



národní  
úložiště  
šedé  
literatury

## **Matematické programování**

Lukšan, Ladislav  
2008

Dostupný z <http://www.nusl.cz/ntk/nusl-39678>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 02.06.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní [nusl.cz](http://www.nusl.cz) .



**Institute of Computer Science**  
**Academy of Sciences of the Czech Republic**

## **Matematické programování**

Ladislav Lukšan

Technical report No. 1043

Prosinec 2008



## **Matematické programování**

Ladislav Lukšan <sup>1</sup>

Technical report No. 1043

Prosinec 2008

Abstrakt:

Tato zpráva obsahuje učební text pro předmět matematické programování na fakultě mechatroniky Technické university v Liberci.

Keywords:

Matematické programování, numerická optimalizace, nelineární aproximace, systémy nelineárních rovnic, algoritmy.

---

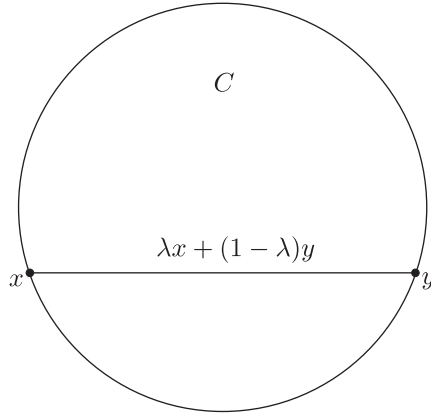
<sup>1</sup>This work was supported by the Grant Agency of the Czech Academy of Sciences, project code IAA1030405. L.Lukšan is also from Technical University of Liberec, Hálkova 6, 461 17 Liberec.

# Konvexní množiny

**Definice 1** Řekněme, že množina  $C \in R^n$  je konvexní, jestliže z  $x \in C$ ,  $y \in C$  plyne

$$\lambda x + (1 - \lambda)y \in C,$$

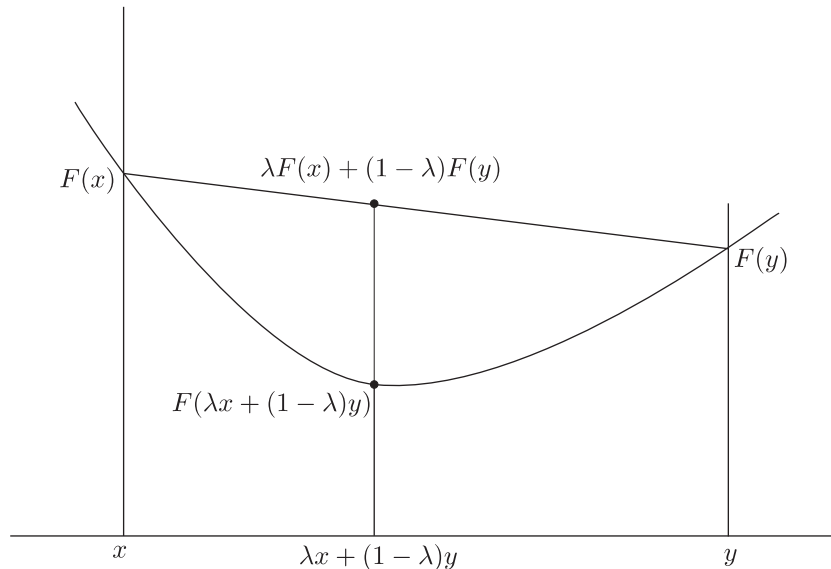
pokud  $0 \leq \lambda \leq 1$ .



**Definice 2** Řekněme, že funkce  $F : R^n \rightarrow R$  je konvexní na konvexní množině  $C \subset R^n$ , jestliže platí

$$F(\lambda x + (1 - \lambda)y) \leq \lambda F(x) + (1 - \lambda)F(y),$$

pokud  $x \in C$ ,  $y \in C$  a  $0 \leq \lambda \leq 1$ .



**Definice 3** Řekněme, že funkce  $F : R^n \rightarrow R$  je lipschitzovská na množině  $C \subset R^n$ , jestliže platí

$$|F(y) - F(x)| \leq L\|y - x\|,$$

pokud  $x \in C$ ,  $y \in C$ .

**Poznámka 1** Vztahy z definic 1 a 2 můžeme zapsat ve tvaru

$$y + \lambda(x - y) \in C \quad \text{a} \quad F(y + \lambda(x - y)) \leq F(y) + \lambda(F(x) - F(y)),$$
$$x + \lambda(y - x) \in C \quad \text{a} \quad F(x + \lambda(y - x)) \leq F(x) + \lambda(F(y) - F(x)).$$

**Příklad 1**

$$\begin{aligned} F(x) &= \sum_i |x_i| \\ F(\lambda x + (1 - \lambda)y) &= \sum_i |\lambda x_i + (1 - \lambda)y_i| = \\ &\leq \sum_i |\lambda x_i| + \sum_i |(1 - \lambda)y_i| = \\ &= \lambda \sum_i |x_i| + (1 - \lambda) \sum_i |y_i| = \\ &= \lambda F(x) + (1 - \lambda)F(y) \end{aligned}$$

**Příklad 2**

$$\begin{aligned} F(x) &= \max_i |x_i| \\ F(\lambda x + (1 - \lambda)y) &= \max_i |\lambda x_i + (1 - \lambda)y_i| \\ &\leq \max_i (|\lambda x_i| + |(1 - \lambda)y_i|) \leq \\ &\leq \max_i |\lambda x_i| + \max_i |(1 - \lambda)y_i| = \\ &= \lambda \max_i |x_i| + (1 - \lambda) \max_i |y_i| = \\ &= \lambda F(x) + (1 - \lambda)F(y) \end{aligned}$$

**Příklad 3**

$$C = \{x \in R_n : F(x) \leq a\}$$

$F$  – konvexní  $\Rightarrow$   $C$  – konvexní

$$\begin{aligned} F(x) &\leq a \\ F(y) &\leq a \\ F(\lambda x + (1 - \lambda)y) &\leq \lambda F(x) + (1 - \lambda)F(y) \leq \lambda a + (1 - \lambda)a = a \end{aligned}$$

**Příklad 4** Množiny

$$C_1 = \{x \in R_n : \sum_i |x_i| \leq 1\}$$
$$C_2 = \{x \in R_n : \max_i |x_i| \leq 1\}$$

jsou konvexní

**Definice 4** Necht  $m \geq 1$ ,  $x_i \in R^n$ ,  $\lambda_i \geq 0$ ,  $1 \leq i \leq m$ ,  $\lambda_1 + \dots + \lambda_m = 1$ . Pak bod

$$x = \sum_{i=1}^m \lambda_i x_i,$$

nazveme konvexní kombinací bodů  $x_i \in R^n$ ,  $1 \leq i \leq m$ .

**Věta 1** Množina  $C \subset R^n$  je konvexní právě tehdy, obsahuje-li všechny konvexní kombinace svých bodů.

**Poznámka 2** Konvexní kombinace konvexních kombinací je opět konvexní kombinací.

**Poznámka 3** Necht  $x_i \in R^n$ ,  $1 \leq i \leq m$ , a  $x = \sum_{i=1}^m \lambda_i x_i$ . Pak bod  $x$  nazveme:

- (a) Lineární kombinací bodů  $x_i \in R^n$ , jsou-li koeficienty  $\lambda_i \in R$  libovolné.
- (b) Nezápornou lineární kombinací bodů  $x_i \in R^n$ , platí-li  $\lambda_i \geq 0$ ,  $1 \leq i \leq m$ .
- (c) Afinní kombinací bodů  $x_i \in R^n$ , platí-li  $\sum_{i=1}^m \lambda_i = 1$ .
- (d) Konvexní kombinací bodů  $x_i \in R^n$ , platí-li  $\sum_{i=1}^m \lambda_i = 1$  a  $\lambda_i \geq 0$ ,  $1 \leq i \leq m$ .

Tyto kombinace definují po řadě (a) lineární podprostory, (b) konvexní kužely, (c) afinní množiny a (d) konvexní množiny. Všechny uvedené množiny jsou konvexní. Afinní množina je posunutým lineárním podprostorem. Lze tedy definovat dimenzi afinní množiny jako dimenzi odpovídajícího lineárního podprostoru a jelikož konvexní množinu lze vnořit do afinní množiny (vynecháním omezení  $\lambda_i \geq 0$ ,  $1 \leq i \leq m$ ) i dimenzi konvexní množiny.

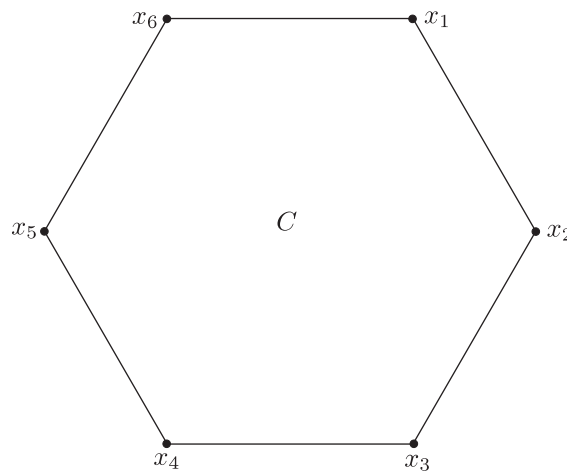
**Věta 2** Průnik konvexních množin je konvexní množinou.

**Věta 3** Lineární kombinace konvexních množin je konvexní množinou.

**Definice 5** Konvexním obalem množiny  $C \subset R^n$  nazveme průnik

$$\text{conv } C = \bigcap_{\alpha} C_{\alpha}$$

všech konvexních množin  $C_{\alpha} \subset R^n$  obsahujících  $C$ .



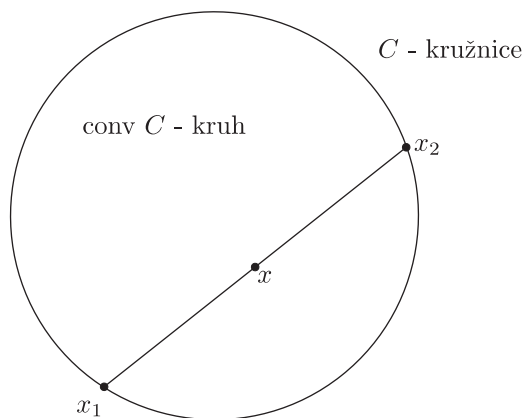
$$C = \text{conv} \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

**Poznámka 4** Zřejmě platí  $C \subset \text{conv } C$ .

**Věta 4** *Konvexní obal množiny  $C \subset R^n$  je množina všech konvexních kombinací bodů z  $C$ , tedy všech bodů tvaru*

$$y = \sum_{i=1}^m \lambda_i x_i,$$

kde  $m \geq 1$ ,  $x_i \in C$ ,  $\lambda_i \geq 0$ ,  $1 \leq i \leq m$ ,  $\lambda_1 + \dots + \lambda_m = 1$ .



**Věta 5** (Caratheodory) *Nechť  $y \in \text{conv } C$ , kde  $C \subset R^n$ . Pak existuje  $n + 1$  bodů  $x_i \in C$ ,  $1 \leq i \leq n + 1$ , takových, že  $y$  je jejich konvexní kombinací.*

**Věta 6** *Je-li množina  $C$  uzavřená, je i množina  $\text{conv } C$  uzavřená.*

**Poznámka 5** *Z věty 4 plyne, že je-li množina  $C$  omezená, je i množina  $\text{conv } C$  omezená. To spolu s větou 6 ukazuje, že je-li množina  $C$  kompaktní, je i množina  $\text{conv } C$  kompaktní.*

**Definice 6** *Nechť  $C \subset R^n$ . Pak funkci*

$$d_C(x) = \inf_{y \in C} \|y - x\|$$

*nazveme vzdáleností bodu  $x$  od množiny  $C$  (nebo vzdálenostní funkcí množiny  $C$ ).*

**Poznámka 6** *Je-li množina  $C \subset R^n$  uzavřená, platí*

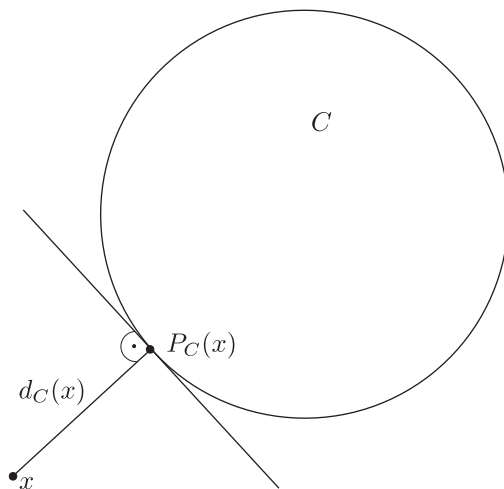
$$d_C(x) = \min_{y \in C} \|y - x\|.$$

V dalším výkladu se omezíme na uzavřené množiny i když většina tvrzení má obecnější charakter.

**Věta 7** *Nechť množina  $C \subset R^n$  je uzavřená. Pak vzdálenostní funkce  $d_C$  je lipschitzovská v  $R^n$  s koeficientem  $L = 1$ . Je-li  $C$  konvexní, je  $d_C$  konvexní v  $R^n$  a ke každému bodu  $x \in R^n$  existuje právě jeden bod  $y \in C$  takový, že*

$$\|y - x\| = d_C(x).$$

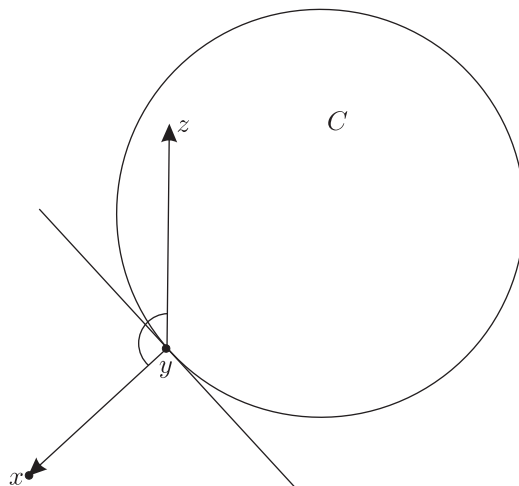
**Definice 7** Necht  $C \subset \mathbb{R}^n$  je uzavřená konvexní množina,  $x \in \mathbb{R}^n$  a  $y \in C$  je bod takový, že  $\|y - x\| = d_C(x)$ . Pak řekneme, že  $y$  je projekcí bodu  $x$  do množiny  $C$  a píšeme  $y = P_C(x)$ .



**Věta 8** Necht  $C \subset \mathbb{R}^n$  je uzavřená konvexní množina a  $x \notin C$ . Pak bod  $y = P_C(x)$  je hraničním bodem množiny  $C$ .

**Lemma 1** Necht  $C \subset \mathbb{R}^n$  je uzavřená konvexní množina,  $x \in \mathbb{R}^n$  a  $y = P_C(x)$ . Pak platí

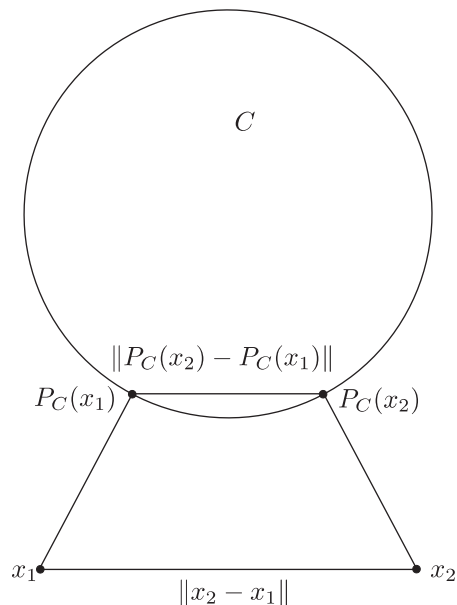
$$(x - y)(z - y) \leq 0 \quad \forall z \in C$$



**Věta 9** Necht  $C \subset \mathbb{R}^n$  je uzavřená konvexní množina. Pak

$$\|P_C(x_2) - P_C(x_1)\| \leq \|x_2 - x_1\| \quad \forall x_1, x_2 \in \mathbb{R}^n.$$





**Definice 8** Necht  $a \in R^n$  a  $\alpha \in R$ . Pak množinu

$$H(a, \alpha) = \{y \in R^n : a^T y \leq \alpha\}$$

nazveme poloprostorem určeným normálovým vektorem  $a$  a číslem  $\alpha$ .

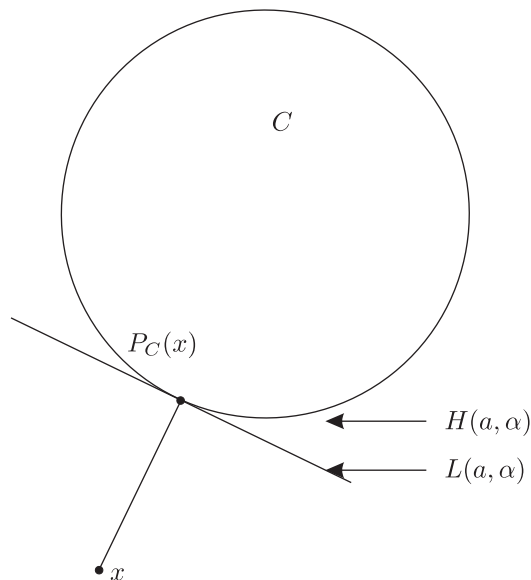
**Poznámka 7** Hranicí poloprostoru  $H(a, \alpha)$  je nadrovina

$$L(a, \alpha) = H(a, \alpha) \cap H(-a, \alpha) = \{y \in R^n : a^T y = \alpha\}.$$

Číslo  $\alpha$  určuje vzdálenost nadroviny  $L(a, \alpha)$  od počátku. Tato vzdálenost se rovná podílu  $\alpha/\|a\|$ . Odtud plyne, že bod  $y = 0$  je hraničním bodem poloprostoru  $H(a, \alpha)$  (leží v hraniční nadrovině  $L(a, \alpha)$ ) právě tehdy, když  $\alpha = 0$ .

**Věta 10** Poloprostor  $H(a, \alpha)$  je uzavřenou konvexní množinou.

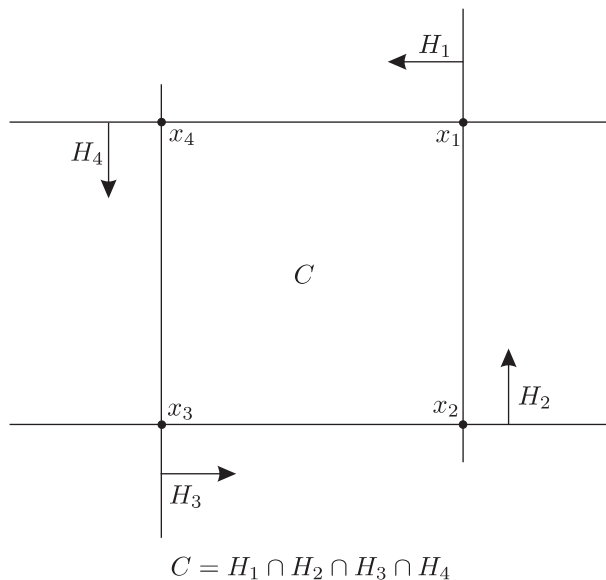
**Věta 11** Necht  $C$  je uzavřená konvexní množina a necht  $x \notin C$ . Pak existuje poloprostor  $H(a, \alpha)$  takový, že  $C \subset H(a, \alpha)$  a  $x \notin H(a, \alpha)$ . Tento poloprostor lze volit tak, že platí  $a = x - P_C(x)$  a  $\alpha = (x - P_C(x))^T P_C(x)$ . Pak  $P_C(x) \in L(a, \alpha)$ , takže  $C \cap L(a, \alpha) \neq \emptyset$ .



**Důsledek 1** Necht  $C_1, C_2$  jsou uzavřené konvexní množiny takové, že  $C_1 \cap C_2 = \emptyset$ . Pak existuje poloprostor  $H(a, \alpha)$  takový, že  $C_1 \subset H(a, \alpha)$  a  $C_2 \cap H(a, \alpha) = \emptyset$

**Věta 12** Uzavřená konvexní množina  $C \subset \mathbb{R}^n$  je průnikem všech poloprostorů obsahujících  $C$ .

**Definice 9** Konvexní množina, která je průnikem konečného počtu poloprostorů, se nazývá polyedrál ní množinou.



**Definice 10** Necht  $C$  je uzavřená konvexní množina a  $H(a, \alpha)$  je poloprostor s hranicí  $L(a, \alpha)$  takový, že  $C \subset H(a, \alpha)$  a  $C \cap L(a, \alpha) \neq \emptyset$ . Pak řekneme, že  $H(a, \alpha)$  je tečným poloprostorem a  $L(a, \alpha)$  tečnou nadrovinou množiny  $C$ .

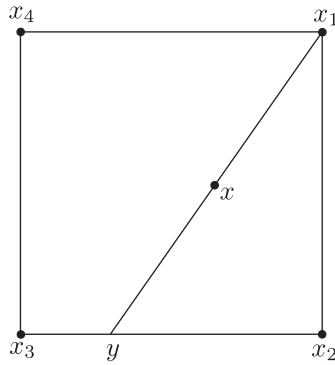
**Poznámka 8** Ve Větě 12 se můžeme omezit na tečné poloprostory (uzavřená konvexní množina je průnikem svých tečných poloprostorů). Obsahuje-li poloprostor  $H(a, \alpha)$  konvexní množinu  $C$ , přičemž  $C \cap L(a, \alpha) = \emptyset$ , lze volbou  $\alpha' = \max_{y \in C} a^T y$  docílit toho, že  $C \subset H(a, \alpha') \subset H(a, \alpha)$  a  $C \cap L(a, \alpha') \neq \emptyset$ .

**Věta 13** Necht bod  $y \in R^n$  je hraničním bodem uzavřené konvexní množiny  $C$ . Pak existuje tečná nadrovina  $L(a, \alpha)$  taková, že  $y \in L(a, \alpha)$ .

**Definice 11** Necht  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Není-li bod  $x$  konvexní kombinací žádných bodů z  $C$  různých od  $x$ , řekneme, že  $x$  je krajním bodem nebo vrcholem množiny  $C$ .

**Poznámka 9** V definici krajních bodů se můžeme omezit na konvexní kombinace dvou bodů z  $C$  různých od  $x$ . Dále se můžeme omezit na průměry dvou bodů z  $C$  různých od  $x$ . Necht  $x = \lambda_1 x_1 + \lambda_2 x_2$ ,  $\lambda_1 + \lambda_2 = 1$ ,  $\lambda_1 \geq \lambda_2 \geq 0$ . Položíme-li  $x_3 = \lambda'_1 x_1 + \lambda'_2 x_2$ , kde  $\lambda'_1 = 2\lambda_1 - 1$ ,  $\lambda'_2 = 2\lambda_2$ , takže  $\lambda'_1 + \lambda'_2 = 1$ ,  $\lambda'_1 \geq 0$ ,  $\lambda'_2 \geq 0$ , platí  $x_3 \in C$  a  $x = (x_1 + x_3)/2$ .

**Věta 14** Kompaktní konvexní množina je konvexním obalem svých krajních bodů.



$$x = \lambda_1 x_1 + (1 - \lambda_1)y \quad y = \lambda_2 x_2 + (1 - \lambda_2)x_3$$

$$x = \lambda'_1 x_1 + \lambda'_2 x_2 + \lambda'_3 x_3$$

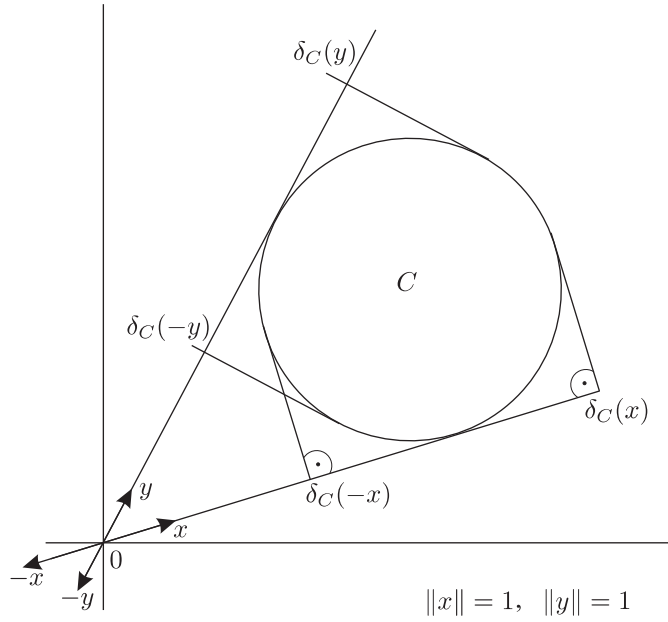
$$\lambda'_1 = \lambda_1 \geq 0, \quad \lambda'_2 = (1 - \lambda_1)\lambda_2 \geq 0, \quad \lambda'_3 = (1 - \lambda_1)(1 - \lambda_2) \geq 0$$

$$\lambda'_1 + \lambda'_2 + \lambda'_3 = 1$$

**Definice 12** Necht  $C \subset R^n$ . Pak funkci

$$\delta_C(x) = \sup_{y \in C} y^T x$$

nazveme opěrnou funkcí množiny  $C$ .



**Poznámka 10** Necht množina  $C \subset R^n$  je kompaktní. Pak platí

$$\delta_C(x) = \max_{y \in C} y^T x.$$

V dalším výkladu se omezíme na kompaktní množiny i když většina tvrzení má obecnější charakter.

**Věta 15** Necht množina  $C \subset R^n$  je kompaktní. Pak opěrná funkce  $\delta_C$  je pozitivně homogenní, subaditivní a lipschitzovská v  $R^n$ .

**Věta 16** Necht množina  $C \subset R^n$  je kompaktní. Pak

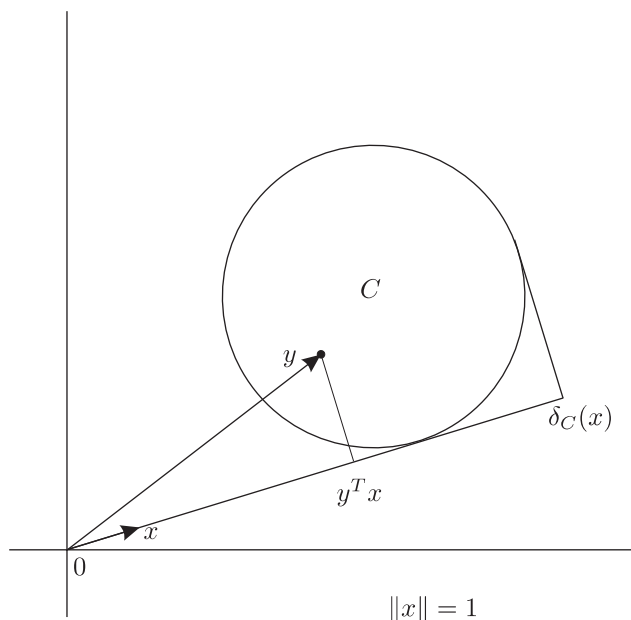
$$\delta_C(x) = \delta_{\text{conv } C}(x) \quad \forall x \in R^n.$$

**Věta 17** Necht množiny  $C_1 \subset R^n$ ,  $C_2 \subset R^n$  jsou konvexní a kompaktní. Pak  $C_1 \subset C_2$  platí právě tehdy, jestliže

$$\delta_{C_1}(x) \leq \delta_{C_2}(x) \quad \forall x \in R^n.$$

**Důsledek 2** Necht množina  $C \subset R^n$  je konvexní a kompaktní. Pak  $y \in C$  právě tehdy, jestliže

$$y^T x \leq \delta_C(x) \quad \forall x \in R^n.$$



**Věta 18** *Nechť množiny  $C_1 \subset R^n$ ,  $C_2 \subset R^n$  jsou kompaktní. Pak*

$$\delta_{C_1+C_2}(x) = \delta_{C_1}(x) + \delta_{C_2}(x).$$

Opěrná funkce množiny  $C \subset R^n$  má bezprostřední vztah k poloprostorům obsahujícím tuto množinu.

**Věta 19** *Množina  $C \subset R^n$  leží v poloprostoru  $H(a, \alpha)$  právě tehdy jestliže  $\alpha \geq \delta_C(a)$ , přičemž  $H(a, \alpha)$  je tečným poloprostorem množiny  $C$  právě tehdy, jestliže  $\alpha = \delta_C(a)$ .*

**Definice 13** *Řekneme, že množina  $K \subset R^n$  je kuželem, jestliže z  $x \in K$  a  $\lambda \geq 0$  plyne  $\lambda x \in K$ .*

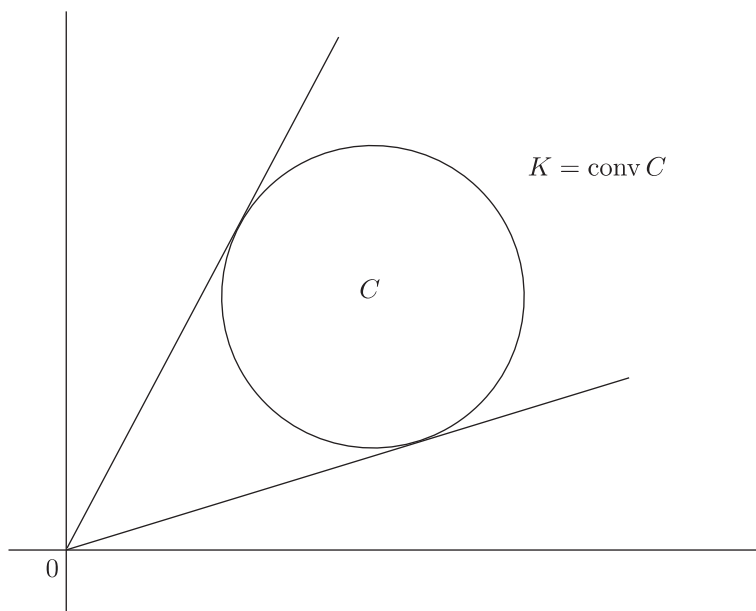
**Věta 20** *Průnik kuželů je kuželem.*

**Věta 21** *Lineární kombinace kuželů je kuželem.*

**Definice 14** *Kuželovým obalem množiny  $C \subset R^n$  nazveme průnik*

$$\text{cone } C = \bigcap_{\alpha} K_{\alpha}$$

*všech kuželů  $K_{\alpha} \subset R^n$  obsahujících  $C$ .*



**Věta 22** *Nechť  $C \subset \mathbb{R}^n$ . Pak platí*

$$\text{cone } C = \bigcup_{\lambda \geq 0} \lambda C = \{x \in \mathbb{R}^n : x = \lambda y, y \in C, \lambda \geq 0\}$$

*Kužel  $\text{cone } C$  je tedy množinou všech nezáporných násobků bodů z  $C$ .*

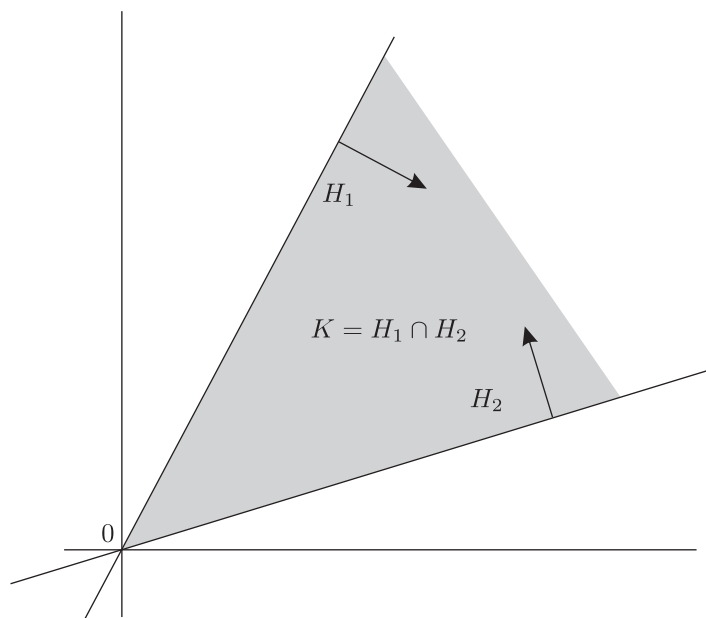
**Věta 23** *Množina  $K \subset \mathbb{R}^n$  je konvexním kuželem právě tehdy, obsahuje-li všechny nezáporné lineární kombinace svých bodů.*

**Důsledek 3** *Množina  $K \subset \mathbb{R}^n$  je konvexním kuželem právě tehdy, obsahuje-li nezáporné násobky a součty svých bodů.*

**Věta 24** *Množina  $\text{cone}(\text{conv } C)$  je množinou všech nezáporných lineárních kombinací bodů z  $C$ .*

Jelikož uzavřený konvexní kužel je uzavřenou konvexní množinou, můžeme studovat tečné poloprostory uzavřených konvexních kuželů.

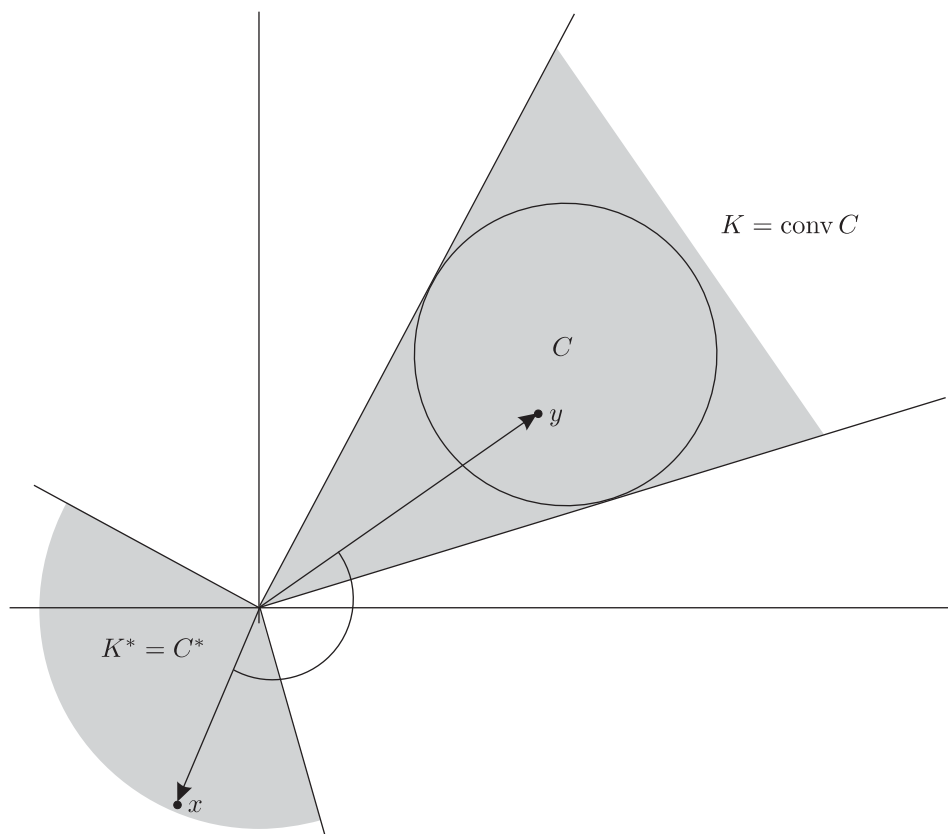
**Věta 25** *Tečný poloprostor uzavřeného konvexního kuželu je uzavřeným konvexním kuželem (takže obsahuje počátek souřadnic). Uzavřený konvexní kužel je průnikem svých tečných poloprostorů.*



**Definice 15** Necht  $C \in R^n$ . Množinu

$$C^* = \{x \in R^n : y^T x \leq 0 \quad \forall y \in C\}$$

nazveme *polárním kuželem množiny C*.



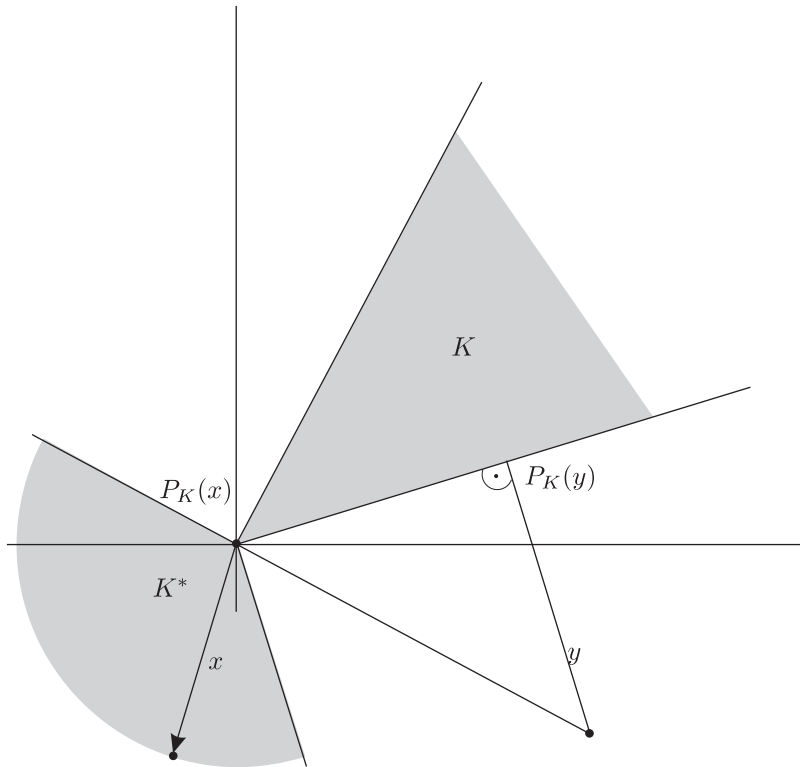
**Poznámka 11** Z definice 15 lze snadno usoudit, že z  $C_1 \subset C_2$  plyne  $C_2^* \subset C_1^*$ .

**Věta 26** *Nechť  $C \subset R^n$ . Pak množina  $C^*$  je uzavřeným konvexním kuželem.*

**Věta 27** *Je-li  $K \subset R^n$  uzavřeným konvexním kuželem, platí  $(K^*)^* = K$ .*

**Věta 28** *Nechť  $K \subset R^n$  je uzavřený konvexní kužel. Pak*

$$K^* = \{x \in R^n : P_K(x) = 0\}.$$



**Věta 29** *Nechť  $K \subset R^n$  je uzavřený konvexní kužel. Pak  $K^*$  je sjednocením normálových vektorů tečných poloprostorů kuželu  $K$ , neboli*

$$K^* = \bigcup_{K \subset H(a,0)} a.$$

**Definice 16** *Kužel  $K \subset R^n$ , který je průnikem konečného počtu tečných poloprostorů, se nazývá polyedrálním kuželem*

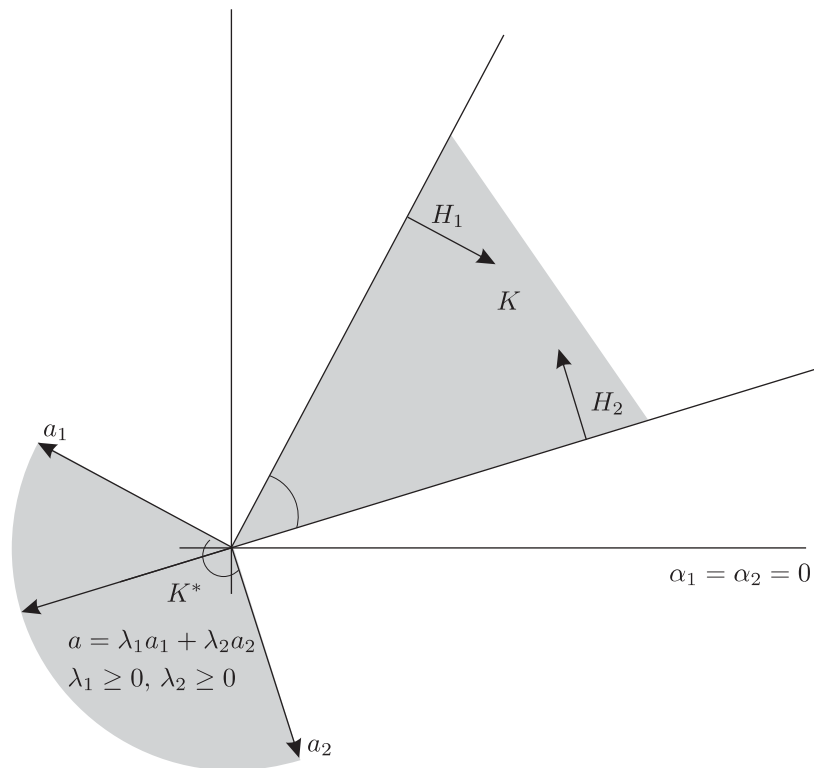
**Věta 30** *Nechť  $K \in R^n$  je polyedrální kužel takový, že*

$$K = \bigcap_{i=1}^m H(a_i, 0)$$

*Pak*

$$K^* = \text{cone}(\text{conv}\{a_i : 1 \leq i \leq m\}).$$





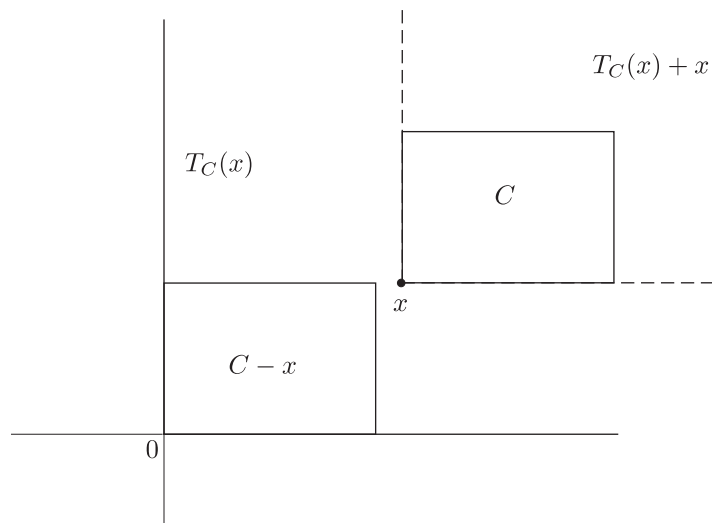
**Definice 17** Necht  $C \subset R^n$  je uzavřená množina a  $x \in C$ . Tečným kuželem množiny  $C$  v bodě  $x$  nazveme množinu

$$T_C(x) = \{y \in R^n : \text{existují posloupnosti } y_i \rightarrow y, t_i \downarrow 0 \text{ takové, že } x + t_i y_i \in C\}$$

**Věta 31** Necht  $C \subset R^n$  je uzavřená množina a  $x \in C$ . Pak  $T_C(x)$  je uzavřeným kuželem. Je-li  $C$  konvexní, je i  $T_C(x)$  konvexní.

**Věta 32** Necht  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Pak

$$T_C(x) = \overline{\bigcup_{\lambda \geq 0} \lambda(C - x)}$$



**Věta 33** Necht  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$  je jejím hraničním bodem. Pak  $T_C(x)$  je půnikem všech tečných poloprostorů množiny  $C - x$  obsahujících počátek souřadnic, neboli

$$T_C(x) = \bigcap_{C-x \subset H(a,0)} H(a,0).$$

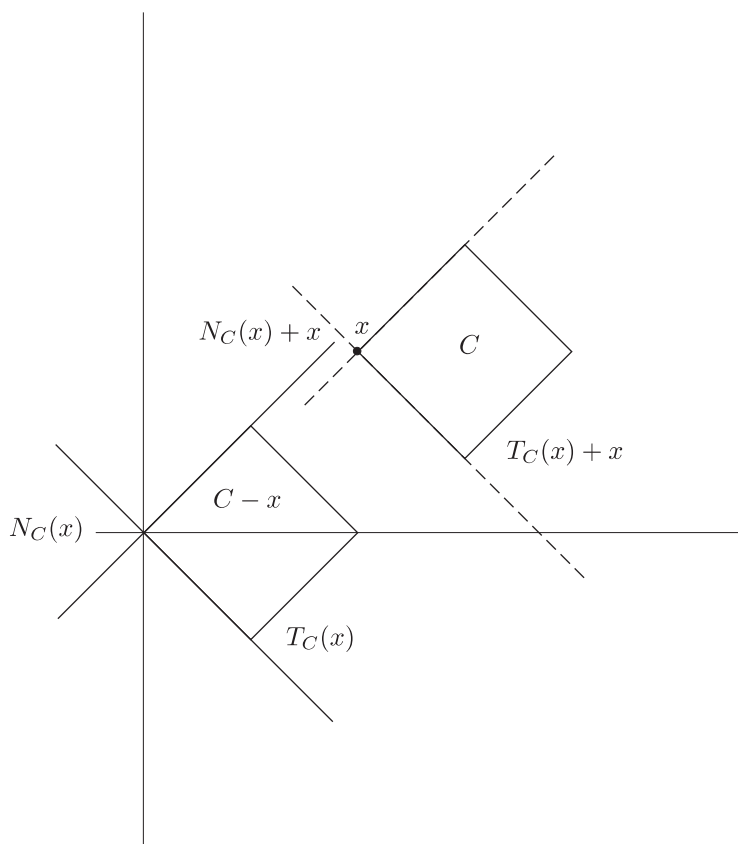
Je-li množina  $C \in R^n$  polyedrál ní, je i tečný kužel  $T_C(x)$  polyedrál ní a existují tečné poloprostory  $H(a_i, 0)$ ,  $1 \leq i \leq m$ , takové, že

$$T_C(x) = \bigcap_{i=1}^m H(a_i, 0).$$

**Definice 18** Necht  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Normálovým kuželem množiny  $C$  v bodě  $x$  nazveme množinu

$$N_C(x) = T_C^*(x),$$

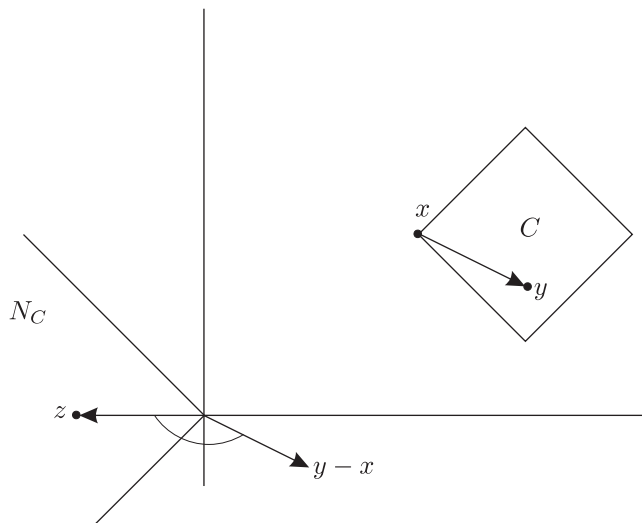
kde  $T_C^*(x)$  je polární kužel tečného kuželu  $T_C(x)$ .



**Poznámka 12** Podle Věty 26 je množina  $N_C(x)$  uzavřeným konvexním kuželem.

**Věta 34** Necht  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Pak

$$N_C(x) = \{z \in R^n : (y - x)^T z \leq 0 \quad \forall y \in C\}.$$



**Věta 35** *Nechť  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$  je jejím hraničním bodem. Pak  $N_C(x)$  je sjednocením normálových vektorů tečných poloprostorů množiny  $C - x$  obsahujících počátek souřadnic, neboli*

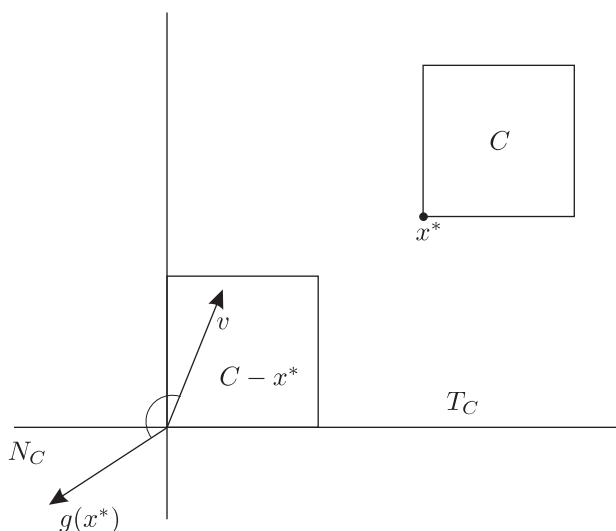
$$N_C(x) = \bigcup_{C-x \subset H(a,0)} a.$$

*Je-li množina  $C \subset R^n$  polyedrání, je i normálový kužel  $T_C(x)$  polyedrání a existují tečné poloprostory  $H(a_i, 0)$ ,  $1 \leq i \leq m$ , takové, že*

$$N_C(x) = \text{cone}(\text{conv}\{a_i : 1 \leq i \leq m\}).$$

**Věta 36** *Nechť  $F \in C^1 : R^n \rightarrow R$  (spojitě diferencovaná)  $C \subset R^n$  a  $x^* \in C$ . Pak je-li  $x^*$  lokálním minimem funkce  $F$  na  $C$ , platí*

$$-g(x^*) \in N_C(x^*)$$



**Věta 37** *Nechť jsou splněny předpoklady Věty 36 a*

$$C = \{x \in R^n : c_i(x) \leq 0, i \in I\}$$

kde  $c_i : R^n \rightarrow R, i \in I$ , jsou konvexní funkce (takže  $C \in R^n$  je konvexní množina). Nechť  $x^* \in C$  a  $\bar{I} = \{i \in I : c_i(x^*) = 0\} \subset I$ . Pak

$$T_C(x^*) = \{s \in R^n : a_i^T s \leq 0, i \in \bar{I}\}$$

a

$$N_C(x^*) = \left\{ v \in R^n : v = \sum_{i \in \bar{I}} u_i a_i, u_i \geq 0 \right\} = \text{cone}(\text{conv} \{a_i : i \in \bar{I}\})$$

kde  $a_i = \nabla c_i(x^*)$ .

## Konvexní funkce

**Definice 19** *Řekneme, že funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$ , jestliže existuje číslo  $\varepsilon > 0$  takové, že  $F$  je definovaná a konvexní v  $B(x, \varepsilon) = \{y : \|y - x\| < \varepsilon\}$ , neboli*

$$F(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda F(x_1) + (1 - \lambda)F(x_2),$$

pokud  $x_1 \in B(x, \varepsilon), x_2 \in B(x, \varepsilon)$  a  $0 \leq \lambda \leq 1$ .

**Poznámka 13** Indukcí snadno dokážeme, že z  $x_i \in B(x, \varepsilon), \lambda_i \geq 0$  a  $\sum_{i=1}^m \lambda_i = 1$  plyne

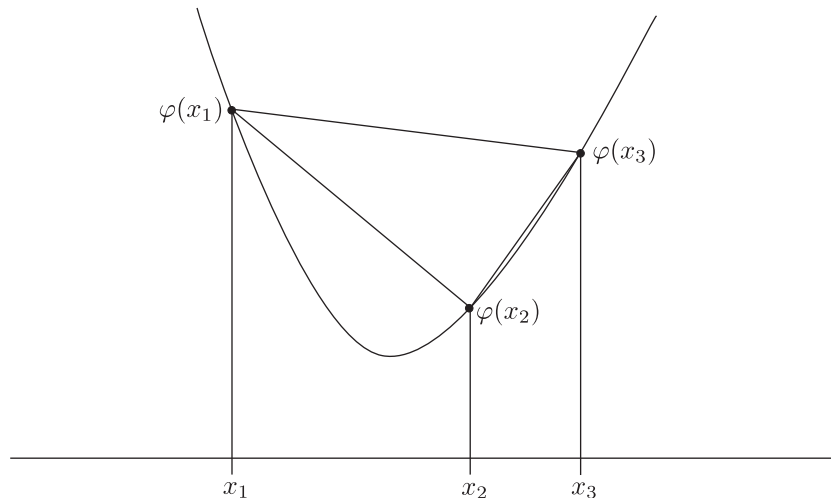
$$F\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i F(x_i),$$

pokud  $F$  je konvexní na  $B(x, \varepsilon)$ .

**Věta 38** *Nechť funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$ . Pak  $F$  je lipschitzovská v okolí bodu  $x$ .*

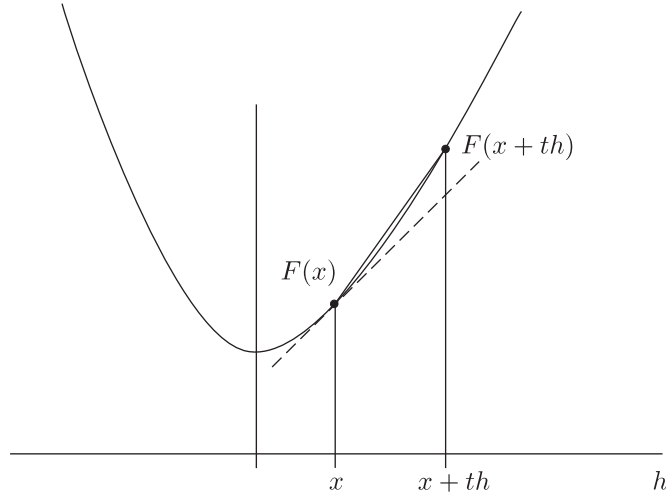
**Lemma 2** *Nechť funkce  $\varphi : R \rightarrow R$  je konvexní na intervalu  $[a, b]$  a nechť  $a \leq t_1 < t_2 < t_3 \leq b$ . Pak platí*

$$\frac{\varphi(t_2) - \varphi(t_1)}{t_2 - t_1} \leq \frac{\varphi(t_3) - \varphi(t_1)}{t_3 - t_1} \leq \frac{\varphi(t_3) - \varphi(t_2)}{t_3 - t_2}.$$



**Definice 20** Řekneme, že funkce  $F : R^n \rightarrow R$  má v bodě  $x \in R^n$  směrovou derivaci ve směru  $h \in R^n$ , existuje-li konečná limita

$$F'(x, h) = \lim_{t \downarrow 0} \frac{F(x + th) - F(x)}{t}.$$



$$F'(x, h) = \lim_{x \downarrow 0} \frac{F(x + th) - F(x)}{x}$$

**Věta 39** Nechť funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$  (takže je Lipschitzovská s nějakou konstantou  $L$  v okolí tohoto bodu). Pak:

(a) Směrová derivace  $F'(x, h)$  existuje pro každé  $h \in R^n$ . Navíc existuje číslo  $\varepsilon > 0$  takové, že

$$F'(x, h) = \inf_{0 < t \|h\| < \varepsilon} \frac{F(x + th) - F(x)}{t}.$$

(b) Funkce  $F'(x, \cdot) : R^n \rightarrow R$  je pozitivně homogenní, subaditivní a Lipschitzovská s konstantou  $L$ .

(c) Funkce  $F'(\cdot, \cdot) : R^n \times R^n \rightarrow R$  je shora polospojité, neboli

$$\limsup_{i \rightarrow \infty} F'(x_i, h_i) \leq F'(x, h),$$

kdykoliv  $x_i \rightarrow x$  a  $h_i \rightarrow h$ .

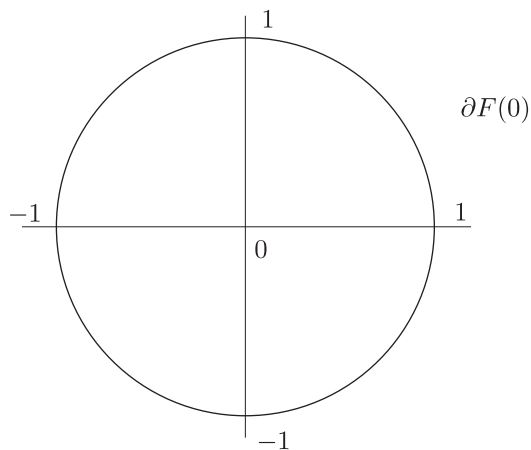
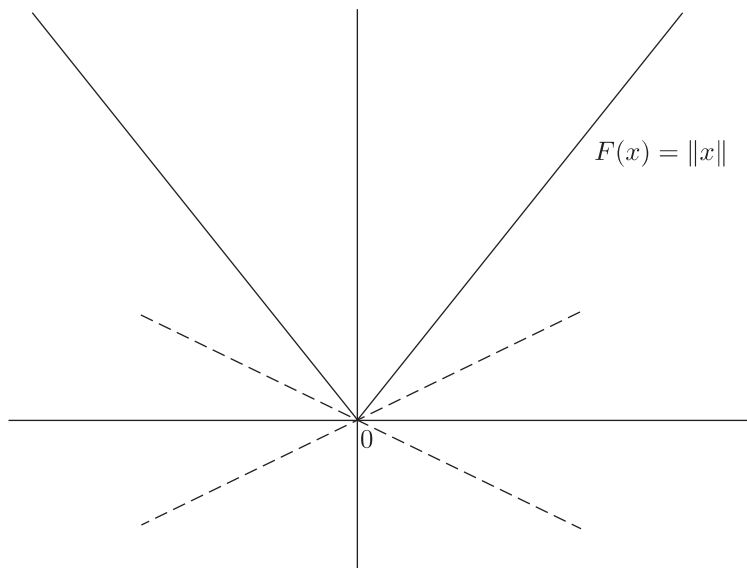
**Poznámka 14** Podle Definice 20 platí  $F'(x, 0) = 0$ , takže podle Věty 39 (b) dostaneme

$$|F'(x, h)| = |F'(x, h) - F'(x, 0)| \leq L \|h\|.$$

**Definice 21** Necht funkce  $F : R^n \rightarrow R$  je konverní v okolí bodu  $x \in R^n$ . Pak množinu

$$\partial F(x) = \{g \in R^n : F'(x, h) \geq g^T h \quad \forall h \in R^n\}$$

nazveme subdiferenciálem funkce  $F$  v bodě  $x$ . Elementy  $g \in \partial F(x)$  budeme nazývat subgradienty funkce  $F$  v bodě  $x$ .



**Věta 40** Necht funkce  $F : R^n \rightarrow R$  je konverní v okolí bodu  $x \in R^n$  (takže je v tomto okolí lipschitzovská s nějakou konstantou  $L$ ). Pak:

- (a) Subdiferenciál  $\partial F(x)$  je neprázdná konverní kompaktní množina taková, že  $\|g\| \leq L \quad \forall g \in \partial F(x)$ .
- (b) Pro libovolný vektor  $h \in R^n$  platí

$$F'(x, h) = \max \{g^T h : g \in \partial F(x)\}.$$

- (c) Jestliže  $x_i \rightarrow x$ ,  $g_i \in \partial F(x_i)$  a  $g_i \rightarrow g$ , pak  $g \in \partial F(x)$  (polospojitost shora).

(d) Existuje číslo  $\varepsilon > 0$  takové, že pro libovolný vektor  $g \in \partial F(x)$  platí

$$F(x+h) - F(x) \geq g^T h \quad \forall h \in B(0, \varepsilon).$$

**Poznámka 15** Podle věty 40 (b) je směrová derivace opěrnou funkcí subdiferenciálu, neboli

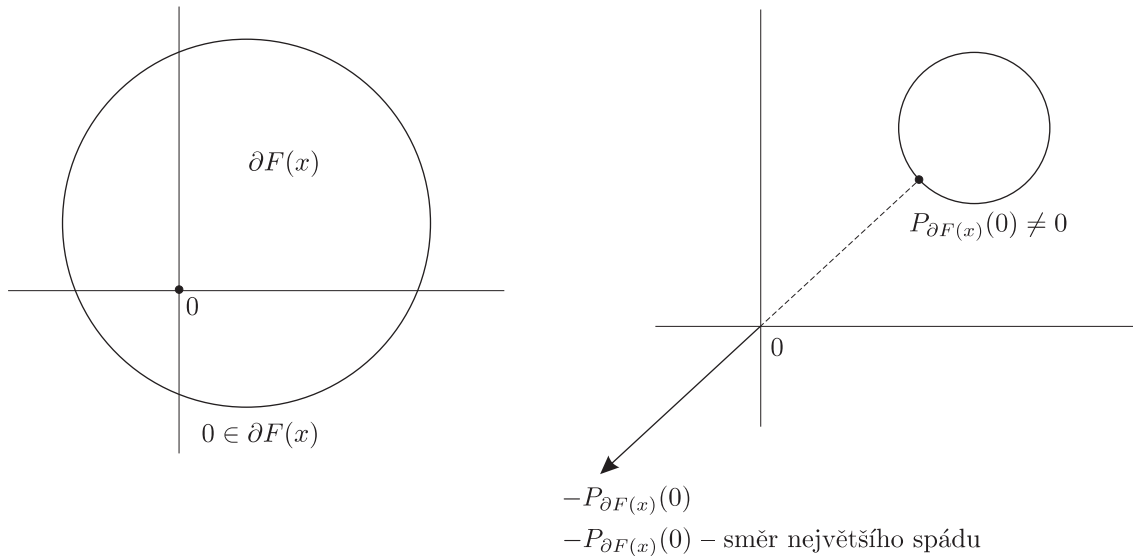
$$F'(x, h) = \delta_{\partial F(x)}(h).$$

**Věta 41** Nechť funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$  a diferencovatelná v bodě  $x \in R^n$ . Pak platí

$$\partial F(x) = \{\nabla F(x)\}.$$

**Věta 42** Nechť funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$ . Pak  $F$  má v bodě  $x$  lokální minimum právě tehdy, jestliže

$$0 \in \partial F(x).$$



**Věta 43** Nechť  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Pak platí

$$T_C(x) = \{y \in R^n : d'_C(x, y) = 0\}$$

( $d'_C(x, y)$  je směrová derivace funkce  $d_C(x)$  ve směru  $y \in R^n$ ).

**Věta 44** Nechť  $C \subset R^n$  je uzavřená konvexní množina a  $x \in C$ . Pak platí

$$N_C(x) = \overline{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)}$$

**Věta 45** Nechť funkce  $F : R^n \rightarrow R$  je dvakrát spojitě diferencovatelná v bodě  $x \in R^n$ . Je-li  $F$  konvexní v okolí bodu  $x \in R^n$ , je její Hessova matice pozitivně semidefinitní v bodě  $x$ . Je-li Hessova matice funkce  $F$  pozitivně definitní v bodě  $x$ , je tato funkce konvexní v okolí bodu  $x$ .

## Nepodmíněná minimalizace

Budeme hledat minimum funkce  $F : R^n \rightarrow R$  na jejím definičním oboru  $\mathcal{D}(F) \subset R^n$  (obvykle  $\mathcal{D}(F) = R^n$ ). Předpokládáme, že  $F \in C^2$  (je dvakrát spojitě diferencovatelná). Označení:

$$g(x) = \begin{bmatrix} \frac{\partial F(x)}{\partial x_1} \\ \dots \\ \frac{\partial F(x)}{\partial x_n} \end{bmatrix}, \quad G(x) = \begin{bmatrix} \frac{\partial^2 F(x)}{\partial x_1^2} & \dots & \frac{\partial^2 F(x)}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 F(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 F(x)}{\partial x_n^2} \end{bmatrix}$$

$g$  – gradient,  $G$  – Hessova matice

**Definice 22**  $F : R^n \rightarrow R$  je zdola omezená, existuje-li  $\underline{F}$  tak, že

$$F(x) \geq \underline{F} \quad \forall x \in R^n \quad (\text{F1})$$

**Definice 23**  $F : R^n \rightarrow R$  má kompaktní hladiny, jestliže

$$\mathcal{L}(\overline{F}) = \{x \in R^n : F(x) \leq \overline{F}\} \quad (\text{F2})$$

je kompaktní  $\forall \overline{F} \in R$  (prázdná množina je kompaktní).

**Definice 24**  $F \in C^2 : R^n \rightarrow R$  má omezené druhé derivace, existuje-li  $\overline{G} > 0$  tak, že  $\|G(x)\| \leq \overline{G} \forall x \in R^n$ , čili

$$|d^T G(x) d| \leq \overline{G} \|d\|^2 \quad \forall x, d \in R^n \quad (\text{F3})$$

**Definice 25**  $F \in C^2 : R^n \rightarrow R$  je stejnoměrně konvexní, existuje-li  $\underline{G} > 0$  tak, že

$$d^T G(x) d \geq \underline{G} \|d\|^2 \quad \forall x, d \in R^n \quad (\text{F4})$$

**Definice 26**  $F \in C^2 : R^n \rightarrow R$  má lipschitzovské druhé derivace, jestliže existuje  $\overline{L} > 0$  tak, že

$$\|G(x+d) - G(x)\| \leq \overline{L} \|d\| \quad \forall x, d \in R^n \quad (\text{F5})$$

**Věta 46** (o střední hodnotě). Nechť  $F \in C^2 : R^n \rightarrow R$  a  $x, d \in R^n$ . Pak platí

$$F(x+d) = F(x) + d^T g(x) + \frac{1}{2} d^T G(\tilde{x}) d$$

kde  $\tilde{x} = x + \tilde{\lambda} d$  a  $0 \leq \tilde{\lambda} \leq 1$ .

**Důsledek 4** Podmínka (F3) implikuje

$$F(x+d) - F(x) \leq d^T g(x) + \frac{1}{2} \overline{G} \|d\|^2$$

Podmínka (F4) implikuje

$$F(x+d) - F(x) \geq d^T g(x) + \frac{1}{2} \underline{G} \|d\|^2$$



**Věta 47** (o střední hodnotě). Nechť  $F \in C^2 : R^n \rightarrow R$  a  $x, d \in R^n$ . Pak platí

$$g(x+d) = g(x) + \int_0^1 G(x+\lambda d) d \lambda$$

**Důsledek 5** Podmínka (F3) implikuje

$$\begin{aligned} \|g(x+d) - g(x)\| &\leq \overline{G} \|d\| \\ d^T (g(x+d) - g(x)) &\leq \overline{G} \|d\|^2 \end{aligned}$$

Podmínka (F4) implikuje

$$\begin{aligned} \|g(x+d) - g(x)\| &\geq \underline{G} \|d\| \\ d^T (g(x+d) - g(x)) &\geq \underline{G} \|d\|^2 \end{aligned}$$

## Podmínky optimality

**Definice 27** Řekneme, že bod  $x^* \in R^n$  je lokálním minimem funkce  $F : R^n \rightarrow R$  jestliže existuje  $\varepsilon > 0$  tak, že

$$F(x^*) \leq F(x) \quad \forall x \in B(x^*, \varepsilon) = \{x \in R^n : \|x - x^*\| < \varepsilon\}$$

Jestliže  $F(x^*) < F(x) \quad \forall x \in B(x^*, \varepsilon) \setminus \{x^*\}$ , řekneme, že bod  $x^*$  je ostrým (izolovaným) lokálním minimem funkce  $F$ .

**Věta 48** (nutné podmínky). Nechť bod  $x^* \in R^n$  je lokálním minimem funkce  $F \in C^2 : R^n \rightarrow R$ . Pak platí

$$\begin{aligned} g(x^*) &= 0 \\ G(x^*) &\succeq 0 \end{aligned}$$

( $G(x^*)$  je pozitivně semidefinitní).

**Věta 49** (postačující podmínky). Nechť  $F \in C^2 : R^n \rightarrow R$  na  $B(x^*, \varepsilon)$  a nechť

$$\begin{aligned} g(x^*) &= 0 \\ G(x^*) &\succ 0 \end{aligned}$$

( $G(x^*)$  je pozitivně definitní). Pak bod  $x^*$  je ostrým (izolovaným) minimem funkce  $F : R^n \rightarrow R$ .

**Poznámka 16** Nechť  $F \in C^2 : R^n \rightarrow R$  na  $B(x^*, \varepsilon)$  a nechť

$$\begin{aligned} g(x^*) &= 0 \\ G(x^*) &\succeq 0 \end{aligned}$$

Je-li  $G(x^*)$  singulární, nelze rozhodnout, je-li  $x^*$  lokálním minimem.

**Příklad 5** Nechť funkce  $F : R \rightarrow R$  je zadaná předpisem  $F(x) = x^3$  a  $x^* = 0$ . Pak

$$\begin{aligned} g(x^*) &= [F'(x^*)] = [3x^*] = [0], \quad \text{pro } x^* = 0 \\ G(x^*) &= [F''(x^*)] = [6x^*] = [0], \quad \text{pro } x^* = 0 \end{aligned}$$

Tedy  $g(x^*) = 0$ ,  $G(x^*) \succeq 0$ , ale  $F$  má v bodě  $x^*$  inflexní bod.

## Konvergence iteračních procesů

**Definice 28** *Nechť  $\{x_i\} \subset R^n$  je posloupnost bodů z  $R^n$ . Jestliže pro libovolné číslo  $\varepsilon > 0$  existuje index  $k(\varepsilon) \in N$  takový, že  $x_i \in B(x^*, \varepsilon) \forall i \geq k(\varepsilon)$ , řekneme, že posloupnost  $\{x_i\}$  konverguje k bodu  $x^* \in R^n$  a píšeme  $x_i \rightarrow x^*$ .*

**Poznámka 17** Jestliže  $\xi_i \rightarrow 0$  (a  $\xi_i \neq 0$ ), je výhodné používat symboly  $o(\xi_i)$  a  $O(\xi_i)$  pro posloupnosti takové, že

$$\|o(\xi_i)\| / \|\xi_i\| \rightarrow 0, \quad \text{a} \quad \|O(\xi_i)\| / \|\xi_i\| \leq C.$$

Pro libovolný exponent  $r \in R$  platí  $(1+o(\xi_i))^r = 1+o(\xi_i)$  a  $(1+O(\xi_i))^r = 1+O(\xi_i)$ , pokud  $o(\xi_i) \rightarrow 0$  a  $O(\xi_i) \rightarrow 0$ . Jestliže  $\xi_i = O(\eta_i)$  a  $\eta_i = O(\xi_i)$ , budeme říkat, že posloupnosti  $\{\xi_i\}$  a  $\{\eta_i\}$  jsou ekvivalentní a budeme psát  $\xi_i \sim \eta_i$ .

**Věta 50** *Nechť  $x_i \rightarrow x^*$  a  $d_i \rightarrow 0$ , kde  $x^* \in R^n$  je stacionární bod funkce  $F \in C^2 : R^n \rightarrow R$  (tj.  $g(x^*) = 0$ ). Označme  $e_i = x_i - x^*$ . Pak platí*

$$\begin{aligned} F(x_i + d_i) - F(x_i) &= d_i^T g_i + \frac{1}{2} d_i^T G_i d_i + o(\|d_i\|^2) = d_i^T g_i + \frac{1}{2} d_i^T G^* d_i + o(\|d_i\|^2) \\ g(x_i + d_i) - g(x_i) &= G_i d_i + o(\|d_i\|) = G^* d_i + o(\|d_i\|) \end{aligned}$$

a

$$\begin{aligned} F(x_i) - F(x^*) &= \frac{1}{2} e_i^T G^* e_i + o(\|e_i\|^2) \\ g(x_i) &= G^* e_i + o(\|e_i\|) \end{aligned}$$

**Poznámka 18** Platí-li (F5), můžeme  $o(\|d_i\|^2)$  a  $o(\|d_i\|)$  nahradit  $O(\|d_i\|^3)$  a  $O(\|d_i\|^2)$ .

**Definice 29** *Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$  (alespoň)  $R$ -lineárně, jestliže*

$$\limsup_{i \rightarrow \infty} \|x_i - x^*\|^{\frac{1}{i}} < 1.$$

**Věta 51** *Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$  (alespoň)  $R$ -lineárně právě tehdy, existují-li čísla  $0 < q < 1$ ,  $M > 0$  a index  $k \in N$  taková, že*

$$\|x_i - x^*\| \leq M q^{i-k} \|x_k - x^*\|$$

$\forall i \geq k$ .

**Definice 30** *Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$   $R$ -superlineárně, jestliže*

$$\lim_{i \rightarrow \infty} \|x_i - x^*\|^{\frac{1}{i}} = 0$$

**Definice 31** *Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$  (alespoň)  $Q$ -lineárně, jestliže*

$$\limsup_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} < 1.$$

**Věta 52** Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$  (alespoň)  $Q$ -lineárně právě tehdy, existuje-li číslo  $0 < q < 1$  a index  $k \in N$  tak, že

$$\frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} \leq q$$

$\forall i \geq k$ .

**Definice 32** Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$   $Q$ -superlineárně, jestliže

$$\lim_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} = 0$$

**Věta 53** Nechť  $x_i \rightarrow x$   $Q$ -lineárně ( $Q$ -superlineárně). Pak  $x_i \rightarrow x$   $R$ -lineárně ( $R$ -superlineárně).

**Definice 33** Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$   $m$ -krokově  $Q$ -superlineárně, jestliže existuje číslo  $m \in N$  takové, že

$$\lim_{i \rightarrow \infty} \frac{\|x_{i+m} - x^*\|}{\|x_i - x^*\|} = 0.$$

Posloupnost  $\{x_i\} \subset R^n$  konverguje k bodu  $x^*$  cyklicky  $m$ -krokově  $Q$ -superlineárně, jestliže existuje číslo  $m \in N$  takové, že

$$\lim_{k \rightarrow \infty} \frac{\|x_{(k+1)m+1} - x^*\|}{\|x_{km+1} - x^*\|} = 0.$$

**Poznámka 19**  $m$ -kroková  $Q$ -superlineární konvergence implikuje cyklickou  $m$ -krokově  $Q$ -superlineární konvergenci.

**Věta 54** Nechť  $x_i \rightarrow x^*$  cyklicky  $m$ -krokově  $Q$ -superlineárně a nechť existuje konstanta  $C > 0$  taková, že  $\|e_{i+1}\| \leq C\|e_i\| \forall i \in N$ . Pak  $x_i \rightarrow x^*$   $R$ -superlineárně.

## Základní optimalizační metody

**Definice 34** Základní optimalizační metodou nazýváme iterační proces, jehož výsledkem je posloupnost  $\{x_i\} \subset R^n$  taková, že

$$x_{i+1} = x_i + \alpha_i s_i$$

kde směrový vektor  $s_i \in R^n$  se určuje na základě hodnot  $x_j, F_j, g_j, G_j, 1 \leq j \leq i$ , a délka kroku  $\alpha_i > 0$  se určuje na základě chování funkce  $F : R^n \rightarrow R$  v okolí bodu  $x_i \in R^n$ .

**Definice 35** Řekneme, že základní optimalizační metoda je globálně konvergentní, jestliže pro libovolný počáteční bod  $x_1 \in R^n$  platí

$$\liminf_{i \rightarrow \infty} \|g(x_i)\| = 0.$$

**Poznámka 20** Definice 35 nezaručuje, že posloupnost  $\{x_i\}$  konverguje (může divergovat nebo mít více hromadných bodů, z nichž některé nemusí být stacionárními body). Dokonce, když  $x_i \rightarrow x^*$ , nemusí  $x^* \in R^n$  být lokálním minimem funkce  $F$  (je to vždy stacionární bod, který může být například sedlovým bodem).

**Metoda největšího spádu:**

$$\begin{aligned} s_i &= -g(x_i) \\ \alpha_i &= \arg \min_{\alpha \geq 0} F(x_i + \alpha s_i) \end{aligned}$$

Výhody:

- je globálně konvergentní
- používá pouze vektory ( $O(n)$  paměťových míst a operací)

Nevýhody:

- vyžaduje přesný výběr délky kroku
- je pouze  $R$ -lineárně konvergentní

$$\limsup_{i \rightarrow \infty} \|x_i - x^*\|^{\frac{1}{i}} \leq \frac{\kappa(G(x^*)) - 1}{\kappa(G(x^*)) + 1}$$

**Newtonova metoda:**

$$\begin{aligned} s_i &= -G^{-1}(x_i) g(x_i) \\ \alpha_i &= 1 \end{aligned}$$

Výhody:

- je superlineárně konvergentní
- používá jednoduchý výběr délky kroku

Nevýhody:

- není globálně konvergentní, je-li  $x_1$  daleko od  $x^*$ , nemusí konvergovat
- používá matici řádu  $n$  a je potřeba řešit soustavu lineárních rovnic ( $O(n^2)$  paměťových míst,  $O(n^3)$  operací)
- je třeba počítat druhé derivace

## Metody spádových směrů

Předpokládáme, že  $g_i \neq 0$ ,  $s_i \neq 0 \forall i \in N$  a označíme

$$c_i = -\frac{s_i^T g_i}{\|s_i\| \|g_i\|}$$

(směrový kosinus).

**Definice 36** Řekneme, že směrové vektory  $s_i \in R^n$ ,  $i \in N$ , jsou spádové, jestliže platí

$$c_i > 0 \quad \forall i \in N \quad (\text{S1a})$$

Řekneme, že směrové vektory  $s_i \in R^n$ ,  $i \in N$ , jsou stejnoměrně spádové, jestliže existuje konstanta  $\underline{c} > 0$  taková, že

$$c_i \geq \underline{c} \quad \forall i \in N \quad (\text{S1b})$$

**Věta 55** Necht'  $s_i = -H_i g_i$  (nebo  $B_i s_i = -g_i$ ) kde  $H_i$  (nebo  $B_i$ ) je symetrická pozitivně definitní matice. Pak platí

$$c_i^2 \geq \frac{1}{\kappa_i}$$

kde  $\kappa_i$  je spektrální číslo podmíněnosti matice  $H_i$  (nebo  $B_i$ ).

**Definice 37** Řekneme, že délky kroku  $\alpha_i > 0$ ,  $i \in N$ , splňují silnou Wolfeho podmínku nebo slabou Wolfeho podmínku nebo Goldsteinovu podmínku nebo Armijovu podmínku, jestliže existují čísla  $0 \leq \varepsilon_1 < 1/2$ ,  $\varepsilon_1 < \varepsilon_2 < 1$  tak, že

$$F_{i+1} - F_i \leq \varepsilon_1 \alpha_i s_i^T g_i \quad (\text{S2})$$

a buď

$$|s_i^T g_{i+1}| \leq \varepsilon_2 |s_i^T g_i| \quad (\text{S3a})$$

nebo

$$s_i^T g_{i+1} \geq \varepsilon_2 s_i^T g_i \quad (\text{S3b})$$

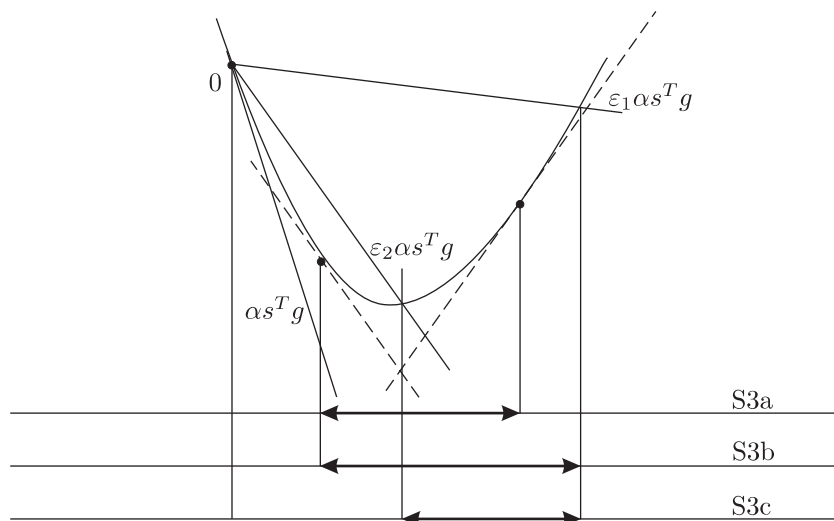
nebo

$$F_{i+1} - F_i \geq \varepsilon_2 \alpha_i s_i^T g_i \quad (\text{S3c})$$

nebo  $\alpha_i > 0$  je první člen vyhovující podmínce (S2) v posloupnosti  $\{\alpha_i^j\}$ ,  $j \in N$ , takové, že  $\underline{\alpha} \|g_i\| / \|s_i\| \leq \alpha_i^1 \leq \bar{\alpha} \|g_i\| / \|s_i\|$  a

$$\underline{\beta} \alpha_i^j \leq \alpha_i^{j+1} \leq \bar{\beta} \alpha_i^j \quad (\text{S3d})$$

$\forall i \in N$ , kde  $0 < \underline{\alpha} < \bar{\alpha}$  a  $0 < \underline{\beta} \leq \bar{\beta} < 1$  jsou konstanty nezávislé na indexu  $i \in N$ .



**Poznámka 21** (Konzistence) Necht' funkce  $F : R^n \rightarrow R$  splňuje podmínky (F1) a (F3) a směrový vektor  $s_i \in R^n$  je spádový. Pak existuje délka kroku  $\alpha_i > 0$  splňující podmínku (S2) a libovolnou z podmínek (S3a)-(S3d).

**Poznámka 22** Použití uvedených podmínek:

- (S3a) - metody sdružených gradientů
- (S3b) - metody s proměnnou metrikou
- (S3c) - numerický výpočet gradientů
- (S3d) - nehladké úlohy

**Definice 38** Základní optimalizační metoda (Definice 34) je metodou spádových směrů (stejněměrně spádových směrů), jestliže směrové vektory  $s_i \in R^n$ ,  $i \in N$ , splňují podmínku (S1a) (podmínku (S1b)) a délky kroku  $\alpha_i > 0$ ,  $i \in N$ , splňují podmínku (S2) a některou z podmínek (S3a)-(S3d).

**Poznámka 23** Metoda největšího spádu  $s_i = -g_i$  je metodou stejněměrně spádových směrů, protože  $s_i^T g_i = -\|g_i\|^2 = -\|s_i\| \|g_i\|$  a (S1b) platí pro  $\underline{c} = 1$ .

**Věta 56** (Globální konvergence) Necht' funkce  $F \in C^2 : R^n \rightarrow R$  splňuje podmínky (F1) a (F3). Pak metoda spádových směrů je globálně konvergentní, jestliže platí

$$\sum_{i=1}^{\infty} c_i^2 = \infty$$

**Poznámka 24** Jestliže  $s_i = -H_i g_i$  (nebo  $B_i s_i = -g_i$ ), můžeme předchozí podmínku nahradit podmínkou

$$\sum_{i=1}^{\infty} \frac{1}{\kappa_i} = \infty$$

kde  $\kappa_i$  jsou spektrální čísla podmíněnosti matic  $H_i$  (nebo  $B_i$ ).

**Poznámka 25** Pro metodu stejněměrně spádových směrů platí  $\|g_i\| \rightarrow 0$ . (Je to silnější tvrzení než globální konvergence podle Definice 35).

**Poznámka 26** Zvolíme  $\underline{c} > 0$  a položíme  $s_i = -g_i$  (restart), kdykoliv vyjde  $c_i < \underline{c}$ . Tato úprava zaručuje globální konvergenci.

**Věta 57** (*Superlineární konvergence*) Necht'  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou spádových směrů taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionárním bodem funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3) a (F4). Necht'

$$\lim_{i \rightarrow \infty} \frac{\|B_i s_i + g_i\|}{\|g_i\|} = 0$$

$$\lim_{i \rightarrow \infty} \frac{\|(B_i - G_i)s_i\|}{\|s_i\|} = 0$$

a necht'  $\alpha_i = 1$  kdykoliv tato hodnota vyhovuje podmínkám (S2) a některé z (S3a)-(S3d). Pak existuje index  $k \in N$  takový, že  $\alpha_i = 1 \forall i \geq k$  a  $x_i \rightarrow x^*$  Q-superlineárně.

**Algoritmus 1** (výběr délky kroku splňující (S2) a (S3b))

Data:  $0 < \underline{\beta} \leq \bar{\beta} < 1 < \underline{\gamma} < \bar{\gamma}$ . Počáteční hodnota  $\alpha > 0$ .

K1: Set  $\bar{\alpha} := 0$ .

K2: Set  $\underline{\alpha} := \bar{\alpha}$ ,  $\bar{\alpha} := \alpha$ . If (S2) and (S3b) stop ( $\alpha_i = \alpha$ ). If not (S2) go to K4.

K3: Určíme  $\alpha$  tak, že  $\underline{\gamma}\bar{\alpha} \leq \alpha \leq \bar{\gamma}\bar{\alpha}$  (extrapolace). Go to K2.

K4: Určíme  $\alpha$  tak, že  $\underline{\beta}(\bar{\alpha} - \underline{\alpha}) \leq (\alpha - \underline{\alpha}) \leq \bar{\beta}(\bar{\alpha} - \underline{\alpha})$  (interpolace).

K5: If (S2) and (S3b) stop ( $\alpha_i = \alpha$ ). If not (S2) set  $\bar{\alpha} := \alpha$  else set  $\underline{\alpha} := \alpha$ . Go to K4.

**Poznámka 27** Extrapolace a interpolace.

$$\varphi(\alpha) = F(x_i + \alpha s_i), \quad \varphi'(\alpha) = s_i^T g(x_i + \alpha s_i)$$

$$A = \frac{\varphi(\bar{\alpha}) - \varphi(\underline{\alpha})}{(\bar{\alpha} - \underline{\alpha})\varphi'(\underline{\alpha})}$$

$$B = \frac{\varphi'(\bar{\alpha})}{\varphi'(\underline{\alpha})}$$

Kvadratická (dvě hodnoty):

$$\alpha - \underline{\alpha} = \frac{\bar{\alpha} - \underline{\alpha}}{2(1 - A)}$$

Kvadratická (dvě derivace):

$$\alpha - \underline{\alpha} = \frac{\bar{\alpha} - \underline{\alpha}}{1 - B}$$

Kubická:

$$\alpha - \underline{\alpha} = \frac{\bar{\alpha} - \underline{\alpha}}{D + \sqrt{D^2 - 3C}}$$

$$C = (B - 1) - 2(A - 1)$$

$$D = (B - 1) - 3(A - 1)$$

**Definice 39** *Jestliže platí*

$$s_i^T g(x_i + \alpha_i s_i) = 0 \quad \forall i \in N$$

*řekneme, že výběr délky kroku je přesný. Jestliže platí*

$$\lim_{i \rightarrow \infty} \frac{s_i^T g(x_i + \alpha_i s_i)}{s_i^T g_i} = 0$$

*řekneme, že výběr délky kroku je asymptoticky přesný.*

**Poznámka 28** Pokud v každé iteraci používáme alespoň jednu kvadratickou nebo kubickou interpolaci (Algoritmus 1, Poznámka 10), je výběr délky kroku asymptoticky přesný.

**Věta 58** (*R-lineární konvergence*) *Nechť  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou stejnoměrně spádových směrů s asymptoticky přesným výběrem délky kroku taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionárním bodem funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3) a (F4). Pak platí*

$$\limsup_{i \rightarrow \infty} \|x_i - x^*\|^{\frac{1}{i}} \leq \frac{\kappa(G^*) - 1 + (\kappa(G^*) + 1)\sqrt{1 - \underline{c}^2}}{\kappa(G^*) + 1 + (\kappa(G^*) - 1)\sqrt{1 - \underline{c}^2}}$$

*kde  $\kappa(G^*)$  je spektrální číslo podmíněnosti matice  $G^*$  a  $\underline{c}$  je konstanta v (S1b).*

**Důsledek 6** *Pro metodu největšího spádu lze položit  $\underline{c} = 1$ , takže platí*

$$\limsup_{i \rightarrow \infty} \|x_i - x^*\|^{\frac{1}{i}} \leq \frac{\kappa(G^*) - 1}{\kappa(G^*) + 1}$$

## Metody sdružených gradientů

**Definice 40** *Řekneme, že metoda spádových směrů je metodou sdružených gradientů (CG), jestliže*

$$s_1 = -g_1 \quad a \quad s_{i+1} = -g_{i+1} + \beta_i s_i \quad \text{pro } i \in N, \quad (\text{CGa})$$

*kde parametr  $\beta_i$  se vybírá tak, aby směrové vektory  $s_i$ ,  $1 \leq i \leq n$ , byly sdružené (nebo  $G$ -ortogonální), aplikujeme-li tuto metodu na ryze konvexní kvadratickou funkci*

$$Q(x) = \frac{1}{2}(x - x^*)^T G(x - x^*)$$

*a používáme-li přesný výběr délky kroku (platí  $s_i^T g_{i+1} = 0 \quad \forall i \in N$ ).*

**Poznámka 29** Označme  $d_i = x_{i+1} - x_i = \alpha_i s_i$  a  $y_i = g_{i+1} - g_i$ . Pak pro kvadratickou funkci  $Q$  platí  $y_i = G d_i$  a podmínku  $G$ -ortogonality vektorů  $s_i$ ,  $s_{i+1}$  lze zapsat ve tvaru  $\alpha_i s_i^T G s_{i+1} = y_i^T s_{i+1} = 0$  (předpokládáme, že  $\alpha_i \neq 0$ ). Odtud prostřednictvím (CGa) dostaneme rovnici  $\beta_i y_i^T s_i - y_i^T g_{i+1} = 0$  neboli

$$\beta_i = \frac{y_i^T g_{i+1}}{y_i^T s_i}. \quad (\text{CGb})$$

Tato volba již zaručuje vzájemnou  $G$ -ortogonálnost směrových vektorů  $s_i$ ,  $1 \leq i \leq n$ .



**Věta 59** (Kvadratické ukončení) Nechť  $Q : R^n \rightarrow R$  je ryze konvexní kvadratická funkce a  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou (CGa)–(CGb) s přesným výběrem délky kroku. Pak pro  $1 \leq j < i \leq n + 1$  platí

$$\begin{aligned} s_j^T g_i &= 0 & (\text{biortogonalita směrů a gradientů}) \\ g_j^T g_i &= 0 & (\text{ortogonalita gradientů}) \\ s_j^T G s_i &= 0 & (\text{sduženost směrů}) \end{aligned}$$

Navíc existuje index  $m \leq n$  takový, že  $g_{m+1} = 0$  a  $x_{m+1} = x^*$ .

**Poznámka 30** Používáme-li přesný výběr délky kroku, můžeme podle (CGa) psát

$$y_i^T s_i = g_{i+1}^T s_i - g_i^T s_i = -g_i^T s_i = g_i^T g_i - \beta_{i-1} g_i^T s_{i-1} = g_i^T g_i.$$

Je-li navíc minimalizovaná funkce kvadratická, platí  $g_i^T g_{i+1} = 0$ , takže

$$y_i^T g_{i+1} = g_{i+1}^T g_{i+1} - g_i^T g_{i+1} = g_{i+1}^T g_{i+1}.$$

Odtud plyne, že ve vzorci pro  $\beta_i$  můžeme použít tři různé jmenovatele a dva různé čitatele, aniž bychom porušili platnost věty 59. Dostaneme tak šest základních metod sdružených gradientů.

$$\beta_i^{HS} = \frac{y_i^T g_{i+1}}{y_i^T s_i}, \quad \beta_i^{PR} = \frac{y_i^T g_{i+1}}{g_i^T g_i}, \quad \beta_i^{LS} = \frac{y_i^T g_{i+1}}{|g_i^T s_i|}$$

(HS – Hestenes a Stiefel, PR – Polak a Ribière, LS – Liu a Storey),

$$\beta_i^{DY} = \frac{g_{i+1}^T g_{i+1}}{y_i^T s_i}, \quad \beta_i^{FR} = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}, \quad \beta_i^{CD} = \frac{g_{i+1}^T g_{i+1}}{|g_i^T s_i|}$$

(DY – Dai a Yuan, FR – Fletcher a Reeves, CD – conjugate descent).

**Poznámka 31** Metody sdružených gradientů můžeme rozdělit do dvou skupin:

- Metody HS, PR, LS jsou výhodnější pro praktické použití, ale nejsou bez nutných úprav globálně konvergentní.
- Metody DY, FR, CD jsou za určitých předpokladů (výběr délky kroku) globálně konvergentní, ale hůře zachovávají sduženost směrových vektorů.
- Metody patřící do téže skupiny se svými vlastnostmi příliš neliší.

## Globální konvergence

**Věta 60** (Globální konvergence metody DY) Nechť  $F : R^n \rightarrow R$  splňuje podmínky (F1) a (F3). Pak metoda CG s výběrem délky kroku splňujícím slabou Wolfeho podmínku (S2), (S3b) je globálně konvergentní, pokud

$$\beta_i = \lambda_i \frac{g_{i+1}^T g_{i+1}}{y_i^T s_i}$$

kde  $-(1 - \varepsilon_2)/(1 + \varepsilon_2) \leq \lambda_i \leq 1$ .

**Věta 61** (Globální konvergence metody FR) Nechť  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  splňuje podmínky (F1) a (F3). Pak metoda CG s výběrem délky kroku splňujícím silnou Wolfeho podmínku (S2), (S3a) s  $\varepsilon_2 < 1/2$  je globálně konvergentní, pokud

$$\beta_i = \lambda_i \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}$$

kde  $|\lambda_i| \leq 1$ .

**Poznámka 32** Metody HS, PR, LS dávají lepší výsledky než metody DY, FR, CD. Konvergenci prvních metod lze zlepšit vyloučením záporných hodnot:

$$\beta_i^{HS+} = \max(0, \beta_i^{HS}), \quad \beta_i^{PR+} = \max(0, \beta_i^{PR}), \quad \beta_i^{LS+} = \max(0, \beta_i^{LS}),$$

nebo kombinováním s druhými metodami:

$$\beta_i^{HS/DY} = \max(0, \min(\beta_i^{HS}, \beta_i^{DY})),$$

$$\beta_i^{PR/FR} = \max(0, \min(\beta_i^{PR}, \beta_i^{FR})),$$

$$\beta_i^{LS/CD} = \max(0, \min(\beta_i^{LS}, \beta_i^{CD})).$$

Kombinované metody jsou globálně konvergentní za stejných předpokladů jako metody DY, FR, CD a dávají stejně dobré výsledky jako metody HS, PR, LS.

### Asymptotická rychlost konvergence

**Lemma 3** Nechť jsou splněny předpoklady věty 59. Nechť  $g_i \neq 0$  pro nějaký index  $1 \leq i \leq n$ . Pak platí

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq \max_{1 \leq k \leq n} P_i^2(\lambda_k), \quad (1)$$

kde  $P_i(\lambda)$  je libovolný polynom stupně  $i$  takový, že  $P_i(0) = 1$ , a  $\lambda_k$ ,  $1 \leq k \leq n$ , jsou vlastní čísla matice  $G$  seřazená vzestupně.

**Věta 62** Nechť jsou splněny předpoklady věty 59. Nechť  $g_i \neq 0$  pro nějaký index  $1 \leq i \leq n$ . Pak platí

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq \left( \frac{\lambda_{m+1-i} - \lambda_1}{\lambda_{m+1-i} + \lambda_1} \right)^2, \quad (2)$$

kde  $\lambda_k$ ,  $1 \leq k \leq m$ , jsou různá vlastní čísla matice  $G$  seřazená vzestupně.

**Důsledek 7** Metoda CG s přesným výběrem délky kroku nalezne minimum ryze konvexní kvadratické funkce po nejvýše  $m$  krocích, kde  $m$  je počet různých vlastních čísel matice  $G$ .

**Věta 63** Nechť jsou splněny předpoklady věty 59. Nechť  $g_i \neq 0$  pro nějaký index  $1 \leq i \leq n$ . Pak platí

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq 4 \left( \frac{\sqrt{\kappa(G)} - 1}{\sqrt{\kappa(G)} + 1} \right)^{2i}. \quad (3)$$

**Definice 41** Řekneme, že metoda CG je cyklicky přerušovaná, jestliže  $\beta_i = 0$ ,  $i \in M$ , kde

$$M = \{(i-1)n + 1 : i \in N\}$$

(vždy po  $n$  krocích se provede restart).

**Věta 64** Necht  $\{x_i\} \subset R^n$  je posloupnost bodů získaná cyklicky přerušovanou metodou CG s asymptoticky přesným výběrem délky kroku taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionárním bodem funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3)-(F5). Pak  $x_i \rightarrow x^*$   $R$ -superlineárně.

**Věta 65** Necht jsou splněny předpoklady věty 64. Pak uvnitř každého cyklu  $x_i \rightarrow x^*$  alespoň  $R$ -lineárně s asymptotickou rychlostí

$$q \leq \frac{\sqrt{\kappa(G^*)} - 1}{\sqrt{\kappa(G^*)} + 1}$$

**Poznámka 33** Odhad  $(\sqrt{\kappa} - 1)/(\sqrt{\kappa} + 1)$  je příznivější než odhad  $(\kappa - 1)/(\kappa + 1)$  platný pro metodu spádových směrů.

$\kappa$	$\varepsilon$	SD	CG
$10^2$	$10^{-4}$	460	46
$10^4$	$10^{-6}$	69077	690
$10^6$	$10^{-8}$	9210340	9210

V tabulce je uvedeno číslo podmíněnosti  $\kappa$ , požadovaná přesnost  $\varepsilon$  a počet iterací potřebných k dosažení požadované přesnosti.

### Modifikace a implementace metod sdružených gradientů

**Věta 66** Uvažujme modifikované metody sdružených gradientů dané předpisem

$$s_1 = -g_1 \quad a \quad s_{i+1} = -\vartheta_i g_{i+1} + \beta_i s_i \quad \text{pro } i \in N,$$

kde  $\beta_i$  nabývá hodnot  $\beta_i^{DY}$ ,  $\beta_i^{FR}$ ,  $\beta_i^{CD}$  a  $\vartheta_i$  hodnot

$$\vartheta_i^{DY} = \frac{y_i^T s_i}{y_i^T s_i} = 1, \quad \vartheta_i^{FR} = \frac{y_i^T s_i}{g_i^T g_i}, \quad \vartheta_i^{CD} = \frac{y_i^T s_i}{|g_i^T s_i|}.$$

Pak platí věta 59. Splňuje-li funkce  $F : R^n \rightarrow R$  podmínky (F1) a (F3), jsou modifikované metody DY, FR, CD s výběrem délky kroku splňujícím slabou Wolfeho podmínku (S2), (S3b) globálně konvergentní,

**Věta 67** Uvažujme modifikované metody sdružených gradientů dané předpisem

$$s_1 = -g_1 \quad a \quad s_{i+1} = -\vartheta_i g_{i+1} + \beta_i s_i \quad \text{pro } i \in N,$$

kde  $\beta_i$  nabývá hodnot  $\beta_i^{HS}$ ,  $\beta_i^{PR}$ ,  $\beta_i^{LS}$  a  $\vartheta_i$  hodnot

$$\vartheta_i^{HS} = \frac{y_i^T s_i}{y_i^T s_i} = 1, \quad \vartheta_i^{PR} = \frac{y_i^T s_i}{g_i^T g_i}, \quad \vartheta_i^{LS} = \frac{y_i^T s_i}{|g_i^T s_i|}.$$

Pak platí věta 59 a navíc  $y_i^T s_{i+1} = 0 \quad \forall i \in N$ .

**Poznámka 34** Účinnost metod CG lze zvýšit restartováním. Je vhodné testovat spádovost směrových vektorů. V tomto případě položíme  $\beta_i = 0$ , pokud neplatí

$$-g_{i+1}^T s_{i+1} \geq \varepsilon_0 \|g_{i+1}\| \|s_{i+1}\|,$$

kde  $\varepsilon_0 \approx 10^{-8}$ . Používáme-li metody DY, FR, CD, je výhodné testovat sdruženost směrových vektorů. V tomto případě položíme  $\beta_i = 0$ , pokud neplatí

$$y_i^T s_{i+1} \leq \eta_1 \|s_{i+1}\| \|y_i\|, \quad (4)$$

kde  $\eta_1 \approx 0.005$ .

**Poznámka 35** V případě metod CG je vhodné používat silnou Wolfeho podmínku s  $\varepsilon_1 = 10^{-4}$  a  $\varepsilon_2 = 10^{-1}$ . Algoritmus 1 je třeba upravit tak, že (S3b) ponecháme beze změny a společně s (S2) vyhodnocujeme podmínku

$$s_i^T g_{i+1} \leq \varepsilon_2 |s_i^T g_i|,$$

která je částí (S3a).

**Algoritmus 2** (Metoda sdružených gradientů)

Data  $\varepsilon_1 = 10^{-4}$ ,  $\varepsilon_2 = 10^{-1}$ ,  $\eta_1 = 0.05$ ,  $\underline{\varepsilon} > 0$ .

K1: Zvolíme  $x_1 \in R^n$ . **Set**  $F_1 := F(x_1)$ ,  $g_1 := g(x_1)$ . **Set**  $i := 1$ .

K2: **If**  $\|g_i\| \leq \underline{\varepsilon}$  **stop**. Určíme směrový vektor  $s_i$  pomocí zvolené metody a rozhodneme o přerušení iteračního procesu podle poznámky 34.

K3: Určíme délku kroku  $\alpha_i > 0$  použitím algoritmu 1 upraveného podle poznámky 35. **Set**  $x_{i+1} := x_i + \alpha_i s_i$ ,  $F_{i+1} := F(x_{i+1})$ ,  $g_{i+1} := g(x_{i+1})$ . **Set**  $i := i + 1$ , **go to** K2.

## Lineární metoda sdružených gradientů

Nechť  $B$  je symetrická pozitivně definitní matice. Pak řešení soustavy rovnic

$$Bs = -g$$

je ekvivalentní minimalizaci kvadratické funkce

$$\frac{1}{2} s^T B s + g^T s$$

Používá se předpokládání:  $C$  je symetrická pozitivně definitní snadno invertovatelná matice (předpokládavač),

$$\begin{aligned} C^{-\frac{1}{2}} B C^{-\frac{1}{2}} C^{\frac{1}{2}} s &= -C^{-\frac{1}{2}} g, \\ \tilde{B} \tilde{s} &= -\tilde{g}, \end{aligned}$$

kde  $\tilde{B}$  je předpokládaná matice. Měla by být co nejlépe podmíněná ( $\kappa(\tilde{B}) \ll \kappa(B)$ ).

**Algoritmus 3** (Metoda PCG pro řešení soustavy lineárních rovnic)

Data:  $0 \leq \omega < 1$

K1: **Set**  $s := 0, r := -g, p := 0, \beta := 0$

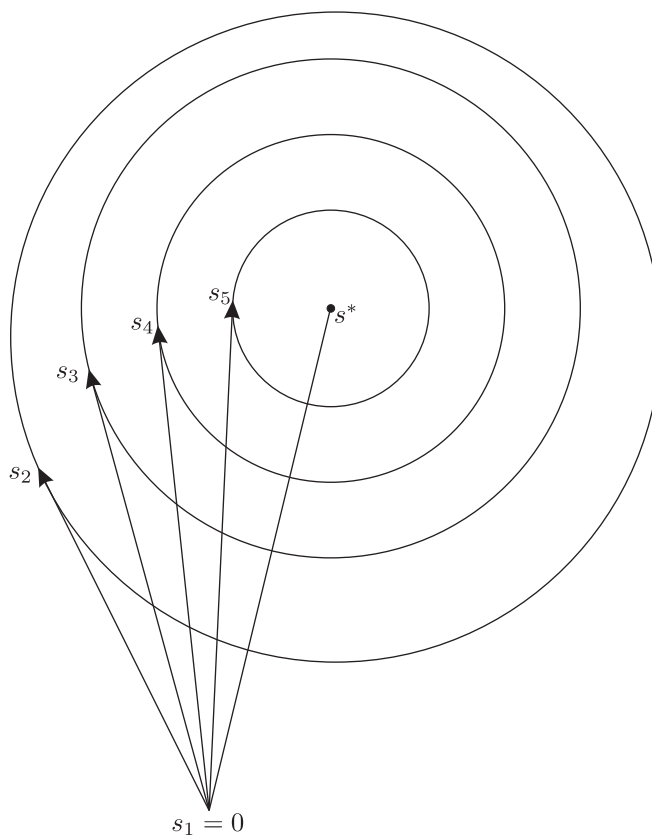
K2: **While**  $\|r\| > \omega\|g\|$  **do**

$$\begin{aligned} \tilde{r} &:= C^{-1}r, & \gamma &:= r^T \tilde{r}, & \beta &:= \beta\gamma \\ p &:= \tilde{r} + \beta p, & q &:= Bp, & \alpha &:= \gamma/p^T q \\ s &:= s + \alpha p, & r &:= r - \alpha q, & \beta &:= 1/\gamma \end{aligned}$$

**end while**

**Věta 68** *Aplikujeme-li metodu PCG s  $C = I$  na kvadratickou funkci  $Q(s)$  a platí-li  $g_j^T g_j > 0, p_j^T B p_j > 0$  pro  $1 \leq j \leq i$ , můžeme pro  $0 < \alpha < \alpha_i$  psát*

$$\begin{aligned} Q(s_{i+1}) &< Q(s_i(\alpha)) < Q(s_i), \\ g^T s_{i+1} &< g^T s_i(\alpha) < g^T s_i, \\ \|s_{i+1}\| &> \|s_i(\alpha)\| > \|s_i\|, \\ \frac{g^T s_{i+1}}{\|g\| \|s_{i+1}\|} &> \frac{g^T s_i(\alpha)}{\|g\| \|s_i(\alpha)\|} > \frac{g^T s_i}{\|g\| \|s_i\|}. \end{aligned}$$



**Věta 69** Jsou-li splněny předpoklady věty 68, platí

$$-Q(s_{i+1}) \geq \frac{1}{2} \frac{\|g\|^2}{\|B\|}, \quad -g^T s_{i+1} \geq \frac{\|g\|^2}{\|B\|}.$$

Je-li navíc matice  $B$  pozitivně definitní, platí

$$-\frac{g^T s_{i+1}}{\|g\| \|s_{i+1}\|} \geq \frac{1}{\sqrt{\kappa(B)}}.$$

## Metody s proměnnou metrikou

**Definice 42** Řekneme, že základní optimalizační metoda je metodou s proměnnou metrikou (VM), jestliže

$$s_i = -H_i g_i \quad \forall i \in N$$

kde  $H_i$ ,  $i \in N$  jsou s.p.d. matice takové, že  $H_1 = I$  a

$$H_{i+1} = \gamma_i (H_i + U_i M_i U_i^T)$$

kde  $U_i \in R^{n \times 2}$  a kde  $M_i \in R^{2 \times 2}$  se vybírá tak, aby byla splněna kvazinewtonovská (QN) podmínka

$$H_{i+1} y_i = \rho_i d_i$$

Přitom  $\gamma_i > 0$ ,  $\rho_i > 0$ ,  $d_i = x_{i+1} - x_i$ ,  $y_i = g_{i+1} - g_i$ .

**Poznámka 36** Nejeфекtivnější metody VM patří do Broydenovy třídy, kde

$$U_i = [d_i, H_i y_i]$$

**Poznámka 37** Zjednodušení:  $i$  vynecháme a  $i + 1$  nahradíme  $+$ .

**Věta 70** Nechť  $H_+ = \gamma(H + U M U^T)$ , kde  $U = [d, H y]$ . Pak QN podmínka  $H_+ y = \rho d$  platí právě tehdy, jestliže

$$M = \begin{bmatrix} \frac{1}{b} \left( \eta \frac{a}{b} + \frac{\rho}{\gamma} \right), & -\frac{\eta}{b} \\ -\frac{\eta}{b}, & \frac{\eta-1}{a} \end{bmatrix}$$

kde  $\eta$  je volný parametr a kde  $a = y^T H y$ ,  $b = y^T d$ ,  $c = d^T H^{-1} d$ .

**Poznámka 38** Po roznásobení dostaneme

$$H_+ = \gamma \left( H + \frac{\rho}{\gamma} \frac{1}{b} d d^T - \frac{1}{a} H y (H y)^T + \frac{\eta}{a} \left( \frac{a}{b} d - H y \right) \left( \frac{a}{b} d - H y \right)^T \right) \quad (\text{H})$$

Nejznámější metody z Broydenovy třídy:

- Davidon, Fletcher, Powell (DFP):  $\eta = 0$ :

$$H_+ = \gamma \left( H + \frac{\rho}{\gamma} \frac{1}{b} d d^T - \frac{1}{a} H y (H y)^T \right)$$

- Broyden, Fletcher, Goldfarb, Shanno (BFGS):  $\eta = 1$

$$H_+ = \gamma \left( H + \left( \frac{\rho}{\gamma} + \frac{a}{b} \right) \frac{1}{b} dd^T - \frac{1}{b} (Hyd^T + d(Hy)^T) \right)$$

- Metoda hodnosti 1 (R1)  $\eta = (\rho/\gamma)/(\rho/\gamma - a/b)$ :

$$H_+ = \gamma \left( H + \frac{\frac{\rho}{\gamma}}{\frac{\rho}{\gamma} - a} \left( \frac{\rho}{\gamma} d - Hy \right) \left( \frac{\rho}{\gamma} d - Hy \right)^T \right)$$

**Poznámka 39** Vlastnosti:

- DFP – velmi špatná, pomalá konvergence
- BFGS – spolehlivá a efektivní
- R1 – matice  $H_+$  nemusí být s.p.d. → selhání metody

**Poznámka 40** Škálování:

- Pro metodu DFP volíme  $\gamma$  tak aby platilo  $\rho/\gamma = b/c$
- Pro metodu BFGS volíme  $\gamma$  tak aby platilo  $\rho/\gamma = a/b$
- Pro metodu R1 volíme  $\gamma$  tak aby platilo  $\rho/\gamma = \sqrt{a/c}$

Tyto hodnoty se používají pouze tehdy, leží-li  $\gamma$  v intervalu  $[1, 6]$ , jinak se volí  $\gamma = 1$ .

**Poznámka 41** Korekce: Parametr  $\rho$  se nejčastěji volí podle vzorce

$$\rho = \frac{d^T y}{2(F - F_+ - s^T g_+)}$$

odvozeného z věty o střední hodnotě (Spedicato).

**Věta 71** *Nechť matice  $H$  je s.p.d.,  $a > 0$ ,  $b > 0$ ,  $c > 0$ ,  $ac - b^2 > 0$  a nechť platí  $(H)$  s  $\gamma > 0$ ,  $\rho > 0$ . Pak  $H_+$  je s.p.d. právě tehdy, jestliže  $\eta \geq \eta^*$ , kde  $\eta^* = -b^2/(ac - b^2) < 0$ .*

**Poznámka 42** Podmínka pro s.p.d. je vždy splněna pro  $\eta \geq 0$  (perfektní část Broydenovy třídy).

**Věta 72** *(Inverzní VM metody) Nechť jsou splněny předpoklady Věty 71 a nechť  $B = H^{-1}$  a  $B_+ = H_+^{-1}$  ( $H_+$  je určena podle  $(H)$ ). Pak platí*

$$B_+ = \frac{1}{\gamma} \left( B + \frac{\gamma}{\rho b} yy^T - \frac{1}{c} Bd(Bd)^T + \frac{\beta}{c} \left( \frac{c}{b} y - Bd \right) \left( \frac{c}{b} y - Bd \right)^T \right) \quad (\text{B})$$

kde

$$\beta\eta(ac - b^2) + (\beta + \eta)b^2 = b^2$$

**Poznámka 43** (Dualita) Vztah (B) dostaneme z (H) záměnou  $\gamma \rightarrow 1/\gamma$ ,  $\rho \rightarrow 1/\rho$ ,  $a \rightarrow c$ ,  $c \rightarrow a$ ,  $d \rightarrow y$ ,  $y \rightarrow d$ ,  $H \rightarrow B$ ,  $\eta \rightarrow \beta$ . Metody DFP a BFGS jsou navzájem duální. Metoda R1 je samoduální.

- DFP:  $\beta = 1$

$$B_+ = \frac{1}{\gamma} \left( B + \left( \frac{\gamma}{\rho} + \frac{c}{b} \right) \frac{1}{b} y y^T - \frac{1}{b} (B d y^T + y (B d)^T) \right)$$

- BFGS:  $\beta = 0$

$$B_+ = \frac{1}{\gamma} \left( B + \frac{\gamma}{\rho} \frac{1}{b} y y^T - \frac{1}{c} B d (B d)^T \right)$$

- R1:  $\beta = (\gamma/\rho)/(\gamma/\rho - c/b)$ :

$$B_+ = \frac{1}{\gamma} \left( B + \frac{\frac{\gamma}{\rho}}{\frac{\gamma}{\rho} b - c} \left( \frac{\gamma}{\rho} y - B d \right) \left( \frac{\gamma}{\rho} y - B d \right)^T \right)$$

**Poznámka 44** (Součinnový tvar) Nechť  $H = S S^T$ ,  $H_+ = S_+ S_+^T$ ,  $B = A A^T$ ,  $B_+ = A_+ A_+^T$ . Nechť  $d = S \tilde{d}$ ,  $\tilde{d} = A d$ ,  $y = A^T \tilde{y}$ ,  $\tilde{y} = S^T y$ , takže  $a = \tilde{y}^T \tilde{y}$ ,  $b = \tilde{y}^T \tilde{d}$ ,  $c = \tilde{d}^T \tilde{d}$ .

- DFP:  $\eta = 0$ ,  $\beta = 1$

$$\begin{aligned} S_+ &= \sqrt{\gamma} \left( S + \frac{1}{a} S \left( \sqrt{\frac{\rho a}{\gamma b}} \tilde{d} - \tilde{y} \right) \tilde{y}^T \right) \\ A_+ &= \frac{1}{\sqrt{\gamma}} \left( A + \frac{1}{b} \left( \sqrt{\frac{\gamma b}{\rho a}} \tilde{y} - \tilde{d} \right) \tilde{y}^T A \right) \end{aligned}$$

- BFGS:  $\eta = 1$ ,  $\beta = 0$

$$\begin{aligned} S_+ &= \sqrt{\gamma} \left( S + \frac{1}{b} S \tilde{d} \left( \sqrt{\frac{\rho b}{\gamma c}} \tilde{d} - \tilde{y} \right)^T \right) \\ A_+ &= \frac{1}{\sqrt{\gamma}} \left( A + \frac{1}{c} \tilde{d} \left( \sqrt{\frac{\gamma c}{\rho b}} \tilde{y} - \tilde{d} \right)^T A \right) \end{aligned}$$

**Věta 73** (Globální konvergence) Nechť  $F : R^n \rightarrow R$  splňuje podmínky (F1), (F3) a (F4). Pak metoda VM s výběrem délky kroku splňujícím slabou Wolfeho podmínkou (S2), (S3b) taková, že  $1 \leq \underline{\gamma}_i \leq \overline{\gamma}$ ,  $0 < \underline{\rho} < \rho_i \leq \overline{\rho}$ ,  $0 < \underline{\eta} < \eta_i \leq \overline{\eta}$ , je globálně konvergentní.

**Poznámka 45** Globální konvergence vyžaduje splnění poměrně silné podmínky (F4) (stejněměrná konvexita). Tuto podmínku lze oslabit modifikací výběru délky kroku nebo modifikací aktualizace VM matice. V praxi se to neprovádí, neboť VM metody (např. BFGS) jsou velmi robustní a efektní.

**Poznámka 46** Zatím nikdo nedokázal, že (nemodifikovaná) DFP metoda je globálně konvergentní.



**Věta 74** (*Superlineární konvergence*) Necht'  $\{x_i\} \subset R^n$  je posloupnost bodů získaná VM metodou s  $\gamma_i = \rho_i = 1$  taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionárním bodem funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3)-(F5). Pokud  $0 \leq \eta_i \leq 1$  a pokud  $\alpha_i = 1$  jestliže tato hodnota vyhovuje podmínkám (S2) a (S3b), pak  $x \rightarrow x^*$  Q-superlineárně.

**Poznámka 47** Věta 74 vyžaduje, aby platilo  $\gamma_i = \rho_i = 1$ . Ačkoliv není dokázána superlineární konvergence pro  $\gamma_i \neq 1$ , škálování obecně zlepšuje účinnost VM metody.

**Věta 75** (*Kvadratické ukončení*) Necht'  $Q : R^n \rightarrow R$  je ryze konvexní kvadratická funkce a necht'  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou VM s přesným výběrem délky kroku. Pak pro  $1 \leq j < i \leq n + 1$  platí

$$\begin{aligned} H_i y_j &= \lambda_j^i d_j \\ s_j^T g_i &= 0 \\ s_j^T G s_i &= 0 \quad (\text{sduženost směrů}) \end{aligned}$$

Navíc existuje index  $k \leq n$  takový, že  $g_{k+1} = 0$  a  $x_{k+1} = x^*$ .

**Poznámka 48** Je-li výběr délky kroku přesný, jsou všechny metody VM ekvivalentní (generují stejné posloupnosti bodů a matic).

**Algoritmus 4** (Metoda s proměnnou metrikou)

Data  $\varepsilon_1 = 10^{-4}$ ,  $\varepsilon_2 = 0.9$ ,  $\underline{\varepsilon} > 0$ ,  $\underline{\rho} = 0.01$ ,  $\bar{\rho} = 100$ ,  $\underline{\gamma} = 1.0$ ,  $\bar{\gamma} = 6.0$ .

K1: Zvolíme  $x_1 \in R^n$  a s.p.d. matici  $H_1$  (obvykle  $H_1 := I$ ). **Set**  $F_1 := F(x_1)$ ,  $g_1 := g(x_1)$ .  
**Set**  $i := 1$ .

K2: **If**  $\|g_i\| \leq \underline{\varepsilon}$  **stop**. **Set**  $s_i := -H_i g_i$ .

K3: Určíme délku kroku  $\alpha_i > 0$  použitím algoritmu 1. **Set**  $x_{i+1} := x_i + \alpha_i s_i$ ,  $F_{i+1} := F(x_{i+1})$ ,  $g_{i+1} := g(x_{i+1})$ .

K4: **Set**  $d_i := x_{i+1} - x_i$ ,  $y_i := g_{i+1} - g_i$ . Určíme parametr  $\rho_i > 0$  podle poznámky 41. **If**  $\rho_i < \underline{\rho}$  **or**  $\rho_i > \bar{\rho}$  **set**  $\rho_i := 1$ . Určíme parametr  $\gamma_i > 0$  podle poznámky 40 (pro metodu BFGS volíme  $\gamma_i/\rho_i = b_i/a_i$ ). **If**  $\gamma_i < \underline{\gamma}$  **or**  $\gamma_i > \bar{\gamma}$  **set**  $\gamma_i := 1$ . Zvolíme parametr  $\eta_i > 0$  a určíme matici  $H_{i+1}$  podle (H) (pro metodu BFGS volíme  $\eta_i = 1$ ).  
**Set**  $i := i + 1$ , **go to** K2.

## Metody s lokálně omezeným krokem

Budeme používat označení

$$Q_i(s) = g_i^T s + \frac{1}{2} s^T B_i s$$

(kvadratická aproximace  $F(x_i + s) - F(x_i)$ )

$$\omega_i(s) = (B_i s + g_i) / \|g_i\|$$

(přesnost určení směrového vektoru)

$$\rho_i(s) = \frac{F(x_i + s) - F(x_i)}{Q_i(s)}$$

(podíl skutečného a předpověděného poklesu funkce  $F : R^n \rightarrow R$ ).

**Definice 43** Řekneme, že základní optimalizační metoda (Definice 34) je metodou s lokálně omezeným krokem, (trust region), jestliže se směrové vektory  $s_i \in R^n$ ,  $i \in N$ , určují tak, že

$$\|s_i\| \leq \Delta_i \quad (\text{T1a})$$

$$\|s_i\| < \Delta_i \Rightarrow \|\omega_i(s_i)\| \leq \bar{\omega} \quad (\text{T1b})$$

$$-Q_i(s_i) \geq \underline{\sigma} \|g_i\| \min(\|s_i\|, \|g_i\|/\|B_i\|) \quad (\text{T1c})$$

kde  $0 < \underline{\sigma} < 1$  a  $0 \leq \bar{\omega} < 1$ , kde délky kroku  $\alpha_i \geq 0$ ,  $i \in N$ , se vybírají tak, že

$$\rho_i(s_i) \leq 0 \Rightarrow \alpha_i = 0 \quad (\text{T2a})$$

$$\rho_i(s_i) > 0 \Rightarrow \alpha_i = 1 \quad (\text{T2b})$$

a kde čísla  $0 < \Delta_i \leq \bar{\Delta}$ ,  $i \in N$ , se volí tak, že

$$\rho_i(s_i) \leq \underline{\rho} \Rightarrow \underline{\beta} \|s_i\| \leq \Delta_{i+1} \leq \bar{\beta} \|s_i\| \quad (\text{T3a})$$

$$\rho_i(s_i) \geq \underline{\rho} \Rightarrow \|s_i\| \leq \Delta_{i+1} \leq \min(\bar{\gamma} \|s_i\|, \bar{\Delta}) \quad (\text{T3b})$$

kde  $0 < \underline{\beta} \leq \bar{\beta} < 1 < \bar{\gamma}$  a  $0 < \underline{\rho} < 1$  (číslo  $\bar{\Delta} > 0$  slouží k tomu, abychom se nedostali mimo definiční obor funkce  $F$ ).

**Poznámka 49** Jestliže  $\bar{\omega} = 0$  nebo  $\bar{\omega} > 0$ , dostaneme přesné nebo nepřesné metody s lokálně omezeným krokem.

**Poznámka 50** Normy v (T1) a (T3) mohou být i jiné než euklidovské. Některé nerovnosti v (T1)–(T3) mohou být upraveny (oslabeny nebo zesíleny).

**Věta 76** (Globální konvergence) Nechť funkce  $F \in C^2 : R^n \rightarrow R$  splňuje podmínky (F1) a (F3). Pak metoda s lokálně omezeným krokem je globálně konvergentní, jestliže platí

$$\sum_{i=1}^{\infty} \frac{1}{M_i} = \infty,$$

kde  $M_i = \max_{1 \leq j \leq i} \|B_j\|$ .

**Poznámka 51** Podmínka  $\sum_{i=1}^{\infty} 1/M_i = \infty$  použitá ve větě 76 je mnohem slabší než podmínka  $\sum_{i=1}^{\infty} 1/\kappa(B_i) = \infty$  použitá v poznámce 24 (matice  $B_i$  nemusí být pozitivně definitní ani dostatečně dobře podmíněná). Předpoklady věty 76 jsou splněny například tehdy, jsou-li matice  $B_i$ ,  $i \in N$ , stejnoměrně omezené, tedy

$$\|B_i\| \leq \bar{B} \quad \forall i \in N$$

**Věta 77** (Superlineární konvergence) Nechť  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou s lokálně omezeným krokem taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionárním bodem funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3) a (F4). Nechť

$$\lim_{i \rightarrow \infty} \omega_i = 0$$

$$\lim_{i \rightarrow \infty} \frac{\|(B_i - G_i)s_i\|}{\|s_i\|} = 0$$

Pak  $x_i \rightarrow x^*$   $Q$ -superlineárně.

## Metody s optimálním lokálně omezeným krokem

**Definice 44** Metody s optimálním lokálně omezeným krokem používají směrové vektory

$$s_i = \arg \min_{\|s\| \leq \Delta_i} Q_i(s) \quad (\overline{T1})$$

přičemž  $\|s_i\| = \Delta_i$ , není-li toto minimum jediné.

**Věta 78** Vektor  $s_i \in R^n$  určený podle  $(\overline{T1})$  vyhovuje podmínkám (T1) s  $\overline{w} = 0$  a  $\underline{\sigma} = 1/2$ .

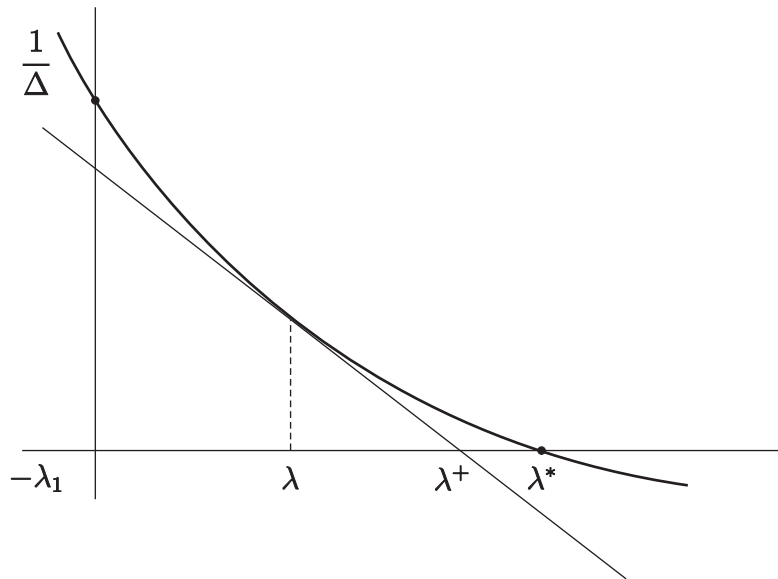
**Věta 79** Vektor  $s_i \in R^n$  je řešením  $(\overline{T1})$  právě tehdy, jestliže  $\|s_i\| \leq \Delta_i$  a jestliže existuje číslo  $\lambda_i \geq 0$  takové, že matice  $B_i + \lambda_i I$  je pozitivně semidefinitní a platí  $(B_i + \lambda_i I)s_i + g_i = 0$  a  $(\|s_i\| - \Delta_i)\lambda_i = 0$ .

**Poznámka 52** Je-li matice  $B_i$  s.p.d. a  $\|B_i^{-1}g_i\| \leq \Delta$ , pokládáme  $s_i = -B_i^{-1}g_i$ . Jinak řešíme nelineární rovnici

$$\varphi(\lambda) = \frac{1}{\Delta_i} - \frac{1}{\|s_i(\lambda)\|} = 0$$

kde

$$(B_i + \lambda I)s_i(\lambda) + g_i = 0$$



**Algoritmus 5** (určení optimálního lokálně omezeného kroku)

Data:  $0 < \underline{\delta} < 1 < \overline{\delta}$ ,  $\Delta > 0$

K1: Set  $\underline{\gamma} := \max_{1 \leq i \leq n} (-B_{ii})$ . Set  $\underline{\Delta} := 0$ ,  $\overline{\lambda} := \|g\|/\Delta + \|B\|$  a  $\lambda := \max(\underline{\gamma}, \underline{\Delta})$ .

K2: Set  $\underline{\lambda} := \max(\underline{\gamma}, \underline{\Delta})$ . If  $\lambda < \underline{\lambda}$  set  $\lambda := \sqrt{\lambda \overline{\lambda}}$ .

K3: If  $B + \lambda I$  je s.p.d., určíme rozklad  $R^T R = B + \lambda I$  ( $R$  je horní trojúhelníková) a go to K4. If  $B + \lambda I$  není s.p.d., určíme  $v \in R^n$  tak, že  $\|v\| = 1$  a  $v^T(B + \lambda I)v \leq 0$ , set  $\underline{\gamma} := \lambda - v^T(B + \lambda I)v$  a go to K2.

*K4:* **Set**  $s := -(R^T R)^{-1}g$ . **If**  $\|s\| > \bar{\delta}\Delta$ , **set**  $\underline{\lambda} := \lambda$  **a go to** *K6*. **If**  $\underline{\delta}\Delta \leq \|s\| \leq \bar{\delta}\Delta$  **stop** ( $s_i = s$ ). **If**  $\|s\| < \underline{\delta}\Delta$  **a**  $\lambda = 0$  **stop** ( $s_i = s$ ). **If**  $\|s\| < \underline{\delta}\Delta$  **a**  $\lambda \neq 0$ , **set**  $\bar{\lambda} := \lambda$  **a go to** *K5*.

*K5:* Určíme  $v \in R^n$  jako aproximaci vlastního vektoru matice  $B$  příslušného vlastnímu číslu  $\underline{\lambda}(B)$  tak, že  $\|v\| = 1$  a  $v^T s > 0$ . Určíme  $\alpha \geq 0$  tak, že  $\|s + \alpha v\| = \Delta$ .

**If**  $\alpha^2 \|Rv\|^2 \leq (1 - \underline{\delta}^2)(\|Rs\|^2 + \lambda\Delta^2)$  **set**  $s := s + \alpha v$  **a stop** ( $s_i = s$ ).

**If**  $\alpha^2 \|Rv\|^2 > (1 - \underline{\delta}^2)(\|Rs\|^2 + \lambda\Delta^2)$  **set**  $\underline{\gamma} := \lambda - \|Rv\|^2$  **a go to** *K6*.

*K6:* **Set**  $v := (R^T)^{-1}s$  **a**

$$\lambda := \lambda + \frac{\|s\|^2}{\|v\|^2} \left( \frac{\|s\| - \Delta}{\Delta} \right)$$

**If**  $\lambda < \underline{\lambda}$ , **set**  $\lambda := \underline{\lambda}$ . **If**  $\lambda > \bar{\lambda}$ , **set**  $\lambda := \bar{\lambda}$ . **Go to** *K2*.

**Poznámka 53** Vektor  $v \in R^n$ , použitý v kroku *K5*, lze určit pomocí programů z knihovny LAPACK; číslo  $\alpha \geq 0$  takové, že  $\|s + \alpha v\| = \Delta$  se určí podle vzorce

$$\alpha = -v^T s + \sqrt{(v^T s)^2 + \Delta^2 - \|s\|^2} = \frac{\Delta^2 - \|s\|^2}{v^T s + \sqrt{(v^T s)^2 + \Delta^2 - \|s\|^2}}$$

**Poznámka 54** Rozklad  $R^T R = B + \lambda I$  v kroku *K3* se provádí průměrně 2-3 krát v každé iteraci. Proto se někdy používá směrový vektor

$$s_i = \arg \min_{\|s(\alpha, \beta)\| \leq \Delta_i} Q_i(s(\alpha, \beta)),$$

kde

$$s(\alpha, \beta) = \alpha g_i + \beta B_i^{-1} g_i.$$

Tato úloha má dimenzi 2 a rozklad matice  $R^T R = B + \lambda I$  se provádí pouze jednou v každé iteraci.

## Nepřesné metody s lokálně omezeným krokem

Aplikujeme-li metodu CG na kvadratickou funkci

$$Q(s) = g^T s + \frac{1}{2} s^T B s,$$

platí (Algoritmus 3 s  $g = -r$ )  $s_1 = 0$ ,  $g_1 = g$ ,  $p_1 = -g$  a

$$\begin{aligned} q_j &= B p_j, \\ \alpha_j &= g_j^T g_j / p_j^T q_j \\ s_{j+1} &= s_j + \alpha_j p_j \\ g_{j+1} &= g_j + \alpha_j q_j \\ \beta_j &= g_{j+1}^T g_{j+1} / g_j^T g_j \\ p_{j+1} &= -g_{j+1} + \beta_j p_j \end{aligned}$$

pro  $1 \leq j \leq n$ .

**Poznámka 55** Aplikujeme-li metodu CG na kvadratickou funkci  $Q(s)$  a je-li  $p_j^T B p_j > 0$ ,  $1 \leq j \leq m \leq n$ , pak pro  $1 \leq j \leq m$  platí

$$\begin{aligned} Q(s_{j+1}) &< Q(s_j) \\ \|s_{j+1}\| &> \|s_j\| \end{aligned}$$

(věta 68) a

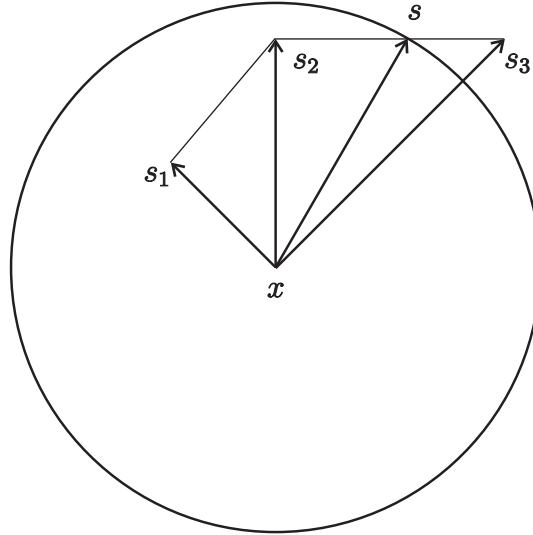
$$Q(s_{j+1}) \leq -\frac{1}{2} \|g\|^2 / \|B\|$$

(věta 69).

**Poznámka 56** Pokud  $p_j^T B p_j \leq 0$  nebo  $\|s_j\| < \Delta$  a  $\|s_{j+1}\| \geq \Delta$ , položíme  $s = s_j + \alpha_j p_j$ , kde  $\alpha_j > 0$  je hodnota taková, že  $\|s_j + \alpha_j p_j\| = \Delta$ , neboli

$$\alpha_j = \frac{\sqrt{(p_j^T s_j)^2 + (\Delta^2 - \|s_j\|^2) \|p_j\|^2} - p_j^T s_j}{\|p_j\|^2}.$$

Pak pro  $j \geq 2$  platí  $Q(s) \leq -(1/2) \|g\|^2 / \|B\|$  a pro  $j = 1$  platí  $Q(s) \leq -(1/2) \|g\| \|s\|$ , takže je splněna podmínka (T1c) s  $\underline{\sigma} = 1/2$ .



**Algoritmus 6** (určení nepřesného lokálně omezeného kroku)

Data:  $0 < \omega < 1$ ,  $\Delta > 0$ .

K1: Set  $s := 0$ ,  $r := -g$ ,  $p := r$ ,  $\sigma := \|r\|^2$ .

K2: Set  $\rho := \sigma$ ,  $q := Bp$ ,  $\delta := p^T q$ . If  $\delta \leq 0$  set

$$\alpha = (\sqrt{(p^T s)^2 + (\Delta^2 - \|s\|^2) \|p\|^2} - p^T s) / \|p\|^2$$

$s := s + \alpha p$  a stop ( $s_i = s$ ). If  $\delta > 0$  go to K3.

K3: Set  $\alpha = \rho / \delta$ . If  $\|s + \alpha p\| \geq \Delta$  set

$$\alpha = (\sqrt{(p^T s)^2 + (\Delta^2 - \|s\|^2) \|p\|^2} - p^T s) / \|p\|^2$$

$s := s + \alpha p$  a stop ( $s_i = s$ ). If  $\|s + \alpha p\| < \Delta$  go to K4.

*K4*: Set  $s := s + \alpha p$ ,  $r := r - \alpha q$ ,  $\sigma := \|r\|^2$ .

If  $\sigma \leq \omega^2 \|g\|^2$  stop ( $s_i = s$ ). If  $\sigma > \omega^2 \|g\|^2$  set  $\beta := \sigma/\rho$ ,  $p := r + \beta p$  a go to **K2**.

## Metody psí nohy

Metody psí nohy kombinují směr největšího spádu se směrem Newtonovy metody

$$s_C = -\frac{g^T g}{g^T B g} g$$

$$s_N = -B^{-1} g$$

**Algoritmus 7** (metoda psí nohy)

Data:  $\Delta > 0$ .

*K1*: If  $g^T B g \leq 0$  nebo  $\|s_C\| \geq \Delta$ , set  $s := -(\Delta/\|g\|)g$  a stop ( $s_i = s$ ).

*K2*: If  $B$  je singulární, určíme vektor  $v \in R^n$  tak, že  $\|v\| = 1$ ,  $v^T B v \leq 0$ ,  $v^T (g + B s_C) \leq 0$  a go to *K4*. If  $B$  není singulární, vypočteme  $s_N$ . If  $s_N = s_C$  set  $s := s_N$  a stop ( $s_i = s$ ).

*K3*: If  $(s_N - s_C)^T s_C \leq 0$  set

$$v := -\frac{s_N - s_C}{\|s_N - s_C\|}$$

a go to *K4*. If  $(s_N - s_C)^T s_C > 0$  a  $\|s_N\| > \Delta$  set

$$v := +\frac{s_N - s_C}{\|s_N - s_C\|}$$

a go to *K4*. If  $(s_N - s_C)^T s_C > 0$  a  $\|s_N\| \leq \Delta$  set  $s := s_N$  a stop ( $s_i = s$ ).

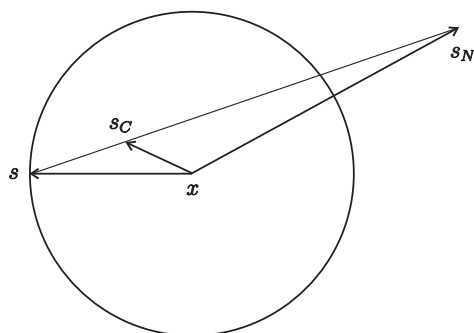
*K4*: Set

$$\alpha := \frac{\Delta^2 - \|s_C\|^2}{v^T s_C + \sqrt{(v^T s_C)^2 + \Delta^2 - \|s_C\|^2}}$$

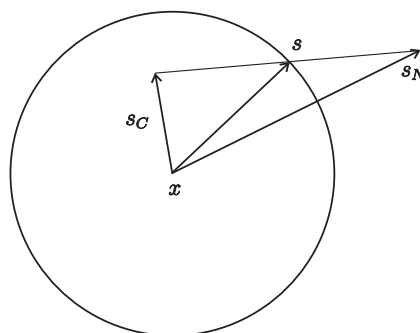
pokud  $v^T s_C \geq 0$  nebo

$$\alpha := -v^T s_C + \sqrt{(v^T s_C)^2 + \Delta^2 - \|s_C\|^2}$$

pokud  $v^T s_C < 0$ , set  $s := s_C + \alpha v$  a stop ( $s_i = s$ ).



$$(s_N - s_C)^T s_C \leq 0$$



$$(s_N - s_C)^T s_C > 0$$

## Newtonova metoda

Newtonova metoda používá matice  $B_i = G(x_i)$ ,  $i \in N$ , takže z podmínky (F3) plyne  $\|B_i\| = \|G(x_i)\| \leq \bar{G}$ ,  $i \in N$ .

**Věta 80** *Nechť funkce  $F \in C^2 : R^n \rightarrow R$  splňuje podmínky (F1) a (F3). Pak Newtonova metoda s lokálně omezeným krokem je globálně konvergentní. Je-li navíc splněna podmínka (F4) a platí-li  $x_i \rightarrow x^*$  a  $\omega_i(s_i) \rightarrow 0$ , pak  $x_i \rightarrow x^*$  Q-superlineárně.*

**Poznámka 57** Realizace Newtonovy metody

- Nepřesná Newtonova metoda: Soustava  $B_i s_i = -g_i$  se řeší nepřesně ( $\omega_i(s_i) > 0$ ) metodou CG, takže je třeba méně než  $O(n^3)$  operací na iteraci. Obvykle

$$\bar{\omega}_i = \min(\bar{\omega}, \|g_i\|, 1/i).$$

- Newtonova metoda s optimálním lokálně omezeným krokem.

**Věta 81** *Nechť funkce  $F \in C^2 : R^n \rightarrow R$  splňuje podmínky (F1)-(F3) a  $\{x_i\} \subset R^n$  je posloupnost generovaná Newtonovou metodou s optimálním lokálně omezeným krokem. Pak existuje hromadný bod  $x^* \in R^n$  posloupnosti  $\{x_i\} \subset R^n$  takový, že  $g(x^*) = 0$  a  $G(x^*) \succeq 0$  (pozitivně semidefinitní). Jestliže bod  $x^*$  vyhovuje postačujícím podmínkám pro extrém (Věta 51:  $g(x^*) = 0$ ,  $G(x^*) \succ 0$ ), pak  $x^*$  je jediným hromadným bodem posloupnosti  $\{x_i\} \subset R^n$  a  $x_i \rightarrow x^*$  Q-superlineárně.*

## Gaussova metoda pro součet čtverců

**Definice 45** *Řekneme, že funkce  $F : R^n \rightarrow R$  je součtem čtverců, jestliže existuje zobrazení  $f : R^n \rightarrow R^m$ ,  $m \geq n$ , takové, že*

$$F(x) = \frac{1}{2} f^T(x) f(x) = \frac{1}{2} \sum_{k=1}^m f_k^2(x)$$

**Poznámka 58** Je-li  $F : R^n \rightarrow R$  součtem čtverců, platí

$$\begin{aligned} g(x) &= J^T(x) f(x) = \sum_{k=1}^m f_k(x) g_k(x) \\ G(x) &= J^T(x) J(x) + C(x) = \sum_{k=1}^m g_k(x) g_k^T(x) + \sum_{k=1}^m f_k(x) G_k(x) \end{aligned}$$

kde  $g_k(x) = \nabla f_k(x)$ ,  $G_k(x) = \nabla^2 f_k(x)$  a

$$J(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Gaussovu-Newtonovu (GN) metodu pro součet čtverců dostaneme z Newtonovy metody zanedbáním členu

$$C(x) = \sum_{k=1}^m f_k(x) G_k(x)$$

v Hessově matici. Tedy  $B_i s_i = -g_i$ ,  $i \in N$ , kde

$$B_i = J_i^T J_i = \sum_{k=1}^m g_k(x_i) g_k^T(x_i)$$

### Zdůvodnění:

- Úlohy s nulovým reziduem  $F(x^*) = 0$ . Z  $x_i \rightarrow x^*$  plyne  $F(x_i) \rightarrow F(x^*) = 0$ , takže  $f_k(x_i) \rightarrow 0$ ,  $1 \leq k \leq m$ . Jestliže  $\|G_k(x_i)\| \leq \bar{G}$ ,  $1 \leq k \leq m$ , pak

$$\|C(x_i)\| = \left\| \sum_{k=1}^m f_k(x_i) G_k(x_i) \right\| \leq \bar{G} \sum_{k=1}^m |f_k(x_i)| \rightarrow 0$$

a tedy  $\|G(x_i) - B_i\| = \|C(x_i)\| \rightarrow 0$  z čehož plyne  $Q$ -superlineární konvergence.

- Linearizace. Platí

$$\begin{aligned} F(x_i + s) &= \frac{1}{2} f^T(x_i + s) f(x_i + s) \\ &\approx \frac{1}{2} (f(x_i) + J(x_i)s)^T (f(x_i) + J(x_i)s) \\ &= \frac{1}{2} f^T(x_i) f(x_i) + f^T(x_i) J(x_i)s + \frac{1}{2} s^T J^T(x_i) J(x_i)s \end{aligned}$$

takže

$$F(x_i + s) - F(x_i) \approx g^T(x_i)s + \frac{1}{2} s^T B_i s$$

což je lokální kvadratická aproximace s maticí  $B_i = J_i^T J_i$ .

**Poznámka 59** Podmínku (F3) nahradíme omezeností druhých derivací funkcí  $f_k(x)$ . Tedy

$$\|G_k(x)\| \leq \bar{G} \quad \forall x \in R^n, \quad 1 \leq k \leq m \quad (\bar{F3}).$$

**Poznámka 60** Podmínka (F1) je splněna vždy, neboť  $F(x) \geq 0 \forall x \in R^n$ . Podmínky (F2) a  $(\bar{F3})$  implikují  $\|f_k(x)\| \leq \bar{f}$  a  $\|g_k(x)\| \leq \bar{g} \forall x \in R^n$ ,  $1 \leq k \leq m$ , a tudíž podmínku (F3).

**Věta 82** *Nechť funkce  $f_k(x) : R^n \rightarrow R$ ,  $1 \leq k \leq m$ , splňují podmínky (F2) a  $(\bar{F3})$ . Pak GN metoda s lokálně omezeným krokem je globálně konvergentní. Je-li navíc splněna podmínka (F4) a platí-li  $x_i \rightarrow x^*$ ,  $F(x^*) = 0$  a  $\omega_i(s_i) \rightarrow 0$ , pak  $x_i \rightarrow x^*$   $Q$ -superlineárně.*

**Poznámka 61** Výpočet směrového vektoru.

- Řešením normální soustavy rovnic

$$J_i^T J_i s_i + J_i^T f_i = 0$$

- Řešením linearizované úlohy nejmenších čtverců (přeurčené soustavy lineárních rovnic)

$$J_i s_i + f_i \approx 0$$



Používá se  $QR$  rozklad

$$J_i = Q_i \begin{bmatrix} R_i \\ 0 \end{bmatrix}, \quad Q_i^T Q_i = I$$

( $R_i$  je horní trojúhelníková matice). Při realizaci GN metody s optimálním lokálně omezeným krokem můžeme soustavu  $(J_i^T J_i + \lambda I)s + J_i^T f_i = 0$  nahradit linearizovanou úlohou

$$\begin{bmatrix} J_i \\ \sqrt{\lambda} I \end{bmatrix} s + \begin{bmatrix} f_i \\ 0 \end{bmatrix} \approx 0$$

- Řešením systémových rovnic. Označíme  $r_i = -(J_i s_i + f_i)$ . Směrový vektor hledáme tak, aby platilo  $J_i^T r_i = 0$ . Tedy

$$\begin{bmatrix} I & J_i \\ J_i^T & 0 \end{bmatrix} \begin{bmatrix} r_i \\ s_i \end{bmatrix} + \begin{bmatrix} f_i \\ 0 \end{bmatrix} = 0$$

což je soustava  $m + n$  rovnic se symetrickou indefinitní maticí (KKT systém, rovnice sedlového bodu). Vhodné pro řídké nebo pro vážené úlohy.

## Hybridní metody pro součet čtverců

**Myšlenka:**

$$F_i \rightarrow F^* = 0 \Rightarrow \text{GN metoda}$$

$$F_i \rightarrow F^* > 0 \Rightarrow \text{VM metoda}$$

**Věta 83** *Nechť  $F_i \rightarrow F^* = 0$   $Q$ -superlineárně. Pak*

$$\lim_{i \rightarrow \infty} \frac{F_i - F_{i+1}}{F_i} = 1$$

*Nechť  $F_i \rightarrow F^* > 0$ . Pak*

$$\lim_{i \rightarrow \infty} \frac{F_i - F_{i+1}}{F_i} = 0$$

**Hybridní metoda:**

Nechť  $B_1 = J_1^T J_1$ . Pro  $i \in N$  pokládáme

$$\begin{aligned} B_{i+1} &= J_{i+1}^T J_{i+1}, & F_i - F_{i+1} &> \underline{\vartheta} F_i \\ B_{i+1} &= B_i + \frac{y_i y_i^T}{y_i^T d_i} - \frac{B_i d_i (B_i d_i)^T}{d_i^T B_i d_i}, & F_i - F_{i+1} &\leq \underline{\vartheta} F_i \end{aligned}$$

kde  $d_i = x_{i+1} - x_i$ ,  $y_i = g_{i+1} - g_i$  (obvykle  $\underline{\vartheta} = 0.01$ ).

## Vektorové diferenční verze Newtonovy metody

Vektorové diferenční verze Newtonovy metody jsou nepřesné metody s lokálně omezeným krokem (Algoritmus 4), kde se násobení maticí  $B = G$  nahradí numerickým výpočtem derivací

$$q = G(x)p \approx \frac{g(x + \delta p) - g(x)}{\delta}$$

kde  $\delta = \sqrt{\varepsilon_M} \|p\|$  ( $\varepsilon_M$  je strojová přesnost).

**Poznámka 62** Vlastnosti:

- Pro nestrukturované úlohy většinou efektivní.
- Pro špatně podmíněné úlohy příliš velký počet gradientů.

## Maticové diferenční verze Newtonovy metody

Předpokládáme, že matice  $G(x)$  je řídká (málo strukturálně nenulových prvků). Je-li matice  $G(x)$  řídká, lze obvykle vypočítat několik sloupců této matice pomocí jedné diference gradientů.

**Příklad 6** *Nechť*

$$G = \begin{bmatrix} G_{11} & G_{12} & 0 & 0 & 0 \\ G_{21} & G_{22} & G_{23} & 0 & 0 \\ 0 & G_{32} & G_{33} & G_{34} & 0 \\ 0 & 0 & G_{43} & G_{44} & G_{45} \\ 0 & 0 & 0 & G_{54} & G_{55} \end{bmatrix}$$

*a*

$$v_1 = [1, 0, 0, 1, 0]^T$$

$$v_2 = [0, 1, 0, 0, 1]^T$$

$$v_3 = [0, 0, 1, 0, 0]^T$$

*Pak*

$$Gv_1 = [G_{11}, G_{21}, G_{34}, G_{44}, G_{54}]^T$$

$$Gv_2 = [G_{12}, G_{22}, G_{32}, G_{45}, G_{55}]^T$$

$$Gv_3 = [0, G_{23}, G_{33}, G_{43}, 0]^T$$

*takže všechny prvky matice  $G$  lze určit pomocí tří diferenčních vzorců*

$$\frac{g(x + \delta v_1) - g(x)}{\delta} \approx Gv_1$$

$$\frac{g(x + \delta v_2) - g(x)}{\delta} \approx Gv_2$$

$$\frac{g(x + \delta v_3) - g(x)}{\delta} \approx Gv_3$$

**Poznámka 63** Je-li matice  $G(x)$  velmi řídká, patří diferenční verze Newtonovy metody pro řídké úlohy k neefektivnějším metodám.

# Metody pro separovatelné úlohy

Separovatelná úloha:

$$F(x) = \sum_{k=1}^m f_k(x)$$

kde  $m = O(n)$  a  $f_k : R^n \rightarrow R$ ,  $1 \leq k \leq m$ , závisí na  $n_k = O(1)$  proměnných. Platí

$$g(x) = \sum_{k=1}^m g_k(x)$$
$$G(x) = \sum_{k=1}^m G_k(x)$$

kde  $g_k(x)$  a  $G_k(x)$  obsahují  $O(1)$  nenulových prvků.

**Označení:**

$N_k$  – množina indexů proměnných funkce  $f_k(x)$  (má  $n_k$  prvků).

$Z_k \in R^{n \times n_k}$  – matice obsahující sloupce jednotkové matice s indexy z  $N_k$ .

$\hat{g}_k = Z_k^T g_k \in R^{n_k}$  – redukované gradienty

$\hat{G}_k = Z_k^T G_k Z_k \in R^{n_k \times n_k}$  – redukované Hessovy matice

**Příklad 7**

$$F(x) = (x_1 - 1)^2 + \sum_{i=2}^n (x_i - x_{i-1})^2$$

$$f_1(x) = (x_1 - 1)^2$$

$$f_i(x) = (x_i - x_{i-1})^2, \quad 2 \leq i \leq n$$

$$Z_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix}, \quad Z_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \dots & \dots \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \dots, \quad Z_n = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dots & \dots \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\hat{g}_1 = [ 2(x_1 - 1) ], \quad \hat{G}_1 = [ 2 ]$$

$$\hat{g}_i = \begin{bmatrix} -2(x_i - x_{i-1}) \\ 2(x_i - x_{i-1}) \end{bmatrix}, \quad \hat{G}_i = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

## Diferenční verze Newtonovy metody pro separovatelné úlohy

Redukované Hessovy matice se počítají numericky pomocí diferencí redukovaných gradientů

$$\hat{G}_k(x)\hat{e}_j \approx \frac{\hat{g}_k(x + \delta Z_k \hat{e}_j) - \hat{g}_k(x)}{\delta}$$

kde  $\hat{e}_j$ ,  $1 \leq j \leq n_k$ , jsou sloupce jednotkové matice řádu  $n_k$ . K určení prvků všech redukovaných Hessových matic je zapotřebí

$$\sum_{k=1}^m n_k^2 = mO(1) = O(n)$$

operací. Výsledná Hessova matice

$$G(x) = \sum_{k=1}^m Z_k \hat{G}_k(x) Z_k^T$$

## Metody s proměnnou metrikou pro separovatelné úlohy

Místo redukovaných Hessových matic se používají jejich aproximace  $\hat{B}_k$ ,  $1 \leq k \leq m$ , které se aktualizují pomocí metod s proměnnou metrikou

$$\hat{B}_k^+ = \frac{1}{\hat{\gamma}_k} \left( \hat{B}_k + \frac{\hat{\gamma}_k}{\hat{\rho}_k} \frac{1}{\hat{b}_k} \hat{y}_k \hat{y}_k^T - \frac{1}{\hat{c}_k} \hat{B}_k \hat{d}_k (\hat{B}_k \hat{d}_k)^T + \frac{\hat{\beta}_k}{\hat{c}_k} \left( \frac{\hat{c}_k}{\hat{b}_k} \hat{y}_k - \hat{B}_k \hat{d}_k \right) \left( \frac{\hat{c}_k}{\hat{b}_k} \hat{y}_k - \hat{B}_k \hat{d}_k \right)^T \right)$$

kde

$$\begin{aligned} \hat{d}_k &= Z_k^T d, \\ \hat{y}_k &= Z_k^T (g_k^+ - g_k) = \hat{g}_k^+ - \hat{g}_k, \end{aligned}$$

$\hat{b}_k = \hat{y}_k^T \hat{d}_k$ ,  $\hat{c}_k = \hat{d}_k^T \hat{B}_k \hat{d}_k$  a  $\hat{\gamma}_k$ ,  $\hat{\rho}_k$ ,  $\hat{\beta}_k$  jsou volené parametry ( $\hat{\beta}_k = 0$  pro metodu BFGS). Aproximace Hessovy matice

$$B = \sum_{k=1}^m Z_k \hat{B}_k Z_k^T$$

Směrový vektor se získá řešením soustavy rovnic  $Bs = -g$ , kde

$$g = \sum_{k=1}^m g_k = \sum_{k=1}^m Z_k \hat{g}_k$$

**Poznámka 64** Nelze zaručit, aby platilo  $\hat{b}_k = \hat{y}_k^T \hat{d}_k > 0 \forall 1 \leq k \leq m \Rightarrow$  některá z matic  $\hat{B}_k$  nemusí být pozitivně definitní  $\Rightarrow B$  nemusí být pozitivně definitní  $\Rightarrow$  nelze dokázat globální konvergenci pro metodu spádových směrů.

**Věta 84** (*Superlineární konvergence*) *Nechť  $\{x_i\} \subset R^n$  je posloupnost získaná VM metodou pro separovatelné úlohy s  $\gamma_k^i = \rho_k^i = 1$ ,  $1 \leq k \leq m$ ,  $\forall i \in N$ , taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je stacionární bod funkce  $F \in C^2 : R^n \rightarrow R$ , která vyhovuje podmínkám (F3)-(F5). Pokud  $0 \leq \eta_k^i \leq 1$ ,  $1 \leq k \leq m$ ,  $\forall i \in N$ , a pokud  $\alpha_i = 1$ , jestliže tato hodnota splňuje podmínku (S2) a (S3b), pak  $x \rightarrow x^*$  Q-superlineárně.*

**Poznámka 65** Metody s proměnnou metrikou pro separovatelné úlohy jsou obvykle velmi efektivní. Je výhodné kombinovat metodou BFGS

$$\begin{aligned}\hat{B}_k^+ &= \hat{B}_k + \frac{1}{\hat{b}_k} \hat{y}_k \hat{y}_k^T - \frac{1}{\hat{c}_k} \hat{B}_k \hat{d}_k (\hat{B}_k \hat{d}_k)^T, & \hat{b}_k > 0 \\ \hat{B}_k^+ &= \hat{B}_k, & \hat{b}_k \leq 0\end{aligned}$$

s metodou hodnosti 1

$$\begin{aligned}\hat{B}_k^+ &= \hat{B}_k + \frac{1}{\hat{b}_k - \hat{c}_k} (\hat{y}_k - \hat{B}_k \hat{d}_k) (\hat{y}_k - \hat{B}_k \hat{d}_k)^T, & |\hat{b}_k - \hat{c}_k| \neq 0 \\ \hat{B}_k^+ &= \hat{B}_k, & |\hat{b}_k - \hat{c}_k| = 0\end{aligned}$$

Na metodu hodnosti 1 přecházíme, pokud z hodnot  $b_k$ ,  $1 \leq k \leq m$ , je alespoň  $m/2$  nekladných.

## Gaussova-Newtonova metoda pro separovatelný součet čtverců

Uvažujme součet čtverců

$$F(x) = \frac{1}{2} \sum_{k=1}^m f_k^2(x)$$

kde  $m = O(n)$  a  $f_k : R^n \rightarrow R$ ,  $1 \leq k \leq m$ , závisí na  $n_k = O(1)$  proměnných. Platí

$$g(x) = \sum_{k=1}^m f_k(x) g_k(x) = \sum_{k=1}^m f_k(x) Z_k \hat{g}_k(x)$$

$$B(x) = J^T(x) J(x) = \sum_{k=1}^m g_k(x) g_k^T(x) = \sum_{k=1}^m Z_k \hat{g}_k(x) \hat{g}_k^T(x) Z_k^T = \sum_{k=1}^m Z_k \hat{B}_k(x) Z_k^T$$

kde  $\hat{B}_k(x) = \hat{g}_k(x) \hat{g}_k^T(x)$  je redukovaná Gaussova-Newtonova matice. V každé iteraci se řeší soustava rovnic  $B(x_i) s_i = -g(x_i)$ .

## Hybridní metody pro separovatelný součet čtverců

- Kombinace GN metody s Newtonovou metodou pro separovatelné úlohy:

Nechť  $\hat{B}_k^1 = \hat{g}_k^1 (\hat{g}_k^1)^T$ ,  $1 \leq k \leq m$ . Pro  $i \in N$  pokládáme

$$\begin{aligned}\hat{B}_k^{i+1} &= \hat{g}_k^{i+1} (\hat{g}_k^{i+1})^T, & F_i - F_{i+1} > \underline{\varrho} F_i \\ \hat{B}_k^{i+1} &= \hat{g}_k^{i+1} (\hat{g}_k^{i+1})^T + f_k^{i+1} \hat{C}_k^{i+1}, & F_i - F_{i+1} \leq \underline{\varrho} F_i\end{aligned}$$

Dále pokládáme

$$\begin{aligned}B_i &= \sum_{k=1}^m Z_k \hat{B}_k^i Z_k^T \\ g_i &= \sum_{k=1}^m f_k^i Z_k \hat{g}_k^i\end{aligned}$$

a řešíme soustavu rovnic  $B_i s_i = -g_i$ .

- Kombinace GN metody s VM metodou hodnosti 1:  
Nechť  $\hat{B}_k^1 = \hat{g}_k^1(\hat{g}_k^1)^T$  a  $\hat{C}_k^1 = I$ ,  $1 \leq k \leq m$ . Pro  $i \in N$  pokládáme

$$\begin{aligned}\hat{C}_k^{i+1} &= \hat{C}_k^i + \frac{\hat{w}_k^i(\hat{w}_k^i)^T}{(\hat{d}_k^i)^T \hat{w}_k^i}, & |(\hat{d}_k^i)^T \hat{w}_k^i| > \underline{\delta} \\ \hat{C}_k^{i+1} &= \hat{C}_k^i, & |(\hat{d}_k^i)^T \hat{w}_k^i| \leq \underline{\delta}\end{aligned}$$

a

$$\begin{aligned}\hat{B}_k^{i+1} &= \hat{g}_k^{i+1}(\hat{g}_k^{i+1})^T, & F_i - F_{i+1} > \underline{\vartheta}F_i \\ \hat{B}_k^{i+1} &= \hat{g}_k^{i+1}(\hat{g}_k^{i+1})^T + f_k^{i+1}\hat{C}_k^{i+1}, & F_i - F_{i+1} \leq \underline{\vartheta}F_i\end{aligned}$$

kde  $\hat{d}_k^i = Z_k^T(x_{i+1} - x_i)$ ,  $\hat{g}_k^i = Z_k^T(g_{i+1} - g_i)$  a  $\hat{w}_k^i = \hat{g}_k^i - \hat{C}_k^i \hat{d}_k^i$ .

**Poznámka 66** Hybridní metody pro separovatelný součet čtverců jsou superlineárně konvergentní. Jsou-li realizovány jeho metody s lokálně omezeným krokem, jsou globálně konvergentní.

## Metody s proměnnou metrikou s omezenou pamětí (vektorový tvar)

**Myšlenka:**

V  $i$ -tém iteračním kroku položíme  $H_1 = I$  a pomocí (nejvýše)  $m$  dvojic vektorů  $d_{i-m}, y_{i-m}, \dots, d_{i-1}, y_{i-1}$  konstruujeme vektory  $H_1 g_i, \dots, H_{m+1} g_i$  (matice  $H_1, \dots, H_{m+1}$  se neukládají). Nakonec pokládáme  $s_i = -H_{m+1} g_i$  ( $m$  je obvykle malé).

**Definice 46** Jsou-li matice  $H_1, \dots, H_{m+1}$  získány metodou s proměnnou metrikou, budeme mluvit o metodě s proměnnou metrikou s omezenou pamětí ( $m$ -krokovou metodou s proměnnou metrikou).

**Poznámka 67** Pro jednoduchost budeme předpokládat, že  $i \leq m+1$  a že matice  $H_1, \dots, H_i$  jsou konstruovány z vektorů  $d_1, y_1, \dots, d_{i-1}, y_{i-1}$  pomocí metody BFGS, jejíž aktualizaci lze zapsat ve tvaru

$$H_{k+1} = \gamma_k V_k^T H_k V_k + \frac{\rho_k}{b_k} d_k d_k^T$$

$1 \leq k \leq i$ , kde  $V_k = I - y_k d_k^T / b_k$  a  $b_k = y_k^T d_k$  ( $d_k = x_{k+1} - x_k$ ,  $y_k = g_{k+1} - g_k$ ). Dále budeme předpokládat, že  $\gamma_k = 1$  pro  $k > 1$ .

**Věta 85** Pro  $m$ -krokovou metodu BFGS platí

$$H_{k+1} = \gamma_1 \left( \prod_{j=1}^k V_j \right)^T \left( \prod_{j=1}^k V_j \right) + \sum_{l=1}^k \frac{\rho_l}{b_l} \left( \prod_{j=l+1}^k V_j \right)^T d_l d_l^T \left( \prod_{j=l+1}^k V_j \right)$$

(Předpokládáme, že  $1 \leq k \leq m$ ).

**Věta 86** (*Strangova formule*) Vektor  $s_i = -H_i g_i$  se určí tak, že se pomocí zpětné rekurence počítají vektory

$$\begin{aligned} u_i &= -g_i \\ u_l &= u_{l+1} - \sigma_l y_l, \quad \sigma_l = d_l^T u_{l+1} / b_l, \quad 1 \leq l \leq i-1 \end{aligned}$$

Pak se pomocí přímé rekurence počítají vektory

$$\begin{aligned} v_1 &= \gamma_1 u_1 \\ v_{k+1} &= v_k + (\rho_k \sigma_k - y_k^T v_k / b_k) d_k, \quad 1 \leq k \leq i-1 \end{aligned}$$

Nakonec se položí  $s_i = v_i$ .

**Poznámka 68** Platí

$$\begin{aligned} u_{l+1} &= - \left( \prod_{j=l+1}^{i-1} V_j \right) g_i \\ v_{k+1} &= \gamma_1 \left( \prod_{j=1}^k V_j \right)^T u_1 + \sum_{l=1}^k \frac{\rho_l}{b_l} \left( \prod_{j=l+1}^k V_j \right)^T d_l d_l^T u_{l+1} \end{aligned}$$

## Metody s proměnnou metrikou s omezenou pamětí (maticový tvar)

Označení:

$$\begin{aligned} D_k &= [d_1, d_2, \dots, d_k] \\ Y_k &= [y_1, y_2, \dots, y_k] \\ L_k &= \begin{bmatrix} d_1^T y_1 & 0 & \dots & 0 \\ d_2^T y_1 & d_2^T y_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ d_k^T y_1 & d_k^T y_2 & \dots & d_k^T y_k \end{bmatrix} \\ C_k &= \text{diag}(d_1^T y_1, \dots, d_k^T y_k) \end{aligned}$$

**Věta 87** Pro  $m$ -krokovou metodu DFP platí

$$B_{k+1} = B_1 - [Y_k, B_1 D_k] \begin{bmatrix} 0, & L_k^T \\ L_k, & C_k - D_k^T B_1 D_k \end{bmatrix}^{-1} \begin{bmatrix} Y_k^T \\ D_k^T B_1 \end{bmatrix}$$

Pro  $m$ -krokovou metodu BFGS platí

$$B_{k+1} = B_1 - [Y_k, B_1 D_k] \begin{bmatrix} -C_k, & L_k^T - C_k^T \\ L_k - C_k, & D_k^T B_1 D_k \end{bmatrix}^{-1} \begin{bmatrix} Y_k^T \\ D_k^T B_1 \end{bmatrix}$$

Pro  $m$ -krokovou metodu hodnosti 1 platí

$$B_{k+1} = B_1 + (Y_k - B_1 D_k)(L_k + L_k^T - C_k - D_k^T B_1 D_k)^{-1}(Y_k - B_1 D_k)^T$$

(Předpokládáme, že  $1 \leq k \leq m$ ).

**Poznámka 69** Podobné vzorce platí i pro matici  $H_{k+1}$ . V tomto případě je však použití Strangovy formule efektivnější.

**Poznámka 70** Matice určené podle Věty 87 lze použít pro výpočet lokálně omezeného kroku nebo v KKT systémech.

**Poznámka 71** Obvykle se pokládá  $B_1 = (1/\gamma_k)I$ . Lze též použít libovolnou řídkou matici  $B_1$ .

## Metody pro řešení soustav nelineárních rovnic

Nechť  $f : R^n \rightarrow R^n$  je spojitě diferencovatelné zobrazení. Budeme hledat bod  $x^* \in R^n$  takový, že  $f(x^*) = 0$ .

**Poznámka 72** Řešení soustav nelineárních rovnic má blízký vztah k minimalizaci funkcí.

- Minimum funkce  $F : R^n \rightarrow R$  můžeme hledat řešením soustavy rovnic  $g(x) = 0$ . Může to však vést k nalezení sedlového bodu nebo maxima.
- Řešení soustavy rovnic  $f(x) = 0$  můžeme hledat jako minimum součtu čtverců  $F(x) = \frac{1}{2}f^T(x)f(x)$ . Může to však vést k nalezení lokálního minima, kdy  $F(x) \neq 0$ .

**Věta 88** *Nechť  $f \in C^1 : R^n \rightarrow R^n$  a nechť bod  $x^* \in R^n$  je lokálním minimem funkce  $F(x) = \frac{1}{2}f^T(x)f(x)$ , přičemž Jacobiova matice  $J(x^*)$  zobrazení  $f$  v bodě  $x^*$  je regulární. Pak platí  $f(x^*) = 0$ .*

**Definice 47** *Zobrazení  $f : R^n \rightarrow R^n$  má stejnoměrně omezené derivace, existuje-li  $\bar{J} > 0$  tak, že*

$$\|J(x)\| \leq \bar{J} \quad \forall x \in R^n \quad (\text{J3})$$

**Definice 48** *Zobrazení  $f : R^n \rightarrow R^n$  je stejnoměrně regulární, existuje-li  $\underline{J} > 0$  tak, že*

$$\|J^{-1}(x)\| \leq \frac{1}{\underline{J}} \quad \forall x \in R^n \quad (\text{J4})$$

**Definice 49** *Zobrazení  $f : R^n \rightarrow R^n$  má lipschitzovské derivace, existuje-li  $\bar{L} > 0$  tak, že*

$$\|J(x+d) - J(x)\| \leq \bar{L}\|d\| \quad \forall x, d \in R^n \quad (\text{J5})$$

**Definice 50** *Základní metodou pro řešení nelineárních rovnic nazýváme iterační proces, jehož výsledkem je posloupnost  $\{x_i\} \subset R^n$  taková, že*

$$x_{i+1} = x_i + \alpha_i s_i$$

*kde směrový vektor  $s_i \in R^n$  se určuje na základě hodnot  $x_j, f_j, J_j, 1 \leq j \leq i$ , a délka kroku  $\alpha_i > 0$  se určuje na základě chování zobrazení  $f : R^n \rightarrow R^n$  v okolí bodu  $x_i \in R^n$ .*

**Definice 51** *Řekneme, že základní metoda pro řešení nelineárních rovnic je globálně konvergentní, jestliže pro libovolný počáteční bod  $x_1 \in R^n$  platí*

$$\liminf_{i \rightarrow \infty} \|f(x_i)\| = 0$$



**Poznámka 73** Newtonova metoda používá směrové vektory  $s_i = -J^{-1}(x_i)f(x_i) \forall i \in N$ . Platí-li (J4), je tento směrový vektor shodný se směrovým vektorem Gaussovy-Newtonovy metody pro minimalizaci součtu čtverců  $F(x) = \frac{1}{2}f^T(x)f(x)$ , neboť v tomto případě platí

$$(J^T(x_i)J(x_i))^{-1}g(x_i) = (J^T(x_i)J(x_i))^{-1}J^T(x_i)f(x_i) = J^{-1}(x_i)f(x_i)$$

**Poznámka 74** Omezíme se na metody takové, že  $s_i = -A_i^{-1}f_i$ , kde matice  $A_i$ ,  $i \in N$ , splňují podmínky

$$\|A_i s\| \leq \bar{A}s \quad \forall s \in R^n, \quad (\text{A3})$$

$$\|A_i s\| \geq \underline{A}s \quad \forall s \in R^n, \quad (\text{A4})$$

$$\|A_i - J_i\| \leq \bar{\vartheta}. \quad (\text{A5})$$

kde  $0 < \underline{A} \leq \bar{A}$  a  $\bar{\vartheta} \geq 0$ , a budeme používat označení  $h_i = A_i^T f_i$ . Pokud  $\bar{\vartheta} < \underline{J}$ , pak z (J3), (J4) a (A5) plyne (A3), (A4) s  $\bar{A} = \bar{J} + \bar{\vartheta}$ ,  $\underline{A} = \underline{J} - \bar{\vartheta}$ .

**Definice 52** Základní metoda pro řešení nelineárních rovnic (Definice 50) je metodou spádových směrů, jestliže se směrové vektory  $s_i \in R^n$ ,  $i \in N$ , určují tak, že

$$\|A_i s_i + f_i\| \leq \bar{\omega} \|f_i\| \quad (\overline{\text{S1}})$$

kde  $0 \leq \bar{\omega} < 1$ , a délky kroku  $\alpha_i > 0$ ,  $i \in N$ , se určují tak, že  $\alpha_i > 0$  je první člen vyhovující podmínce

$$F_{i+1} - F_i \leq -\underline{\rho} \alpha_i F_i \quad (\overline{\text{S2}})$$

v posloupnosti  $\alpha_i^j$ ,  $j \in N$ , takové, že  $\alpha_i^1 = 1$  a  $\underline{\beta} \alpha_i^j \leq \alpha_i^{j+1} \leq \bar{\beta} \alpha_i^j$ , kde  $0 < \underline{\beta} \leq \bar{\beta} < 1$ ,  $0 < \underline{\rho} \leq 1/2$ .

**Věta 89** (Konzistence) Nechť zobrazení  $f : R^n \rightarrow R^n$  splňuje podmínky (J3)-(J5). Nechť matice  $A_i$ ,  $i \in N$ , splňují podmínku (A5) s  $\bar{\vartheta} < (1 - \bar{\omega})\underline{J}/2$ . Pak lze v každém iteračním kroku nalézt směrový vektor  $s_i \in R^n$  vyhovující podmínce  $(\overline{\text{S1}})$  a délku kroku  $\alpha_i > 0$  vyhovující podmínce  $(\overline{\text{S2}})$ . Navíc existuje konstanta  $0 < \underline{\alpha} < 1$  taková, že  $\alpha_i \geq \underline{\alpha} \forall i \in N$ .

**Poznámka 75** Podmínky (A3)-(A5) mají pouze teoretický význam (neznáme-li  $J_i$ , nemůžeme (A5) ověřit). Dávají však podklad pro použití restartů. Nelze-li splnit podmínky  $(\overline{\text{S1}})$  a  $(\overline{\text{S2}})$ , pokládáme  $A_i = J_i$ . Pak lze položit  $\bar{A} = \bar{J}$ ,  $\underline{A} = \underline{J}$ ,  $\bar{\vartheta} = 0$  a najít  $s_i \in R^n$  a  $\alpha_i > 0$  vyhovující podmínkám  $(\overline{\text{S1}})$  a  $(\overline{\text{S2}})$ .

**Věta 90** (Globální konvergence) Nechť jsou splněny předpoklady Věty 89. Nechť  $\{x_i\} \subset R^n$  je posloupnost generovaná metodou spádových směrů (takže platí  $(\overline{\text{S1}})$  a  $(\overline{\text{S2}})$ ). Pak  $x_i \rightarrow x^*$  a  $f(x^*) = 0$ .

**Věta 91** (Superlineární konvergence) Nechť  $\{x_i\} \subset R^n$  je posloupnost bodů generovaná metodou spádových směrů taková, že  $x_i \rightarrow x^*$ , kde  $x^* \in R^n$  je nulovým bodem zobrazení  $f \in C^1 : R^n \rightarrow R^n$ , které vyhovuje podmínkám (J3) a (J4). Nechť

$$\lim_{i \rightarrow \infty} \frac{\|A_i s_i + f_i\|}{\|f_i\|} = 0$$

$$\lim_{i \rightarrow \infty} \frac{\|(A_i - J_i)s_i\|}{\|s_i\|} = 0$$

a nechť  $\alpha_i = 1$ , kdykoliv tato hodnota vyhovuje podmínce  $(\overline{\text{S2}})$ . Pak existuje index  $k \in N$  takový, že  $\alpha_i = 1 \forall i \geq k$  a  $x_i \rightarrow x^*$  Q-superlineárně.

**Poznámka 76** Pro řešení nelineárních rovnic existují též metody s lokálně omezeným krokem. Jelikož Newtonovu metodu lze bez problémů realizovat jako metodu spádových směrů, nebudeme se jimi zabývat.

**Poznámka 77** Newtonova metoda používá matice  $A_i = J_i$ ,  $i \in N$ , takže platí (A3)–(A5) s  $\bar{A} = \bar{J}$ ,  $\underline{A} = \underline{J}$  a  $\bar{\vartheta} = 0$ . Splňuje-li zobrazení  $f : R^n \rightarrow R^n$  podmínky (J3)–(J4), je Newtonova metoda globálně konvergentní. Jestliže  $\|J_i s_i + f_i\|/\|f_i\| \rightarrow 0$ , je tato metoda Q-superlineárně konvergentní.

## Diferenční verze Newtonovy metody

Matice  $A_i$ ,  $i \in N$ , se volí tak, že

$$A_i e_j = \frac{f(x_i + \delta e_j) - f(x_i)}{\delta}, \quad 1 \leq j \leq n$$

kde  $e_j$ ,  $1 \leq j \leq n$ , jsou sloupce jednotkové matice řádu  $n$ .

**Věta 92** *Nechť zobrazení  $f : R^n \rightarrow R^n$  splňuje podmínky (J3)–(J5) a nechť*

$$\delta \leq \frac{(1 - \bar{\omega})\underline{J}}{\bar{L}\sqrt{n}}$$

*Pak matice  $A_i$ ,  $i \in N$ , získané diferenční verzí Newtonovy metody vyhovují podmínkám (A3)–(A5) s  $\bar{\vartheta} < (1 - \bar{\omega})\underline{J}/2$ ,  $\underline{A} > \underline{J} - (1 - \bar{\omega})\underline{J}/2$  a  $\bar{A} < \bar{J} + (1 - \bar{\omega})\underline{J}/2$ .*

**Poznámka 78** Obvykle se pokládá  $\delta = \sqrt{\varepsilon_M}$ , kde  $\varepsilon_M$  je strojová přesnost ( $\approx 10^{-16}$  v dvojnásobné aritmetice). Problémy mohou nastat, jsou-li  $\underline{J}$  malé a  $\bar{J}$ ,  $\bar{L}$  velká.

## Kvazinevtonovské metody

**Definice 53** *Řekneme, že základní metoda pro řešení nelineárních rovnic je kvazinevtonovskou metodou (QN), jestliže*

$$A_i s_i + f_i = 0$$

*kde  $A_i$ ,  $i \in N$ , jsou regulární matice konstruované tak, že*

$$A_{i+1} = A_i + u_i v_i^T$$

*kde  $u_i \in R^n$ ,  $v_i \in R^n$  se vybírají tak, aby byla splněna kvazinevtonovská podmínka*

$$A_{i+1} d_i = y_i$$

*kde  $d_i = x_{i+1} - x_i$  a  $y_i = f_{i+1} - f_i$ .*

**Věta 93** *Nechť  $A_+ = A + uv^T$ . Pak  $A_+ d = y$  právě tehdy, jestliže  $v^T d \neq 0$  a  $u = (y - Ad)/v^T d$ , takže*

$$A_+ = A + \frac{(y - Ad)v^T}{v^T d} \quad (\bar{A})$$

**Poznámka 79** Nejznámější QN metody:

- Broydenova dobrá metoda:  $v = d$

$$A_+ = A + \frac{(y - Ad)d^T}{d^T d}$$

- Broydenova špatná metoda:  $v = A^T y$

$$A_+ = A + \frac{(y - Ad)y^T A}{y^T Ad}$$

- Metoda aktualizace sloupců:  $v = e_k$

$$A_+ = A + \frac{(y - Ad)e_k^T}{e_k^T d}$$

Vektor  $e_k \in R^n$  se vybírá tak, že

$$e_k^T d = \max_{1 \leq i \leq n} e_i^T d$$

( $e_i, 1 \leq i \leq n$ , jsou sloupce jednotkové matice).

**Věta 94** *Nechť matice  $A$  je regulární a platí  $(\bar{A})$ . Pak matice  $A_+$  je regulární právě tehdy, jestliže  $v^T A^{-1} y \neq 0$ .*

**Věta 95** *(Inverzní QN metody) Nechť jsou splněny předpoklady věty 94 a necht'  $S = A^{-1}$  a  $S_+ = A_+^{-1}$  ( $A_+$  je matice určená podle  $(\bar{A})$ ). Pak platí*

$$S_+ = S + \frac{(d - Sy)v^T S}{v^T Sy} \quad (\bar{S})$$

**Poznámka 80** Inverzní QN metody:

- Broydenova dobrá metoda:  $v = d$

$$S_+ = S + \frac{(d - Sy)d^T S}{d^T Sy}$$

- Broydenova špatná metoda:  $S^T v = y$

$$S_+ = S + \frac{(d - Sy)y^T}{y^T y}$$

- Inverzní metoda aktualizace sloupců:  $S^T v = e_k$

$$S_+ = S + \frac{(d - Sy)e_k^T}{e_k^T y}$$

Vektor  $e_k \in R^n$  se vybírá tak, že

$$e_k^T y = \max_{1 \leq i \leq n} e_i^T y$$

**Věta 96** *Nechť zobrazení  $f : R^n \rightarrow R^n$  vyhovuje podmínkám (J3)-(J5) a necht'  $x^* \in R^n$  je bod takový, že  $f(x^*) = 0$ . Pak existují čísla  $\bar{\delta} > 0, \bar{\vartheta} > 0$  taková, že pokud  $\|x_1 - x^*\| \leq \bar{\delta}$  a  $\|A_1 - J_1\| \leq \bar{\vartheta}$  a posloupnost  $\{x_i\} \subset R^n$  je určena dobrou Broydenovou metodou s jednotkovým výběrem délky kroku ( $\alpha_i = 1 \forall i \in N$ ), pak  $x_i \rightarrow x^*$  Q-superlineárně.*

# Odhad parametrů systémů diferenciálních rovnic

Máme minimalizovat funkci

$$F(x) = \int_{t_0}^{t_1} F_A(x, y(x, t), t) dt + F_T(x, y(x, t_1)),$$

kde

$$\frac{dy(x, t)}{dt} = f_S(x, y(x, t), t), \quad y(x, t_0) = f_I(x)$$

(stavový systém). Přitom  $x \in R^n$ , a stavový systém obsahuje  $n_S$  diferenciálních rovnic.

Odstranění integrálu:

$$F(x) = \tilde{F}_A(x, t_1) + F_T(x, y(x, t_1)),$$

kde

$$\frac{dy(x, t)}{dt} = f_S(x, y(x, t), t), \quad y(x, t_0) = f_I(x)$$

$$\frac{d\tilde{F}_A(x, t)}{dt} = F_A(x, y(x, t), t), \quad \tilde{F}_A(x, t_0) = 0$$

Řešíme  $n_S + 1$  diferenciálních rovnic v přímém směru.

Předpoklady:

(A1) Existuje spojitě řešení stavového systému na intervalu  $[t_0, t_1]$  kdykoliv  $x \in X$ .

(A2) Funkce  $F_A, F_T, f_S, f_I$  jsou dvakrát spojitě diferencovatelné na  $X$ .

Přitom  $X \subset R^n$  je oblast obsahující všechny body  $x_i \in R^n$   $i \in N$ , získané během iteračního procesu.

## Přímý výpočet gradientu

Označíme

$$u(x, t) = \frac{dy(x, t)}{dx}.$$

Derivováním dostaneme

$$g^T(x) = \tilde{g}_A^T(x, t_1) + \frac{\partial F_T(x, y(x, t_1))}{\partial x} + \frac{\partial F_T(x, y(x, t_1))}{\partial y} u(x, t_1),$$

kde

$$\frac{du(x, t)}{dt} = \frac{\partial f_S(x, y, t)}{\partial x} + \frac{\partial f_S(x, y, t)}{\partial y} u(x, t), \quad u(x, t_0) = \frac{df_I(x)}{dx},$$

$$\frac{d\tilde{g}_A^T(x, t)}{dt} = \frac{\partial F_A(x, y, t)}{\partial x} + \frac{\partial F_A(x, y, t)}{\partial y} u(x, t), \quad \tilde{g}_A^T(x, t_0) = 0.$$

Řešíme navíc  $n(n_S + 1)$  diferenciálních rovnic v přímém směru.

## Zpětný výpočet gradientu

Nechť  $p(t)$  je libovolné zobrazení a  $y(x, t)$  je řešení stavového systému, takže

$$f_S(x, y, t) - \frac{dy(x, t)}{dt} = 0$$

Můžeme tedy psát

$$F(x) = \int_{t_0}^{t_1} \left[ F_A(x, y, t) + p^T(t) \left( f_S(x, y, t) - \frac{dy(x, t)}{dt} \right) \right] dt + F_T(x, y(x, t_1))$$

Integrovaní per partes  $u^T v' = (u^T v)' - (u^T)' v$ , kde  $u = p(t)$ ,  $v = y(x, t)$

$$\begin{aligned} F(x) &= \int_{t_0}^{t_1} \left[ F_A(x, y, t) + p^T(t) f_S(x, y, t) + \frac{dp^T(t)}{dt} y(x, t) \right] dt \\ &+ p^T(t_0) y(x, t_0) - p^T(t_1) y(x, t_1) + F_T(x, y(x, t_1)). \end{aligned}$$

Derivováním dostaneme

$$\begin{aligned} g^T(x) &= \int_{t_0}^{t_1} \left[ \frac{\partial F_A(x, y, t)}{\partial x} + p^T(t) \frac{\partial f_S(x, y, t)}{\partial x} \right. \\ &+ \left. \left( \frac{\partial F_A(x, y, t)}{\partial y} + p^T(t) \frac{\partial f_S(x, y, t)}{\partial y} + \frac{dp^T(t)}{dt} \right) \frac{dy(x, t)}{dx} \right] dt \\ &+ p^T(t_0) \frac{df_I(x)}{dx} + \frac{\partial F_T(x, y(x, t_1))}{\partial x} + \left( \frac{\partial F_T(x, y(x, t_1))}{\partial y} - p^T(t_1) \right) \frac{dy(x, t_1)}{dx}. \end{aligned}$$

Funkci  $p(t) = p(x, t)$  volíme tak, aby se kulaté závorky vynulovaly. Tedy

$$-\frac{dp^T(x, t)}{dt} = \frac{\partial F_A(x, y, t)}{\partial y} + p^T(x, t) \frac{\partial f_S(x, y, t)}{\partial y}, \quad p^T(x, t_1) = \frac{\partial F_T(x, y(x, t_1))}{\partial y},$$

$$g^T(x) = \int_{t_0}^{t_1} \left[ \frac{\partial F_A(x, y, t)}{\partial x} + p^T(x, t) \frac{\partial f_S(x, y, t)}{\partial x} \right] dt + p^T(x, t_0) \frac{df_I(x)}{dx} + \frac{\partial F_T(x, y(x, t_1))}{\partial x}.$$

Odstranění integrálu

$$g(x) = \tilde{g}_A(x, t_0) + \left( \frac{df_I(x)}{dx} \right)^T p(x, t_0) + \left( \frac{\partial F_T(x, y(x, t_1))}{\partial x} \right)^T,$$

kde

$$-\frac{dp(x, t)}{dt} = \left( \frac{\partial F_A(x, y, t)}{\partial y} \right)^T + \left( \frac{\partial f_S(x, y, t)}{\partial y} \right)^T p(x, t), \quad p(x, t_1) = \left( \frac{\partial F_T(x, y(x, t_1))}{\partial y} \right)^T,$$

$$\frac{d\tilde{g}_A(x, t)}{dt} = \left( \frac{\partial F_A(x, y, t)}{\partial x} \right)^T + \left( \frac{\partial f_S(x, y, t)}{\partial x} \right)^T p(x, t), \quad \tilde{g}_A(x, t_1) = 0.$$

Řešíme navíc  $n_S + n$  diferenciálních rovnic ve zpětném směru.

## Přímý výpočet Hessovy matice

Označíme

$$v(x, t) = \frac{du(x, t)}{dx} = \frac{d^2y(x, t)}{dx^2}.$$

Derivováním dostaneme

$$\begin{aligned} G(x) &= \tilde{G}_A(x, t_1) + \frac{\partial^2 F_T(x, y(x, t_1))}{\partial x^2} \\ &+ \left[ 2 \frac{\partial^2 F_T(x, y(x, t_1))}{\partial x \partial y} + \frac{\partial^2 F_T(x, y(x, t_1))}{\partial y^2} u(x, t_1) \right] u(x, t_1) \\ &+ \frac{\partial F_T(x, y(x, t_1))}{\partial y} v(x, t_1), \end{aligned}$$

kde

$$\begin{aligned} \frac{dv(x, t)}{dt} &= \frac{\partial^2 f_S(x, y, t)}{\partial x^2} + \left[ 2 \frac{\partial^2 f_S(x, y, t)}{\partial x \partial y} + \frac{\partial^2 f_S(x, y, t)}{\partial y^2} u(x, t) \right] u(x, t) \\ &+ \frac{\partial f_S(x, y, t)}{\partial y} v(x, t), \quad v(x, t_0) = \frac{d^2 f_I(x)}{dx^2} \end{aligned}$$

$$\begin{aligned} \frac{d\tilde{G}_A(x, t)}{dt} &= \frac{\partial^2 F_A(x, y, t)}{\partial x^2} + \left[ 2 \frac{\partial^2 F_A(x, y, t)}{\partial x \partial y} + \frac{\partial^2 F_A(x, y, t)}{\partial y^2} u(x, t) \right] u(x, t) \\ &+ \frac{\partial F_A(x, y, t)}{\partial y} v(x, t), \quad \tilde{G}_A(x, t_0) = 0 \end{aligned}$$

Řešíme navíc  $n^2(n_S + 1)$  diferenciálních rovnic v přímém směru.

## Automatické derivování

Existují dva postupy

- Přímé automatické derivování – výpočet derivace zobrazení
- Zpětné automatické derivování – výpočet gradientu funkce

Ukážeme na příkladech použití obou postupů

**Příklad 8** (přímé derivování). Máme nalézt derivaci funkce  $F(x) = x_1 \sin(x_2 x_3 + x_4)$  podle proměnné  $x_3$ .

$x_1$	$=$	$x_1$	$F = x_1 \sin(x_2 x_3 + x_4)$
$x'_1$	$=$	$0$	
$x_2$	$=$	$x_2$	
$x'_2$	$=$	$0$	
$x_3$	$=$	$x_3$	
$x'_3$	$=$	$1$	
$x_4$	$=$	$x_4$	
$x'_4$	$=$	$0$	
$x_5$	$=$	$x_2 x_3$	$= x_2 x_3$
$x'_5$	$=$	$x_2 x'_3 + x'_2 x_3$	$= x_2$
$x_6$	$=$	$x_4 + x_5$	$= x_4 + x_2 x_3$
$x'_6$	$=$	$x'_4 + x'_5$	$= x_2$
$x_7$	$=$	$\sin(x_6)$	$= \sin(x_4 + x_2 x_3)$
$x'_7$	$=$	$\cos(x_6) x'_6$	$= \cos(x_4 + x_2 x_3) x_2$
$x_8$	$=$	$x_1 x_7$	$= x_1 \sin(x_4 + x_2 x_3)$
$x'_8$	$=$	$x_1 x'_7 + x'_1 x_7$	$= x_1 x_2 \sin(x_4 + x_2 x_3) x_2$
$F$	$=$	$x_8$	$= x_1 \sin(x_4 + x_2 x_3)$
$F'$	$=$	$x'_8$	$= x_1 \cos(x_4 + x_2 x_3) x_2$

**Příklad 9** (zpětné derivování). Máme nalézt gradient funkce  $F(x) = x_1 \sin(x_2 x_3 + x_4)$ .

$x_1$	$=$	$x_1$	$\bar{x}_i = \partial F / \partial x_i$
$x_2$	$=$	$x_2$	
$x_3$	$=$	$x_3$	
$x_4$	$=$	$x_4$	
$x_5$	$=$	$x_2 x_3$	$= x_2 x_3$
$x_6$	$=$	$x_4 + x_5$	$= x_4 + x_2 x_3$
$x_7$	$=$	$\sin(x_6)$	$= \sin(x_4 + x_2 x_3)$
$x_8$	$=$	$x_1 x_7$	$= x_1 \sin(x_4 + x_2 x_3)$
$\bar{x}_8$	$=$	$1$	
$\bar{x}_7$	$=$	$\bar{x}_8 x_1$	$= x_1$
$\bar{x}_1$	$=$	$\bar{x}_8 x_7$	$= \sin(x_4 + x_2 x_3)$
$\bar{x}_6$	$=$	$\bar{x}_7 \cos x_6$	$= x_1 \cos(x_4 + x_2 x_3)$
$\bar{x}_5$	$=$	$\bar{x}_6$	$= x_1 \cos(x_4 + x_2 x_3)$
$\bar{x}_4$	$=$	$\bar{x}_6$	$= x_1 \cos(x_4 + x_2 x_3)$
$\bar{x}_3$	$=$	$\bar{x}_5 x_2$	$= x_1 \cos(x_4 + x_2 x_3) x_2$
$\bar{x}_2$	$=$	$\bar{x}_5 x_3$	$= x_1 \cos(x_4 + x_2 x_3) x_3$
$\partial F / \partial x_1$	$=$	$\bar{x}_1$	$= \sin(x_4 + x_2 x_3)$
$\partial F / \partial x_2$	$=$	$\bar{x}_2$	$= x_1 \cos(x_4 + x_2 x_3) x_3$
$\partial F / \partial x_3$	$=$	$\bar{x}_3$	$= x_1 \cos(x_4 + x_2 x_3) x_2$
$\partial F / \partial x_4$	$=$	$\bar{x}_4$	$= x_1 \cos(x_4 + x_2 x_3)$

# Minimalizace s nelineárními omezeními (nelineární programování)

**Definice 54** Úlohou nelineárního programování nazýváme nalezení bodu  $x^* \in R^n$  takového, že

$$x^* = \arg \min_{\substack{c_I(x) \leq 0 \\ c_E(x) = 0}} F(x) \quad (\text{NP})$$

kde  $F : R^n \rightarrow R$ ,  $c_I : R^n \rightarrow R^{m_I}$ ,  $c_E : R^n \rightarrow R^{m_E}$  jsou dvakrát spojitě diferencovatelná zobrazení ( $c_I(x) \leq 0$  je míněno po složkách).

**Poznámka 81** Předpokládáme, že  $I = \{1, \dots, m_I\}$ ,  $E = \{m_I + 1, \dots, m_I + m_E = m\}$  a že minimum v (NP) je lokální.

**Poznámka 82** Množina

$$C = \{x \in R^n : c_I(x) \leq 0, c_E(x) = 0\}$$

se nazývá přípustnou množinou úlohy (NP).

**Definice 55** Necht'  $C \subset R^n$  je přípustná množina úlohy (NP) a  $x \in C$ . Pak množinu

$$\overline{E}(x) = E \cup \{i \in I : c_i(x) = 0\}$$

nazveme množinou indexů omezení aktivních v bodě  $x$ .

**Definice 56** Necht'  $C \subset R^n$  je přípustná množina úlohy (NP) a  $x \in C$ . Jestliže gradienty  $\nabla c_i(x)$ ,  $i \in \overline{E}(x)$ , jsou lineárně nezávislé, řekneme, že jsou splněny podmínky LICQ (linear independence constraint qualification).

**Poznámka 83** Podmínky LICQ v podstatě zajišťují, aby linearizovaná omezení dostatečně dobře vystihovala nelineární omezení. Často se používají slabší podmínky MFCQ (Mangasarian-Fromowitz constraint qualification):

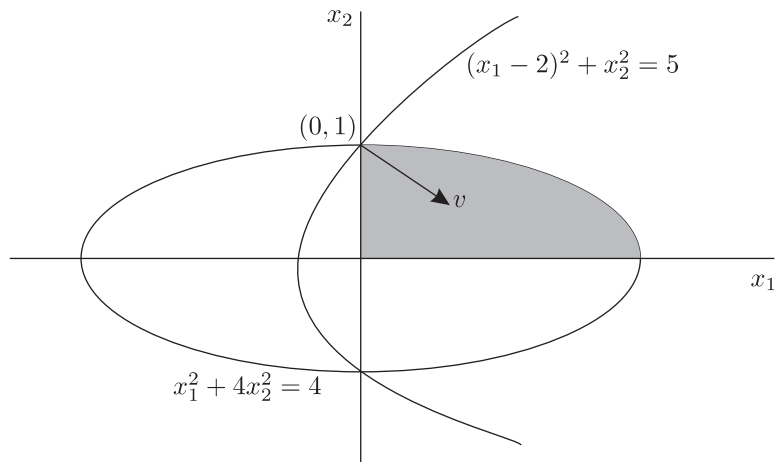
- (a) Gradienty  $\nabla c_i(x)$ ,  $i \in E$ , jsou lineárně nezávislé.
- (b) Existuje vektor  $v \in R^n$  takový, že  $v^T \nabla c_i(x) = 0 \forall i \in E$  a  $v^T \nabla c_i(x) < 0 \forall i \in \overline{E}(x) \setminus E$ .

Z LICQ plyne MFCQ (ale ne naopak). Pro praktické použití jsou výhodnější podmínky LICQ.

**Příklad 10** Platí MFCQ, neplatí LICQ:

$$\begin{aligned} c_1(x) &= -x_1 \leq 0 \\ c_2(x) &= -x_2 \leq 0 \\ c_3(x) &= x_1^2 + 4x_2^2 - 4 \leq 0 \\ c_4(x) &= (x_1 - 2)^2 + x_2^2 - 5 \leq 0 \end{aligned}$$





$$x^* = [0, 1]^T \in R^2$$

$$\bar{E}(x^*) = \{1, 3, 4\}$$

$$\nabla c_1(x^*) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \nabla c_3(x^*) = \begin{bmatrix} 2x_1 \\ 8x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \end{bmatrix}, \quad \nabla c_4(x^*) = \begin{bmatrix} 2(x_1 - 2) \\ 2x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

Tyto vektory nejsou lineárně nezávislé, neboť je jich víc než  $n = 2$ . Pro  $v = [1, -1]^T$  platí

$$v^T \nabla c_1 = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = -1 < 0$$

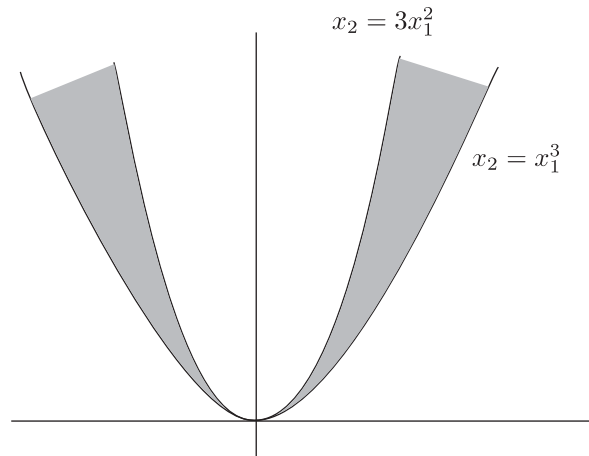
$$v^T \nabla c_3 = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 8 \end{bmatrix} = -8 < 0$$

$$v^T \nabla c_4 = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = -6 < 0$$

**Příklad 11** Neplatí MFCQ

$$c_1(x) = x_1^2 - x_2 \leq 0$$

$$c_2(x) = -3x_1^2 + x_2 \leq 0$$



$$x^* = [0, 0]^T$$

$$\nabla c_1(x^*) = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \nabla c_2(x^*) = \begin{bmatrix} -6x_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Pro libovolný vektor  $v \in R^n$  platí

$$\begin{aligned} v^T \nabla c_1(x^*) &= -v_2 \\ v^T \nabla c_2(x^*) &= +v_2 \end{aligned}$$

Nemůže současně platit  $v_2 > 0$  a  $v_2 < 0$ .

**Definice 57** *Nechť  $C \subset R^n$  je přípustná množina úlohy (NP) a  $x \in C$ . Jestliže existují vektory  $u_I \in R^{m_I}$  a  $u_E \in R^{m_E}$  takové, že*

$$\begin{aligned} \nabla_x L(x, u) &= 0 \\ c_i(x) &\leq 0, \quad u_i \geq 0, \quad u_i c_i(x) = 0, \quad i \in I \\ c_i(x) &= 0, \quad i \in E \end{aligned}$$

kde

$$L(x, u) = f(x) + u_I^T c_I(x) + u_E^T c_E(x)$$

řekneme, že  $x \in C$  je KKT (Karush, Kuhn, Tucker) bodem úlohy (NP).

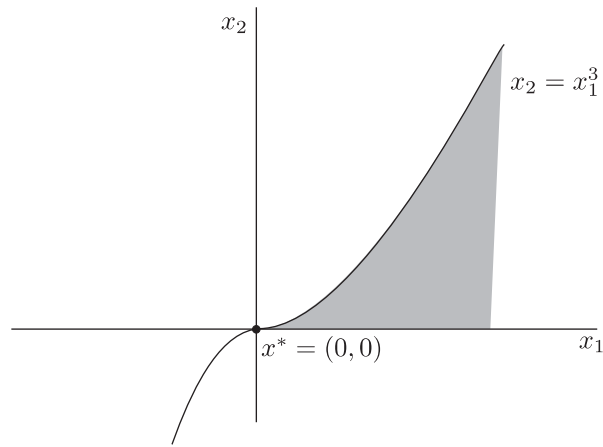
**Poznámka 84** Funkce  $L(x, u)$  se nazývá Lagrangeovou funkcí úlohy (NP). Vektory  $u_I \in R^{m_I}$  a  $u_E \in R^{m_E}$  se nazývají Lagrangeovými multiplikátory. Dvojice  $(x, u) \in R^{n+m}$  se nazývá KKT párem úlohy (NP).

**Věta 97** (Nutné podmínky 1.řádu) *Nechť bod  $x \in C$  je řešením úlohy (NP) a jsou v něm splněny podmínky LICQ (nebo MFCQ). Pak  $x \in C$  je KKT bodem úlohy (NP).*

**Poznámka 85** Jsou-li funkce  $c_i(x)$ ,  $i \in I$ , konvexní a funkce  $c_i(x)$ ,  $i \in E$ , lineární, můžeme podmínky LICQ (nebo MFCQ) vynechat.

**Příklad 12** Uvažujeme úlohu NP, kde

$$\begin{aligned} F(x_1, x_2) &= x_1 \\ c_1(x_1, x_2) &= -x_1^3 + x_2 \leq 0 \\ c_2(x_1, x_2) &= -x_2 \leq 0 \end{aligned}$$



Minimum  $x^* = [0, 0]^T$

$$\nabla c_1(x^*) = \begin{bmatrix} -3x_1^2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla c_2(x^*) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Stejně jako v příkladu 11 neplatí MFCQ.

$$\begin{aligned} L(x, u) &= x_1 + u_1(-x_1^3 + x_2) + u_2(-x_2) \\ \nabla L(x^*, u) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + u_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \neq 0 \end{aligned}$$

(pro libovolný vektor  $u \in R^2$ ).

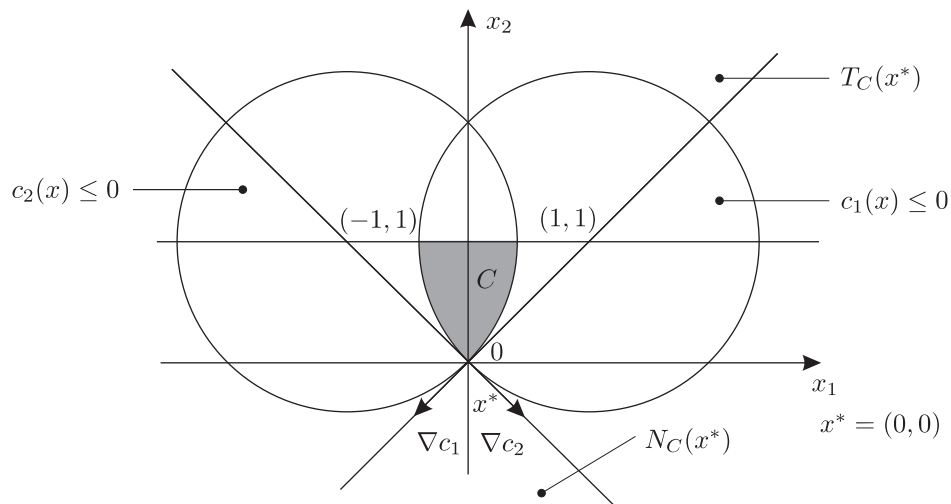
**Příklad 13** Uvažujeme úlohu NP, kde

$$F(x) = x_2$$

$$c_1(x) = \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 1)^2 - 1 \leq 0$$

$$c_2(x) = \frac{1}{2}(x_1 + 1)^2 + \frac{1}{2}(x_2 - 1)^2 - 1 \leq 0$$

$$c_3(x) = x_2 - 1 \leq 0$$



$$\nabla c_1(x^*) = \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla c_2(x^*) = \begin{bmatrix} x_1 + 1 \\ x_2 - 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Linearizovaná omezení

$$\left. \begin{array}{l} -x_1 - x_2 \leq 0 \\ x_1 - x_2 \leq 0 \end{array} \right\} \equiv T_C(x^*)$$

Musí být

$$\begin{array}{ll} g^T v \geq 0 & \forall v \in T_C(x^*) \\ -g^T v \leq 0 & \forall v \in T_C(x^*) \end{array}$$

⇕

$$-g \in N_C(x^*)$$

$$\begin{aligned} N_C(x) &= \text{cone}(\text{conv} \{ \nabla c_1(x^*), \nabla c_2(x^*) \}) \\ &= \{ u_1 \nabla c_1(x^*) + u_2 \nabla c_2(x^*) : u_1 \geq 0, u_2 \geq 0 \} \end{aligned}$$

Tedy

$$-g = u_1^* \nabla c_1(x^*) + u_2^* \nabla c_2(x^*) \quad \Rightarrow \quad g + u_1^* \nabla c_1(x^*) + u_2^* \nabla c_2(x^*) = 0$$

$$\begin{aligned} L(x, u) &= x_2 + u_1 c_1(x) + u_2 c_2(x) + u_3 c_3(x) \\ \nabla L(x, u) &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + u_1 \begin{bmatrix} -1 \\ -1 \end{bmatrix} + u_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \\ u_1 &= \frac{1}{2}, \quad u_2 = \frac{1}{2}, \quad u_3 = 0, \\ c_1 &= 0, \quad c_2 = 0, \quad c_3 = -1 \end{aligned}$$

**Definice 58** *Nechť  $x \in C$  je KKT bodem úlohy (NP). Jestliže platí*

$$c_i(x) < u_i \quad \forall i \in I$$

*řekneme, že v bodě  $x \in C$  jsou splněny podmínky striktní komplementarity (SC).*

**Příklad 14** *Uvažujme úlohu (NP), kde*

$$\begin{aligned} F(x_1, x_2) &= \frac{1}{2}(x_1^2 + x_2^2) \\ c_1(x_1, x_2) &= -x_1 \leq 0 \\ c_2(x_1, x_2) &= -x_2 \leq 0 \end{aligned}$$

Platí

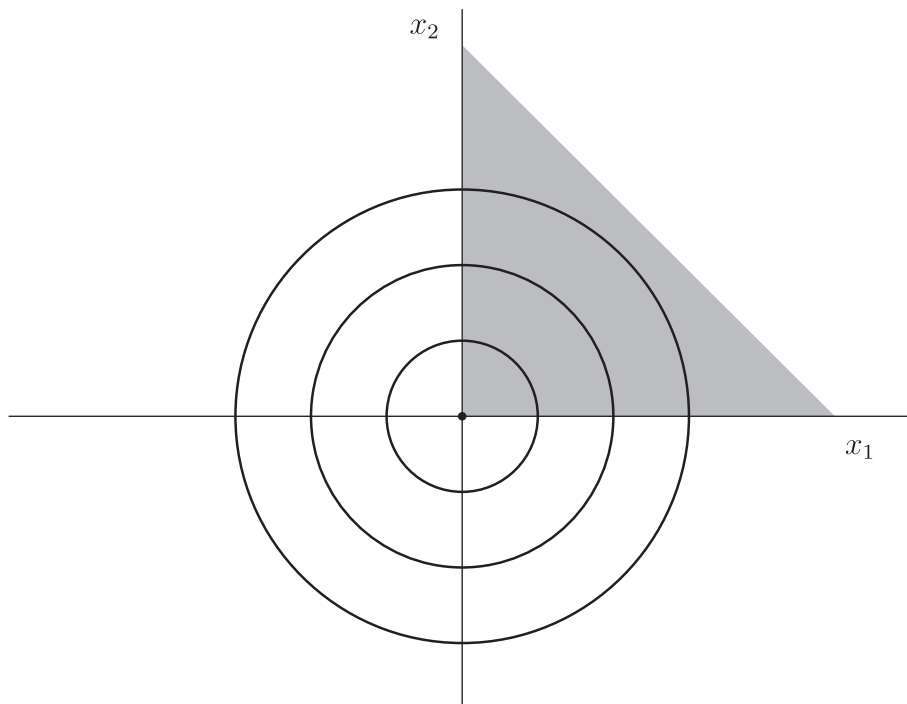
$$\begin{aligned} L(x, u) &= \frac{1}{2}(x_1^2 + x_2^2) - u_1 x_1 - u_2 x_2 \\ \nabla_x L(x, u) &= \begin{bmatrix} x_1 - u_1 \\ x_2 - u_2 \end{bmatrix} \end{aligned}$$

Je-li bod  $x$  KKT bodem úlohy (NP), musí platit  $\nabla_x L(x, u) = 0$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ ,  $u_1 \geq 0$ ,  $u_2 \geq 0$ ,  $u_1 x_1 = 0$ ,  $u_2 x_2 = 0$ , což lze zajistit jedině tehdy, jestliže

$$x_1 = u_1 = 0$$

$$x_2 = u_2 = 0$$

takže nejsou splněny podmínky SC.



Funkce  $F(x_1, x_2)$  má minimum (bez omezení) v bodě  $x = [0, 0]^T$ . Je tedy možné omezení  $x_1 \geq 0$ ,  $x_2 \geq 0$  vypustit.

**Poznámka 86** Nejsou-li splněny podmínky SC, zkomplikují se podmínky 2.řádu pro řešení úlohy (NP). Také nastanou problémy při vyšetřování výpočetních algoritmů. Proto budeme splnění podmínek SC předpokládat.

**Věta 98** (Postačující podmínky 2.řádu) Nechť bod  $x \in C$  je KKT bodem úlohy (NP) a jsou v něm splněny podmínky LICQ (nebo MFCQ) a SC. Jestliže

$$z^T \nabla_{xx}^2 L(x, u) z > 0$$

pro všechny vektory  $z \in R^n$ ,  $z \neq 0$  takové, že

$$z^T \nabla c_i(x) = 0 \quad \forall i \in \overline{E}(x)$$

pak  $x \in C$  je řešením úlohy (NP).

**Definice 59** Budeme používat označení

$$\begin{aligned} a_i(x) &= \nabla c_i(x), \quad i \in I \cup E \\ A_I(x) &= [a_i(x), i \in I] \\ A_E(x) &= [a_i(x), i \in E] \\ \overline{A}(x) &= [a_i(x), i \in \overline{E}(x)] \end{aligned}$$

$$g(x, u) = \nabla_x L(x, u) = \nabla F(x) + \sum_{i \in I \cup E} u_i \nabla c_i(x) = \nabla F(x) + A_I(x)u_I + A_E(x)u_E$$

$$G(x, u) = \nabla_{xx}^2 L(x, u) = \nabla^2 F(x) + \sum_{i \in I \cup E} u_i \nabla^2 c_i(x)$$

$$C_I = \text{diag} \{c_i, i \in I\}$$

$$U_I = \text{diag} \{u_i, i \in I\}$$

a označení  $\bar{Z}(x)$  pro matici, jejíž sloupce tvoří ortonormální bázi v ortogonálním doplňku podprostoru generovaného vektory  $a_i(x)$ ,  $i \in \bar{E}(x)$ , takže (jsou-li splněny podmínky LICQ)  $[\bar{A}(x), \bar{Z}(x)]$  je regulárně čtvercová matice,  $\bar{Z}^T(x)\bar{Z}(x) = I$  a  $\bar{Z}^T(x)\bar{A}(x) = 0$ .

**Poznámka 87** Použijeme-li označení z Definice 59, můžeme postačující podmínky 2.řádu (Věta 98) zapsat ve tvaru

$$g(x, u) = 0$$

$$c_I(x) \leq 0, \quad u_I \geq 0, \quad U_I C_I(x) = 0$$

$$c_E(x) = 0$$

$$\bar{Z}^T(x)G(x, u)\bar{Z}(x) \succ 0 \quad (\text{positivně definitní})$$

## Minimalizace na lineární varietě

Minimalizace s lineárními omezeními ve tvaru rovností. Máme řešit úlohu

$$x^* = \arg \min_{x \in C} F(x)$$

$$C = \{x \in R^n : a_i^T x = \alpha_i, i \in E\}$$

**Poznámka 88** KKT podmínky mají v tomto případě tvar

$$g(x) + Au = 0$$

$$A^T x = b$$

kde  $g(x)$  je gradient funkce  $F(x)$ ,  $A = [a_1, \dots, a_m]$ ,  $b = [\alpha_1, \dots, \alpha_m]^T$ .

**Poznámka 89** Krok modifikované Newtonovy metody pro řešení KKT systému má tvar

$$\begin{bmatrix} B & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} s \\ u \end{bmatrix} = - \begin{bmatrix} g \\ 0 \end{bmatrix}$$

kde  $B$  je (obvykle s.p.d.) aproximace Hessovy matice. Pak  $x_+ = x + \alpha s$ , kde  $\alpha > 0$  je délka kroku.

## Metody promítaných gradientů

Řešíme-li soustavu z poznámky 89, dostaneme

$$\begin{aligned} u &= -(A^T B^{-1} A)^{-1} A^T B^{-1} g \\ s &= -(B^{-1} - B^{-1} A (A^T B^{-1} A)^{-1} A^T B^{-1}) g = -Pg \end{aligned}$$

**Poznámka 90** Metody promítaných gradientů reprezentují lineární varietu maticí  $C = (A^T B^{-1} A)^{-1}$  (nebo trojúhelníkovým rozkladem  $R^T R = A^T B^{-1} A$ ) a maticí  $P = B^{-1} - B^{-1} A (A^T B^{-1} A)^{-1} A^T B^{-1}$

### Přidání omezení

**Věta 99** Nechť  $A^+ = [A, a]$ , kde  $a \notin \mathcal{L}(A)$  a nechť  $C^+ = ((A^+)^T H A^+)^{-1}$ ,  $(R^+)^T R^+ = (A^+)^T H A^+$ ,  $P^+ = H - H A^+ C^+ (A^+)^T H$ . Pak platí

$$\begin{aligned} C^+ &= \begin{bmatrix} C + \frac{C A^T H a a^T H A C}{a^T P a}, & -\frac{C A^T H a}{a^T P a} \\ -\frac{a^T H A C}{a^T P a}, & \frac{1}{a^T P a} \end{bmatrix}, \\ R^+ &= \begin{bmatrix} R & r \\ 0 & \rho \end{bmatrix}, \end{aligned}$$

kde  $R^T r = A^T H a$ ,  $\rho^2 = a^T H a - r^T r$ , a

$$P^+ = P - \frac{P a a^T P}{a^T P a}$$

### Ubrání omezení

**Věta 100** Nechť  $A M = [A^-, a]$ , kde  $M$  je nějaká permutační matice, a nechť  $C^- = ((A^-)^T H A^-)^{-1}$ ,  $(R^-)^T R^- = (A^-)^T H A^-$ . Nechť

$$M^T C M = \begin{bmatrix} \tilde{C} & \tilde{c} \\ \tilde{c}^T & \tilde{\gamma} \end{bmatrix}$$

a nechť  $Q$  je ortogonální matice taková, že

$$Q R M = \begin{bmatrix} \tilde{R} & \tilde{r} \\ 0 & \tilde{\rho} \end{bmatrix}$$

je horní trojúhelníková matice. Pak platí

$$C^- = \tilde{C} - \frac{\tilde{c} \tilde{c}^T}{\tilde{\gamma}}$$

a  $R^- = \tilde{R}$ .

## Metody redukováných gradientů

Předpokládáme, že  $Z$  je matice jejíž sloupce tvoří ortonormální bázi v ortogonálním doplňku podprostoru generovaném sloupci matice  $A$ . Má-li  $A$  plnou hodnost, je  $[A, Z]$  čtvercová regulární a platí  $A^T Z = 0$ ,  $Z^T Z = I$ .

**Poznámka 91** Označme  $\tilde{g} = Z^T g$  redukováný gradient,  $\tilde{G} = Z^T G Z$  a  $\tilde{B} = Z^T B Z$  redukovanou Hessovu matici a její aproximaci. Pak

$$\begin{aligned}\tilde{g} &= Z^T g \\ \tilde{s} &= -\tilde{B}^{-1} \tilde{g} = -(Z^T B Z)^{-1} Z^T g \\ s &= Z \tilde{s} = -Z (Z^T B Z)^{-1} Z^T g\end{aligned}$$

**Poznámka 92** V metodách s proměnnou metrikou se redukované matice  $\tilde{B} = Z^T B Z$  a  $\tilde{H} = \tilde{B}^{-1} = Z^T H Z$  aktualizují pomocí redukováných veličin  $\tilde{y} = \tilde{g}_+ - \tilde{g} = Z^T (g_+ - g) = Z^T y$  a  $\tilde{d} = \alpha \tilde{s}$ .

**Poznámka 93** Metody redukováných gradientů reprezentují lineární variantu trojúhelníkovým rozkladem  $R^T R = A^T A$  a maticí  $Z$  takovou, že  $A^T Z = 0$  a  $Z^T Z = I$ . Lagrangeovy multiplikátory se počítají podle vzorce  $u = -(R^T R)^{-1} A^T g$ .

### Přidání omezení

**Věta 101** Nechť  $A^+ = [A, a]$ , kde  $a \notin \mathcal{L}(A)$ . Nechť  $\tilde{a} = Z^T a$  a  $Q$  je ortogonální matice taková, že

$$\tilde{a}^T Q = [0, \dots, 0, \|\tilde{a}\|]$$

Položme  $ZQ = [Z^+, z]$ . Pak  $(A^+)^T Z^+ = 0$  a  $(Z^+)^T Z^+ = I$ .

### Ubrání omezení

**Věta 102** Nechť  $AM = [A^-, a]$ , kde  $M$  je nějaká permutační matice. Nechť  $Q$  je ortogonální matice taková, že matice  $QRM$  je horní trojúhelníková. Nechť  $Z^- = [Z, z]$ , kde

$$z = AM(QRM)^{-1} [0, \dots, 0, 1]^T$$

Pak platí  $(A^-)^T Z^- = 0$  a  $(Z^-)^T Z^- = I$ .

## Metody aktualizovaných gradientů

Předpokládejme, že  $S$  je matice s lineárně nezávislými sloupci taková, že

$$SS^T = P = H - HA(A^T HA)^{-1} A^T H$$

**Poznámka 94** Metody aktualizovaných gradientů určují směrový vektor podle vzorců

$$\begin{aligned}\tilde{s} &= -\tilde{g} = -S^T g \\ s &= S \tilde{s} = -S \tilde{g} = -SS^T g\end{aligned}$$



Matice  $S$  se aktualizují pomocí metod s proměnnou metrikou v součinném tvaru. Aktualizace metody BFGS má tvar

$$S_+ = \sqrt{\gamma} \left( S_+ + \frac{1}{b} S \tilde{d} \left( \sqrt{\frac{\rho b}{\gamma c}} \tilde{d} - \tilde{y} \right)^T \right)$$

kde  $\tilde{y} = \tilde{g}_+ - \tilde{g} = S^T(g_+ - g) = S^T y$  a  $\tilde{d} = \alpha \tilde{s}$ .

**Poznámka 95** Metody aktualizovaných gradientů reprezentují lineární varietu trojúhelníkovým rozkladem  $R^T R = A^T A$  a maticí  $S$  takovou, že  $SS^T = H - HA(A^T HA)^{-1}A^T H$ . Lagrangeovy multiplikátory se počítají podle vzorce  $u = -(R^T R)^{-1}A^T g$

### Přidání omezení

**Věta 103** Nechť  $A^+ = [A, a]$ , kde  $a \notin \mathcal{L}(A)$ . Nechť  $S = [\tilde{S}, s]$  a

$$S^+ = \tilde{S} - \left( \frac{1 - \lambda a^T s}{a^T S S^T a} S S^T a + \lambda s \right) a^T \tilde{S}$$

kde  $\lambda$  je kořenem kvadratické rovnice

$$\lambda^2 a^T \tilde{S} \tilde{S}^T a + 2\lambda a^T s = 1$$

Pak  $(A^+)^T S^+ = 0$  a platí

$$S^+(S^+)^T = H - HA^+((A^+)^T HA^+)^{-1}(A^+)^T H$$

### Ubrání omezení

**Věta 104** Nechť  $AM = [A^-, a]$ , tak  $M$  je nějaká permutační matice. Nechť  $Q$  je ortogonální matice taková, že  $QRM$  je horní trojúhelníková. Nechť  $S^- = [S, s]$ , kde

$$s = AM(QRM)^{-1}[0, \dots, 0, 1]^T$$

Pak platí  $(A^-)^T S^- = 0$  a

$$S^-(S^-)^T = H - HA^-((A^-)^T HA^-)^{-1}(A^-)^T H$$

## Minimalizace s lineárními omezeními

Máme řešit úlohu

$$\begin{aligned} x^* &= \arg \min_{x \in C} F(x) \\ C &= \{x \in \mathbb{R}^n : a_i^T x \leq \alpha_i, i \in I, a_i^T x = \alpha_i, i \in E\} \end{aligned}$$

Tedy

$$c_i(x) = a_i^T x - \alpha_i, \quad i \in I \cup E$$

## Metody aktivních omezení

### Algoritmus 8 (metoda aktivních omezení)

K1: Najdeme přípustný bod  $x \in C$  (úloha lineárního programování), určíme množinu  $\bar{E}(x)$  indexů omezení aktivních v bodě  $x \in C$  určíme reprezentaci variety

$$\bar{L}(x) = \bigcap_{i \in \bar{E}(x)} L(a_i, \alpha_i)$$

vypočteme hodnotu  $F(x)$  a gradient  $g(x)$ .

K2: Určíme směrový vektor  $s \in \bar{L}(x)$  (minimalizace na lineární varietě) a vektor Lagrangeových multiplikátorů  $u$ . Jestliže  $\|s\| = 0$  a  $u \geq 0$  ukončíme výpočet (máme KKT bod  $\Rightarrow$  řešení)

K3: Jestliže  $\|s\| = 0$  a neplatí  $u \geq 0$  určíme index  $\ell \in \bar{E}(x)$  takový, že

$$u_\ell = \min_{i \in \bar{E}(x)} (u_i),$$

ubereme omezení s indexem  $\ell$ , čímž změníme množinu  $\bar{E}(x)$ , varietu  $\bar{L}(x)$  a její reprezentaci a přejdeme na krok K2.

K4: Jestliže  $\|s\| \neq 0$  určíme délku kroku  $0 < \alpha < \bar{\alpha}$  (například Armijovým výběrem), kde

$$\bar{\alpha} = \min_{i \notin \bar{E}(x), \alpha_i^T s > 0} \frac{\alpha_i - a_i^T x}{a_i^T s}$$

vypočteme hodnotu  $F(x)$  a gradient  $g(x)$ .

K5: Pokud  $\alpha = \bar{\alpha}$  přidáme nová aktivní omezení čímž změníme množinu  $\bar{E}(x)$ , varietu  $\bar{L}(x)$  a její reprezentaci.

K6: Přejdeme na krok K2.

**Poznámka 96** Minimum na lineární varietě se obvykle určuje přibližně. Omezení se ubírá pokud  $\|Pg\| \leq \varepsilon \|g\|$  (metody promítaných gradientů), nebo  $\|\tilde{g}\| \leq \varepsilon \|g\|$  (metody redukováných gradientů), kde  $\varepsilon > 0$  je požadovaná přesnost.

## Lineární programování

Máme řešit úlohu

$$\begin{aligned} x^* &= \arg \min_{x \in C} g^T x \\ C &= \{x \in \mathbb{R}^n : a_i^T x \leq \alpha_i, i \in I, a_i^T x = \alpha_i, i \in E\} \end{aligned}$$

## Simpexová metoda

Je to metoda největšího spádu s lineárními omezeními (takže  $B = I$ ). Délka kroku se vždy vybírá tak, aby platilo  $\alpha = \bar{\alpha}$ .

- Začneme-li v libovolném přípustném bodě, dostaneme se vždy do vrcholu.
- Je-li tento vrchol KKT bodem, ukončíme výpočet.
- V opačném případě ubereme omezení a po hraně se dostaneme do sousedního vrcholu.
- Procházíme vrcholy  $\Rightarrow$  konečný počet kroků, ale exponenciální složitost.

## Nalezení přípustného bodu

Nechť  $x \in R^n$ . Pro zjednodušení budeme předpokládat, že  $a_i^T x \geq \alpha_i$  pro  $i \in E$ . Nechť

$$\begin{aligned} I_1(x) &= \{i \in I \cup E : a_i^T x \leq \alpha_i\} \\ I_2(x) &= \{i \in I \cup E : a_i^T x > \alpha_i\} \end{aligned}$$

Řešíme posloupnost úloh lineárního programování

$$\begin{aligned} x^* &= \arg \min_{x \in C} \sum_{i \in I_2(x)} (a_i^T x - \alpha_i) \\ C &= \{x \in R^n : a_i^T x \leq \alpha_i, i \in I_1(x), a_i^T x \geq \alpha_i, i \in I_2(x)\}. \end{aligned}$$

Řešením každé této úlohy je bod, ve kterém se alespoň jedno porušené omezení stane aktivním  $\Rightarrow$  po konečném počtu kroků dostaneme přípustný bod.

## Kvadratické programování

Máme řešit úlohu

$$\begin{aligned} x^* &= \arg \min_{x \in C} \left( \frac{1}{2} x^T G x + g^T x \right) \\ C &= \{x \in R^n : a_i^T x \leq \alpha_i, i \in I, a_i^T x = \alpha_i, i \in E\} \end{aligned}$$

Použijeme Newtonovu metodu s lineárními omezeními (takže  $B = G$ ).

- Optimální délka kroku  $\alpha = 1$  vede k nalezení minima na lineární varietě.
- Pokud  $\bar{\alpha} \leq 1$  přidáváme nová aktivní omezení

## Metody rekursivního kvadratického programování

**Definice 60** Úlohou kvadratického programování (QP) přiřazenou úloze (NP) v bodě  $x \in R^n$  nazýváme nalezení vektoru  $\Delta x$ , který minimalizuje kvadratickou funkci

$$\frac{1}{2}(\Delta x)^T B \Delta x + (\nabla f(x))^T \Delta x$$

(kde  $B$  je aproximací  $G(x, u)$ ) za podmínek

$$\begin{aligned} c_I(x) + A_I^T(x) \Delta x &\leq 0 \\ c_E(x) + A_E^T(x) \Delta x &= 0 \end{aligned}$$

**Poznámka 97** Nechť  $B = G(x, u)$ . Označíme-li  $u + \Delta u$  Lagrangeovy multiplikátory úlohy (QP), můžeme KKT podmínky úlohy (QP) zapsat ve tvaru

$$\begin{aligned} g(x, u + \Delta u) + G(x, u) \Delta x &= 0 \\ c_i(x) + a_i^T(x) \Delta x &\leq 0, \quad i \in I \\ u_i + \Delta u_i &\geq 0, \quad i \in I \\ (u_i + \Delta u_i)(c_i(x) + a_i^T(x) \Delta x) &= 0, \quad i \in I \\ c_i(x) + a_i^T(x) \Delta x &= 0, \quad i \in E \end{aligned}$$

které vzniknou linearizací KKT podmínek úlohy (NP).

**Poznámka 98** Úloha (QP) nemusí mít řešení i když úloha (NP) má řešení. Tento nedostatek lze odstranit tak, že se omezení v úloze (QP) nahradí oslabenými omezeními

$$\begin{aligned} \beta^l c_I(x) + A_I^T(x) \Delta x &\leq 0 \\ \beta^l c_E(x) + A_E^T(x) \Delta x &= 0 \end{aligned}$$

kde  $0 < \beta < 1$  a  $l$  je nejmenší přirozené číslo takové, že modifikovaná úloha (QP) má řešení.

**Definice 61** Metody rekursivního kvadratického programování (RQP) jsou iterační a jejich iterační krok má tvar

$$\begin{aligned} x^+ &= x + \alpha \Delta x \\ u_I^+ &= u_I + \alpha \Delta u_I \\ u_E^+ &= u_E + \alpha \Delta u_E \end{aligned}$$

kde  $\Delta x$ ,  $\Delta u_I$ ,  $\Delta u_E$  jsou směrové vektory určené řešením úlohy (QP) a  $\alpha > 0$  je délka kroku.

**Poznámka 99** Nechť  $x^* \in C$  je bod, ve kterém jsou splněny postačující podmínky druhého řádu (Věta 98). Pak existuje okolí  $B(x^*, \varepsilon) \subset R^n$  takové, že  $\overline{E}(x) = \overline{E}(x^*) \forall x \in B(x^*, \varepsilon)$ . Proto se v případě lokální konvergence můžeme omezit na omezení ve tvaru rovnosti

$$c_i(x) = 0, \quad i \in \overline{E} = \overline{E}(x^*)$$

Použijeme označení  $\overline{P} = I - \overline{A}(\overline{A}^T \overline{A})^{-1} \overline{A}^T = \overline{Z} \overline{Z}^T$  (kde  $\overline{A} = \overline{A}(x)$  a  $\overline{Z} = \overline{Z}(x)$ ).

**Věta 105** Necht  $\{x_i\} \subset R^n$ ,  $\{u_i\} \subset R^m$  jsou posloupnosti získané metodou RQP (Definice 61) s  $\alpha_i = 1$ ,  $i \in N$ , takové, že  $x_i \rightarrow x^*$ ,  $u_i \rightarrow u^*$ , kde  $(x^*, u^*)$  je KKT pár úlohy (NP) vyhovující postačujícím podmínkám 2.řádu s LICQ a SC. Pak  $x_i \rightarrow x^*$  Q-superlineárně právě tehdy, jestliže

$$\frac{\|\overline{P}_i(B_i - G(x^*, u^*))(x_{i+1} - x^*)\|}{\|x_{i+1} - x^*\|} \rightarrow 0$$

**Důsledek 8** Newtonova metoda s  $B_i = G(x_i, u_i)$  a  $\alpha_i = 1$  je Q-superlineárně konvergentní (za předpokladu, že  $x_i \rightarrow x^*$ ,  $u_i \rightarrow u^*$ ).

**Poznámka 100** Zatímco podmínky pro superlineární konvergenci lze v případě Newtonovy metody snadno splnit, je obtížnější zajistit globální konvergenci této metody. K zajištění globální konvergence se používají exaktní pokutové funkce.

**Definice 62** Funkci

$$P_\sigma(x) = F(x) + \sigma \|c^0(x)\|_P$$

kde  $\sigma > 0$ ,

$$\begin{aligned} c_i^0(x) &= \max(c_i(x), 0), & i \in I \\ c_i^0(x) &= c_i(x), & i \in E \end{aligned}$$

a  $\|\cdot\|_P$  je nějaká (primární) norma, nazveme exaktní pokutovou funkcí úlohy (NP).

**Věta 106** Necht  $(x^*, u^*)$  je KKT pár úlohy (NP) vyhovující postačujícím podmínkám 2.řádu s LICQ a SC a necht  $\sigma > \|u^*\|_D$ , kde  $\|\cdot\|_D$  je norma duální k normě  $\|\cdot\|_P$ . Pak  $P_\sigma(x)$  má lokální minimum v  $x^*$ .

**Poznámka 101** Exaktní pokutovou funkcí úlohy (NP) můžeme použít k výběru délky kroku. Necht

$$P(\alpha) = P_\sigma(x + \alpha\Delta x)$$

Délku kroku volíme tak, aby byla splněna podmínka

$$P(\alpha) - P(0) \leq \varepsilon_1 \alpha P'(0)$$

(Armijo). Musí však platit  $P'(0) < 0$ .

**Věta 107** Necht  $P(\alpha) = P_\sigma(x + \alpha\Delta x)$  a

$$\sigma > \|u^+\|_D - \frac{(\Delta x)^T G(x, u) \Delta x}{\|c^0(x)\|_P}$$

Pak platí  $P'(0) < 0$ .

**Poznámka 102** Metody RQP s exaktní pokutovou funkcí lze realizovat tak, že jsou globálně konvergentní. Exaktní pokutová funkce však kazí superlineární konvergenci, neboť ani v blízkém okolí řešení neplatí  $P(1) < P(0)$  (Maratosův jev), takže nelze použít jednotkovou délku kroku.

**Poznámka 103** K výběru délky kroku je výhodnější použít rozšířenou Lagrangeovu funkci

$$P(\alpha) = F(x + \alpha\Delta x) + (u + \Delta u)^T c^0(x + \alpha\Delta x) + \frac{\sigma}{2} \|c^0(x + \alpha\Delta x)\|^2$$

V tomto případě nelze obecně dokázat globální konvergenci metody.

**Poznámka 104** Metody RQP jsou vhodné pro menší úlohy. Velké řídké úlohy vyžadují kvalitní řešiče velkých řídkých úloh (QP), které nejsou vždy k dispozici.

## Metody vnitřních bodů

**Myšlenka:** Úloha (NP) se převede na posloupnost úloh (IP) závislých na parametru  $\mu$ , přičemž  $\mu \rightarrow 0$ .

**Definice 63** Necht'  $s_I \in R^{m_I}$ ,  $s_I > 0$  a  $\mu > 0$ . V úloze (IP) (závislé na parametru  $\mu$ ) hledáme minimum barierové funkce

$$B(x, s) = F(x) - \mu e^T \ln(S_I)e$$

s omezeními

$$\begin{aligned} c_I(x) + s_I &= 0, & i \in I \\ c_E(x) &= 0, & i \in E \end{aligned}$$

Přitom  $S_I = \text{diag}(s_i, i \in I)$  a  $e$  je vektor, jehož prvky jsou jednotky.

**Věta 108** Vektor  $(x, s_I) \in R^{n+m_I}$  je KKT bodem úlohy (IP), existují-li vektory  $u_I \in R^{m_I}$ ,  $u_E \in R^{m_E}$  takové, že

$$\begin{aligned} \nabla_x L(x, u) &= 0 \\ S_I U_I e &= \mu e \\ c_I(x) + s_I &= 0 \\ c_E(x) &= 0 \end{aligned}$$

kde  $L(x, u)$  je Lagrangeova funkce (stejná jako v Definicí 57).

**Poznámka 105** Derivováním barierové funkce bychom dostali rovnici  $U_I e = \mu S_I^{-1} e$ . Rovnice  $S_I U_I e = \mu e$  je výhodnější (vede na efektivnější algoritmy).

**Poznámka 106** Linearizací KKT podmínek dostaneme krok Newtonovy metody.

$$\begin{bmatrix} G & 0 & A_I & A_E \\ 0 & U_I & S_I & 0 \\ A_I^T & I & 0 & 0 \\ A_E^T & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta s_I \\ \Delta u_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} g \\ S_I U_I e - \mu e \\ c_I + s_I \\ c_E \end{bmatrix}$$

kde  $g = g(x, u)$  a  $G = G(x, u)$ . Tuto rovnici lze redukovat eliminací vektoru  $\Delta s_I$ . Platí

$$\Delta s_I = -M_I(u_I + \Delta u_I) + \mu U_I^{-1} e$$

kde  $M_I = U_I^{-1} S_I$ , takže

$$\begin{bmatrix} G & A_I & A_E \\ A_I^T & -M_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} g \\ c_I + \mu U_I^{-1} e \\ c_E \end{bmatrix}$$

**Definice 64** Necht'  $x \in R^n$ ,  $s_I \in R^{m_I}$  a  $u_I \in R^{m_I}$ . Pak množiny

$$\begin{aligned} \hat{I}(x, s_I, u_I) &= \{i \in I, s_i \leq \varepsilon_I u_i\} \\ \check{I}(x, s_I, u_I) &= \{i \in I, s_i > \varepsilon_I u_i\} \end{aligned}$$

kde  $\varepsilon_I > 0$ , nazveme množinami indexů aktivních a neaktivních omezení.

**Poznámka 107** Eliminací neaktivních omezení dostaneme

$$\Delta\check{u}_I = \check{M}_I^{-1}(\check{c}_I + \check{A}_I^T \Delta x) + \mu \check{S}_I^{-1} e$$

takže

$$\begin{bmatrix} \hat{G} & \hat{A}_I & A_E \\ \hat{A}_I^T & -\hat{M}_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\hat{u}_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} \hat{g} \\ \hat{c}_I + \mu \hat{U}_I^{-1} e \\ c_E \end{bmatrix} \quad (\text{IPS})$$

kde

$$\begin{aligned} \hat{G} &= G + \check{A}_I \check{M}_I^{-1} \check{A}_I^T \\ \hat{g} &= g + \check{A}_I \check{M}_I^{-1} \check{c}_I + \mu \check{A}_I \check{S}_I^{-1} e \end{aligned}$$

Dále

$$\begin{aligned} \Delta\hat{s}_I &= -\hat{M}_I(\hat{u}_I + \Delta\hat{u}_I) + \mu \hat{U}_I^{-1} e \\ \Delta\check{s}_I &= -(\check{c}_I + \check{A}_I^T \Delta x + \check{s}_I) \end{aligned}$$

Nejprve určíme  $\Delta x$ ,  $\Delta\hat{u}_I$ ,  $\Delta u_E$  (nepřesným) řešením soustavy (IPS) a vypočteme vektor  $\Delta\hat{s}_i$ . Pak vypočteme  $\Delta\check{u}_I$  a  $\Delta\check{s}_I$ .

**Poznámka 108** Matice  $\hat{G}$  a  $\hat{M}$  jsou omezené (předpokládáme, že  $G$  a  $\check{A}_I$  jsou omezené) a jsou-li splněny podmínky SC, platí

$$\hat{M}_I \rightarrow 0$$

**Poznámka 109** Matice v (IPS) je symetrická a indefinitní. Soustavu (IPS) lze efektivně řešit metodou sdružených gradientů s předpodmiňovačem

$$C = \begin{bmatrix} \hat{D} & \hat{A}_I & A_E \\ \hat{A}_I^T & -\hat{M}_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix}$$

kde  $\hat{D}$  je diagonální pozitivně definitní matice (aproximace matice  $\hat{G}$ ).

**Poznámka 110** K výběru délky kroku se používá rozšířená barierová Lagrangeova funkce

$$\begin{aligned} P(\alpha) &= F(x + \alpha\Delta x) - \mu e^T \ln(S_I + \alpha\Delta S_I) e \\ &\quad + (u_I + \Delta u_I)^T (c_I(x + \alpha\Delta x) + s_I + \alpha\Delta s_I) + (u_E + \Delta u_E)^T c_E(x + \alpha\Delta x) \\ &\quad + \frac{\sigma}{2} \|c_I(x + \alpha\Delta x) + s_I + \alpha\Delta s_I\|^2 + \frac{\sigma}{2} \|c_E(x + \alpha\Delta x)\|^2 \end{aligned}$$

kde  $\sigma > 0$ .

**Věta 109** Nechť  $\Delta x$  a  $\Delta s_I$  jsou vektory určené podle Poznámky 107 a nechť

$$\sigma > - \frac{(\Delta x)^T G \Delta x + (\Delta s_I)^T S_I^{-1} U_I \Delta s_I}{\|c_E\|^2 + \|c_I + s_I\|^2}$$

Pak řešíme-li soustavu (IPS) dostatečně přesně, platí  $P'(0) < 0$ .

**Poznámka 111** Požadujeme, aby platilo  $s_I + \alpha \Delta s_I > 0$ ,  $u_I + \alpha \Delta u_I > 0$ . Z tohoto důvodu se určí horní meze

$$\begin{aligned}\bar{\alpha}_s &= \gamma \min_{i \in I, \Delta s_i < 0} \left( -\frac{s_i}{\Delta s_i} \right) \\ \bar{\alpha}_u &= \gamma \min_{i \in I, \Delta u_i < 0} \left( -\frac{u_i}{\Delta u_i} \right)\end{aligned}$$

kde  $0 < \gamma < 1$  (např.  $\gamma = 0.99$ ). Pak

$$\begin{aligned}s_I^+ &= s_I + \min(\alpha, \bar{\alpha}_s) \Delta s_I \\ u_I^+ &= u_I + \min(\alpha, \bar{\alpha}_u) \Delta u_I\end{aligned}$$

**Poznámka 112** Parametr  $\mu$  se přepočítává v každém iteračním kroku. Podle Věty 108 by mělo platit  $S_I U_I e = \mu e$ . Proto volíme

$$\mu = \lambda \frac{s_I^T u_I}{m_I}$$

kde  $0 < \lambda < 1$ . Používají se různé heuristické vzorce, například

$$\lambda = 0.1 \left[ \min \left( 0.05 \frac{1 - \rho}{\rho}, 2 \right) \right]^3$$

kde

$$\rho = \frac{\min_{i \in I} s_i u_i}{\frac{s_I^T u_I}{m_i}}$$

(míra centrality).

## Metody nehladkých rovnic

**Myšlenka:** Nelineární KKT systém se (pomocí Fischerovy-Burmeisterovy funkce) převede na systém polohladkých rovnic a ten se řeší pomocí Newtonovy metody.

**Poznámka 113** Fischerova-Burmeisterova funkce má tvar

$$\psi(x_1, x_2) = \sqrt{x_1^2 + x_2^2} - (x_1 + x_2)$$

(a)  $\psi(x_1, x_2) = 0$  platí právě tehdy, jestliže  $x_1 \geq 0$ ,  $x_2 \geq 0$ ,  $x_1 x_2 = 0$

(b) V bodě, kde  $x_1 x_2 \neq 0$  platí

$$\nabla \psi(x_1, x_2) = \begin{bmatrix} \frac{x_1}{\sqrt{x_1^2 + x_2^2}} - 1 \\ \frac{x_2}{\sqrt{x_1^2 + x_2^2}} - 1 \end{bmatrix}$$

V bodě, kde  $x_1 x_2 = 0$ , je  $\psi(x_1, x_2)$  polohladká a  $[-1, -1]^T \in \partial \psi(0, 0)$ , kde  $\partial \psi(0, 0)$  značíme subdiferenciál funkce  $\psi$  v bodě  $(0, 0)$ . Označíme-li  $r(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$  pro  $x_1 x_2 \neq 0$  a  $r(x_1, x_2) = 1$  pro  $x_1 x_2 = 0$ , je vždy

$$\begin{bmatrix} \frac{x_1}{r(x_1, x_2)} - 1 \\ \frac{x_2}{r(x_1, x_2)} - 1 \end{bmatrix} \in \partial \psi(x_1, x_2)$$



**Poznámka 114** Použijeme-li Fischerovu-Burmeisterovu funkci, můžeme nelineární KKT systém

$$\begin{aligned} g(x, u) &= 0 \\ c_i(x) &\leq 0, \quad u_i \geq 0, \quad u_i c_i(x) = 0, \quad i \in I \\ c_i(x) &= 0, \quad i \in E \end{aligned}$$

převést na systém polohladkých rovnic

$$\begin{aligned} g(x, u) &= 0 \\ \psi(u_i, -c_i(x)) &= 0, \quad i \in I \\ c_i(x) &= 0, \quad i \in E \end{aligned}$$

**Poznámka 115** Budeme používat označení z Definice 59 a položíme

$$R_I = \text{diag} \{r_i, i \in I\} = \text{diag} \left\{ \sqrt{u_i^2 + c_i^2}, i \in I \right\}$$

**Poznámka 116** Linearizací polohladkých rovnic dostaneme krok Newtonovy metody

$$\begin{bmatrix} G & A_I & A_E \\ (R_I + C_I)R_I^{-1}A_I^T & -(R_I - U_I)R_I^{-1} & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} g \\ \psi_I \\ c_E \end{bmatrix}$$

kde  $g = g(x, u)$  a  $G = G(x, u)$ . Tuto rovnici lze symetrizovat vynásobením diagonální maticí  $\text{diag} (I, R_I(R_I + C_I)^{-1}, I)$ . Platí

$$\begin{bmatrix} G & A_I & A_E \\ A_I^T & -M_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} g \\ R_I(R_I + C_I)^{-1}\psi_I \\ c_E \end{bmatrix}$$

kde  $M_I = (R_I + C_I)^{-1}(R_I - U_I)$ .

**Definice 65** Necht  $x \in R^n$ ,  $u_I \in R^{m_I}$ . Pak množiny

$$\begin{aligned} \hat{I}(x, u_I) &= \{i \in I : r_i - u_i \leq \varepsilon_I(r_i + c_i)\} \\ \check{I}(x, u_I) &= \{i \in I : r_i - u_i > \varepsilon_I(r_i + c_i)\} \end{aligned}$$

kde  $\varepsilon_I > 0$ , nazveme množinami indexů aktivních a neaktivních omezení.

**Poznámka 117** Podmínka  $r_i - u_i \leq \varepsilon_I(r_i + c_i)$  je ekvivalentní podmínce

$$-\frac{\partial \psi_i}{\partial u_i} \leq \varepsilon_I \frac{\partial \psi_i}{\partial c_i}$$

**Věta 110** Necht  $x^* \in C$  je KKT bodem úlohy (NP), ve kterém jsou splněny podmínky LICQ a SC. Pak existují okolí  $\mathcal{N} \subset R^n$  a  $\mathcal{M}_I \subset R^{m_I}$  vektorů  $x^*$  a  $u_I^*$  taková, že

$$\begin{aligned} c_i(x^*) = 0 &\Rightarrow i \in \hat{I}(x, u_I) \\ u_i^* = 0 &\Rightarrow i \in \check{I}(x, u_I) \end{aligned}$$

pokud  $x \in \mathcal{N}$  a  $u_I \in \mathcal{M}_I$ . Pokud  $x \rightarrow x^*$  a  $u_I \rightarrow u_I^*$ , pak

$$\begin{aligned} c_i(x^*) = 0 &\Rightarrow \frac{r_i - u_i}{r_i + c_i} \rightarrow 0 \\ u_i^* = 0 &\Rightarrow \frac{r_i - u_i}{r_i + c_i} \rightarrow \infty \end{aligned}$$

**Poznámka 118** Eliminací neaktivních omezení dostaneme

$$\Delta\check{u}_I = \check{M}_I^{-1}[(\check{R}_I + \check{C}_I)^{-1}\check{R}_I\check{\psi}_I + \check{A}_I^T\Delta x]$$

takže

$$\begin{bmatrix} \hat{G} & \hat{A}_I & A_E \\ \hat{A}_I^T & -\hat{M}_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\hat{u}_I \\ \Delta u_E \end{bmatrix} = - \begin{bmatrix} \hat{g} \\ (\hat{R}_I + \hat{C}_I)^{-1}\hat{R}_I\hat{\psi}_I \\ c_E \end{bmatrix} \quad (\text{NES})$$

kde

$$\begin{aligned} \hat{G} &= G + \check{A}_I\check{M}_I^{-1}\check{A}_I^T \\ \hat{g} &= g + \check{A}_I\check{M}_I^{-1}(\check{R}_I + \check{C}_I)^{-1}\check{R}_I\check{\psi}_I \end{aligned}$$

Nejprve určíme  $\Delta x$ ,  $\Delta\hat{u}$ ,  $\Delta u_E$  (nepřesným řešením soustavy (NES)). Pak vypočteme  $\Delta\check{u}$ .

**Poznámka 119** Soustavy (NES) a (IPS) mají velmi podobné vlastnosti. Platí opět Poznámky 108 a 109.

**Poznámka 120** K výběru délky kroku se používá rozšířená Lagrangeova funkce

$$\begin{aligned} P(\alpha) &= F(x + \alpha\Delta x) \\ &+ (u_I + \Delta u_I)^T c_I(x + \alpha\Delta x) + (u_E + \Delta u_E)^T c_E(x + \alpha\Delta x) \\ &+ \frac{\sigma}{2}\|\psi_I(u_I + \alpha\Delta u_I, -c_I(x + \alpha\Delta x))\|^2 + \frac{\sigma}{2}\|c_E(x + \alpha\Delta x)\|^2 \end{aligned}$$

kde  $\sigma > 0$ .

**Věta 111** *Nechť  $\Delta x$  a  $\Delta u_I$  jsou vektory určené podle Poznámky 118 a necht'*

$$\sigma > -\frac{(\Delta x)^T G \Delta x}{\|\psi_I\|^2 + \|c_E\|^2}$$

*Pak řešíme-li soustavu (NES) dostatečně přesně, platí  $P'(0) < 0$ .*

## Řešení lineárních KKT systémů

Soustava

$$K\bar{d} = \begin{bmatrix} \hat{G} & \hat{A}_I & A_E \\ \hat{A}_I^T & -\hat{M}_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \begin{bmatrix} d \\ \hat{d}_I \\ d_E \end{bmatrix} = \begin{bmatrix} b \\ \hat{b}_I \\ b_E \end{bmatrix} = \bar{b} \quad (\text{K})$$

Předpodmiňovač

$$C = \begin{bmatrix} \hat{D} & \hat{A}_I & A_E \\ \hat{A}_I^T & -\hat{M}_I & 0 \\ A_E^T & 0 & 0 \end{bmatrix} \quad (\text{P})$$

kde  $\hat{D}$  je pozitivně definitní diagonální aproximace matice  $\hat{G}$  (např. její diagonála).

**Algoritmus 9** (metoda PCG)

Data:  $0 \leq \omega < 1$ ,  $\bar{d} \in R^{n+\hat{m}_I+m_E}$  (např.  $\bar{d} = 0$ ).

Set  $\bar{r} := \bar{b} - K\bar{d}$ ,  $\beta := 0$

While  $\|\bar{r}\| > \omega\|\bar{b}\|$  do

$$\begin{aligned}\tilde{r} &:= C^{-1}\bar{r}, & \gamma &:= \bar{r}^T\tilde{r}, & \beta &:= \beta\gamma \\ \bar{p} &:= \tilde{r} + \beta\bar{p}, & \bar{q} &:= K\bar{p}, & \alpha &:= \gamma/\bar{p}^T\bar{q} \\ \bar{d} &:= \bar{d} + \alpha\bar{p}, & \bar{r} &:= \bar{r} - \alpha\bar{q}, & \beta &:= 1/\gamma\end{aligned}$$

end while

**Poznámka 121** Matice  $K$  je indefinitní. Metoda CG může selhat. Předpoklášovač  $C$  je také indefinitní. Kompenzuje indefinitnost matice  $K$ .

**Věta 112** Uvažujme předpoklášovač ( $P$ ) aplikovaný na soustavu ( $K$ ) a předpokládejme, že  $\hat{G} - \hat{D}$  je regulární. Pak matice  $KC^{-1}$  má alespoň  $\hat{m}_I + 2m_E$  jednotkových vlastních čísel, ale nanejvýš  $\hat{m}_I + m_E$  jim odpovídajících lineárně nezávislých vlastních vektorů. Zbylá vlastní čísla matice  $KC^{-1}$  jsou vlastními čísly matice

$$Z_E^T \tilde{G} Z_E (Z_E^T \tilde{D} Z_E)^{-1}$$

kde  $[A_E, Z_E]$  je regulární,  $Z_E^T A_E = 0$ ,  $Z_E^T Z_E = I$ , a kde

$$\begin{aligned}\tilde{G} &= \hat{G} + \hat{A}_I \hat{M}_I^{-1} \hat{A}_I^T \\ \tilde{D} &= \hat{D} + \hat{A}_I \hat{M}_I^{-1} \hat{A}_I^T\end{aligned}$$

Je-li  $Z_E^T \tilde{G} Z_E$  pozitivně definitní, jsou všechna vlastní čísla kladná.

**Věta 113** Uvažujme předpoklášovač ( $P$ ) aplikovaný na systém ( $K$ ) a předpokládejme, že matice  $\hat{G} - \hat{D}$  je regulární. Pak Krylovův prostor

$$\mathcal{K} = \text{span} \{ \bar{r}, KC^{-1}\bar{r}, (KC^{-1})^2\bar{r}, \dots \}$$

má dimenzi nejvýše  $\min(n+1, n-m_E+2)$ .

**Věta 114** Uvažujme Algoritmus 9 s předpoklášovačem ( $P$ ) aplikovaný na systém ( $K$ ). Předpokládejme, že počáteční  $\bar{d} \in R^{n+\hat{m}_I+m_E}$  je zvolen tak, že  $\hat{r}_I = 0$  a  $r_E = 0$ . Nechť  $Z_E^T \tilde{G} Z_E$  je pozitivně definitní. Pak

- Vektor  $d^*$  (první část vektoru  $\bar{d}^*$ , který je řešením soustavy ( $K$ )) je nalezen po nejvýše  $n - m_E$  iteracích.
- Algoritmus nemůže skončit na dělení nulou dříve, než je nalezen vektor  $d^*$ .
- Chyba  $\|d - d^*\|$  konverguje k nule přinejmenším  $R$ -lineárně s kvocientem

$$q \leq \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

kde  $\kappa$  je spektrální číslo podmíněnosti matice  $Z_E^T \tilde{G} Z_E (Z_E^T \tilde{D} Z_E)^{-1}$ .

- Jestliže  $d = d^*$ , pak také  $\hat{d}_I = \hat{d}_I^*$  a  $d_E^*$  lze určit podle vzorce

$$d_E^* = d_E + (A_E^T \tilde{D}^{-1} A_E)^{-1} A_E^T \tilde{D}^{-1} r$$

# Řešení zobecněných minimaxových úloh

**Definice 66** Řekneme, že  $F : R^n \rightarrow R$  je zobecněnou minimaxovou funkcí, jestliže

$$F(x) = h(F_1(x), \dots, F_m(x)), \quad F_i(x) = \max_{1 \leq j \leq n_i} f_{ij}(x), \quad 1 \leq i \leq m,$$

kde  $h : R^m \rightarrow R$  and  $f_{ij} : R^n \rightarrow R$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n_i$ , jsou hladké funkce splňující tyto předpoklady.

**Předpoklad 1.** Funkce  $F_i(x)$ ,  $1 \leq i \leq m$ , jsou zdola omezené na  $R^n$  (existují čísla  $\underline{F}_i \in R$  taková, že  $F_i(x) \geq \underline{F}_i$ ,  $1 \leq i \leq m$ , pro všechna  $x \in R^n$ ).

**Předpoklad 2.** Funkce  $h \in C^2 : R^m \rightarrow R$  je konvexní a platí

$$\partial h(z)/\partial z_i \geq \underline{h}_i > 0, \quad 1 \leq i \leq m,$$

pro všechna  $z \in Z = \{z \in R^m : z_i \geq \underline{F}_i, 1 \leq i \leq m\}$ .

**Předpoklad 3.** Funkce  $f_{ij}(x)$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n_i$ , jsou dvakrát spojitě diferencovatelné na konvexním obalu množiny

$$\mathcal{L}(\bar{F}) = \{x \in R^n : F_i(x) \leq \bar{F}, 1 \leq i \leq m\}$$

pro dostatečně velkou horní mez  $\bar{F}$  a mají omezené první a druhé derivace na  $\text{conv } \mathcal{L}(\bar{F})$  (existují čísla  $\bar{g}$  a  $\bar{G}$  taková, že  $\|\nabla f_{ij}(x)\| \leq \bar{g}$ ,  $\|\nabla^2 f_{ij}(x)\| \leq \bar{G}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n_i$ , pro všechna  $x \in \text{conv } \mathcal{L}(\bar{F})$ ).

**Příklad 15** Nejjednodušší zobecněnou minimaxovou funkcí je součet

$$F(x) = \sum_{i=1}^m F_i(x) = \sum_{i=1}^m \max_{1 \leq j \leq n_i} f_{ij}(x).$$

V tomto případě  $\partial h(z)/\partial z_i = 1$ ,  $1 \leq i \leq m$ , pro libovolný vektor  $z$  a matice  $H(z)$  je diagonální. Pro  $m = 1$  dostaneme klasickou minimaxovou úlohu.

$$F(x) = \max_{1 \leq i \leq m} f_i(x).$$

**Příklad 16** Volbou  $F_i(x) = |f_i(x)| = \max(f_i(x), -f_i(x))$ ,  $1 \leq i \leq m$  dostaneme funkci obsahující absolutní hodnoty. Klasickým případem je součet absolutních hodnot

$$F(x) = \sum_{i=1}^m |f_i(x)|.$$

**Poznámka 122** Minimalizace funkce  $F(x)$  je ekvivalentní úloze nelineárního programování: minimalizovat funkci

$$h(z_1, \dots, z_m)$$

s omezeními

$$f_{ij}(x) \leq z_i, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n_i$$

(podmínky  $\partial h(z)/\partial z_i \geq \underline{h}_i > 0$ ,  $1 \leq i \leq m$ , pro  $z \in Z$  postačují k tomu, aby platilo  $z_i = F_i(x)$ ,  $1 \leq i \leq m$ , v bodě minima). Nutné KKT podmínky pro řešení této úlohy mají tvar

$$\sum_{i=1}^m \sum_{j=1}^{n_i} u_{ij} \nabla f_{ij}(x) = 0, \quad \sum_{j=1}^{n_i} u_{ij} = \frac{\partial h(z)}{\partial z_i},$$

$$u_{ij} \geq 0, \quad z_i - f_{ij}(x) \geq 0, \quad u_{ij}(z_i - f_{ij}(x)) = 0,$$

kde  $u_{ij}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n_i$ , jsou Lagrangeovy multiplikátory.

**Poznámka 123** Úlohu nelineárního programování lze řešit primární metodou vnitřních bodů: aplikujeme Newtonovu metodu na posloupnost bariérových funkcí

$$B_\mu(x, z) = h(z) - \mu \sum_{i=1}^m \sum_{j=1}^{n_i} \log(z_i - f_{ij}(x)),$$

kde  $\mu > 0$  a  $\mu \rightarrow 0$ .

**Poznámka 124** Nutné KKT podmínky pro extrém bariérové funkce mají tvar

$$\nabla_x B_\mu(x, z) = \sum_{i=1}^m \sum_{j=1}^{n_i} \nabla f_{ij}(x) \frac{\mu}{z_i - f_{ij}(x)} = 0,$$

$$\frac{\partial B_\mu(x, z)}{\partial z_i} = \frac{\partial h(z)}{\partial z_i} + \sum_{j=1}^{n_i} \frac{\mu}{z_i - f_{ij}(x)} = 0, \quad 1 \leq i \leq m.$$

**Poznámka 125** Minimalizaci bariérové funkce můžeme chápat jako dvojúrovňovou optimalizaci

$$z(x) = \arg \min_{z \in Z} B_\mu(x, z),$$

$$x = \arg \min_{x \in R^n} B(x; \mu), \quad B(x; \mu) \triangleq B_\mu(x, z(x)),$$

kde  $Z$  je množina použitá v předpokladu 2.

**Věta 115** *Soustava*

$$\frac{\partial h(z)}{\partial z_i} - \sum_{j=1}^{n_i} \frac{\mu}{z_i - f_{ij}(x)} = 0, \tag{MS}$$

má pro pevné  $x \in R^n$  jednoznačné řešení  $z(x; \mu) \in Z \subset R^m$  takové, že

$$F_i(x) < \underline{z}_i \leq z_i(x; \mu) \leq \bar{z}_i, \quad 1 \leq i \leq m,$$

kde

$$\underline{z}_i = F_i(x) + \mu/\bar{h}_i, \quad \bar{z}_i = F_i(x) + n_i\mu/\underline{h}_i,$$

a kde  $\underline{h}_i > 0$  jsou meze použité v předpokladu 2 a  $\bar{h}_i = h_i(\bar{z}_1, \dots, \bar{z}_m)$ .

Umíme-li najít řešení soustavy nelineárních rovnic, můžeme se omezit na nepodmíněnou minimalizaci funkce  $B(x; \mu) = B_\mu(x, z(x))$ . Je účelné znát gradient a Hessovu matici funkce  $B(x; \mu)$ .

**Věta 116** Označme

$$u_{ij}(x) = \frac{\mu}{z_i - f_{ij}(x)}, \quad v_{ij}(x) = \frac{\mu}{(z_i - f_{ij}(x))^2}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n_i.$$

Pak

$$\nabla B(x; \mu) = \sum_{i=1}^m \sum_{j=1}^{n_i} \nabla f_{ij}(x) u_{ij}(x),$$

$$\nabla^2 B(x; \mu) = W(x, z(x)) - C(x, z(x)) D(x, z(x))^{-1} C^T(x, z(x)),$$

kde

$$W(x, z) = \sum_{i=1}^m \sum_{j=1}^{n_i} \nabla^2 f_{ij}(x) u_{ij}(x) + \sum_{i=1}^m \sum_{j=1}^{n_i} \nabla f_{ij}(x) v_{ij}(x) (\nabla f_{ij}(x))^T,$$

$$C(x, z) = \left[ \sum_{j=1}^{n_1} \nabla f_{1j}(x) v_{1j}(x), \dots, \sum_{j=1}^{n_m} \nabla f_{mj}(x) v_{mj}(x) \right],$$

$$D(x, z) = \nabla^2 h(z) + \text{diag} \left( \sum_{j=1}^{n_1} v_{1j}(x), \dots, \sum_{j=1}^{n_m} v_{mj}(x) \right).$$

**Věta 117** Necht' matice  $G = \sum_{i=1}^m \sum_{j=1}^{n_i} \nabla^2 f_{ij}(x) u_{ij}(x)$  je pozitivně definitní. Pak matice  $\nabla B(x; \mu)$  je pozitivně definitní.

**Poznámka 126** Platí

$$\begin{aligned} (\nabla^2 B(x; \mu))^{-1} &= W(x, z(x))^{-1} - W(x, z(x))^{-1} C(x, z(x)) \\ &\quad \left( C^T(x, z(x)) W^{-1}(x, z(x)) C(x, z(x)) - D(z(x)) \right)^{-1} \\ &\quad C^T(x, z(x)) W(x, z(x))^{-1}. \end{aligned}$$

**Poznámka 127** Iterační krok primární metody vnitřních bodů vypadá takto:

- Na začátku iteračního kroku známe  $x \in R^n$  a  $\mu > 0$ .
- Řešením soustavy (MS) určíme vektor  $z(x; \mu)$ .
- Určíme směrový vektor  $s = -(\nabla^2 B(x; \mu))^{-1} \nabla B(x; \mu)$ .
- Určíme délku kroku  $\alpha > 0$  tak, aby platilo  $B_\mu(x + \alpha s, z(x + \alpha s; \mu)) < B_\mu(x, z(x; \mu))$  (vektor  $z(x + \alpha s; \mu)$  se určuje řešením soustavy (MS)).
- Položíme  $x^+ = x + \alpha \Delta x$  a určíme novou hodnotu  $0 < \mu^+ < \mu$ .

**Poznámka 128** Určení nové hodnoty  $0 < \mu^+ < \mu$ :

- Pokud  $\|g(x_k; \mu_k)\|^2 \geq \rho \mu_k$ , položíme  $\mu_{k+1} = \mu_k$  (bariérový parametr se nemění).
- Pokud  $\|g(x_k; \mu_k)\|^2 < \rho \mu_k$ , položíme  $\mu_{k+1} = \max(\underline{\mu}, \|g_k(x_k; \mu_k)\|^2)$ .

Používají se hodnoty  $\underline{\mu} = 10^{-10}$  a  $\rho = 0.1$ .

**Věta 118** Necht' jsou splněny předpoklady 1–3. Pak primární metoda vnitřních bodů pro minimalizaci zobecněné minimaxové funkce je globálně konvergentní.

## Lipschitzovské funkce

**Definice 67** Řekneme, že funkce  $F : R^n \rightarrow R$  je lipschitzovská v okolí bodu  $x \in R^n$  (s konstantou  $L$ ), jestliže existuje číslo  $\varepsilon > 0$  takové, že platí

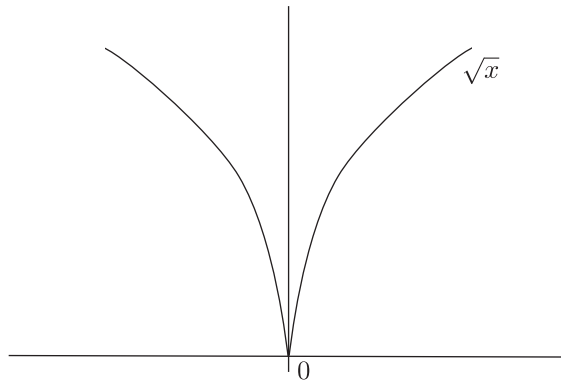
$$|F(x_2) - F(x_1)| \leq L\|x_2 - x_1\|, \quad (5)$$

pokud  $x_1 \in B(x, \varepsilon)$  a  $x_2 \in B(x, \varepsilon)$ .

**Příklad 17** Funkce  $\sqrt{x}$  není lipschitzovská v okolí bodu  $x = 0$ . Platí

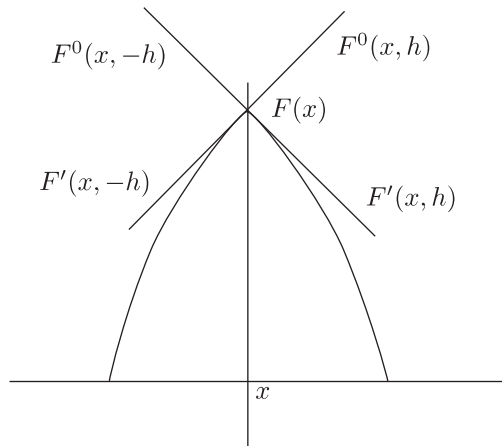
$$\sqrt{x_2} - \sqrt{x_1} = \frac{x_2 - x_1}{\sqrt{x_2} + \sqrt{x_1}},$$

takže pro libovolné číslo  $L > 0$  platí  $\sqrt{x_2} - \sqrt{x_1} > L(x_2 - x_1)$ , pokud  $x_1 < 1/(2L)$  a  $x_2 < 1/(2L)$ .



**Definice 68** Zobecněnou (Clarkovu) směrovou derivaci funkce  $F : R^n \rightarrow R$  v bodě  $x \in R^n$  ve směru  $h \in R^n$  definujeme předpisem

$$F^0(x, h) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{F(y + th) - F(y)}{t}. \quad (6)$$



**Věta 119** Necht funkce  $F : R^n \rightarrow R$  je lipschitzovská (s konstantou  $L$ ) v okolí bodu  $x \in R^n$ . Pak:

(a) Funkce  $F^0(x, \cdot) : R^n \rightarrow R$  je pozitivně homogenní, subaditivní a lipschitzovská s konstantou  $L$ .

(b) Funkce  $F^0(\cdot, \cdot) : R^n \times R^n \rightarrow R$  je shora polospojité, neboli

$$\limsup_{i \rightarrow \infty} F^0(x_i, h_i) \leq F^0(x, h),$$

kdykoliv  $x_i \rightarrow x$  a  $h_i \rightarrow h$ .

(c) Platí  $F^0(x, -h) = (-F)^0(x, h) \forall h \in R^n$ .

**Definice 69** Necht funkce  $F : R^n \rightarrow R$  je lipschitzovská v okolí bodu  $x \in R^n$ . Pak množinu

$$\partial F(x) = \{g \in R^n : F^0(x, h) \geq g^T h \quad \forall h \in R^n\}$$

nazveme subdiferenciálem funkce  $F$  v bodě  $x$ . Prvky  $g \in \partial F(x)$  budeme nazývat subgradienty funkce  $F$  v bodě  $x$ .

**Poznámka 129** Necht funkce  $F : R^n \rightarrow R$  je konvexní v okolí bodu  $x \in R^n$ . Pak platí

$$\begin{aligned} F^0(x, h) &= F'(x, h) \quad \forall h \in R^n, \\ \partial F(x) &= \{g \in R^n : F'(x, h) \geq g^T h \quad \forall h \in R^n\} \end{aligned}$$

(oba subdiferenciály splývají).

**Definice 70** Necht funkce  $F : R^n \rightarrow R$  je lipschitzovská v okolí bodu  $x \in R^n$ . Jestliže  $F'(x, h)$  existuje a platí  $F^0(x, h) = F'(x, h) \forall h \in R^n$ , řekneme, že  $F$  je regulární v bodě  $x$ .

**Věta 120** Spojitě diferencovatelné a konvexní funkce jsou regulární. Dále jsou regulární:  
(a) Nezáporné lineární kombinace regulárních funkcí, tedy funkce tvaru

$$F(x) = \sum_{i=1}^m \lambda_i f_i(x).$$

(b) Bodová maxima regulárních funkcí, tedy funkce tvaru

$$F(x) = \max_{1 \leq i \leq m} f_i(x).$$

(kde  $\lambda_i \geq 0$  a  $f_i$  jsou regulární).

**Věta 121** Necht funkce  $F : R^n \rightarrow R$  je lipschitzovská (s konstantou  $L$ ) v okolí bodu  $x \in R^n$ . Pak:



(a) Subdiferenciál  $\partial F(x)$  je neprázdná konvexní kompaktní množina taková, že  $\|g\| \leq L$   
 $\forall g \in \partial f(x)$ .

(b) Platí

$$F^0(x, h) = \max \{g^T h : g \in \partial F(x)\} \quad \forall h \in R^n.$$

(c) Subdiferenciál je shora polospojité (jestliže  $x_i \rightarrow x$ ,  $g_i \in \partial F(x_i)$  a  $g_i \rightarrow g$ , pak  $g \in \partial F(x)$ ).

(d) Platí  $\partial(-F)(x) = -\partial F(x)$ .

**Poznámka 130** Podle věty 121 (b) je zobecněná směrrová derivace  $F^0(x, h)$  opěrnou funkcí subdiferenciálu  $\partial F(x)$ , neboli

$$F^0(x, h) = \delta_{\partial F(x)}(h).$$

**Příklad 18** Necht  $F(x) = -\|x\|$ . Pak

$$F^0(0, h) = -\|0, h\|^0 = \|0, -h\|^0 = \|0, -h\|' = \|h\|,$$

$$\partial F(0) = -\partial\|0\| = \partial\|0\| = \{g \in R^n : \|g\| \leq 1\}.$$

**Věta 122** Necht funkce  $F : R^n \rightarrow R$  je spojitě diferencovatelná v bodě  $x \in R^n$ . Pak  $F$  je lipschitzovská v okolí bodu  $x$  a platí

$$\partial F(x) = \{\nabla F(x)\}. \quad (7)$$

**Věta 123** Necht funkce  $F : R^n \rightarrow R$  je lokálně lipschitzovská v okolí bodu  $x \in R^n$ , který je jejím lokálním extrémem (minimem nebo maximem). Pak platí

$$0 \in \partial F(x).$$

**Věta 124** (o střední hodnotě). Necht funkce  $F : R^n \rightarrow R$  je lokálně lipschitzovská na otevřené množině obsahující úsečku  $[x, y]$ . Pak existuje bod  $z \in (x, y)$  takový, že

$$F(y) - F(x) \in (\partial F(z))^T(y - x).$$

**Definice 71** Necht funkce  $F : R^n \rightarrow R$  je lokálně lipschitzovská v okolí bodu  $x \in R^n$ . Pak množinu

$$\partial_B F(x) = \left\{ \lim_{i \rightarrow \infty} \nabla F(x_i) : x_i \rightarrow x, \nabla F(x_i) \text{ existuje} \right\}$$

nazveme  $B$ -diferenciálem funkce  $F$ .

**Věta 125** (Clarke). *Nechť funkce  $F : R^n \rightarrow R$  je lipschitzovská v okolí bodu  $x \in R^n$ . Pak platí*

$$\partial F(x) = \text{conv } \partial_B F(x).$$

**Příklad 19** *Nechť  $\|\cdot\| : R^n \rightarrow R$  je euklidovská norma. Pak  $\|\cdot\|$  je spojitě diferencovatelná v každém bodě  $x \neq 0$  (platí  $\nabla\|x\| = x/\|x\|$ ). Jestliže  $x_i = t_i h$ , pak*

$$\lim_{t_i \downarrow 0} \nabla\|x_i\| = \lim_{t_i \downarrow 0} \frac{t_i h}{t_i \|h\|} = \frac{h}{\|h\|}$$

takže

$$\partial_B\|0\| = \left\{ \frac{h}{\|h\|} : h \in R^n \right\}$$

(jednotková sféra) a

$$\partial\|0\| = \text{conv}\partial_B\|0\| = \{h \in R^n : \|h\| \leq 1\}$$

(uzavřená jednotková koule).

## Lipschitzovská zobrazení

Je-li zobrazení  $f : R^n \rightarrow R^m$  spojitě diferencovatelné v bodě  $x \in R^n$ , existuje Jacobiova matice

$$\mathcal{J}f(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix}.$$

Je-li zobrazení  $f : R^n \rightarrow R^m$  lokálně lipschitzovské, má množina bodů v nichž  $\mathcal{J}$  neexistuje míru nula (Rademacherova věta). Proto je opodstatněná tato definice.

**Definice 72** *Nechť zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské v okolí bodu  $x \in R^n$ . Pak množinu*

$$\partial f(x) = \text{conv } \partial_B f(x),$$

kde

$$\partial_B f(x) = \left\{ \lim_{i \rightarrow \infty} \mathcal{J}f(x_i) : x_i \rightarrow x, \mathcal{J}f(x_i) \text{ existuje} \right\},$$

nazveme zobecněným Jakobiánem zobrazení  $f$ .

**Věta 126** *Nechť zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské (s konstantou  $L$ ) v okolí bodu  $x \in R^n$ . Pak*

(a) *Platí*

$$\partial f(x) \subset \begin{bmatrix} \partial f_1(x) \\ \vdots \\ \partial f_m(x) \end{bmatrix},$$

kde  $\partial f_i(x)$  jsou subdiferenciály složek  $f_i : R^n \rightarrow R$ ,  $1 \leq i \leq m$ , v bodě  $x \in R^n$ .

- (b) Zobecněný Jakobián  $\partial f(x)$  je neprázdná konvexní kompaktní množina taková, že  $\|J\| \leq L \forall J \in \partial f(x)$ .
- (c) Zobecněný Jakobián  $\partial f(x)$  je shora polospojité (jestliže  $x_i \rightarrow x$ ,  $J_i \in \partial f(x_i)$  a  $J_i \rightarrow J$ , pak  $J \in \partial f(x)$ ).

**Poznámka 131** Nechť  $[x, y]$  je uzavřený interval. Zavedeme označení

$$\partial f([x, y]) = \text{conv} \bigcup_{z \in [x, y]} \partial f(z).$$

Množina  $\partial f([x, y])$  je kompaktní a konvexní.

**Věta 127** (o střední hodnotě). Nechť zobrazení  $f : R^n \rightarrow R^m$  je lokálně lipschitzovské na otevřené množině obsahující úsečku  $[x, y]$ . Pak platí

$$f(y) - f(x) \in \partial f([x, y])(y - x).$$

**Věta 128** (o složeném zobrazení). Nechť zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské v okolí bodu  $x \in R^n$  a funkce  $F : R^m \rightarrow R$  je lipschitzovská v okolí bodu  $f(x)$ . Pak funkce  $\varphi = F \circ f : R^n \rightarrow R$  (t.j.  $\varphi(x) = F(f(x))$ ) je lipschitzovská v okolí bodu  $x \in R^n$  a platí

$$\partial \varphi(x) \subset \text{conv} \left\{ J^T v : J \in \partial f(x), v \in \partial F(f(x)) \right\},$$

příčemž rovnost nastává například tehdy, když:

- (a) Funkce  $F$  je spojitě diferencovatelná v bodě  $f(x)$ .
- (b) Funkce  $F$  je regulární v bodě  $f(x)$  a zobrazení  $f$  je spojitě diferencovatelné v bodě  $x$ .
- (c) Funkce  $F$  je regulární v bodě  $f(x)$ , složky  $f_i : R^n \rightarrow R$ ,  $1 \leq i \leq m$ , jsou regulární v bodě  $x$  a pro libovolný prvek  $v \in \partial F(f(x))$  platí  $v \geq 0$  (t.j.  $v_i \geq 0$ ,  $1 \leq i \leq m$ ).

**Důsledek 9** Nechť funkce  $f_1 : R^n \rightarrow R$ ,  $f_2 : R^n \rightarrow R$  jsou lipschitzovské v okolí bodu  $x \in R^n$ . Pak funkce  $\varphi = f_1 f_2$  je lipschitzovská v okolí bodu  $x$  a platí

$$\partial \varphi(x) \subset \partial f_1(x) f_2(x) + f_1(x) \partial f_2(x)$$

příčemž rovnost nastává, jsou-li funkce  $f_1, f_2$  regulární a platí-li  $f_1(x) \geq 0$ ,  $f_2(x) \geq 0$ . V tomto případě je funkce  $\varphi = f_1 f_2$  regulární.

**Důsledek 10** Nechť zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské v okolí bodu  $x \in R^n$ . Pak funkce  $\varphi = (1/2) f^T f$  je lipschitzovská v okolí bodu  $x$  a platí

$$\partial \varphi(x) = \left\{ J^T f(x) : J \in \partial(f(x)) \right\}.$$

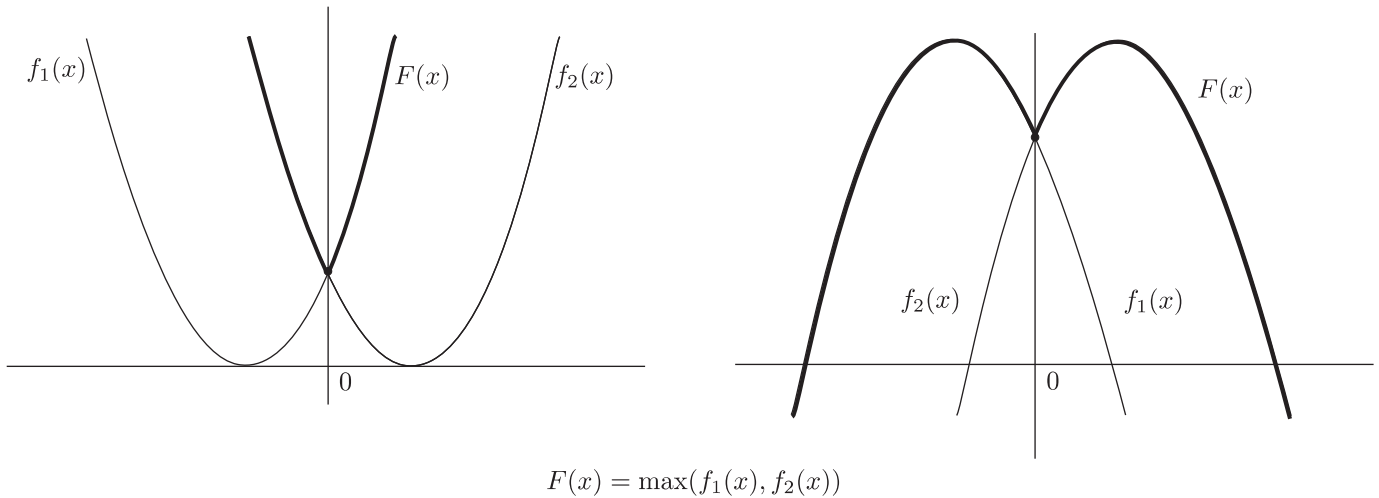
**Věta 129** Necht funkce  $f_i : R^n \rightarrow R$ ,  $1 \leq i \leq m$ , jsou lipschitzovské v okolí bodu  $x \in R^n$ . Pak funkce

$$\varphi(x) = \max_{1 \leq i \leq m} f_i(x)$$

je lipschitzovská v okolí bodu  $x$  a platí

$$\partial\varphi(x) \subset \text{conv} \{ \partial f_i(x) : i \in I(x) \},$$

kde  $I(x) = \{i \in \{1, \dots, m\} : f_i(x) = \varphi(x)\}$ . Jsou-li funkce  $f_i$ ,  $1 \leq i \leq m$ , regulární v bodě  $x$ , je funkce  $\varphi$  regulární v bodě  $x$  a místo inkluze platí rovnost.



## Polohladká zobrazení

**Definice 73** Necht zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské v okolí bodu  $x \in R^n$ . Jestliže pro každé  $h \in R^n$  existuje limita

$$\lim_{\substack{J \in \partial f(x+th') \\ h' \rightarrow h, t \downarrow 0}} Jh'$$

(nezávislá na volbě  $J \in \partial f(x+th')$ ), řekneme, že zobrazení  $f$  je polohladké v bodě  $x$ .

**Věta 130** Necht zobrazení  $f : R^n \rightarrow R^m$  je polohladké v bodě  $x \in R^n$ . Pak pro libovolný vektor  $h \in R^n$  existuje směrová derivace

$$f'(x, h) = \lim_{t \downarrow 0} \frac{f(x+th) - f(x)}{t}$$

a platí

$$f'(x, h) = \lim_{\substack{J \in \partial f(x+th') \\ h' \rightarrow h, t \downarrow 0}} Jh'.$$

**Věta 131** Zobrazení, jehož složky jsou spojitě diferencovatelné, konvervní nebo polohladké je polohladké.

**Věta 132** (o složeném zobrazení). Necht' zobrazení  $f : R^n \rightarrow R^m$  je polohladké v bodě  $x \in R^n$  a funkce  $F : R^m \rightarrow R$  je polohladká v bodě  $f(x)$ . Pak složené zobrazení  $\varphi = F \circ f$  je polohladké v bodě  $x$ .

**Důsledek 11** Necht' funkce  $f_i : R^n \rightarrow R$ ,  $1 \leq i \leq m$ , jsou polohladké v bodě  $x \in R^n$  a  $\lambda_i \in R$ ,  $1 \leq i \leq m$ . Pak například tyto funkce jsou polohladké:

(a) Lineární kombinace

$$\varphi = \sum_{i=1}^m \lambda_i f_i$$

(b) Součin

$$\varphi = \prod_{i=1}^m f_i$$

(c) Libovolná norma

$$\varphi = \left\| \begin{bmatrix} f_1(x) \\ \dots \\ f_m(x) \end{bmatrix} \right\|,$$

tedy například

$$\varphi = \sum_{i=1}^m |f_i| \quad \text{nebo} \quad \varphi = \max_{1 \leq i \leq m} |f_i|$$

(d) Bodové maximum

$$\varphi = \max_{1 \leq i \leq m} f_i$$

**Důsledek 12** Složky polohladkého zobrazení jsou polohladké funkce. Lineární kombinace polohladkých zobrazení je polohladké zobrazení. Skalární součin polohladkých zobrazení je polohladká funkce.

**Příklad 20** Fischerova–Burmeisterova funkce  $\psi : R^2 \rightarrow R$  je definovaná předpisem

$$\psi(x) = \sqrt{x_1^2 + x_2^2} - (x_1 + x_2)$$

(a)  $\psi(x) = 0$  platí právě tehdy, když  $x_1 \geq 0$ ,  $x_2 \geq 0$  a  $x_1 x_2 = 0$ . Skutečně

$$x_1 < 0 \Rightarrow \psi(x) = \sqrt{|x_1|^2 + x_2^2} + |x_1| - x_2 \geq |x_1| + |x_2| - x_2 > 0,$$

$$x_2 < 0 \Rightarrow \psi(x) = \sqrt{x_1^2 + |x_2|^2} - x_1 + |x_2| \geq |x_1| - x_1 + |x_2| > 0,$$

$$x_1 > 0, x_2 > 0 \Rightarrow \psi(x) = \sqrt{x_1^2 + x_2^2} - (x_1 + x_2) < \sqrt{x_1^2 + 2x_1 x_2 + x_2^2} - (x_1 + x_2) = 0,$$

$$x_1 = 0, x_2 \geq 0 \Rightarrow \psi(x) = |x_2| - x_2 = 0,$$

$$x_1 \geq 0, x_2 = 0 \Rightarrow \psi(x) = |x_1| - x_1 = 0.$$

(b) V bodě  $x \neq 0$  je  $\psi(x)$  spojitě diferencovatelná a platí

$$\nabla \psi(x) = \begin{bmatrix} \frac{x_1}{\sqrt{x_1^2 + x_2^2}} - 1 \\ \frac{x_2}{\sqrt{x_1^2 + x_2^2}} - 1 \end{bmatrix}.$$

V bodě  $x = 0$  je  $\psi(x)$  polohladká (je rozdílem konvexní funkce  $\sqrt{x_1^2 + x_2^2}$  a lineární funkce  $x_1 + x_2$ ). Platí

$$\partial_B \psi(0) = \bigcup_{\varphi \in [0, 2\pi]} [\cos \varphi - 1, \sin \varphi - 1] = S(-e, 1)$$

(jednotková sféra se středem v bodě  $[-1, -1]$ ) a

$$\partial \psi(0) = \text{conv} \partial_B \psi(0) = \text{conv} S(-e, 1) = \overline{B(-e, 1)}$$

(uzavřená jednotková koule se středem v bodě  $[-1, -1]$ ).

**Věta 133** Necht zobrazení  $f : R^n \rightarrow R^m$  je lipschitzovské v okolí bodu  $x \in R^n$ . Pak  $f$  je polohladké v bodě  $x$  právě tehdy, existuje-li směrová derivace  $f'(x, h)$  a platí-li

$$Jh - f'(x, h) = o(\|h\|)$$

pokud  $h \rightarrow 0$  a  $J \in \partial f(x + h)$ .

**Věta 134** Necht zobrazení  $f : R^n \rightarrow R^m$  je polohladké v bodě  $x \in R^n$ . Pak platí

$$f(x + h) - f(x) - f'(x, h) = o(\|h\|)$$

a

$$f(x + h) - f(x) - Jh = o(\|h\|).$$

pokud  $h \rightarrow 0$  a  $J \in \partial f(x + h)$ .

**Poznámka 132** Věty 133–134 mají velký význam pro řešení polohladkých rovnic. Vyplývá z nich, že prvky zobecněného Jakobiánu  $\partial f(x)$  dobře vystihují chování polohladkého zobrazení  $f(x)$ , takže k řešení rovnice  $f(x) = 0$  lze použít Newtonovu metodu, která místo Jacobiovy matice používá libovolný prvek zobecněného Jakobiánu.