

# MASARYKOVA UNIVERZITA

Přírodovědecká fakulta



Diplomová práce

Pavel Valkoun

## Numerické metody nepodmíněné minimalizace

Vedoucí práce: prof. RNDr. Ondřej Došlý, DrSc.

Studijní program Aplikovaná matematika

Studijní obor Matematika-ekonomie

2010



Masarykova univerzita

Přírodovědecká fakulta



## ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Pavel Valkoun**

Studijní program - obor: **Aplikovaná matematika - Matematika - ekonomie**

Ředitel Ústavu matematiky a statistiky PŘF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje diplomovou práci s tématem:

### **Numerické metody nepodmíněné minimalizace**

#### **Numerical method of unconstrained minimization**

*Oficiální zadání:* Vypracujte text zaměřený na numerické metody hledání nepodmíněných extrémů funkcí více proměnných. Nejprve se zaměřte na obecnou charakterizaci těchto metod a potom se věnujte speciálním technikám, zejména metodě proměnné metriky, metodě sdružených směrů a modifikacemi těchto metod.

#### *Literatura:*

*Hamala, Man. Nelineárne programovanie. 2. dopl. vyd. Bratislava : Alfa, 1976. 240 s.*

*Künzi, Hans Paul - Krelle, Whelm - Oettli, Werner. Nichtlineare Programmierung. Berlin : Springer-Verlag, 1962. 221 s.*

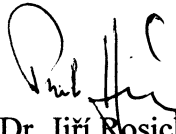
*Došlý, Ondřej. Základy konvexní analýzy a optimalizace v  $R^n$ . 1. vyd. Brno : Masarykova univerzita, 2005. viii, 185. ISBN8021039051.*

*Vedoucí diplomové práce:* prof. RNDr. Ondřej Došlý, DrSc.

*Datum zadání diplomové práce:* prosinec 2008

*Datum odevzdání diplomové práce:* dle harmonogramu ak. roku 2009/2010

V Brně dne 1. 12. 2008

v. r.   
prof. RNDr. Jiří Rosický, DrSc.  
ředitel Ústavu matematiky a statistiky

Zadání diplomové práce převzal dne: **17.2.2009**

Podpis studenta



Tímto bych rád poděkoval prof. RNDr. Ondřeji Došlému, DrSc. a RNDr. Ladislavu Adamcovi, CSc. za odborné vedené, pomoc a rady při vytváření mé práce.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně za použití uvedených pramenů a literatury.

V Brně, dne 13. 5. 2010

Pavel Valkoun

# Abstrakt

Název práce: Numerické metody nepodmíněné minimalizace

Autor: Pavel Valkoun

Ústav matematiky a statistiky Přírodovědecké fakulty MU

Vedoucí diplomové práce: prof. RNDr. Ondřej Došlý, DrSc., RNDr. Ladislav Adamec, CSc.

Abstrakt: Text je zaměřený na popis numerických metod hledání nepodmíněných extrémů funkcí více proměnných. Za tímto účelem je popsána teorie týkající se optimalizace a jsou ukázány některé metody jednorozměrné minimalizace. V druhé části práce jsou probírány samotné metody vícerozměrné minimalizace a jejich modifikace.

Klíčová slova: minimalizační metody, numerické metody, nepodmíněná optimalizace

Title: Numerical method of unconstrained minimization

Author: Pavel Valkoun

Department of Mathematics and Statistics, Faculty of Science, MU

Supervisor: prof. RNDr. Ondřej Došlý, DrSc., RNDr. Ladislav Adamec, CSc.

Abstract: The text gives description of numerical methods of unconstrained minimization of functions of several variables. For this reason the first part of the text contains basic theory of optimization as well as some methods of one-dimensional minimization. The methods of minimization of functions containing more than one variable and modifications of these methods are described in the second part of the text.

Keywords: minimization, numerical methods, unconstrained optimization

# Obsah

<b>Předmluva</b>	<b>6</b>
<b>1 Koncept metod nepodmíněné minimalizace</b>	<b>7</b>
<b>2 Metody jednorozměrné optimalizace</b>	<b>9</b>
2.1 Komparativní metody . . . . .	10
2.1.1 Rovnoměrné dělení intervalu . . . . .	10
2.1.2 Fibonacciho metoda . . . . .	13
2.1.3 Metoda zlatého řezu . . . . .	19
2.2 Gradientní metody . . . . .	22
2.2.1 Metoda tečen (Newtonova metoda) . . . . .	22
2.3 Interpoláčn� metody . . . . .	26
2.3.1 Metoda kvadratick� interpolace . . . . .	26
<b>3 Metody v�cerozm�rn� optimalizace</b>	<b>32</b>
3.1 Komparativn� metody . . . . .	33
3.1.1 Metoda Hooka a Jeevese . . . . .	33
3.2 Gradientn� metody . . . . .	36
3.2.1 Metoda nejrychlejs�ho sp�du . . . . .	37
3.2.2 Metoda nejrychlejs�ho sp�du pro kvadratick� funkce . . . . .	39
3.2.3 Metoda paraleln�ch tečen . . . . .	46
3.2.4 Metoda nejrychlejs�ho sp�du s droben�m kroku . . . . .	48
3.3 Metody sdruzen�ch sm�r� . . . . .	51
3.3.1 Metoda sdruzen�ch sm�r� pro kvadratick� funkce . . . . .	51
3.3.2 Metoda sdruzen�ch sm�r� pro nekvadratick� funkce . . . . .	59
3.3.3 Metoda sdruzen�ch sm�r� nult�ho r�du . . . . .	62
3.4 Newtonovsk� metody . . . . .	65
3.4.1 Newtonova Metoda . . . . .	65
3.4.2 Metoda prom�nn� metriky . . . . .	70
<b>Seznam pouz�t� literatury</b>	<b>77</b>
<b>Rejst�r�k pojm�</b>	<b>78</b>
<b>Rejst�r�k prom�nn�ch</b>	<b>79</b>

---

# Předmluva

Ačkoliv se to patrně většině čtenářů několika stran tohoto díla nemusí jevit jako něco z reálného života, tak neustálé optimalizování je pro lidskou mysl zcela přirozené. Naprosto triviální situace, o kterých ani nepřemýšlíme jako o matematice, jsou ve své podstatě minimalizační nebo maximalizační procesy. Jdeme-li do obchodu cokoli koupit, vybrat dárek pro někoho blízkého, nebo když jen zvažujeme, jak strávíme volný čas, tak vlastně v naší hlavě probíhají série optimalizací a porovnávání výsledků.

Naše myšlení je však zároveň geniální i omezené. Dokáže obejít mnohdy složité matematické postupy a říci nám, která volba je pro nás zřejmě nejvýhodnější. Bez použití tužky, papíru nebo počítače a mnohdy tak automaticky, že si ani neuvědomíme, že jsme vůbec o něčem rozhodovali. Avšak k tomu, aby naše optimalizace byly efektivní, potřebujeme již dříve získané zkušenosti. Jak by člověk věděl, jestli je třeba daná cena mrkve adekvátní, kdyby nikdy mrkev nekupoval? V případě mrkve by člověk asi obětoval nějaké peníze a nové zkušenosti si tímto způsobem získal. K čemu však dospějeme, když nabytí těchto zkušeností je příliš nákladné, časově náročné nebo jinak rizikové? V tomto případě nastupuje "opravdová" matematika, jak si ji představuje většina lidí.

Tato práce popisuje některé tyto matematické optimalizační postupy a metody. Začíná teoretickým základem optimalizace, pokračuje metodami jednorozměrnými a poslední kapitola je věnována metodám vícerozměrné optimalizace a jejím modifikacím. Toto dílo však nebylo vytvořeno s cílem odhalit nějakou doposud nepoznanou vědomost, ale nastínit čtenáři určité principy či postupy optimalizace. Mým cílem není předat technické znalosti, ale pomoci přemýšlet o běžných věcech v trochu jiné rovině než je obvyklé.

„Nejpozoruhodnější na člověku je jeho schopnost myšlení.“

Aristotelés ze Stageiry  
(384-322 př.n.l)

---

# Kapitola 1

## Koncept metod nepodmíněné minimalizace

Nechť  $X \subset \mathbb{R}^n$  je konvexní množina a uvažujme na  $X$  konvexní funkci. Naší snahou je najít bod  $x^* \in X$  v němž funkce  $f$  nabývá na  $X$  svého minima. Tuto minimalizovanou funkci budeme nazývat funkcí účelovou. V mnoha praktických případech neznáme explicitní vyjádření funkce  $f$  ani její derivace, nemůžeme tedy použít k hledání minima standardní minimalizační postup pomocí derivací. Jelikož však neexistuje obecný postup, kterým by se řídily všechny metody v tomto textu probírané, je zde popsán obecný postup řešení jedné z jednodušších skupin metod. Obecný postup při řešení takového problému nám nastíní teoretickou koncepci stojící za všemi metodami nepodmíněné minimalizace.

Zvolíme počáteční bod  $x_0 \neq x^*$ . Tato volba je často určena přímo použitou metodou. Jelikož se v bodě  $x^*$  nachází minimum funkce  $f(x)$ , musí platit

$$\Delta f(x_0) = f(x^*) - f(x_0) < 0. \quad (1.1)$$

Nyní se pokusíme z bodu  $x_0$  posunout co nejbližší bodu  $x^*$ , přitom jeden krok tohoto posunu budeme definovat jako

$$\Delta x_k = x_{k+1} - x_k = \alpha_k s_k, \quad (1.2)$$

kde  $\alpha_k$  udává délku kroku a  $s_k$  jeho směr. Krok nás přiblížil minimu, pokud bude splněno

$$\Delta f(x_k) = f(x_{k+1}) - f(x_k) < 0. \quad (1.3)$$

Posloupnosti bodů splňující tuto podmínku se říká minimalizující posloupnost.

Je důležité si uvědomit, že jedním krokem jsme se k minimu jen přiblížili a při konečném počtu kroků se tedy výsledný bod nemusí rovnat bodu minima<sup>1</sup>. Metody nepodmíněné minimalizace ve většině případů dávají jen odhad bodu minima. Proto je často pro ukončení procesu výpočtu použita předem zvolená podmínka, za které se proces ukončí a poslední

---

<sup>1</sup>Existují i případy, kdy minima není možno danou metodou dosáhnout v konečném počtu kroků vůbec.

## KAPITOLA 1. KONCEPT METOD NEPODMÍNĚNÉ MINIMALIZACE

---

zjištěný bod je prohlášen za minimum účelové funkce. Toto omezení je zpravidla reprezentováno skalárem, který označíme  $\Delta$  a nazveme koeficientem přesnosti. Samotná podmínka může být definována různě - často pracuje se vzdáleností bodů ve dvou po sobě následujících krocích.

Naším cílem tedy v každém kroku bude najít takový vektor  $\Delta x_k$ , který by nás přiblížil co nejbližší k hledanému minimu, v nejlepším případě přímo tímto minimem procházel. Metody opakující dokola stejný postup jednoho kroku se nazývají metodami iteračními a jeden krok označíme za jednu iteraci. Možností jak volit délku koru  $\alpha_k$  a směr kroku  $s_k$  je několik - liší se v jednoduchosti, rychlosti konvergence a přesnosti určení minima. Obecně lze říci, že čím je proces nalezení délky a směru kroku složitější a časově náročnější, tím je metoda přesnější v odhadu minima.



---

## Kapitola 2

# Metody jednorozměrné optimalizace

Metody jednorozměrné optimalizace jsou nejjednodušší optimalizační metody a díky své snadno pochopitelné grafické interpretaci vhodné pro osvojení si prvních znalostí o optimalizaci. Samostatně nemají mnoho praktických použití, ale využívají se u metod vícerozměrné optimalizace, které se výpočtem převádějí právě na optimalizaci jednorozměrnou.

Důležitým aspektem při dělení metod do jednotlivých skupin je, zda je metoda aktivní nebo pasivní. Pasivní metody mají předem daný průběh a výsledky získané před dokončením iterace (například naměřené funkční hodnoty) nemají vliv na žádný z dalších kroků v průběhu dané iterace. Tyto metody jsou zpravidla náročnější na počet zjišťovaných hodnot nebo při stejném počtu zjišťovaných hodnot méně přesné v porovnání s metodami aktivními, kde se rozhodujeme o dalším postupu na základě informací zjištěných v dané iteraci. Pasivní metody však mohou mít své využití. Například, pokud je zjišťování funkčních hodnot časově náročné, ale je možné zjišťovat více funkčních hodnot zároveň. V takovém případě by byl výpočet pomocí aktivní metody velice zdlouhavý, jelikož bychom v každém kroku čekali na výsledek, zatímco při použití metody pasivní bychom prohledali celý uvažovaný interval najednou. Jelikož jsou pasivní metody obecně jednodušší, podíváme se jako první na metodu pasivní a dále již na metody aktivní.

Nejdříve se zaměříme na metody, které nevyžadují znalost samotné účelové funkce  $f(x)$  ve všech bodech intervalu, kde hledáme minimum ani jejích derivací, tedy na metody n.gradientní. Tyto metody slouží k hledání intervalu lokalizace minima - tento interval bude nejmenším intervalem, ve kterém se minimum s určitostí nachází. Velikost tohoto intervalu bude záviset na zvolené metodě a na dalších parametrech metody samotné. Dále si ukážeme metodu, která využívá derivace účelové funkce, tedy patří do skupiny metod gradientních, nakonec se seznámíme s metodou, která bude přímo odhadovat minimum pomocí aproximace účelové funkce a hledání minima této aproximace. Jedná se o metodu interpolační.

Jelikož konvexnost funkcí je příliš silná a omezující podmínka, zavedeme si slabší pojem, který bude vyhovovat většině zde popisovaných metod:

**Definice 1.** Řekneme, že funkce  $f: [a, b] \rightarrow \mathbb{R}$  je unimodální, jestliže existuje  $x^*$  náležící

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

do intervalu  $[a, b]$  takové, že  $f$  je klesající (rostoucí) na  $[a, x^*)$  a rostoucí (klesající) na  $(x^*, b]$ .

## 2.1 Komparativní metody

První metody, kterými se budeme zabývat, slouží k řešení problému, kdy neznáme ani funkci  $f(x)$  ani její derivace. Jediné co můžeme zjistit je funkční hodnota v určitém bodě a tyto hodnoty porovnávat. Z tohoto důvodu jsou tyto metody nazývány metodami komparativními. Metody se setkávají s protichůdnými požadavky na minimalizaci počtu bodů, jejichž funkční hodnotu musíme zjistit a na minimalizaci intervalu lokalizace minima zaručující větší přesnost.

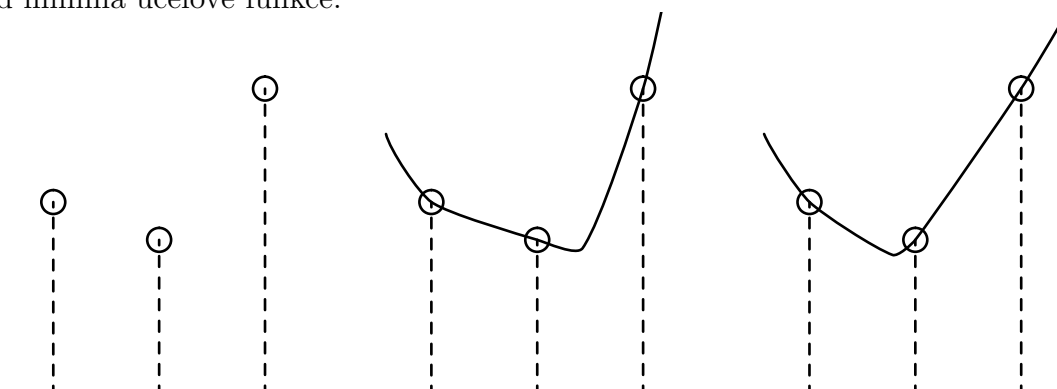
### 2.1.1 Rovnoměrné dělení intervalu

Mějme na intervalu  $\langle a, b \rangle$  definovanou spojitou unimodální účelovou funkci  $f(x)$  s minimem v bodě  $x^*$ . Naším úkolem je najít takový interval lokalizace minima, aby bod minima nebyl vzdálený od středu tohoto intervalu více než o zadaný koeficient přesnosti  $\Delta$ .

Jako výchozí bod  $x_0$  zvolíme hraniční bod  $a$  a budeme vytvářet posloupnost bodů  $x_{k+1} = x_k + \alpha_k s_k$ , v nichž budeme zjišťovat funkční hodnotu. Rozdělením intervalu  $\langle a, b \rangle$  pomocí  $N$  bodů získáme  $N + 1$  podintervalů. Jelikož se volba  $x_{k+1}$  odehrává na přímce, tedy v jednorozměrném prostoru, směr kroku bude vždy  $s_k = 1$ , vyjadřující postup z bodu  $a$  do bodu  $b$ . Pokud bychom se rozhodli postupovat naopak, byl by směr  $s_k = -1$ . Délku jednoho kroku lze vyjádřit

$$\alpha_k = \frac{b - a}{N + 1}. \quad (2.1)$$

Pro objasnění optimálního zvolení počtu bodů  $N$  musíme zjistit více o metodě samotné. Jak můžeme vidět na obrázku níže, při nalezení minimální funkční hodnoty v bodech vytvořené posloupnosti nemůžeme jednoznačně určit jediný interval  $\langle x_k, x_{k+1} \rangle$  jako interval obsahující bod minima účelové funkce.



## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Tato nejednoznačnost plyne z toho, že nevíme, zda se bod opravdového minima nachází blíže k bodu  $a$  nebo k bodu  $b$  od bodu s minimální zjištěnou funkční hodnotou. Proto za konečný interval lokalizace minima zvolíme takový interval  $\langle x_{k-1}, x_{k+1} \rangle$ , pro který platí

$$f(x_k) \leq f(x_{k-1}), \quad f(x_k) \leq f(x_{k+1}),$$

tj.  $x_k$  je takový bod, že

$$f(x_k) = \min \{f(x_j), j = 1, \dots, N\}.$$

Tento interval je nejmenším podintervalem původního intervalu, ve kterém se bude bod minima s jistotou nacházet. Jelikož  $x_{k+1}$  je středem konečného intervalu lokalizace minima a vzdálenost bodu minima od středu tohoto intervalu nesmí být větší než přesnost  $\Delta$  je vidět, že maximální velikost intervalu  $\langle x_k, x_{k+2} \rangle$  musí být  $2\Delta$ . Z toho vyplývá, že velikost intervalu mezi dvěma sousedními body dělení musí být menší nebo rovna přesnosti  $\Delta$ , tedy

$$\alpha_k = x_{k+1} - x_k \leq \Delta. \quad (2.2)$$

Z rovnic (2.1) a (2.2) můžeme odvodit omezení pro počet bodů  $N$  dělících interval  $\langle a, b \rangle$

$$\begin{aligned} \frac{b-a}{N+1} &\leq \Delta, \\ b-a &\leq \Delta(N+1), \\ \frac{b-a}{\Delta} - 1 &\leq N. \end{aligned} \quad (2.3)$$

V poslední rovnici nastává rovnost v případě, že délka intervalu  $\langle a, b \rangle$  je dělitelná koeficientem udávajícím požadovanou přesnost. Pokud není tato podmínka splněna, tak optimální  $N$  je nejbližší celé číslo vyšší než  $\frac{b-a}{\Delta} - 1$ , jelikož bude vyhovovat požadavku přesnosti a bude nejmenší možné. Tato podmínka lze zapsat jako

$$\frac{b-a}{\Delta} - 1 \leq N < \frac{b-a}{\Delta} \quad (2.4)$$

Nyní máme veškeré teoretické znalosti potřebné k výpočtu pomocí rovnoměrného dělení intervalu.

**Příklad 1.** Na intervalu  $\langle 0, 15 \rangle$  najděte bod minima funkce  $f(x)$  s požadovanou přesností  $\Delta = 2$  pomocí metody rovnoměrného dělení intervalu. Víte, že  $f(0) = 7$  a  $f(15) = 7,84$ .

**Řešení.** Nejdříve zjistíme počet bodů rozdělujících interval  $\langle 0, 15 \rangle$ .

$$\frac{b-a}{\Delta} - 1 = \frac{15-0}{2} - 1 = 6,5 \leq N.$$

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Jelikož délka původního intervalu není dělitelná koeficientem přesnosti musíme doplnit druhou nerovnici.

$$N < \frac{b-a}{\Delta} = 7,5.$$

Z těchto dvou nerovnic získáváme  $N = 7$ . Nyní víme, že interval bude rozdělen pomocí sedmi bodů na osm podintervalů. Délka jednoho podintervalu bude

$$\alpha_k = \frac{b-a}{N+1} = \frac{15-0}{7+1} = 1,875.$$

Při znalosti délky podintervalu můžeme rozdělit interval  $\langle 0, 15 \rangle$  a zjistit funkční hodnoty v jednotlivých bodech. Předpokládejme, že jsme nějakým způsobem, například měřením, našli funkční hodnoty tak, jak jsou uspořádány v následující tabulce.

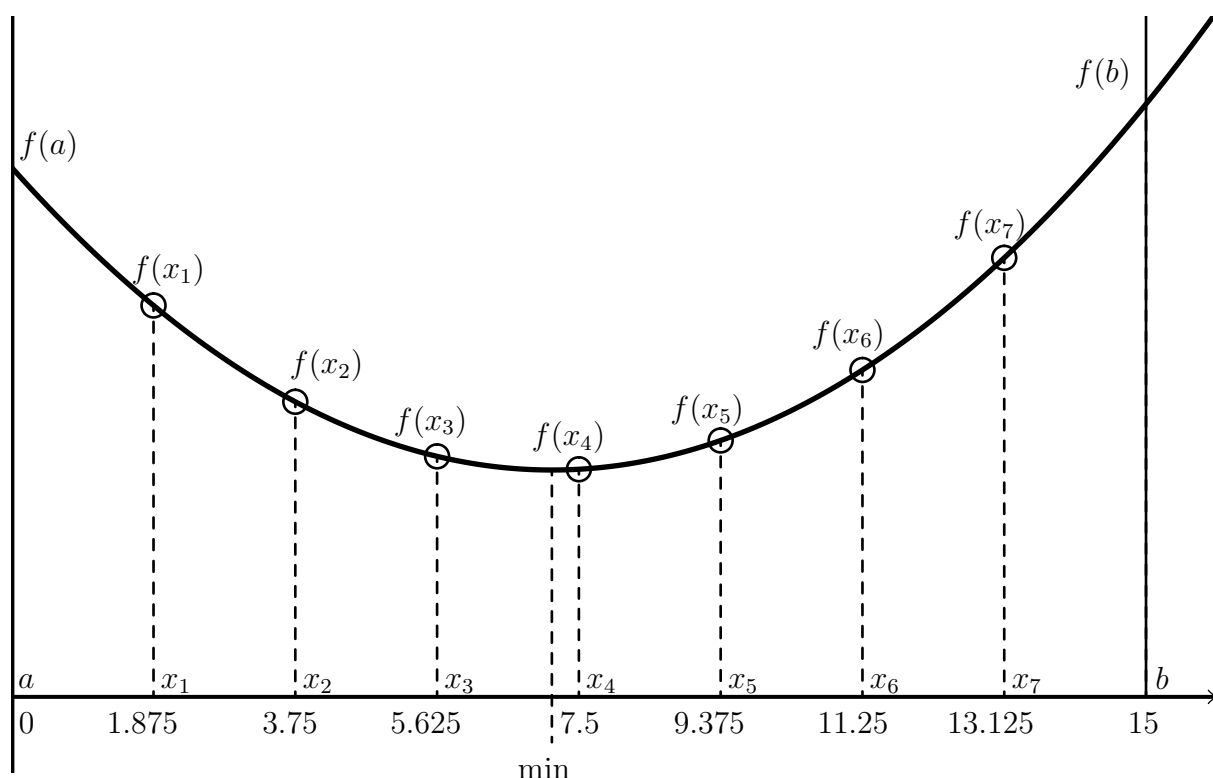
k	a	1	2	3	4	5	6	7	b
$x_k$	0	1,875	3,75	5,625	7,5	9,375	11,25	13,125	15
$f(x_k)$	7,000	5,175	3,902	3,180	3,01	3,390	4,322	5,805	7,84

Bod minima je  $x_4 = 7,5$  s funkční hodnotou  $f(x_4) = 3,01$ . Interval lokalizace minima bude v tomto případě  $\langle x_3, x_5 \rangle$ , tedy  $\langle 5,625; 9,375 \rangle$ .

Nyní předpokládejme, že hodnoty v tabulce jsou funkční hodnoty hledané funkce. Pak můžeme porovnat nalezené minimum s minimem skutečným. Účelová funkce má rovnici

$$f(x) = (2 - 0,28x)^2 + 3.$$

Skutečný bod minima je v bodě 7,1428 s funkční hodnotou  $f(7,1428) = 3$ . Skutečný bod minima jsme minuli o 0,357. Ve funkční hodnotě je naše nepřesnost rovna 0,01.



### 2.1.2 Fibonacciho metoda

Mějme opět spojitou unimodální účelovou funkci  $f(x)$  definovanou na intervalu  $\langle a, b \rangle$ . Dále popsaná metoda je založená na Fibonacciho posloupnosti

$$\{F_n\}_{n=0}^{\infty} = \{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \dots\}$$

Posloupnost lze též zapsat rekurentním vzorcem s počáteční podmínkou

$$F_{n+2} = F_{n+1} + F_n, \quad F_0 = F_1 = 1.$$

pro  $n = 0, 1, 2, \dots$

Prvotním, u této metody velice důležitým krokem bude zvolení počtu kroků metody. Jelikož však doposud nemáme dostatek znalostí o funkčnosti metody odložíme tuto volbu na později. Zatím jen označíme  $N$  pozici v Fibonacciho posloupnosti, kterou bychom zvolili pomocí požadované přesnosti, kde  $F_N$  je hodnota na této pozici ve Fibonacciho posloupnosti.

Jak již bylo zmíněno na počátku, všechny metody nemají stejný obecný postup a postup řešení Fibonacciho metody je lehce odlišný od předchozí metody - jedná se o metodu aktivní, zatímco metoda předcházející je metodou pasivní. Metoda stále pracuje v krocích, avšak nyní při každém kroku zmenšíme interval lokalizace minima a v kroku následujícím pracujeme jen s tímto zmenšeným intervalem. Interval, který po uplynutí zvoleného počtu kroků zůstane, je hledaným intervalem lokalizace minima.

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Označme  $a_0 = a, b_0 = b$ . Nyní výraz

$$|b_0 - a_0|$$

označuje velikost intervalu  $\langle a, b \rangle$ .

Postupně budeme tvořit posloupnost reálných čísel

$$\lambda_i = \frac{F_{N-i}}{F_N} |b_0 - a_0|, \quad (2.5)$$

kde  $\lambda_0$  označuje celý interval  $\langle a, b \rangle$  a následující  $\lambda_i$  označují postupně se zmenšující podintervaly. Na začátku výpočtu vypočteme  $\lambda_1$  a  $\lambda_2$  a položíme

$$y_0 = a_0 + \lambda_2, \quad z_0 = a_0 + \lambda_1. \quad (2.6)$$

Nyní nastává část metody využívající možnosti zjistit funkční hodnotu v daných bodech. Zjistíme funkční hodnotu v bodech  $y_0$  a  $z_0$  a tyto funkční hodnoty porovnáme. Mohou nastat tři následující situace. V každé z nich se provede přeznačení některých z již nalezených bodů a doplnění chybějícího čtvrtého bodu s tím, že se interval lokalizace minima zmenší o  $\lambda_2 = 1 - \lambda_1$  (obecně o  $\lambda_{k+1}$ , kde  $k$  označuje číslo kroku, přičemž první krok má číslo 1).

Následující možnosti budou zapsány pro obecný  $k$ -tý krok.

**Možnost a)**  $f(y_{k-1}) \leq f(z_{k-1})$

Provedeme přeznačení

$$a_k = a_{k-1}, \quad b_k = z_{k-1}, \quad z_k = y_{k-1}.$$

Jelikož nám v tomto případě vypadl bod  $y$ , doplníme ho s pomocí  $\lambda_{k+2}$

$$y_k = a_k + \lambda_{k+2}.$$

**Možnost b)**  $f(y_{k-1}) \geq f(z_{k-1})$

Provedeme přeznačení

$$a_k = y_{k-1}, \quad b_k = b_{k-1}, \quad y_k = z_{k-1}.$$

Jelikož nám v tomto případě vypadl bod  $z$  doplníme ho s pomocí  $\lambda_{k+2}$

$$z_k = b_k - \lambda_{k+2}.$$

**Možnost c)**  $f(y_{k-1}) = f(z_{k-1})$

V tomto případě nezáleží na tom, jaký bod zvolíme. Můžeme se libovolně rozhodnout mezi postupem a) nebo b).

Tímto postupem jsme interval lokalizace minima zmenšili v  $k$ -tém kroku o  $\lambda_{k+1}$  a máme připraveny všechny body  $a_k, b_k, y_k, z_k$  pro další krok.

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Na počátku kroku  $k = N - 1$  se výpočet dostane do situace, kdy

$$y_{k-1} = z_{k-1}.$$

Pro konečné zmenšení intervalu lokalizace minima však potřebujeme tyto body rozdílné, proto místo vypočteného  $z_{k-1}$  využijeme

$$z_{k-1} = y_{k-1} + \varepsilon^1,$$

kde  $\varepsilon \in \mathbb{R}$  je dostatečně malé a provede se krok  $N - 1$ , v tomto případě již bez zjišťování polohy nových  $y_k, z_k$ .

Nyní už známe metodu dostatečně k tomu, abychom v závislosti na předem zadané přesnosti  $\Delta$  mohli adekvátně zvolit  $N$ .

**Věta 2.1.** *Nechť  $N \in \mathbb{N}$  je libovolné. Po uplynutí  $N - 1$  kroků Fibonacciho metody zůstane interval lokalizace minima o velikosti*

$$\lambda_N = \frac{F_{N-N}}{F_N} |b_0 - a_0| = \frac{1}{F_N} |b_0 - a_0|. \quad (2.7)$$

*Z tohoto vyplývá, že parametr  $N$  musí být zvolen tak, aby*

$$\frac{\frac{1}{F_N} |b_0 - a_0|}{2} \leq \Delta, \quad (2.8)$$

*kde  $\Delta$  udává požadovanou přesnost výpočtu.*

*Důkaz.* Délka výsledného intervalu lokalizace minima lze zapsat jako

$$\lambda_N = \lambda_0 - \lambda_2 - \lambda_3 - \dots - \lambda_N.$$

Tuto rovnici můžeme s použitím vzorce (2.5) rozepsat do tvaru

$$\frac{F_0}{F_N} |b_0 - a_0| = \frac{F_N}{F_N} |b_0 - a_0| - \frac{F_{N-2}}{F_N} |b_0 - a_0| - \frac{F_{N-3}}{F_N} |b_0 - a_0| - \dots - \frac{F_0}{F_N} |b_0 - a_0|.$$

Jelikož však mají všechny zlomky stejný jmenovatel a všechny závisí na stejné délce původního intervalu můžeme ekvivalentně dokázat rovnici

$$F_0 = F_N - (F_{N-2} + F_{N-3} + \dots + F_0). \quad (2.9)$$

Protože součet dvou po sobě jdoucích čísel Fibonacciho posloupnosti dává číslo následující musí platit

$$F_N = F_{N-1} + F_{N-2}.$$

---

<sup>1</sup>Stejně můžeme využít i  $y_{k-1} = z_{k-1} - \varepsilon$ . Znaménko v této rovnici nehraje důležitou roli. Významné je jen, aby bylo možné porovnat funkční hodnoty v dvou rozdílných bodech.

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Dosazením do rovnice (2.9) a odečtením členu  $F_{N-2}$  získáme

$$F_0 = F_{N-1} - (F_{N-3} + F_{N-4} + \dots + F_0).$$

Dále platí

$$F_{N-1} = F_{N-2} + F_{N-3}.$$

Opět dosadíme a získáváme

$$F_0 = F_{N-2} - (F_{N-4} + F_{N-5} + \dots + F_0).$$

Při opakování tohoto postupu dostatečně mnohokrát pro zvolené  $N$  dostaneme

$$F_0 = F_4 - (F_2 + F_1 + F_0).$$

Z definice platí

$$F_4 = F_3 + F_2,$$

z čeho vyplývá poslední rovnice

$$F_0 = F_3 - F_1 - F_0,$$

tedy

$$1 = 1.$$

Dokázali jsme, že v čitateli námi hledaného zlomku udávající velikost intervalu lokalizace minima bude  $F_0 = 1$ , tedy délka tohoto intervalu bude

$$\frac{F_0}{F_N} |b_0 - a_0| = \frac{1}{F_N} |b_0 - a_0|.$$

Tímto jsme dokázali rovnici (2.7). Nerovnici (2.8) získáme, když požadujeme, aby bod skutečného minima nebyl od středu námi zvoleného intervalu lokalizace minima vzdálen o více jak zadanou přesnost  $\Delta$ .

□

**Příklad 2.** Na intervalu  $\langle 0, 15 \rangle$  najděte bod minima funkce  $f(x)$  pomocí Fibonacciho metody.

**Řešení.** Jak již bylo zmíněno výše, počet kroků Fibonacciho metody přímo závisí na předem zvoleném parametru  $N$ . Abychom mohli porovnat výsledky jednotlivých metod, zvolíme  $N$  tak, aby počet zjišťovaných hodnot (např. měření) byl stejný jako u metod ostatních. Pro náš příklad tedy  $N = 6$ ,  $F_N = 13$ .

Pro začátek metody víme, že  $a_0 = 0$ ,  $b_0 = 15$ . První dva body dělení získáme dosazením do vzorce

$$y_0 = a_0 + \lambda_2 = \frac{F_{N-2}}{F_N} |b_0 - a_0| = \frac{F_4}{F_6} |b_0 - a_0| = \frac{5}{13} \cdot 15 = 5,769$$



## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

$$z_0 = a_0 + \lambda_1 = \frac{F_{N-1}}{F_N} |b_0 - a_0| = \frac{F_5}{F_6} |b_0 - a_0| = \frac{8}{13} \cdot 15 = 9,230$$

Porovnáním funkčních hodnot získáme

$$f(y_0) = f(5,769) = 3,147 < f(z_0) = f(9,230) = 3,341.$$

Nastává možnost **a)** v našem rozdělení. Pomocí přeznačení omezíme interval lokalizace minima na interval  $\langle 0; 9,230 \rangle$ .

$$a_1 = a_0 = 0, \quad b_1 = z_0 = 9,230, \quad z_1 = y_0 = 5,769.$$

Následující  $y_1$  vypočteme ze vzorce

$$y_1 = a_1 + \lambda_3 = \frac{F_{N-3}}{F_N} |b_0 - a_0| = \frac{F_3}{F_6} |b_0 - a_0| = \frac{3}{13} \cdot 15 = 3,461.$$

Porovnáním funkčních hodnot získáme

$$f(y_1) = f(3,461) = 4,062 > f(z_1) = f(5,769) = 3,147.$$

Nastává možnost **b)** v našem rozdělení. Pomocí přeznačení omezíme interval lokalizace minima na interval  $\langle 3,461; 9,230 \rangle$ .

$$a_2 = y_1 = 3,461, \quad b_2 = b_1 = 9,230, \quad y_2 = z_1 = 5,769.$$

Následující  $z_2$  vypočteme ze vzorce

$$z_2 = b_2 - \lambda_4 = b_2 - \frac{F_{N-4}}{F_N} |b_0 - a_0| = b_2 - \frac{F_2}{F_6} |b_0 - a_0| = 9,230 - \frac{2}{13} \cdot 15 = 6,922.$$

Krok 3:

Porovnání funkčních hodnot:

$$f(y_2) = f(5,769) = 3,147 > f(z_2) = f(6,922) = 3,003.$$

Získání nových hraničních a dělicích bodů:

$$a_3 = y_2 = 5,769, \quad b_3 = b_2 = 9,230, \quad y_3 = z_2 = 6,922,$$

$$z_3 = b_3 - \lambda_5 = b_3 - \frac{F_{N-5}}{F_N} |b_0 - a_0| = b_3 - \frac{F_1}{F_6} |b_0 - a_0| = 9,230 - \frac{1}{13} \cdot 15 = 8,076.$$

Krok 4:

Porovnání funkčních hodnot:

$$f(y_3) = f(6,922) = 3,003 < f(z_3) = f(8,076) = 3,068,$$

Získání nových hraničních a dělicích bodů:

$$a_4 = a_3 = 5,769, \quad b_4 = z_3 = 8,076, \quad z_4 = y_3 = 6,922,$$

$$y_4 = a_4 + \lambda_6 = a_4 + \frac{F_{N-6}}{F_N} |b_0 - a_0| = a_4 + \frac{F_0}{F_6} |b_0 - a_0| = 5,769 + \frac{1}{13} \cdot 15 = 6,922.$$

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

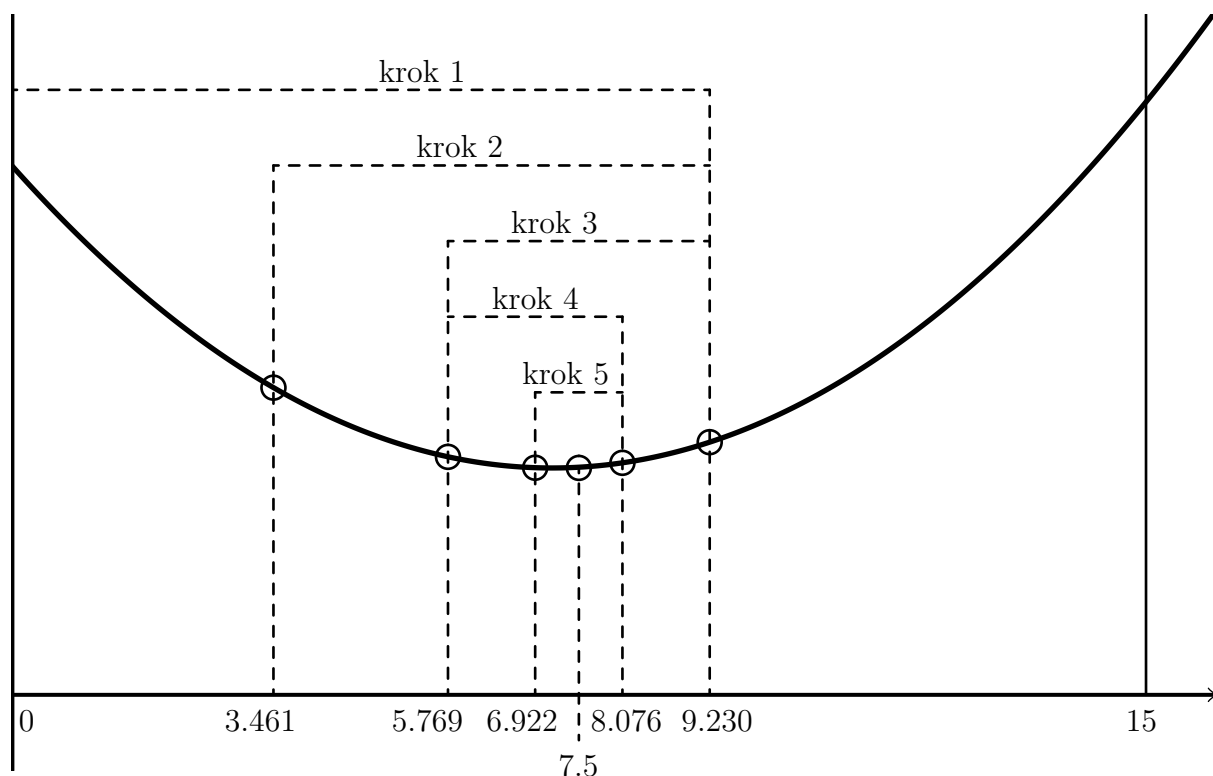
Nyní náš příklad dospěl do stavu, kdy  $y_4 = z_4$ . Můžeme si zvolit, zda zvýšíme hodnotu  $z_4$  (laicky řečeno posuneme bod  $z_4$  doprava od bodu  $y_4$ ), nebo snížíme hodnotu  $y_4$  (laicky řečeno posuneme bod  $y_4$  doleva od bodu  $z_4$ ). My zde zvolíme přeznačení  $z_4 = y_4 + \varepsilon$ . V našem konkrétním příkladě bude adekvátní  $z_4 = 6,953$ .

$$y_4 = f(6,922) = 3,0037 > z_4 = f(6,953) = 3,0028.$$

Posledním přeznačením získáme konečné ohraničení intervalu lokalizace minima.

$$a_5 = y_4 = 6,922, \quad b_5 = b_4 = 8,076.$$

Pokud si jako odhad minima hledané funkce zvolíme střed takto získaného intervalu lokalizace minima získáme hodnotu  $x_{min} = 7,5$  s funkční hodnotou  $f(x_{min}) = 3,01$ .



Jelikož jsme použili tuto metodu na stejnou funkci jako v předešlém případě i zde je minimum v bodě 7,1428 s funkční hodnotou  $f(7,1428) = 3$ . Skutečný bod minima jsme minuli o 0,357. Ve funkční hodnotě je naše nepřesnost rovna 0,01.

Využili jsme stejně, jako u předchozího případu, zjištění sedmi funkčních hodnot (pro přehlednost je z grafu vynechána hodnota 6,953). Mezi zjišťované hodnoty počítáme i hodnotu finálního odhadu minima. Pokud bychom nevyžadovali znalost tohoto odhadu, je možné počítat příklad pro hodnotou parametru  $N = 7$  pro vyšší výslednou přesnost. Bod, ve kterém odhadujeme minimum je sice stejný jako v případě metody předešlé, ale interval

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

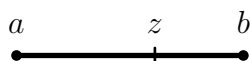
---

lokalizace minima je rozdílný. Zatímco Fibonacciho metoda skončila s výsledným intervalem  $\langle 6, 922; 8, 076 \rangle$  tak interval získaný metodou rovnoměrného dělení byl  $\langle 5, 625; 9, 375 \rangle$ , tedy více jak třikrát větší. Fibonacciho metoda nám tedy dává daleko spolehlivější odhad umístění bodu minima.

#### 2.1.3 Metoda zlatého řezu

Metoda zlatého řezu se nazývá podle poměru, ve kterém dělí interval lokalizace minima. Při rozdělení úsečky v poměru zlatého řezu je poměr délky celé úsečky k délce delší části stejný jako poměr délky delší části k délce části menší.

$$\frac{|b - a|}{|z - a|} = \frac{|z - a|}{|b - z|} \quad (2.10)$$



Metoda zlatého řezu se od Fibonacciho metody liší jen umístěním bodů dělení. Postup porovnávání funkčních hodnot a zkracování intervalu funguje naprosto stejně.

Umístění prvního bodu  $z_0$  v intervalu bude  $z_0 = \hat{\tau}|b_0 - a_0|$ , kde  $\hat{\tau}$  získáme z (2.10) aplikací na úsečku o délce 1 následujícím způsobem

$$\frac{1}{\tau} = \frac{\tau}{1 - \tau}$$

Úpravou dostaneme kvadratickou rovnici

$$\tau^2 + \tau - 1 = 0,$$

která má kladné řešení

$$\hat{\tau} = \frac{\sqrt{5} - 1}{2} \doteq 0,61.$$

Druhé (záporné) řešení zde není zapsáno, jelikož dělicí bod nenáleží do uvažovaného intervalu.

Nyní máme první bod dělení. Druhý získáme zobrazením prvního bodu pomocí středové souměrnosti podle středu intervalu  $[a, b]$ . Dále postupujeme jako při Fibonacciho metodě. Při zkrácení intervalu najdeme chybějící bod dělení opět pomocí středové souměrnosti podle středu intervalu, tedy obecně zapsané rovnice pro výpočet bodů dělení po  $k$ -tém kroku jsou

$$y_k = (1 - \hat{\tau})|b_k - a_k| \quad z_k = \hat{\tau}|b_k - a_k|.$$

Metoda zlatého řezu má oproti Fibonacciho metodě tu výhodu, že není vázaná na předem zvolené  $N$  a můžeme provést tolik přiblížení, kolik chceme, případně kolik je potřeba k

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

dosažení potřebné přesnosti. Existuje zde však i nevýhoda a tou je, že při stejném počtu kroků je výsledný interval lokalizace minima delší než při použití metody Fibonacciho.

Tuto skutečnost lze dokázat následujícím způsobem.

Délka intervalu lokalizace minima  $|b_k - a_k|$  se při každém kroku metodou zlatého řezu zmenší na  $\hat{\tau}|b_k - a_k|$ . Z tohoto lze odvodit, že po  $k$  krocích bude interval lokalizace minima

$$|b_k - a_k| = \hat{\tau}^k |b_0 - a_0|. \quad (2.11)$$

Tuto délku intervalu po  $k$ -tém kroku můžeme porovnat s rovnicí (2.7), která udává délku intervalu lokalizace minima po  $k$ -tém kroku Fibonacciho metodou. Poměr těchto dvou délek bude

$$\frac{\hat{\tau}^k |b_0 - a_0|}{\frac{1}{F_{k+1}} |b_0 - a_0|} = \hat{\tau}^k F_{k+1}. \quad (2.12)$$

Tyto poměry jsou i pro relativně malá  $k$  stabilní, pohybující se okolo 1,17. Tedy interval lokalizace minima vypočítaný metodou zlatého řezu bude o 17% delší než interval vypočítaný pomocí Fibonacciho metody, a to při stejném počtu kroků.

**Příklad 3.** Na intervalu  $\langle 0, 15 \rangle$  najděte bod minima funkce  $f(x)$  pomocí metody zlatého řezu. Z důvodu lepšího srovnání je naše účelová funkce stejná jako u předešlých metod.

**Řešení.** Ze zadání víme, že  $a_0 = 0, b_0 = 15$ . První body dělení vypočítáme

$$\begin{aligned} y_0 &= (1 - \hat{\tau})|b_0 - a_0| = 0,39 \cdot 15 = 5,85, \\ z_0 &= \hat{\tau}|b_0 - a_0| = 0,61 \cdot 15 = 9,15. \end{aligned}$$

Porovnáním funkčních hodnot získáme

$$f(y_0) = f(5,85) = 3,131 < f(z_0) = f(9,15) = 3,316.$$

Přeznačení se provede stejně jako u Fibonacciho metody

$$a_1 = a_0 = 0, \quad b_1 = z_0 = 9,15, \quad z_1 = y_0 = 5,85.$$

Musíme dopočítat polohu bodu  $y_1$ , a to zobrazením bodu  $z_1$  pomocí středové souměrnosti podle středu intervalu  $\langle a_1, b_1 \rangle$ .

$$y_1 = b_1 - z_1 = 9,15 - 5,85 = 3,3.$$

Opět máme všechny hraniční i dělicí body potřebné pro další krok

$$f(y_1) = f(3,3) = 4,157 > f(z_1) = f(5,85) = 3,131.$$

Provedeme přeznačení

$$a_2 = y_1 = 3,3, \quad b_2 = b_1 = 9,15, \quad y_2 = z_1 = 5,85.$$

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.1. KOMPARATIVNÍ METODY

---

Dopočítáme polohu bodu  $z_2$

$$z_2 = b_2 - (y_2 - a_2) = 9,15 - (5,85 - 3,3) = 6,6.$$

Krok 3:

Porovnání funkčních hodnot

$$f(y_2) = f(5,85) = 3,131 > f(z_2) = f(6,6) = 3,023.$$

Získání nových hraničních a dělicích bodů

$$a_3 = y_2 = 5,85, \quad b_3 = b_2 = 9,15, \quad y_3 = z_2 = 6,6,$$

$$z_3 = b_3 - (y_3 - a_3) = 9,15 - (6,6 - 5,85) = 8,4.$$

Krok 4:

Porovnání funkčních hodnot

$$f(y_3) = f(6,6) = 3,023 < f(z_3) = f(8,4) = 3,124.$$

Získání nových hraničních a dělicích bodů

$$a_4 = a_3 = 5,85, \quad b_4 = z_3 = 8,4, \quad y_4 = y_3 = 6,6,$$

$$z_4 = b_4 - (z_4 - a_4) = 8,4 - (6,6 - 5,85) = 7,65.$$

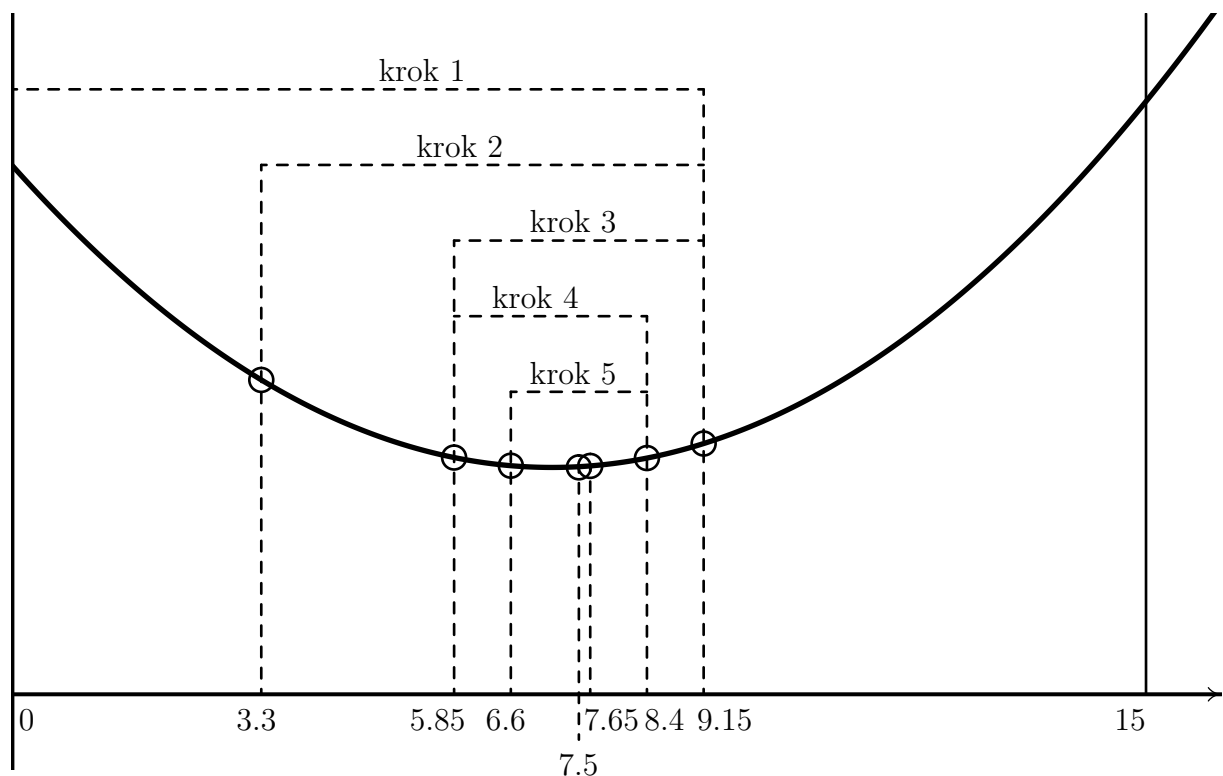
Krok 5:

Porovnání funkčních hodnot

$$f(y_4) = f(6,6) = 3,023 > f(z_4) = f(7,65) = 3,020.$$

Získání nových hraničních bodů

$$a_5 = y_4 = 6,6, \quad b_5 = b_4 = 8,4.$$



## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.2. GRADIENTNÍ METODY

---

Za odhad minima můžeme považovat střed výsledného intervalu  $\langle 6, 6; 8, 4 \rangle$ , tedy  $x_{\min} = 7, 5$ ,  $f(x_{\min}) = 3, 01$ . Odhadované umístění minima je opět stejné jako u předešlých metod.

Můžeme porovnat výše dokazovanou rozdílnost ve velikosti intervalů lokalizace minima u Fibonacciho metody a metody zlatého řezu. U Fibonacciho metody byl konečný interval lokalizace minima  $\langle a_5, b_5 \rangle = \langle 6, 922; 8, 076 \rangle$ , velikost intervalu  $|\langle 6, 922; 8, 076 \rangle| = 1, 154$ . Konečný interval lokalizace minima při použití metody zlatého řezu byl  $\langle a_5, b_5 \rangle = \langle 6, 6; 8, 4 \rangle$ , velikost intervalu  $|\langle 6, 6; 8, 4 \rangle| = 1, 8$ . Jak vidíme, tak interval lokalizace minima, který jsme dostali při použití Fibonacciho metody při stejném počtu zjišťovaných hodnot, je značně menší v porovnání s intervalem získaným při použití metody zlatého řezu. Přesto je interval získaný metodou rovnoměrného dělení více jak dvakrát větší než interval získaný metodou zlatého řezu. Metoda zlatého řezu má však i svoji výhodu, a to, že v ní může být pokračováno bez jakéhokoliv omezení jako je volba parametru  $N$  u Fibonacciho metody. Podívejme se, jak by hledání minima dopadlo, pokud bychom mohli zjistit o jednu funkční hodnotu více.

Nejdříve musíme získat dělicí body  $y_5$  a  $z_5$

$$z_5 = z_4 = 7, 65, \quad y_5 = b_5 - (z_5 - a_5) = 8, 4 - (7, 65 - 6, 6) = 7, 35$$

Porovnáním funkčních hodnot dělicích bodů získáme

$$f(y_5) = f(7, 35) = 3, 003 < f(z_5) = f(7, 65) = 3, 020.$$

Provedeme přeznačení pro nalezení hraničních bodů

$$a_6 = a_5 = 6, 6, \quad b_6 = z_5 = 7, 65.$$

Nyní je odhad bodu minima  $x_{\min} = 7, 125$ ,  $f(x_{\min}) = 3$ . Další krok nás značně přiblížil ke skutečnému minimu  $x^* = 7, 1428$ .

## 2.2 Gradientní metody

V následující metodě budeme kromě funkční hodnoty účelové funkce potřebovat i její derivace. Musíme tedy po funkci požadovat diferencovatelnost na intervalu  $\langle a, b \rangle$  a znalost jejích derivací.

### 2.2.1 Metoda tečen (Newtonova metoda)

Jak již název napovídá, budeme k vyřešení příkladu touto metodou potřebovat tečny účelové funkce, tedy první derivaci. V této metodě však vyžadujeme znalost i druhé derivace a pomůže znalost i derivace třetí. I zde není předem určován počet kroků metody, ukončení iteračního postupu je zaručeno sledováním vzdálenosti po sobě jdoucích bodů, ve kterých se sestřoují tečny.

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.2. GRADIENTNÍ METODY

---

Metoda staví na faktu, že minimalizace

$$f(x) \rightarrow \min$$

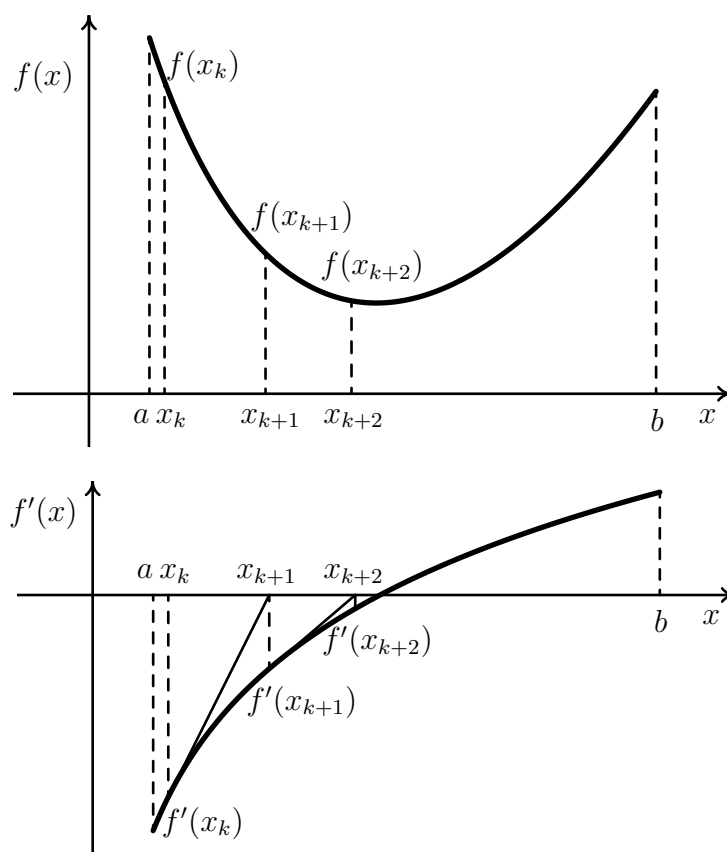
lze ekvivalentně zapsat

$$f'(x) \rightarrow 0.$$

Pokud bychom uvažovali stále jen funkci unimodální, tak by tento předpoklad nemusel platit. Avšak jak si ukážeme v následujícím výkladu, pro korektní funkčnost metody budeme vyžadovat funkci ryze konvexní, kdy je tento předpoklad vždy splněn.

Metoda využívá funkci první derivace účelové funkce a v každé iteraci k jejímu grafu v bodě  $x_k$  vytvoří tečnu. Průsečík této tečny s osou  $x$  bude bod následující iterace. Pro získání bodu  $x_{k+1}$  se použije vzorec

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (2.13)$$



Tento vzorec odvodíme z rovnice tečny k funkci derivace účelové funkce v bodě  $x_k$ . Pro lepší názornost provedeme substituci  $g(x) = f'(x)$ . Potom tečna bude mít vyjádření

$$g(x_k) + g'(x_k)(x - x_k)$$

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.2. GRADIENTNÍ METODY

---

Bodem další iterace  $x_{k+1}$  bude bod na této přímce, kde je funkční hodnota rovna nule, tedy

$$\begin{aligned}g(x_k) + g'(x_k)(x_{k+1} - x_k) &= 0 \\ \frac{g(x_k)}{g'(x_k)} &= -x_{k+1} + x_k \\ x_{k+1} &= x_k - \frac{g(x_k)}{g'(x_k)}\end{aligned}$$

Po vrácení substituce získáváme

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

Na začátku počítání zvolíme parametr  $\Delta > 0$ , který bude určovat přesnost, s jakou se chceme přiblížit k minimu. Jakmile dojdeme k bodu  $x_{k+1}$  takovému, že platí  $|x_{k+1} - x_k| \leq \Delta$ , ukončíme výpočet a bod  $x_{k+1}$  budeme považovat za bod minima.

Aby posloupnost bodů  $\{x_k\}$  konvergovala k bodu minima účelové funkce  $x^* = \arg \min f(x)$ , musí platit:

a)  $f'(a)f'(b) < 0$ ,

tedy funkce musí být v bodě  $a$  klesající a v bodě  $b$  rostoucí. (Případ, že by funkce byla v bodě  $a$  rostoucí a v bodě  $b$  klesající sice splní podmínku, ale nemusí být uvažován, jelikož v takovém případě není potřeba používat jakoukoliv výpočetní metodu pro hledání minima. Minimum bude v jednom z hraničních bodů. Metoda však stejně dobře může sloužit k hledání maxima)

b)  $\text{sign } f''(x) = k$ ,  
 $x \in \langle a, b \rangle$

kde  $k$  je konstantní, takové že  $k \neq 0$ . Z tohoto důvodu můžeme použít tuto metodu pouze na funkce ryze konvexní. Naše původní omezení na funkce unimodální nyní nebude dostatečné.

c)  $\text{sign } f'''(x) = l$ ,  
 $x \in \langle a, b \rangle$

kde  $l$  je konstantní, takové že  $l \neq 0$ .

Pro vhodné zvolení počátečního bodu posloupnosti  $\{x_k\}$  využijeme třetí derivace účelové funkce. Pro tento první bod by mělo platit:

$$f'(x_1)f'''(x_1) > 0,$$

jelikož v případě, že je počáteční bod posloupnosti  $\{x_k\}$  zvolen špatně, je možné, že další aproximace bude ležet mimo interval  $\langle a, b \rangle$ .



## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.2. GRADIENTNÍ METODY

---

**Příklad 4.** Najděte minimum účelové funkce pomocí metody tečen. Znáte rekurzivní vztah pro výpočet nového bodu posloupnosti

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} = x_k - \frac{\frac{x_k - 3}{5} - \frac{1}{x_k}}{\frac{1}{5} + \frac{1}{x_k^2}}. \quad (2.14)$$

Jako výchozí bod byl zvolen  $x_0 = 1$ , parametr udávající přesnost výpočtu je  $\Delta = 0,1$ .

**Řešení.** Podle (2.14) vypočítáme další bod posloupnosti

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = 1 - \frac{\frac{1 - 3}{5} - \frac{1}{1}}{\frac{1}{5} + \frac{1}{1}} = 2,1\bar{6}.$$

Nový bod posloupnosti nesplňuje požadavek na přesnost výpočtu, a proto pokračujeme

$$x_2 = x_1 - \frac{f'(x_1)}{f''(x_1)} = 2,1\bar{6} - \frac{\frac{2,1\bar{6} - 3}{5} - \frac{1}{2,1\bar{6}}}{\frac{1}{5} + \frac{1}{2,1\bar{6}^2}} = 3,688, .$$

$$x_3 = x_2 - \frac{f'(x_2)}{f''(x_2)} = 3,688 - \frac{\frac{3,688 - 3}{5} - \frac{1}{3,688}}{\frac{1}{5} + \frac{1}{3,688^2}} = 4,176,$$

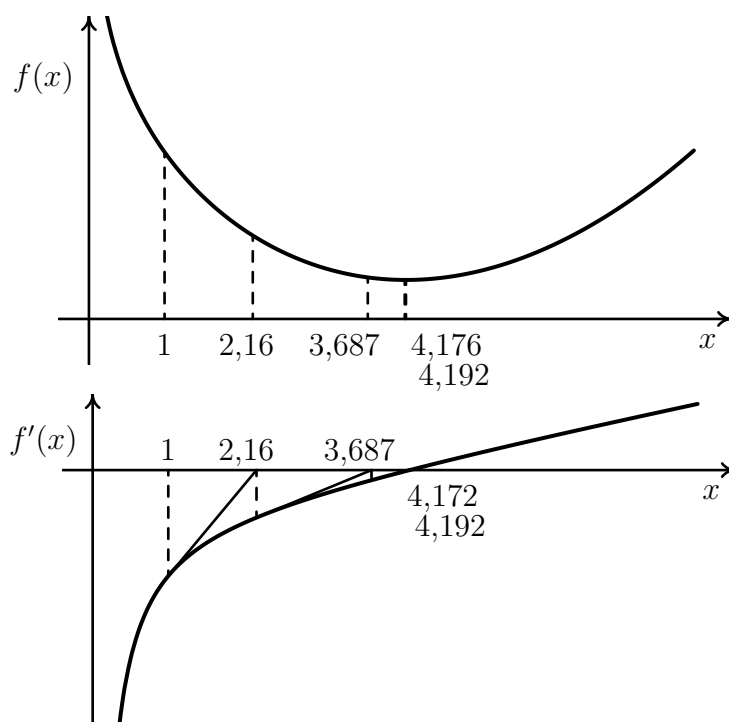
$$x_4 = x_3 - \frac{f'(x_3)}{f''(x_3)} = 4,176 - \frac{\frac{4,176 - 3}{5} - \frac{1}{4,176}}{\frac{1}{5} + \frac{1}{4,176^2}} = 4,192.$$

Vzdálenost mezi dvěma posledními body posloupnosti je  $|x_4 - x_3| = |4,192 - 4,176| = 0,016 \leq \Delta$ . Za hledaný bod minima považujeme  $x_4 = 4,192$ .

Naše účelová funkce byla

$$f(x) = \frac{(x - 3)^2}{10} - \ln(2x),$$

jejíž skutečné minimum se nachází v bodě  $x^* = 4,192$ , který jsme při přesnosti na tři desetinná místa odhadli naprosto přesně.



## 2.3 Interpolační metody

Metody interpolace, na rozdíl od předcházejících metod, nehledají menší interval lokalizace minima přímo pomocí účelové funkce, ale snaží se co nejlépe nahradit účelovou funkci vhodně zvolenou funkcí tak, aby podle průběhu této nahrazující funkce mohla metoda odhadovat minimum funkce účelové a podle toho zmenšovat interval lokalizace. Opět nemusíme znát vyjádření účelové funkce, ani jejích derivací na celém intervalu, kde hledáme extrém. Jediné, co budeme potřebovat, jsou funkční hodnoty ve zvolených bodech.

### 2.3.1 Metoda kvadratické interpolace

Metoda kvadratické interpolace využívá aproximaci účelové funkce  $f(x)$  pomocí kvadratické funkce  $F(x)$ . Tato kvadratická funkce bude mít vyjádření

$$F(x) = A(x - c)^2 + B(x - c) + C, \quad (2.15)$$

kde  $A, B$  a  $C$  jsou neznámé parametry, které musíme vypočítat, abychom mohli zjistit minimum aproximace naší účelové funkce, zatímco  $c$  je bod.

Tato metoda opět, jako doposud všechny metody, kromě metody tečen, lze použít na unimodální účelovou funkci. Pro větší přehlednost výpočtů vynecháme indexaci proměnných podle kroků metody.

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.3. INTERPOLAČNÍ METODY

---

**Věta 2.2.** *Mějme zadanou účelovou funkci  $f(x)$  a zjištěnou její kvadratickou aproximaci  $F(x)$ . Potom minimum kvadratické interpolace se bude nacházet v bodě  $d$ , který lze vyjádřit*

$$d = c - \frac{B}{2A} \quad (2.16)$$

*Důkaz.* Bod minima kvadratické aproximace nalezneme pomocí položení první derivace podle  $x$  rovno nule a následném vyjádření  $x$

$$\begin{aligned} 2A(x - c) + B &= 0, \\ 2A(x - c) &= -B, \\ x - c &= -\frac{B}{2A}, \\ x &= c - \frac{B}{2A}. \end{aligned}$$

Tento bod označíme jako bod  $d$ . □

Pro vyřešení úlohy budeme potřebovat určit trojici neznámých parametrů  $A, B, C$ . Abychom je určili potřebujeme 3 funkční hodnoty účelové funkce, zde tedy v metodě nastane chvíle, kdy musíme zvolit 3 body a v těchto bodech zjistit funkční hodnoty. Tyto funkční hodnoty účelové funkce prohlásíme zároveň za funkční hodnoty kvadratické aproximace. Tedy

$$f(a) = F(a), \quad f(b) = F(b), \quad f(c) = F(c),$$

přičemž  $a < c < b$ . Zároveň ze vztahu (2.15) musí platit, že

$$F(c) = C,$$

tedy známe jeden z neznámých parametrů a zbylé dva můžeme snadno dopočítat.

**Věta 2.3.** *Nechť máme účelovou funkci  $f(x)$ , její kvadratickou aproximaci  $F(x)$ , zvolené body  $a, b, c$  a zjištěné jejich funkční hodnoty. Nyní můžeme získat analytické vyjádření bodu  $d$ , tedy minima kvadratické interpolace*

$$d = c + \frac{(f(a) - f(c))(b - c)^2 - (f(b) - f(c))(a - c)^2}{(f(a) - f(c))(b - c) - (f(b) - f(c))(a - c)}. \quad (2.17)$$

*Důkaz.* Nechť již máme zvolené body  $a, b, c$  a známe jejich funkční hodnoty. Potom dosazením hodnoty  $a$  za  $x$  do rovnice (2.15) a nahrazením parametru  $C$  funkční hodnotou  $f(c)$  získáme

$$\begin{aligned} f(a) &= A(a - c)^2 + B(a - c) + f(c) \\ \frac{f(a) - f(c)}{a - c} &= A(a - c) + B. \end{aligned}$$

KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE  
2.3. INTERPOLAČNÍ METODY

---

Analogicky i dosazením  $b$  získáme

$$\frac{f(b) - f(c)}{b - c} = A(b - c) + B.$$

Nyní vyjádříme  $A$  a  $B$  a vyřešíme soustavu dvou rovnic s dvěma neznámými.

$$A = \frac{f(a) - f(c)}{(a - c)^2} - \frac{B}{a - c},$$

$$B = \frac{f(b) - f(c)}{b - c} - A(b - c).$$

Dosazením druhé rovnice do první získáváme

$$A = \frac{f(a) - f(c)}{(a - c)^2} - \frac{1}{a - c} \left( \frac{f(b) - f(c)}{b - c} - A(b - c) \right),$$

$$A - \frac{A(b - c)}{a - c} = \frac{f(a) - f(c)}{(a - c)^2} - \frac{f(b) - f(c)}{(b - c)(a - c)},$$

$$\frac{A(a - b)}{a - c} = \frac{f(a) - f(c)}{(a - c)^2} - \frac{f(b) - f(c)}{(b - c)(a - c)},$$

$$A(a - b) = \frac{f(a) - f(c)}{a - c} - \frac{f(b) - f(c)}{b - c},$$

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

Stejným způsobem získáme vyjádření  $B$

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

Obě vyjádření dosadíme do rovnice (2.16)

$$d = c - \frac{1}{2} \left( \frac{\frac{1}{a - b} \left( \frac{(f(b) - f(c))(a - c)}{b - c} - \frac{(f(a) - f(c))(b - c)}{a - c} \right)}{\frac{1}{a - b} \left( \frac{f(a) - f(c)}{a - c} - \frac{f(b) - f(c)}{b - c} \right)} \right).$$

Zlomek rozšíříme  $(a - b)(b - c)(a - c)$

$$d = c - \frac{1}{2} \left( \frac{(f(b) - f(c))(a - c)^2 - (f(a) - f(c))(b - c)^2}{(f(a) - f(c))(b - c) - (f(b) - f(c))(a - c)} \right).$$

Pro lepší přehlednost využijeme znaménko před zlomkem

$$d = c + \frac{1}{2} \frac{(f(a) - f(c))(b - c)^2 - (f(b) - f(c))(a - c)^2}{(f(a) - f(c))(b - c) - (f(b) - f(c))(a - c)}.$$

□

## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.3. INTERPOLAČNÍ METODY

---

Nyní máme 3 zvolené body  $a, b, c$  a jeden bod vypočítaný  $d$ . Můžeme tedy přistoupit k zužování intervalu lokalizace minima. Při tomto postupu bude primárně záležet na pozici bodu  $d$  vůči bodu  $c$ , proto si možnosti rozdělíme na dvě situace, které vyjádříme pro obecný  $k$ -tý krok.

$$\begin{array}{ll}
 d_k < c_k : & d_k > c_k : \\
 f(d_k) < f(c_k) \Rightarrow & f(d_k) < f(c_k) \Rightarrow \\
 \begin{array}{l} a_{k+1} = a_k \\ c_{k+1} = d_k \\ b_{k+1} = c_k \end{array} & \begin{array}{l} a_{k+1} = c_k \\ c_{k+1} = d_k \\ b_{k+1} = b_k \end{array} \\
 f(d_k) > f(c_k) \Rightarrow & f(d_k) > f(c_k) \Rightarrow \\
 \begin{array}{l} a_{k+1} = d \\ c_{k+1} = c_k \\ b_{k+1} = b_k \end{array} & \begin{array}{l} a_{k+1} = a_k \\ c_{k+1} = c_k \\ b_{k+1} = d_k \end{array}
 \end{array}$$

V této chvíli jsme dokončili jeden krok metody a máme opět připraveny body  $a, b, c$  na krok další. Ukončení procesu metody není závislý na počtu kroků, ale na vzdálenosti po sobě nalezených bodů minima kvadratické funkce.

$$|d_{k+1} - d_k| < \Delta.$$

Kde  $\Delta$  je požadovaná přesnost výpočtu.

**Příklad 5.** Pro danou účelovou funkci najděte bod minima pomocí metody kvadratické interpolace. Víte, že  $f(a_0) = f(0,1) = 5,45$ ,  $f(b_0) = f(10) = 4,9$ ,  $f(c_0) = f(2) = 1,714$ . Koeficient udávající požadovanou přesnost je  $\Delta = 0,1$ .

**Řešení.** První výpočet provedeme pomocí neznámých parametrů  $A$  a  $B$ . Následující výpočty již provedeme pomocí zjednodušeného vzorce.

$$\begin{aligned}
 A &= \frac{f(a_0) - f(c_0)}{(a_0 - c_0)(a_0 - b_0)} - \frac{f(b_0) - f(c_0)}{(b_0 - c_0)(a_0 - b_0)} = \frac{5,45 - 1,714}{(0,1 - 2)(0,1 - 10)} - \frac{4,9 - 1,714}{(10 - 2)(0,1 - 10)} = \\
 &= 0,239,
 \end{aligned}$$

$$\begin{aligned}
 B &= \frac{(f(b_0) - f(c_0))(a_0 - c_0)}{(b_0 - c_0)(a_0 - b_0)} - \frac{(f(a_0) - f(c_0))(b_0 - c_0)}{(a_0 - c_0)(a_0 - b_0)} = \frac{(4,9 - 1,714)(0,1 - 2)}{(10 - 2)(0,1 - 10)} - \\
 &- \frac{(5,45 - 1,714)(10 - 2)}{(0,1 - 2)(0,1 - 10)} = -1,512.
 \end{aligned}$$

Takto vypočtené hodnoty můžeme použít pro získání hledaného dělicího bodu

$$d_0 = c_0 - \frac{B}{2A} = 2 - \frac{(-1,512)}{2 \cdot 0,239} = 5,163.$$

Nyní porovnáme funkční hodnoty naší účelové funkce v bodech  $c_0$  a  $d_0$ , abychom zjistili jakým způsobem provést přeznačení stávajících bodů a zúžili tak interval lokalizace minima

$$f(c_0) = f(2) = 1,714 > 1,133 = f(5,163) = f(d_0).$$

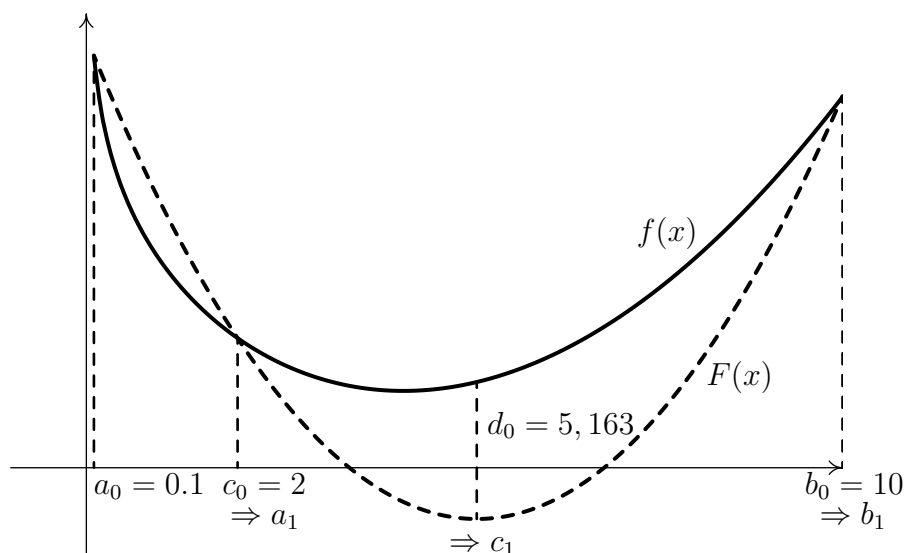
## KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE

### 2.3. INTERPOLAČNÍ METODY

---

Jelikož  $d_0 > c_0$  a  $f(c_0) > f(d_0)$ , provedeme přeznačení

$$a_1 = c_0, \quad c_1 = d_0, \quad b_1 = b_0.$$



Další krok provedeme již pomocí zjednodušeného vzorce

$$\begin{aligned} d_1 &= c_1 + \frac{1}{2} \frac{(f(a_1) - f(c_1))(b_1 - c_1)^2 - (f(b_1) - f(c_1))(a_1 - c_1)^2}{(f(a_1) - f(c_1))(b_1 - c_1) - (f(b_1) - f(c_1))(a_1 - c_1)} = \\ &= 5,163 + \frac{1}{2} \frac{(1,174 - 1,133)(10 - 5,163)^2 - (4,9 - 1,133)(2 - 5,163)^2}{(1,174 - 1,133)(10 - 5,163) - (4,9 - 1,133)(2 - 5,163)} = \\ &= 4,344. \end{aligned}$$

Nyní zkontrolujeme, zda jsme nedosáhli již dostatečné přesnosti

$$|d_1 - d_0| = |4,344 - 5,163| = 0,819 > 0,1 = \Delta.$$

Podmínka pro ukončení výpočtu není splněna, proto pokračujeme v přeznačení stávajících bodů a nalezení nového bodu  $d$ . Začneme opět porovnáním funkčních hodnot v dělicích bodech

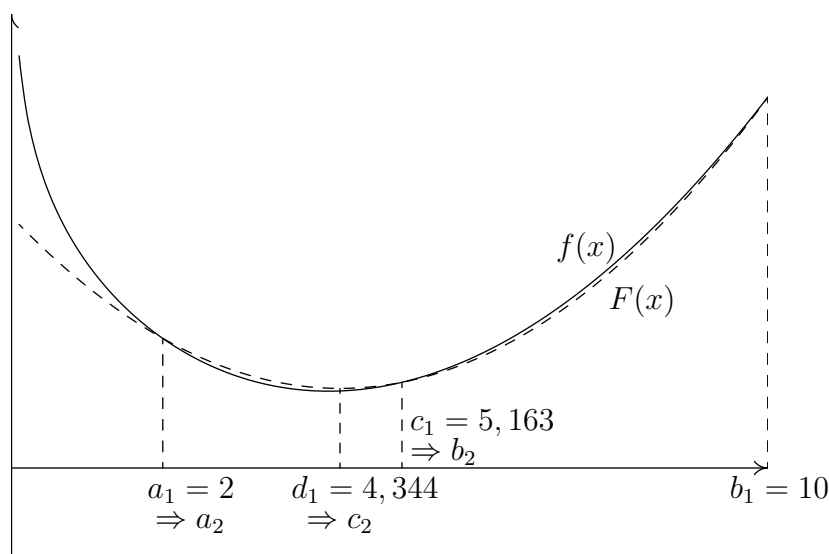
$$f(c_1) = f(5,163) = 1,133 > 1,018 = f(4,344) = f(d_1).$$

Jelikož  $d_1 < c_1$  a  $f(c_1) > f(d_1)$ , provedeme přeznačení

$$a_2 = a_1, \quad c_2 = d_1, \quad b_2 = c_1.$$

KAPITOLA 2. METODY JEDNOROZMĚRNÉ OPTIMALIZACE  
2.3. INTERPOLAČNÍ METODY

---



Z obrázku vidíme, že naše aproximace se v relevantní oblasti již velice blíží naší neznámé funkci. Získáme nový bod  $d$  a zkontrolujeme, zda jsme nedosáhli dostatečné přesnosti.

$$\begin{aligned} d_2 &= c_2 + \frac{1}{2} \frac{(f(a_2) - f(c_2))(b_2 - c_2)^2 - (f(b_2) - f(c_2))(a_2 - c_2)^2}{(f(a_2) - f(c_2))(b_2 - c_2) - (f(b_2) - f(c_2))(a_2 - c_2)} = \\ &= 4,344 + \frac{1}{2} \frac{(1,174 - 1,018)(5,163 - 4,344)^2 - (1,133 - 1,018)(2 - 4,344)^2}{(1,174 - 1,018)(10 - 5,163) - (4,9 - 1,133)(2 - 5,163)} = \\ &= 4,246. \end{aligned}$$

$$|d_2 - d_1| = |4,246 - 4,344| = 0,098 < 0,1 = \Delta.$$

Podmínka požadované přesnosti nám říká, že jsme již dostatečné přesnosti dosáhli a můžeme výpočet ukončit. Za hledaný bod minima budeme považovat bod  $x_{\min} = d_2 = 4,246$  s funkční hodnotou  $f(4,246) = 1,0161$ . Námi minimalizovaná funkce byla

$$f(x) = \frac{(x-3)^2}{10} - \ln(2x) + 3.$$

Tato funkce má minimum v bodě  $x^* = 4,193$  s funkční hodnotou  $f(4,193) = 1,0157$ . Náš výpočet můžeme tedy považovat za velice přesný. Ačkoliv byl příklad početně náročný, tak počet potřebných zjišťování funkčních hodnot nebyl nikterak velký. Pokud započteme i zadané hodnoty, tak jsme potřebovali zjistit celkem šest funkčních hodnot.

---

## Kapitola 3

# Metody vícerozměrné optimalizace

Metody vícerozměrné optimalizace můžeme, stejně jako metody jednorozměrné, rozdělit do podkategorií podle toho, kolik informací potřebujeme k tomu, abychom je mohli použít. Metody vícerozměrné optimalizace dělíme na

- metody komparativní,
- metody gradientní,
- metody newtonovské.

Názvy jednotlivých kategorií vypovídají o fungování daných metod stejně jako při jednorozměrné optimalizaci. Při použití metody komparativní nebudeme potřebovat znalost derivací účelové funkce, avšak obecně se budeme muset smířit s nejpomalejší konvergencí. Metody gradientní se vyznačují vyšší rychlostí konvergence, ale můžeme je použít jen v případech, že jsme schopni získat hodnoty gradientu. Při použití metody z poslední kategorie, z metod newtonovských, nám již nebude stačit znalost hodnot gradientu, ale k výpočtu musíme využít i hodnoty Hesseho matice. Stejně jako u metod jednorozměrné optimalizace i zde obecně platí, že čím výpočetně složitější postup jedné iterace, tím je rychlejší konvergence metody.

Postup řešení úloh vícerozměrné optimalizace je téměř pro každou úlohu obecně stejný. Ať již s pomocí porovnávání funkčních hodnot, nebo z hodnot gradientu, se pro danou metodu určí nejvhodnější směr pro jednorozměrnou optimalizaci a ta se následně provede. Tímto získáme nový bod nacházející se blíže minima a celý postup opakujeme. Metody se od sebe liší ve dvou skutečnostech. První, tou významnější, je způsob hledání směru, ve kterém bude prováděna jednorozměrná optimalizace. Druhý rozdíl tkví v odlišném provádění samotné jednorozměrné optimalizace, kdy se metody komparativní liší, jelikož nemůžeme použít minimalizačního postupu pomocí derivace.



### 3.1 Komparativní metody

Mějme  $n$ -rozměrnou funkci  $y = f(x_1, \dots, x_n) = f(\mathbf{x})$ , pro kterou neznáme hodnoty gradientu, ať již z důvodu nemožnosti, či vysoké nákladnosti jejich pořízení. Jediným zdrojem informací pro nás bude zjišťování funkčních hodnot v námi zvolených bodech.

#### 3.1.1 Metoda Hooka a Jeevese

Zvláštností první, z námi probíraných, metod vícerozměrné optimalizace je to, že samotné hledání směru pro jednorozměrnou optimalizaci připomíná několik po sobě jdoucích jednorozměrných optimalizací. Metoda využívá souřadnicového hledání, tedy vždy ve směru jedné ze souřadnicových os podle určitého klíče zvolí bod, pro který je zjištěna funkční hodnota, která je následně porovnána s výchozím bodem. Pokud je funkční hodnota nově nalezeného bodu menší, je považován za dosavadní minimum pro danou osu a nadále se postupuje z něj. Když je takto nalezeno minimum ve směru jedné souřadnicové osy, provede se hledání a porovnání podél další osy a tak dále. Jakmile provedeme minimalizace podél veškerých os, získáme druhý bod tvořící společně s výchozím bodem přímkou, po které se pokusíme přiblížit co nejvíce k minimu dané účelové funkce. Zde je tato optimalizace omezena na hledání dalších bodů ve směru od  $\mathbf{x}_i$  k  $\mathbf{x}_{i+1}$ , který jsme našli poslední optimalizací podél os. Jakmile narazíme na bod, který má vyšší funkční hodnotu než bod předešlý, tak postup ukončíme a budeme pokračovat z předposledního nalezeného bodu pomocí souřadnicového hledání. Metoda začíná volbou prvního opěrného, výchozího, bodu  $\mathbf{x}_0 \in \mathbb{R}^n$ . Tento bod je zvolen náhodně, případně odhadem, který náleží mimo metodu samotnou.

Jak jsme již naznačili,  $\mathbf{x}_i$  budeme značit opěrný bod po  $i$  krocích, dále označíme  $\hat{\mathbf{x}}$  za umístění doposud vypočtené minimální hodnoty účelové funkce a jako poslední  $\bar{\mathbf{x}}$  budeme označovat bod, ve kterém budeme při každé optimalizaci zjišťovat funkční hodnotu.

Na začátku metody musíme zvolit dostatečně malé  $\Delta > 0$ , které bude využito pro ukončení iterování při dosažení určité vyžadované přesnosti. Další důležitou volbou je volba vektoru  $\mathbf{d} = (d_1, \dots, d_n) \in \mathbb{R}^n$ . Jednotlivé hodnoty  $d_i$  budou určovat délku kroku při optimalizaci podél  $i$ -té osy, proto je považováno za standardní, že jsou všechny hodnoty  $d_i$  stejné.

Nyní máme vše připravené na provedení samotného kroku metodou Hooka a Jeevese. Mějme výchozí bod  $\mathbf{x}_0$ . Pro tento bod vypočítáme hodnotu účelové funkce  $f(\mathbf{x}_0)$ . Jelikož je  $\mathbf{x}_0$  naším prvním bodem, u kterého známe funkční hodnoty, provedeme označení  $\mathbf{x}_0 = \bar{\mathbf{x}} = \hat{\mathbf{x}}$ , tedy platí  $f(\mathbf{x}_0) = f(\bar{\mathbf{x}}) = f(\hat{\mathbf{x}})$ . Nyní nastává fáze, kdy budeme postupně porovnávat funkční hodnoty bodů nalezených ve směru všech  $n$  os. Při prvním vyhledávání provedeme

$$\bar{\mathbf{x}} = \hat{\mathbf{x}} + d_1 \mathbf{e}_1, \quad (3.1)$$

kde  $\mathbf{e}_1$  je jednotkový vektor s jedničkou a první pozici, tedy  $\mathbf{e}_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ . Vypočítáme funkční hodnotu nově nalezeného bodu  $f(\bar{\mathbf{x}})$  a porovnáme s dosavadním mi-

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.1. KOMPARATIVNÍ METODY

---

nimem  $f(\hat{\mathbf{x}})$ . Nyní mohou nastat dvě možnosti:

Možnost **a)**

V případě, že  $f(\bar{\mathbf{x}}) < f(\hat{\mathbf{x}})$  jsme se tímto posunem dostali blíže k minimu a provedeme přeznačení  $\hat{\mathbf{x}} = \bar{\mathbf{x}}$ .

Možnost **b)**

Pokud ovšem platí  $f(\bar{\mathbf{x}}) > f(\hat{\mathbf{x}})$ , tak daný bod  $\bar{\mathbf{x}}$  zavrhneme a spočteme  $\bar{\mathbf{x}} = \hat{\mathbf{x}} - d_1 \mathbf{e}_1$ <sup>1</sup>. Pro tento bod také zjistíme funkční hodnotu, porovnáme s dosavadním minimem a případně přeznačíme. Pokud ani tento bod nevyhovuje, zůstáváme na bodě stávajícím a pokračujeme prohledáváním podél další osy.

Tímto jsme ukončili prohledávání podél první z os a můžeme pokračovat. Vypočteme následující bod  $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = \hat{\mathbf{x}} + d_2 \mathbf{e}_2,$$

kde  $\mathbf{e}_2 = (0, 1, 0, \dots, 0)^T \in \mathbb{R}^n$  a postupujeme při porovnání funkčních hodnot stejně jako u prvního prohledávání.

Prověřením posledního bodu  $\bar{\mathbf{x}} = \hat{\mathbf{x}} + d_n \mathbf{e}_n$ , případně  $\bar{\mathbf{x}} = \hat{\mathbf{x}} - d_n \mathbf{e}_n$ , je ukončen jeden krok hledání minima podél os. Pokud jsme během tohoto kroku nezjistili zlepšení oproti výchozímu bodu  $\hat{\mathbf{x}} = \mathbf{x}_0$ , bude minimum zřejmě tak blízko tohoto bodu, že zvolené hodnoty  $\mathbf{d} = (d_1, \dots, d_n)$  jsou příliš velké. Provedeme proto  $\mathbf{d} = \frac{1}{2} \mathbf{d}$  a pokračujeme nadále s poloviční délkou kroku. Postup opakujeme dokud nebude splněno

$$\max \mathbf{d} \leq \Delta, \tag{3.2}$$

kde  $\Delta$  je na začátku zvolený skalár určující vyžadovanou přesnost výpočtu.

Pokud je ovšem během daného kroku zlepšení dosaženo, tedy  $f(\mathbf{x}_0) > f(\hat{\mathbf{x}})$ , označíme bod  $\hat{\mathbf{x}}$  za  $\mathbf{x}_1 = \hat{\mathbf{x}}$ , takto nalezený bod nazveme opěrným bodem. Nyní však nezačneme další krok hledání minima podél os, ale provedeme posun po přímce spojující poslední bod minimalizující posloupnosti s nově nalezeným opěrným bodem (v našem případě mezi  $\mathbf{x}_0$  a  $\mathbf{x}_1$ ). Provedeme výpočet

$$\bar{\mathbf{x}} = 2\mathbf{x}_1 - \mathbf{x}_0.$$

Pro daný bod zkontrolujeme, zda  $f(\bar{\mathbf{x}}) < f(\mathbf{x}_1)$ . Pokud tato podmínka platí, tak nás krok přiblížil k minimu a provedeme přeznačení  $\mathbf{x}_2 = \bar{\mathbf{x}}$  a pokračujeme v hledání dalšího bodu posunem

$$\bar{\mathbf{x}} = 2\mathbf{x}_2 - \mathbf{x}_1.$$

Jakmile některý z bodů nebude splňovat podmínku  $f(\bar{\mathbf{x}}) < f(\mathbf{x}_i)$ , tak tento postup zastavíme a budeme pokračovat hledáním podél os v posledním bodě, který nás přiblížil minimu.

---

<sup>1</sup>Použití znamének v opačném pořadí by nijak neovlivnilo fungování metody

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.1. KOMPARATIVNÍ METODY

---

**Příklad 6.** Proved'te výpočet metodou Hooka a Jeevese pro funkci  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Znáte následující informace:

Výchozí bod  $\mathbf{x}_0 = \widehat{\mathbf{x}} = (0,8; 3)^T$ , vektor velikosti kroku  $\mathbf{d} = (0,8; 0,8)$ .

**Řešení.** Jako první zjistíme funkční hodnotu  $f(\mathbf{x}_0) = 8,34$ . Z počátečního bodu  $\mathbf{x}_0$  vyhledáme testovaný bod ve směru osy  $x$

$$\bar{\mathbf{x}} = \widehat{\mathbf{x}} + d_1 \mathbf{e}_1 = \begin{pmatrix} 0,8 \\ 3 \end{pmatrix} + 0,8 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1,6 \\ 3 \end{pmatrix}.$$

Zjistíme funkční hodnotu a porovnáme s funkční hodnotou v  $\widehat{\mathbf{x}}$

$$f(\bar{\mathbf{x}}) = f\left(\begin{pmatrix} 1,6 \\ 3 \end{pmatrix}\right) = 5,46 < 8,34 = f(\widehat{\mathbf{x}}) = f(\mathbf{x}_0).$$

Jelikož má nový bod funkční hodnotu nižší než  $\mathbf{x}_0$ , označíme ho za dosavadní umístění minima  $\widehat{\mathbf{x}} = \bar{\mathbf{x}}$ . Nyní musíme ještě provést hledání podél druhé osy

$$\bar{\mathbf{x}} = \widehat{\mathbf{x}} + d_2 \mathbf{e}_2 = \begin{pmatrix} 1,6 \\ 3 \end{pmatrix} + 0,8 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1,6 \\ 3,8 \end{pmatrix}.$$

Opět zjistíme a porovnáme funkční hodnotu s  $\widehat{\mathbf{x}}$

$$f(\bar{\mathbf{x}}) = f\left(\begin{pmatrix} 1,6 \\ 3,8 \end{pmatrix}\right) = 7,7 > 5,46 = f(\widehat{\mathbf{x}}).$$

Zjištěná funkční hodnota pro nově nalezený bod  $\bar{\mathbf{x}}$  je větší než u bodu  $\widehat{\mathbf{x}}$ , proto je bod nevyhovující a musíme provést druhý výpočet

$$\bar{\mathbf{x}} = \widehat{\mathbf{x}} - d_2 \mathbf{e}_2 = \begin{pmatrix} 1,6 \\ 3 \end{pmatrix} - 0,8 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1,6 \\ 2,2 \end{pmatrix},$$

$$f(\bar{\mathbf{x}}) = f\left(\begin{pmatrix} 1,6 \\ 2,2 \end{pmatrix}\right) = 4,5 < 5,46 = f(\widehat{\mathbf{x}}).$$

Pro tento vypočítaný bod je funkční hodnota nižší než u  $\widehat{\mathbf{x}}$ , proto provedeme přeznačení  $\widehat{\mathbf{x}} = \bar{\mathbf{x}}$ .

Nyní jsme ukončili fázi vyhledávání bodů podle souřadnicových os a získali jsme bod, jehož funkční hodnota je menší než funkční hodnota našeho posledního opěrného bodu

$$f(\widehat{\mathbf{x}}) = f\left(\begin{pmatrix} 1,6 \\ 2,2 \end{pmatrix}\right) = 4,5 < 8,34 = f(\mathbf{x}_0).$$

Provedeme tedy přeznačení bodu  $\widehat{\mathbf{x}}$  na nový opěrný bod  $\mathbf{x}_1 = \widehat{\mathbf{x}} = (1,6; 2,2)^T$ .

Nastává fáze posunu po přímce spojující dva opěrné body. Podle vzorce vypočítáme další bod  $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = 2\mathbf{x}_1 - \mathbf{x}_0 = 2 \begin{pmatrix} 1,6 \\ 2,2 \end{pmatrix} - \begin{pmatrix} 0,8 \\ 3 \end{pmatrix} = \begin{pmatrix} 2,4 \\ 1,4 \end{pmatrix}.$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.2. GRADIENTNÍ METODY

---

Našli jsme bod posunu a zkontrolujeme, zda funkční hodnota v tomto bodě je menší nebo větší než funkční hodnota v  $\mathbf{x}_1$

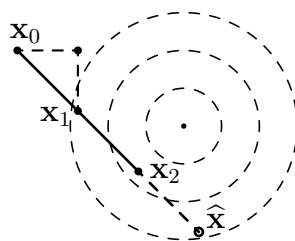
$$f(\bar{\mathbf{x}}) = f\left(\begin{array}{c} 2,4 \\ 1,4 \end{array}\right) = 3,22 < 4,5 = f(\mathbf{x}_1).$$

Provedeme tedy přeznačení  $\mathbf{x}_2 = \bar{\mathbf{x}}$ , vyhledáme další bod porovnáme funkční hodnoty

$$\bar{\mathbf{x}} = 2\mathbf{x}_2 - \mathbf{x}_1 = 2\left(\begin{array}{c} 2,4 \\ 1,4 \end{array}\right) - \left(\begin{array}{c} 1,6 \\ 2,2 \end{array}\right) = \left(\begin{array}{c} 3,2 \\ 0,6 \end{array}\right),$$

$$f(\bar{\mathbf{x}}) = f\left(\begin{array}{c} 3,2 \\ 0,6 \end{array}\right) = 4,5 > 3,22 = f(\mathbf{x}_2).$$

Došli jsme k závěru, že funkční hodnota dalšího nalezeného bodu je vyšší než funkční hodnota bodu předcházejícího, tudíž ukončíme hledání po přímce spojující opěrné body a opakujeme souřadnicové hledání z bodu  $\mathbf{x}_2 = (2,4; 1,4)^T$



Dodejme, že funkce, kterou jsme se snažili minimalizovat v tomto příkladě, byla  $f(x, y) = (x - 3)^2 + (y - 2)^2 + 2,5$  (soustředné kružnice v obrázku výše jsou vrstevnicemi tohoto konvexního paraboloidu). Z obrázku můžeme vidět, že při dalším souřadnicovém hledání bychom se dostali do bodu  $(3,2; 2,2)^T$ , tedy tak blízko minima, že bychom pravděpodobně výpočet ukončili pro dosažení potřebné přesnosti.

## 3.2 Gradientní metody

Na rozdíl od předešlé metody budeme nyní vyžadovat diferencovatelnost účelové funkce  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ . Tedy budeme mít k dispozici gradient

$$\text{grad } f(\mathbf{x}) = \left(\begin{array}{c} f_{x_1}(\mathbf{x}) \\ \vdots \\ f_{x_n}(\mathbf{x}) \end{array}\right) = f'(\mathbf{x}) = \nabla f(\mathbf{x}).$$

### 3.2.1 Metoda nejrychlejšího spádu

Pravděpodobně nejjednodušší gradientní metoda, a proto v této sekci začneme právě s ní. Myšlenka za touto metodou staví na faktu, že pokud v každém bodě najdeme směr, ve kterém funkce nejrychleji klesá a posuneme se do minima nacházejícího se v tomto směru, tak náš postup bude konvergovat k minimu účelové funkce. Problém s využitím této metody může nastat, pokud má funkce více bodů lokálního minima. Jakmile se dostaneme do určitého okolí jednoho z minim, tak již touto metodou nemůžeme uniknout, nehledě na to, zda je minimum globální nebo ne.

Metoda nejrychlejšího spádu bere za směr každého obecně  $k$ -tého kroku zápornou hodnotu gradientu, jakožto směr, ve kterém funkce nejrychleji klesá, tedy  $\mathbf{s}_k = -f'(\mathbf{x}_k)$ . Délku kroku  $\alpha_k$  v této metodě zjistíme až při nalezení následujícího bodu, jelikož tento bod je nalezen pomocí jednorozměrné minimalizace podél takto určeného směru. Obecně v kroku  $k + 1$  získáme nový bod minima následujícím výpočtem

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k = \mathbf{x}_k - \alpha_k f'(\mathbf{x}_k) = \arg \min_{\alpha > 0} f(\mathbf{x}_k - \alpha f'(\mathbf{x}_k)). \quad (3.3)$$

Takto upravenou minimalizaci můžeme zapsat

$$\min_{\alpha > 0} f(\mathbf{x}_k - \alpha f'(\mathbf{x}_k)) = \min_{\alpha > 0} \varphi_k(\alpha) = \varphi_k(\alpha_k) \quad (3.4)$$

**Věta 3.1.** *Ve dvou po sobě jdoucích krocích získané směry postupu  $\mathbf{s}_k$ ,  $\mathbf{s}_{k+1}$  nalezené pomocí metody nejrychlejšího spádu, pro něž platí  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k$ , jsou k sobě kolmé.*

*Důkaz.* Z (3.4) plyne, že

$$\left[ \frac{\partial \varphi_k(\alpha)}{\partial \alpha} \right]_{\alpha_k} = 0.$$

Můžeme tedy tento výraz rozepsat

$$\left[ \frac{\partial \varphi_k(\alpha)}{\partial \alpha} \right]_{\alpha_k} = \left[ (\nabla_{\mathbf{x}} f(\mathbf{x}))^T \cdot \frac{\partial \mathbf{x}}{\partial \alpha} \right]_{\alpha_k}^2 = -[\mathbf{s}(\mathbf{x}_k - \alpha_k \mathbf{s}_k)]^T \mathbf{s}(\mathbf{x}_k) = -\mathbf{s}(\mathbf{x}_{k+1})^T \mathbf{s}(\mathbf{x}_k) = 0$$

□

I podmínka ukončení iteračního procesu využívá hodnot gradientu

$$\sqrt{\mathbf{s}_k^T \mathbf{s}_k} = \sigma < \varepsilon.$$

---

<sup>2</sup>Tento výraz udává derivaci celé funkce násobený derivací vnitřní složky.  $\alpha$  je jednou z částí  $\mathbf{x}$ , a proto funkci  $f$  derivujeme podle  $\mathbf{x}$ . Následně derivujeme vnitřní složku, tedy samotné  $\mathbf{x}$  podle  $\alpha$ .

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.2. GRADIENTNÍ METODY

---

**Příklad 7.** Proveďte jeden krok minimalizace metodou nejrychlejšího spádu a ověřte kolmost následujícího směru. Minimalizovanou účelovou funkcí je

$$f(\mathbf{x}) = f(x, y) = \frac{x^2}{25} + \frac{y^2}{4} + 1.$$

Počáteční bodem je  $\mathbf{x}_0 = (4; 1, 2)^T$ .

**Řešení.** Jako první vypočítáme směr, ve kterém budeme hledat následující bod minimalizující posloupnosti. Derivace účelové funkce jsou  $f_x = \frac{2x}{25}$  a  $f_y = \frac{y}{2}$ , tedy směrový vektor se rovná

$$s_0 = -f'(\mathbf{x}_0) = \begin{pmatrix} -0,32 \\ -0,6 \end{pmatrix}.$$

Nyní víme, že následující bod bude na polopřímce

$$\mathbf{x} = (4 - 0,32\alpha; 1, 2 - 0,6\alpha)^T.$$

Po dosazení získáme

$$\varphi_0(\alpha) = \frac{(4 - 0,32\alpha)^2}{25} - \frac{(1, 2 - 0,6\alpha)^2}{4} + 1$$

Tuto rovnici upravíme do tvaru

$$\varphi_0(\alpha) = 2 - 0,462\alpha + 0,094\alpha^2.$$

Máme najít minimum této funkce vzhledem k  $\alpha$ . Funkci tedy zderivujeme a derivaci položíme rovnu nule

$$\varphi_0'(\alpha) = 0,188\alpha - 0,462 = 0.$$

Tato rovnice nám dává jediný kladný výsledek  $\alpha_0 = 2,457$ . Dosazením získáme následující bod minimalizující posloupnosti

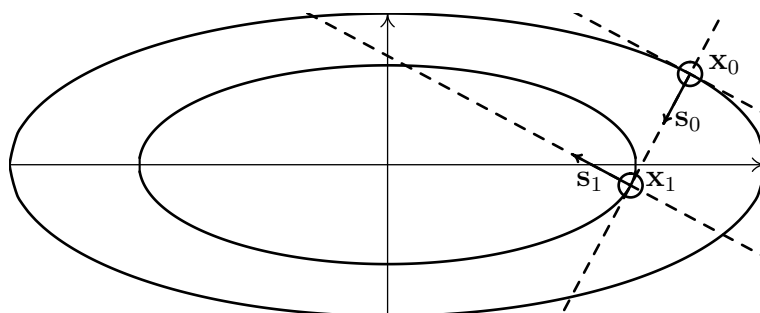
$$\mathbf{x}_1 = \begin{pmatrix} 4 - 0,32\alpha_0 \\ 1, 2 - 0,6\alpha_0 \end{pmatrix} = \begin{pmatrix} 3,214 \\ -0,274 \end{pmatrix}.$$

Pro směr dalšího postupu víme

$$s_1 = -f'(\mathbf{x}_1) = \begin{pmatrix} -0,257 \\ 0,137 \end{pmatrix}.$$

Zkontrolujeme kolmost dvou po sobě následujících směrů

$$s_1^T s_0 = f'(\mathbf{x}_1)^T f'(\mathbf{x}_0) = (-0,257; 0,137) \begin{pmatrix} -0,32 \\ -0,6 \end{pmatrix} = 0$$



### 3.2.2 Metoda nejrychlejšího spádu pro kvadratické funkce

**Věta 3.2.** Pro případ kvadratické účelové funkce

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{b}, \quad (3.5)$$

kde  $\mathbf{A}$  je pozitivně definitní symetrická matice je minimum této kvadratické funkce zároveň řešením soustavy lineárních rovnic  $\mathbf{A} \mathbf{x} = \mathbf{b}$ .

*Důkaz.* Pokud je  $\mathbf{x}^*$  řešením soustavy lineárních rovnic  $\mathbf{A} \mathbf{x} = \mathbf{b}$ , tak porovnáním jeho funkční hodnoty a funkční hodnoty v bodě  $\mathbf{x}^* + \mathbf{h}$  získáme

$$\begin{aligned} f(\mathbf{x}^* + \mathbf{h}) - f(\mathbf{x}^*) &= \frac{1}{2} (\mathbf{x}^* + \mathbf{h})^T \mathbf{A} (\mathbf{x}^* + \mathbf{h}) - (\mathbf{x}^* + \mathbf{h})^T \mathbf{b} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x}^* - \mathbf{x}^{*T} \mathbf{b} = \\ &= \frac{1}{2} \mathbf{h}^T \mathbf{A} \mathbf{h} - \mathbf{h}^T (\mathbf{A} \mathbf{x}^* - \mathbf{b}). \end{aligned}$$

Jelikož jsme předpokládali, že  $\mathbf{x}^*$  je řešením dané soustavy lineárních rovnic, tak je výraz  $\mathbf{h}^T (\mathbf{A} \mathbf{x}^* - \mathbf{b})$  nulový a výsledek je zredukován do tvaru  $\frac{1}{2} \mathbf{h}^T \mathbf{A} \mathbf{h}$ . Daný výraz je vždy kladný, tudíž  $f(\mathbf{x}^*)$  musí být minimum účelové funkce.  $\square$

Pokud je metoda použita právě na kvadratickou funkci, je možné využít vypočteného vzorce pro určení délky posunu mezi dvěma následujícími body minimalizující posloupnosti. K tomuto vzorci jde dojít dvěma možnými postupy, které se liší základní myšlenkou.

**Postup 1)**

Dle metody nejrychlejšího spádu platí  $\mathbf{s}_k = -f'(\mathbf{x}_k) = \mathbf{b} - \mathbf{A} \mathbf{x}_k$ . Dále využijeme rovnice (3.3) a (3.5)

$$\begin{aligned} f(\mathbf{x}_k + \alpha \mathbf{s}_k) - f(\mathbf{x}_k) &= \frac{1}{2} (\mathbf{x}_k + \alpha \mathbf{s}_k)^T \mathbf{A} (\mathbf{x}_k + \alpha \mathbf{s}_k) - (\mathbf{x}_k + \alpha \mathbf{s}_k)^T \mathbf{b} - \frac{1}{2} \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k + \mathbf{x}_k^T \mathbf{b} = \\ &= \frac{1}{2} (\alpha^2 \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k + \alpha \mathbf{s}_k^T \mathbf{A} \mathbf{x}_k + \mathbf{x}_k^T \mathbf{A} \alpha \mathbf{s}_k) - \alpha \mathbf{s}_k^T \mathbf{b} = \\ &= \frac{\alpha^2}{2} \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k + \alpha \mathbf{s}_k^T (\mathbf{A} \mathbf{x}_k - \mathbf{b}) = \frac{\alpha^2}{2} \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k - \alpha \mathbf{s}_k^T \mathbf{s}_k. \end{aligned}$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.2. GRADIENTNÍ METODY

---

Pokud nyní převedeme  $f(\mathbf{x}_k)$  na pravou stranu rovnice, zderivujeme podle  $\alpha$  a výsledek položíme roven nule, tak získáme minimalizaci  $f(\mathbf{x}_k + \alpha \mathbf{s}_k)$  podle koeficientu určujícího délku kroku.

$$\frac{\partial f}{\partial \alpha} = \alpha \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k - \mathbf{s}_k^T \mathbf{s}_k = 0.$$

Úpravou dostaneme vzorec pro výpočet  $\alpha_k$

$$\alpha_k = \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}. \quad (3.6)$$

Při použití metody nejrychlejšího spádu na funkci tvaru (3.5) tedy můžeme definovat minimalizující posloupnost rekurentně jako

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \left( \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} \right) \mathbf{s}_k. \quad (3.7)$$

**Postup 2)**

V blízkém okolí  $|\mathbf{x}^* - \mathbf{x}_k| < \varepsilon$  extrému  $\mathbf{x}^*$  můžeme účelovou funkci aproximovat prvními třemi členy Taylorova mnohočlenu

$$F(\mathbf{x}) = f(\mathbf{x}_k) - (\mathbf{x} - \mathbf{x}_k)^T \mathbf{s}_k + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T \mathbf{H} (\mathbf{x} - \mathbf{x}_k).$$

Vzorec je na rozdíl od vzorce předešlého možné použít i na jiné než kvadratické funkce, avšak vzorec v takovém případě není přesný. Přesnost vzorce pro kvadratickou účelovou funkci vychází z toho, že matice druhých derivací  $\mathbf{H}$ , označovaná jako Hesseho matice, je v tomto případě konstantní. Nyní využijeme této aproximace a dosadíme  $\mathbf{x} = \mathbf{x}_k + \alpha \mathbf{s}_k$

$$F(\mathbf{x}) = f(\mathbf{x}_k) - \alpha \mathbf{s}_k^T \mathbf{s}_k + \frac{1}{2} \alpha^2 \mathbf{s}_k^T \mathbf{H} \mathbf{s}_k.$$

Když tuto rovnici zderivujeme a položíme rovnou nule dostaneme minimalizaci aproximace účelové funkce vzhledem ke koeficientu určujícímu délku kroku  $\alpha$  mezi jednotlivými body minimalizující posloupnosti

$$\frac{\partial F}{\partial \alpha} = \alpha \mathbf{s}_k^T \mathbf{H} \mathbf{s}_k - \mathbf{s}_k^T \mathbf{s}_k = 0.$$

Úpravou získáme vzorec

$$\alpha_k = \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{H} \mathbf{s}_k}.$$

Opět můžeme vyjádřit minimalizující posloupnost rekurentním vzorcem

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \left( \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{H} \mathbf{s}_k} \right) \mathbf{s}_k.$$

Jelikož při kvadratickou účelovou funkci platí  $\mathbf{A} = \mathbf{H}$ , dostali jsme se ke stejnému vzorci jako prvním postupem.



KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

---

**Věta 3.3.** *Nechť je dána účelová funkce  $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x} - \mathbf{x}^T \mathbf{b}$ , kde  $\mathbf{A}$  je pozitivně definitní symetrická matice. Je-li posloupnost  $\{\mathbf{x}_k\}$  vytvořená metodou nejrychlejšího spádu podle vzorce (3.7), pak je tato posloupnost minimalizující.*

*Důkaz.* Provedeme dosazením do výpočtu vzorce

$$\begin{aligned} f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) &= \frac{\alpha_k^2}{2} \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k - \alpha_k \mathbf{s}_k^T \mathbf{s}_k = \frac{1}{2} \left( \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} \right)^2 \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k - \left( \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} \right) \mathbf{s}_k^T \mathbf{s}_k = \\ &= \frac{1}{2} \frac{(\mathbf{s}_k^T \mathbf{s}_k)^2}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} - \frac{(\mathbf{s}_k^T \mathbf{s}_k)^2}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = -\frac{1}{2} \frac{(\mathbf{s}_k^T \mathbf{s}_k)^2}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}. \end{aligned}$$

Pro lepší přehlednost lze tuto rovnice zapsat

$$f(\mathbf{x}_{k+1}) = f(\mathbf{x}_k) - \frac{1}{2} \frac{(\mathbf{s}_k^T \mathbf{s}_k)^2}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}. \quad (3.8)$$

Z této rovnice je již vidět, že díky pozitivní definitnosti matice  $\mathbf{A}$  bude vždy platit  $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$ , a tedy posloupnost  $\{\mathbf{x}_k\}$  je minimalizující.  $\square$

**Příklad 8.** Provedte jeden krok minimalizace pomocí metody nejrychlejšího spádu a ověřte kolmost následujícího směru pomocí vypočítaného vzorce. Hodnoty jsou stejné jako v předchozím příkladu pro možnost porovnání. Počáteční bodem je  $\mathbf{x}_0 = (4; 1, 2)^T$ . Účelová funkce je definována maticí  $\mathbf{A}$  a vektorem  $\mathbf{b}$

$$\mathbf{A} = \begin{pmatrix} \frac{2}{25} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

**Řešení.** Derivace účelové funkce jsou  $f_x = \frac{2x}{25}$  a  $f_y = \frac{y}{2}$ . Opět jako první vypočítáme směrový vektor, ve kterém budeme postupovat při hledání dalšího bodu minimalizující posloupnosti

$$\mathbf{s}_0 = -f'(\mathbf{x}_0) = \begin{pmatrix} -0,32 \\ -0,6 \end{pmatrix}.$$

Nyní již můžeme použít vzorec

$$\alpha_0 = \frac{\mathbf{s}_0^T \mathbf{s}_0}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} = \frac{(-0,32; -0,6) \begin{pmatrix} -0,32 \\ -0,6 \end{pmatrix}}{(-0,32; -0,6) \begin{pmatrix} \frac{2}{25} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} -0,32 \\ -0,6 \end{pmatrix}} = 2,457.$$

$$\mathbf{x}_1 = \begin{pmatrix} 4 - 0,32\alpha_0 \\ 1,2 - 0,6\alpha_0 \end{pmatrix} = \begin{pmatrix} 3,214 \\ -0,274 \end{pmatrix}.$$

Jelikož nám vyšlo tímto postupem  $\alpha_0$  stejné jako, kdy jsme počítali daný příklad bez použití vzorce, tak již víme, že nový směr je kolmý k směru  $\mathbf{s}_0$ .

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

---

Pro porovnání rychlosti konvergence s dalším příkladem přikládáme hodnoty dalších iterací

$$\mathbf{x}_5 = \begin{pmatrix} 0,5996 \\ -0,0512 \end{pmatrix}, \quad \mathbf{x}_{10} = \begin{pmatrix} 0,0601 \\ 0,0180 \end{pmatrix}, \quad \mathbf{x}_{15} = \begin{pmatrix} 0,0090 \\ -0,0008 \end{pmatrix}. \quad (3.9)$$

Od skutečného bodu minima  $\mathbf{x}^* = (0,0)^T$  se již patnáctá iterace vzdaluje jen minimálně. Konvergenci tedy můžeme považovat za relativně rychlou.

Metoda se potýká s několika problémy s konvergencí. První z těchto problémů je, že pokud bude minimalizovaná funkce dostatečně protáhlá ve směru jedné z os (pro  $n = 2$  by vrstevnice takovéto funkce vypadaly jako protažené elipsy), tak posun ve směru gradientu nás posune k minimu jen málo a konvergence tak bude značně pomalá. Abychom si mohli tuto vlastnost dokázat, musíme nejdříve dokázat jednu pomocnou větu.

**Věta 3.4.** *Nechť  $\mathbf{A}$  je symetrická pozitivně definitní matice o rozměrech  $n \times n$ . Označme  $\lambda_1$  a  $\lambda_n$  její nejmenší a největší vlastní hodnotu. Pak pro každé  $\mathbf{x} \in \mathbb{R}^n$  platí*

$$\|\mathbf{x}\|^4 = (\mathbf{x}^T \mathbf{x})^2 \leq \mathbf{x}^T \mathbf{A} \mathbf{x} \cdot \mathbf{x}^T \mathbf{A} \mathbf{x} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n} (\mathbf{x}^T \mathbf{x})^2.$$

*Důkaz.* Jelikož matici  $\mathbf{A}$  je možné vyjádřit ve tvaru  $\mathbf{A} = \mathbf{R}^T \text{diag} \{ \lambda_1, \dots, \lambda_n \} \mathbf{R}$ , kde  $\mathbf{R}^{-1} = \mathbf{R}^T$ , můžeme bez újmy na obecnosti předpokládat, že  $\mathbf{A} = \text{diag} \{ \lambda_1, \dots, \lambda_n \}$ , tedy diagonální maticí. Potom také  $\mathbf{A}^{-1} = \text{diag} \{ \lambda_1^{-1}, \dots, \lambda_n^{-1} \}$ . Pro  $\mathbf{x} = (x_1, \dots, x_n)^T$  označíme

$$w_i = \frac{x_i^2}{\|\mathbf{x}\|^2}, \quad \lambda = \sum_{i=1}^n \lambda_i w_i, \quad \mu = \sum_{i=1}^n \lambda_i^{-1} w_i. \quad (3.10)$$

Potom platí

$$\frac{(\mathbf{x}^T \mathbf{A} \mathbf{x}) \cdot (\mathbf{x}^T \mathbf{A}^{-1} \mathbf{x})}{\|\mathbf{x}\|^4} = \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\|\mathbf{x}\|^2} \cdot \frac{\mathbf{x}^T \mathbf{A}^{-1} \mathbf{x}}{\|\mathbf{x}\|^2} = \frac{\sum_{i=1}^n \lambda_i x_i^2}{\|\mathbf{x}\|^2} \cdot \frac{\sum_{i=1}^n \lambda_i^{-1} x_i^2}{\|\mathbf{x}\|^2} = \lambda \mu. \quad (3.11)$$

K nalezení maximální hodnoty  $\lambda \mu$  nám pomůže geometrická interpretace. Vytvoříme proto souřadnou soustavu s osami  $\lambda$ ,  $\mu$ , do které budeme značit body  $\mathbf{A}_k$ . Každý bod  $\mathbf{A}_k$  bude mít souřadnice  $[\lambda_k, \lambda_k^{-1}]$ , kde  $k = 1, \dots, n$ . Z takto určených bodů vytvoříme konvexní množinu  $\text{conv} \{ \mathbf{A}_1, \dots, \mathbf{A}_n \}$ , ve které každý bod můžeme vyjádřit ve tvaru

$$[\lambda, \mu] = \left[ \sum_{i=1}^n w_i \lambda_i, \sum_{i=1}^n w_i \lambda_i^{-1} \right],$$

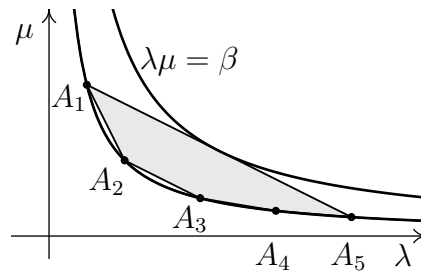
kde  $w_i$  splňují  $w_i \geq 0$ ,  $\sum_{i=1}^n w_i = 1$ .

Nyní si všimněme, že naše  $w_i$  definované jako  $w_i = \frac{x_i^2}{\|\mathbf{x}\|^2}$  přesně splňuje tyto vztahy.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

Hledáme tedy takový bod, kdy součin  $\lambda\mu$  definovaný vztahem (3.11) dosáhne své maximální hodnoty  $\beta$ . Geometrický pohled nám říká, že hledáme takový bod v konvexní množině  $\text{conv}\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ , který leží na hyperbole  $\lambda\mu = \beta$  nejbližší od středu souřadnicového systému, tedy s maximálním  $\beta$ . Je zřejmé, že takovýto bod bude ležet na úsečce spojující body  $\mathbf{A}_1 = [\lambda_1, \lambda_1^{-1}]$  a  $\mathbf{A}_n = [\lambda_n, \lambda_n^{-1}]$ .

Obrázek níže nám ukazuje příklad pro  $n = 5$ . Vyplněná oblast určuje konvexní množinu  $\text{conv}\{\mathbf{A}_1, \dots, \mathbf{A}_5\}$  a máme zde znázorněnu hyperbolu  $\lambda\mu = \beta$  pro maximální  $\beta$ .



Nyní nám zbývá vypočítat toto maximální  $\beta$ . Použijeme znalost směrnic úsečky mezi  $A_1$  a  $A_n$  a tečny k funkci  $\lambda\mu = \beta$ , kterou získáme derivací. Tyto směrnice se musí rovnat, aby  $\beta$  bylo maximální.

$$\begin{aligned} \frac{\partial\mu}{\partial\lambda} &= \frac{\lambda_1^{-1} - \lambda_n^{-1}}{\lambda_n - \lambda_1}, \\ -\frac{\beta}{\lambda^2} &= \frac{-\left(\frac{\lambda_n - \lambda_1}{\lambda_1\lambda_n}\right)}{\lambda_n - \lambda_1}, \\ \beta &= \frac{\lambda^2}{\lambda_1\lambda_n}. \end{aligned}$$

Zde využijeme toho, že  $\lambda = \frac{1}{2}(\lambda_1 + \lambda_n)$ , což znamená, že bod dotyku hyperboly a úsečky spojující  $A_1$  a  $A_n$  se nachází přesně ve středu zmiňované úsečky. Tato skutečnost platí díky tomu, že hyperboly nejsou navzájem nijak posunuty mimo posun způsobený změnou  $\beta$ . Dosazením získáváme

$$\beta = \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}.$$

Tímto je dokázána pravá strana nerovnosti (3.10). Levá strana nerovnosti plyne z Cauchy-Schwarzovy nerovnosti dosazením  $a_i = x_i\sqrt{\lambda_i}$ ,  $b_i = x_i/\sqrt{\lambda_i}$ , kde  $\mathbf{a} = (a_1, \dots, a_n)$ ,  $\mathbf{b} = (b_1, \dots, b_n)$

$$\begin{aligned} (\mathbf{a}^T\mathbf{b})^2 &\leq \|\mathbf{a}\|^2 \cdot \|\mathbf{b}\|^2, \\ (\mathbf{x}^T\mathbf{x})^2 &\leq \mathbf{x}^T \text{diag}\{\lambda_1, \dots, \lambda_n\} \mathbf{x} \cdot \mathbf{x}^T \text{diag}\{\lambda_1^{-1}, \dots, \lambda_n^{-1}\} \mathbf{x}, \\ (\mathbf{x}^T\mathbf{x})^2 &\leq (\mathbf{x}^T\mathbf{A}\mathbf{x}) \cdot (\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}). \end{aligned}$$

□

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.2. GRADIENTNÍ METODY

---

Nyní již můžeme rozebrat rychlost konvergence metody nejrychlejšího spádu. Opět využijeme dokázanou rovnici (3.8)

$$f(\mathbf{x}_{k+1}) = f(\mathbf{x}_k) - \frac{1}{2} \frac{(\mathbf{s}_k^T \mathbf{s}_k)^2}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}.$$

Nyní využijeme toho, že

$$\frac{1}{2} \mathbf{s}_k^T \mathbf{A}^{-1} \mathbf{s}_k = \frac{1}{2} (\mathbf{A} \mathbf{x}_k - b)^T \mathbf{A}^{-1} (\mathbf{A} \mathbf{x}_k - b) = \frac{1}{2} \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k - \mathbf{x}_k^T b = f(\mathbf{x}_k),$$

a pomocí věty (3.4) dostaneme

$$\frac{f(\mathbf{x}_k)}{f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})} = \frac{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k \cdot \mathbf{s}_k^T \mathbf{A}^{-1} \mathbf{s}_k}{(\mathbf{s}_k^T \mathbf{s}_k)^2} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n}. \quad (3.12)$$

Úpravou této nerovnice získáme

$$\begin{aligned} \frac{f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})}{f(\mathbf{x}_k)} &\geq \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2}, \\ 1 - \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2} &\geq \frac{f(\mathbf{x}_{k+1})}{f(\mathbf{x}_k)}, \\ \frac{(\lambda_n - \lambda_1)^2}{(\lambda_n + \lambda_1)^2} &\geq \frac{f(\mathbf{x}_{k+1})}{f(\mathbf{x}_k)}. \end{aligned}$$

Když si nyní označíme

$$a = \frac{(\lambda_n - \lambda_1)^2}{(\lambda_n + \lambda_1)^2} = \left( \frac{1 - \frac{\lambda_1}{\lambda_n}}{1 + \frac{\lambda_1}{\lambda_n}} \right)^2,$$

tak máme vztah  $f(\mathbf{x}_{k+1})/f(\mathbf{x}_k) \leq a$ . Z něho víme, že

$$\begin{aligned} f(\mathbf{x}_1) &\leq a f(\mathbf{x}_0) \\ f(\mathbf{x}_2) &\leq a f(\mathbf{x}_1) \leq a^2 f(\mathbf{x}_0), \\ f(\mathbf{x}_3) &\leq a f(\mathbf{x}_2) \leq a^2 f(\mathbf{x}_1) \leq a^3 f(\mathbf{x}_0), \end{aligned}$$

z čehož lze odvodit vztah

$$f(\mathbf{x}_k) \leq a^k f(\mathbf{x}_0).$$

Tato nerovnice nám objasní problém s konvergencí metody nejrychlejšího spádu. Pokud nastane případ, že  $\lambda_n \gg \lambda_1$ , z čehož plyne, že hodnota  $a$  se bude blížit 1, tak bude metoda konvergovat poměrně pomalu.

Tento problém si ukážeme na následujícím příkladu, jehož rychlost konvergence porovnáme s příkladem minulým.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

---

**Příklad 9.** Metodou nejrychlejšího spádu vypočtete minimum funkce typu (3.5), pro kterou platí

$$\mathbf{A} = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{27}{200} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}.$$

Výchozí bod je  $\mathbf{x}_0 = (0; 3, 9)^T$ .

**Řešení.** Z derivací účelové funkce vypočítáme směrový vektor

$$\mathbf{s}_0 = -f'(\mathbf{x}_0) = \mathbf{b} - \mathbf{A}\mathbf{x}_0 = \begin{pmatrix} 0,05 \\ -0,0265 \end{pmatrix}.$$

S pomocí vzorce získáme  $\alpha_0$

$$\alpha_0 = \frac{\mathbf{s}_0^T \mathbf{s}_0}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} = \frac{(0,05; -0,0265) \begin{pmatrix} 0,05 \\ -0,0265 \end{pmatrix}}{(0,05; -0,0265) \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{27}{200} \end{pmatrix} \begin{pmatrix} 0,05 \\ -0,0265 \end{pmatrix}} = 0,849.$$

Nyní máme vše potřebné pro nalezení dalšího bodu minimalizující posloupnosti

$$\mathbf{x}_1 = \begin{pmatrix} 0 + 0,05\alpha_0 \\ 3,9 - 0,0265\alpha_0 \end{pmatrix} = \begin{pmatrix} 0,0425 \\ 3,8775 \end{pmatrix}.$$

Už při prvním kroku vidíme, že posun z bodu  $\mathbf{x}_0 = (0; 3, 9)^T$  do  $\mathbf{x}_1 = (0,0425; 3,8775)^T$  značí pomalou konvergenci.

$$\mathbf{s}_1 = -f'(\mathbf{x}_1) = \mathbf{b} - \mathbf{A}\mathbf{x}_1 = \begin{pmatrix} -0,0237 \\ -0,0447 \end{pmatrix}.$$

$$\alpha_1 = \frac{\mathbf{s}_1^T \mathbf{s}_1}{\mathbf{s}_1^T \mathbf{A} \mathbf{s}_1} = 1,0441.$$

Další bod minimalizační posloupnosti bude

$$\mathbf{x}_2 = \begin{pmatrix} 0,0425 - 0,0237\alpha_1 \\ 3,8775 - 0,0447\alpha_1 \end{pmatrix} = \begin{pmatrix} 0,0177 \\ 3,8308 \end{pmatrix}.$$

Po dvou krocích jsme se posunuli z  $\mathbf{x}_0 = (0; 3, 9)^T$  do  $\mathbf{x}_2 = (0,0177; 3,8308)^T$  a již můžeme jasně vidět, že v tomto případě není neupravená metoda nejrychlejšího spádu vhodná pro tento výpočet. Abychom ukázali jak pomalu metoda konverguje ke skutečnému minimu v bodě  $\mathbf{x}^* = (1; 0)^T$ , jsou dodány další hodnoty bodů minimalizující posloupnosti.

$$\begin{aligned} \mathbf{x}_5 &= \begin{pmatrix} 0,0761 \\ 3,7412 \end{pmatrix}, & \mathbf{x}_{10} &= \begin{pmatrix} 0,0856 \\ 3,5662 \end{pmatrix}, & \mathbf{x}_{25} &= \begin{pmatrix} 0,2275 \\ 3,1281 \end{pmatrix}, & \mathbf{x}_{100} &= \begin{pmatrix} 0,5913 \\ 1,5938 \end{pmatrix}, \\ \mathbf{x}_{200} &= \begin{pmatrix} 0,8330 \\ 0,6513 \end{pmatrix}, & \mathbf{x}_{400} &= \begin{pmatrix} 0,9721 \\ 0,1088 \end{pmatrix}, & \mathbf{x}_{600} &= \begin{pmatrix} 0,9953 \\ 0,0182 \end{pmatrix}, & \mathbf{x}_{800} &= \begin{pmatrix} 0,9992 \\ 0,0030 \end{pmatrix}. \end{aligned}$$

Druhou nedokonalostí metody nejrychlejšího spádu je fakt, že pokud se netrefíme volbou výchozího bodu do bodu, ze kterého bychom se jedním krokem dostali do bodu minima, tak se nikdy v konečném počtu kroků do bodu minima nedostaneme a metoda nejrychlejšího spádu je i pro kvadratickou funkci nekonečným iteračním procesem. Tato nedokonalost však nemusí znamenat problémem pro kvalitu odhadu.

### 3.2.3 Metoda paralelních tečen

Metoda paralelních tečen je nejjednodušší odpovědí na problém pomalé konvergence metody nejrychlejšího spádu. Dalo by se říci, že tato metoda je skloubením dvou předchozích metod vícerozměrné optimalizace, tedy metody nejrychlejšího spádu a metody Hooka a Jeevese.

Předpokládejme stejnou účelovou funkci typu (3.5). Necht' máme  $\mathbf{x}_0 \in \mathbb{R}^n$  jako výchozí bod. Následující dva body  $\mathbf{x}_1$  a  $\mathbf{x}_2$  získáme pomocí metody nejrychlejšího spádu. Nyní nastává část podobná metodě Hooka a Jeevese. Ne zvolíme za další směr postupu  $\mathbf{s}_2 = -f'(\mathbf{x}_2)$ , ale použijeme  $\mathbf{s}_2 = \mathbf{x}_2 - \mathbf{x}_0$ . Následující bod minimalizující posloupnosti  $\mathbf{x}_3$  najdeme opět standardní jednorozměrnou minimalizací. Jelikož jsme v daném směru našli neoptimálnější bod, neopakujeme hledání v tomto směru jako u metody Hooka a Jeevese, nýbrž opět najdeme dva následující body minimalizující posloupnosti  $\mathbf{x}_4$  a  $\mathbf{x}_5$  pomocí metody nejrychlejšího spádu. Bod  $\mathbf{x}_6$  získáme jednorozměrnou optimalizací ve směru  $\mathbf{s}_5 = \mathbf{x}_5 - \mathbf{x}_3$ . Celý postup opakujeme do splnění podmínky ukončení iterování.

**Příklad 10.** Získejte bod  $\mathbf{x}_3$  minimalizující posloupnosti metodou paralelních tečen pro účelovou funkci z minulého příkladu, tedy funkci typu (3.5), pro kterou platí

$$\mathbf{A} = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{27}{200} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}.$$

Výchozí bod je  $x_0 = (0; 3, 9)^T$ .

**Řešení.** Body minimalizující posloupnosti  $\mathbf{x}_1$  a  $\mathbf{x}_2$  jsme v příkladu dokazujícím pomalou konvergenci metody nejrychlejšího spádu. Nyní budeme potřebovat jen bod  $\mathbf{x}_2$

$$\mathbf{x}_2 = \begin{pmatrix} 0,0177 \\ 3,8308 \end{pmatrix}.$$

Dle metody paralelních tečen získáme nový směr pro jednorozměrnou optimalizaci jako směr z bodu  $\mathbf{x}_0$  do bodu  $\mathbf{x}_2$

$$\mathbf{s}_2 = \mathbf{x}_2 - \mathbf{x}_0 = \begin{pmatrix} 0,0177 \\ 3,8308 \end{pmatrix} - \begin{pmatrix} 0 \\ 3,9 \end{pmatrix} = \begin{pmatrix} 0,0177 \\ -0,0692 \end{pmatrix}.$$

V této chvíli se můžeme rozhodnout, zda budeme považovat za výchozí bod této jednorozměrné optimalizace bod  $\mathbf{x}_0$  nebo bod  $\mathbf{x}_2$ . Je zřejmé že výsledek by se lišil o 1, tedy

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.2. GRADIENTNÍ METODY

---

přesně o velikost vektoru  $\mathbf{s}_2$ . Jelikož volba bodu  $\mathbf{x}_0$  bude výpočetně jednodušší, provedeme tuto volbu.

Musíme tedy optimalizovat účelovou funkci  $f(\mathbf{x}_0 + \alpha \mathbf{s}_2)$  podle  $\alpha$ . Ze zadání víme, že funkce  $f(x)$ , kde provedeme označení  $\mathbf{x} = (u, v)^T$ , je tvaru

$$f(x) = u^2 + \frac{1}{2}uv + \frac{27}{400}v^2 - 2u - \frac{1}{2}v$$

Dosadíme  $u = 0,0177\alpha$ ,  $v = 3,9 - 0,0692\alpha$

$$\begin{aligned} f(x) &= (0,0177\alpha)^2 + \frac{1}{2}(0,0177\alpha)(3,9 - 0,0692\alpha) + \frac{27}{400}(3,9 - 0,0692\alpha)^2 - 2(0,0177\alpha) - \\ &\quad - \frac{1}{2}(3,9 - 0,0692\alpha) = 0,000024\alpha^2 - 0,002719\alpha - 0,923325. \end{aligned}$$

Vyjádření účelové funkce v této formě zderivujeme podle  $\alpha$  a položíme rovno nule, abychom dostali bod minima ve směru  $\mathbf{s}_2$

$$\frac{\partial f(x)}{\partial \alpha} = 0,000048\alpha_2 - 0,002719 = 0,$$

$$\alpha_2 = 56,399.$$

Následující bod minimalizující posloupnosti tedy bude

$$\mathbf{x}_3 = \mathbf{x}_0 + \alpha_2 \mathbf{s}_2 = \begin{pmatrix} 0 + 0,0177\alpha_2 \\ 3,9 - 0,0692\alpha_2 \end{pmatrix} = \begin{pmatrix} 0,9982 \\ -0,0028 \end{pmatrix}$$

Pro porovnání dodáme bod  $\mathbf{x}_3$  získaný neupravenou metodou nejrychlejšího spádu  $\mathbf{x}_3 = (0,0595; 3,8087)^T$ . Je lehce vidět, že bod  $\mathbf{x}_3$  získaný metodou paralelních tečen je mnohem blíže skutečnému bodu minima  $\mathbf{x}^* = (1; 0)^T$ .

**Věta 3.5.** *Nechť  $\mathbf{s}_{3k-1} = \mathbf{x}_{3k-1} - \mathbf{x}_{3(k-1)}$ , pro  $k \in \mathbb{N}$ , představuje směr jednorozměrné optimalizace v  $k$ -tém kroku metody paralelních tečen. Potom bod minimalizující posloupnosti  $\mathbf{x}_{3k}$  lze vypočítat jako*

$$\mathbf{x}_{3k} = \mathbf{x}_{3k-1} + \left( \frac{\mathbf{s}_{3k-1}^T (\mathbf{b} - \mathbf{A}\mathbf{x}_{3k-1})}{\mathbf{s}_{3k-1}^T \mathbf{A} \mathbf{s}_{3k-1}} \right) \mathbf{s}_{3k-1}. \quad (3.13)$$

*Důkaz.* Důkaz je velice podobný důkazu vzorce (3.6) pro metodu nejrychlejšího spádu, proto zde jen ukážeme rozdíl. Již víme, že

$$f(\mathbf{x}_{3k-1} + \alpha \mathbf{s}_{3k-1}) - f(\mathbf{x}_{3k-1}) = \frac{\alpha^2}{2} \mathbf{s}_{3k-1}^T \mathbf{A} \mathbf{s}_{3k-1} + \alpha \mathbf{s}_{3k-1}^T (\mathbf{A}\mathbf{x}_{3k-1} - \mathbf{b}).$$

Zde se již však  $\mathbf{s}_{3k-1}$  nerovná  $\mathbf{b} - \mathbf{A}\mathbf{x}_{3k-1}$ , tudíž rovnice nadále zůstane v tomto tvaru. Převedením  $f(\mathbf{x}_{3k-1})$  na pravou stranu, derivací a položením rovno nule získáme minimalizaci funkční hodnoty nově nalezeného bodu

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

---

$$\frac{\partial f(\mathbf{x}_{3k-1} + \alpha \mathbf{s}_{3k-1})}{\partial \alpha} = \alpha \mathbf{s}_{3k-1}^T \mathbf{A} \mathbf{s}_{3k-1} + \mathbf{s}_{3k-1}^T (\mathbf{A} \mathbf{x}_{3k-1} - \mathbf{b}) = 0.$$

Úpravou získáme

$$\alpha = \frac{\mathbf{s}_{3k-1}^T (\mathbf{b} - \mathbf{A} \mathbf{x}_{3k-1})}{\mathbf{s}_{3k-1}^T \mathbf{A} \mathbf{s}_{3k-1}}.$$

□

Nyní se můžeme přesvědčit, zda pomocí vzorce vyjde  $\alpha$  v minulém příkladě stejně jak jsme předpokládali.

$$\alpha = \frac{(0, 0177; -0, 0692) \left( \left( \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix} - \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{27}{200} \end{pmatrix} \begin{pmatrix} 0, 0177 \\ 3, 8308 \end{pmatrix} \right)}{(0, 0177; -0, 0692) \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{27}{200} \end{pmatrix} \begin{pmatrix} 0, 0177 \\ -0, 0692 \end{pmatrix}} = \frac{0, 00267}{0, 000048} = 55, 399.$$

Pro případ, kdy jsme počítali pomocí vzorce, a tedy jsme jako počátek naší jednorozměrné minimalizace brali bod  $\mathbf{x}_2$ , nám vyšlo  $\alpha$  o 1 menší než pro případ předešlý, kdy jsme jako počátek zvolili  $\mathbf{x}_0$ . Náš původní výpočet je tedy korektní.

### 3.2.4 Metoda nejrychlejšího spádu s drobením kroku

V této sekci probereme řešení případu, kdy z nějakého důvodu není možné pro jednorozměrnou optimalizaci použít standardní minimalizační postup pomocí derivací. Jako první odvodíme obecné principy, které poté upřesníme na použití pro metodu nejrychlejšího spádu.

Nechť jsou dána  $d, \vartheta \in \mathbb{R}$ , pro která platí  $d > 0$ ,  $\vartheta \in (0, 1)$ . Jelikož nemůžeme provést optimalizaci pomocí derivací využijeme postup podobný metodám komparativním. Zvolíme  $\alpha = d$  a porovnáme funkční hodnoty účelové funkce v bodech  $f(\mathbf{x}_k + \alpha \mathbf{s}_k)$  a  $f(\mathbf{x}_k)$ . Pokud bude platit  $f(\mathbf{x}_k + \alpha \mathbf{s}_k) < f(\mathbf{x}_k)$ , tak naše  $\alpha$  vyhovuje a určíme ho jako délku  $k$ -tého kroku, tedy  $\alpha_k = d$ .

V případě, že tato podmínka splněna není, tak předefinujeme  $\alpha = \vartheta \alpha$  a celý postup opakujeme. Z tohoto důvodu se koeficient  $\vartheta$  nazývá koeficient drobení kroku a je často volen  $\vartheta = \frac{1}{2}$ . Jelikož je zvolený směr pro tuto formu hledání vybrán jako při metodě nejrychlejšího spádu, musí nastat splnění dané podmínky v konečně krocích.

Pokud si pro testování zvolíme silnější podmínku

$$f(\mathbf{x}_k + \alpha \mathbf{s}_k) - f(\mathbf{x}_k) < \varepsilon \alpha \mathbf{s}_k^T f'(\mathbf{x}_k), \quad (3.14)$$



## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.2. GRADIENTNÍ METODY

---

kde  $\varepsilon \in (0, 1)$ , jedná se modifikovanou metodu drobení kroku. Z důvodu lepšího pochopení této podmínky ji přepíšeme do tvaru

$$f(\mathbf{x}_k) - f(\mathbf{x}_k + \alpha \mathbf{s}_k) > \varepsilon \alpha \mathbf{s}_k^T (-f'(\mathbf{x}_k)).$$

Nyní se podmínka dá lépe interpretovat. Naznačuje, že čím bližší je směr naší jednorozměrné optimalizace  $\alpha \mathbf{s}_k$  ke směru v vektoru  $-f'(\mathbf{x}_k)$  tím větší očekáváme rozdíl mezi funkčními hodnotami nového a předchozího bodu minimalizující posloupnosti. Koeficient  $\varepsilon$  určuje tedy míru ambiciozности našich očekávání.

**Věta 3.6.** *Nechť je účelová funkce  $f$  diferencovatelná a pro všechna  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $M > 0$  splňuje Lipschitzovu podmínku  $|f(x) - f(y)| \leq M\|x - y\|$ . Je-li  $\varepsilon \in (0, 1)$  a  $\mathbf{s}_k^T f'(\mathbf{x}_k) < 0$ , potom je podmínka (3.14) splněna pro  $\alpha = (0, \bar{\alpha})$ , kde*

$$\bar{\alpha} = -\frac{(1 - \varepsilon) \mathbf{s}_k^T f'(\mathbf{x}_k)}{M \|\mathbf{s}_k\|^2}. \quad (3.15)$$

*Důkaz.* Vyjdeme z Lagrangeovy věty o střední hodnotě, která pro  $\alpha \in (0, \bar{\alpha})$  a  $\kappa \in (0, 1)$  říká

$$f'(\mathbf{x}_k + \alpha \kappa \mathbf{s}_k) = \frac{f(\mathbf{x}_k + \alpha \mathbf{s}_k) - f(\mathbf{x}_k)}{\alpha \mathbf{s}_k}$$

Úpravou získáme

$$\begin{aligned} f(\mathbf{x}_k + \alpha \mathbf{s}_k) - f(\mathbf{x}_k) &= \alpha \mathbf{s}_k^T f'(\mathbf{x}_k + \alpha \kappa \mathbf{s}_k) = \\ &= \alpha \mathbf{s}_k^T f'(\mathbf{x}_k) + \alpha \mathbf{s}_k^T (f'(\mathbf{x}_k + \alpha \kappa \mathbf{s}_k) - f'(\mathbf{x}_k)). \end{aligned}$$

Tento výraz je podle Cauchy-Schwarzovy nerovnosti menší nebo roven výrazu

$$\alpha \mathbf{s}_k^T f'(\mathbf{x}_k) + \|\alpha \mathbf{s}_k\| \|f'(\mathbf{x}_k + \alpha \kappa \mathbf{s}_k) - f'(\mathbf{x}_k)\|.$$

Jelikož funkce splňuje Lipschitzovu podmínku je tento výraz menší nebo roven výrazu

$$\begin{aligned} &\alpha (\mathbf{s}_k^T f'(\mathbf{x}_k) + M \|\alpha \mathbf{s}_k\| \|\alpha \kappa \mathbf{s}_k\|) \leq \\ &\leq \alpha (\mathbf{s}_k^T f'(\mathbf{x}_k) + M \alpha \|\mathbf{s}_k\|^2). \end{aligned}$$

Podle vzorce (3.15) je poslední výraz roven

$$\begin{aligned} &\alpha (\mathbf{s}_k^T f'(\mathbf{x}_k) - (1 - \varepsilon) \mathbf{s}_k^T f'(\mathbf{x}_k)) = \\ &= \alpha \varepsilon \mathbf{s}_k^T f'(\mathbf{x}_k), \end{aligned}$$

což dokazuje platnost podmínky (3.14). □

Nyní tyto obecné podmínky převedeme do konkrétního případu metody nejrychlejšího spádu. Naším hlavním rozdílem oproti obecným vzorcům je fakt, že již máme zvolený směr jednorozměrné minimalizace jako  $\mathbf{s} = -f'(\mathbf{x})$ , tedy podmínku (3.14) převedeme na

$$f(\mathbf{x}_k + \alpha \mathbf{s}_k) - f(\mathbf{x}_k) < -\varepsilon \alpha \|f'(\mathbf{x}_k)\|^2. \quad (3.16)$$

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.2. GRADIENTNÍ METODY

---

Tato podmínka je podle Věty 3.6 splněna pro  $\alpha = (0, \bar{\alpha})$ , kde

$$\bar{\alpha} = \frac{(1 - \varepsilon)}{M}. \quad (3.17)$$

Z teoretické konstrukce kroků metodou nejrychlejšího spádu s modifikovaným drobením kroku lze odvodit, že

$$\bar{\alpha} \geq \alpha_k \geq \hat{\alpha} = \min \left\{ d, \frac{\vartheta(1 - \varepsilon)}{M} \right\}. \quad (3.18)$$

Tato skutečnost vychází z toho, že pokud je podmínka (3.16) splněna hned pro  $\alpha = d$ , tak platí  $\alpha_k = d$ . Pokud toto neplatí tak pro první  $m$  takové, že  $\alpha_k = d\vartheta^m$ , pro které je podmínka splněna, musí platit, že tato podmínka není splněna pro  $\alpha = d\vartheta^{m-1}$ . To podle (3.17) znamená

$$\alpha = d\vartheta^{m-1} > \bar{\alpha} = \frac{(1 - \varepsilon)}{M}.$$

Toto pro  $\alpha_k$  implikuje

$$\alpha_k = d\vartheta^m = \vartheta d\vartheta^{m-1} > \vartheta \bar{\alpha} = \frac{\vartheta(1 - \varepsilon)}{M}.$$

**Věta 3.7.** *Předpokládejme, že účelová funkce  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  je spojitě diferencovatelná, zdola ohraničená a její derivace splňuje Lipschitzovu podmínku, tedy  $\|f'(\mathbf{x}) - f'(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\|$  pro všechna  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  a  $M > 0$ . Potom pro libovolnou výchozí hodnotu  $\mathbf{x}_0$  při splnění podmínky dané nerovnicí (3.16) platí*

$$\lim_{k \rightarrow \infty} \|f'(\mathbf{x}_k)\|^2 = 0.$$

*Důkaz.* Z vyžadované ohraničenosti zdola a z (3.16) vyplývá, že

$$\lim_{k \rightarrow \infty} (f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})) = 0.$$

Potom úpravou (3.16) dostaneme

$$\|f'(\mathbf{x}_k)\|^2 \leq \frac{f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})}{\varepsilon \alpha_k} \leq \frac{f(\mathbf{x}_k) - f(\mathbf{x}_{k+1})}{\varepsilon \hat{\alpha}}.$$

Pro poslední výraz víme, že jeho limita pro  $k \rightarrow \infty$  se rovná nule, tedy dle pravidel pro limity posloupností platí

$$\lim_{k \rightarrow \infty} \|f'(\mathbf{x}_k)\|^2 = 0.$$

□

Všimněme si, že tato věta nevyžadovala pro svou platnost konvexnost účelové funkce. Pro takovou funkci tedy touto metodou najdeme pouze stacionární bod. Pokud budeme uvažovat funkci konvexní je stacionárním bodem vždy jen jedině globální minimum.

### 3.3 Metody sdružených směrů

Jak jsme si již ukázali, tak metoda nejrychlejšího spádu není velice efektivní. Toto je zejména z důvodu, že nevyužívá plně informace získané během výpočetního procesu. Každý krok je informačně oddělen od ostatních. Nevyužívá se nijak směrů získaných v minulých krocích ani se nijak neplánují kroky další. Metoda sdružených směrů toto plýtvání informacemi odstraňuje.

O odstranění této informační nevyžitosti se snažila již metoda paralelních tečen, avšak v tomto případě je princip do značné míry odlišný. Když metoda paralelních tečen využívá zpětnou informaci, tedy body, které jsme již vypočítali, tak metoda sdružených gradientů se dívá do budoucna a plánuje jak budou následující směry postupu vypadat.

#### 3.3.1 Metoda sdružených směrů pro kvadratické funkce

Opět jako u metody nejrychlejšího spádu si nejdříve předvedeme jak metoda funguje, pokud je účelová funkce ve tvaru

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{b}. \quad (3.19)$$

Pro danou metodu bude opět platit

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k, \quad (3.20)$$

kde

$$f(\mathbf{x}_k + \alpha_k \mathbf{s}_k) = \min_{\alpha} f(\mathbf{x}_k + \alpha \mathbf{s}_k), \quad (3.21)$$

pro  $k = 0, 1, \dots, n - 1$ .

Nechť máme bázi  $n$ -rozměrného eukleidovského prostoru  $\mathbf{s}_1, \dots, \mathbf{s}_n$ . Dále nechť máme zadaný výchozí bod  $\mathbf{x}_0$  a pro vysvětlení principu známe i přesné řešení našeho minimalizačního problému, které označíme  $\mathbf{x}^*$ . Potom můžeme směr z výchozího bodu do skutečného minima zapsat jako lineární kombinaci vektorů báze, tedy

$$\mathbf{x}^* - \mathbf{x}_0 = \sum_{k=0}^{n-1} \alpha_k \mathbf{s}_k. \quad (3.22)$$

Pokud budou  $\mathbf{s}_k$  ortogonální tak můžeme vyjádřit koeficienty určující délku jednotlivých vektorů

$$\alpha_k = \frac{\mathbf{s}_k^T (\mathbf{x}^* - \mathbf{x}_0)}{\mathbf{s}_k^T \mathbf{s}_k}, \quad k = 0, 1, \dots, n - 1. \quad (3.23)$$

Tento vzorec by nám k ničemu nepomohl, pokud by zůstal v tomto tvaru, jelikož obsahuje během výpočtu neznámé  $\mathbf{x}^*$ . Proto si zadefinujeme nový pojem, který nám pomůže. Tento nejdůležitější termín, který budeme na následujících stranách využívat, nám napovídá název metody. Jedná se o sdruženost. V tomto případě půjde o  $\mathbf{A}$ -sdruženost směrových vektorů.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.3. METODY SDRUŽENÝCH SMĚRŮ

---

**Definice 2.** Nechť  $\mathbf{A}$  je pozitivně definitní symetrická matice. Potom nenulové vektory  $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^n$  nazveme  $\mathbf{A}$ -sdrúžené, pokud platí  $\mathbf{s}_1^T \mathbf{A} \mathbf{s}_2 = 0$ . Systém nenulových vektorů  $\mathbf{s}_0, \dots, \mathbf{s}_{n-1} \in \mathbb{R}^n$  se nazývá  $\mathbf{A}$ -sdrúžený, pokud platí

$$\mathbf{s}_i^T \mathbf{A} \mathbf{s}_j = 0, \quad \text{pro } i \neq j.$$

Místo termínu  $\mathbf{A}$ -sdrúženost se některá literatura zmiňuje o  $\mathbf{A}$ -ortogonalitě. Tento termín vychází z toho, že vektory  $\mathbf{s}_k$  jsou během výpočtu počítány pomocí Gram-Schmidtova ortogonalizačního procesu. Jelikož už pro tuto metodu neplatí, že se pohybujeme ve směru antigradientu provedeme si označení  $\mathbf{r}_k = -f'(\mathbf{x}_k)$ .

Pokud budeme pro vektory  $\mathbf{s}_i$  předpokládat  $\mathbf{A}$ -sdrúženost místo ortogonality, tak se náš vzorec pro  $\alpha_i$  změní na

$$\alpha_k = \frac{\mathbf{s}_k^T \mathbf{A} (\mathbf{x}^* - \mathbf{x}_0)}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = \frac{\mathbf{s}_k^T (\mathbf{b} - \mathbf{A} \mathbf{x}_0)}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = \frac{-\mathbf{s}_k^T f'(\mathbf{x}_0)}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = \frac{\mathbf{s}_k^T \mathbf{r}_0}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}. \quad (3.24)$$

Jakmile tedy budeme mít posloupnost  $\mathbf{A}$ -sdrúžených vektorů, tak se bez problému dostaneme z výchozího bodu do bodu minima.

Všimněme si, že pro pozitivně definitní matici  $\mathbf{A}$  a systém  $\mathbf{A}$ -sdrúžených vektorů  $\mathbf{s}_1, \dots, \mathbf{s}_{n-1} \in \mathbb{R}^n$  platí, že vektory  $\mathbf{s}_i$  a  $\mathbf{s}_j$  jsou navzájem nezávislé pro  $i \neq j$ . Toto je lehce vidět z pozitivní definitnosti matice  $\mathbf{A}$ , protože platí, že  $\mathbf{s}_i^T \mathbf{A} \mathbf{s}_i \neq 0$ . Pokud by pro nějaké  $i \in \{0, \dots, n-1\}$  platilo  $\mathbf{s}_i = \sum_{j \neq i} \lambda_j \mathbf{s}_j$  tak by to znamenalo, že  $\mathbf{s}_i^T \mathbf{A} \mathbf{s}_j = \sum_{j \neq i} \lambda_j \mathbf{s}_j^T \mathbf{A} \mathbf{s}_i = 0$ , což je ve sporu s pozitivní definitností matice  $\mathbf{A}$ .

Pro sestrojení posloupnosti  $\mathbf{A}$ -sdrúžených vektorů budeme potřebovat posloupnost vektorů  $\mathbf{u}_0, \dots, \mathbf{u}_{n-1}$ , které tvoří bázi. Takovou posloupnost již však máme ve formě  $\mathbf{u}_k = \mathbf{r}_k$ , pro  $k \in \{0, 1, \dots, n-1\}$ . Využijeme tedy Gram-Schmidtovy ortogonalizace k výpočtu  $\mathbf{A}$ -sdrúžených vektorů

$$\begin{aligned} \mathbf{s}_0 &= \mathbf{r}_0, \\ \mathbf{s}_{k+1} &= \mathbf{r}_{k+1} + \sum_{l=1}^k \beta_{k+1,l} \mathbf{s}_l, \quad k = 0, 1, \dots, n-1, \end{aligned} \quad (3.25)$$

kde

$$\beta_{k+1,l} = \frac{-(\mathbf{s}_l^T \mathbf{A} \mathbf{r}_{k+1})}{\mathbf{s}_l^T \mathbf{A} \mathbf{s}_l}.$$

Jelikož jsme za vektory  $\mathbf{u}_k$  zvolili body  $\mathbf{r}_k$  pro výchozí bod víme

$$\mathbf{r}_0 = -f'(\mathbf{x}_0) = \mathbf{b} - \mathbf{A} \mathbf{x}_0,$$

z toho vyplývá, že

$$\mathbf{s}_0 = \mathbf{r}_0.$$

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Dle vzorce (3.24) můžeme poté vypočítat  $\alpha_0$  a využít ho pro výpočet nového bodu minimalizující posloupnosti

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{s}_0.$$

Další vektor posloupnosti  $\mathbf{A}$ -sdružených vektorů vypočítáme

$$\mathbf{s}_1 = \mathbf{r}_1 + \beta_{1,0} \mathbf{s}_0, \quad (3.26)$$

kde

$$\beta_{1,0} = \frac{-(\mathbf{s}_0^T \mathbf{A} \mathbf{r}_1)}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0}.$$

Z tohoto postupu jdou odvodit obecné rovnice pro metodu sdružených směrů, pro které platí  $\mathbf{s}_1 = \mathbf{r}_1 = \mathbf{b} - \mathbf{A} \mathbf{x}_1$

$$\alpha_k = \frac{\mathbf{s}_k^T \mathbf{r}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}, \quad (3.27)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{s}_k, \quad (3.28)$$

$$\mathbf{r}_{k+1} = \mathbf{b} - \mathbf{A} \mathbf{x}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{s}_k, \quad (3.29)$$

$$\beta_k = \frac{-(\mathbf{s}_k^T \mathbf{A} \mathbf{r}_{k+1})}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}, \quad (3.30)$$

$$\mathbf{s}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{s}_k, \quad (3.31)$$

pro  $k = 0, 1, \dots, n-1$ .

V těchto vztazích se nachází dva rozdíly od vztahů vypočtených dříve. Tyto rozdíly dokážeme následujícími větami.

**Věta 3.8.** Výraz (3.24) je ekvivalentní výrazu (3.27), tedy

$$\frac{\mathbf{s}_k \mathbf{r}_0}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = \frac{\mathbf{s}_k^T \mathbf{r}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}$$

*Důkaz.* Potřebujeme dokázat, že  $\mathbf{s}_k \mathbf{r}_0 = \mathbf{s}_k \mathbf{r}_k$ .

Pro  $k = 0$  platí věta triviálně. Pro libovolné  $k > 0$  víme, že

$$\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha \mathbf{A} \mathbf{s}_{k-1} = \mathbf{r}_{k-2} - \alpha \mathbf{A} \mathbf{s}_{k-1} - \alpha \mathbf{A} \mathbf{s}_{k-2} = \mathbf{r}_0 - \alpha \mathbf{A} \mathbf{s}_{k-1} - \alpha \mathbf{A} \mathbf{s}_{k-2} \cdots - \alpha \mathbf{A} \mathbf{s}_0.$$

Jelikož jsou vektory  $\mathbf{s}_k$   $\mathbf{A}$ -sdružené, tak z toho vyplývá

$$\mathbf{s}_k \mathbf{r}_k = \mathbf{s}_k (\mathbf{r}_0 - \alpha \mathbf{A} \mathbf{s}_{k-1} - \alpha \mathbf{A} \mathbf{s}_{k-2} \cdots - \alpha \mathbf{A} \mathbf{s}_0) = \mathbf{s}_k \mathbf{r}_0.$$

□

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

K důkazu rozdílného vzorce pro  $\mathbf{s}_{k+1}$  musíme nejdříve dokázat ortogonalitu vektorů  $\mathbf{r}_k \in \mathbb{R}^n$ . Toto dokážeme zároveň v větě, která nás ujistí v našem původním předpokladu, že vektory  $\mathbf{r}_k \in \mathbb{R}^n$  tvoří bázi prostoru.

**Věta 3.9.** *Nechť bod  $\mathbf{x}_0 \in \mathbb{R}^n$  je zvolen libovolně a následující body minimalizující posloupnosti  $\mathbf{x}_1, \dots, \mathbf{x}_{n-1}$  a vektory  $\mathbf{s}_0, \dots, \mathbf{s}_n$  jsou dány vztahy (3.20), (3.28), (3.31). Potom vektory  $\mathbf{r}_0, \dots, \mathbf{r}_{n-1}$  tvoří bázi  $\mathbb{R}^n$ .*

*Důkaz.* Vektory  $\mathbf{r}_0, \dots, \mathbf{r}_{n-1}$  budou tvořit bázi prostoru  $\mathbb{R}^n$ , právě tehdy když budou ortogonální. Jelikož  $\mathbf{r}_k = -f'(\mathbf{x}_k)$ , tak pokud dokážeme ortogonalitu vektorů  $f'(\mathbf{x}_k)$ , tak tím dokážeme i ortogonalitu vektorů  $\mathbf{r}_k$ ,  $k = 0, 1, \dots, n-1$ .

Každý následující bod minimalizující posloupnosti  $\{\mathbf{x}_k\}$  je hledán dle (3.20) jednorozměrnou minimalizací pomocí vzorce

$$\alpha_k = \frac{\mathbf{s}_k^T \mathbf{r}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = -\frac{\mathbf{s}_k^T f'(\mathbf{x}_k)}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k}.$$

Z tohoto důvodu platí  $\mathbf{s}_k^T f'(\mathbf{x}_{k+1}) = 0^3$ . Rovnici (3.20) upravíme vynásobením  $\mathbf{A}$  a odečtením  $\mathbf{b}$  od obou stran rovnice. Tímto získáme tvar

$$f'(\mathbf{x}_{k+1}) = f'(\mathbf{x}_k) + \alpha_k \mathbf{A} \mathbf{s}_k,$$

který využijeme pro výpočet

$$\begin{aligned} f'(\mathbf{x}_{k+1})^T f'(\mathbf{x}_k) &= (f'(\mathbf{x}_k) + \alpha_k \mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_k) = \\ &= f'(\mathbf{x}_k)^T f'(\mathbf{x}_k) - \frac{\mathbf{s}_k^T f'(\mathbf{x}_k)}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} (\mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_k) = \\ &= \|f'(\mathbf{x}_k)\|^2 - \frac{f'(\mathbf{x}_k)^T (-f'(\mathbf{x}_k) + \beta_{k-1} \mathbf{s}_{k-1})}{(\mathbf{A} \mathbf{s}_k)^T (-f'(\mathbf{x}_k) + \beta_{k-1} \mathbf{s}_{k-1})} (\mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_k) = \\ &= \|f'(\mathbf{x}_k)\|^2 - \|f'(\mathbf{x}_k)\|^2 \frac{(\mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_k)}{(\mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_k)} = 0. \end{aligned}$$

□

Nyní máme dokázanou ortogonalitu vektorů  $\mathbf{r}_k$ , můžeme tedy dokázat rozdílnost vzorců (3.25) a (3.31). Rozdílnost těchto vzorců tkví v tom, že rovnice (3.31) obsahuje jen jedno  $\beta$ . Nyní dokážeme, že veškerá ostatní  $\beta$  jsou nulová.

**Věta 3.10.** *Veškerá  $\beta_{k+1,l}$ , pro která platí  $k \neq l$  jsou nulová.*

---

<sup>3</sup>Je prakticky dokázáno v metodě nejrychlejšího spádu. Následující směr z bodu, který je nalezen pomocí jednorozměrné optimalizace je kolmý k směru, kterým jsme se do daného bodu dostali.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.3. METODY SDRUŽENÝCH SMĚRŮ

---

*Důkaz.* Vyjdeme z rovnice (3.29), kterou upravíme do tvaru

$$\mathbf{r}_{k+1}^T \mathbf{A} \mathbf{s}_l = \frac{1}{\alpha_l} \mathbf{r}_{k+1}^T (\mathbf{r}_l - \mathbf{r}_{l+1}),$$

dosazením za  $\alpha_l$  a úpravou získáme

$$\beta_{k+1,l} = -\frac{\mathbf{r}_{k+1} (\mathbf{r}_l - \mathbf{r}_{l+1})}{\mathbf{s}_l^T \mathbf{r}_l}.$$

Nyní jelikož  $l \in \{1, \dots, k\}$ , tak je důkaz prakticky dokončen. Pro jakékoliv  $l \in \{1, \dots, k-1\}$  je výraz  $\mathbf{r}_{k+1} (\mathbf{r}_l - \mathbf{r}_{l+1})$  roven nule. Jen pro  $l = k$  se bude výraz rovnat

$$\beta_{k+1,k} = -\frac{\mathbf{r}_{k+1}^T (\mathbf{r}_k - \mathbf{r}_{k+1})}{\mathbf{s}_k^T \mathbf{r}_k} = -\frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{s}_k^T \mathbf{r}_k},$$

což se po úpravě rovná vzorci (3.30). □

**Věta 3.11.** *Nechť  $m \in \{1, \dots, n\}$ , vektory  $\mathbf{s}_0, \dots, \mathbf{s}_{m-1}$  jsou  $\mathbf{A}$ -sdružené,  $\mathbf{x}_0 \in \mathbb{R}^n$  je libovolné a body  $\mathbf{x}_1, \dots, \mathbf{x}_m$  jsou dány vztahy (3.20) a (3.21). Potom pro kvadratickou funkci  $f$  typu (3.19) platí*

$$f(\mathbf{x}_m) = \min_{\mathbf{x} \in X_m} f(\mathbf{x}),$$

kde  $X_m = \mathbf{x}_0 + \text{Lin}\{\mathbf{s}_0, \dots, \mathbf{s}_{m-1}\}$ . Speciálně, je-li  $m = n$ , potom

$$f(\mathbf{x}_m) = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}),$$

Z důvodu délky důkazu je důkaz vynechán. Je ho možné najít v [3], str. 109.

Nyní je vidět, že

$$\mathbf{x}_n = \mathbf{x}^*.$$

Metoda sdružených gradientů je tedy pro kvadratickou funkci metodou konečně krokovou (přesněji řečeno  $n$  krokovou). Při normálním výpočtu však v důsledku zaokrouhlovacích chyb nedostaneme přesný výsledek. Výpočet je možné provádět i pro  $k > n$ , kde platí

$$\|\mathbf{e}_{k+1}\| < \|\mathbf{e}_k\|,$$

pro všechna  $k$ , kde  $\mathbf{e}_k = \mathbf{x}^* - \mathbf{x}_k$ .

**Příklad 11.** Pomocí metody sdružených směrů najděte minimum funkce tvaru (3.19), pro kterou platí

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Výchozím bodem je  $\mathbf{x}_0 = (0; 0; 0)^T$ .

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

**Řešení.** Zjistíme si derivace účelové funkce podle jednotlivých proměnných. Pro přehlednost označíme  $\mathbf{x} = (u_1, u_2, u_3)^T$ .

$$f'(\mathbf{x}) = \begin{pmatrix} u_1 + u_3 - 1 \\ 2u_2 + u_3 \\ u_1 + u_2 + 2u_3 \end{pmatrix}.$$

První směr jednorozměrné optimalizace najdeme jako v metodě nejrychlejšího spádu, tedy jako směr antigradientu.

$$\mathbf{s}_0 = -f'(\mathbf{x}_0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Ze vzorce získáme  $\alpha_0$  a zjistíme nový bod minimalizující posloupnosti

$$\alpha_0 = \frac{\mathbf{s}_0^T \mathbf{r}_0}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} = \frac{(1; 0; 0) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}{(1; 0; 0) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} = 1.$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{s}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Pro výpočet následujícího směru jednorozměrné minimalizace již potřebujeme zjistit  $\beta_0$

$$\beta_0 = \frac{-(\mathbf{s}_0^T \mathbf{A} \mathbf{r}_1)}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} = \frac{-(1; 0; 0) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}}{(1; 0; 0) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} = 1.$$

Nový směr jednorozměrné optimalizace bude

$$\mathbf{s}_1 = \mathbf{r}_1 + \beta_0 \mathbf{s}_0 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Pomocí  $\alpha_1$  vypočítáme nový bod minimalizující posloupnosti

$$\alpha_1 = \frac{\mathbf{s}_1^T \mathbf{r}_1}{\mathbf{s}_1^T \mathbf{A} \mathbf{s}_1} = \frac{(1; 0; -1) \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}}{(1; 0; -1) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} = 1$$



KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{s}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Poslední krok provedeme již zrychleně.

$$\begin{aligned} \beta_1 &= \frac{-(\mathbf{s}_1^T \mathbf{A} \mathbf{r}_2)}{\mathbf{s}_1^T \mathbf{A} \mathbf{s}_1} = 1, \\ \mathbf{s}_2 &= \mathbf{r}_2 + \beta_1 \mathbf{s}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \\ \alpha_2 &= \frac{\mathbf{s}_2^T \mathbf{r}_2}{\mathbf{s}_2^T \mathbf{A} \mathbf{s}_2} = 1, \\ \mathbf{x}_3 &= \mathbf{x}_2 + \alpha_2 \mathbf{s}_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}. \end{aligned}$$

Bod minima byl nalezen v  $n = 3$  krocích.

Doposud jsme neustále předpokládali, že matice  $\mathbf{A}$  je pozitivně definitní. Nyní se podíváme na případ, kdy bude tato matice jen pozitivně semidefinitní a zjistíme, zda bude metoda použitelná i pro tento modifikovaný případ.

**Věta 3.12.** *Nechť matice  $\mathbf{A}$  je pozitivně semidefinitní matice o rozměrech  $n \times n$  a hodností  $r \leq n$ . Pak při použití metody sdružených směrů najdeme minimum dané účelové funkce  $f$  v nejvýše  $r$  krocích (výpočet skončí po  $k$  krocích, kde  $k \in (1, \dots, r)$ , pak  $f'(\mathbf{x}_k) = 0$  a  $\mathbf{s}_0, \dots, \mathbf{s}_{k-1} \neq 0$ ). V opačném případě výpočet ukáže, že minimum neexistuje, tak, že  $\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k = 0$ , přičemž  $f'(\mathbf{x}_0), \dots, f'(\mathbf{x}_{k-1}) \neq 0$  pro nějaké  $k \in \{0, \dots, r\}$ .*

*Důkaz.* Nejdříve budeme předpokládat, že kvadratická účelová funkce typu (3.19) nabývá na  $\mathbb{R}^n$  svého minima. Dále budeme předpokládat, že máme vektor  $\mathbf{y} \in \mathbb{R}^n$ , pro který platí  $\mathbf{A} \mathbf{y} = 0$ . Z tohoto vyplývá, že bude platit i  $\mathbf{y}^T \mathbf{b} = 0$ . Pokud by toto neplatilo, tak by funkční hodnota  $f(\alpha \mathbf{y}) = -\alpha \mathbf{y}^T \mathbf{b} \rightarrow -\infty$  pro  $\alpha \rightarrow \pm\infty$  v závislosti na znaménku  $\mathbf{y}^T \mathbf{b}$ , což je spor s předpokladem, že funkce nabývá na  $\mathbb{R}^n$  svého minima. Potom pro libovolné  $\mathbf{x} \in \mathbb{R}^n$  platí

$$\mathbf{y}^T f'(\mathbf{x}) = \mathbf{y}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = \mathbf{y}^T \mathbf{A} \mathbf{x} - \mathbf{y}^T \mathbf{b} = 0. \quad (3.32)$$

Mějme prostor

$$\ker \mathbf{A} := \{\mathbf{y} \in \mathbb{R}^n : \mathbf{A} \mathbf{y} = 0\},$$

pokud označíme  $r$  hodnost matice  $\mathbf{A}$ , tak tento prostor má dimenzi  $n - k$ . Z rovnice (3.32) víme, že vektor parciálních derivací bude pro libovolné  $\mathbf{x}$  náležet v ortogonálním doplňku  $(\ker \mathbf{A})^\perp$ . Je zřejmé, že tento prostor bude mít dimenzi  $r$ . Podle (3.31) musí i vektory  $\mathbf{s}_k \in (\ker \mathbf{A})^\perp$ , tudíž  $\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k = 0$  právě tehdy když  $\mathbf{s}_k = 0$ .

Naší snahou nyní bude ukázat, že metoda sdružených gradientů najde bod minima

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

nejvýše v  $r$  krocích, tedy bude platit  $f'(\mathbf{x}_i) = 0$ , pro nějaké  $i \in \{1, \dots, r\}$ . Předpokládejme, že máme  $\mathbf{s}_i \neq 0$ , platí tedy  $\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k \neq 0$  pro  $j = 1, \dots, r-1$ . Již víme, že vektory parciálních derivací  $f'(\mathbf{x}), j = 0, \dots, r$  jsou navzájem ortogonální<sup>4</sup>. Tyto vektory však náležejí do  $r$ -dimenzionálního prostoru  $(\ker \mathbf{A})^\perp$ . Toto naznačuje, že alespoň jeden z nich musí být nulový a právě v tomto bodě se bude nacházet minimum účelové funkce.

V této části budeme naopak předpokládat, že účelová funkce není na  $\mathbb{R}^n$  zdola ohraničená. Tedy musí nastat rovnost  $\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k = 0$  pro některé  $k \in \{0, \dots, n-1\}$ . V tomto případě tedy není možné pokračovat ve výpočtu, jelikož nemůžeme vypočítat následující  $\beta$ .

Pokud by platilo  $\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k > 0$ , pak by byly vektory  $\mathbf{s}_0, \dots, \mathbf{s}_{n-1}$  navzájem  $\mathbf{A}$ -sdružené, z čehož, jak jsme si ukázali, plyne lineární nezávislost. Potom by libovolný nenulový  $\mathbf{x} \in \mathbb{R}^n$  šlo vyjádřit jako lineární kombinace vektorů  $\mathbf{s}$ , tedy  $\mathbf{x} = \sum_{k=1}^{n-1} \eta_k \mathbf{s}_k$ , kde alespoň jeden z koeficientů  $\eta_k$  je nenulový. Potom by platilo

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{k=1}^{n-1} \eta_k^2 \mathbf{s}_k^T \mathbf{A} \mathbf{s}_k > 0.$$

Toto by značilo, že matice  $\mathbf{A}$  je pozitivně definitní, a tedy rovnice  $f'(\mathbf{x}) = \mathbf{A} \mathbf{x} - \mathbf{b} = 0$  má řešení. Toto řešení by bylo minimum účelové funkce  $f$ , což je spor s předpokladem, že účelová funkce je zdola neohraničená na  $\mathbb{R}^n$ . □

Metoda má ještě jednu výhodu a tou je, že pokud je počáteční bod  $\mathbf{x}_0$  zvolen vhodně pro danou účelovou funkci, tak se do skutečného bodu minima funkce dostane v méně než  $n$  krocích. Tato vlastnost je možná demonstrovat například při minimalizaci účelové funkce tvaru (3.19), pro kterou platí

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix},$$

kde jako výchozí bod zvolíme  $\mathbf{x}_0 = (1; 4; 5; 1)^T$ .

Jelikož jsme výpočet touto metodou již ukazovali na jiném příkladě, tak zde nyní pouze znázorníme, jak metoda konverguje po méně jak  $n$  krocích, v tomto případě po krocích dvou. Výpočtem metodou sdružených směrů získáme

$$\mathbf{x}_1 = \begin{pmatrix} 1,75 \\ 3 \\ 1,75 \end{pmatrix},$$

$$\mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

---

<sup>4</sup>Můžeme ukázat podobně jako ve Větě 3.9.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Pro  $\mathbf{x} = (u_1; u_2; u_3)$  při dalším výpočtu narazíme na

$$\begin{aligned} f_{u_1}(\mathbf{x}_2) &= 2u_1 - u_2 = 0, \\ f_{u_2}(\mathbf{x}_2) &= 2u_2 - u_1 - u_3 - 2 = 0, \\ f_{u_3}(\mathbf{x}_2) &= 2u_3 - u_2 - 1 = 0, \end{aligned}$$

což naznačuje, že jsme našli bod minima účelové funkce. Postup je tedy ukončen po kroku  $n - 1$  místo kroku  $n$ .

### 3.3.2 Metoda sružených směrů pro nekvadratické funkce

Jelikož v následující podkapitole budeme uvažovat funkce v jiném tvaru než tvar (3.19), tak již nemůžeme využít v předchozí sekci definované  $\mathbf{A}$ -sruženosti. Metoda však jde modifikovat, aby tvořila posloupnost  $\{\mathbf{x}_k\}$ , která bude také minimalizující i pro nekvadratické funkce.

Ze stejného důvodu proč již nemůžeme využít  $\mathbf{A}$ -sruženosti tak si již nemůžeme ulehčit výpočet  $\alpha_k$  pomocí vzorce (3.27). Pro výpočet koeficientu udávající délku kroku je nyní nutné použít standardní jednorozměrnou minimalizaci. Po této úpravě se veškeré modifikace omezí pouze na úpravu vzorce pro  $\beta_k$ , který nemůžeme na rozdíl od vzorce pro  $\alpha_k$  vypustit a nahradit jiným postupem.

Vyjdeme z vzorce využívaného pro nalezení  $\mathbf{A}$ -sružených vektorů při metodě sružených směrů pro kvadratické funkce, tedy vzorce

$$\beta_k = \frac{-(\mathbf{s}_k^T \mathbf{A} \mathbf{r}_{k+1})}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} = \frac{f'(\mathbf{x}_{k+1})^T \mathbf{A} \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k},$$

Dále upravíme vzorec  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{s}_k$  do tvaru

$$\mathbf{A} \mathbf{s}_k = \frac{1}{\alpha_k} (\mathbf{r}_k - \mathbf{r}_{k+1}) = \frac{1}{\alpha_k} (f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k)).$$

Nyní když spojíme oba takto získané vztahy, tak dostaneme vzorec, který již nebude striktně vyžadovat kvadratický tvar účelové funkce. Tohoto docílíme odstraněním matice  $\mathbf{A}$  z tohoto vztahu

$$\beta_k = \frac{f'(\mathbf{x}_{k+1})^T (f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k))}{\mathbf{s}_k^T (f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k))}.$$

Jelikož víme, že  $\mathbf{s}_k$  jde vyjádřit pomocí  $\mathbf{r}_i$ , pro  $i = \{0, 1, \dots, k\}$  a z ortogonalit vektorů  $f'(\mathbf{x})$  se tento výraz bude rovnat

$$\beta_k = \frac{\|f'(\mathbf{x}_{k+1})\|}{\|f'(\mathbf{x}_k)\|}. \quad (3.33)$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Z tohoto postupu je zřejmé, že daný vzorec je pro kvadratické funkce ekvivalentní se vztahem (3.30). Avšak na rozdíl od tohoto vzorce je použitelný nejen pro kvadratickou účelovou funkci, ale má smysl i pro funkce nekvadratické.

Jak bylo řečeno v sekci zabývající se metodou sružených směrů pro kvadratické funkce, tak důvodem nedokonalé konvergence jsou zaokrouhlovací chyby při jednorozměrné optimalizaci. Tento problém se snaží odstranit modifikace, která upravuje výpočet  $\beta_{k-1}$

$$\beta_{k-1} = \begin{cases} 0 & \text{pro } k = n, 2n, 3n, \dots \\ \frac{\|f'(\mathbf{x}_k)\|}{\|f'(\mathbf{x}_{k-1})\|} & \text{pro } k \neq n, 2n, 3n, \dots \end{cases}$$

Tento vzorec znamená, že se vždy na konci  $l$ -tého cyklu o  $n$  krocích jako směr dalšího postupu vybere  $\mathbf{s}_{ln} = -f'(\mathbf{x}_{ln})$ . Tímto se proces ve své podstatě očistí o možné nahromaděné chyby.

**Příklad 12.** Pomocí metody sružených směrů pro nekvadratické funkce vypočítejte minimum  $\mathbf{x}^* = (u_1^*; u_2^*)$  účelové funkce

$$f(\mathbf{x}) = f(u_1, u_2) = 2u_1^4 - 2u_1^2u_2 + 4u_1^2 - 8u_1 + \frac{1}{2}u_2^2.$$

Výchozím bodem bude  $\mathbf{x}_0 = (0; 0)$ .

**Řešení.** První bod bude veden ve směru antigradientu, tudíž musíme vypočítat jednotlivé parciální derivace pro danou účelovou funkci

$$\begin{aligned} \frac{\partial f}{\partial u_1} &= 8u_1^3 - 4u_1u_2 + 8u_1 - 8, \\ \frac{\partial f}{\partial u_2} &= -2u_1^2 + u_2. \end{aligned}$$

Z těchto derivací nyní určíme náš první směr postupu pomocí dosazení bodu  $\mathbf{x}_0$

$$\mathbf{s}_0 = \mathbf{r}_0 = -f'(\mathbf{x}_0) = \begin{pmatrix} 8 \\ 0 \end{pmatrix}.$$

Jelikož nám v této úpravě metody chybí vzorec pro vypočítání koeficientu udávajícího délku kroku  $\alpha$ , musíme tento koeficient vypočítat standardní jednorozměrnou minimalizací z bodu  $\mathbf{x}_0$  ve směru  $\mathbf{s}_0$

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha\mathbf{s}_0 = \begin{pmatrix} 8\alpha \\ 0 \end{pmatrix}$$

Tento výraz dosadíme do účelové funkce a provedeme minimalizaci podle  $\alpha_0$

$$\begin{aligned} 2(8\alpha)^4 + 4(8\alpha)^2 - 8(8\alpha) &\longrightarrow \min, \\ 8192\alpha^4 + 256\alpha^2 - 64\alpha &\longrightarrow \min, \\ 32768\alpha^3 + 512\alpha - 64 &= 0. \end{aligned}$$

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Naším výpočtem tedy získáme  $\alpha_0 = 0,0853$ , tudíž nový bod minimalizující posloupnosti se rovná  $\mathbf{x}_1 = (0,6823; 0)$ .

Nyní pro další postup potřebujeme z nově získaného vzorce vypočítat  $\beta_0$

$$\beta_0 = \frac{\|f'(\mathbf{x}_1)\|}{\|f'(\mathbf{x}_0)\|} = \frac{\sqrt{(0; -0,931) \begin{pmatrix} 0 \\ -0,931 \end{pmatrix}}}{\sqrt{(-8; 0) \begin{pmatrix} -8 \\ 0 \end{pmatrix}}} = \frac{0,931}{8} = 0,116.$$

Nyní již můžeme zjistit nový směr jednorozměrné minimalizace

$$\mathbf{s}_1 = \mathbf{r}_1 + \beta_0 \mathbf{s}_0 = \begin{pmatrix} 0 \\ 0,931 \end{pmatrix} + 0,116 \begin{pmatrix} 8 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,931 \\ 0,931 \end{pmatrix}.$$

Pro výpočet délky kroku použijeme substituci  $\mathbf{x}_2 = \mathbf{x}_1 + \alpha \mathbf{s}_1$ , kterou vložíme do účelové funkce a tuto funkci minimalizujeme podle  $\alpha$

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha \mathbf{s}_1 = \begin{pmatrix} 0,6823 + 0,931\alpha \\ 0,931\alpha \end{pmatrix}.$$

Úpravou a optimalizací získáváme

$$\begin{aligned} & 2(0,6823 + 0,931\alpha)^4 - 2(0,931\alpha)(0,6823 + 0,931\alpha)^2 + 4(0,6823 + 0,931\alpha)^2 - \\ & - 8(0,6823 + 0,931\alpha) + \frac{1}{2}(0,931\alpha)^2 \longrightarrow \min, \\ & 1,5029\alpha^4 + 2,7913\alpha^3 + 6,3778\alpha^2 - 0,8674\alpha - 3,1628 \longrightarrow \min, \end{aligned}$$

$$\frac{\partial f}{\partial \alpha} = 6,0119\alpha^3 + 8,3741\alpha^2 + 12,7556\alpha - 0,8674 = 0$$

Tímto výpočtem získáme  $\alpha_1 = 0,065$ , tedy nový bod minimalizující posloupnosti se rovná

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha \mathbf{s}_1 = \begin{pmatrix} 0,6823 + 0,931 \cdot 0,065 \\ 0,931 \cdot 0,065 \end{pmatrix} = \begin{pmatrix} 0,7428 \\ 0,0606 \end{pmatrix}.$$

Opět získáme  $\beta$  a zjistíme další směr postupu

$$\beta_1 = \frac{\|f'(\mathbf{x}_2)\|}{\|f'(\mathbf{x}_1)\|} = \frac{\sqrt{(1,043; -1,043) \begin{pmatrix} 1,043 \\ -1,043 \end{pmatrix}}}{\sqrt{(0; -0,931) \begin{pmatrix} 0 \\ -0,931 \end{pmatrix}}} = \frac{1,4753}{0,931} = 1,5846.$$

$$\mathbf{s}_2 = \mathbf{r}_2 + \beta_1 \mathbf{s}_1 = \begin{pmatrix} -1,043 \\ 1,043 \end{pmatrix} + 1,5846 \begin{pmatrix} 0,931 \\ 0,931 \end{pmatrix} = \begin{pmatrix} 0,4321 \\ 2,5184 \end{pmatrix}.$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Nový bod minimalizující posloupnost získáme další jednorozměrnou minimalizací

$$\mathbf{x}_3 = \mathbf{x}_2 + \alpha \mathbf{s}_2 = \begin{pmatrix} 0,7428 + 0,4321\alpha \\ 0,0606 + 2,5184\alpha \end{pmatrix}.$$

Po dosazení a úpravě získáme minimalizaci

$$0,6972\alpha^4 - 0,4610\alpha^3 + 1,8983\alpha^2 - 2,1766\alpha - 3,1915 \longrightarrow \min,$$

kteřá má řešení  $\alpha = 0,5588$ . Nový bod minimalizující posloupnosti tedy bude

$$\mathbf{x}_3 = \mathbf{x}_2 + \alpha \mathbf{s}_2 = \begin{pmatrix} 0,7428 + 0,4321 \cdot 0,5588 \\ 0,0606 + 2,5184 \cdot 0,5588 \end{pmatrix} = \begin{pmatrix} 0,9843 \\ 1,4680 \end{pmatrix}.$$

Ve třech krocích jsme se dostali z bodu  $\mathbf{x}_0 = (;0)^T$  s funkční hodnotou  $f(\mathbf{x}_0) = 0$  do bodu  $\mathbf{x}_3 = (0,9843; 1,4680)^T$  s funkční hodnotou  $f(\mathbf{x}_3) = -3,8887$ . Skutečným bodem minima účelové funkce je přitom bod  $\mathbf{x}^* = (1; 2)$  s funkční hodnotou  $f(\mathbf{x}^*) = -4$ .

Jak můžeme z příkladu vidět, tak již metoda sdružených gradientů pro nekvadratické funkce nekonverguje nejvýše v  $n$  krocích, kde  $n$  je dimenze prostoru, jako tomu bylo u kvadratických funkcí.

#### 3.3.3 Metoda sdružených směrů nultého řádu

V této sekci se opět vrátíme k minimalizaci kvadratické účelové funkce, tedy funkce typu (3.19), kde  $\mathbf{A}$  je symetrická, pozitivně definitní matice. Následující metoda se ve svém výpočtu podobá metodě Hooka a Jeevese. Podobnost vyplývá z toho, že na rozdíl od modifikací metody sdružených směrů, které jsme probírali doteď, se při této variantě k získání  $\mathbf{A}$ -sdružených vektorů používá kroků o  $n$  jednorozměrných optimalizacích a již ne výpočet pomocí gradientu účelové funkce. Pokud by k vyřešení jednorozměrných minimalizací byla využita některá z negradientní metod, tak pro výpočet metodou sdružených směrů nultého řádu nepotřebujeme znát derivace účelové funkce, což je také důvod názvu této modifikace.

Jak jsme již naznačili, tak výpočet stejně jako u metody Hooka a Jeevese začíná tím, že se z výchozího bodu provede  $n$  jednorozměrných minimalizací. Standardně se za směry pro tyto první minimalizace berou vektory kanonické báze  $\mathbb{R}^n$ . Jakmile tímto způsobem nalezneme konečný bod, tak ho spojíme s bodem výchozím a v tomto směru provádíme opět jednorozměrnou minimalizaci za účelem nalezení dalšího bodu minimalizující posloupnosti. Na rozdíl od metody Hooka a Jeevese však neprovádíme jednorozměrné optimalizace v stále stejných směrech, avšak nahrazujeme původní směry těmi směry, které získáme spojením posledního bodu minimalizující posloupnosti a bodu nalezeného jednorozměrnými optimalizacemi. Toto nahrazování směrů jednorozměrných optimalizací je v této metodě stěžejní, jelikož jak si ukážeme později, tak tímto způsobem získané vektory budou  $\mathbf{A}$ -sdružené.

Mějme tedy výchozí bod  $\mathbf{x}_0 \in \mathbb{R}^n$ . Jako první směry jednorozměrných optimalizací můžeme zvolit libovolné navzájem lineárně nezávislé vektory  $\mathbf{s}_0, \dots, \mathbf{s}_{n-1}$ . Nyní se provede

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

postupně všech  $n$  jednorozměrných minimalizací. Výsledek  $i$ -té minimalizace v kroku  $k + 1$  označíme  $\mathbf{y}_{k,i}$ . Po provedení všech  $n$  minimalizací získáme tedy bod  $\mathbf{y}_{k,n}$ . Obecně pro krok  $k + 1$  zapíšeme

$$f(\mathbf{y}_{k,0}) = \min_{\alpha \in \mathbb{R}^n} f(\mathbf{x}_k + \alpha \mathbf{s}_0), \quad (3.34)$$

$$f(\mathbf{y}_{k,i+1}) = \min_{\alpha \in \mathbb{R}^n} f(\mathbf{y}_{k,i} + \alpha \mathbf{s}_i), \quad i = 1, \dots, n-1. \quad (3.35)$$

Bod  $\mathbf{y}_{k,n}$  však není bodem minimalizující posloupnosti, jen nám poslouží k nalezení tohoto bodu. Následující jednorozměrnou minimalizaci provedeme ve směru

$$\mathbf{z}_k = \mathbf{y}_{k,n} - \mathbf{x}_k. \quad (3.36)$$

Samotný nový bod minimalizující posloupnosti získáme jako

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{z}_k, \quad (3.37)$$

$$f(\mathbf{x}_k + \alpha_k \mathbf{z}_k) = \min_{\alpha \in \mathbb{R}} f(\mathbf{x}_k + \alpha \mathbf{z}_k). \quad (3.38)$$

Nyní jsme téměř dokončili jeden krok této metody. Zbývá jen získání nových směrů, ve kterých bude prováděno  $n$  jednorozměrných minimalizací v následujícím kroku. Pravidlo pro tuto výměnu je jednoduché. První z těchto vektorů směru vyřadíme a místo něj jako poslední ze všech směrů přidáme vektor, který jsme získali jako  $\mathbf{z}_k = \mathbf{y}_{k,n} - \mathbf{x}_k$ . Toto je možné znázornit pomocí změny množiny vektorů směru, používaných v daných krocích

$$\{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{n-2}, \mathbf{s}_{n-1}\} \longrightarrow \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{n-1}, \mathbf{z}_k\}.$$

Pro další postup si dokážeme jedno pomocné tvrzení, které později využijeme v důkazu **A**-sdrůženosti směrových vektorů.

**Věta 3.13.** *Nechť je dána kvadratická funkce typu (3.19), kde **A** je symetrická, pozitivně definitní matice a vektory  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1} \in \mathbb{R}^n$ ,  $a, b \in \mathbb{R}^n$ , takové, že  $a \neq b$ . Označme  $\mathfrak{L} = \text{Lin}\{\mathbf{z}_0, \dots, \mathbf{z}_{k-1}\}$  lineární obal směrových vektorů  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1}$ . Je-li*

$$f(\mathbf{u}) = \min_{\mathbf{x} \in a + \mathfrak{L}} f(\mathbf{x}), \quad f(\mathbf{v}) = \min_{\mathbf{x} \in b + \mathfrak{L}} f(\mathbf{x}),$$

potom

$$(\mathbf{u} - \mathbf{v})^T \mathbf{A} \mathbf{z}_i = 0, \quad i = 0, 1, \dots, l-1.$$

*Důkaz.* Z předpokladů víme, že  $f'(\mathbf{u}) = 0$ ,  $f'(\mathbf{v}) = 0$ , pro libovolné  $\mathbf{s}_i$  musí tedy platit i  $f'(\mathbf{u})^T \mathbf{z}_i = 0$ ,  $f'(\mathbf{v})^T \mathbf{z}_i = 0$ . Odečtením těchto dvou rovnic získáme  $f'(\mathbf{u})^T \mathbf{z}_i - f'(\mathbf{v})^T \mathbf{z}_i = 0$ . Nyní využijeme toho, že  $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$  a získáme

$$\begin{aligned} (\mathbf{A}\mathbf{u} - \mathbf{b})^T \mathbf{z}_i - (\mathbf{A}\mathbf{v} - \mathbf{b})^T \mathbf{z}_i &= 0, \\ (\mathbf{u} - \mathbf{v})^T \mathbf{A} \mathbf{z}_i &= 0, \end{aligned}$$

□

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
 3.3. METODY SDRUŽENÝCH SMĚRŮ

---

Nyní již máme vše připraveno na nejdůležitější větu této metody.

**Věta 3.14.** *Nechť máme kvadratickou účelovou funkci  $f$  tvaru (3.19), kde  $\mathbf{A}$  je symetrická, pozitivně definitní matice. Dále nechť vektory  $\mathbf{z}_0, \dots, \mathbf{z}_{n-1}, \mathbf{x}_0, \dots, \mathbf{x}_n$  jsou sestrojeny metodou sdružených směrů nultého řádu, tedy pomocí vzorců (3.34) - (3.3.3). Jsou-li vektory  $\mathbf{z}_0, \dots, \mathbf{z}_{n-1}$  nenulové, pak jsou  $\mathbf{A}$ -sdružené, tedy  $f(\mathbf{x}_n) = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$*

*Důkaz.* Indukcí.

Budeme předpokládat  $\mathbf{A}$ -sdruženost vektorů  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1}$ . Jako směry, ve kterých budeme provádět jednorozměrné optimalizace tedy máme  $\{\mathbf{s}_{k,k}, \dots, \mathbf{s}_{k,n-1}, \mathbf{z}_0, \dots, \mathbf{z}_{k-1}\}$ . Potom podle (3.34) - (3.3.3) je bod  $\mathbf{y}_{k,n}$  sestrojen jednorozměrnými minimalizacemi ve směrech vektorů  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1}$ , které jsou  $\mathbf{A}$ -sdružené. Tyto jednorozměrné minimalizace vycházejí z bodu  $\mathbf{y}_{k,n-k}$ . To podle Věty 3.11 znamená, že platí

$$f(\mathbf{y}_{k,n}) = \min_{\mathbf{x} \in \mathbf{y}_{k,n-k} + \mathcal{L}} f(\mathbf{x}),$$

kde  $\mathcal{L} = \text{Lin}\{\mathbf{z}_0, \dots, \mathbf{z}_{k-1}\}$ . Stejně  $\mathbf{x}_k$  je sestrojen pomocí minimalizací ve směrech  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1}$  vycházející z bodu  $\mathbf{y}_{k-1,n-k+1}$ , tedy

$$f(\mathbf{x}_k) = \min_{\mathbf{x} \in \mathbf{y}_{k-1,n-k+1} + \mathcal{L}} f(\mathbf{x}).$$

Z Věty 3.13 tedy vyplývá, že  $\mathbf{z}_k = \mathbf{y}_{k,n} - \mathbf{x}_k$  je  $\mathbf{A}$ -sdružený s  $\mathbf{z}_0, \dots, \mathbf{z}_{k-1}$  a z Věty 3.11 vyplývá  $f(\mathbf{x}_n) = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$ . □

**Příklad 13.** Pomocí metody sdružených směrů nultého řádu zjistěte minimum funkce typu (3.19), kde

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Výchozí bod  $\mathbf{x}_0 = (0; 0)^T$  a prvotní směry pro jednorozměrné minimalizace budou  $\mathbf{s}_0 = (1; 0)^T, \mathbf{s}_1 = (0; 1)^T$ .

**Řešení.** Abychom si ulehčili práci, tak budeme jednorozměrnou minimalizaci provádět pomocí vzorce (3.27), tudíž budeme využívat derivace účelové funkce, i když bychom nemuseli. Provedeme tedy po sobě  $n = 2$  jednorozměrné minimalizace

$$\mathbf{y}_{0,0} = \mathbf{x}_0 + \frac{\mathbf{s}_0^T \mathbf{r}_0}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} \mathbf{s}_0 = \mathbf{x}_0 - \frac{\mathbf{s}_0^T (\mathbf{A} \mathbf{x}_0 - \mathbf{b})}{\mathbf{s}_0^T \mathbf{A} \mathbf{s}_0} \mathbf{s}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\mathbf{y}_{0,1} = \mathbf{y}_{0,0} - \frac{\mathbf{s}_1^T (\mathbf{A} \mathbf{y}_{0,0} - \mathbf{b})}{\mathbf{s}_1^T \mathbf{A} \mathbf{s}_1} \mathbf{s}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Nyní můžeme najít směr, ve kterém provedeme jednorozměrnou minimalizaci za účelem nalezení následujícího bodu minimalizující posloupnosti

$$\mathbf{z}_0 = \mathbf{y}_{0,1} - \mathbf{x}_0 = \frac{1}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix},$$



## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.4. NEWTONOVSKÉ METODY

---

$$\mathbf{x}_1 = \mathbf{x}_0 + \frac{\mathbf{z}_0^T \mathbf{r}_0}{\mathbf{z}_0^T \mathbf{A} \mathbf{z}_0} \mathbf{z}_0 = \mathbf{x}_0 - \frac{\mathbf{z}_0^T (\mathbf{A} \mathbf{x}_0 - b)}{\mathbf{z}_0^T \mathbf{A} \mathbf{z}_0} \mathbf{z}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 2 \cdot \frac{1}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \frac{2}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Tímto jsme dokončili jeden krok metody a přiřadíme vektor  $\mathbf{z}_0$  mezi vektory směrů, ve kterých je prováděna jednorozměrná minimalizace

$$\{\mathbf{s}_0, \mathbf{s}_1\} \longrightarrow \{\mathbf{s}_1, \mathbf{z}_0\}.$$

Následuje druhý krok

$$\mathbf{y}_{1,0} = \mathbf{x}_1 - \frac{\mathbf{s}_1^T (\mathbf{A} \mathbf{x}_1 - b)}{\mathbf{s}_1^T \mathbf{A} \mathbf{s}_1} \mathbf{s}_1 = \frac{2}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 10 \\ -3 \end{pmatrix},$$
$$\mathbf{y}_{1,1} = \mathbf{y}_{1,0} - \frac{\mathbf{z}_0^T (\mathbf{A} \mathbf{y}_{1,0} - b)}{\mathbf{z}_0^T \mathbf{A} \mathbf{z}_0} \mathbf{z}_0 = \frac{1}{5} \begin{pmatrix} 10 \\ -3 \end{pmatrix} + \frac{1}{2} \cdot \frac{1}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 25 \\ -7 \end{pmatrix}.$$

Nyní zjistíme druhý  $\mathbf{A}$ -sdružený vektor  $\mathbf{z}_1$

$$\mathbf{z}_1 = \mathbf{y}_{1,1} - \mathbf{x}_1 = \frac{1}{10} \begin{pmatrix} 25 \\ -7 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 5 \\ -3 \end{pmatrix}.$$

Následný vypočítaný bod minimalizující posloupnosti již bude naším hledaným minimem, jelikož bude sestaven pomocí  $\mathbf{A}$ -sdružených vektorů  $\mathbf{z}_0$  a  $\mathbf{z}_1$

$$\mathbf{x}_2 = \mathbf{x}_1 + \frac{\mathbf{z}_1^T \mathbf{r}_1}{\mathbf{z}_1^T \mathbf{A} \mathbf{z}_1} \mathbf{z}_1 = \mathbf{x}_1 - \frac{\mathbf{z}_1^T (\mathbf{A} \mathbf{x}_1 - b)}{\mathbf{z}_1^T \mathbf{A} \mathbf{z}_1} \mathbf{z}_1 = \frac{2}{5} \begin{pmatrix} 5 \\ -1 \end{pmatrix} + 2 \cdot \frac{1}{10} \begin{pmatrix} 5 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Můžeme jen ukázat, že vektory  $\mathbf{z}_0$  a  $\mathbf{z}_1$  jsou skutečně  $\mathbf{A}$ -sdružené

$$\mathbf{z}_0^T \mathbf{A} \mathbf{z}_1 = \frac{2}{5} (5; -1) \cdot \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \cdot \frac{1}{10} \begin{pmatrix} 5 \\ -3 \end{pmatrix} = \frac{2}{50} (3; 5) \begin{pmatrix} 5 \\ -3 \end{pmatrix} = 0.$$

## 3.4 Newtonovské metody

V této části se budeme zabývat metodami, které ke svému průběhu budou potřebovat nejen znalost první derivace účelové funkce  $f'(x)$ , ale také vyžadují znalost derivace druhé  $f''(x)$ . Začneme s metodou, podle které je tato sekce nazvána, tedy metodou Newtonovou.

### 3.4.1 Newtonova Metoda

Jak již bylo zmíněno v úvodu, budeme nyní vyžadovat, aby druhá derivace účelové funkce  $f''(x)$  existovala a byla spojitá na  $\mathbb{R}^n$ . Zatím tedy máme na Hesseho matici druhých derivací jen podmínku symetrie. V průběhu výpočtu uvidíme jaké další podmínky metoda na tuto matici bude klást.

Nechť máme bod  $\mathbf{x}_k$  jako výchozí bod v obecném  $k$ -tém kroku. Metoda v každém kroku nahrazuje minimalizovanou účelovou funkci Taylorovým polynomem 2. stupně se

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

středem v bodě  $\mathbf{x}_k$  a hledá minimum této aproximace na místo minima účelové funkce. Za předpokladu, že  $f$  je silně konvexní, je tento bod jediným bodem globálního minima.

Taylorův polynom se středem v  $\mathbf{x}_k$ , kterým budeme nahrazovat danou účelovou funkci, zapíšeme jako

$$F_{\mathbf{x}_k}(\mathbf{x}) = f(\mathbf{x}_k) + (\mathbf{x} - \mathbf{x}_k)^T f'(\mathbf{x}_k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T f''(\mathbf{x}_k) (\mathbf{x} - \mathbf{x}_k).$$

Dále  $\mathbf{x}_{k+1}$  najdeme jako řešení úlohy  $F_{\mathbf{x}_k}(\mathbf{x}) \rightarrow \min$ . Tuto úlohu získáme derivací Taylorova polynomu podle  $\mathbf{x}$  a položením rovno nule. Úpravou dostaneme

$$\begin{aligned} \frac{\partial F_{\mathbf{x}_k}(\mathbf{x})}{\partial \mathbf{x}} &= f'(\mathbf{x}_k) + (\mathbf{x} - \mathbf{x}_k)^T f''(\mathbf{x}_k) = 0, \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - (f''(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k). \end{aligned} \quad (3.39)$$

Z tohoto výrazu vidíme, že Hesseho matice druhých derivací musí být regulární, abychom mohli výpočet touto metodou provést. Dále pokud je  $f$  konvexní, tak je matice  $f''(\mathbf{x}_k)$  pozitivně semidefinitní, a tedy  $\mathbf{x}_{k+1}$  bude globálním minimem  $F_{\mathbf{x}_k}(\mathbf{x})$  na  $\mathbb{R}^n$ .

Všimněme si, že pokud bude zadaná účelová funkce  $f$  funkcí kvadratickou, tak bude Hesseho matice konstantní a bod minima aproximace se bude rovnat skutečnému bodu minima účelové funkce. Potom je možné najít tento bod minima v jediném kroku.

Nyní si dokážeme větu, která nám napoví ohledně rychlosti konvergence Newtonovy metody.

**Věta 3.15.** *Nechť účelová funkce  $f$  je na  $\mathbb{R}^n$  dvakrát diferencovatelná, silně konvexní s konstantou  $\theta > 0$  a splňuje*

$$\|f''(\mathbf{x}) - f''(\mathbf{y})\| \leq M \|\mathbf{x} - \mathbf{y}\|, \quad (3.40)$$

pro libovolné  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , tedy  $f''(\mathbf{x})$  je Lipschitzovská s konstantou  $M > 0$ . Pokud splňuje výchozí bod  $\mathbf{x}_0$

$$\|f''(\mathbf{x}_0)\| < \frac{8\theta^2}{M}, \quad (3.41)$$

tedy existuje  $q \in (0, 1)$ , pro které platí

$$\|f''(\mathbf{x}_0)\| = \frac{8\theta^2 q}{M}, \quad (3.42)$$

potom posloupnost  $\{\mathbf{x}_k\}$  daná vztahem (3.39) konverguje k jedinému bodu minima účelové funkce  $\mathbf{x}^*$  s kvadratickou rychlostí, tedy

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \frac{4\theta q^{2^k}}{M}.$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.4. NEWTONOVSKÉ METODY

---

*Důkaz.* Jelikož daná účelová funkce je silně konvexní, tak má právě jedno globální minimum na  $\mathbb{R}^n$  a platí pro ni

$$\mathbf{s}^T f''(\mathbf{x}) \mathbf{s} \geq 2\theta \|\mathbf{s}\|^2, \quad (3.43)$$

pro  $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n$ . Z tohoto víme, že matice  $f''(\mathbf{x})$  je pozitivně definitní pro každé  $\mathbf{x} \in \mathbb{R}^n$ , tedy je regulární a můžeme tudíž využít výpočtu pomocí Newtonovy metody.

Dále pro silně konvexní účelovou funkci víme, že platí

$$(f'(\mathbf{x}_k) - f'(\mathbf{x}^*))^T (\mathbf{x}_k - \mathbf{x}^*) \geq 2\theta \|\mathbf{x}_k - \mathbf{x}^*\|^2,$$

jelikož platí  $f'(\mathbf{x}^*) = 0$  a využitím Cauchy-Schwarzovy nerovnosti získáváme

$$\|\mathbf{x}_k - \mathbf{x}^*\|^2 \leq \frac{1}{2\theta} (f'(\mathbf{x}_k) - f'(\mathbf{x}^*))^T (\mathbf{x}_k - \mathbf{x}^*) \leq \frac{1}{2\theta} \|f'(\mathbf{x}_k)\| \|\mathbf{x}_k - \mathbf{x}^*\|,$$

tedy

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \frac{1}{2\theta} \|f'(\mathbf{x}_k)\|. \quad (3.44)$$

Nyní odhadneme  $\|f'(\mathbf{x}_k)\|$ . Mějme identitu

$$\frac{\partial}{\partial \alpha} f'(\mathbf{x}_k + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)) = f''(\mathbf{x}_k + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)) (\mathbf{x}_{k+1} - \mathbf{x}_k).$$

Tuto identitu integrujme podle  $\alpha$  od 0 do 1 a od obou stran odečteme  $f''(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k)$

$$\begin{aligned} \left[ f'(\mathbf{x}_k + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)) \right]_0^1 - f''(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) &= \\ &= \int_0^1 f''(\mathbf{x}_k + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)) (\mathbf{x}_{k+1} - \mathbf{x}_k) d\alpha - f''(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k), \end{aligned}$$

úpravou získáme

$$\begin{aligned} f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k) - f''(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) &= \\ &= \int_0^1 [f''(\mathbf{x}_k + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)) - f''(\mathbf{x}_k)] (\mathbf{x}_{k+1} - \mathbf{x}_k) d\alpha. \end{aligned}$$

Jelikož máme na pravé straně této rovnice rozdíl dvou druhých derivací v rozdílných bodech, můžeme použít (3.40) a získáme

$$\|f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k) - f''(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k)\| \leq \frac{M}{2} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|^2 \quad (3.45)$$

---

<sup>5</sup>Důkaz tohoto tvrzení je mimo zaměření této práce a tudíž zde není napsán. Je ho možné najít například v [3], str. 38.

<sup>6</sup>Důkaz neuveden ze stejného důvodu jako u výrazu výše. Je ho možné najít například v [9], str. 91.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

Dosazením do obou stran z rovnice (3.39) ve tvaru  $\mathbf{x}_{k+1} - \mathbf{x}_k = -(f''(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)$  dostaneme nerovnost

$$\|f'(\mathbf{x}_{k+1})\| \leq \frac{M}{2} \|(f''(\mathbf{x}_k))^{-1}\|^2 \|f'(\mathbf{x}_k)\|^2. \quad (3.46)$$

Nyní se pokusíme odhadnout normu matice  $\|(f''(\mathbf{x}_k))^{-1}\|$ . Položme  $\mathbf{s} = (f''(\mathbf{x}_k))^{-1} \mathbf{y}$ , pak využijeme (3.45), Cauchy-Schwarzovu nerovnost a dostaneme

$$\|(f''(\mathbf{x}_k))^{-1} \mathbf{y}\|^2 \leq \frac{1}{2\theta} \mathbf{y}^T (f''(\mathbf{x}_k))^{-1} \mathbf{y} \leq \frac{1}{2\theta} \|\mathbf{y}\| \|(f''(\mathbf{x}_k))^{-1} \mathbf{y}\|.$$

Úpravou první a třetí části této nerovnice dostaneme

$$\frac{\|(f''(\mathbf{x}_k))^{-1} \mathbf{y}\|}{\|\mathbf{y}\|} \leq \frac{\|(f''(\mathbf{x}_k))^{-1}\| \|\mathbf{y}\|}{\|\mathbf{y}\|} = \|(f''(\mathbf{x}_k))^{-1}\|.$$

Tedy získáváme

$$\|(f''(\mathbf{x}_k))^{-1}\| = \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|(f''(\mathbf{x}_k))^{-1} \mathbf{y}\|}{\|\mathbf{y}\|} \leq \frac{1}{2\theta}.$$

Dosazením do (3.46) dostaneme

$$\|f'(\mathbf{x}_{k+1})\| \leq \frac{M}{8\theta^2} \|f'(\mathbf{x}_k)\|^2.$$

Tato nerovnost je odvozena pro libovolné  $k$ , tedy můžeme postupně dosazovat zpět do (3.46). Jakmile se dostaneme k  $k = 0$ , použijeme rovnost (3.42)

$$\begin{aligned} \|f'(\mathbf{x}_k)\| &\leq \frac{M}{8\theta^2} \|f'(\mathbf{x}_{k-1})\|^2 \leq \left(\frac{M}{8\theta^2}\right)^3 \|f'(\mathbf{x}_{k-2})\|^4 \leq \dots \leq \\ &\leq \left(\frac{M}{8\theta^2}\right)^{2^k-1} \|(f''(\mathbf{x}_0))^{-1}\|^{2^k} = \frac{8\theta^2}{M} q^{2^k}. \end{aligned}$$

Tímto jsme dokázali první nerovnici a druhou získáme dosazením tohoto výsledku do (3.44)

$$\begin{aligned} \|\mathbf{x}_k - \mathbf{x}^*\| &\leq \frac{1}{2\theta} \cdot \frac{8\theta^2}{M} q^{2^k}, \\ \|\mathbf{x}_k - \mathbf{x}^*\| &\leq \frac{4\theta}{M} q^{2^k}. \end{aligned}$$

□

Tato věta nám objasnila výhody a komplikace, které plynou z využití Newtonovy metody. Výhodou je konvergence kvadratickou rychlostí, kterou se metoda řadí mezi rychle konvergující metody. Nevýhodou výpočetního procesu je však značná numerická náročnost výpočtu  $(f''(\mathbf{x}_k))^{-1}$  při větším  $n$ . Tuto zásadní nevýhodu se pokusíme odstranit v další probírané metodě.

Dále je také vidět, že pro korektní fungování Newtonovy metody máme omezený výběr výchozího bodu podmínkou (3.41). Při nesplnění této podmínky existují i silně konvexní funkce, které nebudou konvergovat.

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

**Příklad 14.** Proved'te jeden krok Newtonovou metodou. Pro přehlednost označíme  $\mathbf{x} = (u_1, u_2)$  Minimalizovaná účelová funkce má vyjádření

$$f(\mathbf{x}) = f(u_1, u_2) = (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{3}{2}},$$

výchozím bodem bude  $\mathbf{x}_0 = (2; 1)^T$ .

**Řešení.** Jako první si vypočítáme první a druhé derivace účelové funkce

$$\begin{aligned} f'_{u_1} &= 3 (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{1}{2}} \cdot (u_1 + u_1 u_2^2), \\ f'_{u_2} &= \frac{3}{2} (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{1}{2}} \cdot (1 + 2u_1^2 u_2), \\ f'_{u_1, u_1} &= 3 (u_1^2 + u_2 + u_1^2 u_2^2)^{-\frac{1}{2}} \cdot (u_1 + u_1 u_2^2)^2 + 3 (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{1}{2}} \cdot (1 + u_2^2), \\ f'_{u_2, u_2} &= \frac{3}{4} (u_1^2 + u_2 + u_1^2 u_2^2)^{-\frac{1}{2}} \cdot (1 + 2u_1^2 u_2)^2 + 3 (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{1}{2}} \cdot u_1^2, \\ f'_{u_1, u_2} &= f'_{u_2, u_1} = \frac{3}{2} (u_1^2 + u_2 + u_1^2 u_2^2)^{-\frac{1}{2}} \cdot (u_1 + u_1 u_2^2) \cdot (1 + 2u_1^2 u_2) + \\ &\quad + 3 (u_1^2 + u_2 + u_1^2 u_2^2)^{\frac{1}{2}} \cdot 2u_1 u_2. \end{aligned}$$

Můžeme tedy zapsat gradient a Hesseho matici v bodě  $\mathbf{x}_0 = (2; 1)$

$$f''(\mathbf{x}_0) = \begin{pmatrix} 34 & 54 \\ 54 & \frac{225}{4} \end{pmatrix}, \quad f'(\mathbf{x}_0) = \begin{pmatrix} 36 \\ \frac{81}{2} \end{pmatrix}$$

Nyní můžeme vyřešit příklad dvěma způsoby. První způsob zahrnuje vypsání Taylorova mnohočlenu dané funkce středem v bodě  $\mathbf{x}_0 = (2; 1)^T$  a poté derivování tohoto mnohočlenu a položení derivací rovno nule. Tímto standardním minimalizačním postupem se dostaneme bez složitých výpočtů k minimu naší aproximace, a tedy k dalšímu bodu minimalizující posloupnosti. Druhou možností je dosadit přímo do vzorce (3.39). Při tomto postupu musíme ovšem počítat inverzi Hesseho matice, avšak pro malé  $n$  může být tento postup znatelně rychlejší.

Postup 1)

Zjistíme vyjádření Taylorova mnohočlenu se středem v bodě  $\mathbf{x}_0 = (2; 1)^T$

$$\begin{aligned} F_{\mathbf{x}_0}(\mathbf{x}) &= f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T f'(\mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T f''(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) = \\ &= 27 + (u_1 - 2; u_2 - 1) \begin{pmatrix} 36 \\ \frac{81}{2} \end{pmatrix} + \frac{1}{2} (u_1 - 2; u_2 - 1) \begin{pmatrix} 34 & 54 \\ 54 & \frac{225}{4} \end{pmatrix} \begin{pmatrix} u_1 - 2 \\ u_2 - 1 \end{pmatrix} = \\ &= 27 + 36(u_1 - 2) + \frac{81}{2}(u_2 - 1) + 17(u_1 - 2)^2 + 54(u_1 - 2) \cdot (u_2 - 1) + \frac{225}{8}(u_2 - 1)^2. \end{aligned}$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.4. NEWTONOVSKÉ METODY

---

Nyní máme Taylorův mnohočlen. Můžeme provést parciální derivace a položit je rovny nule. Tímto získáme bod minima aproximace

$$\begin{aligned}\frac{\partial F_{\mathbf{x}_0}(\mathbf{x})}{\partial u_1} &= 36 + 34(u_1 - 2) + 54(u_2 - 1) = 0, \\ \frac{\partial F_{\mathbf{x}_0}(\mathbf{x})}{\partial u_2} &= \frac{81}{2} + 54(u_1 - 2) + \frac{225}{4}(u_2 - 1) = 0.\end{aligned}$$

Úpravou dostaneme soustavu lineárních rovnic

$$\begin{aligned}34u_1 + 54u_2 - 86 &= 0, \\ 54u_1 + \frac{225}{4}u_2 - \frac{495}{4} &= 0.\end{aligned}$$

Řešením této soustavy a tedy náš hledaný další bod minimalizující posloupnosti je bod  $\mathbf{x}_1 = \left(\frac{410}{223}; \frac{97}{223}\right)^T = (1, 84; 0, 43)^T$ .

Postup 2)

Při tomto postupu použijeme přímo vzorce pro výpočet dalšího bodu minimalizující posloupnosti

$$\mathbf{x}_1 = \mathbf{x}_0 - (f''(\mathbf{x}_0))^{-1} f'(\mathbf{x}_0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} -\frac{25}{446} & \frac{12}{223} \\ \frac{12}{223} & -\frac{68}{2007} \end{pmatrix} \cdot \begin{pmatrix} 36 \\ \frac{81}{2} \end{pmatrix} = \begin{pmatrix} \frac{410}{223} \\ \frac{97}{223} \end{pmatrix}.$$

Je vidět, že druhý postup zabral jediný řádek oproti prvnímu postupu, jehož výpočet trval mnohem déle. Musíme si však uvědomit, že jsme měli malý rozměr matice  $(f''(\mathbf{x}))^{-1}$  a výpočet této inverzní matice nebyl výpočetně náročný.

#### 3.4.2 Metoda proměnné metriky

Jak již bylo zmíněno, tak se Newtonova metoda potýká s problémem numerické náročnosti výpočtu  $(f''(\mathbf{x}))^{-1}$ . Metoda proměnné metriky se tedy nesnaží vypočítat tuto inverzi najednou, ale postupným výpočtem, stejně jako se v metodě sdružených směrů postupně získávají  $\mathbf{A}$ -sdružené vektory. Teoretický základ je rovněž odvozen z této metody, tu však již známe, tak si můžeme tuto ideu lehce osvětlit.

Budeme opět uvažovat kvadratickou účelovou funkci tvaru (3.19), platí tedy  $\mathbf{A} = f''(\mathbf{x})$ . Pro libovolný bod  $\mathbf{x} \in \mathbb{R}^n$  pomocí úpravy derivace účelové funkce platí

$$\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b},$$

dosazením za  $\mathbf{b}$  a úpravou získáme

$$\mathbf{x}^* = \mathbf{A}^{-1}(\mathbf{A}\mathbf{x} - f'(\mathbf{x})) = \mathbf{x} - \mathbf{A}^{-1}f'(\mathbf{x}).$$

## KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

### 3.4. NEWTONOVSKÉ METODY

---

Tato rovnice nám říká, že při provedení minimalizace z libovolného  $\mathbf{x} \in \mathbb{R}^n$  ve směru  $-\mathbf{A}^{-1}f'(\mathbf{x})$  se dostaneme tímto krokem do bodu minima účelové funkce. Nyní ukážeme, že pro metodu sdružených směrů toto platí pro krok  $k = n - 1$ . Pro tento krok v metodě sdružených gradientů z rovnice (3.29) platí

$$\mathbf{s}_{n-1} = \frac{1}{\alpha_{n-1}} \mathbf{A}^{-1} (\mathbf{r}_{n-1} - \mathbf{r}_n),$$

jelikož se jedná o poslední krok metodou, tak víme, že  $\mathbf{r}_n = 0$ , potom tedy platí

$$\mathbf{s}_{n-1} = \frac{1}{\alpha_{n-1}} \mathbf{A}^{-1} \mathbf{r}_{n-1} = -\frac{1}{\alpha_{n-1}} \mathbf{A}^{-1} f'(\mathbf{x}_{n-1}),$$

Při dosazení této rovnice do (3.28) dostaneme

$$\mathbf{x}_n = \mathbf{x}_{n-1} - \mathbf{A}^{-1} f'(\mathbf{x}_{n-1}),$$

Což je přesně naše minimalizace, která směřuje do bodu minima dané účelové funkce.

Metoda proměnné metriky tohoto využívá. Nechť  $\mathbf{Z}_0$  je libovolná symetrická, pozitivně definitní matice. Nový bod minimalizující posloupnosti pro obecný  $k$ -tý krok je získán jako

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{Z}_k f'(\mathbf{x}_k), \quad (3.47)$$

z čehož je vidět, že se metoda bude snažit o postupné upravování matice  $\mathbf{Z}_k$  tak, aby platilo  $\mathbf{Z}_n = \mathbf{A}^{-1}$ . Stejně jako v metodě sdružených gradientů nyní můžeme vypočítat  $\alpha_k$ , tedy koeficient udávající délku kroku při jednorozměrné minimalizaci ve směru  $-\mathbf{Z}_k f'(\mathbf{x}_k)$

$$\begin{aligned} \mathbf{x}_n &= \mathbf{x}_0 - \sum_{k=0}^{n-1} \alpha_k \mathbf{Z}_k f'(\mathbf{x}_k), \\ \mathbf{x}_n - \mathbf{x}_0 &= -\sum_{k=0}^{n-1} \alpha_k \mathbf{Z}_k f'(\mathbf{x}_k), \end{aligned}$$

Z toho pro  $\mathbf{A}$ -sdružené vektory  $-\mathbf{Z}_k f'(\mathbf{x}_k)$  platí

$$\begin{aligned} \alpha_k &= -\frac{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{x}_n - \mathbf{x}_0)}{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{Z}_k f'(\mathbf{x}_k))} = -\frac{(\mathbf{Z}_k f'(\mathbf{x}_k))^T (\mathbf{b} - \mathbf{A} \mathbf{x}_0)}{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{Z}_k f'(\mathbf{x}_k))} = \\ &= \frac{f'(\mathbf{x}_k)^T \mathbf{Z}_k f'(\mathbf{x}_0)}{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{Z}_k f'(\mathbf{x}_k))}. \end{aligned}$$

Rovnost

$$\alpha_k = \frac{f'(\mathbf{x}_k)^T \mathbf{Z}_k f'(\mathbf{x}_0)}{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{Z}_k f'(\mathbf{x}_k))} = \frac{f'(\mathbf{x}_k)^T \mathbf{Z}_k f'(\mathbf{x}_k)}{(\mathbf{Z}_k f'(\mathbf{x}_k))^T \mathbf{A} (\mathbf{Z}_k f'(\mathbf{x}_k))},$$

je možné dokázat stejně jako ve Větě 3.8. Dále využijeme vektoru  $\mathbf{s}_k$  definovaného jako

$$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k = -\alpha \mathbf{Z}_k f'(\mathbf{x}_k),$$

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

Nyní již nepočítáme celý nový směrový vektor, ale musíme vypočítat matici  $\mathbf{Z}_{k+1}$ , abychom mohli pokračovat ve výpočtu

$$\mathbf{Z}_{k+1} = \mathbf{Z}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} + \mathbf{C}_k. \quad (3.48)$$

Když porovnáme tento vzorec s vzorcem využívaným pro výpočet nového směru při použití metody sdružených směrů, je očividně jeden člen navíc. Tímto členem je  $\mathbf{C}_k$ , které má za úlohu, stejně jako  $\beta_k \mathbf{s}_k$  u metody sdružených směrů, zajistit  $\mathbf{A}$ -sdruženost vektorů  $\mathbf{s}_0, \dots, \mathbf{s}_k$ . Matice  $\mathbf{C}_k$  musí být určena tak, aby bylo splněno

$$\mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_k = \mathbf{s}_k.$$

Dosazením této rovnice do rovnice výše a úpravou získáváme

$$\begin{aligned} \mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_k &= \mathbf{Z}_k \mathbf{A} \mathbf{s}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{A} \mathbf{s}_k} \mathbf{A} \mathbf{s}_k + \mathbf{C}_k \mathbf{A} \mathbf{s}_k, \\ \mathbf{s}_k &= \mathbf{Z}_k \mathbf{A} \mathbf{s}_k + \mathbf{s}_k + \mathbf{C}_k \mathbf{A} \mathbf{s}_k, \\ (\mathbf{Z}_k + \mathbf{C}_k) \mathbf{A} \mathbf{s}_k &= 0. \end{aligned}$$

Pro lepší přehlednost provedeme substituci  $\mathbf{h}_k = \mathbf{A} \mathbf{s}_k$  z definice  $\mathbf{s}_k$  můžeme tuto rovnost přepsat též jako  $\mathbf{h}_k = f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k)$ . Podmínku výše můžeme upravit do tvaru

$$\mathbf{C}_k = -\frac{(\mathbf{Z}_k \mathbf{h}_k)(\mathbf{Z}_k \mathbf{h}_k)^T}{\mathbf{h}_k^T \mathbf{Z}_k \mathbf{h}_k}.$$

Z této rovnice je dále jasné, že musí platit  $\mathbf{h}_k^T \mathbf{Z}_k \mathbf{h}_k \neq 0$ . Vložením poslední rovnice do (3.48) a z definice  $\mathbf{h}_k$  získáme rovnici

$$\mathbf{Z}_{k+1} = \mathbf{Z}_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{h}_k} - \frac{(\mathbf{Z}_k \mathbf{h}_k)(\mathbf{Z}_k \mathbf{h}_k)^T}{\mathbf{h}_k^T \mathbf{Z}_k \mathbf{h}_k}. \quad (3.49)$$

Tato transformace nám navíc zachovává pozitivní definitnost matice  $\mathbf{Z}_{k+1}$ .

**Věta 3.16.** *Nechť je účelová funkce tvaru (3.19), kde  $\mathbf{A}$  je symetrická, pozitivně definitní matice. Pak vektory  $\mathbf{s}_0, \dots, \mathbf{s}_{k-1}$  jsou v libovolném  $k$ -tém kroku  $\mathbf{A}$ -sdružené a  $\mathbf{Z}_n = \mathbf{A}^{-1} = (f''(\mathbf{x}_n))^{-1}$ .*

*Důkaz.* Platí

$$\mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_k = \mathbf{s}_k.$$

pro  $k = 0$  tedy spolu s definicí  $\mathbf{h}_k$  platí

$$\mathbf{Z}_1 \mathbf{A} \mathbf{s}_0 = \mathbf{Z}_1 \mathbf{h}_0 = \mathbf{s}_0.$$

Napíšeme si pomocnou rovnici

$$\mathbf{Z}_1 f'(\mathbf{x}_1) = \frac{1}{\alpha_1} (\mathbf{x}_1 - \mathbf{x}_2) = -\frac{\mathbf{s}_1}{\alpha_1},$$



KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

Jelikož stejně jako u předešlých metod platí, že směr derivace je kolmý k směru, kterým jsme se do daného bodu dostali, tak můžeme dokázat

$$0 = \mathbf{s}_0^T f'(\mathbf{x}_1) = (\mathbf{Z}_1 \mathbf{A} \mathbf{s}_0)^T f'(\mathbf{x}_1) = (\mathbf{A} \mathbf{s}_0)^T \mathbf{Z}_1 f'(\mathbf{x}_1) = -\frac{1}{\alpha_1} \mathbf{s}_0^T \mathbf{A} \mathbf{s}_1.$$

Toto nám dokazuje platnost  $\mathbf{A}$ -sdruženosti pro první dva vektory  $\mathbf{s}$ . Pro dokázání této vlastnosti pro  $\mathbf{s}_1$  a  $\mathbf{s}_2$  položíme

$$\mathbf{Z}_2 \mathbf{A} \mathbf{s}_1 = \mathbf{Z}_2 \mathbf{h}_1 = \mathbf{s}_1.$$

V důkazu bychom pokračovali stejně jako u předchozího kroku. Nyní se pokusíme dokázat  $\mathbf{A}$ -sdruženost vektorů  $\mathbf{s}_0$  a  $\mathbf{s}_2$ . Z  $\mathbf{A}$ -sdruženosti vektorů  $\mathbf{s}_0, \mathbf{s}_1$  a z rovnosti  $\mathbf{s}_0 f'(\mathbf{x}_1) = 0$  vyvodíme rovnici

$$\mathbf{s}_0^T (f'(\mathbf{x}_1) + \mathbf{A} \mathbf{s}_1) = 0.$$

Z definice  $\mathbf{h}_k$  víme, že  $f'(\mathbf{x}_2) = f'(\mathbf{x}_1) + \mathbf{A} \mathbf{s}_1$ , tedy platí

$$\mathbf{s}_0 f'(\mathbf{x}_2) = 0.$$

Opět vytvoříme pomocnou rovnici

$$\mathbf{Z}_2 f'(\mathbf{x}_2) = \frac{1}{\alpha_2} (\mathbf{x}_2 - \mathbf{x}_3) = -\frac{\mathbf{s}_2}{\alpha_2}$$

Potom využijeme toho, že pokud vynásobíme rovnici (3.49) členem  $\mathbf{A} \mathbf{s}_0$ , tak nám vyjde rovnice  $\mathbf{Z}_2 \mathbf{A} \mathbf{s}_0 = \mathbf{Z}_1 \mathbf{A} \mathbf{s}_0 = \mathbf{s}_0$ , jelikož jsou ostatní členy díky  $\mathbf{A}$ -sdruženosti vektorů  $\mathbf{s}_0, \mathbf{s}_1$  nulové. Nyní můžeme upravit

$$0 = \mathbf{s}_0 f'(\mathbf{x}_2) = (\mathbf{Z}_2 \mathbf{A} \mathbf{s}_0)^T f'(\mathbf{x}_2) = (\mathbf{A} \mathbf{s}_0)^T \mathbf{Z}_2 f'(\mathbf{x}_2) = -\frac{1}{\alpha_2} \mathbf{s}_0^T \mathbf{A} \mathbf{s}_2.$$

Nyní jsme dokázali, že vzorce

$$\mathbf{Z}_k \mathbf{A} \mathbf{s}_i = \mathbf{s}_i, \quad \text{pro } 0 \leq i \leq k, \quad (3.50)$$

$$\mathbf{s}_i^T \mathbf{A} \mathbf{s}_j = 0, \quad \text{pro } 0 \leq i < j < k, \quad (3.51)$$

$$\mathbf{s}_i^T f'(\mathbf{x}_k) = 0, \quad \text{pro } 0 \leq i < k, \quad (3.52)$$

mohou být dokázány indukcí pro  $k$ . Pro  $k = 0, 1$  jsme platnost již dokázali. Nyní tedy předpokládejme, že vzorce platí pro  $k$  a dokážeme, že platí pro  $k + 1$ . Pro  $i = k$  platí vzorec (3.52) z definice metody. Prozkoumáme tedy celou možnost  $i < k$ . Vyjdeme z rovnice

$$f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_{i+1}) = \mathbf{A}(\mathbf{x}_{k+1} - \mathbf{x}_{i+1}) = \mathbf{A}(\mathbf{x}_{k+1} - \mathbf{x}_k + \mathbf{x}_k - \mathbf{x}_{k-1}) + \dots + \mathbf{x}_{i+2} - \mathbf{x}_{i+1}).$$

Pomocí definice  $\mathbf{s}_k$  získáme

$$f'(\mathbf{x}_{k+1}) = f'(\mathbf{x}_{i+1}) + \mathbf{A}(\mathbf{s}_{i+1} + \dots + \mathbf{s}_k)$$

KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE  
3.4. NEWTONOVSKÉ METODY

---

Celou rovnici vynásobíme  $\mathbf{s}_i$

$$\mathbf{s}_i f'(\mathbf{x}_{k+1}) = \mathbf{s}_i^T f'(\mathbf{x}_{i+1}) + \mathbf{s}_i^T \mathbf{A} (\mathbf{s}_{i+1} + \dots + \mathbf{s}_k) = 0.$$

Tato rovnost platí z našeho předpokladu platnosti  $\mathbf{A}$ -sdrúženosti vektorů  $\mathbf{s}_0, \dots, \mathbf{s}_k$ . Tímto jsme dokázali rovnici (3.52). Nyní dokážeme rovnici (3.50). Opět z  $\mathbf{A}$ -sdrúženosti vektorů  $\mathbf{s}_0, \dots, \mathbf{s}_k$  platí

$$(\mathbf{Z}_k \mathbf{h}_k)^T \mathbf{A} \mathbf{s}_i = \mathbf{h}_k^T (\mathbf{Z}_k \mathbf{A} \mathbf{s}_i) = \mathbf{s}_k^T \mathbf{A} \mathbf{s}_i = 0.$$

Pokud tedy obě strany rovnice (3.49) vynásobíme  $\mathbf{A} \mathbf{s}_i$  získáme při použití rovnice výše

$$\mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_i = \mathbf{Z}_k \mathbf{A} \mathbf{s}_i = \mathbf{s}_i,$$

jelikož kromě členu  $\mathbf{Z}_k$  budou oba zbývající členy na pravé straně rovnice (3.49) po vynásobení rovny nule. Nyní stačí dokázat rovnici (3.51). Vyjdeme z již dokázané rovnice (3.52)

$$0 = \mathbf{s}_k^T f'(\mathbf{x}_{k+1}) = (\mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_k)^T f'(\mathbf{x}_{k+1}) = \mathbf{s}_k^T \mathbf{A} \mathbf{Z}_{k+1} f'(\mathbf{x}_{k+1}) = -\frac{1}{\alpha_{k+1}} \mathbf{s}_k^T \mathbf{A} \mathbf{s}_{k+1}.$$

Dále platí

$$\begin{aligned} \mathbf{s}_i^T \mathbf{A} \mathbf{s}_{k+1} &= (\mathbf{A} \mathbf{s}_i)^T (-\alpha_{k+1} \mathbf{Z}_{k+1} f'(\mathbf{x}_{k+1})) = -\alpha_{k+1} (\mathbf{Z}_{k+1} \mathbf{A} \mathbf{s}_i)^T f'(\mathbf{x}_{k+1}) = \\ &= -\alpha_{k+1} \mathbf{s}_i^T (f'(\mathbf{x}_k) + \mathbf{A} \mathbf{s}_k) = 0. \end{aligned}$$

Tímto je dokázána  $\mathbf{A}$ -sdrúženost vektorů  $\mathbf{s}_0, \dots, \mathbf{s}_{k-1}$  pro libovolný  $k$ -tý krok. Z tohoto vyplývá, že vektor  $f'(\mathbf{x}_n)$  musí být roven nule, jelikož je ortogonální k  $n$  lineárně nezávislým vektorům  $\mathbf{s}_0, \dots, \mathbf{s}_{n-1}$ . Tímto způsobem pro

$$\mathbf{Z}_n \mathbf{A} \mathbf{s}_i = \mathbf{s}_i, \quad \text{pro } i = 0, 1, \dots, n-1,$$

musí platit

$$\mathbf{Z}_n \mathbf{A} = \mathbf{I}_n, \quad \text{neboli} \quad \mathbf{Z}_n = \mathbf{A}^{-1}.$$

□

**Příklad 15.** Metodou proměnné metriky najděte minimum funkce tvaru (3.19), pro kterou platí

$$\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

za výchozí bod si zvolíme  $\mathbf{x}_0 = (0; 0)^T$  a úvodní aproximaci matice  $\mathbf{A}^{-1}$  jako

$$\mathbf{Z}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.4. NEWTONOVSKÉ METODY

---

**Řešení.** Jako první zjistíme hodnotu derivace v bodě  $\mathbf{x}_0$

$$f'(\mathbf{x}_0) = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$

Pomocí této derivace můžeme zjistit koeficient udávající délku posunu do následujícího bodu minimalizující posloupnosti a následně i tento bod

$$\alpha_0 = \frac{f'(\mathbf{x}_0)^T \mathbf{Z}_0 f'(\mathbf{x}_0)}{(\mathbf{Z}_0 f'(\mathbf{x}_0))^T \mathbf{A} (\mathbf{Z}_0 f'(\mathbf{x}_0))} = 0,25,$$
$$\mathbf{x}_1 = \mathbf{x}_0 - \alpha_0 \mathbf{Z}_0 f'(\mathbf{x}_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - 0,25 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,5 \\ 0 \end{pmatrix}.$$

Nyní musíme spočítat pomocné vektory  $\mathbf{h}_0$  a  $\mathbf{s}_0$ , abychom mohli zjistit následující aproximaci matice  $\mathbf{A}^{-1}$

$$\mathbf{s}_0 = \mathbf{x}_1 - \mathbf{x}_0 = \begin{pmatrix} 0,5 \\ 0 \end{pmatrix},$$
$$\mathbf{h}_0 = \mathbf{A}\mathbf{s}_0 = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$
$$\mathbf{Z}_1 = \mathbf{Z}_0 + \frac{\mathbf{s}_0 \mathbf{s}_0^T}{\mathbf{s}_0^T \mathbf{h}_0} - \frac{(\mathbf{Z}_0 \mathbf{h}_0) (\mathbf{Z}_0 \mathbf{h}_0)^T}{\mathbf{h}_0^T \mathbf{Z}_0 \mathbf{h}_0} = \begin{pmatrix} 0,45 & -0,4 \\ -0,4 & 0,8 \end{pmatrix}.$$

Nyní máme jak nový bod minimalizující posloupnosti tak novou aproximaci, můžeme tedy začít druhý krok.

$$f'(\mathbf{x}_1) = \mathbf{A}\mathbf{x}_1 - \mathbf{b} = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

V tomto kroku vypočítaný bod minimalizující posloupnosti již bude hledaný bod minima

$$\alpha_1 = \frac{f'(\mathbf{x}_1)^T \mathbf{Z}_1 f'(\mathbf{x}_1)}{(\mathbf{Z}_1 f'(\mathbf{x}_1))^T \mathbf{A} (\mathbf{Z}_1 f'(\mathbf{x}_1))} = 0,625,$$
$$\mathbf{x}_2 = \mathbf{x}_1 - \alpha_1 \mathbf{Z}_1 f'(\mathbf{x}_1) = \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} - 0,625 \begin{pmatrix} 0,45 & -0,4 \\ -0,4 & 0,8 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} =$$
$$= \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} + \begin{pmatrix} 0,25 \\ -0,5 \end{pmatrix} = \begin{pmatrix} 0,75 \\ -0,5 \end{pmatrix}.$$

Pro ověření, že jsme skutečně v bodě minima vypočítáme hodnotu derivace účelové funkce v tomto bodě

$$f'(\mathbf{x}_2) = \mathbf{A}\mathbf{x}_2 - \mathbf{b} = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0,75 \\ -0,5 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

### KAPITOLA 3. METODY VÍCEROZMĚRNÉ OPTIMALIZACE

#### 3.4. NEWTONOVSKÉ METODY

---

Pokud spočítáme matici  $\mathbf{Z}_2$ , tak můžeme ověřit, že platí  $\mathbf{Z}_2 = \mathbf{A}^{-1}$

$$\mathbf{Z}_2 = \mathbf{Z}_1 + \frac{\mathbf{s}_1 \mathbf{s}_1^T}{\mathbf{s}_1^T \mathbf{h}_1} - \frac{(\mathbf{Z}_1 \mathbf{h}_1)(\mathbf{Z}_1 \mathbf{h}_1)^T}{\mathbf{h}_1^T \mathbf{Z}_1 \mathbf{h}_1} = \begin{pmatrix} 0,375 & -0,25 \\ -0,25 & 0,5 \end{pmatrix} = \mathbf{A}^{-1}.$$

Tato konkrétní varianta metody byla vyvinuta matematiky Davidonem, Fletcherem a Powellem. Často je tedy nazývána DFP metodou. Spolu s ostatními kwazinewtonovskými metodami patří mezi nejefektivnější výpočetní postupy z hlediska praktických výsledků.

---

## Seznam použité literatury

- [1] M. Aoki, *Vvedenije v metody optimizacii*, Nauka, Moskva 1977.
- [2] B.P. Demidovič, I.A. Maron, *Osnovy vyčísleitelnoj matematiky*, Nauka, 3. vydání, Moskva 1966.
- [3] O. Došlý, *Základy konvexní analýzy a optimalizace v  $\mathbb{R}^n$* , Skriptum MU, Brno 2005.
- [4] D.M. Himmelblau, *Prikladnoje nelinejnoje programirovanije*, Mir, Moskva 1975.
- [5] P. Hudzovič, *Optimalizácia*, Slovenská technická univerzita, Bratislava 2001.
- [6] B.G. Karmanov, *Matematičeskoje programirovanije*, Nauka, Moskva 1986.
- [7] M. Mañas, *Optimalizační metody*, Nakladatelství technické literatury, Praha 1979.
- [8] A. Ralston *Základy numerické matematiky*, Academia, Praha 1973.
- [9] A.G. Sucharev, A.V. Timochov, V.V. Fedorov, *Kurs metodov optimizacii*, Nauka, Moskva 1986.

---

# Rejstřík pojmů

- A-sdruženost, 49
- aproximace účelové funkce, 24
- báze prostoru, 49
- Cauchy-Schwarcova nerovnost, 41
- Fibonacciho posloupnost, 11
- gradient, 34
- Gram-Schmidův ortogonalizační proces, 50
- Hesseho matice, 38
- interval lokalizace minima, 7
- iterace, 6
- jádro zobrazení, 55
- kernel, 55
- koeficient drobení kroku, 46
- koeficient délky kroku, 38
- koeficient přesnosti, 6
- Lagrangeova věta o střední hodnotě, 47
- lineární obal, 61
- Lipschitzova podmínka, 47
- metoda
  - aktivní, 7
  - Fibonacciho, 11
  - gradientní, 20
  - gradientní (vícerozměrná), 34
  - Hooka a Jeevese, 31
  - interpolační, 24
  - komparativní, 8
  - komparativní (vícerozměrná), 31
  - kvadratické interpolace, 24
  - nejrychlejšího spádu, 35
    - pro kvadratické funkce, 37
    - s drobením kroku, 46
  - newtonova, 63
  - newtonovská, 63
  - paralelních tečen, 44
  - pasivní, 7
  - proměnné metriky, 68
  - sdužených směrů, 49
    - nultého řádu, 60
    - pro kvadratické funkce, 49
    - pro nekvadratické funkce, 57
  - tečen, 20
  - vícerozměrné optimalizace, 30
  - zlatého řezu, 17
- minimalizující posloupnost, 5
- nezávislost vektorů, 50
- opěrný bod, 32
- pozitivní semidefinitnost, 55
- požadavek přesnosti výpočtu, 6
- silná kovexitá, 64
- soustava lineárních rovnic, 37
- souřadnicové hledání, 31
- stacionární bod, 48
- Taylorův mnohočlen, 38
- unimodálnost, 7
- účelová funkce, 5
- vlastní hodnoty matice, 40
- zlatý řez, 17

---

# Rejstřík proměnných

(Seřazeny podle prvního použití v práci, udávané s číslem stránky)

$x_0, \mathbf{x}_0$	- výchozí bod, počáteční aproximace, (7)
$x_k, \mathbf{x}_k$	- bod minimalizující posloupnosti (zpravidla nalezený po $k$ -tém kroku), (7)
$\alpha_k$	- koeficient udávající délku kroku $k + 1$ , (7)
$s_k, \mathbf{s}_k$	- směr jednorozměrné minimalizace pro krok $k + 1$ , (7)
$x^*, \mathbf{x}^*$	- skutečný bod minima účelové funkce, (7)
$\Delta$	- koeficient přesnosti, (8)
$a, b$	- hraniční body intervalu lokalizace minima, (10)
$F_n$	- člen Fibonacciho posloupnosti ležící na místě $n + 1$ , (13)
$y, z$	- dělicí body intervalu lokalizace minima, (14)
$\lambda_k$	- podinterval původního intervalu tvořený při výpočtu Fibonacciho metodou, (14)
$\hat{\tau}$	- poměr zlatého řezu, (19)
$c, d$	- dělicí body intervalu lokalizace minima pro metodu kvadratické interpolace, (26, 27)
$F(x), F(\mathbf{x})$	- aproximace účelové funkce, (26, 66)
$\mathbf{d}$	- vektor délek kroků při jednorozměrných optimalizacích, (33)
$\bar{\mathbf{x}}$	- testovaný bod při souřadnicovém vyhledávání nebo hledání ve směru opěrného bodu, (33)
$\hat{\mathbf{x}}$	- doposud nalezené minimum souřadnicovým vyhledáváním, (33)
$\mathbf{A}$	- matice kvadratické účelové funkce, (39)
$\mathbf{b}$	- lineární člen v kvadratické účelové funkci, (39)
$\mathbf{H}$	- Hesseho matice, matice druhých derivací, (40)
$\lambda_n$	- vlastní hodnoty matice $\mathbf{A}$ , (42)
$\vartheta$	- koeficient drobení kroku, (48)
$\mathbf{r}$	- antigradient, záporná hodnota gradientu, (52)
$\beta_k$	- člen zajišťující $\mathbf{A}$ -sdruženost vektorů směru jednorozměrné optimalizace, (52)
$\mathbf{y}_{k,i}$	- pomocný bod nalezený jednorozměrnou optimalizací, (63)
$\mathbf{z}_k$	- nově nalezené směry jednorozměrných minimalizací v metodě sdružených směrů nultého řádu, (63)
$\mathbf{Z}_k$	- aproximace Hesseho matice, (71)