionalmatrisen för den sammansatta funktionen till de enskilda funktionernas funktionalmatriser. För att få fram ett sådant uttryck kan man betrakta, $\frac{\partial}{\partial t_j}(\mathbf{F} \circ \mathbf{G})_i(\mathbf{t})$, vilket kan med kedjeregeln för skalärvärda funktioner uttryckas som följer:

$$\frac{\partial}{\partial t_j}(\mathbf{F} \circ \mathbf{G})_i(\mathbf{t}) = \sum_{k=1}^p \frac{\partial F_i}{\partial x_k}(\mathbf{G}(\mathbf{t})) \frac{\partial G_k}{\partial t_j}(\mathbf{t}).$$

Summan i högerledet kan även betraktas som följande matrisprodukt där den vänstra matrisen har samma element som $\nabla F_i(\mathbf{G}(\mathbf{t}))$ och den högra har samma element som $\frac{\partial}{\partial t_j} \mathbf{G}(\mathbf{t})$:

$$\frac{\partial}{\partial t_j} (\mathbf{F} \circ \mathbf{G})_i(\mathbf{t}) = \begin{bmatrix} \frac{\partial F_i}{\partial x_1}(\mathbf{G}(\mathbf{t})) & \dots & \frac{\partial F_i}{\partial x_p}(\mathbf{G}(\mathbf{t})) \end{bmatrix} \begin{bmatrix} \frac{\partial G_1}{\partial t_j}(\mathbf{t}) \\ \vdots \\ \frac{\partial G_p}{\partial t_j}(\mathbf{t}) \end{bmatrix}.$$

Detta är per definition lika med elementet i i :te raden och j :te kolumnen av funktionalmatrisen $D(\mathbf{F} \circ \mathbf{G})(\mathbf{t})$. Från detta samband kan hela funktionalmatrisen av $(\mathbf{F} \circ \mathbf{G})(\mathbf{t})$ uttryckas som:

$$D(\mathbf{F} \circ \mathbf{G})(\mathbf{t}) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{G}(\mathbf{t})) & \dots & \frac{\partial F_1}{\partial x_p}(\mathbf{G}(\mathbf{t})) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{G}(\mathbf{t})) & \dots & \frac{\partial F_m}{\partial x_p}(\mathbf{G}(\mathbf{t})) \end{bmatrix} \begin{bmatrix} \frac{\partial G_1}{\partial t_1}(\mathbf{t}) & \dots & \frac{\partial G_1}{\partial t_n}(\mathbf{t}) \\ \vdots & \ddots & \vdots \\ \frac{\partial G_p}{\partial t_1}(\mathbf{t}) & \dots & \frac{\partial G_p}{\partial t_n}(\mathbf{t}) \end{bmatrix}.$$

Genom att notera att den vänstra matrisen är funktionalmatrisen av $\mathbf{F}$ i punkten $\mathbf{G}(\mathbf{t})$ och den högra matrisen funktionalmatrisen av $\mathbf{G}$ i punkten $\mathbf{t}$ kan detta skrivas om som:

$$D(\mathbf{F} \circ \mathbf{G})(\mathbf{t}) = D\mathbf{F}(\mathbf{G}(\mathbf{t}))D\mathbf{G}(\mathbf{t}).$$

Detta är den mest generella versionen av kedjeregeln.

5 Vektorfält

Ett n -dimensionellt vektorfält definieras som en vektorvärd funktion med definitionsmängd i $\mathbb{R}^n$ och sina värden i $\mathbb{R}^n$ (observera att skillnaden från generella vektorvärda funktioner är att $n = m$ i vektorfält).

5.1 Funktionaldeterminanter

Determinanten av funktionalmatrisen av ett differentierbart vektorfält $\mathbf{F}(\mathbf{x})$ kallas för *funktionaldeterminanten* av $\mathbf{F}$ och betecknas med $\frac{d(\mathbf{F})}{d(\mathbf{x})}$

$$\frac{d(\mathbf{F})}{d(\mathbf{x})} = \det \left(\frac{\partial(\mathbf{F})}{\partial(\mathbf{x})} \right) = \begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{vmatrix}$$

Funktionaldeterminanten av en sammansättning av vektorfält i samma dimension är lika med produkten av funktionaldeterminanterna av de ingående funktionerna innan sammansättningen. Detta innebär att för två n -dimensionella vektorfält

$\mathbf{F}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ och $\mathbf{x}(\mathbf{t}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (obs. samma n) gäller

$$\frac{d(\mathbf{F})}{d(\mathbf{t})} = \frac{d(\mathbf{F})}{d(\mathbf{x})} \cdot \frac{d(\mathbf{x})}{d(\mathbf{t})}.$$

Observera att denna formel ser exakt ut som kedjeregeln. Detta beror på att det i princip är kedjeregeln fast skrivet med notationen av funktionaldeterminanter.

5.2 Inversa funktionssatsen

Ibland kan det vara av intresse att bedöma om ett vektorfält är bijektivt, speciellt då man letar efter ett koordinatbyte eftersom ett vektorfält måste vara bijektivt för att vara ett giltigt koordinatbyte. För att få en idé om huruvida ett vektorfält är bijektivt kan man skapa en linjär approximation i en punkt där funktionen är differentierbar. Om funktionaldeterminanten är nollskild kan funktionen approximeras med en bijektiv avbildning, vilket tyder på att funktionen själv bör vara bijektiv nära denna punkt. Slutsatsen av denna tanke stöds av och formuleras i inversa funktionssatsen.

Sats: (Inversa funktionssatsen)

Om $\mathbf{F}$ är ett $\mathcal{C}^1$ -vektorfält och $\mathbf{a}$ är en punkt i $\mathbf{F}$:s definitionsmängd sådan att

$$\det(D\mathbf{F}(\mathbf{a})) \neq 0$$

så finns det öppna omgivningar U och V av $\mathbf{a}$ resp. $\mathbf{F}(\mathbf{a})$ så att $\mathbf{F} : U \rightarrow V$ är bijektiv och dess invers $\mathbf{F}^{-1} : V \rightarrow U$ tillhör klassen $\mathcal{C}^1$.