Tavelpresentation 2C

Hugo Lom Thea Mattsson Arvid Andersson
Linus Sundström Mats Richardson Lina Wincent

Februari 2019

1 Föreläsning 1

1.1 Taylors Formel i flervariebelanalys

För att använda sig av Taylors formel (även kallad Taylorutveckling) i flervariabelsanalys utnyttjas kombination av Taylors formel i en variabel samt kedjeregeln.

Taylors formel i en variabel fås av följande **Sats**:
Låt $k \in \mathbb{N}$, $a \in \mathbb{R}$ samt $f : \mathbb{R} \to \mathbb{R}$, $f \in C^{k+1}(\mathbb{D})$ där $\mathbb{D}$ är en omgivning till a. Då gäller att $\forall x \in \mathbb{D}$:

$$f(x) = \overbrace{\left[\sum_{j=0}^{k} \frac{f^j(a)(x-a)^j}{j!}\right]}^{\text{Taylorpolynomet i } x=a} + \frac{f^{k+1}(\xi)(x-a)^{k+1}}{(k+1!)}, \tag{1}$$

för något $\xi \in (x, a)$, där $f^0(a) = f(a)$. Funktionen f måste vara $k+1$ gånger differentierbar och dess $k+1$:a derivata ä kontinuerlig i definitionsmängden. Ett notationsbyte med variabelbytet $x = a + h$ ger ett mer användbart uttryck,

$$f(a+h) = \left[\sum_{j=0}^{k} \frac{f^j(a)\,h^j}{j!}\right] + \frac{f^{(k+1)}(\xi)\,h^{k+1}}{(k+1)!}, \tag{2}$$

vilket sen kan förenklas till,

$$f(a+h) = \left[\sum_{j=0}^{k} \frac{f^j(a)\,h^j}{j!}\right] + \mathcal{O}\left(h^{k+1}\right), \tag{3}$$

där $\mathcal{O}\left(h^{k+1}\right)$ är en restterm. k kallas Taylorutvecklingens ordning.
Vid $k = 1$, blir utvecklingen för en funktion $f \in C^2(\mathbb{D})$:

$$f(a+h) = f(a) + f'(a)\,h + \mathcal{O}\left(h^2\right), \tag{4}$$

vilket kan ses som en förfining av linjäriseringsformeln där $\mathcal{O}\left(h^2\right) = h\rho(h)$. Tayloruttvecklingar av högre ordning ger en mer "förfinad" felterm, vilket ger en bättre approximation.

För att utöka Taylors formel till funktioner i flera variabler krävs ett variabelbyte för att kunna använda sig av envariabels-formeln.

Låt $f : \mathbb{R}^n \to \mathbb{R}$, och sätt $F(t) = f(\mathbf{a} + t\mathbf{h})$ där $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ och $\mathbf{h} = (h_1, h_2, \cdots, h_n)$ s.a $\mathbf{a}$ är en punkt i, och har som omgivning, $\mathbb{D} \subseteq \mathbb{R}^n$. Enligt Taylors formel för en variabel (4) får vi:

$$F(0+t) = f(\mathbf{a} + t\mathbf{h}) = \sum_{j=0}^{k} \frac{F(0)^{(j)}\,t^j}{j!} + \mathcal{O}\left(t^{k+1}\right), \tag{5}$$

där $F^{(j)}$ är funktionen Fs j:te derivata. Från kedjeregeln får vi $F' = Df$ där $D = \left(h_1 \frac{\partial}{\partial x_1}, \cdots, h_k \frac{\partial}{\partial x_k}\right) \left(= \mathbf{h} \bullet \triangledown, \triangledown = \left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\right)\right)$. Induktivt fås då $F^{(j)} = D^j f$. Vi ersätter $F^{(j)}(0)$ med $D^j f(\mathbf{a})$ vilket för $t = 1$ ger uttrycket som formar Taylors Sats

$$f \in C^{k+1}(\mathbb{D}), \ \mathbf{a} \in \mathbb{D}$$

$$\Rightarrow f(\mathbf{a} + \mathbf{h}) = \sum_{j=0}^{k} \frac{(\mathbf{h} \bullet \triangledown)^j f(\mathbf{a})}{j!} + \mathcal{O}\left(\|\mathbf{h}\|^{k+1}\right), \tag{6}$$

,där en viktig utveckling är av ordning 2 för $f : \mathbb{R}^2 \to \mathbb{R}$ då man får Taylorutvecklingen;

$$\begin{aligned} f(a+h, b+k) = & f(a,b) + hf_x(a,b) + kf_y(a,b) \\ & + \frac{1}{2}\left[h^2 f_{xx}(a,b) + 2hkf_{xy} + k^2 f_{yy}(a,b)\right] \\ & + \mathcal{O}\left(\|\mathbf{h}\|^3\right), \end{aligned} \tag{7}$$

där $\mathbf{h} = (h, k)$, vilken används för klassificering av kritiska punkter.

2 Föreläsning 2

2.1 Klassificering av kritiska punkter

I en variabel defineras en kritisk punkt som en punkt på en funktionskurva då derivatan av funktionen i den punkten är lika med noll. Klassificeringen (lokalt max eller min) ges av andraderivatans tecken. Vi ser att för klassificeringen kan man använda sig av en andra grads Taylorutveckling för att klassificera den kritiska punkten som antingen max eller min genom att kolla på lutningen till en närliggande punkt $f(a+h) - f(a) = f'(a)h + \frac{f''(a)}{2}h^2 + \mathcal{O}\left(h^3\right)$ där a är en kritisk punkt s.a $f'(a) = 0$. Detta ger att differensen $f(a+h) - f(a)$ är lika med $\frac{f''(a)}{2}h^2 + \mathcal{O}\left(h^3\right)$. Då ser vi att för tillräckligt små h:

1. $\frac{f''(a)}{2}h^2 > 0$ blir differensen strängt positiv $\Rightarrow x = a$är ett strängt lokalt minima.
2. $\frac{f''(a)}{2}h^2 < 0$ blir differensen strängt negativ $\Rightarrow x = a$ är ett strängt lokalt maxima.

På liknande sätt vill man klassificera kritiska punkter i flera dimensioner.
Låt $f : \mathbb{R}^n \to \mathbb{R}, \ \mathbf{a} \in \mathbb{R}^n$ och $\mathbb{D}$ vara en öppen omgivning till a. Om f har en lokal

extrempunkt i $\mathbf{x} = \mathbf{a}$ och är differentierbar i $\mathbf{x} = \mathbf{a}$, då gäller att $\nabla f(\mathbf{a}) = \mathbf{0}$. Givet detta kan vi nu se på Taylorutvecklingen för $f : \mathbb{R}^2 \to \mathbb{R}$ av ordning 2;

$$f(a+h, b+k) - f(a,b) = \frac{1}{2}\left[h^2 f_{xx}(a,b) + 2hk f_{xy} + k^2 f_{yy}(a,b)\right] \tag{8}$$
$$+ \mathcal{O}\left(\|\mathbf{h}\|^3\right), \tag{9}$$

där $\frac{1}{2}\left[h^2 f_{xx}(a,b) + 2hk f_{xy} + k^2 f_{yy}(a,b)\right]$ sägs vara på binär kvadratisk form, alltså att $Q(h,k) = Ah^2 + 2Bhk + Ck^2$ där

$$A = \frac{1}{2} f_{xx}(a,b)$$
$$B = \frac{1}{2} f_{xy}(a,b)$$
$$C = \frac{1}{2} f_{yy}(a,b)$$

3 Föreläsning 3

3.1 Klassificering av kritiska punkter, fortsättning

Vi känner till att Q är den så kallade *binära kvadratiska formen*. Per definition sägs formen $Q(x,y)$ vara

- *positiv definit* om $Q(x,y) > 0$ då $(x,y) \neq (0,0)$.
- *negativ definit* om $Q(x,y) < 0$ då $(x,y) \neq (0,0)$.
- *indefinit* om $Q(x_1,y_1) > 0$ och $Q(x_2,y_2) < 0$ för några punkter (x_1,y_1) och (x_2,y_2).

Därtill kan noteras att $Q(x,y)$ är *positiv semidefinit* om $Q(x,y) \geq 0$ för alla $(x,y) \in \mathbb{R}^2$ och $Q(x_1,y_1) = 0$ för något $(x_1,y_1) \neq (0,0)$. I det fall att $Q(x,y) \leq 0$ $\forall (x,y) \in \mathbb{R}^2$ och $Q(x_1,y_1) \leq 0$ för något $(x_1,y_1) \neq (0,0)$ sägs $Q(x,y)$ istället vara *negativ semidefinit*.
En i sammanhanget relevant sats relaterar ovanstående definition till klassificeringen av kritiska punkter.

SATS 2
Om (a,b) är en kritisk, inre punkt i defintionsmängden $\mathbb{D}$ tillhörande funktionen $f(x,y)$ och Q betecknar den kvadratiska formen i Taylorutvecklingen kring (a,b) så gäller:
i. Om $Q(h,k)$ är positiv definit är (a,b) ett strängt lokalt minimum.
ii. Om $Q(h,k)$ är negativ definit är (a,b) ett strängt lokalt maximum.
iii. Om $Q(h,k)$ är indefinit är (a,b) en s.k. sadelpunkt, dvs. varken lokalt maximum eller minimum.[1]

[1] Detta kan jämföras med en s.k. terasspunkt i envariabelfallet.

BEVIS (i)
Beviset kräver följande lemma:

Lemma: Då $Q(x, y)$ är en positiv definit binär kvadratisk form finns ett $\delta > 0$ s.a.

$$\forall(x, y) \in \mathbb{R}^2 : Q(x, y) \geq \delta \cdot (x^2 + y^2) \tag{10}$$

BEVIS FÖR LEMMAT
Låt s^1 beteckna enhetscirkeln, vi har då: $s^1 = \{(x, y) \mid x^2 + y^2 = 1\}$. s^1 är en kompakt mängd och $Q(x, y)$ är kontinuerlig i hela $\mathbb{R}^2$, vilket medför att Q antar ett minsta värde på cirkeln. Då vi har att Q är positiv definit är detta minsta värde strängt positivt. Vi betecknar detta värde ε. För en godtycklig punkt $(x, y) \in \mathbb{R}^2$ finns ett tal $t \geq 0$ s.a. $(x, y) = (tx_1, ty_1)$ för en punkt $(x_1, y_1) \in s^1$. Den kvadratiska formen $Q \Rightarrow Q(tx_1, ty_2) = t^2 Q(x_1, y_1)$. Därigenom fås $Q(x, y) = t^2 Q(x_1, y_1) \geq t^2(\delta(x_1^2 + y_1^2)) = \delta[(tx_1)^2 + (ty_1)^2] = \delta(x^2 + y^2)$, vsv.

BEVIS FÖR SATSEN
Antag $Q(h, k)$ positiv definit. Enligt lemmat finns ett tal $\delta > 0$ sådant att $Q(h, k) \geq \delta(h^2 + k^2)\ \forall(h, k) \in \mathbb{R}^2$.

Detta kan vi formulera som att det gäller att $Q(h, k) \geq \delta \cdot \|\mathbf{h}\|^2$ för alla $\mathbf{h} \in \mathbb{R}^2$. Av lemma och Taylor får vi

$$f(a + h, b + k) - f(a, b) = Q(h, k) + O(\|\mathbf{h}\|^3) \geq \delta\|\mathbf{h}\|^2 + O(\|\mathbf{h}\|^3) \tag{11}$$

Detta observerar vi är strängt större än noll då $\|\mathbf{h}\|$ är tillräckligt liten. ■

En annan viktig sats för klassificeringen av kritiska punkter är följande

SATS 1
Låt $Q(x, y) = Ax^2 + 2Bxy + Cy^2$.
Då gäller
$AC - B^2 > 0$ och $A > 0 \Rightarrow Q$ positiv definit
$AC - B^2 > 0$ och $A < 0 \Rightarrow Q$ negativ definit
$AC - B^2 < 0 \Rightarrow Q$ indefinit
$AC - B^2 = 0 \Rightarrow Q$ semidefinit.
I det sista fallet, då $AC - B^2 = 0$, går det inte att enbart utifrån detta att utläsa vilken slags kritisk punkt (x, y) är.

3.2 Bevis av sats 1

Givet $Q(x, y) = Ax^2 + 2Bxy + Cy^2$, antag först $A \neq 0$. Då kan $Q(x, y)$ skrivas om till

$$Q(x, y) = A((x + \frac{B}{A}y)^2 + (AC - B^2)(\frac{y}{A})^2). \tag{12}$$

Då gäller att för $AC - B^2 > 0$, alltså när A avgör tecknet, är $Q(x, y)$ positiv definit för $A > 0$ samt negativ definit för $A < 0$. Det är definit ty uttrycket är nollskilt utanför origo. Det bevisar de första två påståendena, ty om $A = 0$ kan inte $AC - B^2 > 0$.
För $AC - B^2 < 0$ kan x och y alltid väljas så att parentesen blir antingen negativ eller positiv. Oavsett tecken på A kommer alltså $Q(x, y)$ vara indefinit. Det bevisar det tredje påståendet.
Om $A = 0$ kan ovanstående resonemang göras för C istället. Om $A = C = 0$ är det inte längre en kvadratisk form. Då är $Q(x, y) = 2Bxy$ och $AC - B^2 < 0 \Leftrightarrow B \neq 0$. Men då är det klart att $Q(x, y) = 2Bxy$ är indefinit, ty xy är positivt då x och y har samma tecken, och negativt då x och y har olika tecken.
För $AC - B^2 = 0$ blir den andra termen i parentesen 0 och $Q(x, y)$ kommer vara 0 längs linjen $x + \frac{B}{A}y = 0$, vilket visar det fjärde påståendet.

3.3 Sats 2, andra formulering

Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f \in \mathbb{C}^3$ och (a, b) vara en kritisk punkt till f.
Då följer att;
I) $f_{xx}f_{yy} - {f_{xy}}^2|_{(a,b)} > 0$ och $f_{xx}(a, b) > 0 \implies (a, b)$ är ett strängt lokalt minimum.
II) $f_{xx}f_{yy} - {f_{xy}}^2|_{(a,b)} > 0$ och $f_{xx}(a, b) < 0 \implies (a, b)$ är ett strängt lokalt maximum.
III) $f_{xx}f_{yy} - {f_{xy}}^2|_{(a,b)} < 0 \implies (a, b)$ är en sadelpunkt.

3.4 Vektorvärda funktioner

Definition: Låt $m, n \in \mathbb{N}$. En funktion $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ kallas en vektorvärd funktion, med specialfallet $m = 1$ som kallas en skalärvärd funktion.
Notation: $\mathbf{F}(x_1, ..., x_n) = (F_1(x_1, ..., x_n), ..., F_m(x_1, ..., x_n))$
Definition: $\mathbf{F}$ sägs vara differentierbar eller tillhöra C om varje F_i är eller gör det.

3.5 Linjärisering

Säg att varje F_i är differentierbar i en punkt $\mathbf{x} = \mathbf{a} \in \mathbb{R}^n$. Då gäller:

$$F_i(\mathbf{a} + \mathbf{h}) = F_i(\mathbf{a}) + \mathbf{h} \bullet \nabla F_i(\mathbf{a}) + |\mathbf{h}|\, \rho(\mathbf{h}) \tag{13}$$

där $\rho(|\mathbf{h}|) \to 0$ då $\mathbf{h} \to 0$. Skalärprodukten i ekvationen ersätts med hjälp av matrismultiplikation, och alla F_i för $i = 1, ..., m$ placeras i en kolumn ovanför varandra. Det ger uttrycket:

$$\mathbf{F}(\mathbf{a} + \mathbf{h}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a}) \cdot \mathbf{h} + |\mathbf{h}|\, \boldsymbol{\rho}(\mathbf{h}) \tag{14}$$

där $\boldsymbol{\rho}(\mathbf{h}) \to \mathbf{0}_m$ då $\mathbf{h} \to \mathbf{0}_n$.
Terminologi: Matrisen $\mathbf{F}'(\mathbf{a})$ kallas för funktionalmatrisen eller Jacobimatrisen

för den vektorvärda funktionen $\mathbf{F} \in C^1$, och har element (i, j)

$$\frac{\partial F_i}{\partial x_j}(\mathbf{a}) \tag{15}$$

4 Föreläsning 4

4.1 Kedjeregeln för vektorvärda funktioner

För två funktioner $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ och $\mathbf{G} : \mathbb{R}^p \to \mathbb{R}^q$ som antas vara C^1 är den sammansatta funktionen $\mathbf{F} \circ \mathbf{G} : \mathbb{R}^p \to \mathbb{R}^m$ definierad då $n = q$.
Påstående:

A. $\mathbf{F} \circ \mathbf{G}$ är också C^1

B. $(\mathbf{F} \circ \mathbf{G})' = \mathbf{F}'\mathbf{G}' \to (\mathbf{F} \circ \mathbf{G})'(\mathbf{a}) = \mathbf{F}'((\mathbf{G}(\mathbf{a})) * \mathbf{G}'(\mathbf{a})$

För en punkt $\mathbf{a} \in \mathbb{R}^p$ så kan värdet på $\mathbf{G}(\mathbf{a} + \mathbf{h})$ approximeras genom den allmänna linjariseringsformeln till $\mathbf{G}(\mathbf{a} + \mathbf{h}) \approx \mathbf{G}(\mathbf{a}) + \mathbf{G}'(\mathbf{a}) * \mathbf{h}$. På samma vis kan $\mathbf{F}$ linjäriseras kring $\mathbf{G}(\mathbf{a})$ enligt

$$\mathbf{F}(\mathbf{G}(\mathbf{a}) + \mathbf{K}) \approx \mathbf{F} \circ \mathbf{G} + \mathbf{F}'(\mathbf{G}(\mathbf{a})) * \mathbf{K}. \tag{16}$$

Följande förenklingar kan göras då $\mathbf{K} = \mathbf{G}(\mathbf{a}) * \mathbf{h}$,

$$\mathbf{F}(\mathbf{G}(\mathbf{a}) + \mathbf{K}) \approx \mathbf{F}(\mathbf{G}(\mathbf{a} + \mathbf{h})) = (\mathbf{F} \circ \mathbf{G})(\mathbf{a} + \mathbf{h}) \tag{17}$$

$$\mathbf{F}(\mathbf{G}(\mathbf{a})) = (\mathbf{F} \circ \mathbf{G})(\mathbf{a}) \tag{18}$$

vilket leder till att den sammansatta funktionen kan approximeras till

$$(\mathbf{F}\circ\mathbf{G})(\mathbf{a}+\mathbf{h}) \approx (\mathbf{F}\circ\mathbf{G})(\mathbf{a})+[(\mathbf{F}'(\mathbf{G}(\mathbf{a}))+\mathbf{G}(\mathbf{a})]*\mathbf{h} \approx (\mathbf{F}\circ\mathbf{G})(\mathbf{a})+(\mathbf{F}\circ\mathbf{G})'(\mathbf{a})*\mathbf{h}. \tag{19}$$

Detta visar en linjariseringsformel för $(\mathbf{F} \circ \mathbf{G})'(\mathbf{a})$ utan restermen vilket visar att den sammansatta funktionen också är C^1 vilket visar A i påståendet. Del B ges av härledningen att $(\mathbf{F} \circ \mathbf{G})'(\mathbf{a}) = \mathbf{F}'((\mathbf{G}(\mathbf{a})) * \mathbf{G}'(\mathbf{a})$.

4.2 Specialfallet då $m = n$, variabelbyten

Då $m = n$ kallas $\mathbf{F}$ för ett n-dimensionellt vektorfält och definieras som $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$. Detta $\mathbf{F}$ kan betraktas som ett variabelbyte. $\mathbf{F}$ definieras i detta fall som $\mathbf{F}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ där $y_1, \ldots, y_n$ är funktioner av $\mathbf{x}$. $\mathbf{F}'$ betecknas enligt

$$\mathbf{F}' = \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}. \tag{20}$$

$\mathbf{F}$ måste vara bijektiv för att variabelbytet skall vara giltigt. Ett exempel är bytet mellan Cartesiska och polära koordinater där man går från inputvariablerna (x, y) till (r, θ). $\mathbf{F}(x, y) = (\sqrt{x^2 + y^2}, arctan(y/x)) = (r, \theta)$ är funktionen för att göra ett variabelbyte mellan Cartesiska och polära koordinater. $\mathbf{F}$ är bijektiv och definierad då $\mathbf{F} : \mathbb{R}^2 \backslash \{(0,0)\} \to (0,\infty) \times [0, 2\pi)$. Detta ger inversen $\mathbf{F}^{-1}(r,\theta) = (r\cos(\theta), r\sin(\theta)) = (x,y)$. Variabelbytets funktionalmatris kan beräknas enligt följande:

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \tag{21}$$

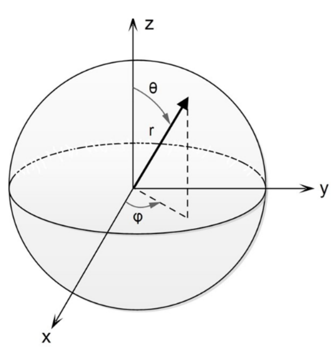
$$\frac{\partial(r,\theta)}{\partial(x,y)} = \left[\frac{\partial(x,y)}{\partial(r,\theta)}\right]^{-1} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta)/r & \cos(\theta)/r \end{pmatrix} = \begin{pmatrix} x/r & y/r \\ -y/r^2 & y/r^2 \end{pmatrix}. \tag{22}$$

4.3 Inversa funktionssatsen

SATS
Låt $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ vara en C^1-funktion från $\mathbb{R}^n$ till $\mathbb{R}^n$ (alltså ett variabelbyte), och $\boldsymbol{a}$ en punkt i definitionsmängden s.a. det $\boldsymbol{f}'(\boldsymbol{a}) \neq 0$. Då finns det öppna omgivningar $\mathbb{U}$ och $\mathbb{V}$ av $\boldsymbol{a}$ respektive $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$ sådana att avbildningen
$\boldsymbol{f} : \mathbb{U} \to \mathbb{V}$ är bijektiv och inversen $\boldsymbol{f}^{-1} : \mathbb{V} \to \mathbb{U}$ en funktion av klassen C^1.

Satsens bevis utelämnas här, men kontentan är att ett varabelbyte $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ är giltigt i en omgivning av punkten $\boldsymbol{a}$ om determinanten av dess Jacobi-matris är nollskild i $\boldsymbol{a}$.
Ett exempel på detta är variabelbytet från Cartesiska till sfäriska koordinater;

$$(x,y,z) \leftrightarrow (\rho,\theta,\phi)$$



$$\theta \in [0,\pi]$$
$$\phi \in [0,2\pi)$$
$$\rho \in [o,\infty)$$

$$x = \rho\sin\theta\cos\phi$$
$$y = \rho\sin\theta\sin\phi$$
$$z = \rho\cos\theta$$

där vi m.h.a. inversa funktionssatsen och trigonometriska ettan kan visa att

$$det\left(\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}\right) = \rho^2\sin\theta \tag{23}$$

som är nollskild då $\sin\theta \neq 0$ och $\rho \neq 0$, d.v.s. utanför z-axeln.

4.4 Implicita funktionssatsen för två variabler

SATS
Låt $F(x, y)$ vara en C^1-funktion av två variabler, och låt (a, b) vara en punkt på nivåkurvan $F(x, y) = C$. Om

$$\frac{\partial F}{\partial y}(a, b) \neq 0 \tag{24}$$

så finns det en öppen omgivning $\mathbb{U}$ av (a, b) sådan att restriktionen av nivåkurvan till $\mathbb{U}$ implicit definierar en C^1-funktion $y = f(x)$. För derivatan till funktionen gäller att

$$f'(x) = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))} \tag{25}$$

Beviset till denna sats bygger på inversa funktionssatsen, och från den första delen av implicita funktionssatsen följer den andra m.h.a. kedjeregeln.