Nr 2, 22 feb -05, Amelia 2

2 Funktioner från $\mathbf{R}^n$ till $\mathbf{R}^m$, linjära, inversa och implicita funktioner

2.1 Funktioner från $\mathbf{R}^n$ till $\mathbf{R}^m$

Vi har i tidigare föreläsningar sett olika tolkningar av funktioner från $\mathbf{R}^n$ till $\mathbf{R}^m$ för olika värden på n och m, där de viktigaste är kurvor ($n = 1$), ytor ($n = 2$) och koordinattransformationer ($n = m$) Vi behöver vissa krav på ingående funktioner för att den geometriska formen ska vara väldefinierad, vilket i allmänhet är att funktionerna som förekommer är kontinuerligt deriverbara, och att derivatan i någon mening inte får vara noll (som jacobianens determinant – determinanten av den matris som innehåller alla möjliga partiella förstaderivator). Att derivatan ska vara skild från noll är ett villkor som vi ska se är viktigt även gällande inversa och implicita funktioner.

Exempel 1 *(Uppgift 301b + $\mathbf{r}'(t)$) Bestäm* $\lim_{t\to 0} \mathbf{r}(t)$ *och* $\mathbf{r}'(t)$ *om*

$$\mathbf{r}(t) = (\frac{\sin 2t}{t}, t\cos t).$$

Lösning: Vi beräknar de två gränsvärdena, för x-komponenten och y-komponenten var för sig. Båda måste existera för att $\lim_{t\to 0} \mathbf{r}(t)$ ska existera.

$$\begin{aligned} \lim_{t\to 0} \frac{\sin 2t}{t} &= \lim_{t\to 0} 2\frac{\sin 2t}{2t} = 2 \cdot 1 = 2 \text{ (standardgränsvärde: } \lim_{t\to 0} \frac{\sin t}{t} = 1) \ , \\ t\cos t &\to 0 \cdot 1 = 0. \end{aligned}$$

Dessutom är

$$\begin{aligned} \mathbf{r}'(t) &= (\frac{d}{dt}\frac{\sin 2t}{t}, \frac{d}{dt} t\cos t) \\ &= (\frac{1}{t^2}\left(-\sin 2t + 2t\cos 2t\right), \cos t - t\sin t)). \end{aligned}$$

Svar: Så $\lim_{t\to 0} \mathbf{r}(t) = (2, 0)$ och $\mathbf{r}'(t) = (\frac{1}{t^2}\left(-\sin 2t + 2t\cos 2t\right), \cos t - t\sin t))$.

Exempel 2 *(Uppgift 304 + acceleration) Bestäm hastighet, fart och acceleration för*

$$\mathbf{r}(t) = (\sin t, \cos t, \cos 2t), 0 \le t \le \frac{\pi}{2}$$

i punkten $t = \frac{\pi}{4}$. *I vilken punkt är farten störst?*

Lösning: Vi får hastigeheten genom att derivera m.a.p. t

$$\begin{aligned} \mathbf{r}'(t) &= \frac{d}{dt}(\sin t, \cos t, \cos 2t) \\ &= (\cos t, -\sin t, -2\sin 2t) \end{aligned}$$

och accelerationen genom att derivera en gång till:

$$\begin{aligned}\mathbf{r}''(t) &= \frac{d}{dt}\mathbf{r}'(t) = \frac{d}{dt}(\cos t, -\sin t, -2\sin 2t) \\ &= (-\sin t, -\cos t, -4\cos 2t).\end{aligned}$$

Farten är beloppet av hastigheten:

$$\begin{aligned}|\mathbf{r}'(t)| &= |(\cos t, -\sin t, -2\sin 2t)| \\ &= \sqrt{\cos^2 t + \sin^2 t + 4\sin^2 2t} = \{\text{trig. ettan}\} \\ &= \sqrt{1 + 4\sin^2 2t}.\end{aligned}$$

Vid tiden $\frac{\pi}{4}$ har vi alltså

$$\begin{aligned}\text{Hastighet} \quad &: \quad \mathbf{r}'(\frac{\pi}{4}) = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -2) \\ \text{Acceleration} \quad &: \quad \mathbf{r}''(\frac{\pi}{4}) = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0) \\ \text{Fart} \quad &: \quad \left|\mathbf{r}'(\frac{\pi}{4})\right| = \sqrt{1 + 4\sin^2 2 \cdot \frac{\pi}{4}} = \sqrt{1 + 4\sin^2 \frac{\pi}{2}} = \sqrt{5}.\end{aligned}$$

Var är farten maximal? Man kan derivera farten $\sqrt{1 + 4\sin^2 2t}$ m.a.p. t och sätta denna derivata lika med noll. Men enklare är att observera att maximum för $\sqrt{1 + 4\sin^2 2t}$ inträffar när $\sin^2 2t = 1$. Det ger $t = \frac{\pi}{4}$, ty vi har parameterintervallet $0 \le t \le \frac{\pi}{2}$. Detta t-värde ger punkten

$$\begin{aligned}\mathbf{r}(\frac{\pi}{4}) &= (\sin\frac{\pi}{4}, \cos\frac{\pi}{4}, \cos 2 \cdot \frac{\pi}{4}) \\ &= (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0).\end{aligned}$$

Svar:

$$\begin{aligned}\text{Hastighet i } \frac{\pi}{4}: \quad & \mathbf{r}'(\frac{\pi}{4}) = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -2) \\ \text{Acceleration i } \frac{\pi}{4}: \quad & \mathbf{r}''(\frac{\pi}{4}) = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0) \\ \text{Fart i } \frac{\pi}{4}: \quad & \left|\mathbf{r}'(\frac{\pi}{4})\right| = \sqrt{1 + 4\sin^2 2 \cdot \frac{\pi}{4}} = \sqrt{1 + 4\sin^2 \frac{\pi}{2}} = \sqrt{5}, \\ \text{Maximal fart i } \frac{\pi}{4} \quad : \quad & \mathbf{r}(\frac{\pi}{4}) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0).\end{aligned}$$

Exempel 3 *Bestäm jacobianen* J_f, $\frac{df}{dy}$ *och* $\frac{df}{d(x,z)}$ *för* $\mathbf{f}(x, y, z) = (x+y+x, xy+xz+yz, xyz)$.

Vi betecknar $\mathbf{f}(x,y,z) = (f_1(x,y,z), f_2(x,y,z), f_3(x,y,z))$. Jacobianen är matrisen av alla partiella derivator av funktionerna $x+y+x, xy+xz+yz$ och xyz :

$$\begin{aligned} \frac{d(f_1,f_2,f_3)}{d(x,y,z)} &= \begin{pmatrix} f'_{1x} & f'_{1y} & f'_{1z} \\ f'_{2x} & f'_{2z} & f'_{2z} \\ f'_{2x} & f'_{3y} & f'_{3z} \end{pmatrix} = \{\text{nio partiella deriveringar}\} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ y+z & x+z & x+y \\ yz & xz & yz \end{pmatrix}. \end{aligned}$$

På analogt sätt är

$$\frac{d(f_1,f_2,f_3)}{dy} = \begin{pmatrix} f'_{1y} \\ f'_{2z} \\ f'_{3y} \end{pmatrix} = \begin{pmatrix} 1 \\ x+z \\ xz \end{pmatrix}$$

och

$$\frac{d(f_1,f_2,f_3)}{d(x,z)} = \begin{pmatrix} f'_{1x} & f'_{1z} \\ f'_{2x} & f'_{2z} \\ f'_{2x} & f'_{3z} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ y+z & x+y \\ yz & yz \end{pmatrix}.$$

Svar:

$$\begin{aligned} \frac{d(f_1,f_2,f_3)}{d(x,y,z)} &= \begin{pmatrix} 1 & 1 & 1 \\ y+z & x+z & x+y \\ yz & xz & yz \end{pmatrix}, \\ .\frac{d(f_1,f_2,f_3)}{dy} &= \begin{pmatrix} 1 \\ x+z \\ xz \end{pmatrix}, \frac{d(f_1,f_2,f_3)}{d(x,z)} = \begin{pmatrix} 1 & 1 \\ y+z & x+y \\ yz & yz \end{pmatrix}. \end{aligned}$$

Exempel 4 *(405b) Bestäm jacobianen till* $\mathbf{f}(x,y) = (xy, x^2+y^2)$.

$$J_f = \frac{d(f_1,f_2)}{d(x,y)} = \begin{pmatrix} f'_{1x} & f'_{1y} \\ f'_{2x} & f'_{2z} \end{pmatrix} = \begin{pmatrix} y & x \\ 2x & 2y \end{pmatrix}.$$

Exempel 5 *(406) Antag att* $\mathbf{g}(x,y) = (x^2, xy)$ *och* $\mathbf{f}(u,v) = (uv, u-v)$.*Bestäm* $\mathbf{f}\circ\mathbf{g}$, J_f, J_f *och* $J_{f\circ g}$.

(a)I sammansättningen $\mathbf{f}\circ\mathbf{g}$, får vi $\mathbf{f}(u,v) = \mathbf{f}(g_1(x,y), g_2(x,y))$, så byt ut u mot $g_1(x,y) = x^2$ och v mot $g_2(x,y) = xy$. Det ger

$$\begin{aligned} \mathbf{f}\circ\mathbf{g} &= (uv, u-v) \\ &= (x^2xy, x^2-xy) = (x^3y, x^2-xy). \end{aligned}$$

(c) Vi får genast jacobianen till $\mathbf{f}\circ\mathbf{g}$ genom att derivera partiellt m.a.p. x och y :

$$J_{f\circ g} = \frac{d(x^3y, x^2-xy)}{d(x,y)} = \begin{pmatrix} 3x^2y & x^3 \\ 2x-y & -x \end{pmatrix}.$$

(b) Kedjeregeln (jämför: $\frac{d}{dx} f \circ g = f'(g(x))g'(x)$) säger att

$$J_{f\circ g} = J_f(\mathbf{g}(\mathbf{x}))J_g(\mathbf{x}).$$

Vi har också

$$J_f(u,v) = \frac{d(uv, u-v)}{d(u,v)} = \begin{pmatrix} v & u \\ 1 & -1 \end{pmatrix}$$

och

$$J_g = \frac{d(x^2, xy)}{d(x,y)} = \begin{pmatrix} 2x & 0 \\ y & x \end{pmatrix},$$

så sätter vi in $u = x^2$ och $v = xy$ i J_f så får vi $J_f(\mathbf{g}(\mathbf{x}))$

$$J_f(x^2, xy) = \begin{pmatrix} xy & x^2 \\ 1 & -1 \end{pmatrix}.$$

Matrisprodukten $J_f(\mathbf{g}(\mathbf{x}))J_g(\mathbf{x})$, som alltså är ett alternativt sätt att beräkna $J_{f\circ g}$, beräknad ovan, blir då

$$\begin{aligned} J_f(\mathbf{g}(\mathbf{x}))J_g(\mathbf{x}) &= \begin{pmatrix} xy & x^2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2x & 0 \\ y & x \end{pmatrix} \\ &= \begin{pmatrix} 3x^2y & x^3 \\ 2x-y & -x \end{pmatrix}, \end{aligned}$$

mycket riktigt.

2.2 Linjära funktioner

2.2.1 Linjära och affina funktioner

En viktigt typ av funktion från $\mathbf{R}^n$ till $\mathbf{R}^m$ är en linjär funktion, som defineras av en matris av typ $m \times n$. En vektor $\mathbf{x}$, skriven som kolonnvektor, avbildas på produkten $A\mathbf{x}$, alltså $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$. Detta är en **linjär funktion**.

En translation, som $\mathbf{f}(\mathbf{x}) = \mathbf{x} + \mathbf{a}$, från $\mathbf{R}^n$ till $\mathbf{R}^n$, där a är en fix vektor, kallas en **affin funktion**. Med ovanstående definition är avbildningen $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{a}$ varken en linjär eller affin funktion. Notera emellertid att i vissa böcker som inte är helt matematiska kallas även detta för en linjär funktion.

Vi har också sett att jacobianen för en linjär funktion $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$ är just matrisen A. Detta hänger på att derivatan av cx m.a.p. x är c, helt enkelt.

2.2.2 Matrisnorm

Vi definierar härnäst **normen** av en matris A, som avbildning av $\mathbf{x}$ på $A\mathbf{x}$. Normen anger hur mycket matrisen förstorar en vektor i den riktning där förstoringen är som mest. Därav maximat i definitionen.

Definition 6 *Normen av A är*

$$\|A\| = \max_{\mathbf{x}\neq 0} \frac{|A\mathbf{x}|}{|\mathbf{x}|} = \max_{|\mathbf{x}|=1} A\mathbf{x}.$$

Vi har givetvis vektorbelopp, $|\mathbf{x}| = \sqrt{x_1^2 + ... + x_n^2}$, både i nämnare och täljare. Definitionen kan också tolkas som att $\|A\|$ är det minsta tal C så att

$$|A\mathbf{x}| \leq C|\mathbf{x}|$$

för alla $\mathbf{x}$. Dividera denna olikhet med talet $|\mathbf{x}|$,så fås $\frac{|A\mathbf{x}|}{|\mathbf{x}|} \leq C$.

De två maxima i definitionen är lika, ty enligt räknelagar för vektorer gäller att $\frac{|A\mathbf{x}|}{|\mathbf{x}|} = |A\frac{\mathbf{x}}{|\mathbf{x}|}|$. Här är $\frac{\mathbf{x}}{|\mathbf{x}|}$ är en vektor av längd 1 (en enhetsvektor), som i det högra maximat. Det räcker alltså att studera enhetsvektorer. Därför kan man säga att man söker den *riktning* som maximerar $|A\mathbf{x}|$. Den ger $\|A\|$.

Vi har trivialt exempelvis

$$\begin{aligned} \left\| \begin{pmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix} \right\| &= 5, \text{ (maximerande riktning: } (0,1,0)), \\ \left\| \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right\| &= 1 \text{ (en rotation ändrar inte vektorns längd)} \end{aligned}$$

Som det första exemplet antyder, så gäller att om matrisen är kvadratisk och diagonal får vi $\sqrt{a_{11}^2 v_1^2 + ... + v_n^2 a_{nn}^2}$ att maximera, där $v_1^2 + ... + v_n^2 = 1$. Denna storhet är maximal om vi väljer $v_i = 1$ för det största elementet a_{ii}^2, dvs det gäller att $\|A\| = \max_i |a_{ii}|$.

Om matrisen är diagonaliserbar, och om vi numrerar egenvärdena efter storleken på deras absolutbelopp: $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_n|$ så kommer egenvektorn svarande mot λ_1 att vara maximerande riktning (eller vi kan göra ett basbyte så matrisen blir diagonal). Så för en matris A som är kvadratisk och diagonaliserbar känner vi normen. Den är beloppet av det (till beloppet) största egenvärdet:

$$\|A\| = |\lambda_1|.$$

Detta är ännu en kraftfull tillämpning på egenvärden och diagonalisering, utöver att förstå linjära avbildningar, att exakt bestämma karaktären hos andragradskurvor och att (med lämpligt koordinatbyte) skaffa sig en matris som har många nollor (diagonal).

Det sistnämnda innebär att datorprogram som innehåller matrisberäkningar, som är mycket vanligt i ingenjörsverksamhet, blir mycket snabbare. Man kan bl.a. använda denna ökade effektivitet till att undersöka problem med betydligt högre noggrannhet (öka matrisens dimension $n \times n$).

2.2.3 Värdemängen till en linjär funktion

Notera att värdemängden till en linjär funktion alltid innehåller $\mathbf{0}$, ty $\mathbf{f}(\mathbf{0}) = A\mathbf{0} = \mathbf{0}$. Om matrisen A är nollmatrisen, är värdemängden endast denna punkt: $V_0 = \{\mathbf{0}\}$. I annat fall kommer vi att ha egenskapen att om $\mathbf{a}$ och $\mathbf{b}$ ligger i V_A så gör även $\mathbf{a} + \mathbf{b}$ det. Ty $\mathbf{a}$ och $\mathbf{b}$ i V_A betyder att det finns punkter $\mathbf{x}$ och $\mathbf{y}$ så att $A\mathbf{x} = \mathbf{a}$ och $A\mathbf{y} = \mathbf{b}$. Men då avbildas $\mathbf{x} + \mathbf{y}$ på $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y} = \mathbf{a} + \mathbf{b}$.

På samma sätt kan man visa att om $\mathbf{a} \in V_A$ så är även $\alpha\mathbf{a} \in V_A$ för alla reella tal α.

Så om $\mathbf{a} \in V_A$ så ligger hela räta linjen genom $\mathbf{a}$ och origo i V_A ($\alpha = 0$ ger origo, och alla andra α alla andra punkter på den räta linjen). Om vi har två linjärt oberoende punkter $\mathbf{a}$ och $\mathbf{b}$ som båda ligger i V_A så kommer alla punkter som ligger i det plan som spänns av $\mathbf{a}$ och $\mathbf{b}$ att ligga i V_A. Slutsatsen är att värdemängen till en linjär avbildning är antingen

1. $\{\mathbf{0}\}$ (om A är nollmatrisen)
2. En rät linje genom origo (exempel: projektion på x-axeln)
3. Ett plan genom origo.
4. Ett hyperplan genom origo (alla punkter i $\mathbf{R}^4$ som uppfyller $ax_1 + bx_2 + cx_3 + dx_4 + e = 0$)

Alla dessa mängder är **linjära rum**, som är viktiga i abstrakt linjär algebra. Ett linjärt rum V kännetecknas just av att om $\mathbf{a} \in V$ så $\alpha\mathbf{a} \in V$ för alla reella tal α, och $\mathbf{a} \in V$ och $\mathbf{b} \in V$ implicerar $\mathbf{a} + \mathbf{b} \in V$.

2.2.4 Sammansättning och invers för linjära funktioner

Det är lätt att sätta samman två linjara funktioner $\mathbf{f} = \mathbf{Ax} : \mathbf{R}^p \to \mathbf{R}^m$ och $\mathbf{g} = \mathbf{Bx} : \mathbf{R}^n \to \mathbf{R}^r$. Funktionen $\mathbf{f}(\mathbf{g}(\mathbf{x}))$, också betecknad $\mathbf{f} \circ \mathbf{g}(\mathbf{x})$, är väldefinerad om $r = p$ (då går matrismultiplikationen ihop) och har matrisen AB. Ty

$$\begin{aligned} \mathbf{f}(\mathbf{g}(\mathbf{x})) &= \{\text{sätt in } \mathbf{f} = \mathbf{Ax} \text{ och } \mathbf{g} = \mathbf{Bx}\} \\ &= \mathbf{A}(\mathbf{Bx}) = \{\text{vi har fått en ren} \\ \text{matrisprodukt}\} &= (\mathbf{AB})\mathbf{x}. \end{aligned}$$

För att $\mathbf{f}(\mathbf{g}(\mathbf{x}))$ ska vara definerad kan vi alltså inte ha vilka dimensioner som helst, utan det krävs att $\mathbf{f} = \mathbf{Ax} : \mathbf{R}^p \to \mathbf{R}^m$ och $\mathbf{g} = \mathbf{Bx} : \mathbf{R}^n \to \mathbf{R}^p$. I den sammansatta funktionen $\mathbf{f}(\mathbf{g}(\mathbf{x}))$ krävs att värdena som $\mathbf{g}$ ger ifrån sig ligger i definitionsmängden för $\mathbf{f}$.

Om en linjär funktion $\mathbf{f} = \mathbf{Ax} : \mathbf{R}^n \to \mathbf{R}^n$ är inverterbar så är dess invers också en linjär funktion, och har matrisen A^{-1}.

Bevis: $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) = \mathbf{A}^{-1}(\mathbf{Ax}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{x}$. Att sammansättningen $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))$ är $\mathbf{x}$ betyder att $\mathbf{f}^{-1}$ avbildar tillbaka till $\mathbf{x}$ igen från det värde dit $\mathbf{f}$ avbildade. Därmed är denna funktion invers till $\mathbf{f}$.

2.3 Inversa funktioner

2.3.1 En variabel

Låt oss först betrakta envariabelanalysfallet, vi betraktar funktionen $f(x)$. Om $f(x)$ är strängt monoton, dvs strängt växande eller strängt avtagande, så kan den inte ta samma värde i två punkter (för i så fall måste den ju vara växande någonstans och avtagande någonstans mellan punkterna). Därför är en strängt monoton funktion inverterbar. Så det finns en invers $f^{-1}(x)$ som upphäver f. I formler: $f^{-1}(f(x)) = x$ och $f(f^{-1}(x)) = x$.

Är funktionen dessutom deriverbar så räcker det med att $f'(x) \neq 0$ och att f' är kontinuerlig. För då är antingen $f'(x) > 0$ (som för $f(x) = e^x$) eller $f'(x) < 0$ (som för $f(x) = e^{-x}$). Då är den alltså strängt monoton.

Emellertid är $f'(x) \neq 0$ inte nödvändigt för inverterbarhet, som exemplet $f(x) = x^3$ visar. Derivatan $f'(x) = 3x^2$ är noll då $x = 0$ ($f'(0) = 3 \cdot 0^2 = 0$) men $f(x) = x^3$ är ändå inverterbar. Men om vi kräver att även inversen ska vara deriverbar så är $f'(x) \neq 0$ nödvändigt och tillräckligt. Inversen till x^3 är $\sqrt[3]{x}$, som inte är deriverbar i $x = 0$.

En rät linje $f(x) = y = kx + l$, där $k = f'(x) \neq 0$ har invers (lös ut x ur $y = kx + l$) $x = \frac{1}{k}y - \frac{l}{k}$, så lutningen för inversen är $\frac{1}{k}$. Då fungerar inte $k = 0$. Samma sak gäller generellt, ty om vi deriverar likheten $f(f^{-1}(x)) = x$ m.a.p. x så får vi med kedjeregeln

$$f'(f^{-1}(x)) \left[f^{-1}(x)\right]' = 1.$$

Så $f'(x) = 1/\left[f^{-1}(x)\right]' \neq 0$ (som $1/k$ för en rät linje), ty $\left[f^{-1}(x)\right]'$ är något tal. Detta tal kan inte vara 0, ty vi har förutsatt att f är deriverbar, så f' existerar.

2.3.2 Flera variabler

Som beskrivits tidigare är en linjär funktion $\mathbf{f}(\mathbf{x}) = A\mathbf{x}\ \mathbf{R}^n \to \mathbf{R}^n$ inverterbar om A är inverterbar. Inversen är då $\mathbf{f}^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$. Vi påminner om att för en linjär funktion är jacobianen matrisen A själv, dvs $J_f = A$. Kravet för invers är alltså i det linjära fallet att $\det J_f \neq 0$.

Kravet $\det J_f \neq 0$ gäller inte bara för linjära funktioner, som vi ska se härnäst. Betrakta en funktion $\mathbf{f}(\mathbf{x}) : \mathbf{R}^n \to \mathbf{R}^n$ med kontinuerliga parallella derivator. Att funktionen har kontinuerliga partiella derivator innebär att den är differentierbar, vilket i sin tur innebär att funktionen lokalt (i någon (liten) omgivning runt punkten) är mycket nära en linjär funktion. Då kan man vänta sig att samma villkor gäller som i det linjära fallet, dock endast lokalt, nära punkten:

$$\det J_f \neq 0.$$

Man kan bevisa detta: en differentierbar funktion f har en differentierbar invers *om och endast om* $\det J_f \neq 0$.

Exempel 7 *(501bc) Har* $\mathbf{f}(x, y) = (xy, x+y)$ *differentierbar invers nära* $(-1, 1)$*? Nära* $(1, 1)$*?*

Lösning: Jacobianen är

$$J_f = \begin{pmatrix} y & 1 \\ x & 1 \end{pmatrix},$$

så vi får determinanten

$$\det J_f = \left| \begin{pmatrix} y & 1 \\ x & 1 \end{pmatrix} \right| = y - x.$$

Således har f en differentierbar invers då $y \neq x$, annars inte.

Svar: Så svaret är ja i punkten $(-1, 1)$ men nej i punkten $(1, 1)$.

Exempel 8 *(501d) I vilka punkter har* $\mathbf{f}(x, y) = (x^2 - y^2, \arctan xy)$ *en differentierbar invers?*

Lösning: Jacobianen är

$$J_f = \begin{pmatrix} 2x & -2y \\ \frac{y}{1+x^2y^2} & \frac{x}{1+x^2y^2} \end{pmatrix},$$

å vi får determinanten

$$\det J_f = \left| \begin{pmatrix} 2x & -2y \\ \frac{y}{1+x^2y^2} & \frac{x}{1+x^2y^2} \end{pmatrix} \right| = \frac{2x^2 + 2y^2}{1 + x^2y^2}.$$

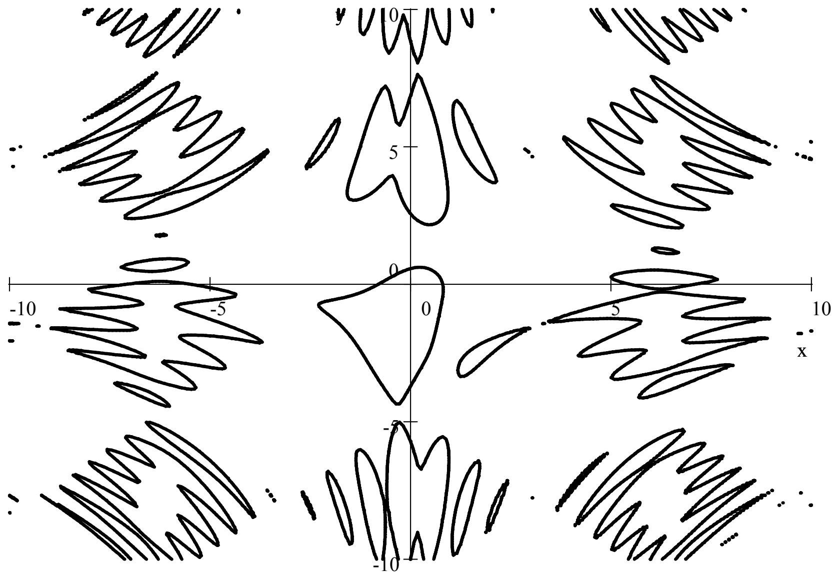
Denna storhet är noll endast då $2x^2 + 2y^2 = 0$, vilket är i origo.

Svar: Således har $\mathbf{f}(x, y) = (x^2 - y^2, \arctan xy)$ en differentierbar invers överallt utom i origo.

2.4 Implicita funktioner

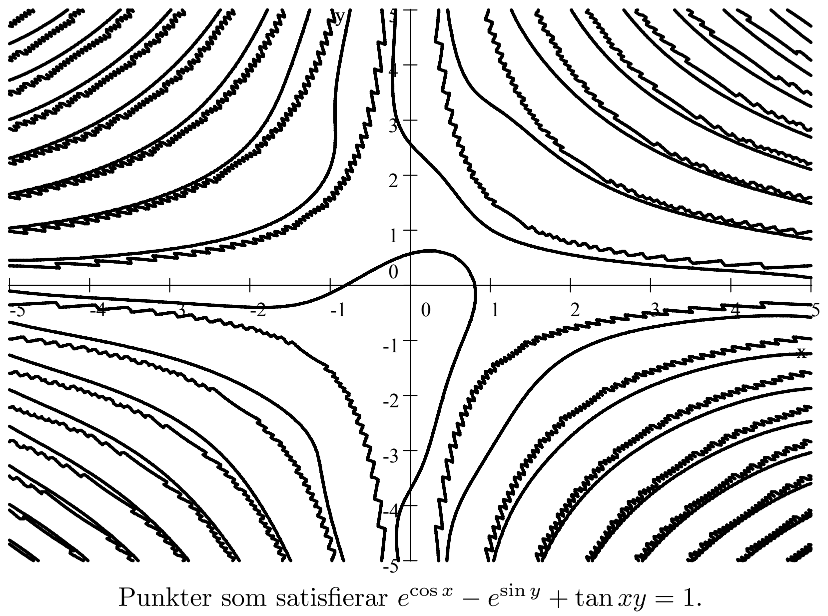
De punkter som satisfierar en ekvation i två variabler bildar ofta en kurva. Ett enkelt exempel är ekvationen $y = f(x)$, där kurvan är grafen till funktionen $f(x)$. Men ofta kan varken x eller y lösas ut ur ekvationen, som är fallet för ekvationen $e^{\cos x} - e^{\sin y} + \sin(xy) - 1 = 0$.. Är det möjligt att ange i vilka punkter kurvan kan beskrivas med en graf till en funktion? Det är den grundläggande frågan i detta avsnitt, vars svar är implicita funktioner.

Ekvationen $e^{\cos x} - e^{\sin y} + \sin(xy) = 1$ är ett (kanske något extremt) exempel på en ekvation där varken x eller y kan lösas ut ur ekvationen. De punkter (x, y) som satisfierar denna ekvation bildar kurvor i planet enligt följande figur:



Punkter som satisfierar $e^{\cos x} - e^{\sin y} + \sin(yx) = 1$.

Här är ett annat säreget exempel:



Punkter som satisfierar $e^{\cos x} - e^{\sin y} + \tan xy = 1$.

Avsnittet om inverterbara funktioner handlar om under vilka villkor det går att definiera en funktion $\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$ från $\mathbf{y} = \mathbf{f}(\mathbf{x})$, dvs från ekvationen

$\mathbf{y} - \mathbf{f}(\mathbf{x}) = \mathbf{0}$. Vi kommer här att generalisera detta problem till en generellare typ av ekvation:

När det gäller implicita funktioner vill vi undersöka när det går att definiera en funktion i någon variabel, $\mathbf{x}$ eller $\mathbf{y}$ från ekvationen

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}.$$

Vi har redan sett exemplen $F(x, y) = y - f(x)$, $F(x, y) = e^{\cos x} - e^{\sin y} + \sin(xy) - 1$ och $F(x, y) = e^{\cos x} - e^{\sin y} + \tan(xy) - 1$. Frågan är: vilket samband definerar $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ mellan variablerna $\mathbf{x}$ och $\mathbf{y}$? Ekvationen ger ett *implicit* samband mellan $\mathbf{x}$ och $\mathbf{y}$. Vi frågar oss vad som krävs för att det ska finnas ett explicit samband.

Generellt är $\mathbf{F}(\mathbf{x}, \mathbf{y})$ en funktion från $\mathbf{R}^{n+m}$ till $\mathbf{R}^n$, så ekvationen $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ är en vektorekvation som svarar mot n ekvationer. Vi vill använda de n ekvationerna till att finna en funktion $\mathbf{f} : \mathbf{R}^m$ till $\mathbf{R}^n$ som uppfyller $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, dvs $\mathbf{F}(\mathbf{x}, \underbrace{\mathbf{f}(\mathbf{x})}_{\mathbf{y}}) = \mathbf{0}$.

Man kan säga att vi ur ekvationen $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ löser ut y som funktion av x.Rättare sagt frågar vi oss när en sådan funktion existerar, snarare än om den kan lösas ut. Det visar sig att en lämplig determinant av en jacobian ska vara skild från noll, även i detta fall.

Vi kan som exempel betrakta ett linjärt fall. Antag att $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ är ett ekvationssystem med fem ($m = 3$, $n = 2$) variabler och två ($n = 2$) linjära ekvationer. Ekvationerna bestämmer ett visst samband mellan de fem variablerna, frågan är vilket. Vi kan här lösa ut två variabler som funktion av de tre övriga, så ekvationssystemet med fem variabler och två ekvationer definerar en funktion $\mathbf{f}$ från $\mathbf{R}^3$ (tre variabler) till $\mathbf{R}^2$ (två variabler). I optimeringslära talar man om **oberoende** ($\mathbf{x}$) och **beroende** variabler ($\mathbf{y}$, beroende av $\mathbf{x}$ enligt $\mathbf{y} = \mathbf{f}(\mathbf{x})$). Naturligtvis kan man på olika sätt välja vilka variabler som ska vara beroende och vilka som ska vara oberoende.

Villkoret från inverterbara funktioner var att jacobianen J_f har determinant skild från noll. Villkoret för en implicit funktion är att jacobianen är skild från noll där vi endast tar med de kolonner i jacobianen som svarar mot de *beoende* variablerna. Denna jacobian blir kvadratisk ($n \times n$), så den har en determinant. Dvs, kravet är att

$$\det \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}) \neq 0.$$

Då definierar ekvationen $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ en funktion $\mathbf{y} = \mathbf{f}(\mathbf{x})$ i någon omgivning till punkten $(\mathbf{x}, \mathbf{y})$.

Exempel 9 *(520) Visa att det finns en funktion* $(y(x), z(x))$ $\mathbf{R}^1$ *till* $\mathbf{R}^2$ *som defineras av*

$$\begin{cases} xy + yz + xz = 2 \\ x^3 + y^3 + z^3 = 9 \end{cases}.$$

Lösning: Här har vi

$$\mathbf{F}(x, y, z) = (xy + yz + xz, x^3 + y^3 + z^3).$$

Beroende variabler ska vara y och z, så intressant jacobian är

$$\begin{aligned}\frac{\partial \mathbf{F}}{\partial(y,z)} &= \begin{pmatrix} \frac{\partial}{\partial y}(xy+yz+xz) & \frac{\partial}{\partial z}(xy+yz+xz) \\ \frac{\partial}{\partial y}(x^3+y^3+z^3) & \frac{\partial}{\partial z}(x^3+y^3+z^3) \end{pmatrix} \\ &= \begin{pmatrix} x+z & x+y \\ 3y^2 & 3z^2 \end{pmatrix},\end{aligned}$$

som är funktion av tre variabler. Dess determinant är

$$(x+z)3z^2 - 3y^2(x+y).$$

I punkten $(0, 1, 2)$ har vi alltså

$$\det \frac{\partial \mathbf{F}}{\partial(y,z)}(0,1,2) = (0+2)3 \cdot 2^2 - 3 \cdot 1^2(0+1) = 21.$$

Detta är inte noll, så svaret är:

Svar: ja.

Exempel 10 *(511) Undersök om* $x^4 + y^4 + x + y = 0$ *definerar en funktion* $y = y(x)$ *nära* $(0, 0)$.

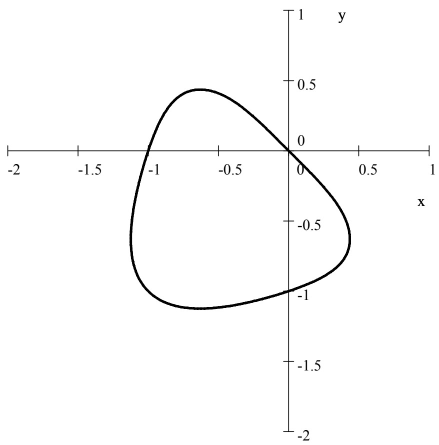
Lösning: Jacobianen av $F(x, y) = x^4 + y^4 + x + y$ är

$$\frac{\partial \mathbf{F}}{\partial(x,y)} = (4x^3 + 1, 4y^3 + 1)$$

så vi har i punkten $(0, 0)$

$$\frac{\partial \mathbf{F}}{\partial(x,y)}(0,0,) = (4 \cdot 0^3 + 1, 4 \cdot 0^3 + 1) = (1, 1).$$

Det går således bra att lösa ut $y = y(x)$.



Kurvan $x^4 + y^4 + x + y = 0$.

Exempel 11 *(513b) Visa att i någon omgivning till* $(1,0)$ *definerar* $x^y + \sin y + x - 2 = 0$ *en kontinuerligt deriverbar funktion* $y = y(x)$.

Lösning: Vi har $F(x,y) = x^y + \sin y + x - 2$ och

$$\frac{\partial F}{\partial(x,y)} = (yx^{y-1} + 1, x^y \ln x + \cos y),$$

så i $(x,y) = (1,0)$ får vi

$$\frac{\partial F}{\partial(x,y)}(1,0) = (1,1)$$

Eftersom de partiella derivatorna är skilda från noll i punkten så existerar $y = y(x)$ i någon omgivning till $(1,0)$.

Låt oss avslutningsvis betrakta sambandet mellan implicit och invers funktion i envariabelfallet. Här är vi intresserade av ekvationen $y = f(x)$. Den kan alternativt formuleras som ekvationen $F(x,y) = 0$, där $F(x,y) = y - f(x)$.Här har vi jacobianen

$$\frac{\partial F}{\partial(x,y)} = \begin{pmatrix} -f'(x) & 1 \end{pmatrix},$$

Så om x ska vara oberoende variabel, dvs vi vill lösa ut y, så är vi intresserade av jacobianen m.a.p. den beroende variabeln y, som är

$$\frac{\partial F}{\partial y} = 1.$$

Den är aldrig är noll, och vi har heller inga problem med att lösa ut y, ty det är ett helt trivialt fall. Vi har redan $y = f(x)$.

Om y ska vara oberoende variabel så vill vi lösa ut x, och vi är intresserade av den inversa funktionen. Aktuell jacobian är då den som svarar mot variabeln x

$$\frac{\partial F}{\partial(x)} = -f'(x).$$

Kravet är att denna jacobian inte ska vara noll, dvs att $f'(x) \neq 0$. I så fall finns lokalt en funktion $f^{-1}(y)$.