Obsah

KAPITOLA 8

Spojité modely s více proměnnými

jedna proměnná nám k modelování nestačí?
– nevadí, stačí vzpomenout na vektory ...

Na samotném počátku našeho putování matematickou krajinou jsme hned viděli, že pracovat současně s více parametry nebylo obtížné, protože s vektory šlo počítat velice podobně jako se skaláry. Jen je třeba si věci dobře rozmyslet. Budeme se nyní znovu zabývat situacemi, kdy matematicky vyjádřené vztahy závisí na více (ale konečně mnoha) parametrech. Uvidíme, že vlastně ani není třeba překvapivých nových nápadů, stačí vždy šikovně redukovat problémy na takové, které už řešit umíme.

Zároveň se konečně budeme umět vrátit k diskusi situací, kdy hodnoty funkcí popisujeme pomocí jejich okamžitých změn – tj. malinko se zastavíme i u obyčejných a parciálních diferenciálních rovnic. Úplně závěrem zmíníme tzv. variační problémy.

Průběžně se budeme také jako obvykle snažit komentovat diskrétní varianty přístupů či problémů.

1. Funkce a zobrazení na $\mathbb{R}^n$

8.1

8.1. Funkce více proměnných. Pro praktické modelování

procesů (nebo objektů v grafice) jen velice zřídka vystačíme s funkcemi $\mathbb{R} \to \mathbb{R}$ jedné proměnné. Přinejmenším bývají potřebné funkce závislé na parametrech a často právě změna výsledků v závislosti na parametrech bývá důležitější než výsledek samotný. Budeme proto uvažovat funkce

$$f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \to \mathbb{R}$$

a budeme se snažit co nejlépe rozšířit naše metody pro sledování hodnot a jejich změn do této situace. Říkáme jim *funkce více proměnných*.

Pro snažší pochopení pojmů budeme často pracovat s případy $n = 2$ nebo $n = 3$ a přitom budeme místo číslovaných proměnných používat písmena x, y, z. To znamená, že funkce f definované v „rovině“ $\mathbb{R}^2$ budou značeny

$$f : \mathbb{R}^2 \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$$

a podobně v „prostoru“ $\mathbb{R}^3$

$$f : \mathbb{R}^3 \ni (x, y, z) \mapsto f(x, y, z) \in \mathbb{R}.$$

Podobně jako u funkcí jedné proměnné hovoříme o definičním oboru $A \subset \mathbb{R}^n$, na kterém je ta která funkce definována. Při zkoumání funkce zadané konkrétním výrazem bývá prvním úkolem zjistit co největší definiční obor, na kterém má tento výraz smysl.

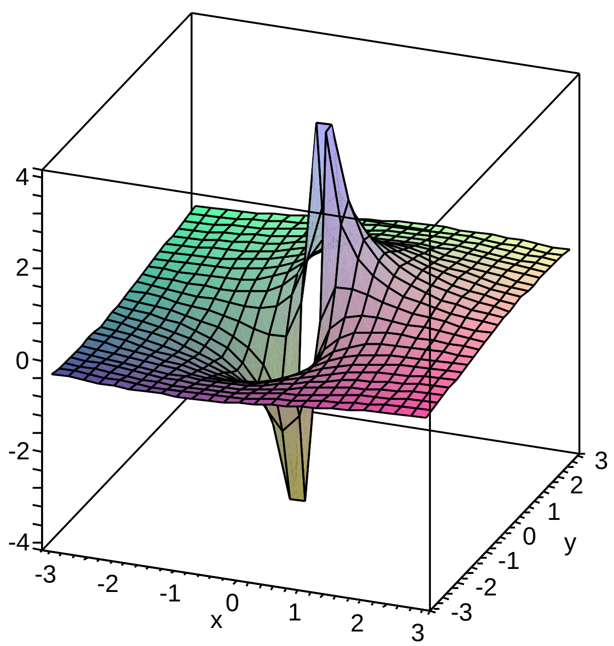
S každou takovou funkcí více proměnných bývá užitečné uvažovat její *graf*, tj. podmnožinu $G_f \subset \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ definovanou vztahem

$$G_f = \{(x_1, \dots, x_n, f(x_1, \dots, x_n));\ (x_1, \dots, x_n) \in A\},$$

kde A je definiční obor f. Např. grafem funkce definované v rovině vztahem

$$f(x, y) = \frac{x + y}{x^2 + y^2}$$

je docela pěkná plocha na obrázku a jejím maximálním definičním oborem jsou všechny body roviny kromě počátku $(0, 0)$.



Při definici a zejména při kreslení obrázku grafu jsme použili pevně zvolené *souřadnice* v rovině. Pokud pro některou z nich zvolíme pevnou hodnotu, zbude nám jen jedna proměnná. Pro pevně zvolenou hodnotu x tak např. dostáváme zobrazení

$$\mathbb{R} \to \mathbb{R}^3,\ y \mapsto (x, y, f(x, y)),$$

tj. *křivku* v prostoru $\mathbb{R}^3$. Křivky jsou vektorové funkce jediné proměnné, se kterými jsme již pracovali v šesté kapitole (viz 6.14). Na obrázku jsou čarami vyneseny obrazy takovýchto křivek pro některé pevně zvolené hodnoty souřadnic x a y.

Křivky $c : \mathbb{R} \to \mathbb{R}^n$ jsou vedle funkcí více proměnných nejjednoduššími příklady *zobrazení* $F : \mathbb{R}^m \to \mathbb{R}^n$, ke kterým se dostaneme brzy také.

U funkcí jedné proměnné jsme celý diferenciální a integrální počet vybudovali na základě pojmů konvergence, otevřených okolí, spojitosti atd. Tyto pojmy jsme poté v druhé části sedmé kapitoly zobecnili nejen pro euklidovské prostory

$\mathbb{R}^n$, ale i obecněji pro tzv. metrické prostory. Před čtením následujících odstavců bude vhodné si tyto pasáže pečlivě připomenout, případně dohledávat si tam potřebné pojmy a výsledky průběžně. Pro jistotu tady jen velice rychle shrneme aspoň něco málo.

8.2

8.2. Euklidovské prostory. Euklidovský prostor E_n vnímáme jako množinu bodů v $\mathbb{R}^n$ bez volby souřadnic a na jeho zaměření $\mathbb{R}^n$ pohlížíme jako na vektorový prostor možných přírůstků, které umíme k bodům prostoru E_n přičítat.

Navíc je na $\mathbb{R}^n$ zvolen standardní skalární součin

$$u \cdot v = \sum_{i=1}^{n} x_i y_i,$$

kde $u = (x_1, \dots, x_n)$ a $v = (y_1, \dots, y_n)$ jsou libovolné vektory. Tím je na E_n dána *metrika*, tj. funkce vzdálenosti $\|P - Q\|$ dvojic bodů P, Q předpisem

$$\|P - Q\|^2 = \|u\|^2 = \sum_{i=1}^{n} x_i^2,$$

kde u je vektor, jehož přičtením k bodu Q obdržíme bod P. Např. v rovině E_2 je tedy vzdálenost bodů $P_1 = (x_1, y_1)$ a $P_2 = (x_2, y_2)$ dána

$$\|P_1 - P_2\|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

Takto definovaná metrika splňuje trojúhelníkovou nerovnost pro každé tři body P, Q, R

$$\|P-R\| = \|(P-Q)+(Q-R)\| \le \|(P-Q)\|+\|(Q-R)\|,$$

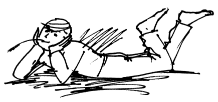
viz 3.25(1) nebo stejnou nerovnost (5.4) pro skaláry. Můžeme proto bez problému přenést (rozšířit) pro body P_i libovolného Euklidovského prostoru pojmy zavedené dosud pro reálné a komplexní skaláry:

Topologie euklidovského prostoru

- *Cauchyovská posloupnost*: posloupnost bodů P_i taková, že pro každé pevně zvolené $\epsilon > 0$ je $\|P_i - P_j\| < \epsilon$ pro všechny indexy, až na konečně mnoho výjimečných hodnot i, j,
- *konvergentní posloupnost*: posloupnost bodů P_i konverguje k bodu P, jestliže pro každé pevně zvolené $\epsilon > 0$ je $\|P_i - P\| < \epsilon$, až na konečně mnoho výjimečných hodnot i, j; bod P pak nazýváme *limitou* posloupnosti P_i,
- *hromadný bod P množiny $A \subset E_n$*: existuje posloupnost bodů v A konvergující k P a vesměs různých od P,
- *uzavřená množina*: obsahuje všechny své hromadné body,
- *otevřená množina*: její doplněk je uzavřený,
- *otevřené δ–okolí bodu P*: množina $\mathcal{O}_\delta(P) = \{Q \in E_n;\ \|P - Q\| < \delta\}$, $\delta \in \mathbb{R}$, $\delta > 0$,
- *hraniční bod P množiny A*: každé δ–okolí bodu P má neprázdný průnik s A i s komplementem $E_n \setminus A$,
- *vnitřní bod P množiny A*: existuje δ–okolí bodu P, které celé leží uvnitř A,

- *ohraničená množina*: leží celá v nějakém δ–okolí některého svého bodu (pro dostatečně velké δ),
- *kompaktní množina*: uzavřená a ohraničená množina.

Čtenář by měl investovat přiměřené úsilí do pročtení od-



stavců 3.25, 5.14–5.17 a 7.14–7.16 a 7.22 a zkusit si promyslet/připomenout definice a souvislosti všech těchto pojmů.

Zejména by mělo být z definic přímo zřejmé, že posloupnosti bodů P_i mají vlastnosti zmiňované v prvních dvou bodech předchozího výčtu tehdy a jen tehdy, když stejně nazvané vlastnosti mají reálné posloupnosti vzniklé z jednotlivých souřadnic bodů P_i ve kterékoliv kartézské souřadné soustavě. Proto také z Lemma 5.12 vyplývá, že každá Caychovská posloupnost bodů v E_n je konvergentní. Zejména je tedy E_n vždy úplným metrickým prostorem.

8.2a

8.3. Kompaktní množiny. Naše hrátky s otevřenými, uzavřenými nebo kompaktními množinami mohly v případě reálné přímky E_1 vypadat jako zbytečné, protože nakonec jsme stejně skoro vždy mluvili jen o intervalech.

U metrických prostorů ve ve druhé části kapitoly sedmé to možná bylo až moc složité. Stejný přístup je ale v případě euklidovských prostorů $\mathbb{R}^n$ docela jednoduchý a zároveň velmi užitečný a podstatný (a je to samozřejmě speciální případ obecných metrických prostorů).

Stejně jako v případě E_1 definujeme otevřené pokrytí množiny (tj. systém otevřených množin, v jejichž sjednocení je daná množina obsažena) a platí s drobnými formulačními úpravami i Věta 5.17:

Věta. *Pro podmnožiny $A \subset E_n$ v euklidovských prostorech platí:*

(1) A je otevřená, právě když je sjednocením nejvýše spočetného systému δ–okolí,
(2) každý bod $a \in A$ je buď vnitřní nebo hraniční,
(3) každý hraniční bod je buď izolovaným nebo hromadným bodem A,
(4) A je kompaktní, právě když každá v ní obsažená nekonečná posloupnost má podposloupnost konvergující k bodu v A,
(5) A je kompaktní, právě když každé její otevřené pokrytí obsahuje konečné pokrytí.

Důkaz. Důkaz z 5.17 lze bez úprav použít v případě

tvrzení (1)–(3), byť s novým chápání pojmů a nahrazením „otevřených intervalů“ jejich vícerozměrnými δ–okolími vhodných bodů.

Důkaz pro zbylá dvě tvrzení je však třeba dosti zásadně upravit. Bude proto dobré si projít důkaz příslušných obecných tvrzení pro metrické prostory v 7.22 a přitom přemýšlet, co je v případě euklidovských prostorů možné zjednodušit. $\square$

8.3

8.4. Křivky v E_n. Skoro celá naše diskuse kolem limit, derivací a integrálů funkcí v 5. a 6. kapitole se týkala funkcí s jednou reálnou proměnnou a reálnými nebo komplexními hodnotami s odůvodněním, že používáme pouze trojúhelníkovou nerovnost platnou pro velikosti reálných i komplexních čísel. Již tehdy jsme si povšimli, že se tento argument do značné míry přenáší na jakékoliv funkce jedné reálné proměnné s hodnotami v euklidovském prostoru $\mathbb{R}^n$ a uvedli jsme několik nástrojů pro práci s křivkami v odstavcích 6.14–6.17.

Připomeňme proto, že pro každou (parametrizovanou) křivku[1], tj. zobrazení $c : \mathbb{R} \to \mathbb{R}^n$ v n–rozměrném prostoru, můžeme pracovat s pojmy, které jednoduše rozšiřují naše úvahy z funkcí jedné proměnné:

- *limita:* $\lim_{t\to t_0} c(t) \in \mathbb{R}^n$
- *derivace:* $c'(t_0) = \lim_{t\to t_0} \frac{1}{|t-t_0|} \cdot (c(t) - c(t_0)) \in \mathbb{R}^n$
- *integrál:* $\int_a^b c(t)dt \in \mathbb{R}^n$.

Všimněme si také, že jak limita tak derivace křivek mají smysl v afinním prostoru, aniž bychom volili souřadnice (přičemž limitou posloupnosti je opět bod v původním prostoru, zatímco derivace je vektor v zaměření!). V případě integrálu ale musíme uvažovat křivky ve vektorovém prostoru $\mathbb{R}^n$. Důvod je vidět už v jednorozměrném případě, kde potřebujeme znát počátek, abychom mohli vidět „plochu pod grafem funkce“.

Opět je přímo z definice zjevné, že limity, derivace i integrály lze spočíst po jednotlivých n souřadných složkách v $\mathbb{R}^n$ a stejně se rozpozná i jejich existence.

U integrálu můžeme také přímo formulovat pro křivky analogii souvislosti Riemannova integrálu a primitivní funkce (viz 6.25):

Tvrzení. *Nechť c je křivka v $\mathbb{R}^n$, spojitá na intervalu $[a, b]$. Pak existuje její Riemannův integrál $\int_a^b c(t)dt$. Navíc je křivka*

$$C(t) = \int_a^t c(s)ds \in \mathbb{R}^n$$

dobře definovaná, diferencovatelná a platí $C'(t) = c(t)$ pro všechny hodnoty $t \in [a, b]$.

Horší je to s větou o střední hodnotě a obecněji s Taylorovou větou, viz 5.38 a 6.4. Ve zvolených souřadnicích je můžeme aplikovat na jednotlivé souřadné funkce diferencovatelné křivky $c(t) = (c_1(t), \dots, c_n(t))$ na konečném intervalu $[a, b]$. Dostaneme např. u věty o střední hodnotě existenci čísel t_i takových, že

$$c_i(b) - c_i(a) = (b - a) \cdot c_i'(t_i).$$

Tato čísla t_i ale budou obecně různá, nemůžeme proto vyjádřit rozdílový vektor koncových bodů $c(b) - c(a)$ jako násobek

[1]V geometrii se většinou rozlišuje mezi křivkou jakožto podmnožinou v E_n a její parametrizací $\mathbb{R} \to \mathbb{R}^n$. My zde pod pojmem „křivka“ rozumíme výhradně parametrizované křivky. Těm se v české geometrické literatuře často říká „dráha“

derivace křivky v jediném bodě. Např. v rovině E_2 pro diferencovatelnou křivku $c(t) = (x(t), y(t))$ takto dostáváme

$$\begin{aligned} c(b) - c(a) &= (x'(\xi)(b-a), y'(\eta)(b-a)) \\ &= (b-a) \cdot (x'(\xi), y'(\eta)) \end{aligned}$$

pro dvě (obecně různé) hodnoty $\xi, \eta \in [a, b]$. Pořád nám ale tato úvaha stačí na následující odhad

Lemma. *Je-li c křivka v E_n se spojitou derivací na kompaktním intervalu $[a, b]$, pak pro všechny $a \le s \le t \le b$ platí*

$$\|c(t) - c(s)\| \le \sqrt{n}(\max_{r \in [a,b]} \|c'(r)\|) \cdot |t - s|.$$

Důkaz. Přímým použitím věty o střední hodnotě dostáváme pro vhodné body r_i uvnitř intervalu $[s, t]$:

$$\begin{aligned} \|c(t) - c(s)\|^2 &= \sum_{i=1}^{n} (c_i(t) - c_i(s))^2 \le \sum_{i=1}^{n} (c_i'(r_i)(t-s))^2 \\ &\le (t-s)^2 \sum_{i=1}^{n} \max_{r \in [s,t]} c_i'(r)^2 \\ &\le n(\max_{r \in [s,t],\ i=1,\dots,n} |c_i'(r)|)^2 (t-s)^2 \\ &\le n \max_{r \in [s,t]} \|c'(r)\|^2 (t-s)^2. \end{aligned}$$

□

Důležitým pojmem je *tečný vektor* ke křivce $c : \mathbb{R} \to E_n$ v bodě $c(t_0) \in E_n$, který definujeme jako vektor v prostoru zaměření $\mathbb{R}^n$ daný derivací $c'(t_0) \in \mathbb{R}^n$. Přímka T zadaná parametricky

$$T : \quad c(t_0) + t \cdot c'(t_0)$$

se nazývá *tečna ke křivce* c v bodě t_0. Na rozdíl od tečného vektoru, tečna T coby neparametrizovaná přímka zjevně nezávisí na parametrizaci křivky c, protože při změně parametrizace dostaneme díky větě o derivování složených funkcí znovu stejný tečný vektor, až na násobek.

8.4

8.5. Parciální derivace. Pro každou funkci $f : \mathbb{R}^n \to \mathbb{R}$

a libovolnou křivku $c : \mathbb{R} \to \mathbb{R}^n$ máme k dispozici jejich kompozici $(f \circ c)(t) : \mathbb{R} \to \mathbb{R}$. Tato složená funkce $F \circ c$ vypovídá o chování funkce f podél křivky c. Nejjednodušší bude použít přímky.

SMĚROVÉ A PARCIÁLNÍ DERIVACE

Definice. Řekneme, že $f : \mathbb{R}^n \to \mathbb{R}$ má *derivaci ve směru vektoru* $v \in \mathbb{R}^n$ v bodě $x \in E_n$, jestliže existuje derivace $d_v f(x)$ složeného zobrazení $t \mapsto f(x + tv)$ v bodě $t = 0$, tj.

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

Hodnotě $d_v f$ také říkáme *směrová derivace*.

Speciální volbou přímek ve směru souřadných os dostáváme tzv. *parciální derivace funkce* f, které značíme $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, nebo bez odkazu na samotnou fukci jako operace $\frac{\partial}{\partial x_i}$.

Pro funkce v rovině tak dostáváme

$$\frac{\partial}{\partial x} f(x, y) = \lim_{t\to 0} \frac{1}{t}(f(x+t, y) - f(x, y))$$
$$\frac{\partial}{\partial y} f(x, y) = \lim_{t\to 0} \frac{1}{t}(f(x, y+t) - f(x, y)).$$

Zejména je vidět, že parciálně podle vybrané proměnné derivujeme tak, že prostě všechny ostatní proměnné považujeme za konstanty a postupujeme jako u funkcí jedné proměnné.

8.4a

8.6. Diferenciál funkce $f : \mathbb{R}^n \to \mathbb{R}$. Se samotnými parciálními nebo směrovými derivacemi nevystačíme pro dobrou aproximaci chování funkce lineárními výrazy. Asi bychom přirozeně očekávali, že „diferencovatelná“ funkce více proměnných bude složením s jakoukoliv diferencovatelnou křivkou dávat diferencovatelné funkce jedné proměnné, které už dobře známe.

Podívejme se ale např. na funkce v rovině zadané výrazy

$$g(x, y) = \begin{cases} 1 & \text{když } yx = 0 \\ 0 & \text{jinak} \end{cases}$$
$$h(x, y) = \begin{cases} 1 & \text{když } y = x^2 \neq 0 \\ 0 & \text{jinak} \end{cases}.$$

Evidentně žádná z nich neprodlužuje všechny hladké křivky procházející bodem (0, 0) na hladké funkce. Přitom ale pro g existují obě parciální derivace v (0, 0) a jiné směrové derivace neexistují, zatímco pro h existují všechny směrové derivace v bodě (0, 0) a je dokonce $d_v h(0) = 0$ pro všechny směry v, takže jde o lineární závislost na $v \in \mathbb{R}^2$.

Snadno si také představíme funkci f, která bude mít podél přímek $(r\cos\theta, r\sin\theta)$ s pevným úhlem θ hodnoty $k(\theta)r$, přičemž $k(\theta)$ je periodická lichá funkce v úhlu θ, s periodou 2π. Její směrové derivace $d_v f$ v (0, 0) všechny existují, ale pro obecné funkce $k(\theta)$ zcela jistě nepůjde o lineární výrazy v závislosti na směrech v.

Budeme proto sledovat případ funkcí jedné proměnné co nejdůsledněji a podobné patologické chování funkcí vyloučíme přímo definicí:

DIFERENCIÁL

Definice. Funkce $f : \mathbb{R}^n \to \mathbb{R}$ je *diferencovatelná v bodě x*, jestliže zároveň platí tři vlastnosti:

(1) v bodě x existují směrové derivace $d_v f(x)$ pro všechny vektory $v \in \mathbb{R}^n$,
(2) $d_v f(x)$ je lineární v závislosti na přírůstku v,
(3) $\lim_{v\to 0} \frac{1}{\|v\|}\big(f(x+v) - f(x) - d_v f(x)\big) = 0.$

Lineární výraz $d_v f$ (ve vektorové proměnné v) nazýváme *diferenciál funkce f* vyčíslený na přírůstku v.

Řečeno slovy, požadujeme, aby v bodě x existovalo dobré přiblížení přírůstků funkce f pomocí lineární funkce přírůstků proměnných veličin.

Přímo z definice směrových derivací vyplývá, že můžeme také diferenciál definovat pouze pomocí vlastnosti (3). Skutečně, pokud existuje nějaká lineární forma $df(x)$ taková, že pro přírůstky v v bodě x platí vlastnost (3) s $d_v f(x) = df(x)(v)$, pak je zjevně $df(x)(v)$ právě směrovou derivací funkce f v bodě x a vlastnosti (1) a (2) jsou tedy splněny automaticky.

Podívejme se, co umíme říci o diferenciálu funkce $f(x, y)$ v rovině za předpokladu, že obě parciální derivace $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ existují a jsou spojité v okolí bodu (x_0, y_0).

Uvažme za tím účelem jakoukoliv hladkou křivku $t \mapsto (x(t), y(t))$ s $x_0 = x(0)$, $y_0 = y(0)$. S použitím věty o střední hodnotě na funkce jedné proměnné v obou sčítancích zvlášť dovodíme, že

$$\begin{aligned}
&\tfrac{1}{t}\big(f(x(t), y(t)) - f(x_0, y_0)\big) = \\
&\tfrac{1}{t}\big(f(x(t), y(t)) - f(x_0, y(t))\big) + \tfrac{1}{t}\big(f(x_0, y(t)) - f(x_0, y_0)\big) \\
&= \tfrac{1}{t}(x(t) - x_0)\cdot\frac{\partial f}{\partial x}(x(\xi), y(t)) + \tfrac{1}{t}(y(t) - y_0)\cdot\frac{\partial f}{\partial y}(x_0, y(\eta))
\end{aligned}$$

pro vhodná čísla ξ a η mezi 0 a t.

Zejména tedy pro každou posloupnost čísel t_n jdoucí k nule získáme příslušné posloupnosti čísel ξ_n a η_n, které také budou konvergovat k nule, a pro všechny bude platit vyjádření výše.

Limitním přechodem $t \to 0$ proto díky spojitosti parciálních derivací dostáváme (viz test konvergence funkce pomocí vybraných posloupností hodnot argumentů, 5.23, a Věta 5.22 o limitách součtů a součinů funkcí)

$$\frac{d}{dt} f(x(t), y(t))|_{t=0} = x'(0)\frac{\partial f}{\partial x}(x_0, y_0) + y'(0)\frac{\partial f}{\partial y}(x_0, y_0),$$

což je příjemné rozšíření platnosti věty o derivování složených funkcí jedné proměnné pro vektorově hodnotové funkce.

Samozřejmě, speciální volbou parametrizovaných přímek

$$(x(t), y(t)) = (x_0 + t\xi, y_0 + t\eta)$$

přechází náš výpočet při $v = (\xi, \eta)$ na rovnost

$$d_v f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\xi + \frac{\partial f}{\partial y}(x_0, y_0)\eta$$

a tento vztah můžeme pěkně vyjádřit způsobem, kterým jsme v lineární algebře zapisovali souřadná vyjádření lineárních funkcí na vektorových prostorech:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

Jinými slovy, směrová derivace $d_v f$ je skutečně lineární funkce $\mathbb{R}^n \to \mathbb{R}$ na přírůstcích, se souřadnicemi danými právě parciálními derivacemi.

Podobným postupem nyní budeme umět dokázat, že předpoklad spojitých parciálních derivací v daném bodě zajišťuje i aproximační vlastnosti diferenciálu.

Budeme už rovnou uvažovat obecné funkce více proměnných:

8.4b **8.7. Věta.** *Nechť $f : E_n \to \mathbb{R}$ je funkce n proměnných, která má v okolí bodu $x \in E_n$ spojité parciální derivace. Pak existuje její diferenciál df v bodě x a jeho souřadné vyjádření je dáno výrazem.*

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n.$$

Důkaz. Odvození věty je naprosto analogické výše uvedenému postupu v případě $n = 2$. Musíme být jen být opatrní v detailech a dokončit úvahu o aproximačních vlastnostech. Úplně stejně jako výše uvažujeme křivku

$$c(t) = (c_1(t), \dots, c_n(t)),$$

$c(0) = (0, ..., 0)$, a bod $x \in \mathbb{R}^n$ a vyjádříme pro složenou funkci $f(c(t))$ rozdíl $f(x + c(t)) - f(x)$ takto

$$\begin{aligned}
&f(x_1 + c_1(t), \dots, x_n + c_n(t)) - f(x_1, x_2 + c_2(t), \dots) \\
&\quad + f(x_1, x_2 + c_2(t), \dots)) - f(x_1, x_2, \dots, x_n + c_n(t)) \\
&\quad \vdots \\
&\quad + f(x_1, x_2, \dots, x_n + c_n(t)) - f(x_1, x_2, \dots, x_n).
\end{aligned}$$

Na všech n sčítanců teď můžeme uplatnit větu o střední hodnotě a, stejně jako v případě dvou proměnných, dostáváme

$$\begin{aligned}
&(c_1(t) - c_1(0)) \frac{\partial f}{\partial x_1}(x_1 + c_1(\theta_1), x_2 + c_2(t), \dots, x_n + c_n(t)) \\
&+ (c_2(t) - c_2(0)) \frac{\partial f}{\partial x_2}(x_1, x_2 + c_2(\theta_2), \dots, x_n + c_n(t)) \\
&\vdots \\
&+ (c_n(t) - c_n(0)) \frac{\partial f}{\partial x_n}(x_1, x_2, \dots, x_n + c_1(\theta_n)),
\end{aligned}$$

pro vhodné hodnoty $0 \le \theta_i \le t$. Jde o konečný součet, proto stejnou argumentací jako v případě dvou proměnných ověříme

$$\frac{d}{dt} f(x + c(t))_{t=0} = c_1'(0) \frac{\partial f}{\partial x_1}(x) + \dots + c_n'(0) \frac{\partial f}{\partial x_n}(x).$$

Speciální volbou křivek $c(t) = x + tv$ pro směrový vektor v máme ověřeno tvrzení o existenci a linearitě směrových derivací v bodě x.

Zároveň ale můžeme úplně stejně aplikovat větu o střední hodnotě na rozdíl

$$\begin{aligned}
f(x + v) - f(x) &= d_v f(x + \theta v) \\
&= v_1 \frac{\partial f}{\partial x_1}(x + \theta v) + \dots + v_n \frac{\partial f}{\partial x_n}(x + \theta v)
\end{aligned}$$

s vhodným $0 \leq \theta \leq 1$, kde druhá rovnost platí, podle výše odvozeného výrazu pro směrové derivace, pro dostatečně malá v díky spojitosti parciálních derivací na okolí bodu x.

Protože jsou všechny parciální derivace spojité v bodě x, víme, že pro libovolně malé $\epsilon > 0$ můžeme najít okolí U počátku v $\mathbb{R}^n$ takové, že se pro $w \in U$ budou všechny parciální derivace $\frac{\partial f}{\partial x_i}(x+w)$ lišit od $\frac{\partial f}{\partial x_i}(x)$ o méně než ϵ. Dostaneme pak odhad

$$\frac{1}{\|w\|}\big(f(x+w) - f(x) - d_w f(x+\theta w)\big) \leq \frac{n}{\|w\|}\|w\|\epsilon$$

a tedy i aproximační vlastnost diferenciálu je splněna. □

8.5

8.8. Tečná rovina ke grafu funkce. Lineární přiblížení chování funkce diferenciálem můžeme také obdobně k funkcím jedné proměnné vyjádřit ve vztahu k jejímu grafu. Jen místo tečen musíme pracovat s nadrovinami.

Pro případ funkce na E_2 a pevně zvoleného bodu $(x_0, y_0) \in E_2$ uvažme rovinu v E_3 zadanou rovnicí

$$\begin{aligned} z &= f(x_0, y_0) + df(x_0, y_0)(x - x_0, y - y_0) \\ &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \end{aligned}$$

Již jsme viděli, že přírůstek funkčních hodnot diferencovatelné funkce $f : E_n \to \mathbb{R}$ v bodech $x + tv$ a x je vždy vyjádřen pomocí směrové derivace $d_v f$ ve vhodném bodě na jejich spojnici. Tato rovina má tedy jako jediná ze všech rovin procházejících bodem (x_0, y_0) vlastnost, že v ní leží derivace a tedy i tečny všech křivek

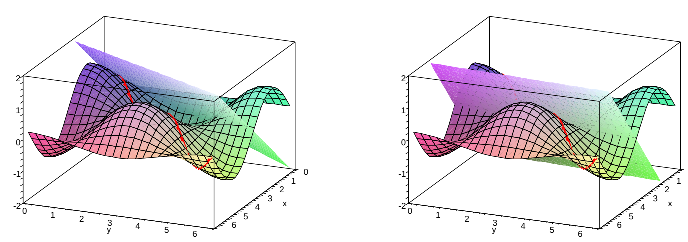
$$c(t) = (x(t), y(t), f(x(t), y(t))).$$

Říkáme jí *tečná rovina* ke grafu funkce f.

Na obrázku jsou zobrazeny dvě tečné roviny ke grafu funkce

$$f(x, y) = \sin(x)\cos(y).$$

Diagonálně vedená čára je obrazem křivky $c(t) = (t, t, f(t, t))$.



Pro funkce n proměnných definujeme tečnou rovinu jako analogii k tečné rovině k ploše v trojrozměrném prostoru. Místo zaplétání se do spousty indexů bude snad užitečná vzpomínka na afinní geometrii, kde jsme s tzv. nadrovinami již pracovali, viz odstavec 4.3.

Tečná (nad)rovina grafu funkce v bodě

Definice. *Tečná nadrovina* ke grafu funkce $f : \mathbb{R}^n \to \mathbb{R}$ v bodě $x \in \mathbb{R}^n$ je nadrovina procházející bodem $(x, f(x))$ se zaměřením, které je grafem lineárního zobrazení $df(x) : \mathbb{R}^n \to \mathbb{R}$, tj. diferenciálu v bodě $x \in E_n$.

Definice vychází ze skutečnosti, že směrová derivace $d_v f$ je dána přírůstkem na tečné (nad)rovině odpovídajícím přírůstku argumentu v.

Z těchto úvah vyplývá řada analogií s funkcemi jedné proměnné. Zejména má diferencovatelná funkce f na E_n v bodě $x \in E_n$ nulový diferenciál tehdy a jen tehdy, když její složení s libovolnou křivkou procházející tímto bodem zde má stacionární bod, tj. ani neroste ani neklesá v lineárním přiblížení.

Jinak řečeno, tečná rovina je v takovém bodě rovnoběžná s nadrovinou proměnných (tj. její zaměření je $E_n \subset E_{n+1}$ s přidanou nulovou poslední souřadnicí). To samozřejmě neznamená, že v takovém bodě musí mít f aspoň lokálně buď maximum nebo minimum. Stejně jako u funkcí jedné proměnné můžeme rozhodovat teprve podle derivací vyšších.

8.6

8.9. Derivace vyšších řádů. Stejně jako v přídpadě jedné proměnné, operaci derivování je možné iterovat. Tentokrát si můžeme pro každou iteraci vybrat jiný směr.

Jestliže vybereme pevný přírůstek $v \in \mathbb{R}^n$, zadává vyčíslení diferenciálů na tomto přírůstku (diferenciální) operaci na diferencovatelných funkcích $f : E_n \to \mathbb{R}$

$$f \mapsto d_v f = df(v)$$

a výsledkem je opět funkce $df(v) : E_n \to \mathbb{R}$. Jestliže je tato funkce opět diferencovatelná, může opakovat totéž s jiným přírůstkem atd. Zejména tedy můžeme pracovat s iteracemi parciálních derivací. Pro *parciální derivace druhého řádu* píšeme

$$\left(\frac{\partial}{\partial x_j} \circ \frac{\partial}{\partial x_i}\right) f = \frac{\partial^2}{\partial x_i \partial x_j} f = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

V případě opakované volby $i = j$ píšeme také

$$\left(\frac{\partial}{\partial x_i} \circ \frac{\partial}{\partial x_i}\right) f = \frac{\partial^2}{\partial x_i^2} f = \frac{\partial^2 f}{\partial x_i^2}.$$

Úplně stejně postupujeme při dalších iteracích a hovoříme o *parciálních derivacích k-tého řádu*

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}.$$

Obecněji můžeme iterovat (u dostatečně diferencovatelných funkcí) také libovolné směrové derivace, např. $d_v \circ d_w f$ pro dva pevné přírůstky $v, w \in \mathbb{R}^n$.

k–KRÁT DIFERENCOVATELNÉ FUNKCE

Řekneme, že je funkce $f : E_n \to \mathbb{R}$ *k–krát diferencovatelná* v bodě x, jestliže všechny parciální derivace až do řádu k včetně existují na nějakém okolí bodu x a jsou v tomto bodě spojité.

Řekneme, že funkce f je k–diferencovatelná, jestliže je k–krát diferencovatelná ve všech bodech svého definičního oboru.

Abychom si vše ukázali v co nejjednodušší formě, budeme opět pracovat chvíli v rovině E_2 za přepokladu spojitosti parciálních derivací druhého řádu. V rovině a prostoru se často stručně značí iterované derivace pouhými odkazy jmen proměnných v pozici indexů u funkce, např.

$$f_x = \frac{\partial f}{\partial x},\ f_{xx} = \frac{\partial^2 f}{\partial x^2},\ f_{xy} = \frac{\partial^2 f}{\partial x \partial y},\ \ f_{yx} = \frac{\partial^2 f}{\partial y \partial x}.$$

Ukážeme, že ve skutečnosti spolu za rozumných podmínek parciální derivace komutují, tzn. není potřeba dbát na pořadí, ve kterém je provádíme.

Dle předpokladu existence a spojitosti parciálních derivací existují limity

$$\begin{aligned} f_{xy}(x, y) &= \lim_{t\to 0} \frac{1}{t}\big(f_x(x, y+t) - f_x(x, y)\big) \\ &= \lim_{t\to 0} \frac{1}{t}\Bigg(\lim_{s\to 0} \frac{1}{s}\big(f(x+s, y+t) - f(x, y+t) \\ &\qquad - f(x+s, y) + f(x, y)\big)\Bigg). \end{aligned}$$

Protože ale limity můžeme vyjádřit pomocí libolného výběru hodnot $t_n \to 0$ a $s_n \to 0$ a limit příslušných posloupností, bude jistě také platit

$$\begin{aligned} f_{xy}(x, y) = \lim_{t\to 0} \frac{1}{t^2}\Bigg(&\big(f(x+t, y+t) - f(x, y+t)\big) \\ &- \big(f(x+t, y) - f(x, y)\big)\Bigg) \end{aligned}$$

a tato limitní hodnota je spojitá v (x, y).

Označme si výraz, ze kterého bereme poslední limitu, jako funkci $\varphi(x, y, t)$ a zkusme jej vyjádřit pomocí parciálních derivací. Pro dočasně pevné t si označme $g(x, y) = f(x+t, y) - f(x, y)$. Pak výraz v poslední velké závorce je díky větě o střední hodnotě roven

$$g(x, y+t) - g(x, y) = t \cdot g_y(x, y+t_0).$$

pro nějaké vhodné t_0, které je mezi nulou a t (a hodnota t_0 závisí na t).

Nyní $g_y(x, y) = f_y(x+t, y) - f_y(x, y)$ a proto můžeme psát φ jako

$$\begin{aligned} \varphi(x, y, t) &= \frac{1}{t} g_y(x, y+t_0) \\ &= \frac{1}{t}\big(f_y(x+t, y+t_0) - f_y(x, y+t_0)\big). \end{aligned}$$

Opětovnou aplikací věty o střední hodnotě,

$$\varphi(x, y, t) = f_{yx}(x + t_1, y + t_0)$$

pro vhodné t_1 mezi nulou a t. Když ale velkou závorku rozdělíme na $(f(x+t, y+t) - f(x+t, y)) - (f(x, y+t) - f(x, y))$, dostaneme stejným postupem s funkcí $h(x, y) = f(x, y+t) - f(x, y)$ vyjádření

$$\varphi(x, y, t) = f_{xy}(x + s_0, y + s_1)$$

s obecně jinými konstantami s_0 a s_1. Protože jsou druhé parciální derivace podle našeho předpoklady spojité, musí i limita pro $t \to 0$ zaručit požadovanou rovnost

$$f_{xy}(x, y) = f_{yx}(x, y)$$

ve všech bodech (x, y).

Stejný postup pro funkce n proměnných dokazuje následující základní výsledek:

Záměnnost parciálních derivací

8.6a **8.10. Věta.** *Nechť $f : E_n \to \mathbb{R}$ je k-krát diferencovatelná funkce se spojitými parciálními derivacemi až do řádu k včetně v okolí bodu $x \in \mathbb{R}^n$. Pak všechny parciální derivace nezávisí na pořadí derivování.*

Důkaz. Důkaz pro druhý řád byl proveden výše pro $n = 2$ a postup v obecném případě se nijak neliší. Formálně můžeme obecný případ u dvou derivací odbýt i tvrzením, že se vždy celá argumentace odehraje ve dvourozměrném afinním podprostoru, tj. všechny ostatní proměnné považujeme za konstantní a v argumentaci nijak aktivně nevystoupí.

U derivací vyššího řádu důkaz dokončíme indukcí podle řádu. Skutečně, každé pořadí indexů lze vytvořit záměnami sousedících dvojic. □

8.6b

8.11. Hessián. Tak jako jsme u derivací prvního řádu zavedli diferenciál coby lineární formu $df(x)$ přibližující nejlépe v daném bodu x funkci f, budeme nyní chtít porozumět kvadratickému přiblížení funkcí $f : E_n \to \mathbb{R}$.

Hessián

Definice. Je-li $f : \mathbb{R}^n \to \mathbb{R}$ libovolná dvakrát diferencovatelná funkce, nazýváme symetrickou matici funkcí

$$Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

Hessián funkce f v bodě x.

Z předchozích úvah jsme již viděli, že vynulování diferenciálu v bodě $(x, y) \in E_2$ zaručuje stacionární chování podél všech křivek v tomto bodu. Hessián

$$Hf(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{xy}(x, y) & f_{yy}(x, y) \end{pmatrix}$$

hraje roli druhé derivace.

Pro každou parametrizovanou přímku

$$c(t) = (x(t), y(t)) = (x_0 + \xi t, y_0 + \eta t)$$

budou totiž mít funkce jedné proměnné

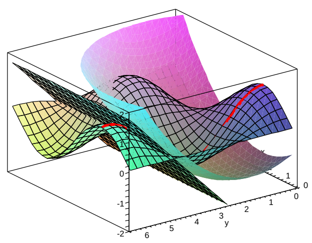
$$\alpha(t) = f(x(t), y(t))$$
$$\beta(t) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)\xi + \frac{\partial f}{\partial y}(x_0, y_0)\eta + \frac{1}{2}\Big(f_{xx}(x_0, y_0)\xi^2 + 2f_{xy}(x_0, y_0)\xi\eta + f_{yy}(x_0, y_0)\eta^2 \Big)$$

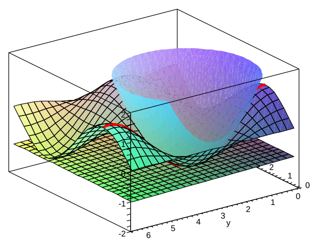
stejné derivace do druhého řádu včetně (přepočtěte!). Funkci β přitom můžeme zapsat vektorově jako

$$\beta(t) = f(x_0, y_0) + df(x_0, y_0)\cdot\begin{pmatrix}\xi \\ \eta\end{pmatrix} + \frac{1}{2}(\xi\ \eta)\cdot Hf(x_0, y_0)\cdot\begin{pmatrix}\xi \\ \eta\end{pmatrix}$$

nebo $\beta(t) = f(x_0, y_0) + df(x_0, y_0)(v) + \frac{1}{2}Hf(x_0, y_0)(v, v)$, kde $v = (\xi, \eta)$ je přírůstek zadaný derivací křivky $c(t)$ a Hessián je použit jako symetrická 2–forma.

To je vyjádření, které již určitě připomíná Taylorovu větu funkcí jedné proměnné, přesněji řečeno kvadratické přiblížení funkce Taylorovým polynomem druhého řádu. Na následujícím obrázku je vynesena jak tečná rovina tak toto kvadratické přiblížení pro dva různé body a funkci $f(x, y) = \sin(x)\cos(y)$.





8.7

8.12. Taylorova věta. Vícerozměrná verze Taylorovy věty je také příkladem matematického tvrzení, kde složitou částí je nalezení správné formulace. Důkaz je už pak docela snadný.

Budeme postupovat ve výše naznačeném směru a zavedeme si značení pro jednotlivé části $D^k f$ aproximací vyšších řádů pro funkce $f : E_n \to \mathbb{R}^n$. Budou to vždy k–lineární výrazy v přírůstcích a nás bude zajímat jen jejich vyčíslení na k stejných hodnotách. Již jsme diskutovali diferenciál $D^1 f = df$ v prvním řádu a hessián $D^2 f = Hf$ v řádu druhém. Obecně pro funkce $f : E_n \to \mathbb{R}$, body $x = (x_1, \dots, x_2) \in E_n$ a přírůstky $v = (\xi_1, \dots, \xi_n)$ klademe

$$D^k f(x)(v) = \sum_{1 \le i_1, \dots, i_k \le n} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x_1, \dots, x_n)\cdot \xi_{i_1} \cdots \xi_{i_k}.$$

Názorným příkladem (s využitím symetrií parciálních derivací) je pro E_2 výraz třetího řádu

$$D^3 f(x,y)(\xi,\eta) = \frac{\partial^3 f}{\partial x^3}\xi^3 + 3\frac{\partial^3 f}{\partial x^2 \partial y}\xi^2\eta + 3\frac{\partial^3 f}{\partial x \partial y^2}\xi\eta^2 + \frac{\partial^3 f}{\partial y^3}\eta^3$$

a obecně

$$D^k f(x,y)(\xi,\eta) = \sum_{\ell=0}^{k} \binom{k}{\ell} \frac{\partial^k f}{\partial x^{k-\ell}\partial y^{\ell}}\xi^{k-\ell}\eta^{\ell}.$$

Taylorův rozvoj se zbytkem

Věta. *Nechť $f : E_n \to \mathbb{R}$ je k–krát diferencovatelná funkce v okolí $\mathcal{O}_\delta(x)$ bodu $x \in E_n$. Pro každý přírůstek $v \in \mathbb{R}^n$ s velikostí $\|v\| < \delta$ pak existuje číslo $0 \le \theta \le 1$ takové, že*

$$f(x+v) = f(x) + D^1 f(x)(v) + \frac{1}{2!}D^2 f(x)(v) + \dots + \frac{1}{(k-1)!}D^{k-1} f(x)(v) + \frac{1}{k!}D^k f(x+\theta\cdot v)(v).$$

Důkaz. Pro přírůstek $v \in \mathbb{R}^n$ uvažujme parametrizovanou přímku $c(t) = x + tv$ v E_n a zkoumejme funkci $\varphi : \mathbb{R} \to \mathbb{R}$ definovanou složením $\varphi(t) = f \circ c(t)$. Taylorova věta pro funkce jedné proměnné říká (viz Věta 6.4)

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \dots + \frac{1}{(k-1)!}\varphi^{(k-1)}(0)t^{k-1} + \frac{1}{k!}\varphi^{(k)}(\theta)t^k.$$

Zbývá nám tedy jen ověřit, že postupným derivováním složené funkce φ dostaneme právě požadovaný vztah. To lze vcelku snadno provést indukcí přes řád k.

Pro $k = 1$ splývá Taylorova věta s již několikrát využitým důsledkem věty o střední hodnotě aplikované na směrovou derivaci. Při jeho odvození jsme vyšli ze vztahu

$$\frac{d}{dt}\varphi(t) = \frac{\partial f}{\partial x_1}(x(t)) \cdot x_1'(t) + \dots + \frac{\partial f}{\partial x_n}(x(t)) \cdot x_n'(t),$$

který platí pro každou křivku a funkci f. To znamená, že

$$D^1 f(c(t))(v) = D^1 f(c(t))(c'(t))$$

pro všechna t v okolí nuly. Stejně budeme postupovat pro funkce $D^\ell f$. Místo přírůstku v můžeme psát $c'(t)$ a zapamatujme si, že další derivování $c(t)$ již vede identicky na nulu všude, tj. $c''(t) = 0$ pro všechna t (protože jde o parametrizovanou přímku).

Předpokládejme, že

$$D^\ell f(x)(v) = \sum_{1\le i_1,\dots,i_\ell\le n}\left(\frac{\partial^\ell f}{\partial x_{i_1}\dots\partial x_{i_\ell}}(x_1(t),\dots,x_n(t))\cdot x'_{i_1}(t)\cdots x'_{i_\ell}(t)\right)$$

a spočtěme totéž pro $\ell+1$. Derivování složené funkce dá podle výše odvozeného vztahu pro derivaci prvního řádu v daném směru a podle pravidla o derivání součinu (viz Věta 5.33)

$$\begin{aligned}\frac{d}{dt}D^\ell f(c(t))(c'(t)) &= \\ &= \frac{d}{dt}\sum_{1\le i_1,\dots,i_\ell\le n}\left(\frac{\partial^\ell f}{\partial x_{i_1}\dots\partial x_{i_\ell}}(x_1(t),\dots,x_n(t))\cdot x'_{i_1}(t)\cdots x'_{i_\ell}(t)\right)\\ &= \sum_{1\le i_1,\dots,i_\ell\le n}\left(\sum_{j=1}^n\frac{\partial^{\ell+1} f}{\partial x_{i_1}\dots\partial x_{i_\ell}\partial x_j}(x_1(t),\dots,x_n(t))\cdot x'_j(t)\cdot x'_{i_1}(t)\cdots x'_{i_\ell}(t)\right)+0\end{aligned}$$

a to skutečně je požadovaný vztah pro řád $\ell+1$. Taylorova věta nyní vyplývá z vyčíslení v bodě $t=0$ a dosazení do rovnosti pro φ na začátku tohoto důkazu. □

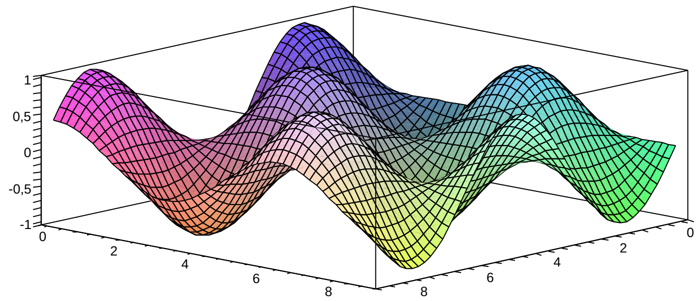
8.13. Lokální extrémy funkcí více proměnných. Zkusme

se nyní s pomocí diferenciálu a hessiánu podívat na lokální maxima a minima funkcí na E_n. Stejně jako v případě funkce jedné proměnné řekneme o vnitřním bodu $x_0\in E_n$ definičního oboru funkce f, že je (lokálním) *maximem* nebo *minimem*, jestliže existuje jeho okolí U takové, že pro všechny body $x\in U$ splňuje funkční hodnota $f(x)\le f(x_0)$ nebo $f(x)\ge f(x_0)$. Pokud nastává v předchozích nerovnostech ostrá nerovnost pro všechny $x\ne x_0$, hovoříme o *ostrém extrému*.

Pro jednoduchost budeme nadále předpokládat, že naše funkce f má spojité parciální derivace prvního i druhého řádu na svém definičním oboru. Nutnou podmínkou pro existenci maxima nebo minima v bodě x_0 je vymizení diferenciálu v tomto bodě, tj. $df(x_0)=0$. Skutečně, pokud je $df(x_0)\ne 0$, pak existuje směr v, ve kterém je $d_vf(x_0)\ne 0$. Pak ovšem nutně je podél přímky x_0+tv na jednu stranu od bodu x_0 hodnota funkce roste a na druhou klesá, viz (5.32).

Vnitřní bod $x\in E_n$ definičního oboru funkce f, ve kterém je diferenciál $df(x)$ nulový nazýváme *stacionární bod funkce f*.

Budeme opět chvíli pracovat s jednoduchou funkcí v E_2 abychom závěry přímo mohli ilustrovat. Uvažme funkci $f(x,y)=\sin(x)\cos(y)$, která už byla předmětem diskuse a obrázků v odstavcích 8.9 a 8.8. Svým tvarem tato funkce připomíná známá kartonová plata na vajíčka, je tedy předem zřejmé, že najdeme řadu extrémů, ale ještě více stacionárních bodů, která ve skutečnosti extrémy nebudou (ta „sedýlka“ viditelná na obrázku).



Spočtěme si tedy první a poté druhé derivace:

$$f_x(x,y) = \cos(x)\cos(y), \quad f_y(x,y) = -\sin(x)\sin(y),$$

takže obě derivace budou nulové pro dvě sady bodů

(1) $\cos(x) = 0$, $\sin(y) = 0$, to je $(x,y) = (\frac{2k+1}{2}\pi, \ell\pi)$, pro libovolné $k, \ell \in \mathbb{Z}$

(2) $\cos(y) = 0$, $\sin(x) = 0$, to je $(x,y) = (k\pi, \frac{2\ell+1}{2}\pi)$, pro libovolné $k, \ell \in \mathbb{Z}$.

Druhé parciální derivace jsou

$$\begin{aligned} Hf(x,y) &= \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}(x,y) \\ &= \begin{pmatrix} -\sin(x)\cos(y) & -\cos(x)\sin(y) \\ -\cos(x)\sin(y) & -\sin(x)\cos(y) \end{pmatrix}. \end{aligned}$$

V našich dvou sadách stacionárních bodů tedy dostáváme následující hessiány:

(1) $Hf(k\pi + \frac{\pi}{2}, \ell\pi) = \pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, přičemž znaménko $-$ nastává, když parity k a ℓ jsou stejné a naopak pro $+$;

(2) $Hf(k\pi, \ell\pi + \frac{\pi}{2}) = \pm\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, přičemž znaménko $-$ nastává, když parity k a ℓ jsou stejné a naopak pro $+$.

Když se nyní podíváme na tvrzení Taylorovy věty pro řád $k = 2$, dostáváme v okolí jednoho ze stacionárních bodů (x_0, y_0)

$$\begin{aligned} &f(x,y) = f(x_0,y_0) + \\ &+ \frac{1}{2}Hf(x_0 + \theta(x-x_0), y_0 + \theta(y-y_0))(x-x_0, y-y_0), \end{aligned}$$

kde Hf nyní vnímáme jako kvadratickou formu vyčíslenou na přírůstku $(x-x_0, y-y_0)$. Protože naše funkce má spojitý hessián (tj. spojité parciální derivace do druhého řádu včetně), a matice hessiánu jsou nedegenerované, nastane lokální maximum tehdy a jen tehdy, když náš bod (x_0, y_0) patří do první skupiny se stejnými paritami k a ℓ. Když budou parity opačné, pak bod z první skupiny bude naopak bodem lokálního minima.

Naopak, hessián u druhé skupiny bodů se vždy vyčíslí kladně na některých přírůstcích a záporně na jiných. Proto se tak bude chovat i celá funkce f v malém okolí daného bodu.

Abychom mohli zformulovat obecné tvrzení o hessiánu a lokálních extrémech ve stacionárních bodech, musíme připomenout diskusi o kvadratických formách v odstavcích 4.31–4.32 v kapitole o afinní geometrii. Zavedli jsme tam pro kvadratickou formu $h : E_n \to \mathbb{R}$ následující přívlastky

- *positivně definitní*, je-li $h(u) > 0$ pro všechny $u \neq 0$
- *positivně semidefinitní*, je-li $h(u) \geq 0$ pro všechny $u \in V$
- *negativně definitní*, je-li $h(u) < 0$ pro všechny $u \neq 0$
- *negativně semidefinitní*, je-li $h(u) \leq 0$ pro všechny $u \in V$
- *indefinitní*, je-li $h(u) > 0$ a $f(v) < 0$ pro vhodné $u, v \in V$.

Zavedli jsme také nějaké metody, které umožňují přímo zjistit, zda daná forma má některý z těchto přívlastků.

Taylorův rozvoj se zbytkem okamžitě dává platnost následující věty:

Věta. *Nechť $f : E_n \to \mathbb{R}$ je dvakrát spojitě diferencovatelná funkce a $x \in E_n$ nechť je stacionární bod funkce f. Potom*

(1) f má v x ostré lokální minimum, je-li $Hf(x)$ positivně definitní,
(2) f má v x ostré lokální maximum, je-li $Hf(x)$ negativně definitní,
(3) f nemá v bodě x lokální extrém je-li $Hf(x)$ indefinitní.

Důkaz. Taylorův rozvoj druhého řádu se zbytkem pro funkci $f(x_1, \dots, x_n)$, bod $x = (x_1, \dots, x_n)$ a přírůstek $v = (v_1, \dots, v_n)$ říká

$$f(x+v) = f(x) + df(x)(v) + \frac{1}{2}Hf(x+\theta \cdot v)(v).$$

Dle předpokladu o nulové hodnotě diferenciálu je tedy

$$f(x+v) = f(x) + \frac{1}{2}Hf(x+\theta \cdot v)(v).$$

Podle našeho předpokladu je kvadratická forma $Hf(x)$ spojitě závislá na bodu x a definitnost, resp. indefinitnost, kvadratických forem je rozhodnutelná podle znaménka hlavních subdeterminantů matice Hf, viz Sylvestrovo kritérium v odstavci 4.32. Samotný determinant je ale coby polynomiální výraz v koeficientech matice spojitou funkcí, proto nenulovost a znaménka zkoumaných determinantů v dostatečně malém okolí bodu x budou stejná jako v bodě x samotném.

Zejména tedy pro pozitivně definitní $Hf(x)$ máme zajištěno, že $f(x+v) > f(x)$ pro dostatečně malá v, jde tedy o ostré minimum funkce f v bodě x. Analogicky pro negativní definitnost. V případě indefinitní formy $Hf(x)$ budou existovat směry v, w ve kterých $f(x+v) > f(x)$ a $f(x+w) < f(x)$ a tedy extrém žádný nenastává. □

Všimněme si, že věta nedává žádný výsledek, pokud je hessián funkce ve zkoumaném bodě degenerovaný a přitom není indefinitní. Důvod je opět stejný jako u funkcí jedné proměnné. V takových případech totiž existují směry, ve kterých první i druhá derivace zmizí a my proto v tomto řádu přiblížení neumíme poznat, zda se funkce bude chovat jako

t^3 nebo jako $\pm t^4$ dokud nespočteme alespoň v potřebných směrech derivace vyšší.

Zároveň si povšimněme, že i v bodech, kde je diferenciál nenulový, má definitnost hessiánu $Hf(x)$ podobné důsledky jako nenulovost druhé derivace u funkce jedné proměnné. Skutečně, výraz

$$z(x+v) = f(x) + df(x)(v)$$

zadává právě tečnou nadrovinu ke grafu funkce f a proto Taylorova věta druhého řádu se zbytkem, tak jak byla využita v důkazu, ukazuje, že při pozitivní definitnosti hessiánu jsou všechny hodnoty funkce f v dostatečně malém okolí bodu x nad hodnotami na tečné nadrovině, tj. celý graf je v dostatečně malém okolí nad tečnou nadrovinou. V případě negativní definitnosti je tomu naopak. U indefinitních hodnot hesiánu opět graf funkce přechází z jedné strany tečné nadroviny na druhou, to se ale obecně děje podél objektů nižší dimenze v tečné nadrovině, nemáme tedy k dispozici přímočaré zobecnění inflexních bodů.

8.9

8.14. Zobrazení. Koncept derivace a diferenciálu lze snadno rozšířit na zobrazení $F : E_n \to E_m$. Při zvolených kartézských souřadnicích na obou stranách je takové zobrazení obyčejná m–tice

$$F(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

funkcí $f_i : E_n \to \mathbb{R}$. Řekneme, že F je *diferencovatelné* nebo *k–krát diferencovatelné zobrazení*, jestliže tuto vlastnost mají všechny funkce $f_1, \dots, f_m$.

Diferenciál a Jacobiho matice

Diferenciály $df_i(x)$ jednotlivých funkcí f_i zobrazení

$$F(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

poskytují lineární přiblížení přírůstků jejich hodnot. Lze proto očekávat, že budou společně dávat také souřadné vyjádření lineárního zobrazení $D^1F(x) : \mathbb{R}^n \to \mathbb{R}^m$ mezi zaměřeními, které bude lineárně aproximovat přírůstky našeho zobrazení. Výsledná matice

$$D^1F(x) = \begin{pmatrix} df_1(x) \\ df_2(x) \\ \vdots \\ df_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}(x)$$

se nazývá *Jacobiho matice zobrazení* F v bodě x.

Lineární zobrazení $D^1F(x)$ definované na přírůstcích $v = (v_1, \dots, v_n)$ pomocí stejně značené Jacobiho matice nazýváme *diferenciál zobrazení* F v bodě x z definičního oboru, jestliže platí

$$\lim_{v \to 0} \frac{1}{\|v\|} \big(F(x+v) - F(x) - D^1F(x)(v) \big) = 0.$$

Přímé použití Věty 8.5 o existenci diferenciálu pro funkce n proměnných na jednotlivé souřadné funkce zobrazení F a sama definice euklidovské vzdálenosti vede k následujícímu tvrzení:

Důsledek. *Nechť $F : E_n \to E_m$ je zobrazení, jehož všechny souřadné funkce mají spojité parciální derivace v okolí bodu $x \in E_n$. Pak existuje diferenciál $D^1F(x)$ zadaný Jacobiho maticí.*

8.9a

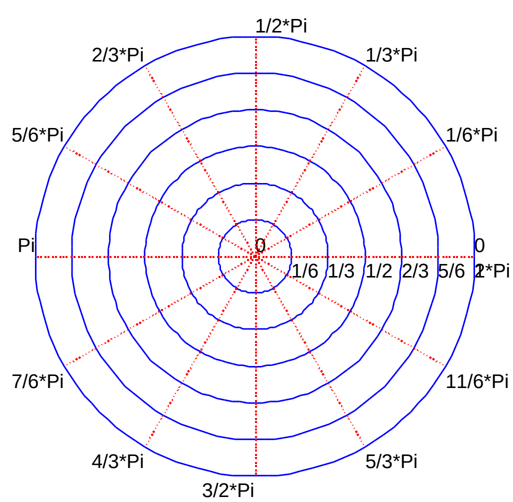
8.15. Transformace. Zobrazení $F : E_n \to E_n$, která mají

inverzní zobrazení $G : E_m \to E_n$ definované na celém svém obrazu, se nazývají *transformace*. Každé takové zobrazení je možné vnímat jako změnu souřadnic. Zpravidla požadujeme, aby F i G byla diferencovatelná.

Stejně jako u vektorových prostorů, volba našeho „pohledu na věc“, tj. volba souřadnic, může zdánlivě zjednodušit nebo zhoršit naše porozumění studovanému objektu. Změnu souřadnic nyní diskutujeme v daleko obecnější formě než jen u afinních zobrazení v kapitole čtvrté. Velice názorný příklad je změna nejobvyklejších souřadnic v rovině na tzv. polární, tj. polohu bodu P zadáváme pomocí jeho vzdálenosti od počátku souřadnic $r = \sqrt{x^2 + y^2}$ a úhlu $\varphi = \arctan(y/x)$ (pokud je $x \neq 0$) mezi spojnicí s počátkem a osou x. Přechod z polárních souřadnic do standardních je

$$P_{\text{polární}} = (r, \varphi) \mapsto (r\cos\varphi, r\sin\varphi) = P_{\text{kartézské}}$$

Je přitom zjevné, že je nutné polární souřadnice vhodně omezit na podmnožinu bodů (r, φ) v rovině, aby existovalo i zobrazení inverzní. Kartézský obraz přímek v polárních souřadnicích s konstantními souřadnicemi r nebo φ je na následujícím obrázku:



Následující věta formuluje velmi užitečné zobecnění pravidla pro derivání složených funkcí jedné proměnné. Je vlastně, až na složitější koncept samotného diferenciálu, úplně stejná, jako už u jedné proměnné viděli. Pro funkce jedné proměnné je totiž Jacobiho matice jediné číslo a to derivace funkce v bodě, násobení Jacobiho matic je tedy prosté násobení derivací vnější a vnitřní složky funkce.) Speciálním případem jsou samozřejmě také vztahy, které jsme odvodili pro derivaci kompozice funkce více proměnných s křivkou.

Diferenciál složeného zobrazení

8.10 **8.16. Věta.** *Nechť $F : E_n \to E_m$ a $G : E_m \to E_r$ jsou dvě diferencovatelná zobrazení, přičemž definiční obor G obsahuje celý obor hodnot F. Pak také složené zobrazení $G \circ F$ je diferencovatelné a jeho diferenciál je v každém bodě z definičního obodu F kompozicí diferenciálů*

$$D^1(G \circ F)(x) = D^1G(F(x)) \circ D^1F(x).$$

Příslušná Jacobiho matice je dána součinem příslušných Jacobiho matic.

Důkaz. V odstavci 8.5 a při důkazu Taylorovy věty jsme odvodili, jak se chová diferencování složených zobrazení vzniklých z funkcí a křivek. Tím jsme dokázali speciální případy této věty s $n = r = 1$. Obecný případ se odvodí prakticky stejným postupem, jen budeme pracovat více s vektory.

Zvolme libovolný pevný přírůstek v a počítejme směrovou derivaci pro kompozici $G \circ F$ v bodě $x \in E_n$. Ve skutečnosti to znamená spočíst postupně diferenciály pro jednotlivé souřadné funkce zobrazení G složené s F. Pišme tedy rovnou jednodušeji $g \circ F$ pro kteroukoliv z nich.

$$d_v(g \circ F)(x) = \lim_{t \to 0} \frac{1}{t}\big(g(F(x + tv)) - g(F(x))\big).$$

Výraz v závorce můžeme ovšem z definice diferenciálu g vyjádřit jako

$$\begin{aligned} g(F(x + tv)) - g(F(x) &= dg(F(x))(F(x + tv) - F(x)) \\ &\qquad + \alpha(F(x + tv) - F(x)), \end{aligned}$$

kde α je funkce definovaná na okolí bodu $F(x)$, která je spojitá a splňuje $\lim_{v \to 0} \frac{1}{\|v\|}\alpha(v) = 0$. Dosazením do rovnosti pro směrovou derivaci dostáváme

$$\begin{aligned} d_v(g \circ F)(x) &= \lim_{t \to 0} \frac{1}{t}\big(dg(F(x))(F(x + tv) - F(x)) \\ &\qquad\qquad + \alpha(F(x + tv) - F(x))\big) \\ &= dg(F(x))\left(\lim_{t \to 0} \frac{1}{t}\left(F(x + tv) - F(x)\right)\right) \\ &\qquad\qquad + \lim_{t \to 0} \frac{1}{t}\left(\alpha(F(x + tv) - F(x))\right) \\ &= dg(F(x)) \circ D^1F(x)(v) + 0, \end{aligned}$$

kde jsme využili skutečnosti, že lineární zobrazení mezi konečněrozměrnými prostory jsou vždy spojitá a vlastnosti funkce α.

Dokázali jsme tedy tvrzení pro jednotlivé funkce $g_1, \dots, g_r$ zobrazení G. Celá věta nyní vyplývá z toho, jak se násobí matice. $\square$

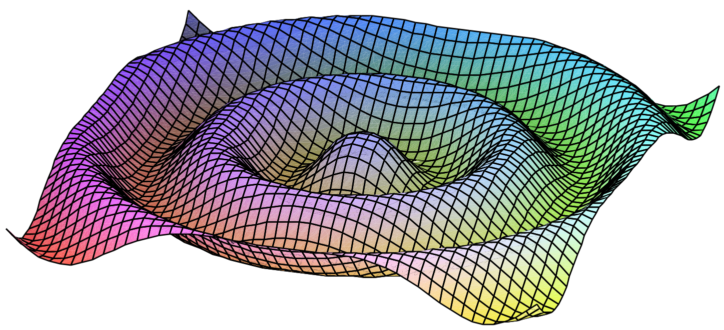
Ilustrujme teď využití konceptu transformace a věty o derivání složených zobrazení na jednoduchém příkladě. Viděli jsme, že polární souřadnice vzniknou z kartézských transformací $F : \mathbb{R}^2 \to \mathbb{R}^2$, kterou v souřadnicích (x, y) a (r, φ) zapíšeme takto (např na definičním oboru všech bodů v prvním kvadrantu roviny mimo body s $x = 0$)

$$r = \sqrt{x^2 + y^2}, \ \varphi = \arctan \frac{y}{x}.$$

Uvažme funkci $g_t : E_2 \to \mathbb{R}$, která má v polárních souřadnicích vyjádření

$$g(r, \varphi, t) = \sin(r - t).$$

Taková funkce nám snad dobře přibližuje vlnění povrchu hladiny po bodovém vzruchu v počátku v čase t, viz obrázek s hodnotou $t = -\pi/2$. Zatímco v polárních souřadnicích bylo snadné ji zadat, v kartézských bychom asi tápali.



Spočtěme nyní derivaci této funkce v kartézských souřadnicích. Použitím naší věty dostaneme

$$\begin{aligned}\frac{\partial g}{\partial x}(x, y, t) &= \frac{\partial g}{\partial r}(r, \varphi)\frac{\partial r}{\partial x}(x, y) + \frac{\partial g}{\partial \varphi}(r, \varphi)\frac{\partial \varphi}{\partial x}(x, y)\\ &= \cos(\sqrt{x^2 + y^2} - t)\frac{x}{\sqrt{x^2 + y^2}} + 0\end{aligned}$$

a podobně

$$\begin{aligned}\frac{\partial g}{\partial y}(x, y, t) &= \frac{\partial g}{\partial r}(r, \varphi)\frac{\partial r}{\partial y}(x, y) + \frac{\partial g}{\partial \varphi}(r, \varphi)\frac{\partial \varphi}{\partial y}(x, y)\\ &= \cos(\sqrt{x^2 + y^2} - t)\frac{y}{\sqrt{x^2 + y^2}}.\end{aligned}$$

8.11

8.17. Věta o inverzním zobrazení. U funkcí jedné proměnné rozhodovala nenulovost první derivace o tom, je-li funkce rostoucí či klesající. Pak takovou musela být i na nějakém okolí zvoleného bodu a tudíž tam existovala i inverzní funkce. Její derivace pak byla převrácenou hodnotou derivace funkce původní.

Když tuto situaci interpretujeme z pohledu zobrazení $E_1 \to E_1$ a lineárních zobrazení $\mathbb{R} \to \mathbb{R}$ coby jejich diferenciálů, je nenulovost nutnou a dostatečnou podmínkou k invertibilitě příslušného diferenciálu. Takto obdržíme tvrzení platné pro konečněrozměrné prostory obecně:

Věta o inverzním zobrazení

Věta. *Nechť $F : E_n \to E_n$ je diferencovatelné zobrazení na nějakém okolí bodu $x_0 \in E_n$ a nechť je Jacobiho matice $D^1 f(x_0)$ invertibilní. Pak na nějakém okolí bodu x_0 existuje inverzní zobrazení F^{-1} a jeho diferenciál v bodě $F(x_0)$ je inverzním zobrazením k diferenciálu $D^1 F(x_0)$, tzn. je zadán inverzní maticí k Jacobiho matici zobrazení F v bodě x_0.*

Důkaz. Nejdříve si zkusme ověřit, že tvrzení je rozumné a očekávatelné. Pokud bychom předpokládali, že inverzní zobrazení existuje a je diferencovatelné v bodě $F(x_0)$, věta o derivování složených funkcí si vynucuje vztah

$$\mathrm{id}_{\mathbb{R}^n} = D^1(F^{-1} \circ F)(x_0) = D^1(F^{-1}) \circ D^1 F(x_0),$$

což ověřuje vztah v závěru věty. Víme proto od začátku, jaký diferenciál pro F^{-1} hledat.

V dalším kroku předpokládejme, že inverzní zobrazení F^{-1} na okolí bodu $F(x_0)$ existuje a je spojité. Budeme v této situaci ověřovat existenci diferenciálu. Z diferencovatelnosti F na okolí x_0 vyplývá, že

$$F(x) - F(x_0) - D^1 F(x_0)(x - x_0) = \alpha(x - x_0)$$

s funkcí $\alpha : \mathbb{R}^n \to 0$ splňující $\lim_{v\to 0} \frac{1}{\|v\|}\alpha(v) = 0$. Pro ověření aproximační vlastnosti lineárního zobrazení $(D^1 F(x_0))^{-1}$ je třeba pouze spočíst následující limitu pro $y = F(x)$ jdoucí k $y_0 = F(x_0)$

$$\lim_{y\to y_0} \frac{1}{\|y - y_0\|}\big(F^{-1}(y) - F^{-1}(y_0) - (D^1 F(x_0))^{-1}(y - y_0)\big).$$

Dosazením z předchozí rovnosti dostáváme

$$\begin{aligned}
&\lim_{y\to y_0} \frac{1}{\|y - y_0\|}\Bigg(x - x_0 - \\
&\qquad (D^1 F(x_0))^{-1}(D^1 F(x_0)(x - x_0) + \alpha(x - x_0))\Bigg) \\
&\qquad = \lim_{y\to y_0} \frac{-1}{\|y - y_0\|}(D^1 F(x_0))^{-1}(\alpha(x - x_0)) \\
&\qquad = (D^1 F(x_0))^{-1} \lim_{y\to y_0} \frac{-1}{\|y - y_0\|}(\alpha(x - x_0)),
\end{aligned}$$

kde poslední rovnost vyplývá ze skutečnosti, že lineární zobrazení mezi konečněrozměrnými prostory jsou vždy spojitá a díky invertibilitě diferenciálu jeho předřazení limitnímu procesu neovlivní ani existenci limity.

Všimněme si, že jsme zdánlivě s důkazem skoro hotoví. Limita na konci našeho výrazu je v důsledku vlastností funkce α nulová, pokud jsou velikosti $\|F(x) - F(x_0)\|$ větší než $C\|x - x_0\|$ pro nějakou konstantu C. To je o trochu silnější vlastnost, než že je F^{-1} spojité, v literatuře se této vlastnosti říká, že je funkce *Lipschitzovsky spojitá*. Zbývá nám tedy už „jenom“ dokázat existenci Lipschitzovsky spojitého inverzního zobrazení k zobrazení F.

Pro další úvahy si zjednodušíme práci převedením obecného případu na o něco jednodušší tvrzení.

Zejména bez újmy na obecnosti lze vhodnou volbou kartézských souřadnic dosáhnout $x_0 = 0 \in R^n$, $y_0 = F(x_0) = 0 \in \mathbb{R}^n$.

Složením zobrazení F s jakýmkoliv lineárním zobrazením G dostateme opět diferencovatelné zobrazení a víme také, jak se změní diferenciál. Volbou $G(x) = (D^1F(0))^{-1}(x)$ dostáváme $D^1(G \circ F)(0) = \mathrm{id}_{\mathbb{R}^n}$. Můžeme tedy zrovna předpokládat

$$D^1F(0) = \mathrm{id}_{\mathbb{R}^n}\,.$$

Uvažme za těchto předpokladů zobrazení $K(x) = F(x) - x$. Toto zobrazení je opět diferencovatelné a jeho diferenciál v bodě 0 je zjevně nulový.

Pro libovolné spojitě diferencovatelné zobrazení K v okolí počátku $\mathbb{R}^n$ platí díky Taylorovu rozvoji prvního řádu se zbytkem jednotlivých souřadných funkcí K_i a díky definici euklidovské vzdálenosti odhad

$$\|K(x) - K(y)\| \le C\sqrt{n}\|x - y\|,$$

kde C je ohraničeno maximem všech absolutních hodnot parciálních derivací v Jacobiho matici zobrazení K na sledovaném okolí.[2]

Protože v našem případě je diferenciál zobrazení K v bodě $x_0 = 0$ nulový, můžeme volbou dostatečně malého okolí U počátku dosáhnout platnosti ohraničení

$$\|K(x) - K(y)\| \le \frac{1}{2}\|x - y\|.$$

Dále dosazením za definici $K(x) = F(x) - x$ a použitím trojúhelníkové nerovnosti $\|(u - v) + v\| \le \|u - v\| + \|v\|$, tj. také $\|u\| - \|v\| \le \|u - v\|$, dostáváme

$$\begin{aligned}\|y - x\| - \|F(x) - F(y)\| &\le \|F(x) - F(y) + y - x\| \\ &\le \frac{1}{2}\|y - x\|.\end{aligned}$$

Odtud konečně

$$\frac{1}{2}\|x - y\| \le \|F(x) - F(y)\|.$$

Tímto odhadem jsme dosáhli opravdu pěkného pokroku: jsou-li na našem malém okolí U počátku $x \neq y$, pak nutně musí být také $F(x) \neq F(y)$. Je tedy naše zobrazení vzájemně

[2] Z této úvahy okamžitě plyne, že funkce, která má spojité parciální derivace na kompaktní množině, je na ní i Lipschitzovsky spojitá.

jednoznačné. Pišme F^{-1} pro jeho inverzi definovanou na obrazu U. Pro ni náš odhad říká

$$\|F^{-1}(x) - F^{-1}(y)\| \leq 2\|x - y\|,$$

je tedy toto zobrazení určitě nejen spojité ale dokonce Lipschitzovsky spojité, tak jak jsme v předchozí části důkazu potřebovali.

Zdánlivě jsme tedy již úplně hotoví (s důkazem), to ale

není pravda. Abychom skutečně dokončili důkaz, musíme ukázat, že je zobrazení F zúžené na dostatečně malé okolí nejen vzájemně jednoznačné, ale že také zobrazuje otevřené okolí nuly na otevřené okolí nuly.[3]

Zvolme si δ tak malé, aby okolí $V = \mathcal{O}_\delta(0)$ leželo v U včetně své hranice a zároveň aby Jacobiho matice zobrazení F byla na celém V invertibilní. To je jistě možné, protože determinant je spojité zobrazení. Označme B hranici množiny V (tj. příslušnou sféru). Protože je B kompaktní a F spojité, má funkce

$$\rho(x) = \|F(x)\|$$

na B maximum i minimum. Označme $a = \frac{1}{2}\min_{x \in B} \rho(x)$ a uvažujme libovolné $y \in \mathcal{O}_a(0)$. Samozřejmě je $a > 0$. Chceme ukázat, že existuje alespoň jedno $x \in V$ takové, že $y = F(x)$, čímž bude celá věta o inverzní funkci dokázána.

Za tímto účelem uvažme funkci (y je náš pevně zvolený bod)

$$h(x) = \|F(x) - y\|^2.$$

Opět obraz $h(V) \cup h(B)$ musí mít minimum. Ukážeme nejprve, že toto minimum nemůže nastat pro $x \in B$. Platí totiž $F(0) = 0$ a proto $h(0) = \|y\| < a$. Zároveň podle naší definice a je pro $y \in \mathcal{O}_a(0)$ vzdálenost y od $F(x)$ pro $x \in B$ alespoň a (protože a jsme volili jako polovinu minima z velikosti $F(x)$ na hranici). Minimum tedy nastává uvnitř V a musí být ve stacionárním bodě z funkce h. To ale znamená že pro všechna $j = 1, \dots, n$ platí

$$\frac{\partial h}{\partial x^j}(z) = \sum_{i=1}^{n} 2(f_i(z) - y_i)\frac{\partial f_i}{\partial x_j}(z) = 0.$$

Na tento systém rovnic se můžeme dívat jako na systém lineárních rovnic s proměnnými $\xi_i = f_i(z) - y_i$ a koeficienty zadanými dvojnásobkem Jacobiho matice $D^1F(z)$. Pro každé $z \in V$ má takový systém ovšem pouze jedno řešení a to je nulové, protože Jacobiho matice je podle našeho předpokladu invertibilní.

Tím jsme našli hledaný bod $x = z \in V$ splňující pro všechna $i = 1, \dots, n$ rovnost $f_i(z) = y_i$. $\square$

[3] V literatuře lze snadno dohledat příklady zobrazení, která třeba spojitě a bijektivně zobrazí úsečku na čtverev apod.

8.12

8.18. Věta o implicitní funkci. Naším dalším cílem je využít větu o inverzním zobrazení pro práci s implicitně definovanými funkcemi. Pro začátek uvažujme diferencovatelnou funkci $F(x, y)$ definovanou v rovině E_2 a hledejme body (x, y), ve kterých platí $F(x, y) = 0$.

Příkladem může být třeba obvyklá (implicitní) definice přímek a kružnic:

$$F(x, y) = ax + by + c = 0$$
$$F(x, y) = (x - s)^2 + (y - t)^2 - r^2 = 0,\ r > 0.$$

Zatímco v prvém případě je (při $b \neq 0$) předpisem zadaná funkce

$$y = f(x) = -\frac{a}{b}x - \frac{c}{b}$$

pro všechna x, ve druhém případě můžeme pro libovolný bod (x_0, y_0) splňující rovnici kružnice a takový, že $y_0 \neq t$ (to jsou totiž krajní body kružnice ve směru souřadnice x), najít okolí bodu x_0, na kterém bude buď

$$y = f(x) = t + \sqrt{(x - s)^2 - r}$$

nebo

$$y = f(x) = t - \sqrt{(x - s)^2 - r},$$

podle toho na kterou polokružnici patří bod (x_0, y_0). Při načrtnutí obrázku je důvod zřejmý – nemůžeme chtít pomocí funkce $y = f(x)$ postihnout horní i dolní půlkružnici zároveň. Zajímavější jsou krajní body intervalu $[s - r, s + r]$. Ty také vyhovují rovnici kružnice, platí v nich ale $F_y(s \pm r, t) = 0$, což vystihuje polohu tečny ke kružnici v těchto bodech rovnoběžnou s osou y. V těchto bodech skutečně neumíme najít okolí, na němž by kružnice byla popsána jako funkce $y = f(x)$.

Navíc umíme i derivace naší funkce $y = f(x) = t + \sqrt{(x - s)^2 - r^2}$, tam kde je definována, vyjádřit pomocí parciálních derivací funkce F:

$$f'(x) = \frac{1}{2}\frac{2(x - s)}{\sqrt{(x - s)^2 - r^2}} = \frac{x - s}{y - t} = -\frac{F_x}{F_y}.$$

Když prohodíme roli proměnných x a y a budeme chtít najít závislost $x = f(y)$ takovou, aby $F(f(y), y) = 0$, pak v okolí bodů $(s \pm r, t)$ bez problémů uspějeme. Všimněme si, že v těchto bodech je parciální derivace F_x nenulová.

Naše pozorování tedy (pro pouhé dva příklady) říká: pro funkci $F(x, y)$ a bod $(a, b) \in E_2$ takový, že $F(a, b) = 0$, umíme jednoznačně najít funkci $y = f(x)$ splňující $F(x, f(x)) = 0$, pokud je $F_y(a, b) \neq 0$. V takovém případě umíme i vypočíst $f'(a) = -F_x(a, b)/F_y(a, b)$. Dokážeme, že ve skutečnosti toto tvrzení platí vždy. Poslední tvrzení o derivaci přitom je dobře zapamatovatelné (a při pečlivém vnímání věcí i pochopitelné) z výrazu pro diferenciál funkce $g(x) = F(x, y(x))$ a diferenciál $dy = f'(x)dx$

$$0 = dg = F_x dx + F_y dy = (F_x + F_y f'(x))dx.$$

Obdobně bychom mohli pracovat s implicitními výrazy $F(x, y, z) = 0$, přičemž můžeme hledat funkci $g(x, y)$ takovou, že $F(x, y, g(x, y)) = 0$. Jako příklad uvažme třeba funkci $f(x, y) = x^2 + y^2$, jejímž grafem je rotační paraboloid s počátkem v bodě $(0, 0)$. Ten můžeme implicitně zadat také rovnicí

$$0 = F(x, y, z) = z - x^2 - y^2.$$

Než sformulujeme výsledek rovnou pro obecnou situaci, všimněme si ještě, jaké dimenze se mohou/mají v problému vyskytovat. Pokud bychom pro tuto funkci F chtěli najít křivku $c(x) = (c_1(x), c_2(x))$ v rovině takovou, že

$$F(x, c(x)) = F(x, c_1(x), c_2(x)) = 0,$$

pak to jistě budeme umět (dokonce pro všechny počáteční podmínky $x = a$) také, ale výsledek nebude jednoznačný pro danou počáteční podmínku. Stačí totiž uvážit libovolnou křivku na rotačním paraboloidu, jejíž průmět do první souřadnice má nenulovou derivaci. Pak považujeme x za parametr křivky a za $c(x)$ zvolíme její průmět do roviny yz.

Očekáváme tedy, že jedna funkce $m + 1$ proměnných

zadává implicitně nadplochu v $\mathbb{R}^{m+1}$, kterou chceme vyjádřit alespoň lokálně jako graf jedné funkce v m proměnných. Lze očekávat, že n funkcí v $m + n$ proměnných bude zadávat průnik n nadploch v $\mathbb{R}^{m+n}$, což je ve „většině" případů m–rozměrný objekt.

Uvažujme proto diferencovatelné zobrazení

$$F = (f_1, \dots, f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n.$$

Jacobiho matice tohoto zobrazení bude mít n řádků a $m + n$ sloupců a můžeme si ji symbolicky zapsat jako

$$\begin{aligned} D^1F &= (D^1_x F, D^1_y F) \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_m} & \frac{\partial f_1}{\partial x_{m+1}} & \dots & \frac{\partial f_1}{\partial x_{m+n}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_m} & \frac{\partial f_n}{\partial x_{m+1}} & \dots & \frac{\partial f_n}{\partial x_{m+n}} \end{pmatrix}, \end{aligned}$$

kde $(x_1, \dots, x_{m+n}) \in \mathbb{R}^{m+n}$ zapisujeme jako $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$, $D^1_x F$ je matice s n řádky a prvními m sloupci v Jacobiho matici, zatímco $D^1_y F$ je čtvercová matice řádu n se zbylými sloupci. Vícerozměrnou analogií k předchozí úvaze s nenulovou parciální derivací podle y je požadavek, aby matice $D^1_y F$ byla invertibilní.

Věta o implicitním zobrazení

Věta. *Nechť $F : \mathbb{R}^{m+n} \to \mathbb{R}^n$ je diferencovatelné zobrazení na otevřeném okolí bodu $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$, ve kterém je $F(a, b) = 0$ a $\det D^1_y F \neq 0$. Potom existuje diferencovatelné zobrazení $G : \mathbb{R}^m \to \mathbb{R}^n$ definované na nějakém okolí U bodu $a \in \mathbb{R}^m$ s obrazem $G(U)$, který obsahuje bod b, a takové, že $F(x, G(x)) = 0$ pro všechny $x \in U$.*

Navíc je Jacobiho matice D^1G zobrazení G na okolí bodu a zadána součinem matic

$$D^1G(x) = -(D^1_y F)^{-1}(x, G(x)) \cdot D^1_x F(x, G(x)).$$

DŮKAZ. Pro zvýšení srozumitelnosti uvedeme napřed

kompletní důkaz pro nejjednodušší případ rovnice $F(x, y) = 0$ s funkcí F dvou proměnných. Bude zdánlivě složitý, protože jej schválně vedeme tak, jak jej bude možné použít i pro obecné dimenze z věty. Rozšíříme funkci F na

$$\tilde{F} : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x, F(x, y)).$$

Jacobiho matice zobrazení $\tilde{F}$ je

$$D^1\tilde{F}(x, y) = \begin{pmatrix} 1 & 0 \\ F_x(x, y) & F_y(x, y) \end{pmatrix}.$$

Z předpokladu $F_y(a, b) \neq 0$ vyplývá, že totéž platí i na nějakém okolí bodu (a, b) a tedy je na tomto okolí funkce $\tilde{F}$ invertibilní podle věty o inverzním zobrazení. Vezměme tedy jednoznačně definované a diferencovatelné inverzní zobrazení $\tilde{F}^{-1}$ na nějakém okolí bodu $(a, 0)$.

Nyní označme $\pi : \mathbb{R}^2 \to \mathbb{R}$ projekci na druhou souřadnici a uvažujme funkci $f(x) = \pi \circ \tilde{F}^{-1}(x, 0)$. To je dobře definovaná a diferencovatelná funkce. Máme ověřit, že následující výraz

$$F(x, f(x)) = F(x, \pi(\tilde{F}^{-1}(x, 0)))$$

bude na okolí bodu $x = a$ nulový. Přitom z definice $\tilde{F}(x, y) = (x, F(x, y))$ vyplývá, že i její inverze musí mít tvar $\tilde{F}^{-1}(x, y) = (x, \pi\tilde{F}^{-1}(x, y))$. Můžeme proto pokračovat v předchozím výpočtu:

$$\begin{aligned} F(x, f(x)) &= \pi(\tilde{F}(x, \pi(\tilde{F}^{-1}(x, 0)))) = \\ &= \pi(\tilde{F}(\tilde{F}^{-1}(x, 0))) = \pi(x, 0) = 0. \end{aligned}$$

Tím máme dokázánu první část věty a zbývá spočíst derivaci funkce $f(x)$. Tuto derivaci můžeme odečíst opět z věty o inverzním zobrazení pomocí matice $(D^1\tilde{F})^{-1}$.

Následující výsledek je snadné ověřit roznásobením matic. (Spočíst lze také přímo explicitní formulí pro inverzní matici s pomocí determinantu a algebraicky adjungované matice, viz odstavec 2.23)

$$\begin{pmatrix} 1 & 0 \\ F_x(x, y) & F_y(x, y) \end{pmatrix}^{-1} = (F_y(x, y))^{-1} \begin{pmatrix} F_y(x, y) & 0 \\ -F_x(x, y) & 1 \end{pmatrix}.$$

Dle definice $f(x) = \pi\tilde{F}^{-1}(x, 0)$ nás z této matice zajímá první položka na druhém řádku, která je právě Jacobiho maticí $D^1 f$. V našem jednoduchém případě je to právě požadovaný skalár $-F_x(x, f(x))/F_y(x, f(x))$.

Obecný důkaz je bezezbytku stejný, není v něm potřeba

změnit žádný z uvedených vztahů (všechny položky v nich jen dostanou vektorový smysl), kromě posledního výpočtu derivace funkce f, kde místo jednotlivých parciálních derivací budou vystupovat příslušné části Jacobiho matice $D^1_x F$ a $D^1_y F$. Samozřejmě je přitom třeba místo se skaláry pracovat s vektory a maticemi.

Pro výpočet Jacobiho matice zobrazení G opět použijeme výpočtu inverzní matice, není ale až tak vhodné přímo využít postupu z odstavce 2.23. Snadnější je nechat se přímo inspirovat případem v dimenzi $m+n=2$, označit si matici

$$(D^1\tilde{F}^{-1}) = \begin{pmatrix} \mathrm{id}_{\mathbb{R}^m} & 0 \\ D_x^1 F(x,y) & D_y^1 F(x,y) \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

s bloky danými dělením na m a n řádků i sloupců (tj. např. A má rozměr $m \times m$, zatímco C je rozměru $n \times m$) a přímo spočíst matice A, B, C, D z definiční rovnosti pro inverzi:

$$\begin{pmatrix} \mathrm{id}_{\mathbb{R}^m} & 0 \\ D_x^1 F(x,y) & D_y^1 F(x,y) \end{pmatrix} \cdot \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \mathrm{id}_{\mathbb{R}^m} & 0 \\ 0 & \mathrm{id}_{\mathbb{R}^n} \end{pmatrix}.$$

Zjevně odtud plyne $A = \mathrm{id}_{\mathbb{R}^m}$, $B = 0$, $D = (D_y^1 F)^{-1}$ a konečně $D_x^1 F + D_y^1 F \cdot C = 0$. Z poslední rovnosti pak dostáváme požadovaný vztah

$$D^1 G = C = -(D_y^1 F)^{-1} \cdot D_x^1 F.$$

Tím je věta dokázána. □

8.13

8.19. Gradient funkce. Jak jsme viděli v minulém odstavci, je-li F spojitě diferencovatelná funkce n proměnných, zadává předpis $F(x_1, \dots, x_n) = b$ s nějakou pevnou hodnotou $b \in \mathbb{R}$ podmnožinu $M \subset \mathbb{R}^n$, která mívá vlastnosti $(n-1)$–rozměrné nadplochy.

Přesněji řečeno, pokud je vektor parciálních derivací

$$D^1 F = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

nenulový, můžeme lokálně množinu M popsat jako graf spojitě diferencovatelné funkce v $n-1$ proměnných. Hovoříme v této souvislosti také o *úrovňových množinách* M_b. Vektor $D^1 F \in \mathbb{R}^n$ se nazývá *gradient funkce* F. V technické a fyzikální literatuře se často zapisuje také jako grad F.

Protože je M_b zadáno pomocí konstantní hodnoty funkce F, budou derivace křivek ležících v M mít jistě tu vlastnost, že na nich bude diferenciál dF vždy vyčíslen nulově – skutečně, pro každou takovou křivku bude $F(c(t)) = b$ a tedy i

$$\frac{d}{dt} F(c(t)) = dF(c'(t)) = 0.$$

Naopak uvažme obecný vektor $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ a velikost příslušné směrové derivace

$$|d_v F| = \left| \frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_n} v_n \right| = \cos\varphi \|D^1 F\| \|v\|$$

kde φ je odchylka vektoru v od gradientu F, viz pojednání o odchylkách vektorů a přímek ve čtvrté kapitole (definice 4.18).

Odtud ovšem vyplývá, že nulové jsou právě ty směrové derivace, které jsou kolmé na gradient, zatímco směr zadaný gradientem je právě ten směr, ve kterém funkce f nejrychleji roste.

Je tedy zřejmé, že tečná rovina k neprázdné úrovňové množině M_b v okolí jejího bodu s nenulovým gradientem

D^1F je určena ortogonálním doplňkem ke gradientu a samotný gradient je tzv. *normálovým vektorem* nadplochy M_b.

Např. pro sféru v $\mathbb{R}^3$ o poloměru $r > 0$ a středu (a, b, c) zadanou rovnicí

$$F(x, y, z) = (x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$$

dostáváme normálové vektory v bodě $P = (x_0, y_0, z_0)$ jako nenulový násobek gradientu, tj. násobek průvodiče

$$D^1F = (2(x_0 - a), 2(y_0 - b), 2(z_0 - c)),$$

a tečné vektory budou právě všechny vektory kolmé na gradient. Implicitně proto jde vždy tečnou rovinu ke sféře v bodě P popsat s pomocí gradientu rovnicí

$$0 = (x_0 - a)(x - x_0) + (y_0 - b)(y - y_0) + (z_0 - c)(z - z_0).$$

To je speciální případ obecné formule:

Tečná nadrovina implicitně zadané nadplochy

Věta. *Pro funkci $F(x_1, \dots, x_n)$ v n proměnných a bod $P = (a_1, \dots, a_n)$ v úrovňové množině M_b funkce F, v jehož okolí je M_b grafem funkce $(n-1)$ proměnných, je implicitní rovnice pro tečnou nadrovinu k M_b*

$$0 = \frac{\partial f}{\partial x_1}(P) \cdot (x_1 - a_1) + \dots + \frac{\partial f}{\partial x_n}(P) \cdot (x_n - a_n).$$

Důkaz. Tvrzení je zřejmé z předchozího výkladu. Tečná nadrovina totiž musí být $(n-1)$–rozměrná, její zaměření je proto zadané jako jádro lineární formy dané gradientem (nulové hodnoty příslušného lineárního zobrazení $\mathbb{R}^n \to \mathbb{R}$ zadaného násobení sloupce souřadnic řádkovým vektorem grad F). Zvolený bod P přitom naší rovnici zjevně vyhovuje. □

8.13a

8.20. Model osvětlení 3D objektů. Uvažujme osvětlení 3D objektu, kde známe směr v dopadu světla na 2D povrch tohoto objektu, tj. množinu M zadanou implicitně nějakou rovnicí $F(x, y, z) = 0$. Intenzitu osvětlení bodu $P \in M$ definujme jako $I \cos\varphi$, kde φ je úhel mezi normálou k M a vektorem opačným ke směru toku světla.. Jak jsme viděli, normála je určena gradientem funkce F. Znaménko našeho výrazu pak bude označovat, kterou stranu plochy osvětlujeme.

Uvažujme např. osvětlení o intezitě I_0 ve směru vektoru $v = (1, 1, -1)$ (tj. „šikmo dolů“) a za objekt zvolme kouli zadanou rovnicí $F(x, y, z) = x^2 + y^2 + z^2 - 1 \le 0$. Pro povrchový bod $P = (x, y, z) \in M$ proto dostaneme intenzitu

$$I(P) = \frac{\operatorname{grad} F \cdot v}{\|\operatorname{grad} F\| \|v\|} I_0 = \frac{-2x - 2y + 2z}{2\sqrt{3}} I_0.$$

Všimněme si, že dle očekávání je maximální (plnou) intenzitou I_0 osvětlen bod $P = \frac{1}{\sqrt{3}}(-1, -1, 1)$ na povrchu koule.

8.14 **8.21. Tečné a normálové prostory.** Přejděme nyní s našimi

úvahami o tečnách a normálách k obecným dimenzím. Máme-li zobrazení $F : \mathbb{R}^{m+n} \to \mathbb{R}^n$, tj. n rovnic pro $n + m$ proměnných

$$f_i(x_1, \dots, x_{m+n}) = b_i, \quad i = 1, \dots, n,$$

pak, za podmínek věty o implicitní funkci je množina všech řešení $(x_1, \dots, x_{m+n}) \in \mathbb{R}^{m+n}$ alespoň lokálně grafem zobrazení $G : \mathbb{R}^m \to \mathbb{R}^n$.

Pro pevnou volbu $b = (b_1, \dots, b_n)$ je samozřejmě množinou všech řešení průnik nadploch $M(b_i, f_i)$ příslušejících jednotlivým funkcím f_i. Totéž musí platit pro tečné směry, zatímco normálové směry jsou generovány jednotlivými gradienty. Proto je-li D^1F Jacobiho matice zobrazení implicitně zadávajícího množinu M s bodem $P = (a_1, \dots, a_{m+n}) \in M$, v jehož okolí je M grafem zobrazení,

$$D^1F = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_{m+n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_{m+n}} \end{pmatrix},$$

potom bude afinní podprostor v $\mathbb{R}^{m+n}$ obsahující právě všechny tečny procházející bodem P dán implicitně rovnicemi:

$$\begin{aligned} 0 &= \frac{\partial f_1}{\partial x_1}(P) \cdot (x_1 - a_1) + \dots + \frac{\partial f_1}{\partial x_n}(P) \cdot (x_{m+n} - a_{m+n}) \\ &\vdots \\ 0 &= \frac{\partial f_n}{\partial x_1}(P) \cdot (x_1 - a_1) + \dots + \frac{\partial f_n}{\partial x_n}(P) \cdot (x_{m+n} - a_{m+n}). \end{aligned}$$

Tento podprostor se nazývá *tečný prostor* k (implicitně zadané) ploše M v bodě P.

Normálový prostor v bodě P je afinní podprostor generovaný bodem P a gradienty všech funkcí $f_1, \dots, f_n$ v bodě P, tj. řádky Jacobiho matice D^1F.

Jako jednoduchý příklad si spočtěme tečnu a normálový prostor ke kuželosečce v $\mathbb{R}^3$. Uvažujme rovnici kuželu s vrcholem v počátku

$$0 = f(x, y, z) = z - \sqrt{x^2 + y^2}$$

a rovinu zadanou

$$0 = g(x, y, z) = z - 2x + y + 1.$$

Bod $P = (1, 0, 1)$ patří jak kuželu tak rovině a průnik M těchto dvou ploch je křivka (namalujte si obrázek). Její tečnou v bodě P bude přímka zadaná rovnicemi

$$\begin{aligned}
0 &= -\left.\frac{1}{2\sqrt{x^2+y^2}}2x\right|_{x=1,y=0} \cdot (x-1)\\
&\quad -\left.\frac{1}{2\sqrt{x^2+y^2}}2y\right|_{x=1,y=0} \cdot y + 1\cdot(z-1)\\
&= -x+z\\
0 &= -2(x-1)+y+(z-1) = -2x+y+z+1,
\end{aligned}$$

zatímco rovina kolmá k naší křivce bodem P bude parametricky dána výrazem

$$(1,0,1)+\tau(-1,0,1)+\sigma(-2,1,1)$$

s parametry τ a σ.

8.15 **8.22. Vázané extrémy.** Nyní se dostáváme k první opravdu vážné aplikaci diferenciálního počtu více proměnných. Typickou úlohou optimalizace nebo řízení je najít extrémy hodnot závisejících na několika (ale konečně mnoha) parametrech, ovšem za nějakých dalších podmínek na vzájemné vztahy parametrů.

Velice často má řešená úloha $m+n$ parametrů, které jsou vázány n podmínkami. V našem jazyce diferenciálního počtu tedy hledáme extrémy diferencovatelné funkce h na množině bodů M zadaných implicitně rovnicí $F(x_1,\dots,x_{m+n})=0$. K tomu již máme připraveny účinné postupy.

Pokud je M ve všech svých bodech grafem hladkého zobrazení v m proměnných, musí být každý extrém $P\in M$ stacionárním bodem, tj. pro každou křivku $c(t)\subset M$ procházející přes $P=c(0)$ musí být $h(c(t))$ extrémem pro tuto funkci jedné proměnné. Proto také musí být derivace

$$\frac{d}{dt}h(c(t))_{|t=0} = d_{c'(0)}h(P) = dh(P)(c'(0)) = 0.$$

To ale znamená, že diferenciál funkce h se v bodě P nuluje na všech tečných přírůstcích k M v bodě P. Tato vlastnost je ekvivalentní tvrzení, že gradient h leží v normálovém podprostoru (přesněji v jeho zaměření). Takové body $P\in M$ budeme nazývat *stacionární body* funkce H vzhledem k vazbám F.

Jak jsme viděli v minulém odstavci, normálový prostor k naší množině M je generován řádky Jacobiho matice zobrazení F a stacionární body jsou proto ekvivalentně určeny následujícím tvrzením:

Metoda Lagrangeových multiplikátorů

Věta. *Nechť $F=(f_1,\dots,f_n):\mathbb{R}^{m+n}\to\mathbb{R}^n$ je diferencovatelná v okolí bodu P, $F(P)=0$. Dále nechť M je zadána implicitně rovnicí $F(x,y)=0$ a hodnost matice D^1F v bodě P je n. Pak P je stacionárním bodem spojitě diferencovatelné funkce $h:\mathbb{R}^{m+n}\to\mathbb{R}$ vzhledem k podmínkám F, právě když existují reálné parametry $\lambda_1,\dots,\lambda_n$ takové, že*

$$\operatorname{grad} h = \lambda_1 \operatorname{grad} f_1 + \cdots + \lambda_n \operatorname{grad} f_n.$$

Všimněme si, že *metoda Langrangeových multiplikátorů* je algoritmická. Podívejme se nejprve na

počty neznámých a rovnic: gradienty jsou vektory o $m+n$ souřadnicích, tedy požadavek z věty dává $m+n$ rovnic. Jako proměnné máme jednak souřadnice $x_1, \dots, x_{m+n}$ hledaných stacionárních bodů P vzhledem k vazbám, ale navíc také n parametrů λ_i v hledané lineární kombinaci. Zbývá však požadavek, že hledaný bod P patří implicitně zadané množině M, což představuje dalších n rovnic. Celkem tedy máme $n+m$ rovnic pro $n+m$ proměnných a proto lze očekávat, že řešením bude diskrétní množina bodů P (tj. každý z nich bude izolovaným bodem).

8.15a

8.23. Nerovnost mezi aritmetickým a geometrickým průměrem. Jako příklad praktického použití metody Lagrangeových multiplikátorů dokážeme nerovnost

$$\frac{1}{n}(x_1 + \dots + x_n) \geq \sqrt[n]{x_1 \cdots x_n}$$

pro jakýchkoliv n kladných čísel $x_1, \dots, x_n$, přičemž rovnost nastane, právě když jsou si všechna x_i rovna.

Uvažme tedy součet $x_1 + \dots + x_n = c$ jako vazebnou podmínku pro nějakou blíže neurčenou nezápornou konstantu c. Budeme hledat maxima a minima funkce

$$f(x_1, \dots, x_n) = \sqrt[n]{x_1 \cdots x_n}$$

za naší vazební podmínky a přepdokladu $x_1 > 0, \dots, x_n > 0$.

Normálový vektor k nadrovině definované podmínkou je $(1, \dots, 1)$. Extrém funkce f tedy může nastat pouze v bodech, kdy je její gradient násobkem tohoto normálového vektoru. Pro hledané body tedy dostáváme soustavu rovnic

$$\frac{1}{n}\frac{1}{x_i}\sqrt[n]{x_1 \cdots x_n} = \lambda,$$

pro $i = 1, \dots, n$ a $\lambda \in \mathbb{R}$.

Tato soustava má zjevně na zkoumané množině jediné řešení $x_1 = \dots = x_n$. Pokud bychom uvažovali i nulové hodnoty x_i, byla by naše množina M zadaná omezením kompaktní a proto by na ní musela mít funkce f jak maximum, tak minimum. Minimum však zjevně dosahuje, právě když je některá zhodnot x_i nulová, v našem bodě s $x_i = \frac{c}{n}$, $i = 1, \dots, n$, nabývá tedy nutně ostrého maxima.

Ve všech ostatních bodech s daným součtem souřadnic c je pak hodnotota jejich geometrického průměru menší a nerovnost je dokázána.