Normált térbeli differenciálszámítás alkalmazásai

Szakdolgozat

Írta

Nagy Zsuzsanna

Matematika B.S.c., Matematikai elemző szakirány

Témavezető:

Simon Péter

Egyetemi docens

Alkalmazott Analízis és Számításmatematikai Tanszék

Eötvös Loránd Tudományegyetem

Budapest

2012

Tartalomjegyzék

1. Bevezetés

A differenciálszámítás az integrálszámítással együtt a matematikai analízis egyik legfőbb területe, amely leegyszerűsítve a változás pillanatnyi sebességének vizsgálatával foglalkozik.

A differenciálszámítás múltja hosszú évszázadokra tekint vissza - a tudósok először a XVI. század vége felé kezdtek foglalkozni ezzel a területtel. A matematikusok sokszor egymástól függetlenül dolgozva a XVII. század során kidolgozták a terület alapjait és jelentős lökést adtak az egyetemes matematika fejlődésének is. A témával a kor legnagyobb elméi, így például Leibniz és maga Isaac Newton foglalkoztak.

A differenciálszámítás napjainkra kiforrott és megszilárdult, az analízis egyik alappillérévé vált, és mint ilyet, számos tudományterületen alkalmazzák. A fizikában például a mozgási törvények esetében, a kémiában a reakcióidők, az operációkutatásban a gazdaságosság kiszámításakor használják, nem említve számtalan egyéb területet.

Szakdolgozatom témájául azért választottam a differenciálszámítást, mivel tanulmányaim során az analízis, azon belül is ez a terület mindig is különösen vonzott, érdekelt. Ezen felül az egyetemi éveim alatt már alaposan megismert egy- és többváltozós deriváláson túl olyasmivel akartam foglalkozni, ami részben számomra is új - ezért választottam a normált térbeli differenciálszámítást.

2. fejezet

2.1. Alapfogalmak

Az egyváltozós differenciálszámítás számos matematikai, fizikai és egyéb területről adódó problémára megoldást nyújt. Az egyik ilyen probléma az egyenes vonalú egyenletes mozgást végző pont. Legyen $s(t)$ a t időpillanatban a pontnak a helyzete a számegyenesen. Ekkor a t_0-beli pillanatnyi sebességet úgy definiáljuk, mint a

$$\lim_{x\to a} \frac{s(t)-s(t_0)}{t-t_0}$$

határértéket (feltéve, hogy ez létezik és véges).

Egy matematikai területről származó probléma az f függvény grafikonjának az $(a, f(a))$ pontbeli érintőjének meghatározása. Ezt az érintőt úgy definiáljuk, mint azt az egyenest, ami átmegy az $(a, f(a))$ ponton és meredeksége

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a},$$

(feltéve, hogy ez a határérték létezik és véges.)

Egy újabb példa, ha egy test hőmérsékletének változására vagyunk kíváncsiak egy t_0 időpillanatban. Jelölje $H(t)$ a t időpillanatban a test hőmérsékletét, ekkor azt vizsgáljuk, hogy milyen gyorsan változik a hőmérséklet egy t_0 pillanatban. A $[t_0, t]$ intervallumban a változás pillanatnyi átlagsebessége

$$\lim_{t\to t_0} \frac{H(t)-H(t_0)}{t-t_0}.$$

(feltéve, hogy létezik és véges).

A példák általánosításaként az alábbi definíciót fogalmazhatjuk meg:

1. Definíció. *Legyen f értelmezve az a pont egy környezetében. Azt mondjuk, hogy az f függvény az a pontban differenciálható, ha a*

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$$

véges határérték létezik. Ez a határérték az f függvény a pontbeli differenciálhányadosa vagy deriváltja.

A fent említett matematikai területről származó problémára precíz megoldást ad a következő definíció:

2. Definíció. *Legyen f differenciálható az a pontban. Az f függvénygrafikon (a,f(a)) pontbeli érintőjén az* $y = f'(a) \cdot (x - a) + f(a)$ *egyenletű egyenest értjük.*

A következő tétel azt mutatja meg, hogy a differenciálhatóság erősebb fogalom, mint a folytonosság. Ha egy függvény differenciálható a megfelelő intervallumon, akkor abból következik, hogy ezen intervallumon a függvény folytonos, ez viszont visszafelé nem teljesül. Vagyis, ha egy függvény folytonos, nem biztos, hogy differenciálható lesz, ezt mutatja meg a megjegyzés.

1. Tétel. *Ha f differenciálható a-ban, akkor f folytonos* [1] *a-ban.*

Bizonyítás: Ha f differenciálható a-ban, akkor

$$\lim_{x\to a}(f(x) - f(a)) = \lim_{x\to a}\left[\frac{f(x) - f(a)}{x - a} \cdot (x - a)\right] = f'(a) \cdot 0 = 0.$$

Ez éppen azt jelenti, hogy f folytonos a-ban.

1. Megjegyzés. *A folytonosság a differenciálhatóságnak szükséges, de nem elégséges feltétele. Van olyan f függvény, amely egy a pontban folytonos, de ott nem differenciálható. Egyszerű példa erre az* $f(x) = |x|$ *függvény a=0-ban. Valóban,*

$$f(x) = \begin{cases} x, & ha\ x \geq 0, \\ -x, & ha\ x < 0, \end{cases}$$

tehát

$$\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1, & ha\ x > 0, \\ -1, & ha\ x < 0. \end{cases}$$

Így

$$\lim_{x\to 0+0} \frac{f(x) - f(0)}{x - 0} = 1$$

és

$$\lim_{x\to 0-0} \frac{f(x) - f(0)}{x - 0} = -1,$$

ezért f nem differenciálható 0-ban.

Az előbbi megjegyzés során megmutattuk, hogy $f(x) = |x|$ függvény nem differenciálható 0-ban, ezért nem lesz a függvény grafikonjának érintője 0-ban. Viszont, ha 0-nak csak a jobboldali vagy csak a baloldali környezetét tekintjük, akkor itt a differenciálhányadosok féloldali határértéke már létezik.

---

[1] Az f folytonos a-ban, ha f értelmezve van az a pont egy környezetében és

$$\lim_{x\to a} f(x) = f(a).$$

2.2. Elemi függvények deriváltjai és differenciálási szabályok

Dolgozatom terjedelmi korlátai miatt az elemi függvények derivált függvényeinek létezését nem bizonyítom, hanem táblázatba foglalom.

$D(f)$ és $D(f')$	$f(x)$	$f'(x)$
$\mathbb{R}$	c $(c \in \mathbb{R})$	0
$x \in \mathbb{R}$, de $x \neq 0$	x^n $(n \in \mathbb{Z})$	nx^{n-1}
$\mathbb{R}$	e^x	e^x
$(0, +\infty)$	$\ln x$	$\frac{1}{x}$
$\mathbb{R}$	a^x,$a \in (0, +\infty)$	$a^x \cdot \ln a$
$(0, +\infty)$	$\log_a x (a \in \mathbb{R}^+$, de $a \neq 1)$	$\frac{1}{x \cdot \log a}$
$\mathbb{R}$	$\sin x$	$\cos x$
$\mathbb{R}$	$\cos x$	-$\sin x$
$x \in \mathbb{R}$, de $x \neq \frac{\pi}{2} + k\pi$ $(k \in \mathbb{Z})$	tg$x$	$\frac{1}{\cos^2 x}$
$x \in \mathbb{R}$, de $x \neq k\pi$ $(k \in \mathbb{Z})$	ctg$x$	-$\frac{1}{\sin^2 x}$

A következő tétel során a differenciálási szabályokat tekintjük át.

2. Tétel. *Ha az f és g függvények differenciálhatóak a-ban, akkor $cf(c \in \mathbb{R})$, $f + g$, $f \cdot g$ is differenciálható a-ban, és*

(i) $(cf)'(a) = cf'(a)$,

(ii) $(f + g)'(a) = f'(a) + g'(a)$,

(iii) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

Ha $g(a) \neq 0$, akkor $\frac{1}{g}$ és $\frac{f}{g}$ is differenciálható a-ban, és

(iv) $(\frac{1}{g})'(a) = -\frac{g'(a)}{g^2(a)}$,

(iiv) $(\frac{f}{g})'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$.

A következő tételben a láncszabály pontos meghatározása szerepel. Ez a láncszabály az összetett függvények deriválásánál szükséges.

3. Tétel. *Ha a g függvény differenciálható a-ban és az f függvény differenciálható $g(a)$-ban, akkor a $h = f \circ g$ függvény is differenciálható a-ban, és*

$$h'(a) = f'(g(a)) \cdot g'(a).$$

Ez a tétel az ún. láncszabály.

Az $y = g(x)$ és $z = f(y)$ jelölést alkalmazva az alábbi módon könnyen kiszámolható

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

2.3. A lokális tulajdonságok és derivált kapcsolata

Mint a fejezet elején is láttuk, a differenciálszámítást számos tudományos területen alkalmazzák. Matematikai alkalmazások közül a függvény vizsgálata is egy differenciálszámítási alkalmazás. Ebben az alfejezetben a függvények lokális tulajdonságai és deriváltjaik kapcsolatát tekintjük át.

3. Definíció. *Legyen az f függvény értelmezve az a pont egy környezetében. Azt mondjuk, hogy f lokálisan növekedő a-ban, ha van olyan $\delta > 0$, hogy minden $a-\delta < x < a$ esetén $f(x) \leq f(a)$, és minden $a < x < a+\delta$ esetén $f(x) \geq f(a)$.*

Analóg módon értelmezzük a szigorúan lokálisan növekedő, lokálisan csökkenő és szigorúan lokálisan csökkenő függvényt. Tovább folytatva a lokális tulajdonságok áttekintését:

4. Definíció. *Azt mondjuk, hogy az f függvénynek az a pontban lokális maximuma (illetve minimuma) van, ha a-nak van olyan U környezete, amelyben f értelmezve van, és minden $x \in U$-ra $f(x) \leq f(a)$ (illetve $f(x) \geq f(a)$). Ekkor az a pontot az f függvény lokális maximumhelyének (illetve lokális minimumhelyének) nevezzük.*

A következő tétel nagyon fontos a függvények vizsgálatánál. A derivált előjeléből tudunk következtetni a függvény lokális tulajdonságaira. Fontos megemlíteni, hogy a tételben szereplő állítások csak egyik irányban teljesülnek.

4. Tétel. *Tegyük fel, hogy f differenciálható a-ban.*

(i) Ha $f'(a) > 0$, akkor f szigorúan lokálisan növekedő a-ban.
(ii) Ha $f'(a) < 0$, akkor f szigorúan lokálisan csökkenő a-ban.
(iii) Ha f lokálisan növekedő a-ban, akkor $f'(a) \geq 0$.
(iv) Ha f lokálisan csökkenő a-ban, akkor $f'(a) \leq 0$.
(v) Ha f-nek lokális szélsőértékhelye van a-ban, akkor $f'(a) = 0$.

2.4. Középértéktételek

Ebben a részben kettő fontos tétellel fogunk megismerkedni, amelyek a differenciálszámítás során leggyakrabban használtak közül valók. Ezek azok a tételek, amelyeket a leggyakrabban akkor szoktunk alkalmazni, ha kapcsolatot keresünk egy függvény és deriváltjainak tulajdonságai között.

5. Tétel. *(Rolle tétele) Tegyük fel, hogy az f függvény folytonos $[a,b]$-ben és differenciálható (a,b)-ben. Ha $f(a) = f(b)$, akkor létezik olyan $c \in (a,b)$, amelyre $f'(c) = 0$.*

Bizonyítás Ha $f(x) = f(a)$ minden $x \in (a, b)$-re, akkor f konstans (a, b)-ben, tehát $f'(x) = 0$ minden $x \in (a, b)$ esetén. Ekkor c gyanánt bármelyik (a, b)-beli számot választhatjuk.

Feltehetjük tehát, hogy van olyan $x_0 \in (a, b)$, amelyre $f(x_0) \neq f(a)$. Tekintsük először azt az esetet, amikor $f(x_0) > f(a)$. Weierstrass tétele szerint f-nek létezik abszolút maximumhelye $[a, b]$-ben. (Mert f folytonos $[a, b]$-ben). Mivel $f(x_0) > f(a) = f(b)$, ezért sem az a pont, sem a b pont nem lehet abszolút maximumhely. Ha tehát c egy abszolút maximumhely, akkor $c \in (a, b)$, és így c egyszersmind lokális maximumhely is. Az 4. Tétel (v) állítása szerint ebből következik, hogy $f'(c) = 0$.

Ha $f(x_0) < f(a)$, akkor ugyanígy okoskodhatunk, csak az f függvény abszolút minimumhelyeit kell tekintenünk. □

A másik fontos tétel a Lagrange-középértéktétel. Ezt nem bizonyítjuk, de Rolle-tétel segítségével könnyen belátható.

6. Tétel. *(Lagrange-középértéktétel) Ha az f függvény folytonos $[a, b]$-ben és differenciálható (a, b)-ben, akkor létezik olyan $c \in (a, b)$, amelyre*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Az előbbi tétel fontos következménye a következő tétel:

7. Tétel. *Ha f folytonos $[a, b]$-ben, differenciálható (a, b)-ben és $f'(x) = 0$ minden $x \in (a, b)$-re, akkor az f függvény konstans $[a, b]$-ben.*

Bizonyítás A Lagrange-középértéktételből következik, hogy minden (a, b)-beli x-hez létezik olyan $c \in (a, b)$, amelyre

$$f'(c) = \frac{f(x) - f(a)}{x - a}.$$

Így $f'(c) = 0$ miatt $f(x) = f(a)$. □

2.5. Szélsőérték keresése

A következő tétel segít nekünk abban, hogy egy tetszőleges deriválható függvénynek a lokális és abszolút szélsőértékeit megkeressük. Ugyanis a derivált előjeléből leolvashatjuk, hogy a függvény mely intervallumokon nő és melyeken csökken, és ez általában elegendő a szélsőértékek megkereséséhez.

8. Tétel. *Legyen f folytonos $[a, b]$-ben és differenciálható (a,b)-ben.*

(i) f akkor és csak akkor monoton növekedő (ill. monoton csökkenő) $[a,b]$*-ben, ha f'(x)* $\geq$*0 (ill. f'(x)* $\leq$ *0) minden* $x \in (a,b)$*-re.*
(ii) f akkor és csak akkor szigorúan monoton növő (ill. szigorúan monoton csökkenő) $[a,b]$*-ben, ha f'(x)*$\geq$*0 (ill. f'(x)* $\leq$ *0) minden* $x \in$ *(a,b)-re, és ha* $[a,b]$*-nek nincs olyan részintervalluma, amelyen f' azonosan nulla.*

Bizonyítás:

(i) Tegyük fel, hogy $f'(x) \geq 0$ minden $x \in (a,b)$-re. A Lagrange-középértéktétel szerint tetszőleges $a \leq x_1 < x_2 \leq b$-hez létezik olyan $c \in (x_1, x_2)$, hogy

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(c).$$

Mivel $f'(c) \geq 0$, így $f(x_1) \leq f(x_2)$, ami éppen azt jelenti, hogy f monoton növekedő $[a,b]$-ben.

Megfordítva, ha f monoton növekedő $[a,b]$-ben, akkor minden (a,b)-beli x helyen lokálisan növekedő. Ezért a 4. Tétel (iii) állításából következik, hogy $f'(x) \geq 0$.

Hasonló bizonyítás a monoton csökkenő függvény esetére.

(ii) Könnyen látható, hogy egy f függvény akkor és csak akkor szigorúan monoton $[a,b]$-n, ha monoton, és $[a,b]$-nek nincs olyan részintervalluma, amelyen f konstans. Így az állítás következik a 7. Tételből.

Az 4. tételből megtudtuk, hogy ha f deriválható a-ban, akkor ahhoz, hogy f-nek a-ban lokális szélsőértékhelye legyen, szükséges, de nem elégséges feltétele, hogy $f'(a)$=0. Az alábbi tételek *elégséges* feltételt adnak a lokális szélsőértékhely létezésére.

9. Tétel. *Legyen f differenciálható az a pont egy környezetében. Ha f'(a)=0 és f' az a pontban előjelet váltva 0, azaz a egy bal oldali környezetében nempozitív és egy jobb oldali környezetében nemnegatív (vagy fordítva), akkor az a pont f-nek lokális minimumhelye (ill. maximumhelye).*

A függvény második deriváltja megmutatja, hogy a függvénynek egy adott pontban van-e lokális szélsőértéke.

10. Tétel. *Legyen* f *kétszer differenciálható* a*-ban. Ha* $f'(a) = 0$ *és* $f''(a) > 0$*, akkor* f*-nek* a*-ban szigorú lokális minimuma van. Ha* $f'(a) = 0$ *és* $f''(a) < 0$*, akkor* f*-nek szigorú lokális maximuma van.*

2. Megjegyzés. *Ha $f'(a) = 0$ és $f''(a) = 0$, akkor nem tudjuk biztosan megmondani, hogy f-nek van vagy nincs lokális szélsőértékhelye a-ban. A különböző lehetőségeket mutatják pl. az $f(x) = x^3$, $f(x) = x^4$ és $f(x) = -x^4$ függvények az $a = 0$ helyen. Ebben az esetben a magasabb rendű deriváltak értékéből lehet elégséges feltételt megtudni arról, hogy az a pont f-nek lokális szélsőértékhelye lesz-e.*

11. Tétel. *(i) Legyen az f függvény $2k$-szor differenciálható az a pontban, ahol $k \geq 1$. Ha*

$$f'(a) = \ldots = f^{(2k-1)}(a) = 0 \quad \text{és} \quad f^{(2k)} > 0,$$

akkor f-nek a-ban szigorú lokális minimuma van. Ha

$$f'(a) = \ldots = f^{(2k-1)}(a) = 0 \quad \text{és} \quad f^{(2k)} < 0,$$

akkor f-nek a-ban szigorú lokális maximuma van.

(ii) Legyen az f függvény $2k+1$-szer differenciálható a-ban, akol $k \geq 1$. Ha

$$f'(a) = \ldots = f^{(2k)}(a) = 0 \quad \text{és} \quad f^{(2k+1)}(a) \neq 0,$$

akkor f szigorú monoton az a pont egy környezetében, tehát a-ban nincs lokális szélsőértékhelye.

2.6. Alkalmazás

Mint láttuk a differenciálszámítás a különböző tudományterületeken sokféleképpen alkalmazható. A matematika területén például a L'Hospital-szabály esetében, ami a kritikus határértékek meghatározásához nyújt segítséget. Egy másik példa a polinomapproximáció, amiben többek között a Taylor-polinom segít abban, hogy jó közelítést kapjunk egy adott függvényhez. A fizika és a kémia területén is számos alkalmazását találhatjuk a differenciálszámításnak. A következő feladat példa egy egyszerű kémiai folyamatra, az oldódásra, amit közönséges differenciálegyenletekkel meg tudunk oldani.

1. Feladat. *A 100 l vizet tartalmazó edénybe literenként 10 kg sót tartalmazó oldat folyik be folyamatosan 5 l/perc sebességgel. Az edénybe belépő folyadék összekeveredik a vízzel, és a keverék ugyanolyan sebességgel kifolyik az edényből. Mennyi só lesz az edényben 1 óra múlva?*

Megoldás: Jelölje a vízben lévő só mennyiségét $x(t)$ (kg-ban mérve). Jelölje $\triangle t$ az eltelt időt. Felírható a következő egyenlet:

$$x(t + \triangle t) = x(t) - \triangle t \cdot \frac{5 \cdot x(t)}{100}$$

Az egyenlet baloldalából kivonva $x(t)$-t és utána elosztva $\triangle t$-vel ($\triangle t \to 0$), a következőt kapjuk:

$$\frac{x(t+\triangle t)-x(t)}{\triangle t}=-\frac{x(t)}{20}$$

A baloldalon a differenciálhányados szerepel, vagyis a következő differenciálegyenletet kapjuk:

$$\dot{x}(t)=-\frac{1}{20}x(t)$$

Ennek az egyenletnek a megoldása:

$$x(t)=c\cdot e^{-\frac{1}{20}t}$$

A kezdeti feltétel $x(0)=10$ alkalmazva, vagyis a fenti egyenletben a t helyére 0-t írva kapjuk, hogy $c=10$. Vagyis azt kapjuk, hogy

$$x(t)=10\cdot e^{-\frac{1}{20}t}$$

Most már ki tudjuk számolni, hogy egy óra elteltével mennyi só marad a tartályban. A fenti egyenletbe t helyére 60 percet helyettesítve kapjuk $10\cdot e^{-3}$.

3. fejezet

3.1. $\mathbb{R}^p \to \mathbb{R}$ függvények differenciálszámítása

3.1.1. Parciális deriváltak

Az egyváltozós differenciálszámítás után, ebben a fejezetben a többváltozós függvények differenciálszámításával foglalkozunk, az [1] hivatkozásban szereplő könyv anyaga alapján. Hasonlóan az egyváltozós esethez itt is számos tudományos területen felmerülő probléma esetében adnak segítséget a parciális differenciálegyenletek. Ilyen például a hővezetési egyenlet, aminek a segítségével ki tudjuk számolni hogyan változik a hőmérséklet egy helyiségben. A másik legismertebb fizikai példa a hullámegyenletek, ami többek között a hangok terjedésével is foglalkozik. Most pedig nézzük meg, hogy mi is az a parciális derivált.

5. Definíció. *Legyen az f függvény értelmezve az $a = (a_1, \ldots, a_p) \in \mathbb{R}^p$ pont egy környezetében. Rögzítsük az $a = (a_1, \ldots, a_p)$ pont koordinátáit az i-edik kivételével, és tekintsük a megfelelő*

$$t \mapsto f_i(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p)$$

szekciófüggvényt. Az így kapott egyváltozós f_i függvény a_i pontban vett deriváltját (amennyiben létezik) az f függvény a pontban vett i-edik parciális deriváltjának nevezzük. Más szóval

$$D_i f(a) = \lim_{t \to a_i} \frac{f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p) - f(a)}{t - a_i}$$

feltéve, hogy a (véges vagy végtelen) limesz létezik.

Legyen az f függvény értelmezve $\mathbb{R}^p$ egy részhalmazán. Az f függvény i-edik parciálisderivált függvényén azt a $D_i f$ függvényt értjük, amely azon a pontokban van értelmezve, ahol f i-edik parciális deriváltja létezik és véges, és ott az értéke $D_i f(a)$.

Egyváltozós esetben már beláttuk, hogy ha egy f függvénynek az a pontban lokális szélsőértéke van és f ott differenciálható, akkor $f'(a)$=0 (a 4. Tétel (v) állítása). Ezt általánosíthatjuk a többváltozós függvényekre.

6. Definíció. *Azt mondjuk, hogy az f függvénynek az $a \in \mathbb{R}^p$ pontban lokális maximuma van, ha van egy olyan $r > 0$, amelyre f értelmezve van a $B(a, r)$ gömbben (a középpontú r sugarú nyílt gömb) és minden $x \in B(a, r)$-re az $f(x) \leq f(a)$ (ill. $f(x) \geq f(a)$).*

Ha minden $x \in B(a,r) \setminus a$ pontra $f(x) < f(a)$ (ill. $f(x) > f(a)$), akkor szigorú lokális maximumról és maximumhelyről (ill. minimumról és minimumhelyről) beszélünk.

Tegyük fel, hogy f-nek lokális maximuma van az $a = (a_1, \ldots, a_p)$ pontban. Ekkor f minden parciális deriváltjának vagyis f_i-nek is lokális maximuma van az a_i pontban. Ha az f_i függvény differenciálható a_i-ben, akkor $f_i'(a_i) = 0$. A következő tétel ezt írja le:

12. Tétel. *Ha az f függvénynek lokális szélsőértéke van az $a \in \mathbb{R}^p$ pontban, és f-nek léteznek a parciális deriváltjai a-ban, akkor $D_i f(a) = 0$ minden $i = 1, \ldots, p$-re.*

Az előző tétel segítségével a többváltozós függvények szélsőérték vizsgálatát visszavezethetjük több egyváltozós függvény szélsőérték vizsgálatára, amelyek folytonosak egy korlátos és zárt intervallumon és differenciálhatóak az intervallum belsejében. Ezt általánosíthatjuk többváltozós függvényekre.

13. Tétel. *Legyen $A \subset \mathbb{R}^p$ korlátos és zárt, legyen $f : A \to \mathbb{R}$ folytonos, és tegyük fel, hogy f-nek léteznek a parciális deriváltjai A belsejének minden pontjában. Ekkor f a legnagyobb (legkisebb) értékét vagy A határán veszi fel, vagy pedig egy olyan a belső pontban, ahol $D_i f(a) = 0$ minden $i = 1, \ldots, p$-re.*

Bizonyítás: A Weierstrass-tétel szerint f-nek van legnagyobb értéke A-n. Legyen $a \in A$ olyan pont, amelyben f értéke a legnagyobb. Csak az $a \in intA$ esetet kell vizsgáljuk. Ekkor f-nek lokális maximuma van a-ban. A tételben feltettük, hogy f-nek léteznek a parciális deriváltjai a pontban, tehát $D_i f(a) = 0$ minden $i = 1, \ldots, p$-re a 12. Tétel miatt.

2. Feladat. *Mi az $f(x,y) = xy + x^2 + y^2$ függvénynek a minimuma és maximuma a $(0,0)$ középpontú $r = 1$ sugarú zárt körlapon?*

Megoldás: Először megnézzük, hogy a függvényt mely halmazon vizsgáljuk. $H = \{(x,y) : |(x,y)| \leq 1\}$. Ezután megnézzük, hogy hol 0 mindkét parciális derivált.

$$D_1 f(a) = y + 2x$$

$$D_2 f(a) = x + 2y$$

Tehát:

$$y + 2x = 0$$

$$x + 2y = 0$$

egyenletrendszert megoldva kapjuk, hogy $x = 0$, és $y = 0$. Azt kaptuk, hogy az egyetlen jelölt a körlap belsejében $(0,0)$ pont.

Most pedig a körlap határán vizsgáljuk a függvényt. A határ:

$$S = \{(x,y) : |(x,y)| = 1\}$$

körvonal, és ezen keresünk minimumot és maximumot. Ehhez kifejezzük a körvonal pontjait egy paraméterrel. Legyen ez a paraméter a szög.

$$S = \{(\cos\alpha, \sin\alpha) : \alpha \in [0, 2\pi]\}.$$

S pontjaiban f értéke:

$$f(\cos\alpha, \sin\alpha) = \cos\alpha \cdot \sin\alpha + \cos^2\alpha + \sin^2\alpha$$

Tehát

$$f(\cos\alpha, \sin\alpha) = \cos\alpha \cdot \sin\alpha + 1$$

függvénynek kell a minimuma és maximuma $[0, 2\pi]$ intervallumon.

Jelöltek: $f'(\alpha) = 0$ és az intervallum két széle: 0 és 2π.
Vagyis a jelöltek: $\frac{\pi}{4}$,$\frac{3\pi}{4}$,$\frac{5\pi}{4}$,$\frac{7\pi}{4}$,0,2π.
Az f függvény maximum értéke a körvonalon: $\frac{3}{2}$, és ezt az értéket felveszi az $\alpha = \frac{\pi}{4}$ és $\frac{5\pi}{4}$ pontokban. Minimum értéke a körvonalon: $\frac{1}{2}$, amit az $\alpha = \frac{3\pi}{4}$ és $\frac{7\pi}{4}$ pontokban vesz fel.

Az intervallum belsejében az origó volt az egyetlen jelölt, ahol az f függvény értéke: 0.

Tehát maximum: $\frac{3}{2}$, amit f a határon veszi fel $\alpha = \frac{\pi}{4}, \frac{5\pi}{4}$, minimum: 0, amit az origóban vesz fel.

3.1.2. Differenciálhatóság

A többváltozós függvények esetében is szükség lehet arra, hogy az eredeti függvényeket jól közelítő függvényekkel helyettesítsük. A fizikai alkalmazásokban például sokszor élünk azzal a feltételezéssel, hogy egy függvényt egy adott pont kis környezetében jól közelíthető legyen elsőfokú polinommal. Egyváltozós esetben láttuk (2. Definíció), hogy $y = f'(a) \cdot (x - a) + f(a)$ egyenletű egyenessel tudjuk jól közelíteni egy f függvényt adott a pontban. A következő definíció felhasználja ezt a polinomot ahhoz, hogy szükséges és elégséges feltételt adjon arra, hogy mikor lehet egy függvényt deriválni egy adott a pontban.

1. Állítás. *Az f függvény differenciálható a-ban, ha*

$$\lim_{x \to a} \frac{|f(x) - [f'(a) \cdot (x-a) + f(a)]|}{|x-a|} = 0$$

Ez lesz a definíció többváltozós esetben is. Először definiáljuk, hogy egy függvény mikor lineáris.

7. Definíció. *Az $l : \mathbb{R}^p \to \mathbb{R}$ függvényt lineárisnak nevezzük, ha vannak olyan $\alpha_1, \ldots, \alpha_p$ valós számok, hogy $l(x) = \alpha_1 x_1 + \ldots + \alpha_p x_p$ minden $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$-re.*

Az előbbi definíció segítségével megfogalmazhatjuk azt, hogy egy többváltozós f függvény mikor differenciálható egy adott a pontban.

8. Definíció. *Tegyük fel, hogy $H \subset \mathbb{R}^p$ és $f : H \to \mathbb{R}$, az a pont H-nak belső pontja. Az f differenciálható a-ban, ha van olyan $l(x) = \alpha_1 x_1 + \ldots + \alpha_p x_p$ lineáris függvény, amelyre*

$$\lim_{x \to a} \frac{|f(x) - [l(x-a) + f(a)]|}{|x-a|} = 0.$$

A folytonosság és a differenciálhatóság közötti kapcsolatot már egyváltozós esetben vizsgáltuk. Az 1. Tétel szerint, ha f differenciálható a-ban, akkor folytonos is abban a pontban. Azt is láttuk, hogy a folytonosság a differenciálhatóságnak szükséges de nem elégséges feltétele. A következő tétel ezt általánosítja többváltozós függvényekre.

14. Tétel. *Ha az f függvény differenciálható az a pontban, akkor f folytonos a-ban.*

Bizonyítás: Tudjuk, hogyha f differenciálható a-ban, akkor

$$\lim_{x \to a} \frac{|f(x) - [l(x-a) + f(a)]|}{|x-a|} = 0$$

Azt tudjuk, hogy a nevező 0-hoz tart. A számlálóban $l(x-a)$ is 0-hoz tart, ezért az egész számláló $f(x) - f(a) \to 0$, végül azt kapjuk, hogy $\lim_{x \to a} f(x) = f(a)$, ami a folytonosságot jelöli.

Egy tanult tétel szerint, ha f parciális deriváltjai léteznek az a pont egy környezetében és folytonosak az a pontban, akkor az f függvény differenciálható a-ban. Hogyha egy f függvény deriválható a-ban abból következik, hogy f-nek léteznek a parciális deriváltjai a-ban, és az is következik, hogy f folytonos a-ban.

Tudjuk, hogy ha f differenciálható a-ban, akkor

$$g(x) = f(a) + \alpha_1(x_1 - a_1) + \ldots + \alpha_p(x_p - a_p)$$

elsőfokú polinom jól közelíti f függvényt a-ban. Ezért feltehető, hogy f és g parciális deriváltjai az a pontban megegyeznek. Egy tanult tétel szerint, ha f differenciálható a-ban, akkor f-nek létezik az összes parciális deriváltja és ez $g(x)$ függvényben $\alpha_i = D_i f(a)$. Ebből a tételből következik, hogy f függvény a-beli érintő hipersíkja a következő:

$$g(x) = f(a) + D_1 f(a) \cdot (x_1 - a_1) + \ldots + D_p f(a) \cdot (x_p - a_p)$$

A jobboldalon szereplő $D_1 f(a) \cdot (x_1 - a_1) + \ldots + D_p f(a) \cdot (x_p - a_p)$ összeg a

$$(D_1 f(a), \ldots, D_p f(a)) \quad \text{és} \quad x - a$$

vektorok skaláris szorzata. Ez motiválta a következő definíciót.

9. Definíció. *Ha f differenciálható az $a \in \mathbb{R}^p$ pontban, akkor a*

$$(D_1 f(a), \ldots, D_p f(a))$$

vektort az f függvény a pontbeli deriváltvektorának nevezzük és $f'(a)$-val jelöljük.

Ennek a definíciónak a segítségével fel tudjuk írni az iránymenti deriváltat, ami megmutatja, hogy adott pontból kiindulva hogyan változik a függvény. Például hegymászás során a iránymenti deriváltvektorral megállapíthatjuk, hogy lejtőn vagy emelkedőn fogunk továbbhaladni. Ez utóbbit a derivált előjeléből tudjuk meghatározni, ha a derivált negatív, akkor lejtőn ereszkedünk, ha pedig pozitív, akkor emelkedőn mászunk felfelé.

10. Definíció. *Legyen $v \in \mathbb{R}^p$ egy egységvektor. A $t \to f(a+tv)$ függvény 0 pontbeli deriváltját (ha létezik) az f függvény a pontbeli v irányú iránymenti deriváltjának nevezzük, és $D_v f(a)$-val jelöljük. Más szóval,*

$$D_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}$$

feltéve, hogy a limesz létezik.

A következő tétel egyszerűsíti az iránymenti derivált felírását.

15. Tétel. *Ha az f függvény differenciálható az $a \in \mathbb{R}^p$ pontban, akkor minden $v \in \mathbb{R}^p$ vektorra a $t \to f(a+tv)$ egyváltozós függvény differenciálható a 0 pontban, és a deriváltja $\langle f'(a), v \rangle$. Speciálisan, ha $|v| = 1$, akkor a $D_v f(a)$ iránymenti derivált létezik és az értéke $D_v f(a) = \langle f'(a), v \rangle$.*

Bizonyítás: Tegyük fel, hogy f differenciálható a-ban. Tekintsük v vektort egységvektornak. Ekkor $x = a + tv$, ha $x \to a$, akkor $t \to 0$. A 8. és a 9. Definícióból tudjuk, hogy

$$\lim_{x\to a} \frac{\left|f(x) - f(a) - \langle f'(a), x-a\rangle\right|}{|x-a|} = 0$$

Ebből következik, hogy

$$0 = \lim_{t\to 0}\left|\frac{f(a+tv) - f(a) - \langle f'(a), tv\rangle}{tv}\right| =$$

mivel $|v| = 1$

$$= \lim_{t\to 0}\left|\left(\Big(\frac{f(a+tv)-f(a)}{t} - \langle f'(a), v\rangle\Big)\cdot\left|\frac{t}{|t|}\right|\right)\right|$$

$$\frac{f(a+tv)-f(a)}{t} = D_v f(a), \langle f'(a), v\rangle = konstans, \left|\frac{t}{|t|}\right| = 1$$

Vagyis

$$D_v f(a) = \langle f'(a), v\rangle$$

3.1.3. Többszörös differenciálás

Lokális szélsőérték keresésnél egyváltozós esetben a függvény második deriváltjából tudjuk meg biztosan, hogy van-e lokális szélsőértéke adott pontban. Tegyük fel, hogy egy f egyváltozós függvény többszörösen deriválható. Ha $f'(a) = 0$ és $f''(a) > 0$, akkor lokális minimuma van a pontban. Többváltozós esetben is egy függvény második deriváltját kell vizsgálni, ahhoz hogy megtudjuk biztosan, hogy a függvénynek ott van-e lokális szélsőértéke.

11. Definíció. *Legyen f értelmezve az $a \in \mathbb{R}^p$ pont egy környezetében. Ha a $D_j f$ parciális derivált létezik az a pont egy környezetében és a $D_j f$ parciálisderivált függvénynek létezik az i-edik parciális deriváltja az a pontban, akkor ezt az f függvény a-beli ij-edik másodrendű parciális deriváltjának nevezzük, és $D_{ij}f(a)$ jelöljük. (Ez csak az egyik jelölés a sok közül.)*

3. Feladat. *Számoljuk ki az $f(x, y) = x \cdot \sin(y)$ függvény első- és másodrendű parciális deriváltját!*

Megoldás: Elsőrendű parciális deriváltak: $D_1 f(x, y) = \sin(y)$, és $D_2 f(x, y) = x \cdot \cos(y)$.
Másodrendű parciális deriváltak: $D_{11} f(x, y) = 0$, $D_{21} f(x, y) = \cos(y)$, $D_{12} = \cos(y)$ és $D_{22} = -x \cdot \sin(y)$. Látjuk, hogy $D_{21} f(x, y) = D_{12} f(x, y)$.

A következő tétel megmutatja, hogy miért teljesül ez az egyenlőség.

16. Tétel. *Ha a kétváltozós $f(x, y)$ függvény $D_1 f(x, y)$ és $D_2 f(x, y)$ parciális deriváltjai léteznek az $(a, b) \in \mathbb{R}^2$ pont egy környezetében és differenciálhatóak az (a, b) pontban, akkor $D_{12} f(a, b) = D_{21} f(a, b)$. Ez az ún. Young-tétel.*

Ha f differenciálható $a \in \mathbb{R}^p$ pont egy környezetében és f parciálisderivált függvényei differenciálhatóak a pontban, akkor f kétszer differenciálható. Ha minden másodrendű parciális derivált folytonos a-ban, akkor minden elsőrendű parciális derivált differenciálható a-ban. Ebből következik, hogy minden elsőrendű parciális derivált folytonos a-ban. Ebből pedig az következik, hogy f differenciálható a-ban.

A k-adrendű parciális deriváltakat k szerinti indukcióval határozzuk meg. Ha már k-ra meghatároztuk, akkor ebből számoljuk ki a $k+1$-edrendűt, és így tovább. A k-szor differenciálhatóságot is hasonlóképpen vezetjük le. Ha f k-szor differenciálható a pont egy környezetében és f mindegyik k-adrendű parciális deriváltja differenciálható a-ban, akkor f $k+1$-szer is differenciálható.

3.2. $\mathbb{R}^p \to \mathbb{R}^q$ függvények

3.2.1. Differenciálhatóság

Ahhoz, hogy az $\mathbb{R}^q$ értékű függvények differenciálhatóságát megismerjük, először át kell tekintenünk a lineáris algebra néhány alapfogalmát. A $H : \mathbb{R}^p \to \mathbb{R}^q$ függvényt lineáris leképezésnek nevezzük, ha $A(x + y) = A(x) + A(y)$ és $A(\lambda x) = \lambda A(x)$ teljesül minden $x, y \in \mathbb{R}^p$ és $\lambda \in \mathbb{R}$ esetén. Az $A : \mathbb{R}^p \to \mathbb{R}^q$ leképezés akkor és csak akkor lineáris, ha minden koordinátafüggvénye [2] lineáris függvény.

Legyen az A leképezés i-edik koordinátafüggvénye az $a_{i1}x_1 + \ldots + a_{ip}x_p$ lineáris függvény minden $i = 1, \ldots, q$-ra. Ekkor az

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{q1} & a_{q2} & \ldots & a_{qp} \end{pmatrix}$$

[2] Legyen $f : \mathbb{R}^p \to \mathbb{R}^q$, és minden $x \in \mathbb{R}^p$ esetén jelöljük $f(x) \in \mathbb{R}^q$ koordinátáit a következőképpen $f(x) = f_1(x), \ldots, f_q(x)$. Ekkor $i = 1, \ldots, q$-ra $f_i = \mathbb{R}^p \to \mathbb{R}$, neve f i-edik koordinátafüggvénye.

mátrixot az A leképezés mátrixának nevezzük. A mátrixnak q sora, és p oszlopa van, és az i-edik sorban az A leképezés i-edik koordinátafüggvényében szereplő együtthatók állnak.

A rövid kitekintés után megfogalmazzuk a következő definícióban, hogy mikor folytonos és differenciálható egy $f : \mathbb{R}^p \to \mathbb{R}^q$ függvény.

12. Definíció. *Az $f : \mathbb{R}^p \to \mathbb{R}^q$ függvény folytonos/differenciálható $a \in \mathbb{R}^p$ pontban, ha f minden f_i koordináta függvénye folytonos/differenciálható a-ban.*

A 8. Definícióban megismertük, hogy mikor differenciálható f egy adott pontban. Ez $\mathbb{R}^p \to \mathbb{R}^q$ esetben úgy módosul, hogy l lineáris függvény helyett $\mathbb{R}^p \to \mathbb{R}^q$ lineáris transzformáció szerepel. Vagyis

$$\lim_{x \to a} \frac{\big|f(x) - f(a) - A(x-a)\big|}{|x-a|} = 0$$

Az előző fejezetben az érintő hipersík segítségével megismertük az $f'(a)$ gradiens vektort. Ebben a fejezetben pedig $f'(a)$ az A mátrix Jacobi-mátrixával egyezik meg.

13. Definíció. *Legyen $f : H \to \mathbb{R}^q$ ahol $H \subset \mathbb{R}^p$, és legyen f differenciálható az $a \in intH$ pontban. Az $A : \mathbb{R}^p \to \mathbb{R}^q$ lineáris leképezést az f függvény a pontbeli deriváltjának nevezzük és $f'(a)$-val jelöljük. Az $f'(a)$ lineáris leképezés mátrixát, tehát a $D_j f_i(a) (j = 1, \ldots, p, i = 1, \ldots, q)$ parciális deriváltakból álló*

$$\begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \ldots & D_p f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \ldots & D_p f_2(a) \\ \vdots & \vdots & & \vdots \\ D_1 f_q(a) & D_2 f_q(a) & \ldots & D_p f_q(a) \end{pmatrix}$$

mátrixot az f függvény a pontbeli Jacobi-mátrixának nevezzük.

A következő tétel az $f : \mathbb{R}^p \to \mathbb{R}^q$ függvény differenciálhatóságáról szóló állításokat foglalja össze.

17. Tétel. *(i) Ha az f függvény differenciálható az a pontban, akkor az f folytonos a-ban, továbbá f mindegyik koordinátafüggvényének mindegyik változó szerinti parciális deriváltja létezik és véges az a pontban.*

(ii) Ha f mindegyik koordinátafüggvényének mindegyik változó szerinti parciális deriváltja létezik és véges az a pont egy környezetében és folytonos az a pontban, akkor f differenciálható az a pontban.

3.2.2. Differenciálási szabályok

18. Tétel. *Ha az $\mathbb{R}^q$-ba képező függvények differenciálhatóak az $a \in \mathbb{R}^p$ pontban, akkor az $f+g$ és λf függvények is differenciálhatóak a-ban, továbbá $(f+g)'(a) = f'(a) + g'(a)$ és $(\lambda f)'(a) = \lambda f'(a)$ minden $\lambda \in \mathbb{R}$-re.*

A következő tétel a leképezések kompozíciójának differenciálhatóságára és deriváltjára vonatkozik.

19. Tétel. *Tegyük fel, hogy*
(i) $H \subset \mathbb{R}^p$, $g : H \to \mathbb{R}^q$ és g differenciálható az $a \in intH$ pontban;
(ii) $g(a) \in intE \subset \mathbb{R}^s$ és f differenciálható a $g(a)$ pontban.
Ekkor $(f \circ g)'(a) = f'(g(a)) \circ g'(a)$.

1. Következmény. ***(az összetett függvény differenciálási szabálya).*** *Tegyük fel, hogy a valós értékű f függvény differenciálható a $b = (b_1, \ldots, b_q) \in \mathbb{R}^q$ pontban, a valós értékű $g_1, \ldots, g_q$ függvények differenciálhatóak az $a \in \mathbb{R}^p$ pontban, és $g_i(a) = b_i$ minden $i = 1, \ldots, q$-ra. Ekkor az $F(x) = f(g_1(x), \ldots, g_q(x))$ függvény differenciálható az a pontban, és*

$$D_j F(a) = \sum_{i=1}^{q} D_i f(b) \cdot D_j g_i(a)$$

minden $j = 1, \ldots, p$-re.

3. Megjegyzés. *Az előbbi összefüggés könnyebben megjegyezhető a következő alakban. Jelöljük f változóit $y_1, \ldots, y_q$-val, és g_i helyett is írjunk y_i-t. Ekkor azt kapjuk, hogy*

$$\frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_j} + \frac{\partial f}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_j} + \ldots + \frac{\partial f}{\partial y_q} \cdot \frac{\partial y_q}{\partial x_j}.$$

4. Feladat. *Bizonyítsuk be illetve határozzuk meg a láncszabály következményének segítségével az $f \cdot g$ és az f/g ($\mathbb{R} \to \mathbb{R}$) deriválási szabályokat!*

Megoldás:

(a) $f(x,y) = x \cdot y$, $D_1 f(x,y) = y$ és $D_2 f(x,y) = x$

$$F(t) = f(g(t), h(t)) = g(t) \cdot h(t)$$

$$F'(t) = D_1 f(g(t), h(t)) \cdot g'(t) + D_2 f(g(t), h(t)) \cdot h'(t)$$

Ebből következik, hogy

$$(g(t) \cdot h(t))' = h(t) \cdot g'(t) + g(t) \cdot h'(t)$$

(b) $f(x,y) = \frac{x}{y}$, $D_1 f(x,y) = \frac{1}{y}$, $D_2 f(x,y) = -\frac{x}{y^2}$

$$F(t) = f(x(t), y(t)) = \frac{x(t)}{y(t)}$$

$$F'(t) = D_1 f(x(t), y(t)) \cdot x'(t) + D_2 f(x(t), y(t)) \cdot y'(t) =$$

$$\frac{1}{y(t)} \cdot x'(t) - \frac{x(t)}{y^2(t)} \cdot y'(t) = \frac{x'(t) \cdot y(t) - x(t) \cdot y'(t)}{y^2(t)}$$

3.2.3. Implicit függvények

Egy egyenlet megoldásánál az ismeretlen implicit módon van meghatározva, és kíváncsiak vagyunk, hogy ezt az ismeretlent ki tudjuk-e fejezni explicit módon, vagyis a többi paraméter segítségével. Például az $ax^2 + bx + c = 0$ másodfokú egyenlet az x ismeretlent implicite meghatározza, és a megoldás során x-et explicit módon kell kifejeznünk az a, b, c paraméterekkel. Jelöljük az a, b, c-t, x_1, x_2, x_3-al és x helyett írjunk y-t. Ekkor az $f(x_1, x_2, x_3, y) = x_1 y^2 + x_2 y + x_3$ négyváltozós függvény, és olyan $g(x_1, x_2, x_3)$ függvényt keresünk, amelyre

$$f(x_1, x_2, x_3, g(x_1, x_2, x_3)) = 0$$

teljesül; ekkor azt mondjuk, hogy az $y = g(x_1, x_2, x_3)$ függvény az előbbi egyenlet megoldása. Tudjuk, hogy a $H = \{(x_1, x_2, x_3) : x_2^2 - 4x_1x_3 < 0\} \subset \mathbb{R}^3$ halmazon nincs megoldás, a $K = \{(x_1, x_2, x_3) : x_1 \neq 0, x_2^2 - 4x_1x_3 \geq 0\} \subset \mathbb{R}^3$ halmazon pedig több folytonos megoldása is van, hiszen a

$$g_{1,2} = \frac{-x_2 \pm \sqrt{x_2^2 - 4x_1x_3}}{2x_1}$$

függvények mindegyike folytonos megoldás K halmazon.

Nem minden esetben tudjuk az y megoldást képlettel vagy zárt formulával felírni a paraméterek segítségével. Van olyan eset, amikor már az $f(x,y)$ függvény sincs ilyen módon megadva. De feltéve, hogy $f(x,y)$-t valamilyen képlet adja meg, még akkor sem biztos, hogy az y megoldás ugyanabba a függvénycsaládba tartozik, amelynek elemeivel f-et kifejeztük. Például az $f(x,y) = x - y^3$ függvény polinom, de az $f(x,y) = 0$ egyenlet megoldása nem az. Tehát csak az a feladatunk, hogy megállapítsuk, hogy létezik-e y függvény, és ha létezik, akkor írjuk le a tulajdonságait. A következő tétel ezt írja le.

20. Tétel. *Tegyük fel, hogy a kétváltozós valós értékű f függvény eltűnik az $(a,b) \in \mathbb{R}^2$ pontban, és folytonos az $[a-\eta, a+\eta] \times [b-\eta, b+\eta]$ téglalapon egy alkalmas $\eta > 0$-ra. Ha az f_x szekciófüggvény szigorúan monoton minden $x \in [a-\eta, a+\eta]$-re, akkor létezik egy δ pozitív szám, úgy, hogy* *(i) minden $x \in (a-\eta, a+\eta)$-hoz létezik egyetlen $g(x) \in (b-\eta, b+\eta)$ szám, amelyre $f(x, g(x)) = 0$, továbbá*

(ii) az így definiált g függvény folytonos az $(a-\delta, a+\delta)$ intervallumban.

A következő tétel következménye az előbbi tételnek.

2. Következmény. ***(egyváltozós implicitfüggvény-tétel).***
Tegyük fel, hogy a kétváltozós valós értékű f függvény eltűnik az $(a, b) \in \mathbb{R}^2$ pontban, és folytonos (a, b) egy környezetében. Tegyük fel továbbá, hogy a $D_2 f$ parciális derivált létezik, véges és nullától különböző az (a, b) pont egy környezetében. Ekkor léteznek olyan δ és η pozitív számok, hogy

(i) minden $x \in (a-\delta, a+\delta)$-hoz létezik egyetlen $\varphi(x) \in (b-\eta, b+\eta)$ szám, amelyre $f(x, g(x)) = 0$, továbbá

(ii) az így definiált g függvény folytonos az $(a-\delta, a+\delta)$ intervallumban.

Egyváltozós esetben a $g'(a)$ deriváltat könnyen megkaphatjuk az összetett függvény deriválási szabályából. Ugyanis $f(x, g(x)) = 0$ az a pont egy környezetében, ezért itt a deriváltja is nulla. Így

$$D_f(a, b) \cdot 1 + D_2 f(a, b) \cdot g'(a) = 0,$$

amiből

$$g'(a) = \frac{-D_1 f(a, b)}{D_2 f(a, b)}$$

5. Feladat. *Tekintsük az $f(x, y) = x^2 + y^2 - 1$ függvényt. Fejezzük ki y-t a többi paraméter segítségével!*

Megoldás. Az egyváltozós implicitfüggvény-tétel feltételeit kell ellenőriznünk. Az f függvény mindenütt folytonos és mindenütt (akárhányszor) differenciálható. Ha $a^2 + b^2 = 1$ és $-1 < a < 1$, akkor $D_2 f(a, b) = 2b \neq 0$. Vagyis teljesülnek a tétel feltételei. Ennek megfelelően van olyan, az a pont egy környezetében folytonos g függvény, amelyre $g(a) = b$ és $x^2 + g(x)^2 - 1 = 0$. Tehát, ha $b > 0$, akkor ilyen a $g(x) = \sqrt{1 - x^2}$ függvény a $(-1, 1)$ intervallumban, ha pedig $b < 0$, akkor a $g(x) = -\sqrt{1 - x^2}$ függvény elégíti ki a feltételeket $(-1, 1)$-ben.

Ha viszont $a = 1$, akkor a semmilyen környezetében nem létezik ilyen függvény, hiszen $x > 1$ esetén $x^2 + y^2 - 1 > 0$ minden y-ra. Viszont itt az egyváltozós implicitfüggvény-tétel feltételei nem is teljesülnek, hiszen $a = 1$ esetén $b = 0$ és így $D_2 f(a, b) = 0$. Ha $a = -1$, akkor ugyanez teljesül.

3.2.4. Feltételes szélsőérték

Egy többváltozós függvénynek a szélsőértékeit csak azon pontok között keressük, amelyek egy adott egyenletet vagy egyenleteket kielégítenek. Ezeket az egyenleteket feltételeknek hívjuk.

14. Definíció. *Legyen $a \in H \subset \mathbb{R}^p$, $F : \mathbb{R}^p \to \mathbb{R}^q$, és legyen $F(a) = 0$. Tegyük fel, hogy a p-változós és valós értékű f függvény értelmezve van az a pont egy környezetében, és létezik egy $\delta > 0$ szám úgy, hogy $f(x) \leq f(a)$ minden olyan $x \in B(a, \delta)$ pontra, amelyre $F(x) = 0$. Ekkor azt mondjuk, hogy az f függvénynek feltételes lokális maximuma van az a pontban az $F = 0$ feltétel mellett. Hasonlóan definiáljuk a feltételes lokális minimumot. Ha f-nek feltételes lokális maximuma vagy minimuma van az a pontban az $F = 0$ feltétel mellett, akkor azt mondjuk, hogy f-nek feltételes lokális szélsőértéke van az a pontban az $F = 0$ feltétel mellett.*

Akkor tudunk feltételes szélsőértéket számolni, hogyha az f függvény folytonos, és az $x \in H$, $F_1(x) = \ldots, F_k(x) = 0$-nak eleget tevő pontok halmaza korlátos és zárt. Mert ekkor Weierstrass tétele szerint ezen a halmazon f felveszi a minimumát és a maximumát, ami definíció szerint a feltételes minimum/maximum. A következő ún. Lagrange-féle multiplikátor-módszer segít nekünk abban, hogy megtaláljuk egy többváltozós függvény feltételes szélsőértékét.

21. Tétel. ***(Lagrange-féle multiplikátor-módszer)***
Tegyük fel, hogy $H \subset \mathbb{R}^p$, $a \in intH$, és $f, F_1, \ldots, F_q : H \to \mathbb{R}$ függvények parciális deriváltjai folytonosak a-ban. Tegyük fel, hogy f-nek feltételes lokális minimuma vagy maximuma van a-ban, akkor vannak olyan $\mu, \lambda_1, \ldots, \lambda_q \in \mathbb{R}$ számok, melyek nem mind 0-k, és amelyekre a

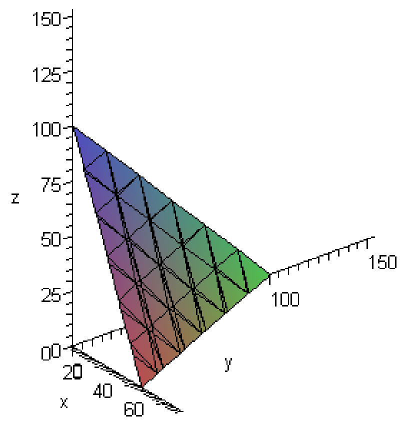
$$\mu \cdot f + \lambda_1 \cdot F_1 + \ldots + \lambda_q \cdot F_q$$

függvény minden parciális deriváltja a-ban 0.

A következő feladat egy egyszerű, de annál hasznosabb alkalmazása a Lagrange-multiplikátor módszernek. A [3] hivatkozásban szereplő könyvből származnak a következő feladatok:

6. Feladat. *Az $f(x, y, z) = 0,05x^2 + 0,16xy + 0,25z^2$ függvénnyel modellezhetjük kedvenc fagylaltunk grammban mért mennyiségét. Jelölje x az egy grammban lévő zsír mennyiségét, y a szénhidrát mennyiségét, z pedig a fehérje mennyiségét. Tegyük fel, hogy egy gramm zsírban 9 kalória, egy gramm szénhidrátban 4 kalória, egy gramm proteinben pedig ugyancsak négy kalória található. Számoljuk ki, hogy hány gramm fagyit ehetünk meg anélkül, hogy 400 kalóriánál többet vinnénk be a szervezetünkbe.*

Megoldás: Az $f(x, y, z) = 0,05x^2 + 0,16xy + 0,25z^2$ függvény folytonos, mivel polinom. Az $F(x, y, z) = 9x + 4y + 4z - 400 = 0$ feltététel egy sík felszínét írja le, amit a következő ábra mutat:



A feltételt leíró halmaz korlátos és zárt, vagyis ezen a halmazon f felveszi a minimumát és a maximumát. Az előbbi tétel szerint a következőknek kell teljesülnie:

$$9x + 4y + 4z - 400 = 0$$

$$\mu D_i f(x, y, z) + \lambda D_i F(x, y, z) = 0$$

Továbbá feltehető, hogy $\mu^2 + \lambda^2 = 1$.
Tehát

$$\mu \cdot (0,1x + 0,16y) + \lambda \cdot 9 = 0$$

$$\mu \cdot 0,16x + \lambda \cdot 4 = 0$$

$$\mu \cdot 0,5z + \lambda \cdot 4 = 0$$

Végül $\mu \cdot 0,4x + \mu \cdot 0,64y = \mu \cdot 1,44x = \mu \cdot 4,5z = 0$ adódik.

(i) Ha $\mu = 0$ akkor $\lambda \cdot 9 = \lambda \cdot 4 = \lambda \cdot 4 = 0$. Ez ellentmondás, mert semelyik λ-ra sem lesz ez a három érték egyenlő.

(ii) Ha $\mu \neq 0$, akkor

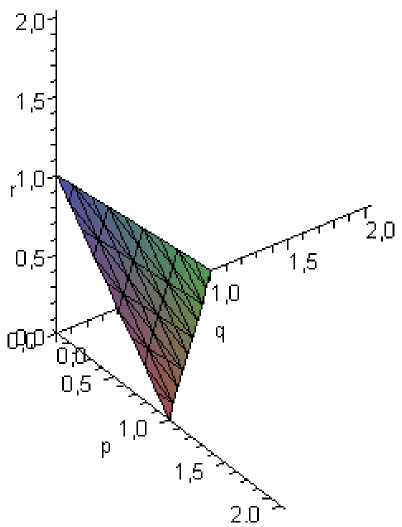
$$0,4x + 0,64y = 1,44x = 4,5z$$

$$9x + 4y + 4z - 400 = 0$$

Az első egyenletből adódik: $x = 3,125z$ és $y = 1,625x$, vagyis $y = 5,078z$. A második egyenletbe behelyettesítve kapjuk: $z = 7,628$, $x = 23,8375$, $y = 38,736$. A kapott értékeket behelyettesítve az f függvénybe: $f(x, y, z) = 190,697$ értéket kapjuk. Ez lesz a feltételes maximum értéke. Vagyis maximum 190 gramm fagylaltot fogyaszthatunk el úgy, hogy 400 kalóriánál kevesebbet viszünk be a szervezetünkbe.

7. Feladat. *Az emberi vércsoportnak genetikailag három allélja van A, B és 0. Vércsoport szempontjából két típust különböztetünk meg: homozigóták, akiknek AA, BB, vagy 00 a vércsoportjuk, a másik csoport: heterozigóták, akiknek AB, A0, B0 a vércsoportjuk. A Hardy-Weinberg szabály szerint a heterozigóták P valószínűségét a* $P(p, q, r) = 2pq + 2pr + 2qr$ *függvény modellezi, ahol p jelöli az A allélt a populációban, q reprezentálja a B allélt, és r pedig a 0 allélt. Mivel valószínűségi eloszlásról van szó, ezért* $p+q+r = 1$ *a feltétel, amin vizsgáljuk a P függvény feltételes valószínűségét. Egy populáción belül a heterozigóta egyedek maximális arányát keressük.*

Megoldás: A P függvény folytonos, mivel polinom. A feltételt leíró halmaz korlátos és zárt.



Vagyis teljesülnek a feltételek, ezért a P függvénynek lesz minimuma vagy maximuma. Az előző feladat mintájára számolunk. A következő egyenleteknek kell teljesülnie:

$$\mu \cdot (2q + 2r) + \lambda \cdot 1 = 0$$

$$\mu \cdot (2p + 2r) + \lambda \cdot 1 = 0$$

$$\mu \cdot (2p + 2q) + \lambda \cdot 1 = 0$$

$$\mu^2 + \lambda^2 = 1$$

Ebből adódik

$$\mu \cdot (2q + 2r) = \mu \cdot (2p + 2r) = \mu \cdot (2p + 2q)$$

(i) Ha $\mu = 0$, akkor $\lambda = 0$, ami ellenmondás.

(ii) Ha $\mu \neq 0$ akkor $2q+2r = 2p+2r = 2p+2q$ amiből következik, hogy $p = q = r$. A feltételbe behelyettesítve kapjuk: $p = 1/3$, $q = 1/3$, $r = 1/3$. Vagyis egy populáción belül a heterozigóta egyedek maximális aránya $2/3$.

4. Normált térbeli differenciálszámítás

Az egyváltozós és többváltozós differenciálszámítás után rátérünk a normált térbeli differenciálszámításra. Azzal, hogy normált térben vizsgáljuk a differenciálszámítást, általánosítjuk azokat a definíciókat és tételeket, amelyeket már egyváltozós és többváltozós esetben megismertünk az előbbi fejezetek során. Elevenítsük fel, hogy mi az a normált tér:

15. Definíció. *Legyen X valós vektortér.*
X-beli normán olyan $\|\cdot\| : X \to \mathbb{R}$ függvényt értünk, amelyik tetszőleges $x, y \in X$ vektorokra és λ valós számra eleget tesz az alábbi követelményeknek:
(a) $\|x\| \geq 0$,
(b) $\|x\| = 0 \iff x = 0$
(c) $\|\lambda x\| = |\lambda| \cdot \|x\|$,
(d) $\|x + y\| \leq \|x\| + \|y\|$.

4.1. Definíciók és elemi tulajdonságok

Az $f : \mathbb{R} \to \mathbb{R}$ függvény

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

deriváltjának a definíciója (1.Definíció) könnyen általánosítható vektorértékű függvényekre. Vektorváltozós $f : X \to \mathbb{R}$ függvényekre azonban $\dim X > 1$ esetén a fenti tört értelmét veszti. Az $f : \mathbb{R} \to \mathbb{R}$ differenciálhatóságára adott következő két ekvivalens definíció azonban már általánosítható lesz:

(a)

$$\frac{|f(a + h) - f(a) - Ah|}{|h|} \to 0 \quad ha \quad h \to 0$$

(b)

$$f(x) - f(a) \equiv u(x)(x - a),$$

ahol $u(x)$ folytonos függvény.

Az (a) Definíció az 1. Definíció általánosítása, vagyis:

$$\frac{|f(a + h) - f(a) - Ah|}{|h|} \to 0 \quad ha \quad h \to 0$$

a h helyére $(x - a)$ helyettesítve kapjuk

$$\frac{|f(a + x - a) - f(a) - A(x - a)|}{|x - a|} \to 0$$

végül a következő relációhoz jutunk

$$\frac{|f(x)-f(a)|}{|x-a|} \to A \quad ha \quad x \to a, \tag{1}$$

ami egyváltozós esetben is a differenciálszámítás definíciója. Itt az A lineáris operátor.

16. Definíció. *Legyen $f : X \to Y$*[3]*. Az f deriválható $a \in D(f)$-ben, ha értelmezve van az a egy környezetében, és az (1) relációval ekvivalens tulajdonságok valamelyike teljesül. Az A leképezést az f függvény a pontbeli (totális) deriváltjának vagy Fréchet-deriváltjának nevezzük, és $f'(a)$-val jelöljük. Vagyis*

$$\lim_{h\to 0} \frac{\|f(a+h)-f(a)-Ah\|}{\|h\|} = 0.$$

Az (1) reláció úgy is fogalmazható, hogy az

$$x \to \frac{f(x)-f(a)}{x-a}$$

függvény folytonosan kiterjeszthető az a pontra. Vagyis ekvivalens egy, az a pontban folytonos és az

$$f(x)-f(a) \equiv u(x)(x-a) \tag{2}$$

azonosságnak eleget tevő $u : D \to \mathbb{R}$ függvény létezésével. Valósértékű függvényeknél ezt neveztük egy függvény határoló síkjának: $g(x) = f(a)+\sum_{i=1}^{n} D_i f(a)(x_i-a_i)$. Vektorértékű esetben pedig definiáltuk egy f függvény a pontbeli Jacobi-mátrixát. Mindkét esetben ekvivalens definíciót kaptunk az előbb megismert azonossággal.

A következő lemma leírja, hogy egy leképezés mikor folytonos lineáris leképezés.

1. Lemma. *Az $A : X \to Y$ lineáris leképezés pontosan akkor folytonos 0-ban, ha létezik olyan $M \geq 0$ konstans, hogy*

$$\|Ax\| \leq M\|x\| \qquad minden \quad x \in X\text{-re.}$$

Ekkor A Lipschitz-folytonos is az M Lipschitz-konstanssal.

Jelöljük $L(X,Y)$-nal a folytonos lineáris $A : X \to Y$ leképezések halmazát. Egyszerűen látható, hogy ha $A \in L(X,Y)$, akkor az előző lemmában lévő egyenlőtlenségnek eleget tevő $M \geq 0$ konstansok alulról zárt intervallumot alkotnak. A legkisebb ilyen konstanst $\|A\|$-val jelölve minden $x \in X$-re fennáll az

$$\|Ax\| \leq \|A\| \cdot \|x\|$$

[3]Az X és Y normált tereket jelölnek ebben a fejezetben.

egyenlőtlenség. Az érdektelen nulla-dimenziós terek kivételével érvényes az

$$\|A\| = \sup_{x \in X} \frac{\|Ax\|}{\|x\|} \qquad \text{és} \quad x \neq 0$$

egyenlőség is.

2. Állítás. $(L(X,Y), \|\cdot\|)$ *normált tér.*

Bizonyítás: Világos, hogy az $L(X,Y)$ a szokásos pontonkénti műveletekre nézve vektortér. Ami nem-triviális normatulajdonság az a háromszög-egyenlőtlenség. Mivel az Y-beli háromszög-egyenlőtlenség alapján minden $x \in X$-re teljesül a

$$\|(A+B)x\| \leq \|Ax\| + \|Bx\| \leq \|A\| \cdot \|x\| + \|B\| \cdot \|x\| = (\|A\| + \|B\|) \cdot \|x\|$$

becslés, $\|A+B\| \leq \|A\| + \|B\|$ baloldali norma definíciója miatt.

Két fontos definíciót kell felelevenítenünk:

17. Definíció. *A valós értékű lineáris leképezéseket lineáris funkcionáloknak is nevezzük.*

18. Definíció. *Az $L(X,\mathbb{R})$ Banach-teret az X normált tér duális terének hívjuk, és általában X'-vel jelöljük.*

A következő állításra szükségünk lesz a a 4. Állítás bizonyítása során.

3. Állítás. *Legyen $(X, \|\cdot\|)$ normált tér.*
(a) Bármely $c \in X$ vektorhoz van olyan $\varphi \in X'$ lineáris funkcionál, hogy $\|\varphi\| = \|c\|$ és $\varphi(c) = \|c\|^2$.
(b) Bármely két különböző $x_1, x_2 \in X$ vektorhoz van olyan $\varphi \in X'$ lineáris funkcionál, hogy $\varphi(x_1) \neq \varphi(x_2)$.

Ezt az állítást most nem bizonyítjuk be, de megtalálható a [4] hivatkozásban szereplő könyv 82. oldalán.

A következő állítás megfogalmazza, hogy a (2) és a (3) kifejezések ekvivalensek.

4. Állítás. *Legyen $f : X \to Y$ és $a \in D(f)$. A következő két tulajdonság ekvivalens:*
(i) van olyan $A \in L(X,Y)$ folytonos lineáris leképezés, amelyre

$$\frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} \to 0 \quad \|h\| \to 0 \tag{3}$$

(ii) van olyan a-ban folytonos $u : D(f) \to L(X,Y)$ függvény, amely eleget tesz a (2) azonosságnak.

Az ekvivalencia akkor is fennáll, ha a (ii)-beli u függvény csak valamely $D(f) \cap B_r(a)$ halmazon van értelmezve.

Továbbá (i) maga után vonja olyan (ii)-beli u függvény létezését is, amelyre $u(a) = A$, és (ii)-ből következik (i) az $A = u(a)$ választással.

Bizonyítás:

(ii) $\Rightarrow$ *(i)*

$$\frac{\|f(a+h) - f(a) - u(a)h\|}{\|h\|} = \frac{\|(u(a+h) - u(a))h\|}{\|h\|} \leq \|u(a+h) - u(a)\|$$

összefüggésből (3) adódik az $A := u(a)$ választással, mert a jobboldali kifejezés u a-beli folytonossága miatt $h \to 0$ esetén nullához tart.

(i) $\Rightarrow$ *(ii)*.

Legyen $\psi \in X'$ funkcionál, amely függ x-től, és minden rögzített $a \neq x \in D(f)$ esetén alkalmazzuk a 3. Állítást $c = x - a$-val, majd osszuk el a kapott funkcionált $\|c\|^2$-tel:

$$\|\varphi\| = \|x - a\| \quad \text{és} \quad \varphi(x-a) = \|x-a\|^2.$$

Ekkor

$$\psi(x-a) = \frac{\varphi(x-a)}{\|x-a\|^2} = 1$$

és

$$\|\psi\| = \frac{\|\varphi\|}{\|x-a\|^2} = \frac{1}{\|x-a\|}$$

adódik.

Akkor az

$$u(x)h := Ah + \psi(h)(f(x) - f(a) - A(x-a)), \quad h \in X$$

képlet olyan $u(x) \in L(X,Y)$ folytonos lineáris leképezést értelmez, amely eleget tesz (2)-nek:

$$u(x)(x-a) := A(x-a) + 1 \cdot (f(x) - f(a) - A(x-a))$$

Be kell még látni, hogy $u(a) := A$ választással az így értelmezett $u : D(f) \to L(X,Y)$ függvény folytonos a-ban. Minden $h \in X$-re fennáll az

$$\|(u(x)-u(a))h\| = \|\psi(h)(f(x)-f(a)-A(x-a))\| \leq \|\psi\| \cdot \|h\| \cdot \|f(x)-f(a)-A(x-a)\|$$

becslés; innen $\|\psi\| = 1/\|x-a\|$ miatt

$$\|u(x) - u(a)\| \leq \frac{\|f(x) - f(a) - A(x-a)\|}{\|x-a\|}$$

Ha $x \to a$, akkor (i) alapján a tört nullához tart, tehát $u(x) \to u(a)$.

A következő definíciókat és állításokat már egy- és többváltozós esetben is definiáltuk, de most általánosabb formában fogalmazzuk meg. Legyen $f : X \to Y$.

19. Definíció. *Jelöljük D_1-gyel azon pontok halmazát, ahol f differenciálható; az $a \to f'(a)$ képlettel értelmezett $f' : D_1 \to L(X,Y)$ függvényt f derivált függvényének hívjuk.*

20. Definíció. *Ha $D_1 = D(f)$, f differenciálható, vagyis ha nyílt halmazon van értelmezve, és az értelmezési tartománya minden pontjában differenciálható.*

21. Definíció. *f folytonosan differenciálható vagy C^1 osztálybeli, ha differenciálható, és ha $f' : D(f) \to L(X,Y)$ (mindenütt) folytonos. Ezt röviden az $f \in C^1$ jelöléssel fejezzük ki.*

5. Állítás. *(a) Ha f differenciálható a-ban, akkor folytonos is a-ban.*
(b) Ha f differenciálható a-ban, akkor

$$\lim_{t \to 0} \frac{f(a+th) - f(a)}{t} = f'(a)h \tag{4}$$

minden $h \in X$-ra. Következésképpen az $f'(a)$ derivált egyértelmű.
(c) A differenciálás lineáris operáció: ha $f, g : X \to Y$ differenciálhatóak a-ban és $\alpha, \beta \in \mathbb{R}$, akkor $\alpha f + \beta g : X \to Y$ is differenciálható a-ban, és

$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a).$$

(d) Ha $f : X \to \mathbb{R}$ és $g : X \to Y$ differenciálhatóak a-ban, akkor $fg : X \to Y$ is differenciálható a-ban, és

$$(fg)'(a)h = (f'(a)h)g(a) + f(a)g'(a)h$$

minden $h \in X$-re.
(e) Ha $g : X \to Y$ differenciálható a-ban és $f : Y \to Z$ differenciálható $g(a)$-ban, és

$$(f \circ g)'(a) = f'(g(a))g'(a).$$

Bizonyítás:

(a) A (2) azonosságból következik, hogy

$$\|f(x) - f(a)\| \le \|u(x)\| \cdot \|x - a\|.$$

Ha $x_n \to a$, akkor $u(x_n) \to u(a)$, tehát $\|u(x_n)\| \cdot \|x_n - a\| \to 0$, és így $f(x_n) \to f(a)$ az előző egyenlőtlenség alapján.

(b) Átalakítva a (4)-t kapjuk: $f(a+th) - f(a) = t \cdot f'(a)h$. Az $u(a) = f'(a)$ egyenlőséget felhasználva kapjuk: $t(f(x) - f(a)) = t(u(a)(x-a))$, ami a (2) azonosság t-szerese.

(c) Léteznek olyan a-ban folytonos $u, v : D \to L(X, Y)$ függvények, hogy

$$f(x) - f(a) \equiv u(x)(x-a), \qquad g(x) - g(a) \equiv v(x)(x-a).$$

Ekkor

$$(\alpha f + \beta g)(x) - (\alpha f + \beta g)(a) \equiv (\alpha u(x) + \beta v(x))(x-a)$$

is teljesül. Mivel $\alpha u + \beta v : D \to L(X, Y)$ folytonos a-ban, és mivel $\alpha f + \beta g$ differenciálható a-ban, ezért teljesül

$$(\alpha f + \beta g)'(a) = (\alpha u + \beta v)(a) = \alpha u(a) + \beta v(a) = \alpha f'(a) + \beta g'(a).$$

(d) Bizonyítás megtalálható a [4] hivatkozásban szereplő könyv 92. oldalán.

(e) állítás bizonyítása ugyancsak megtalálható a fent megnevezett helyen.

4. Megjegyzés. *A (4)-beli határértéket az f függvény h irányú deriváltjának nevezzük a pontban (10. Definíció).*

Egyváltozós esetben is láttuk, hogy a derivált fogalma nagyon hasznos a függvények szélsőértékeinek a keresése során. Már egyváltozós és többváltozós esetben is megnéztük, hogy egy adott függvénynek mikor lesz lokális maximuma/minimuma. Normált térben hasonlóan definiálható a lokális szélsőérték.

22. Definíció. *Az $f : X \to \mathbb{R}$ függvénynek lokális minimuma van az $a \in D(f)$ pontban, ha van olyan $r > 0$, hogy $f(x) \geq f(a)$ minden $x \in D(f) \cap B_r(a)$-ra.*

A (3) relációt gyakran a kényelmesebb

$$f(a+h) = f(a) + Ah + o(h), \qquad h \to 0$$

formába írjuk.

6. Állítás. *Ha f-nek lokális minimuma van a-ban és differenciálható is a-ban, akkor $f'(a) = 0$.*

Bizonyítás: Rögzítsünk egy tetszőleges $h \in X$ vektort. A feltevéseink alapján $t \to 0$ esetén

$$f(a) \leq f(a+th) = f(a) + tf'(a)h + o(t)$$

Innen $t > 0$-val osztva

$$0 \leq f'(a)h + o(1)$$

majd $t \searrow 0$ esetén $f'(a)h \geq 0$ adódik. Hasonlóan, $t \nearrow 0$ esetén a fordított $f'(a)h \leq 0$ egyenlőtlenséget kapjuk, tehát $f'(a)h = 0$ minden $h \in X$ esetén.

4.2. Középértéktételek

Az 1.4. alfejezetben már megismertük a Lagrange-középértéktételt. Látni fogjuk, hogy ez a tétel a vektorértékű függvényekre már érvényét veszti. Ezért egy gyengébb állítást fogunk megfogalmazni ezekre a függvényekre.

7. Állítás. *(Lagrange) Ha $f : X \to \mathbb{R}$ az $[a, b] \subset X$ szakasz minden pontjában differenciálható, akkor van olyan $c \in (a, b)$ pont, hogy*

$$f(b) - f(a) = f'(c)(b - a).$$

5. Megjegyzés. *Vektorértékű függvényekre az állítás nem teljesül. Például az $f(x) := (\cos x, \sin x)$ függvényre*

$$|f'(x)| = |-\sin x, \cos x)| = 1$$

minden $x \in \mathbb{R}$ pontban. Mivel $f(2\pi)-f(0) = 0$, az $f(2\pi)-f(0) = 2\pi f'(c)$ egyenlőség semmilyen c-re nem teljesül.

Vektorértékű függvényekre a következő gyengébb eredmények igazolhatók.

22. Tétel. *Ha $f : X \to Y$ egy $[a, b] \subset X$ szakasz minden pontjában differenciálható, akkor vannak olyan $c_1, c_2 \in (a, b)$ pontok, hogy*

$$\|f(b) - f(a)\| \leq \|f'(c_1)\| \cdot \|b - a\| \tag{5}$$

és

$$\|f(b) - f(a) - f'(a)(b - a)\| \leq \|f'(c_2) - f'(a)\| \cdot \|b - a\| \tag{6}$$

Bizonyítás: A 3. Állítás szerint létezik olyan $\varphi \in Y'$ lineáris funkcionál, hogy

$$\|\varphi\| \leq 1 \qquad\qquad \varphi(f(b) - f(a)) = \|f(b) - f(a)\|.$$

Alkalmazzuk a Lagrange-középértéktételt a $\varphi \circ f : X \to \mathbb{R}$ függvényre. Az állítás szerint van olyan $c_1 \in (a, b)$ pont, hogy

$$\|f(b) - f(a)\| = \varphi(f(b) - f(a)) = \varphi(f(b)) - \varphi(f(a)) = \varphi(f'(c_1)(b - a)).$$

Mivel

$$\varphi(f'(x)(b - a)) \leq \|\varphi\| \cdot \|f'(x)(b - a)\| \leq \|f'(x)\| \cdot \|b - a\|$$

vagyis (5) teljesül.

A (6) egyenlőtlenséget úgy kapjuk, ha f helyett $x \to f(x) - f'(a)x$ függvényre alkalmazzuk (5)-t.

4.3. Gâteaux-differenciál

Tudjuk, hogy a (4)-beli határértéket az f függvény h irányú deriváltjának nevezzük az a pontban. Megfogalmazhatjuk a következő definíciót:

23. Definíció. *Ha $h \mapsto f'(a)h$ folytonos lineáris leképezés, amely az egész X-en értelmezve van, akkor f-et az a pontban gyengén differenciálhatónak vagy G-differenciálhatónak nevezzük, ezt a leképezést $f'(a)$-el jelöljük, és az f függvény a pontbeli gyenge differenciáljának vagy Gâteaux-differenciáljának nevezzük.*

A gyenge differenciál egyértelmű, mert az iránymenti deriváltak egyértelműek. Az előbbi definícióból következik, hogy ha f differenciálható a-ban, és $A : X \to Y$, akkor

$$\lim_{t \to 0} \frac{f(a+th) - f(a)}{t} = Ah$$

minden $h \in X$-re , így A egyértelműen meghatározott, f gyengén differenciálható a-ban, és $A = f'(a)$. Ekkor $f'(a)$-t az f függvény a-beli differenciáljának nevezzük.

Ha $X = \mathbb{R}$, akkor az f függvény a-beli differenciálhatósága ekvivalens

$$c = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

határértékkel. Ha $h = x - a$, akkor az 1. Definícióhoz jutunk.

Ha $X =$ Hilbert tér és $Y = \mathbb{R}$, akkor van olyan

$$\nabla f(x) \in X,$$

az f gradiense, amelyre

$$f'(a)h = \langle h, \nabla f(a) \rangle \quad \text{minden} \quad h \in X - re.$$

A 15. Tételhez jutunk, ha $|h| = 1$. A $\nabla f(x)$ jelölés magyarázata az [5] hivatkozásban szereplő könyv 238. oldalán megtalálható.

A 16. Definícióban megismert Fréchet derivált alkalmazható $f : \mathbb{R}^p \to \mathbb{R}^q$ függvény esetén is, ekkor a Fréchet-derivált, vagyis az $f'(a)$ megegyezik a 13. Definícióban megismert Jacobi-mátrixszal.

4.4. Magasabb rendű deriváltak

Tudjuk, hogy az $(X_1, \|\cdot\|_1), \ldots, (X_m, \|\cdot\|_m)$ normált terek szorzatán az $X := X_1 \times \ldots \times X_m$ vektorteret értjük a

$$\|x\| := \|x_1\|_1 + \ldots + \|x_m\|_m, \qquad x = (x_1, \ldots, x_m) \in X$$

normával ellátva.

24. Definíció. *Az $A : X \to Y$ leképezés m-lineáris, (azaz minden változójában lineáris leképezés)*[4] *ha bármely rögzített $1 \leq j \leq m$ egész és $x_i \in X_i$ pontok esetén $(1 \leq i \leq m, i \neq j)$, az*

$$X_j \ni x_j \mapsto A(x_1, \ldots, x_m)$$

függvény lineáris.

Ha $m = 1$, akkor visszakapjuk a lineáris leképezés fogalmát. A következő lemma leírja, hogy egy m-lineáris leképezés mikor lesz folytonos:

2. Lemma. *Az m-lineáris $A : X \to Y$ leképezés pontosan akkor folytonos, ha van olyan $M \geq 0$ szám, hogy*

$$\|A(x_1, \ldots, x_m)\|_Y \leq M\|x_1\|_1 \cdots \|x_m\|_m$$

minden $x_i \in X_i$-re, $i = 1, \ldots, m$.

Hasonlóan az 1. Lemmához itt is egy $M \geq 0$ szám esetén lesz az *m-lineáris* leképezés folytonos. Most nem bizonyítjuk be ezt a lemmát, de megtalálható a [4] hivatkozásban szereplő könyv 105. oldalán.

A 2. Állítás általánosítása a következő:

8. Állítás. *$L^m(X, Y) = L^m(X_1 \times \cdots \times X_m, Y)$ normált tér.*

Az előbbi állításnál több is igaz.

9. Állítás. *Ha $A \in L^m(X_1 \times \cdots \times X_m, Y)$, akkor az*

$$f(A)(x_1)(x_2, \ldots, x_m) := A(x_1, \ldots, x_m)$$

képlettel definiált

$$f : L^m(X_1 \times \cdots \times X_m, Y) \to L(X_1, L^{m-1}(X_2 \times \cdots X_m, Y))$$

leképezés izometrikus izomorfizmus.

[4] idézve az [5] hivatkozásban szereplő könyv 243. oldaláról.

Ennek a tételnek fontos következményei:

10. Állítás. *(a) Az*

$$L(X_1, L(X_2, \dots L(X_m, Y) \dots) \quad \text{\textit{és}} \quad L^m(X, Y)$$

normált terek izometrikusan izomorfak.
(b) Ha $\dim X < \infty$*, akkor minden* m*-lineáris* $A : X \to Y$ *leképezés folytonos.*

6. Megjegyzés. *Ha* $X_1 = \dots = X_m = Z$*, akkor* $X = Z^m$*-et írunk.*

A 3.1.3. alfejezetben megismertük a többváltozós függvények többszörös differenciálását. Normált térben hasonlóan definiáljuk a magasabb rendű deriváltakat, ezeket ugyancsak rekurzióval definiáljuk:

25. Definíció. *Legyen* $f : X \to Y$*,* $a \in D(f)$ *és* $k \geq 2$*.*
(a) f k*-szor differenciálható* a*-ban, ha differenciálható a környezetében, és ha* $f' : X \to L(X, Y)$ $k-1$*-szer differenciálható* a*-ban;* $(f')^{(k-1)}(a)$*-t az* f *függvény* k*-adik deriváltjának nevezzük* a*-ban, és* $f^{(k)}(a)$*-val jelöljük. Tehát*

$$f^{(k)}(a) \in L^k(X^k, Y) \qquad\qquad f^{(k)} : X \to L^k(X^k, Y).$$

(b) f k*-szor differenciálható, ha (az értelmezési tartomány nyílt, és ha)* f *minden* $a \in D(f)$ *pontban* k*-szor differenciálható.*
(c) f k*-szor folytonosan differenciálható vagy* C^k *osztálybeli, ha differenciálható, és ha* f' $k-1$*-szer folytonosan differenciálható. Ezt az* $f \in C^k$ *szimbólummal jelöljük.*
(d) f *végtelen sokszor differenciálható vagy* C^∞ *osztálybeli, ha minden* $k \in \mathbb{Z}$*-re* C^k*-beli. Ezt az* $f \in C^\infty$ *szimbólummal jelöljük.*

Fontos megjegyezni, hogy f pontosan akkor C^k osztálybeli, ha k-szor differenciálható és f^k folytonos.

Minden konstans függvény differenciálható, és a deriváltja azonosan nulla, tehát szintén konstans függvény. Vagyis a konstans függvények végtelen sokszor differenciálhatóak.

Többváltozós esetben megismert Young-tétel normált térben a következőképpen fogalmazható meg:

23. Tétel. *Ha* $f : X \to Y$ k*-szor differenciálható* a*-ban, ahol* $k \geq 2$*, akkor az* $f^{(k)} \in L^k(X^k, Y)$ *derivált szimmetrikus.*

Az előbbi tétel bizonyítása megtalálható a [4] hivatkozásban szereplő könyv 111. oldalán.

4.5. Alkalmazás

A Fréchet-derivált egyik alkalmazási területe a differenciálegyenleteken belül a variációszámítás, ami speciális típusú függvények szélsőértékeinek vizsgálatával foglalkozik. A variációszámítás feladatok során a minimalizálandó vagy maximalizálandó függvény egy (vagy néhány) függvénytől, más kifejezéssel élve végtelen sok változótól függ. Ezek a feladatok a fizikai alkalmazások kapcsán kerültek be a matematikába.

A variációszámítás az alábbi feladat vizsgálatából fejlődött ki:

8. Feladat. ***A brachisztrochon-probléma***
Legyen $b, d \in \mathbb{R}^+$. Kérdés: Milyen alakú pályán jut el egy anyagi pont a legrövidebb idő alatt egy függőleges sík (b, d) pontjából a $(0, 0)$ pontba, ha csak a gravitáció hatását vesszük figyelembe?

Az energia megmaradásának tételéből levezethető, hogy ha a test kezdősebesség nélkül indul, a gravitációs együtthatót pedig g jelöli, akkor

$$T(x) := \int_0^b \sqrt{\frac{1 + \dot{x}^2(t)}{2gx(t)}} dt$$

idő alatt halad végig a test egy olyan x függvény gráfja által meghatározott pálya mentén, amelyre $x \in C^1([0, b])$, $x(b) = d, x(0) = 0$ teljesül. A matematikai feladat ezek után a T funkcionál minimumának meghatározása.

A fenti példából a következő általános matematikai problémát fogalmazhatjuk meg:

26. Definíció. *Legyen $a, b, c, d \in \mathbb{R}$ olyan, hogy $a < b$, és legyen f az $\Omega \subset \mathbb{R}^3$ tartományon értelmezett folytonos, valós értékű függvény. Az*

$$\mathcal{M} := \{x \in C^1([a, b]); x(a) = c, x(b) = d, \mathcal{R}_{(id,x,\dot{x})} \subset \Omega\}$$

megengedett függvényosztályon értelmezett

$$I(x) := \int_a^b f(t, x(t), \dot{x(t)})dt = \int_a^b f \circ (id, x, \dot{x})$$

funkcionál szélsőértékhelyeinek meghatározását a legegyszerűbb variációs problémának nevezzük. Az f függvény neve: alapfüggvény.

A variációszámítás tárgya olyan módszerek kifejlesztése, amelyekkel a 26. Definícióban megadott I funkcionál szélsőértékhelyeit meg lehet keresni. Az $I : \mathcal{M} \to \mathbb{R}$ függvény szélsőértékeinek vizsgálatára alkalmazható a Fréchet-derivált fogalmán alapuló differenciálszámítás.

A 26. Definícióban szereplő I funkcionál Fréchet-deriváltja a következő (a [2] hivatkozásban szereplő könyv 305. oldalán található jelöléseket használva):

$$\delta I_{x_0} : C_0^1([a,b]) \to \mathbb{R},$$

$$\begin{aligned}\delta I_{x_0}(\eta) &:= \int_a^b \partial_2 f(x_0^{[1]}(t))\eta(t) + \partial_3 f(x_0^{[1]}(t))\dot{\eta}(t)dt \\ &= \int_a^b (\partial_2 f \circ x_0^{[1]})\eta + (\partial_3 f \circ x_0^{[1]})\dot{\eta}.\end{aligned}$$

A kapott funkcionált az I funkcionál $x_0 \in M$ függvényre vonatkozó első variációjának nevezzük, a $\delta I_{x_0} = 0$ egyenletet kielégítő x_0 függvényeket pedig stacionárius függvényeknek nevezzük. Az $x_0^{[1]}$ jelöli x_0 első felemeltjét, vagyis ha $x_0 \in C^1([a,b])$, akkor $x_0^{[1]} := (id, x_0, \dot{x_0})$. A $\delta I_{x_0} = 0$ egyenletet Euler-Lagrange-féle differenciálegyenletnek nevezzük.

5. Összefoglalás

Szakdolgozatomban a differenciálszámítás egyváltozós esetéből kiindulva a többváltozós eset vizsgálatán keresztül jutottam el a normált térbeli deriváláshoz.

Az első fejezetben az egyváltozós deriválást tekintettük át. Azon belül a korábbi tanulmányok során már megismert alapfogalmakat idéztük fel újból. Ezután megvizsgáltuk az elemi függvények differenciálhatóságát, valamint szélsőértékek keresésének módjával foglalkoztunk. A fejezet végén egy egyszerű alkalmazást néztük differenciálegyenletek területéről.

A következő fejezetben a többváltozós esetet vizsgálva eljutottunk egy bonyolultabb rendszerhez. Itt már vektorváltozós függvények deriválását tekintettük át. Mivel az egyváltozós esetben kapott általános definícióban szereplő tört itt értelmét veszti, mivel a nevező egy vektor, ezért újabb definíciókat kellett bevezetnünk. Először a valós értékű függvények deriválását idéztük fel, majd a differenciálhatóságot vizsgáltuk. Rátértünk a többszörös differenciálásra is. Majd a vektorértékű függvények differenciálhatóságát tekintettük át. Kitértünk az implicit függvények vizsgálatára is. A fejezet végén a feltételes szélsőérték keresés módjával foglalkoztunk, ami egy igen fontos alkalmazása a differenciálszámításnak. A biológia területéről néztünk példákat, ezzel is hangsúlyozva, hogy milyen sok tudományág alapja a deriválás.

Végül az utolsó fejezetben egy általános képletet kaptunk a deriválásra. Főbb definíciókat, tételeket néztünk át, ami az egy- és többváltozós esetre visszavezethetők. Először az alapfogalmakat, elemi tulajdonságokat általánosítottuk, amiket már az előbbi fejezetek során felidéztünk. Ezután az egyváltozós esetben megismert Lagrange-középértéktételt vizsgáltuk normált térben. Megnéztük a magasabb rendű deriváltakat, amiket rekurzió segítségével definiáltunk. A fejezet végén pedig megnéztük, hogy a variációszámítás hogyan alkalmazza a Fréchet-deriváltat.

Hivatkozások

[1] Laczkovich Miklós, T.Sós Vera: Analízis I-II, Nemzeti Tankönyvkiadó, 2006,2007.

[2] Tóth János, Simon L. Péter: Differenciálegyenletek, Typotex, 2005.

[3] Ron Larson, David C. Falco: Calculus: an applied approach, Brooks Cole, 2007.

[4] Komornik Vilmos: Valós analízis előadások I.-II., Typotex, 2003.

[5] Járai Antal: Modern alkalmazott analízis, Typotex, 2007.

Köszönetnyilvánítás

Végezetül szeretnék köszönetet mondani tanáraimnak, akik tanítottak engem.

Külön szeretném megköszönni témavezetőmnek Simon Péternek, a rengeteg segítséget, türelmet és észrevételt.

Köszönettel tartozom még azoknak, akik segítették munkámat és nem utolsósorban Édesanyámnak és Páromnak.

Nyilatkozat

Név:

ELTE Természettudományi Kar, szak:

ETR azonosító:

Szakdolgozat címe:

A szakdolgozat szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló munkám eredménye, saját szellemi termékem, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 20

a hallgató aláírása