Numerické metody pro nepodmíněnou optimalizaci. Učební text

Lukšan, Ladislav
1994

NUMERICKÉ METODY
PRO NEPODMÍŇENOU OPTIMALIZACI

L. Lukšan

Výzkumná zpráva č. V-590, květen 1994

1. Úvod

1.1. Základní pojmy

Budeme používat označování:

$$x \in R^n \text{ pro } n - \text{dimensionální vektor}$$

$$F(x) \text{ pro funkci } F : R^n \to R$$

$$g(x) = [\partial F/\partial x_1, \ldots, \partial F/\partial x_n]^T$$

$$G(x) = \begin{bmatrix} \partial^2 F/\partial x_1^2, & \ldots, & \partial^2 F/\partial x_1 \partial x_n \\ \vdots & \vdots & \vdots \\ \partial^2 F/\partial x_n \partial x_1, & \ldots, & \partial^2 F/\partial x_n^2 \end{bmatrix}$$

Zde $F(x)$ je účelová funkce, $g(x)$ je její gradient a $G(x)$ je její Hessova matice (matice druhých parciálních derivací). Spojitost druhých parciálních derivací implikuje symetrii matice $G(x)$. Při vyšetřování konvergence optimalizačních metod budeme často používat předpoklady (F1)-(F5):

Definice 1 Řekneme, že funkce $F : R^n \to R$ je zdola omezená jestliže platí

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

Definice 2 Řekneme, že funkce $F : R^n \to R$ má kompaktní hladiny, jestliže množina

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

je kompaktní $\forall \overline{F}$ (prázdná množina se předpokládá kompaktní).

Definice 3 Řekneme, že funkce $F : R^n \to R$ má omezené druhé derivace, jestliže platí

$$d^T G(x) d \leq \overline{G} \parallel d \parallel^2 \tag{F3}$$

$\forall x \in R^n$, $\forall d \in R^n$. Je to ekvivalentní podmínce $\parallel G(x) \parallel \leq \overline{G}$ $\forall x \in R^n$.

Poznámka 1 Místo omezenosti druhých derivací stačí obvykle lipschitzovskost prvních derivací:

$$\parallel g(x+d) - g(x) \parallel \leq \overline{G} \parallel d \parallel$$

$\forall x \in R^n$, $\forall d \in R^n$

Definice 4 Řekneme, že funkce $F : R^n \to R$ je stejnoměrně (nebo silně) konvexní, jestliže platí

$$d^T G(x) d \geq \underline{G} \parallel d \parallel^2 \tag{F4}$$

$\forall x \in R^n$, $\forall d \in R^n$.

Definice 5 Řekneme, že funkce $F : R^n \to R$ má lipschitzovské druhé derivace, jestliže

$$\parallel G(x+d) - G(x) \parallel \leq \overline{L} \parallel d \parallel \tag{F5}$$

$\forall x \in R^n$, $\forall d \in R^n$.

Při konvergenčních důkazech budeme často používat věty o střední hodnotě:

(a) $F(x+d) = F(x) + d^T g(x) + \frac{1}{2} d^T G(\tilde{x}) d$, kde $\tilde{x} = x + \tilde{\lambda} d$ a $0 \leq \tilde{\lambda} \leq 1$.
Z (F3) plyne

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

Z (F4) plyne

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

(b) $g(x+d) = g(x) + \int_0^1 G(x + \lambda d) d d\lambda$
z (F3) plyne

$$\parallel g(x+d) - g(x) \parallel \leq \overline{G} \parallel d \parallel$$
$$d^T (g(x+d) - g(x) \leq \overline{G} \parallel d \parallel^2$$

z (F4) plyne

$$\parallel g(x+d) - g(x) \parallel \geq \underline{G} \parallel d \parallel$$
$$d^T (g(x+d) - g(x) \geq \underline{G} \parallel d \parallel^2$$

Důkaz posledního tvrzení:

$$d^T(g(x+d)-g(x)) = \int_0^1 d^T G(x+\lambda d) d d\lambda \geq \int_0^1 \underline{G} \parallel d \parallel^2 d\lambda = \underline{G} \parallel d \parallel^2$$

$$\underline{G} \parallel d \parallel^2 \leq d^T(g(x+d)-g(x)) \leq \parallel d \parallel \parallel g(x+d)-g(x) \parallel$$

1.2. Podmínky optimality

Definice 6 Řekneme, že bod $x^* \in R^n$ je lokálním minimem funkce $F : R^n \to R$, jestliže existuje číslo $\varepsilon > 0$ takové, že

$$F(x^*) \leq F(x) \quad \forall x \in \mathcal{B}(x^*, \varepsilon)$$

kde $\mathcal{B}(x^*, \varepsilon) = \{x \in R^n : \parallel x - x^* \parallel < \varepsilon\}$. Jestliže navíc $F(x^*) < F(x)$ pokud $x^* \neq x$, řekneme, že bod $x^* \in R^n$ je izolovaným lokálním minimem funkce $F : R^n \to R$.

Tvrzení 1 Nechť bod $x^* \in R$ je lokálním minimem funkce $F : R^n \to R$ a nechť $F \in C^1$ (spojitě diferenciatelná) na $\mathcal{B}(x^*, \varepsilon)$. Pak platí

$$g(x^*) = 0$$

jestliže navíc $F \in C^2$ (dvakrát spojitě diferenciatelná) na $\mathcal{B}(x^*, \varepsilon)$, pak platí

$$G(x^*) \geq 0$$

(matice $G(x^*)$ je positivně semidefinitní)

Tvrzení 2 Nechť $F : R^n \to R \in C^2$ na $\mathcal{B}(x^*, \varepsilon)$ a nechť platí

$$g(x^*) = 0$$

a

$$G(x^*) > 0$$

(matice $G(x^*)$ je pozitivně definitní). Pak bod $x^* \in R^n$ je ostrým lokálním minimem funkce $F : R^n \to R$.

1.3. Základní pojmy z teorie konvergence

Nyní se budeme zabývat vlastnostmi konvergentních posloupností.

Definice 7 Nechť $x_i \in R^n$, $i \in N$, je posloupnost bodů. Jestliže pro libovolné $\varepsilon > 0$ existuje index $k \in N$ tak, že $x_i \in \mathcal{B}(x^*, \varepsilon)$ $\forall k \geq i$, řekneme, že posloupnost $x_i \in R^n$, $i \in N$ konverguje k bodu $x^* \in R^n$ a píšeme $x_i \to x^*$. Používáme značení $F_i = F(x_i)$, $g_i = g(x_i)$, $G_i = G(x_i)$.

Věta 3 Nechť $x_i \in R^n$, $d_i \in R^n$, $i \in N$, jsou dvě posloupnosti. Nechť $F \in C^2$ a je splněna podmínka (F5). Pak platí

$$F(x_i + d_i) - F(x_i) = d_i^T g_i + \frac{1}{2} d_i^T G_i d_i + O(\| d_i \|^3)$$

a

$$g(x_i + d_i) - g(x_i) = G_i d_i + O(\| d_i \|^2)$$

(zde $\| O(\xi_i) \| \leq C \| \xi_i \|$ pokud $\| \xi_i \| \to 0$)

Důkaz Z (F5) a z věty (a) o střední hodnotě plyne

$$\begin{aligned} F(x_i + d_i) - F(x_i) &= d_i^T g_i + \frac{1}{2} d_i^T (G(x_i + \tilde{\lambda} d_i)) d_i = \\ &= d_i^T g_i + \frac{1}{2} d_i^T G_i d_i + \frac{1}{2} d_i^T (G(x_i + \tilde{\lambda} d_i) - G_i) d_i \leq \\ &\leq d_i^T g_i + \frac{1}{2} d_i^T G_i d_i + \frac{1}{2} \| G(x_i + \tilde{\lambda} d_i) - G_i \| \| d_i \|^2 \leq \\ &\leq d_i^T g_i + \frac{1}{2} d_i^T G_i d_i + \frac{\tilde{\lambda}}{2} \overline{L} \| d_i \|^3 \end{aligned}$$

kde $0 \leq \tilde{\lambda} \leq 1$. Pobobně z (F5) a z věty (b) o střední hodnotě plyne

$$\begin{aligned} g(x_i + d_i) - g(x_i) &= \int_0^1 G(x_i + \lambda d_i) d_i d\lambda = \\ &= G_i d_i + \int_0^1 (G(x_i + \lambda d_i) - G_i) d_i d\lambda \end{aligned}$$

a

$$\begin{aligned} \| \int_0^1 (G(x_i + \lambda d_i) - G_i) d_i d\lambda \| &\leq \int_0^1 \| G(x_i + \lambda d_i) - G_i \| \| d_i \| d\lambda \leq \\ &\leq \overline{L} \int_0^1 \| d_i \|^2 d\lambda = \overline{L} \| d_i \|^2 \end{aligned}$$

Věta 4 Nechť $x_i \in R^n$, $e_i = x_i - x^*$, $i \in N$, kde $x^* \in R^n$ je lokální minimum funkce $F \in C^2$ která vyhovuje podmínce (F5). Pak platí

$$F(x_i) - F(x^*) = \frac{1}{2} e_i^T G^* e_i + O(\| e_i \|^3)$$

a

$$g(x_i) = G^* e_i + O(\| e_i \|^2)$$

Důkaz Stejný jako v předchozí větě. Provede se záměna x_i místo $x_i + d_i$, x^* místo x_i, $e_i = x_i - x^*$ místo $d_i = x_i + d_i - x_i$ a přihlédne se k tomu, že $g(x^*) = 0$.

Poznámka 2 Bez podmínky (F5) lze odvodit odhady

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

a

$$g(x_i) = g^* e_i + o(\| e_i \|) \| e_i \|$$

(zde $\| o(\xi_i) \| \to 0$ pokud $\| \xi_i \| \to 0$).

Definice 8 Nechť $x_i \to x^*$. Jestliže existuje index $k \in N$ a hodnoty $0 < M_k < \infty$ a $0 < q < 1$, tak že

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

$\forall i \geq k$, řekneme, že posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* R-lineárně.

Věta 5 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* R-lineárně právě tehdy jestliže

$$\limsup_{i \to \infty} \| x_i - x^* \|^{\frac{1}{i}} = \hat{q} < 1$$

(lim sup existuje a je menší než 1).

Důkaz Z definice 8 plyne

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

$\forall i \geq k$, kde $q < 1$, takže

$$\| x_i - x^* \|^{\frac{1}{i}} \leq (M_k \| x_k - x^* \|)^{\frac{1}{i}} q^{1-\frac{k}{i}} \leq \max(1, M_k \| x_k - x^* \|)$$

Tato posloupnost je omezená, takže existuje lim sup a platí

$$\limsup_{i \to \infty} \| x_i - x^* \|^{\frac{1}{i}} = \hat{q} \leq \lim_{i \to \infty} (M_k \| x_k - x^* \|)^{\frac{1}{i}} q^{1-\frac{k}{i}} = q < 1$$

Z druhé strany nechť existuje výraz na levé straně poslední nerovnosti a je roven $\hat{q} < 1$. Pak pro libovolné číslo $\hat{q} < q < 1$ existuje index $k \in N$ tak, že platí

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

neboli

$$\| x_i - x^* \| \leq q^i$$

$\forall i \geq k$. Zvolme

$$M_k = \frac{q^k}{\| x_k - x^* \|}$$

Pak platí

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

Definice 9 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* R-superlineárně, jestliže

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

Definice 10 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* Q-lineárně, jestliže existuje index $k \in N$ a konstanta $0 < q < 1$ tak, že

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

Definice 11 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* Q-superlineárně, jestliže

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

Věta 6 Nechť $x_i \to x^*$ Q-lineárně (Q-superlineárně). Pak $x_i \to x^*$ R-lineárně (R-superlineárně)

Důkaz R-lineární konvergence plyne v Q-lineární konvergence bezprostředně (stačí volit $M_k = 1$). Nechť $0 < q < 1$ je libovolné (malé) číslo. Z Q-superlineární konvergence plyne existence indexu $k \in N$ takového, že

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

takže podle věty 5 platí

$$\limsup_{i \to \infty} \| x_i - x^* \|^{\frac{1}{i}} \leq q$$

Protože q je libovolné (malé) musí platit

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

Poznámka 3 Q-lineární (Q-superlineární) konvergence implikuje monotonnost posloupnosti $\| x_i - x^* \|$, $i \in N$ (počínaje vhodným indexem $k \in N$).

Definice 12 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* Q-kvadraticky, jestliže existuje index $k \in N$ a konstanta $0 < M_k < \infty$ tak, že

$$\| x_{i+1} - x^* \| \leq M_k \| x_i - x^* \|^2 \qquad \forall i \geq k$$

Definice 13 Posloupnost $x_i \in R^n$, $i \in N$, konverguje k bodu x^* Q-m-krokově kvadraticky, jestliže existuje index $k \in N$, číslo $m \in N$ a konstanta $0 < M_k < \infty$ tak, že

$$\| x_{i+m} - x^* \| \leq M_k \| x_i - x^* \|^2 \qquad \forall i \geq k$$

Poznámka 4 Někdy se používá slabší předpoklad $\| x_{jl+m} - x^* \| \leq M \| x_{jl} - x^* \|^2$ $\forall j \in N$ (vybraná posloupnost $i = jl$).

1.4. Základní optimalizační metody

Základní optimalizační metoda je posloupnost $x_i \in R^n$, $i \in N$, 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 α_i se určuje na základě chování funkce $F : R^n \to R$ v okolí bodu $x_i \in R^n$.

Definice 14 Řekneme, že základní optimalizační metoda je globálně konvergentní, jestliže pro libovolný počáteční vektor $x_1 \in R^n$ platí

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

Metoda největšího spádu:

Metoda největšího spádu je definována vztahy

$$s_i = -g(x_i)$$

$$\alpha_i = \arg \min_{\alpha \geq 0} F(x_i + \alpha_i s_i)$$

Výhody:

- je globálně konvergentní
- používá pouze vektory z R^n

  $O(n)$ - paměťových míst

  $O(n)$ - operací na iteraci

Nevýhody

- přesný výběr délky kroku
- je pouze R-lineárně konvergentní s asymptotickou rychlostí

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

Odhad asymptotické rychlosti je obvykle realistický (není nadhodnocený). Například jestliže $\kappa(G(x^*)) = 10^3$ potřebujeme ke snížení chyby $\| x - x^* \|$ o 4 řády zhruba 4600 iterací a jestilže $\kappa(G(x^*)) = 10^6$ potřebujeme ke snížení chyby $\| x - x^* \|$ o 8 řádů zhruba 9200000 iterací!

Newtonova metoda:

Newtonova metoda je definována vztahy

$$s_i = -G^{-1}(x_i)g(x_i)$$

$$\alpha_i = 1$$

Výhody:

- je Q- kvadraticky konvergentní. Pokud konverguje, stačí k nalezení lokálního minima pouze několik iterací
- jednoduchý výběr délky kroku

Nevýhody

- není globálně konvergentní. Pokud x_1 je daleko od x^*, nemusí konvergovat.
- používá matici a je třeba řešit soustavu lineárních rovnic.

  $O(n^2)$ - paměťových míst

  $O(n^3)$ - operací na iteraci
- je třeba počítat druhé derivace

Důmyslnější metody:

Metody spádových směrů vyvinuté z metody největšího spádu

- nepřesný výběr délky kroku

- urychlení konvergence (metody sdružených gradientů, metody s proměnnou metrikou).

Metody s lokálně omezeným krokem vyvinuté z Newtonovy metody

- zajištění globální konvergence
- snížení počtu operací (nepřesná Newtonova metoda)

2. Metody spádových směrů

2.1. Základní vlastnosti metod spádových směrů

Definice 15 Řekneme, že směrové vektory $s_i \in R^n$, $i \in N$, jsou spádové jestliže platí

$$s_i^T g_i < 0 \tag{$\overline{\text{S1}}$}$$

$\forall i \in N$. Řekneme, že směrové vektory $s_i \in R^n$, $i \in N$ jsou stejnoměrně spádové, jestliže existuje konstanta $0 < \varepsilon_0 \leq 1$ taková, že platí

$$s_i^T g_i \leq -\varepsilon_0 \parallel s_i \parallel \parallel g_i \parallel \tag{S1}$$

$\forall i \in N$.

Definice 16 Řekneme, že délka kroku $\alpha_i > 0$, $i \in N$, splňuje buď silnou Wolfeho podmínku nebo slabou Wolfeho podmínku nebo Goldsteinovu podmínku nebo Armijovu podmínku, jestliže existují čísla $0 < \varepsilon_1 < 1/2$ a $\varepsilon_1 < \varepsilon_2 < 1$ tak, že

$$F_{i+1} - F_i \leq \varepsilon_1 \alpha_i s_i^T g_i \tag{S2}$$

a buď

$$|s_i^T g_{i+1}| \leq \varepsilon_2 |s_i^T g_i| \tag{S3a}$$

nebo

$$s_i^T g_{i+1} \geq \varepsilon_2 s_i^T g_i \tag{S3b}$$

nebo

$$F_{i+1} - F_i \geq \varepsilon_2 \alpha_i s_i^T g_i \tag{S3c}$$

nebo

$$\underline{c} \parallel g_i \parallel \leq \alpha_i \parallel s_i \parallel \tag{S3d}$$

kde $0 < \underline{c} < \infty$ je nějaká konstanta.

Definice 17 Řekneme, že základní optimalizační metoda $x_{i+1} = x_i + \alpha_i s_i$ $i \in N$ je metodou spádových směrů jestliže směrové vektory $s_i \in R^n$, $i \in N$, splňují podmínku (S1) (jsou stejnoměrně spádové) a délky kroku $\alpha_i > 0$, $i \in N$ splňují podmínku (S2) a některou z podmínek (S3a)-(S3d).

Poznámka 5 Pro metody sdružených gradientů odvozené z metody největšího spádu se používá podmínka (S3a). Pro metody s proměnnou metrikou odvozené z Newtonovy metody (kde $\alpha_i \to 1$ pro $i \to \infty$) se používá podmínka (S3b). Pro bezderivační metody se používá podmínka (S3c). Pro nehladké úlohy se používá podmínka (S3d).

Poznámka 6 Podmínka (S1) se často nahražuje nějakou slabší podmínkou (směrové vektory nemusí být stejnoměrně spádové). Důkaz globální konvergence je pak obtížnější.

Poznámka 7 Metoda největšího spádu je metodou spádových směrů, neboť $s_i = -g_i$, takže $s_i^T g_i = - \parallel g_i \parallel^2 = - \parallel s_i \parallel \parallel g_i \parallel$ a (S1) platí pro $\varepsilon_0 = 1$.

Lemma 7 (Konzistence) Nechť funkce $F : R^n \to R$ splňuje podmínky (F1) a (F3) a směrový vektor s_i splňuje podmínku $(\overline{S2})$. Pak pro libovolnou z podmínek (S3a)-(S3d) existuje délka kroku $\alpha_i > 0$, která vyhovuje této podmínce i podmínce (S2).

Důkaz (pro (S3b)). Označme

$$\mathcal{M}_i = \{\alpha \geq 0 : F(x_i + \alpha s_i) - F_i \leq \varepsilon_1 \alpha s_i^T g_i\}$$

zřejmě $\mathcal{M}_i \neq \emptyset$ neboť $0 \in \mathcal{M}_i$. Nechť $\alpha_i = \sup \mathcal{M}_i$. Podle (F1) platí $F(x_i + \alpha s_i) \geq \underline{F}$ takže pro $\alpha_i \in \mathcal{M}_i$ dostaneme $\alpha_i \leq (\underline{F} - F_i)/\varepsilon_1 s_i^T g_i < \infty$. Ukážeme nyní, že platí $s_i^T g(x_i + \alpha_i s_i) \geq \varepsilon_1 s_i^T g_i$. Předpokládejme naopak, že

$$s_i^T g_i(x_i + \alpha_i s_i) = \varepsilon s_i^T g_i$$

kde $\varepsilon > \varepsilon_1$. Potom podle (F3) platí

$$\begin{aligned} F(x_i + \alpha s_i) - F_i &\leq F(x_i + \alpha_i s_i) - F_i + s_i^T g(x_i + \alpha_i s_i)(\alpha - \alpha_i) + \frac{1}{2}\overline{G}(\alpha - \alpha_i)^2 \leq \\ &\leq \varepsilon_1 \alpha_i s_i^T g_i + \varepsilon(\alpha - \alpha_i) s_i^T g_i + \frac{1}{2}\overline{G}(\alpha - \alpha_i)^2 = \\ &= \alpha \varepsilon_1 s_i^T g_i - (\varepsilon_1 - \varepsilon)(\alpha - \alpha_i) s_i^T g_i + \frac{1}{2}\overline{G}(\alpha - \alpha_i)^2 \end{aligned}$$

a pro $\alpha = \alpha_i + (\varepsilon_1 - \varepsilon) s_i^T g_i / \overline{G} > \alpha_i$ dostaneme

$$F(x_i + \alpha s_i) - F_i \leq \alpha \varepsilon_1 s_i^T g_i - \frac{1}{2} \frac{(\varepsilon - \varepsilon_1)^2 (s_i^T g_i)^2}{\overline{G}} < \alpha \varepsilon_1 s_i^T g_i$$

což je spor neboť $\alpha_i = \sup \mathcal{M}_i$. Z $s_i^T g(x_i + \alpha_i s_i) \geq \varepsilon_1 s_i^T g_i$ a $s_i^T g_i < 0$ plyne $\alpha_i > 0$..

Věta 8 (Globální konvergence) Nechť funkce $F : R^n \to R$ splňuje podmínky (F1) a (F3). Pak metoda spádových směrů je globálně konvergentní.

Důkaz (pro (S3b)) Dokážeme nejprve, že existuje konstanta $\underline{c} > 0$ tak, že pro libovolný index $i \in N$ platí

$$\alpha_i \geq \underline{c} \parallel g_i \parallel / \parallel s_i \parallel \tag{a}$$

a

$$F_{i+1} - F_i \leq -\varepsilon_0 \varepsilon_1 \underline{c} \parallel g_i \parallel \tag{b}$$

Z podmínky (S3b) dostaneme

$$\varepsilon_2 s_i^T g_i \leq s_i^T g(x_i + \alpha_i s_i) \leq s_i^T g_i + \alpha_i \overline{G} \parallel s_i \parallel^2$$

neboli

$$\alpha_i \geq \frac{(\varepsilon_2 - 1) s_i^T g_i}{\overline{G} \parallel s_i \parallel^2} \geq \frac{\varepsilon_0 (1 - \varepsilon_2) \parallel g_i \parallel}{\overline{G} \parallel s_i \parallel}$$

takže platí (a) pro $\underline{c} = \varepsilon_0 (1 - \varepsilon_2)/\overline{G} > 0$. Dále z podmínky (S2) a z (a) dostaneme

$$F_{i+1} - F_i \leq \varepsilon_1 \alpha_i s_i^T g_i \leq -\varepsilon_0 \varepsilon_1 \alpha_i \parallel s_i \parallel \parallel g_i \parallel \leq -\varepsilon_0 \varepsilon_1 \underline{c} \parallel g_i \parallel^2$$

takže platí (b). Nyní můžeme psát

$$F_{i+1} = F_1 + \sum_{j=1}^{i} (F_{j+1} - F_j) \leq F_1 - \varepsilon_0 \varepsilon_1 \underline{c} \sum_{j=1}^{i} \parallel g_j \parallel^2$$

Podle (b) je posloupnost F_i, $i \in N$ klesající a podle (F1) je zdola omezená. Existuje tedy limita

$$\underline{F} \leq \lim_{i \to \infty} F_i \leq F_1 - \varepsilon_0 \varepsilon_1 \underline{c} \sum_{i=1}^{\infty} \parallel g_i \parallel^2$$

takže

$$\sum_{i=1}^{\infty} \parallel g_i \parallel^2 \leq \frac{F_1 - \underline{F}}{\varepsilon_0 \varepsilon_1 \underline{c}}$$

takž nutně $\parallel g_i \parallel \to 0$.

Důsledek Metoda největšího spádu je globálně konvergentní.

Poznámka 7 Tvrzení $\parallel g_i \parallel \to 0$ je silnější než podmínka v definici 14. K tomu aby metoda spádových směrů byla globálně konvergentní ve smyslu definice 14 stačí, když (S1) platí pro nějakou nekonečnou podmnožinu množiny N.

Poznámka 8 Je-li základní optimalizační metoda globálně konvergentní (ve smyslu definice 14), nemusí ještě platit $x_i \to x^*$. Splňuje-li funkce $F : R^n \to R$ podmínku

(F2) nemůže posloupnost $x_i \in R^n$, $i \in N$ divergovat, může však mít více hromadných bodů. Vyhovuje-li nějaký hromadný bod $x^* \in R^n$ posloupnosti $x_i \in R^n$, $i \in N$, nutným podmínkám pro lokální minimum (Tvrzení 2), pak již platí $x_i \to x^*$. V dalším výkladu budu předpokládat, že globální konvergence automaticky znamená $x \to x^*$.

Věta 9 (Lineární konvergence) Nechť $x_i \in R^n$, $i \in N$ je posloupnost generovaná metodou spádových směrů taková, že $x_i \to x^*$. Nechť funkce $F : R^n \to R$ vyhovuje podmínkám (F3) a (F4). Pak pro libovolný index $k \in N$ platí

$$\frac{\| x_i - x^* \|}{\| x_k - x^* \|} \leq \sqrt{\frac{\overline{G}}{\underline{G}}}\; q^{i-k}$$

$\forall i \geq k$, kde $q = \sqrt{1 - 2\varepsilon_0\varepsilon_1\underline{c}\underline{G}^2/\overline{G}}$.

Důkaz Podle (b) platí

$$F_{i+1} - F^* \leq F_i - F^* - \varepsilon_0\varepsilon_1\underline{c} \| g_i \|^2$$

z vět o střední hodnotě dostaneme

$$F_i - F^* \leq \frac{1}{2}\overline{G} \| x_i - x^* \|^2 \tag{c}$$

$$F_i - F^* \geq \frac{1}{2}\underline{G} \| x_i - x^* \|^2 \tag{d}$$

$$\| g_i \| \leq \overline{G} \| x_i - x^* \| \tag{e}$$

$$\| g_i \| \geq \underline{G} \| x_i - x^* \| \tag{f}$$

Můžeme tedy psát

$$\frac{F_{i+1} - F^*}{F_i - F^*} \leq 1 - \varepsilon_0\varepsilon_1\underline{c}\frac{\| g_i \|^2}{F_i - F^*} \leq 1 - \varepsilon_0\varepsilon_1\underline{c}\frac{2\underline{G}^2 \| x_i - x^* \|^2}{\overline{G} \| x_i - x^* \|^2} = 1 - 2\varepsilon_0\varepsilon_1\underline{c}\underline{G}^2/\overline{G} = q^2 \tag{g}$$

Pro podmínku (S3b) platí $\underline{c} = \varepsilon_0(1 - \varepsilon_2)/\overline{G}$ (důkaz věty 8), takže

$$q^2 = 1 - 2\varepsilon_0^2\varepsilon_1(1 - \varepsilon_2)\underline{G}^2/\overline{G}^2 > 1 - 2\varepsilon_0^2\varepsilon_1(1 - \varepsilon_1)\underline{G}^2/\overline{G}^2 > 1 - \varepsilon_0^2\underline{G}^2/\overline{G}^2 \geq 0$$

neboť $0 < \varepsilon_0 \leq 1$, $0 < \varepsilon_1 < 1/2$, $\varepsilon_1 < \varepsilon_2 < 1$. Několikanásobným použitím nerovnosti (g) dostaneme

$$\frac{F_i - F^*}{F_k - F^*} \leq q^{2(i-k)}$$

což s použitím (c) a (d) dává

$$\frac{\| x_i - x^* \|}{\| x_k - x^* \|} \leq \sqrt{\frac{\overline{G}}{\underline{G}}} \sqrt{\frac{F_i - F^*}{F_k - F^*}} \leq \sqrt{\frac{\overline{G}}{\underline{G}}} q^{i-k}$$

Poznámka 9 Často se (F4) předpokládá až od nějakého indexu k (například, když $x_i \to x^*$ a x^* splňuje postačující podmínky pro lokální minimum, pak předpokládáme (F4) až v okolí bodu x^*).

Poznámka 10 Podle věty 9 je $\| e_{i+1} \| = O(\| e_i \|)$

Definice 18 Řekneme, že výběr délky kroku je asymptoticky přesný, jestliže

$$\alpha_i = -\frac{s_i^T g_i}{s_i^T G^* s_i}(1 + O(\| e_i \|))$$

Lemma 10 Nechť $x_i \in R^n$, $i \in N$, je posloupnost generovaná metodou spadových směrů s asymptoticky přesným výběrem délky kroku, taková, že $x_i \to x^*$. Nechť funkce $F : R^n \to R$ vyhovuje podmínkám (F3)-(F5). Pak platí

$$F_{i+1} - F_i = -\frac{1}{2}\frac{(s_i^T g_i)^2}{s_i^T G^* s_i}(1 + O(\| e_i \|))$$

Důkaz Budeme psát $u_i \sim v_i$ jestliže $u_i = O(v_i)$ a $v_i = O(u_i)$. Podle věty 4 platí

$$g_i = G^* e_i + O(\| e_i \|^2)$$

což s použitím (F3) a (F4) dává $g_i \sim e_i$. Jelikož z (F3) a (F4) plyne $s_i^T G^* s_i \sim \| s_i \|^2$, z (S1) plyne $s_i^T g_i \sim \| s_i \| \| g_i \|$ a předpokládáme, že $\| e_i \| \to 0$, můžeme psát

$$\alpha_i = -\frac{s_i^T g_i}{s_i^T G^* s_i}(1 + O(\| e_i \|)) \sim \frac{\| g_i \|}{\| s_i \|}$$

takže $\| d_i \| = \| \alpha_i s_i \| \sim \| g_i \| \sim \| e_i \|$ a podle věty 3 platí

$$\begin{aligned} F_{i+1} - F_i &= \alpha_i s_i^T g_i + \frac{1}{2}\alpha_i^2 s_i^T G_i s_i + O(\| d_i \|^3) = \\ &= \alpha_i s_i^T g_i + \frac{1}{2}\alpha_i^2 s_i^T G^* s_i + \frac{1}{2} d_i^T (G_i - G^*) d_i + O(\| d_i \|^3) = \\ &= \alpha_i s_i^T g_i + \frac{1}{2}\alpha_i^2 s_i^T G^* s_i + O(\| e_i \|^3) \end{aligned}$$

Dosadíme-li do tohoto vyjádření vztah pro asympticky přesný výběr délky kroku, dostaneme

$$F_{i+1} - F_i = -\frac{1}{2}\frac{(s_i^T g_i)^2}{s_i^T G^* s_i}(1 + O(\| e_i \|))$$

neboť $(s_i^T g_i)^2 / s_i^T G^* s_i \sim \| g_i \|^2 \sim \| e_i \|^2$.

Tvrzení 11 Nechť B je symetrická pozitivně definitní (SPD) matice a vektory $u \in R^n$, $v \in R^n$ vyhovují podmínce

$$\frac{(u^T v)^2}{u^T u v^T v} \geq 1 - \varepsilon^2$$

kde $0 \leq \varepsilon \leq 1$. Pak platí

$$\frac{(u^T v)^2}{u^T B u v^T B^{-1} v} \geq \frac{4\kappa(B)(1-\varepsilon^2)}{(\kappa(B) + 1 + (\kappa(B) - 1)\varepsilon)^2}$$

Věta 12 Nechť jsou splněny podmínky lemmatu 10. Pak platí

$$\limsup_{i\to\infty} \| x_i - x^* \|^{\frac{1}{i}} \leq \frac{\kappa(G^*) - 1 + (\kappa(G^*) + 1)\sqrt{1-\varepsilon_0^2}}{\kappa(G^*) + 1 + (\kappa(G^*) - 1)\sqrt{1-\varepsilon_0^2}}$$

Důkaz Podle věty 4 platí

$$F_i - F^* = \frac{1}{2} e_i^T G^* e_i + O(\| e_i \|^3)$$

$$g_i = G^* e_i + O(\| e_i \|^2)$$

takže s použitím (F3) a (F4) a z toho, že $\| g_i \| \sim \| e_i \|$ dostaneme

$$e_i = (G^*)^{-1} g_i (1 + O(\| e_i \|))$$

a

$$F_i - F^* = \frac{1}{2} g_i^T (G^*)^{-1} g_i (1 + O(\| e_i \|))$$

Použijeme-li lemma 10 můžeme psát

$$\frac{F_{i+1} - F^*}{F_i - F^*} = 1 + \frac{F_{i+1} - F_i}{F_i - F^*} = 1 - \frac{(s_i^T g_i)^2}{s_i^T G^* s_i g_i^T (G^*)^{-1} g_i}(1 + O(\| e_i \|))$$

(platí $(1 + O(\| e_i \|))/(1 + O(\| e_i \|)) = 1 + O(\| e_i \|)$). Podle (S1) platí $(s_i^T g_i)^2 \geq \varepsilon_0^2 \| s_i \|^2 \| g_i \|^2$ takže s použitím tvrzení 11 dostaneme

$$\frac{(s_i^T g_i)^2}{s_i^T G^* s_i g_i^T (G^*)^{-1} g_i} \geq \frac{4\kappa(G^*)\varepsilon_0^2}{(\kappa(G^*) + 1 + (\kappa(G^*) - 1)\sqrt{1-\varepsilon_0^2})^2}$$

takže

$$\begin{aligned}\frac{F_{i+1}-F^*}{F_i-F^*} &\leq \left(\frac{(\kappa(G^*)-1+(\kappa(G^*)+1)\sqrt{1-\varepsilon_0^2}}{(\kappa(G^*)+1+(\kappa(G^*)-1)\sqrt{1-\varepsilon_0^2}}\right)^2 (1+O(\| e_i \|) = \\ &= \hat{q}^2(1+O(\| e_i \|))\end{aligned}$$

K libovolnému číslu $q, \hat{q} < q < 1$, tedy existuje index $k \in N$ tak, že

$$\frac{F_{i+1}-F^*}{F_i-F^*} \leq q^2$$

$\forall i \geq k$. Můžeme tedy postupovat stejně jako v důkazu věty 9, takže

$$\frac{F_i-F^*}{F_k-F^*} \leq q^{2(i-k)}$$

a

$$\frac{\| x_i - x^* \|}{\| x_k - x^* \|} \leq \sqrt{\frac{\overline{G}}{\underline{G}}} q^{i-k}$$

a podle věty 5 platí

$$\limsup_{i\to\infty} \| x_i - x^* \|^{\frac{1}{i}} \leq q$$

Jelikož to platí pro libovolné číslo q, $\hat{q} < q < 1$, dokázali jsme tvrzení věty.

Poznámka 11 Pro metodu největšího spádu je $\varepsilon_0 = 1$, takže

$$\limsup_{i\to\infty} \leq \frac{\kappa(G^*)-1}{\kappa(G^*)+1}$$

Poznámka 12 Asymptoticky přesný výběr délky kroku dostaneme, vybíráme-li délku kroku pomocí kvadratické nebo kubické interpolace (viz poznámku 15).

Věta 13 (Superlineární konvergence). Nechť $x_i \in R^n$, $i \in N$, je posloupnost získaná metodou spádových směrů taková, že $x_i \to x^*$. Nechť funkce $F : R^n \to R$ vyhovuje podmínkám (F3) a (F4). Nechť $\alpha_i = 1$, kdykoliv tato hodnota vyhovuje podmínkám (S2) a (S3). Nechť směrový vektor se určuje tak, že

$$\lim_{i\to\infty} \frac{\| B_i s_i + g_i \|}{\| g_i \|} = 0 \tag{A}$$

kde

$$\lim_{i\to\infty} \frac{\| (G^* - B_i) s_i \|}{\| s_i \|} = 0 \tag{B}$$

Pak existuje index $k \in N$ takový, že $\alpha_i = 1\ \forall i \geq k$ a posloupnost $x_i\ i \in N$, konverguje superlineárně k bodu $x^* \in R^n$. Navíc $-s_i^T g_i \geq \varepsilon_0 \parallel s_i \parallel\parallel g_i \parallel \forall i \geq k$, pokud $\varepsilon_0 > 1/\kappa(G^*)$.

Důkaz (a) ukážeme, že existuje index $k_2 \in N$ tak, že

$$\parallel g_i \parallel / \overline{G} \leq \parallel s_i \parallel \leq \parallel g_i \parallel / \underline{G}$$

pokud $\underline{G} < \underline{\lambda}(G^*)$ a $\overline{G} > \overline{\lambda}(G^*)$. Označme $u_i = (B_i s_i + g_i)/ \parallel g_i \parallel$ a $v_i = (G^* - B_i)s_i / \parallel s_i \parallel$ pak platí

$$s_i^T B_i s_i = s_i^T u_i \parallel g_i \parallel - s_i^T g_i \leq (\parallel u_i \parallel + 1) \parallel s_i \parallel\parallel g_i \parallel$$

a

$$s_i^T B_i s_i = s_i^T G^* s_i - s_i^T v_i \parallel s_i \parallel \geq (\underline{\lambda}(G^*) - \parallel v_i \parallel) \parallel s_i \parallel^2$$

což dohromady dává

$$\parallel s_i \parallel \leq \frac{1 + \parallel u_i \parallel}{\underline{\lambda}(G^*) - \parallel v_i \parallel} \parallel g_i \parallel$$

a jelikož $\parallel u_i \parallel \to 0$ a $\parallel v_i \parallel \to 0$ podle (A) a (B), existuje index $k_1 \in N$ tak, že $\parallel s_i \parallel \leq \parallel g_i \parallel / \underline{G}\ \forall i \geq k$, pokud $\underline{G} < \lambda(G^*)$. Jelikož

$$\frac{\parallel G^* s_i + g_i \parallel}{\parallel g_i \parallel} \leq \frac{\parallel (G^* - B_i)s_i \parallel}{\underline{G} \parallel s_i \parallel} + \frac{\parallel B_i s_i + g_i \parallel}{\parallel g_i \parallel} \to 0$$

podle (A) a (B), existuje index $k_2 \geq k_1$ tak, že

$$| \parallel G^* s_i \parallel - \parallel g_i \parallel | \leq \parallel G^* s_i + g_i \parallel \leq (1 - \overline{\lambda}(G^*)/\overline{G}) \parallel g_i \parallel$$

kde $\overline{G} > \overline{\lambda}(G^*)$, takže buď $\parallel G^* s_i \parallel \geq \parallel g_i \parallel$ nebo $\parallel g_i \parallel - \parallel G^* s_i \parallel \leq (1 - \overline{\lambda}(G^*)/\overline{G})$ $\parallel g_i \parallel$, což obojí dává $\parallel s_i \parallel \geq \parallel g_i \parallel / \overline{G}\ \forall i \geq k_2$.
(b) Ukážeme, že existuje index $k_3 \geq k_2$ takový, že $-s_i^T g_i \geq \varepsilon_0 \parallel s_i \parallel\parallel g_i \parallel \forall i \geq k_3$, pokud $\varepsilon_0 > 1/\kappa(G^*)$. Zvolme $\underline{G} < \underline{\lambda}(G^*)$ a $\overline{G} > \overline{\lambda}(G^*)$ tak, aby platilo $\varepsilon_0 = \underline{G}/\overline{G}$. Z definice u_i a v_i a z (a) plyne, že

$$\begin{aligned} -s_i^T g_i &= s_i^T B_i s_i - s_i^T u_i \parallel g_i \parallel = s_i^T G^* s_i - s_i^T v_i \parallel s_i \parallel - s_i^T u_i \parallel g_i \parallel \geq \\ &\geq \frac{1}{\overline{G}} \parallel g_i \parallel\parallel s_i \parallel (\underline{\lambda}(G^*) - \parallel v_i \parallel - \overline{G} \parallel u_i \parallel) \end{aligned}$$

a jelikož $\parallel u_i \parallel \to 0$ a $\parallel v_i \parallel \to 0$ podle (A) a (B), existuje index $k_3 \geq k_2$ tak, že $-s_i^T g_i \geq (\underline{G}/\overline{G}) \parallel s_i \parallel\parallel g_i \parallel = \varepsilon_0 \parallel s_i \parallel\parallel g_i \parallel \forall i \geq k_3$.
(c) Ukážeme, že existuje index $k \geq k_3$ tak, že hodnota $\alpha_i = 1$ vyhovuje podmínkám (S2) a (S3b). Označme

$$\eta_i = \frac{s_i^T g_i + s_i^T G^* s_i}{s_i^T g_i}$$

Pak podle předchozích výsledků dostaneme pro $i \geq k_3$

$$|\eta_i| = \frac{|s_i^T g_i + s_i^T G^* s_i|}{|s_i^T g_i|} \leq \frac{\| s_i \| \| g_i + G^* s_i \|}{\varepsilon_0 \| s_i \| \| g_i \|} \leq \frac{\| g_i + B_i s_i \|}{(\underline{G}/\overline{G}) \| g_i \|} + \frac{\| (G^* - B_i) s_i \|}{\underline{G} \| s_i \|}$$

a podle (A) a (B) platí $|\eta_i| \to 0$. Nyní použijeme věty o střední hodnotě.

$$F(x_i + s_i) - F(x_i) = s_i^T g_i + \frac{1}{2} s_i^T G^* s_i + o(\| s_i \|) \| s_i \|^2$$

$$s_i^T g(x_i + s_i) = s_i^T g_i + s_i^T G^* s_i + o(\| s_i \|) \| s_i \|^2$$

neboť $\| s_i \| \sim \| g_i \| \sim \| e_i \|$. Můžeme tedy psát

$$\lim_{i \to \infty} \frac{F(x_i + s_i) - F(x_i)}{s_i^T g_i} = \frac{1}{2} + \lim_{i \to \infty} (\frac{1}{2} \eta_i + o(\| s_i \|)) = \frac{1}{2}$$

$$\lim_{i \to \infty} \frac{s_i^T g(x_i + s_i)}{s_i^T g_i} = \lim_{i \to \infty} (\eta_i + o(\| s_i \|)) = 0$$

a protože $0 < \varepsilon_1 < 1/2$ a $\varepsilon_1 < \varepsilon_2 < 1$ existuje index $k \geq k_3$ tak, že (S2) a (S3b) platí $\forall i \geq k$.
(d) Superlineární konvergence. Použijeme větu o střední hodnotě

$$g(x_i + s_i) = g_i + G^* s_i + o(\| s_i \|) \| s_i \|$$

Podle předchozích výsledků dostaneme $x_{i+1} = x_i + s_i \; \forall i \geq k$ a $\| s_i \| \sim \| g_i \| \to 0$
Můžeme tedy psát

$$\begin{aligned}
\frac{\| x_{i+1} - x^* \|}{\| x_i - x^* \|} &\leq \frac{\overline{G}}{\underline{G}} \frac{\| g_{i+1} \|}{\| g_i \|} \leq \\
&\leq \frac{\overline{G}}{\underline{G}} \left(\frac{\| g(x_i + s_i) - g_i - B_i s_i \|}{\| g_i \|} + \frac{\| B_i s_i + g_i \|}{\| g_i \|} \right) \leq \\
&\leq \frac{\overline{G}}{\underline{G}} \left(\frac{\| (G^* - B_i) s_i \|}{\underline{G} \| s_i \|} + \frac{\| B_i s_i + g_i \|}{\| g_i \|} + o(\| s_i \|) \right)
\end{aligned}$$

takže podle (A) a (B) platí

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

Výběr délky kroku:

Algoritmus 1 (S3b) Data $0 < \beta_1 < \beta_2 < 1 < \gamma_1 < \gamma_2$.

Krok 1 Zvolíme počáteční délku kroku $\alpha > 0$. Položíme $\overline{\alpha} = 0$.

Krok 2 Položíme $\underline{\alpha} = \overline{\alpha}$ a $\overline{\alpha} = \alpha$. Jsou-li splněny podmínky (S2) a (S3b), ukončíme výpočet s $\alpha_i = \alpha$. Není-li splněna podmínka (S2), přejdeme na krok 4.

Krok 3 Určíme hodnotu α pomocí extrapolace tak, aby $\gamma_1\overline{\alpha} \le \alpha \le \gamma_2\overline{\alpha}$ a přejdeme na krok 2.

Krok 4 Určíme hodnotu α pomocí interpolace tak, aby $\beta_1(\overline{\alpha} - \underline{\alpha}) \le (\alpha - \underline{\alpha}) \le \beta_2(\overline{\alpha} - \underline{\alpha})$.

Krok 5 Jsou-li splněny podmínky (S2) a (S3b), ukončíme výpočet s $\alpha_i = \alpha$. Není-li splněna podmínka (S2) položíme $\overline{\alpha} = \alpha$ a přejdeme na krok 4. V opačném případě položíme $\underline{\alpha} = \alpha$ a přejdeme na krok 4.

Poznámka 13 Jsou-li splněny podmínky (F1) a (F3) najde algoritmus délku kroku vyhovující podmínkám (S2) a (S3b) po konečném počtu kroků. Je-li splněna podmínka (F4) nezávisí tento počet na indexu $i \in N$.

Poznámka 14 Extrapolace a interpolace. Označme $\varphi(\alpha) = F(x_i + \alpha s_i)$, $\varphi(\alpha) = s_i^T g(x_i + \alpha s_i)$ a

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

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

Kvadratická interpolace (dvě hodnoty):

$$\alpha - \underline{\alpha} = \frac{\overline{\alpha} - \underline{\alpha}}{2(1 - A)} \tag{Ka}$$

Kvadratická interpolace (dvě derivace):

$$\alpha - \underline{\alpha} = \frac{\overline{\alpha} - \underline{\alpha}}{1 - B} \tag{Kb}$$

Kubická interpolace:

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

kde

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

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

Poznámka 15 Určuje-li se délka kroku pomocí (Ka) nebo (Kb) nebo (C) a platí-li (F3) a (F4), je výběr délky kroku asymptoticky přesný.

Poznámka 16 Počáteční výběr délky kroku. Pro superlineárně konvergentní metody volíme $\alpha = 1$. U metod sdružených gradientů volíme $\alpha = 2(F_{i-1} - F_i)/s_i^T g_i$ nebo (v prvním iteračním kroku) $\alpha = 2(\underline{F} - F_i)/s_i^T g_i$.

Shrnutí: Metody spádových směrů $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$.

a) (F1) - (F3) $\rightarrow$ globální konvergence
b) (F3) - (F4) $\rightarrow$ lineární konvergence
c) (F3) - (F5) $\rightarrow$ asymptotický odhad

$$\limsup_{i\to\infty} \| x_i - x^* \|^{\frac{1}{i}} \leq \frac{\kappa(G^*) - 1 + (\kappa(G^*) + 1)\sqrt{1-\varepsilon_0^2}}{\kappa(G^*) + 1 + (\kappa(G^*) - 1)\sqrt{1-\varepsilon_0^2}}$$

d) (F3) - (F4) $\rightarrow$ superlineární konvergence pokud

$$\frac{\| B_i s_i + g_i \|}{\| g_i \|} \rightarrow 0$$

$$\frac{\| (G^* - B_i)s_i \|}{\| s_i \|} \rightarrow 0$$

2.2. Metody sdružených gradientů

Definice 19 Řekneme, že základní optimalizační metoda je metodou sdružených gradientů jestliže $s_1 = -g_1$ a

$$s_{i+1} = -g_{i+1} + \frac{y_i^T g_{i+1}}{y_i^T s_i} s_i \tag{CGa}$$

(Hestenes, Stiefel) nebo

$$s_{i+1} = -g_{i+1} + \frac{y_i^T g_{i+1}}{g_i^T g_i} s_i \tag{CGb}$$

(Polak, Ribiere) nebo

$$s_{i+1} = -g_{i+1} + \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} s_i \tag{CGc}$$

(Fletcher, Reeves) pro $i \in N$. Přitom $y_i = g_{i+1} - g_i$.

Poznámka 17 Metoda (CGa) je teoreticky nejpodloženější. Metoda (CGb) dává nejlepší praktické výsledky. Meotda (CGc) je nejjednodušší a je globálně konvergentní bez přerušování iteračního procesu.

Věta 14 (Globální konvergence). Nechť funkce $F : R^n \rightarrow R$ splňuje podmínky (F1)-(F3). Pak metoda sdružených gradientů Fletchera a Reevese (CGc) s výběrem délky kroku splňujícím silnou Wolfeho podmínku (S2) a (S3a), kde $\varepsilon_1 < \varepsilon_2 < 1/2$, je globálně konvergentní.

Důkaz (a) (Al-Baali) Dokážeme indukcí nerovnost

$$-1 - \frac{\varepsilon_2}{1-\varepsilon_2} \le \frac{s_i^T g_i}{\| g_i \|^2} \le -1 + \frac{\varepsilon_2}{1-\varepsilon_2}$$

$\forall i \in N$. Pro $i = 1$ to platí, neboť $s_1 = -g_1$ a tedy $s_1^T g_1 / \| g_1 \|^2 = -1$. Použijeme-li (CGc), dostaneme

$$\frac{s_{i+1}^T g_{i+1}}{\| g_{i+1} \|^2} = -1 + \frac{\| g_{i+1} \|^2}{\| g_i \|^2} \frac{s_i^T g_{i+1}}{\| g_{i+1} \|^2} = -1 + \frac{s_i^T g_{i+1}}{\| g_i \|^2}$$

Podle (S3a) platí

$$|s_i^T g_{i+1}| \le -\varepsilon_2 s_i^T g_i$$

takže

$$\begin{aligned}
-1 + \varepsilon_2 \frac{s_i^T g_i}{\| g_i \|^2} &\le \frac{s_{i+1}^T g_{i+1}}{\| g_{i+1} \|^2} \le -1 - \varepsilon_2 \frac{s_i^T g_i}{\| g_i \|^2} \\
-1 - \varepsilon_2 \left(1 + \frac{\varepsilon_2}{1-\varepsilon_2}\right) &\le \frac{s_{i+1}^T g_{i+1}}{\| g_{i+1} \|^2} \le -1 + \varepsilon_2 \left(1 + \frac{\varepsilon_2}{1-\varepsilon_2}\right) \\
-1 - \frac{\varepsilon_2}{1-\varepsilon_2} &\le \frac{s_{i+1}^T g_{i+1}}{\| g_{i+1} \|^2} \le -1 + \frac{\varepsilon_2}{1-\varepsilon_2}
\end{aligned}$$

(b) Z (S3a) plyne (S3b). Podle (F3) a (S3b) platí

$$\varepsilon_2 s_i^T g_i \le s_i^T g_{i+1} \le s_i^T g_i + \alpha_i \overline{G} \| s_i \|^2$$

takže

$$\alpha_i \ge -\frac{(1-\varepsilon_2) s_i^T g_i}{\overline{G} \| s_i \|^2}$$

Podle (S2) platí

$$F_i - F_{i+1} \ge -\varepsilon_1 \alpha_i s_i^T g_i \ge \frac{\varepsilon_1 (1-\varepsilon_2)(s_i^T g_i)^2}{\overline{G} \| s_i \|^2} = \frac{\varepsilon_1(1-\varepsilon_2)}{\overline{G}} \frac{(s_i^T g_i)^2}{\| g_i \|^4} \frac{\| g_i \|^4}{\| s_i \|^2}$$

Z Al-Baaliho nerovnosti (pravá část) plyne

$$-\frac{s_i^T g_i}{\| g_i \|^2} \ge 1 - \frac{\varepsilon_2}{1-\varepsilon_2} = \frac{1-2\varepsilon_2}{1-\varepsilon_2} > 0$$

neboť $\varepsilon_1 < \varepsilon_2 < 1/2$. Platí tedy

$$F_1 - \underline{F} \geq \lim_{i\to\infty}(F_1 - F_{i+1}) = \sum_{i=1}^{\infty}(F_i - F_{i+1}) \geq \sum_{i=1}^{\infty} \frac{\varepsilon_1(1-2\varepsilon_2)}{\overline{G}} \frac{\| g_i \|^4}{\| s_i \|^2}$$

Předpokládejme, že neplatí

$$\liminf_{i\to\infty} \| g_i \| = 0$$

Pak existuje $\underline{\varepsilon} > 0$ tak, že $\| g_i \| \geq \underline{\varepsilon}\ \forall i \in N$, takže

$$\infty > \frac{(F_1 - \underline{F})\overline{G}}{\underline{\varepsilon}^4 \varepsilon_1 (1 - 2\varepsilon_2)} \geq \sum_{i=1}^{\infty} \frac{1}{\| s_i \|^2}$$

(c) Z (S3a) plyne (S3b). Podle (S3b) a Al-Baaliho nerovnosti (levá část) platí

$$-s_i^T g_{i+1} \leq -\varepsilon_2 s_i^T g_i \leq \varepsilon_2 \left(1 + \frac{\varepsilon_2}{1-\varepsilon_2}\right) \| g_i \|^2 = \frac{\varepsilon_2}{1-\varepsilon_2} \| g_i \|^2$$

Použijeme-li (CGc), dostaneme

$$\begin{aligned}
\| s_{i+1} \|^2 &\leq \| g_{i+1} \|^2 - 2\frac{\| g_{i+1} \|^2}{\| g_i \|^2} s_i^T g_{i+1} + \frac{\| g_{i+1} \|^4}{\| g_i \|^4} \| s_i \|^2 \leq \\
&\leq \| g_{i+1} \|^2 + \frac{2\varepsilon_2}{1-\varepsilon_2} \| g_{i+1} \|^2 + \frac{\| g_{i+1} \|^4}{\| g_i \|^4} \| s_i \|^2 = \\
&= \frac{1+\varepsilon_2}{1-\varepsilon_2} \| g_{i+1} \|^2 + \frac{\| g_{i+1} \|^4}{\| g_i \|^4} \| s_i \|^2
\end{aligned}$$

Toto je rekurzivní vztah pro $\| s_i \|$. Postupným dosazováním dostaneme

$$\begin{aligned}
\| s_{i+1} \|^2 &\leq \frac{1+\varepsilon_2}{1-\varepsilon_2} \| g_{i+1} \|^2 + \frac{\| g_{i+1} \|^4}{\| g_i \|^4} \| s_i \|^2 \leq \\
&\leq \frac{1+\varepsilon_2}{1-\varepsilon_2} \| g_{i+1} \|^2 + \frac{\| g_{i+1} \|^4}{\| g_i \|^4} \left(\frac{1+\varepsilon_2}{1-\varepsilon_2} \| g_i \|^2 + \frac{\| g_i \|^4}{\| g_{i-1} \|^4} \| s_{i-1} \|^2 \right) \leq \\
&\leq \frac{1+\varepsilon_2}{1-\varepsilon_2} \| g_{i+1} \|^4 \sum_{j=1}^{i+1} \frac{1}{\| g_j \|^2}
\end{aligned}$$

Předpokládejme, že neplatí

$$\liminf_{i\to\infty} \| g_i \| = 0$$

Pak existuje $0 < \underline{\varepsilon} < \overline{\varepsilon} < \infty$ tak, že $\underline{\varepsilon} \leq \| g_i \| \leq \overline{\varepsilon}\ \forall i \in N$ (existence $\overline{\varepsilon}$ plyne z (F2) a (F3)), takže

$$\| s_i \|^2 \leq \frac{(1+\varepsilon_2)\overline{\varepsilon}^4}{(1-\varepsilon_2)\underline{\varepsilon}^2} i$$

a můžeme psát

$$\sum_{i=1}^{\infty} \frac{1}{\| s_i \|^2} \geq \frac{(1-\varepsilon_2)\underline{\varepsilon}^2}{(1+\varepsilon_2)\overline{\varepsilon}^4} \sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

což je spor, neboť v (b) jsme dokázali že tento součet je konečný.

Poznámka 18 Označme β_i koeficient u s_i v (CGa)-(CGc). Dá se dokázat, že metoda sdružených gradientů je globálně konvegentní, pokud

$$\beta_i \leq \frac{1}{2\varepsilon_3} \frac{\| g_{i+1} \|^2}{\| g_i \|^2} \quad \forall i \in N$$

kde $0 < \varepsilon_2 < \varepsilon_3 < 1/2$. Toto však nelze obecně zaručit. Proto se používá přerušování iteračního procesu ($s_i = -g_i$ pokud podmínka není splněna).

Věta 15 (Kvadratické ukončení) Nechť $Q : R^n \to R$ je ryze konvexní kvadratická funkce, $Q(x) = \frac{1}{2}(x-x^*)^T G(x-x^*)$. Nechť x_i, $i \in N$, je posloupnost generovaná metodou sdružených gradientů s přesným výběrem délky kroku (platí $s_i^T g_{i+1} = 0$ $\forall i \in N$). Pak existuje index $k \leq n$ tak, že $g_{k+1} = 0$ a $x_{k+1} = x^*$.

Důkaz (Pro CGa). Předpokládejme, že $g_i \neq 0$ $\forall 1 \leq i \leq n$. Dokážeme indukcí, že $s_i \neq 0$ a $\alpha_i \neq 0$ $\forall 1 \leq i \leq n$ a že platí

(α) $$s_j^T g_i = 0 \qquad \forall 1 \leq j < i \leq n+1$$

(β) $$g_j^T g_i = 0 \qquad \forall 1 \leq j < i \leq n+1$$

(γ) $$s_j^T G s_i = 0 \qquad \forall 1 \leq j < i \leq n$$

Z (β) plyne, že gradienty g_i, $1 \leq i \leq n$, jsou nenulové a vzájemně ortogonální, tudíž lineárně nezávislé, takže nutně $g_{n+1} = 0$.

Pro $i = 1$ platí $s_1^T g_1 = -g_1^T g_1 < 0$ takže $s_1 \neq 0$ a $\alpha_1 > 0$ a dále není co dokazovat. Indukční krok:
(a) Nechť $i \leq n$. Podle indukčního předpokladu (γ) platí:

$$s_j^T g_{i+1} = s_j^T g_i + s_j^T y_i = s_j^T g_i + \alpha_i s_j^T G s_i = 0$$

$\forall 1 \leq j < i$, neboť pro kvadratickou funkci $Q(x)$ platí $y_i = g_{i+1} - g_i = \alpha_i G s_i$. Z přesného výběru délky kroku plyne $s_i^T g_{i+1} = 0$. Je tedy $s_j^T g_{i+1} = 0$ $\forall 1 \leq j \leq i$.

(b) Nechť $i \leq n$. Z (CGa) a (a) dostaneme

$$\begin{aligned} g_1 &= -s_1 \\ g_j &= -s_j + \beta_{j-1} s_{j-1} \qquad \forall 1 < j \leq i \end{aligned}$$

Takže podle (a) platí

$$\begin{aligned} g_1^T g_{i+1} &= -s_1^T g_{i+1} = 0 \\ g_j^T g_{i+1} &= -s_j^T g_{i+1} + \beta_{j-1} s_{j-1}^T g_{i+1} = 0 \qquad \forall 1 < j \leq i \end{aligned}$$

(c) Nechť $i < n$. Z (CGa) a (a) dostaneme

$$s_{i+1}^T g_{i+1} = -g_{i+1}^T g_{i+1} + \frac{y_i^T g_{i+1}}{y_i^T s_i} s_i^T g_{i+1} = -g_{i+1}^T g_{i+1} < 0$$

takže $s_{i+1} \neq 0$ a $\alpha_{i+1} \neq 0$. Z (CGa) a (b) dostaneme

$$s_j^T G s_{i+1} = -s_j^T G g_{i+1} + \beta_i s_j^T G s_i = -s_j^T G g_{i+1} = -\frac{1}{\alpha_j} y_j^T g_{i+1} = -\frac{1}{\alpha_j}(g_{j+1} - g_j)^T g_{i+1} = 0$$

$\forall 1 \leq j < i$ neboť podle předpokladu (γ) platí $s_j^T G s_i = 0$ $\forall 1 \leq j < i$. Dále podle (CGa) platí

$$s_i^T G s_{i+1} = -s_i^T G g_{i+1} + \frac{y_i^T g_{i+1}}{y_i^T s_i} s_i^T G s_i = -s_i^T G g_{i+1} + \frac{s_i^T G g_{i+1}}{y_i^T s_i} y_i^T s_i = 0$$

takže $s_j^T G s_{i+1} = 0$ $\forall 1 \leq j \leq i$.

Poznámka 19 Důkaz byl proveden pro CGa. Věta 15 platí i pro ostatní metody sdružených gradientů neboť podle (β) platí

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

a z (α) plyne

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

Poznámka 20 Metoda sdružených gradientů s přesným výběrem délky kroku najde minimum kvadratické funkce po nejvýše n krocích. Neplatí to však jestliže

- výběr délky kroku není přesný
- funkce není kvadratická
- Hessova matice je špatně podmíněná a projevují se zaokrouhlovací chyby.

Pak je třeba pokračovat ve výpočtu. Aby byly i nadále splněny předpoklady věty 15, je třeba iterační proces přerušit ($s_{n+1} = -g_{n+1}$).

Definice 20 Řekneme, že základní optimalizační metoda je přerušovanou metodou sdružených gradientů, jestliže $s_i = -g_i$ pro $i \in M$ a jestliže platí některý ze vzorců (CGa), (CGb), (CGc) pro $i \notin M$, kde $M = \{nk + 1 : k \in N\}$.

Poznámka 21 Přerušovaná metoda sdružených gradientů je globálně konvergentní (stačí modifikovat důkaz věty 8 tak jak je to naznačeno v poznámce 7).

Věta 16 Nechť $x_i \in R^n$, $i \in N$, je posloupnost získaná přerušovanou metodou sdružených gradientů Fletchera a Reevese (CGc) s výběrem délky kroku splňujícím silnou Wolfeho podmínku (S2) a (S3a), kde $\varepsilon_1 < \varepsilon_2 < 1/2$, taková, že $x_i \to x^*$. Nechť funkce $F : R^n \to R$ vyhovuje podmínkám (F3) a (F4). Pak směrové vektory $\| s_i \|$, $i \in N$ jsou stejnoměrně spádové a platí $\| s_i \| \sim \| g_i \|$.

Důkaz (a) Zřejmě $\| e_i \| = O(\| e_{i-1} \|)$ (poznámka 2) a $\| g_i \| \sim \| e_i \|$ (poznámka 10), takže $\| g_i \| = O(\| g_{i-1} \|)$. Existuje tedy konstanta $c < \infty$ tak, že

$$\frac{\| g_i \|}{\| g_{i-1} \|} \leq c \qquad \forall i \notin M$$

Nechť $i \notin M$. Pak podle (CGc) platí

$$\| s_i \| \leq \| g_i \| + \frac{\| g_i \|^2}{\| g_{i-1} \|^2} \| s_{i-1} \|$$

takže

$$\frac{\| s_i \|}{\| g_i \|} \leq 1 + \frac{\| g_i \|}{\| g_{i-1} \|} \frac{\| s_{i-1} \|}{\| g_{i-1} \|} \leq 1 + c \frac{\| s_{i-1} \|}{\| g_{i-1} \|}$$

Nechť $k = \sup\{j \in M, j \leq i\}$. Protože $s_k = -g_k$, platí $\| s_k \| / \| g_k \| = 1$, takže rekurentním použitím poslední nerovnosti dostaneme

$$\frac{\| s_i \|}{\| g_i \|} \leq \sum_{j=0}^{i-k} c^j \leq \sum_{j=0}^{n} c^j \triangleq \overline{c}$$

(b) Použijeme-li Al-Baaliho nerovnost (levou část) dostaneme

$$-\frac{s_i^T g_i}{\| g_i \|^2} \geq 1 - \frac{\varepsilon_2}{1 - \varepsilon_2} = \frac{1 - 2\varepsilon_2}{1 - \varepsilon_2}$$

což spolu s (a) dává

$$-\frac{s_i^T g_i}{\| s_i \| \| g_i \|} = -\frac{s_i^T g_i}{\| g_i \|^2} \frac{\| g_i \|}{\| s_i \|} \geq -\frac{1}{\overline{c}} \frac{s_i^T g_i}{\| g_i \|^2} \geq \frac{1}{\overline{c}} \frac{1 - 2\varepsilon_2}{1 - \varepsilon_2}$$

takže $-s_i^T g_i \leq \varepsilon_0 \| s_i \| \| g_i \|$ kde $\varepsilon_0 = (1 - 2\varepsilon_2)/(\overline{c}(1 - \varepsilon_2))$.

(c) Použitím Al-Baaliho nerovnosti a Schwartzovy nerovnosti dostaneme

$$\| s_i \| \| g_i \| \geq -s_i^T g_i \geq \frac{1-2\varepsilon_2}{1-\varepsilon_2} \| g_i \|^2$$

což dává $\| s_i \| \geq \underline{c} \| g_i \|$, kde $\underline{c} = (1-2\varepsilon_2)/(1-\varepsilon_2)$.

Tvrzení 17 (n-kroková kvadratická konvergence) Nechť jsou splněny předpoklady věty 16 a nechť výběr délky kroku je asymptoticky přesný. Nechť funkce $F : R^n \to R$ splňuje navíc podmínku (F5). Pak existuje index $k \in M$ a konstanta $\overline{C} < \infty$ tak, že pro $\forall i \in M$, $i \geq k$ platí

$$\| x_{i+n} - x^* \| \leq \overline{C} \| x_i - x^* \|^2$$

Lemma 18 Nechť jsou splněny předpoklady věty 15. Nechť $g_i \neq 0$ pro nějaký index $1 \leq i \leq n$. Pak platí

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq \max_{1 \leq k \leq n} \left(\overline{P}_i(\lambda_k)\right)^2$$

kde $\overline{P}_i(\lambda)$ je libovolný polynom stupně i takový, že $\overline{P}_i(0) = 1$, a λ_k, $1 \leq k \leq n$, jsou vlastní čísla matice G.

Důkaz (a) Dokážeme indukcí, že pro $1 \leq j \leq i$ platí $g_j \in \mathcal{K}_j$ a $s_j \in \mathcal{K}_j$, kde

$$\mathcal{K}_j = \text{span}\{g_1, Gg_1, \ldots, G^{j-1} g_1\}$$

je Krylovův podprostor stupně j generovaný maticí G a vektorem g_1. Pro $j = 1$ je to zřejmé. Předpokládejme, že to platí pro $j = i-1$. Protože z $x_i = x_{i-1} + \alpha_{i-1} s_{i-1}$ plyne $g_i = g_{i-1} + \alpha_{i-1} G s_{i-1}$ (vlastnost kvadratické funkce (Q)) a protože platí $g_{i-1} \in \mathcal{K}_{i-1}$ a $G s_{i-1} \in span(Gg_1, G^2 g_1, \ldots, G^{i-1} g_1) \subset \mathcal{K}_i$ (indukční předpoklad), dostaneme $g_i \in \mathcal{K}_i$. Dále protože $s_i = -g_i + \beta_{i-1} s_{i-1}$ (CG) a protože platí $s_{i-1} \in \mathcal{K}_{i-1} \subset \mathcal{K}_i$ (indukční předpoklad) a $g_i \in \mathcal{K}_i$ (dokázaná inkluze), dostaneme $s_i \in \mathcal{K}_i$

(b) Podle (a) platí

$$\begin{aligned} x_{i+1} - x^* &= x_1 - x^* + \sum_{j=1}^{i} \alpha_j s_j = x_1 - x^* + P_{i-1}^*(G) g_1 = \\ &= x_1 - x^* + P_{i-1}^*(G) G (x_1 - x^*) = (I + G P_{i-1}^*)(x_1 - x^*) \end{aligned}$$

kde $P_{i-1}^*(G)$ je určitý polynom stupně $i-1$ v G. Označme $\overline{P}_i^* = I + G P_{i-1}^*$ (takže $\overline{P}_i^*$ je stupně i a $\overline{P}_i^*(0) = 1$). Jelikož z důkazu věty 15 plyne, že $s_j^T g_{i+1} = 0$ $\forall 1 \leq j \leq i$, je

$$x_{i+1} = x^* + \overline{P}_i^*(G)(x_1 - x^*) = \arg \min_{x = x^* + \overline{P}_i(G)(x_1 - x^*)} Q(x)$$

takže

$$\begin{aligned} Q(x_{i+1}) - Q(x^*) &= \frac{1}{2}(x_{i+1} - x^*)G(x_{i+1} - x^*) = \frac{1}{2}(x_1 - x^*)^T \overline{P}_i^*(G) G \overline{P}_i^*(G)(x_1 - x^*) \leq \\ &\leq \frac{1}{2}(x_1 - x^*)^T \overline{P}_i(G) G \overline{P}_i(G)(x_1 - x^*) \end{aligned}$$

pro libovolný polynom $\overline{P}_i$ stupně i takový, že $\overline{P}_i(0) = 0$. Nechť λ_k a v_k $1 \leq k \leq n$ jsou vlastní čísla (nezáporná) a vlastní vektory (ortonormální) matice G a nechť

$$x_1 - x^* = \sum_{k=1}^{n} \gamma_k v_k$$

Pak

$$Q(x_1) - Q(x^*) = \frac{1}{2}(x_1 - x^*)G(x_1 - x^*) = \frac{1}{2}\left(\sum_{k=1}^{n} \gamma_k v_k\right)^T G \left(\sum_{k=1}^{n} \gamma_k v_k\right) = \frac{1}{2}\sum_{k=1}^{n} \gamma_k^2 \lambda_k$$

a

$$\begin{aligned} Q(x_{i+1}) - Q(x^*) &\leq \frac{1}{2}(x_1 - x^*)^T \overline{P}_i(G) G \overline{P}_i(G)(x_1 - x^*) = \\ &= \frac{1}{2}\left(\sum_{k=1}^{n} \overline{P}_i(\lambda_k)\gamma_k v_k\right)^T G \left(\sum_{k=1}^{n} \overline{P}_i(\lambda_k)\gamma_k v_k\right) = \\ &= \frac{1}{2}\sum_{k=1}^{n} \overline{P}_i^2(\lambda_k)\gamma_k^2\lambda_k \leq \frac{1}{2}\max_{1\leq k\leq n} \overline{P}_i^2(\lambda_k)\sum_{k=1}^{n}\gamma_k^2\lambda_k \end{aligned}$$

Po vydělení dostaneme

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq \max_{1\leq k\leq n} \overline{P}_i^2(\lambda_k)$$

Věta 19 Nechť jsou splněny předpoklady věty 15. Nechť $g_i \neq 0$ pro nějaký index $1 \leq i \leq n$. Pak platí

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

Důkaz Podle lemmatu 18 platí

$$\frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} \leq \max_{1\leq k\leq n} (\overline{P}_i(\lambda_k))^2$$

pro libovolný polynom $\overline{P}_i(\lambda)$ stupně nanejvýš i takový, že $\overline{P}_i(0) = 1$. Zvolíme polynom $\overline{P}_i(\lambda)$ tak, aby minimalizoval hodnotu

$$\max_{\lambda_1 \le \lambda \le \lambda_n} |\overline{P}_i(\lambda)|$$

Tuto vlastnost má Čebyševův polynom transformovaný na interval $\lambda_1 \le \lambda \le \lambda_n$ a normovaný tak, aby nabýval hodnotu 1 pro $\lambda = 0$, tedy polynom

$$\overline{P}_i(\lambda) = \frac{T_i\left(\frac{\lambda_n+\lambda_1-2\lambda}{\lambda_n-\lambda_1}\right)}{T_i\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)}$$

kde $T_i(\xi) = \cos(i \arccos \xi)$ pro $|\xi| \le 1$ a $T_i(\xi) = ((\xi + \sqrt{\xi^2-1})^i + (\xi - \sqrt{\xi^2-1})^i)/2$ pro $|\xi| \ge 1$. Jelikož $|T_i(\xi)| \le 1$ pro $|\xi| \le 1$, platí

$$\max_{\lambda_1 \le \lambda \le \lambda_n} |\overline{P}_i(\lambda)| \le 1/T_i\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)$$

Zbývá tedy vyčíslit hodnotu na pravé straně poslední nerovnosti. Označme $\xi = (\lambda_n + \lambda_1)/(\lambda_n - \lambda_1)$. Zřejmě $|\xi| \ge 1$, takže

$$\begin{aligned} T_i(\xi) &= \frac{1}{2}((\xi + \sqrt{\xi^2-1})^i + (\xi - \sqrt{\xi^2-1})^i) \ge \frac{1}{2}(\xi + \sqrt{\xi^2-1})^i = \\ &= \frac{1}{2}\frac{1}{2^i}(\sqrt{\xi+1} + \sqrt{\xi-1})^{2i} \end{aligned}$$

neboť

$$(\sqrt{\xi+1} + \sqrt{\xi-1})^2 = 2(\xi + \sqrt{\xi^2-1})$$

Dosadíme-li hodnotu ξ, dostaneme

$$\begin{aligned} T_i\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right) &\ge \frac{1}{2}\left(\sqrt{\frac{\lambda_n}{\lambda_n-\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_n-\lambda_1}}\right)^{2i} = \frac{1}{2}\left(\frac{(\sqrt{\lambda_n}+\sqrt{\lambda_1})^2}{\lambda_n-\lambda_1}\right)^i = \\ &= \frac{1}{2}\left(\frac{\sqrt{\lambda_n}+\sqrt{\lambda_1}}{\sqrt{\lambda_n}-\sqrt{\lambda_1}}\right)^i \end{aligned}$$

Platí tedy

$$\begin{aligned} \frac{Q(x_{i+1}) - Q(x^*)}{Q(x_1) - Q(x^*)} &\le \left(\max_{1 \le k \le n} |\overline{P}_i(\lambda_k)|\right)^2 \le \left(\frac{1}{T_i\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)}\right)^2 \le \\ &\le 4\left(\frac{\sqrt{\lambda_n}-\sqrt{\lambda_1}}{\sqrt{\lambda_n}+\sqrt{\lambda_1}}\right)^{2i} = 4\left(\frac{\sqrt{\kappa(G)}-1}{\sqrt{\kappa(G)}+1}\right)^{2i} \end{aligned}$$

Důsledkem Tvrzení 17 (s výsledky získanými při jeho důkazu) a věty 19 je toto tvrzení.

Tvrzení 20 (Asymptotický odhad) Nechť jsou splněny předpoklady věty 16 a nechť výběr délky kroku je asymptoticky přesný. Nechť funkce $F : R^n \to R$ splňuje navíc podmínku (F5). Pak platí

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

Poznámka 22 Odhad $(\sqrt{\kappa}-1)/(\sqrt{\kappa}+1)$ je mnohem příznivější než odhad $(\kappa-1)/(\kappa+1)$ platný pro metodu spádových směrů jak ukazuje tato tabulka:

	SD	CG
$\kappa = 10^2$, $\varepsilon = 10^{-4}$	460	45
$\kappa = 10^4$, $\varepsilon = 10^{-6}$	69077	690
$\kappa = 10^6$, $\varepsilon = 10^{-8}$	9210340	9210

V tabulce je uveden počet iterací potřebný k dosažení požadované přesnosti ε.

Implementace metod sdružených gradientů:

1. Úpravy algoritmu pro výběr délky kroku.

   (a) Počáteční odhad

   $$\alpha = \min\left(1, \frac{2(F_i - F_{i-1})}{s_i^T g_i}, \frac{20(\underline{F} - F_i)}{s_i^T g_i}\right)$$

   kde $\underline{F}$ je dolní odhad pro minimální hodnotu funkce F.

   (b) Silná Wolfeho podmínka (S3a)

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

   kde $\varepsilon_2 = 10^{-1}$. Algoritmus 1 je třeba pozměnit tak, že v něm ponecháme podmínku (S3b) ale k podmínce (S2) přidáme podmínku

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

   která je částí podmínky (S3a).

2. Škálování. Je výhodné metodu sdružených gradientů škálovat (nejjednodušší předpodmínění). Místo

   $$-s_{i+1} = g_{i+1} - \beta_i s_i$$

   Použijeme vzorec

   $$-s_{i+1} = \gamma_{i+1}(g_{i+1} - \beta_i s_i)$$

kde

$$\gamma_{i+1} = \frac{y_i^T d_i}{y_i^T y_i}$$

($y_i = g_{i+1} - g_i, d_i = x_{i+1} - x_i = \alpha_i s_i$) V tomto případě však musíme vzorce pro parametr β_i upravit tak, že

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

$$\beta_i = \frac{1}{\gamma_i} \frac{y_i^T g_{i+1}}{g_i^T g_i} \tag{CGb}$$

$$\beta_i = \frac{1}{\gamma_i} \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} \tag{CGc}$$

Parametr γ_{i+1}, je nutné udržovat v určitých mezích ($0.005 \leq \gamma_{i+1} \leq 200$).

3. Řízené přerušování iteračního procesu.

   (a) Klasický způsob. Iterační proces se přeruší vždy po n iteracích, nebo když $\beta_i \leq 0$.

   (b) Srovnání s (CGc). Iterační proces se přeruší, pokud neplatí

   $$\eta_1 \beta_i^{FR} \leq \beta_i \leq \eta_2 \beta_i^{FR}$$

   kde $\beta_i^{FR} = g_{i+1}^T g_{i+1} / g_i^T g_i$ a $\eta_1 \approx 0.1 - 0.3$ a $\eta_2 \approx 1.1 - 1.3$.

   (c) Test na sdruženost směrů, iterační proces se přeruší pokud

   $$s_{i+1}^T y_i \geq \eta_3 \parallel s_{i+1} \parallel\parallel y_i \parallel$$

   kde $\eta_3 \approx 0.04 - 0.05$.

   (d) Test na ortogonalitu gradientů, iterační proces se přeruší pokud

   $$g_{i+1}^T g_i \geq \eta_4 \parallel g_{i+1} \parallel\parallel g_i \parallel$$

   kde $\eta_4 \approx 0.4 - 0.5$.

Algoritmus 2 (CG) Data $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 10^{-1}$, $\eta_1 = 0.3$, $\eta_2 = 1.2$, $\eta_3 = 0.05$, $\eta_4 = 0.46$, $\underline{\varepsilon} > 0$.

Krok 1 Zvolíme počáteční odhad $x_1 \in R^n$, vypočteme $F_1 = F(x_1)$, $g_1 = g(x_1)$ a položíme $i = 1$.

Krok 2 Pokud $\parallel g_i \parallel \leq \underline{\varepsilon}$ ukončíme výpočet. V opačném případě určíme koeficient β_{i-1} podle některé z metod (CGa), (CGb), (CGc) a rozhodneme o přerušení iteračního procesu podle některé ze strategií 3a, 3b, 3c, 3d. Určíme škálovací koeficient γ_i a určíme směrový vektor s_i podle 2.

Krok 3 Určíme délku kroku α_i použitím algoritmu 1 upraveného podle 1a a 1b. Položíme $x_{i+1} = x_i + \alpha_i s_i$, vypočteme $F_{i+1} = F(x_{i+1})$, $g_{i+1} = g(x_{i+1})$, zvětšíme i o 1 a přejdeme na krok 2.

Numerické výsledky: 100 proměnných, 24 testovacích funkcí.

Metoda	NIT	NFV	Čas
CGc + 3a	10789	20021	4:54 *
CGc + 3c	5652	11210	2:42
CGc + 3d	5292	10683	2:42
CGb + 3a	8940	17624	4:54 *
CGb + 3b	7131	14121	3:09
CGa + 3a	10357	20423	5:32 *
CGa + 3b	6579	12976	3:05

2.3. Metody s proměnnou metrikou

Definice 21 Řekneme, že základní optimalizační metoda je metodou s proměnnou metrikou, jestliže

$$s_i = -H_i g_i \qquad \forall i \in N \tag{VM1}$$

kde H_i, $i \in N$, jsou symetrické pozitivně definitní (SPD) matice konstruované podle rekurentního vztahu

$$H_{i+1} = \gamma_i (H_i + U_i M_i U_i^T) \tag{VM2}$$

kde $U_i \in R^{n\times 2}$, $M_i \in R^{2\times 2}$ (SPD) a $\gamma_i > 0$, a vyhovující podmínce

$$H_{i+1} y_i = \rho_i d_i \tag{VM3}$$

kde $y_i = g_{i+1} - g_i$, $d_i = x_{i+1} - x_i = \alpha_i s_i$ a $\rho_i > 0$.

Poznámka 23 Matice H_{i+1} se získává z matice H_i aktualizací jejíž hodnost je nanejvýš 2. Nejefektivnější metody s proměnnou metrikou patří do Broydenovy třídy, která je charakterizivaná výběrem $U_i = [d_i, H_i y_i]$.

Věta 21 (Kvadratické ukončení) Nechť x_i, $i \in N$, je posloupnost generovaná metodou s proměnnou metrikou z Broydenovy třídy s přesným výběrem délky kroku (platí $s_i^T g_{i+1} = 0$ $\forall i \in N$) aplikovaná na ryze konvexní kvadratickou funkci (Q). Pak existuje index $k \le n$ tak, že $g_{k+1} = 0$ a $x_{k+1} = x^*$.

Důkaz Předpokládejme, že $g_i \neq 0$ $\forall 1 \le i \le n$. Dokážeme indukcí, že $s_i \neq 0$ a $\alpha_i \neq 0$ $\forall 1 \le i \le n$ a že platí

(α) $$H_i y_j = \lambda_i^j d_j \qquad \forall 1 \le j < i \le n+1$$

(β) $$s_j^T g_i = 0 \qquad \forall 1 \le j < i \le n+1$$

(γ) $$s_j^T G s_i = 0 \qquad \forall 1 \le j < i \le n$$

Z (VM1) a (γ) plyne, že s_i, $1 \le i \le n$, jsou nenulové a vzájemně sdružené (G-ortogonální), tudíž lineárně nezávislé, takže podle (β) nutně $g_{n+1} = 0$. Pro $i = 1$ platí $s_1^T g_1 = -s_1^T H_1 s_1 < 0$ (H_1 je SPD) takže $\alpha_1 > 0$ a dále není co dokázat. Indukční krok:

(a) Nechť $i \le n$. Z (γ) a (Q) plyne $d_i^T y_j = d_i^T G d_j = \alpha_i \alpha_j s_i^T G s_j = 0$ a (α) navíc dává $y_i^T H_i y_j = \lambda_i^j y_i^T d_j = \lambda_i^j d_i^T G d_j = 0$, takže $U_i^T y_j = 0$ $\forall 1 \le j < i$. Podle (VM2) a (α) tedy platí

$$H_{i+1} y_j = \gamma_i (H_i y_j + U_i^T M_i U_i^T y_j) = \gamma_i H_i y_i = \gamma_i \lambda_i^j d_j \triangleq \lambda_{i+1}^j d_j$$

$\forall 1 \le j < i$. Použijeme-li (VM3) dostaneme $H_{i+1} y_i = \rho_i d_i \triangleq \lambda_{i+1}^i d_i$, takže $H_{i+1} y_j = \lambda_{i+1}^j d_i$ $\forall 1 \le j \le i$.

(b) Nechť $i \le n$. Z (γ) a (Q) plyne $s_j^T g_{i+1} = s_j^T g_i + s_j^T y_i = s_j^T g_i + \alpha_i s_j^T G s_i = 0$ $\forall 1 \le j < i$. Z přesného výběru délky kroku dostaneme $s_i^T g_{i+1} = 0$, takže celkem $s_j^T g_{i+1} = 0$ $\forall 1 \le j \le i$.

(c) Podle (VM1) je $g_{i+1}^T s_{i+1} = -g_{i+1}^T H_{i+1} g_{i+1} < 0$ takže $s_{i+1} \ne 0$ a $\alpha_{i+1} \ne 0$. Použijeme-li (VM1), (Q), (a), (b) dostaneme

$$s_j^T G s_{i+1} = -\frac{1}{\alpha_j} y_j^T H_{i+1} g_{i+1} = -\frac{\lambda_{i+1}^j}{\alpha_j} d_j^T g_{i+1} = -\lambda_{i+1}^j s_j^T g_{i+1} = 0$$

$\forall 1 \le j \le i$.

Věta 22 (Aproximace Hessovy matice). Nechť jsou splněny předpoklady věty 21 s $\gamma_i = 1$ a $\rho_i = 1$ $\forall i \in N$. Pak platí $H_{n+1} = G^{-1}$.

Důkaz Z důkazu věty 21 plyne, že

$$H_{n+1} y_j = d_j \qquad \forall 1 \le j \le n$$

a že vektory d_j a $y_j = G d_j$, $1 \le j \le n$, jsou lineárně nazávislé. Z tohoto důvodu musí platit $H_{n+1} = G^{-1}$.

Poznámka 24 Nyní se budeme zabývat vyšetřováním aktualizace (VM3). Pro zjednodušení budeme index i vynechávat a index $i+1$ nahradíme symbolem $+$.

Lemma 23 Nechť $H_+ = \gamma(H + UMU^T)$, kde $U = [d, Hy]$. Pak $H_+ y = \rho d$ 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 η je volný parametr a kde

$$a = y^T H y, \qquad b = y^T d, \qquad c = d^T H^{-1} d$$

Důkaz Podle (VM2) a (VM3) musí platit

$$\begin{aligned} H_+ y &= \gamma\left(Hy + [d, Hy]\begin{bmatrix} m_1 & m_2 \\ m_2 & m_3 \end{bmatrix}\begin{bmatrix} b \\ a \end{bmatrix}\right) = \\ &= \gamma(Hy + (m_1 b + m_2 a)d + (m_2 b + m_3 a)Hy) = \rho d \end{aligned}$$

takže nutně

$$m_1 b + m_2 a = \rho/\gamma$$

$$m_2 b + m_3 a = -1$$

Jeden parametr je nadbytečný. Zvolíme $m_2 = -\eta/b$ a zbylé prvky m_1, m_3 určíme řešením soustavy. Tím dostaneme matici M uvedenou v lemmatu 23.

Poznámka 25 Vztah $H_+ = \gamma(H + UMU^T)$ můžeme roznásobit. Pak platí

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

(Broydenova třída). Nejznámější členy Broydenovy třídy dostaneme, položíme-li $\eta = 0$ (metoda DFP):

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

nebo $\eta = 1$ (metoda BFGS):

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

kde $\omega = \rho/\gamma + a/b$, nebo $\eta = (\rho/\gamma)/(\rho/\gamma - a/b)$ (metoda hodnosti 1):

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

Lemma 24 Nechť H je SPD matice, $a > 0$, $b > 0$, $c > 0$, $ac - b^2 > 0$ a nechť platí (H) kde $\gamma > 0$ a $\rho > 0$. Pak matice $(1/\gamma)H^{-\frac{1}{2}}H_+H^{-\frac{1}{2}}$ má $n-2$ jednotkových vlastních čísel a zbylá dvě vlastní čísla jsou řešením kvadratické rovnice.

$$\lambda^2 - p\lambda + q = 0$$

kde

$$p = \frac{1}{b^2}(\eta(ac - b^2) + b^2) + \frac{\rho}{\gamma}\frac{c}{b}$$

$$q = \frac{\rho}{\gamma}\frac{1}{ab}(\eta(ac - b^2) + b^2)$$

Důkaz Podle (VM2) platí

$$\frac{1}{\gamma}H^{-\frac{1}{2}}H_+H^{-\frac{1}{2}} = I + H^{-\frac{1}{2}}UMU^TH^{-\frac{1}{2}}$$

Tato matice má $n-2$ jednotkových vlastních čísel odpovídajících $n-2$ vlastním vektorům kolmým k $H^{-\frac{1}{2}}U$. Zbylé dva vlastní vektory můžeme vyjádřit ve tvaru $H^{-\frac{1}{2}}Uz$ takže odpovídající vlastní čísla musí vyhovovat rovnici

$$H^{-\frac{1}{2}}U(I + MU^TH^{-1}U)z = \lambda H^{-\frac{1}{2}}Uz$$

neboli (po vynásobení $(U^TH^{-1}U)^{-1}U^TH^{-\frac{1}{2}}$ zleva)

$$((1-\lambda)I + MU^TH^{-1}U)z = 0$$

Dosadíme-li M z lemmatu 23 a

$$U^TH^{-1}U = \begin{bmatrix} c & b \\ b & a \end{bmatrix}$$

můžeme psát

$$\begin{aligned}\det((1-\lambda)I + MU^TH^{-1}U) &= \det\left(\begin{bmatrix} 1-\lambda, & 0 \\ 0, & 1-\lambda \end{bmatrix} + \right.\\ &\left. + \begin{bmatrix} \frac{1}{b}\left(\eta\frac{a}{b} + \frac{\rho}{\gamma}\right) & -\frac{\eta}{b} \\ -\frac{\eta}{b}, & -\frac{\eta-1}{a} \end{bmatrix}\begin{bmatrix} c, & b \\ b, & a \end{bmatrix}\right) = 0\end{aligned}$$

což po úpravě dává $\lambda^2 - p\lambda + q$ s koeficienty uvedenými v lemmatu 24.

Věta 25 Nechť jsou splněny předpoklady lemmatu 24. Pak H_+ je SPD právě tehdy, jestliže $\eta(ac-b^2)+b^2>0$.

Důkaz Je třeba najít podmínku pro to, aby rovnice $\lambda^2-p\lambda+q$ s koeficienty uvedenými v lemmatu 24 měla kladné kořeny. Označme λ_1 a λ_2 tyto kořeny. Pak $\lambda_1+\lambda_2=p$ a $\lambda_1\lambda_2=q$ takže $\lambda_1>0$ a $\lambda_2>0$ právě tehdy. když $p>0$ a $q>0$. Z definice p a q plyne, že

$$p=\frac{a}{b}\frac{\gamma}{\rho}+\frac{\rho}{\gamma}\frac{c}{b}$$

Jelikož předpokládáme $a>0$, $b>0$, $c>0$, $\gamma>0$, $\rho>0$, platí $p>0$ kdykoliv $q>0$. Z $q>0$ dostaneme podmínku $\eta(ac-b^2)+b^2>0$.

Poznámka 26 Poslední nerovnost lze zapsat ve tvaru $\eta>\eta^*$, kde

$$\eta^*=-\frac{b^2}{ac-b^2}<0$$

Podmínky $a>0$, $c>0$ jsou splněny, je-li matice H SPD a $d\neq 0$, $y\neq 0$. Podmínka $b>0$ je splněna, vybíráme-li délku kroku podle (S3b), neboť

$$y^Td=\alpha(g_+-g)^Ts\geq\alpha(\varepsilon_2-1)g^Ts>0$$

Jestliže $b\leq 0$, není matice H^+ SPD pro žádné hodnoty parametrů γ, ρ a η.

Věta 26 (Aktualizace matice $B=H^{-1}$). Nechť jsou splněny předpoklady lemmatu 24. Nechť $B=H^{-1}$ a nechť H_+ je matice určená pomocí aktualizace (H). Nechť $B_+=H_+^{-1}$. Pak platí

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

kde

$$\beta\eta(ac-b^2)+(\beta+\eta)b^2=b^2\tag{$*$}$$

Důkaz Inverzí vztahu $H_+=\gamma(H+UMU^T)$ dostaneme

$$B_+=\frac{1}{\gamma}(B-BU(M^{-1}+U^TBU)^{-1}U^TB)\triangleq\frac{1}{\gamma}(B+BUKU^TB)$$

(Woodburyho věta), kde $K\in R^{2\times 2}$. Jelikož podle (VM3) platí $H_+y=\rho d$, musí platit $B_+d=(1/\rho)y$ neboli

$$B_+d=\frac{1}{\gamma}\left(Bd+[Bd,y]\begin{bmatrix}k_1&k_2\\k_2&k_3\end{bmatrix}\begin{bmatrix}c\\b\end{bmatrix}\right)=\frac{1}{\gamma}(Bd+(k_1c+k_2b)Bd+(k_2c+k_3b)y)=\frac{1}{\rho}y$$

takže nutně

$$k_1 c + k_2 b = -1$$

$$k_2 c + k_3 b = \gamma/\rho$$

Zvolíme $k_2 = -\beta/b$ a zbylé prvky k_1, k_3 určíme řešením soustavy. Tím dostaneme

$$K = \begin{bmatrix} \frac{\beta-1}{c}, & -\frac{\beta}{b} \\ -\frac{\beta}{b} & \frac{1}{b}\left(\beta\frac{c}{b} + \frac{\gamma}{\rho}\right) \end{bmatrix}$$

což po dasazení do $B_+ = (1/\gamma)(B+BUKU^TB)$ dává (B). Vztah (*) lze získat například z rovnosti

$$K = -(M^{-1} + U^T BU)^{-1}$$

(nebudeme to provádět).

Poznámka 27 (Dualita) Vztah (B) dostaneme ze vztahu (H) záměnou $\gamma \to 1/\gamma$, $\rho \to 1/\rho$, $a \to c$, $c \to a$, $d \to y$, $y \to d$, $H \to B$, $\eta \to \beta$. Metody DFP a BFGS jsou navzájem duální. Metodu DFP dostaneme pro $\beta = 1$. Metodu BFGS dostaneme pro $\beta = 0$:

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

Metoda hodnosti 1 je samoduální, dostaneme ji pro $\beta = (\gamma/\rho)/(\gamma/\rho - c/b)$:

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

Poznámka 28 Matice B_+ je SPD právě tehdy, jestliže $\beta > \beta^*$, kde

$$\beta^* = -\frac{b^2}{ac - b^2} < 0$$

Lemma 27 Nechť x_i, $i \in N$ je posloupnost generovaná metodou s proměnnou metrikou z Broydenovy třídy takovou, že $\gamma_i = 1$, $\rho_i = 1$ a $(1-\lambda)\beta_i^* \le \beta_i \le 1 - \lambda$. Nechť funkce $F : R^n \to R$ splňuje podmínky (F3) a (F4) pak platí

$$\begin{aligned} Tr(B_{i+1}) &= Tr(B_i) + \frac{\| y_i \|^2}{y_i^T d_i} + \beta_i \frac{\| y_i \|^2}{y_i^T d_i}\frac{d_i^T B_i d_i}{y_i^T d_i} - (1-\beta_i)\frac{\| B_i d_i \|^2}{d_i^T B_i d_i} - 2\beta_i \frac{y_i^T B_i d_i}{y_i^T d_i} \le \\ &\le Tr(B_i) + \overline{G} + \frac{\overline{G}}{1-\varepsilon_2}\alpha_i - \frac{\lambda \underline{G}}{2(1-\varepsilon_1)}\frac{\alpha_i}{c_i^2} + \frac{2\overline{G}}{1-\varepsilon_2}\frac{\alpha_i}{c_i} \end{aligned} \tag{T}$$

a

$$\det(B_{i+1}) = \det(B_i)\frac{y_i^T d_i}{d_i^T B_i d_i}\left(1 - \frac{\beta_i}{\beta_i^*}\right) \geq \det(B_i)\lambda\frac{1-\varepsilon_2}{\alpha_i} \tag{D}$$

kde $\alpha_i > 0$ je délka kroku a $c_i > 0$ je směrový kosinus ($-s_i^T g_i = c_i \parallel s_i \parallel\parallel g_i \parallel$).

Věta 28 (Globální konvergence) Nechť jsou splněny předpoklady lemmatu 27. Pak

$$\liminf_{i\to\infty} \parallel g_i \parallel = 0$$

Důkaz Z první části (T) dostaneme

$$\begin{aligned} Tr(B_{i+1}) &\leq Tr(B_i) + \overline{G} + \overline{G}\frac{\parallel d_i \parallel^2}{y_i^T d_i}\frac{\parallel B_i d_i \parallel}{\parallel d_i \parallel} + 2\frac{\parallel y_i \parallel\parallel d_i \parallel}{y_i^T d_i}\frac{\parallel B_i d_i \parallel}{\parallel d_i \parallel} \leq \\ &\leq Tr(B_i) + \overline{G} + \left(\frac{\overline{G}}{\underline{G}} + 2\frac{\overline{G}}{\underline{G}}\right) \parallel B_i \parallel \leq \overline{G} + \left(1 + 3\frac{\overline{G}}{\underline{G}}\right) Tr(B_i) \end{aligned}$$

(používáme věty o střední hodnotě) Tato rekurze implikuje existenci konstanty $\overline{C} > 1$ takové, je

$$Tr(B_{i+1}) \leq \overline{C}^i$$

Podle druhé části (D) je

$$\det(B_{i+1}) \geq \det(B_i)\lambda\frac{1-\varepsilon_2}{\alpha_i} \geq \det(B_1)\lambda^i(1-\varepsilon_2)^i \prod_{j=1}^{i}\frac{1}{\alpha_j}$$

Použitím nerovnosti pro geometrický a aritmetický průměr dostaneme

$$\det(B_{i+1}) \leq \left[\frac{Tr(B_{i+1})}{n}\right]^n \leq Tr^n(B_{i+1}) \leq (\overline{C}^n)^i$$

takže

$$\prod_{j=1}^{i}\frac{1}{\alpha_j} \leq \frac{(\overline{C}^n)^i}{\lambda^i(1-\varepsilon_2)^i(\min(1,\det(B_1))^i} \triangleq \frac{1}{\underline{C}^i}$$

neboli podle nerovnosti pro geometrický a aritmetický průměr

$$\underline{C} \leq \left(\prod_{j=1}^{i}\alpha_j\right)^{\frac{1}{i}} \leq \frac{1}{i}\sum_{j=1}^{i}\alpha_j$$

Označme

$$\xi_i = \frac{\overline{G}}{1-\varepsilon_2} - \frac{\lambda \underline{G}}{2(1-\varepsilon_1)c_i^2} + \frac{2\overline{G}}{(1-\varepsilon_2)c_i}$$

když podle druhé části (T) platí

$$Tr(B_{i+1}) \leq Tr(B_i) + \overline{G} + \alpha_i \xi_i$$

Přepokládejme, že $\liminf_{i\to\infty} \parallel g_i \parallel > 0$. Pak podle důkazu věty 8 nutně platí $c_i \to 0$ a tedy $\xi_i \to -\infty$. Existuje tedy index $k \in N$ takový, že $\xi_i < -2\overline{G}/\overline{C}\ \forall i \geq k$. Pak ale

$$\begin{aligned} Tr(B_{i+1}) &\leq Tr(B_k) + (i+1-k)\overline{G} + \sum_{j=k}^{i} \alpha_j \xi_j < Tr(B_k) + (i+1-k)\overline{G} - \frac{2\overline{G}}{\underline{C}} \sum_{j=k}^{i} \alpha_j = \\ &= Tr(B_k) + (i+1-k)\overline{G} - \frac{2\overline{G}}{\underline{C}} \sum_{j=1}^{i} \alpha_j + \frac{2\overline{G}}{\underline{C}} \sum_{j=1}^{k-1} \alpha_j \leq \\ &\leq Tr(B_k) + \frac{2\overline{G}}{\underline{C}} \sum_{j=1}^{k-1} \alpha_j - (k-1)\overline{G} - i\overline{G} \to -\infty \end{aligned}$$

což je spor, neboť stopa SPD matice je kladná.

Tvrzení 29 (Superlineární konvergence) Nechť jsou splněny předpoklady lemmatu 27 a nechť navíc $x_i \to x^*$, funkce $F : R^n \to R$ splňuje podmínku (F5) a $\alpha_i = 1$ se vybírá vždy, když tato hodnota vyhovuje slabé Wolfeho podmínce. Pak jestliže

$$\sum_{\substack{i=1 \\ \beta_i < 0}}^{\infty} \frac{\beta_i}{\beta_i^*} < \infty$$

platí

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

Poznámka 29 K důkazu tvrzení 29 se používá invariance Broydenovy třídy metod s proměnnou metrikou k transformaci proměnných. Označíme-li $x = T\hat{x}$ pak $\hat{g} = T^T g$, $\hat{G} = T^T G T$ a také $d = T\hat{d}$, $\hat{y} = T^T y$. Označíme-li $\hat{B} = T^T B T$, $\hat{B}^+ = T^T B^+ T$, pak z (B) plyne

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

kde $\hat{a} = \hat{y}^T \hat{B}^{-1} \hat{y}$, $\hat{b} = \hat{y}^T \hat{d}$, $\hat{c} = \hat{d}^T \hat{B} \hat{d}$ a kde parametry β, γ, ρ jsou stejné jako v (B). Pro teoretické účely se pokládá $T = (G^*)^{-\frac{1}{2}}$.

Implementace metod s proměnnou metrikou:

1. Výběr délky kroku: Metody s proměnnou metrikou nejsou citlivé na výběr délky kroku. Je možné použít algoritmus 1 beze změny. Volí se počáteční odhad

$$\alpha = \min\left(1, \frac{4(\underline{F} - F_i)}{s_i^T g_i}\right)$$

2. Stabilizace (parametr ρ). Označme $\varphi(\lambda) = F(x+\lambda d)$. Parametr ρ se volí tak, aby platilo $d^T B_+ d \approx \varphi''(1)$. Použitím zpětného rozvoje $\varphi(0) = \varphi(1) - \varphi'(1) + \varphi''(\hat{\lambda})/2$, kde $0 \le \hat{\lambda} \le 1$, můžeme psát $\varphi''(\hat{\lambda}) = 2(\varphi(0) - \varphi(1) + \varphi'(1))$, takže po dosazení $d^T B_+ d = \varphi''(\hat{\lambda}) \approx \varphi''(1)$ do (VM3) dostaneme $d^T y/\rho = d^T B_+ d = 2(\varphi(0) - \varphi(1) + \varphi'(1))$, což dává

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

   Tuto hodnotu používáme pouze tehdy, jestliže $0.01 \le \rho \le 100$, v opačném případě pokládáme $\rho = 1$.

3. Škálování (parametr γ). Jelikož podle (VM3) má platit $B_+ d = y/\rho$, je výhodné volit parametr γ tak, aby Bd/γ bylo co nejblíže k y/ρ. Tedy například

$$d^T Bd/\gamma = d^T y/\rho \Rightarrow \gamma/\rho = c/b$$

   nebo

$$y^T d/\gamma = y^T B^{-1} y/\rho \Rightarrow \gamma/\rho = b/a$$

   Další možnost je geometrický střed $\gamma/\rho = (c/a)^{\frac{1}{2}}$. Tedy k danému ρ najdeme $\gamma = \rho c/b$ nebo $\gamma = \rho b/a$ nebo $\gamma = \rho(c/a)^{\frac{1}{2}}$. Pokud pro takto získanou hodnotu neplatí $0.5 \le \gamma \le 4.0$ pokládáme $\gamma = 1$.

Algoritmus 3 (VM) Data $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 0.9$, $\underline{\varepsilon} > 0$, $\underline{\rho} = 0.01$, $\overline{\rho} = 100$, $\underline{\gamma} = 0.5$, $\overline{\gamma} = 4$.

Krok 1 Zvolíme počáteční odhad $x_1 \in R^n$ vypočteme $F_1 = F(x_1)$, $g_1 = g(x_1)$, zvolíme počáteční SPD matici H_1 (obvykle $H_1 = I$) a položíme $i = 1$.

Krok 2 Pokud $\| g_i \| \le \underline{\varepsilon}$ ukončíme výpočet. V opačném případě položíme $s_i = -H_i g_i$ a určíme délku kroku α_i použitím algoritmu 1. Položíme $x_{i+1} = x_i + \alpha_i s_i$ a vypočteme $F_{i+1} = F(x_{i+1})$, $g_{i+1} = g(x_{i+1})$.

Krok 3 Položíme $d_i = x_{i+1} - x_i$ a $y_i = g_{i+1} - g_i$. Určíme parametr ρ_i podle 2. Jestliže $\rho_i < \underline{\rho}$ nebo $\rho_i > \overline{\rho}$ položíme $\rho_i = 1$. Určíme parametr γ_i podle 3. Jestliže $i > 1$ a současně $\gamma_i < \underline{\gamma}$ nebo $\gamma_i > \overline{\gamma}$, položíme $\gamma_i = 1$. Zvolíme $\eta_i = 1$ (metoda BFGS). Určíme matici H_{i+1} podle (H), zvětšíme i o 1 a přejdeme na krok 2.

Numerické výsledky: 14 testovacích funkcí

Metoda	n	NIT	NFV	Čas
CG	20	939	1976	2.36
	40	1460	3014	7.58
	80	2696	5383	33.12
VM	20	806	1029	2.31
	40	1140	1484	8.56
	80	1608	2013	35.48
MN	20	237	290	1.54
	40	276	366	7.04
	80	347	462	47.08

MN - Modifikovaná Newtonova metoda (algoritmus jako VM ale s $B_i = G_i$), používá se nemonotonní výběr délky kroku.

3. Metody s lokálně omezeným krokem

3.1. Základní vlastnosti metod s lokálně omezeným krokem

Při výkladu metod s lokálně omezeným krokem budeme používat označení

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

pro kvadratickou funkci, která lokálně aproximuje funkci $F(x_i + s)$ a označení

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

pro přesnost určení směrového vektoru. Dále budeme používat označení

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

pro podíl skutečného a předpověděného poklesu funkce $F : R^n \to R$.

Definice 22 Řekneme, že základní minimalizační metoda $x_{i+1} = x_i + \alpha_i s_i$, $i \in N$, je metodou s lokálně omezeným krokem, jestliže směrové vektory $s_i \in R^n$, $i \in N$, se určují tak, že

$$\| s_i \| \leq \Delta_i \tag{T1a}$$

$$\| s_i \| < \Delta_i \Rightarrow \omega_i(s_i) \leq \overline{\omega} \tag{T1b}$$

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

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

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

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

Přitom posloupnost $\Delta_i > 0$, $i \in N$, se konstruuje tak, že

$$\rho_i(s_i) < \underline{\rho} \Rightarrow \underline{\beta} \| s_i \| \leq \Delta_{i+1} \leq \overline{\beta} \| s_i \| \tag{T3a}$$

$$\rho_i(s_i) \geq \underline{\rho} \Rightarrow \Delta_i \leq \Delta_{i+1} \leq \overline{\gamma} \Delta_i \tag{T3b}$$

kde $0 < \underline{\beta} \leq \overline{\beta} < 1 < \overline{\gamma}$ a $0 < \underline{\rho} < 1$.

Poznámka 30 Jestliže $\overline{\omega} = 0$ nebo $\overline{\omega} > 0$ dostaneme přesné nebo nepřesné metody s lokálně omezeným krokem.

Poznámka 31 Normy v (T1) a (T3) mohou být i jiné než euklidovské. Některé podmínky mohou být oslabeny.

Poznámka 32 Označme $N_1 \subset N$ množinu indexů takových, že platí (T1b), $N_2 \subset N$ množinu indexů takových, že platí (T2b) a $N_3 \subset N$ množinu indexů takových, že platí (T3b).

Lemma 30 Aplikujeme-li metodu s lokálně omezeným krokem (T1)-(T3) na funkci $F : R^n \to R$, která splňuje podmínku (F3), existuje konstanta $\underline{c} > 0$ taková, že

$$\| s_i \| \geq \underline{c} m_i / M_i \tag{$*$}$$

kde

$$m_i = \min_{1 \leq j \leq i} \| g_j \|$$

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

Důkaz (a) Nechť $i \in N_1$. Pak podle (T1b) platí

$$| \| B_i s_i \| - \| g_i \| | \leq \| B_i s_i + g_i \| = \omega_i(s_i) \| g_i \| \leq \overline{\omega} \| g_i \|$$

takže buď $\| B_i s_i \| \geq \| g_i \|$ nebo $\| B_i s_i \| < \| g_i \|$ a $\| B_i s_i \| \geq (1-\overline{\omega}) \| g_i \|$. Spojením těchto nerovností dostaneme $\| B_i \| \| s_i \| \geq \| B_i s_i \| \geq (1-\overline{\omega}) \| g_i \|$, což dává $\| s_i \| \geq (1-\overline{\omega}) m_i/M_i$.
(b) Nechť $i \notin N_3$. Pak podle definice množiny N_3 a funkce $Q_i(s)$ platí

$$F(x_i+s_i)-F(x_i) \geq \underline{\rho} Q_i(s_i) = \underline{\rho}\left(g_i^T s_i + \frac{1}{2} s_i^T B_i s_i\right) \geq \underline{\rho}\left(g_i^T s_i - \frac{1}{2} \| B_i \| \| s_i \|^2\right)$$

Z druhé strany (věta o střední hodnotě) dostaneme

$$F(x_i+s_i)-F(x_i) \leq g_i^T s_i + \frac{1}{2}\overline{G} \| s_i \|^2$$

což dohromady dává

$$\frac{1}{2}(\overline{G}+\underline{\rho} \| B_i \|) \| s_i \|^2 \geq (\underline{\rho}-1) g_i^T s_i$$

z (T1c) dostaneme

$$-\underline{\sigma} \| g_i \| \min(\| s_i \|, \| g_i \| / \| B_i \|) \geq Q_i(s_i) \geq g_i^T s_i - \frac{1}{2} \| B_i \| \| s_i \|^2$$

což spolu s předchozí nerovností dává

$$\begin{aligned} \frac{1}{2}(\overline{G}+\underline{\rho} \| B_i \|) \| s_i \|^2 &\geq (\underline{\rho}-1) g_i^T s_i \geq \frac{1}{2}(\underline{\rho}-1) \| B_i \| \| s_i \|^2 - \\ &- \underline{\sigma}(\underline{\rho}-1) \| g_i \| \min(\| s_i \|, \| g_i \| / \| B_i \|) \end{aligned}$$

neboli

$$\frac{1}{2}(\overline{G}+ \| B_i \|) \| s_i \|^2 \geq \underline{\sigma}(1-\underline{\rho}) \| g_i \| \min(\| s_i \|, \| g_i \| / \| B_i \|)$$

takže buď $\| s_i \| \geq \| g_i \| / \| B_i \| \geq m_i/M_i$ nebo

$$\frac{1}{2}(\overline{G} M_i / \| B_1 \| + M_i) \| s_i \| \geq \underline{\sigma}(1-\underline{\rho}) \| g_i \|$$

což dává $\| s_i \| \geq [2\underline{\sigma}(1-\underline{\rho}) \| B_1 \| / (\overline{G}+ \| B_1 \|)] m_i/M_i$
(c) Nechť $i = 1$. Pokud $\| g_1 \| = 0$, platí zřejmě $\| s_1 \| \geq \| g_1 \| / \| B_1 \| \geq m_1/M_1$. Pokud $\| g_1 \| \neq 0$ můžeme psát

$$\| s_1 \| = \frac{\| s_1 \| \| B_1 \|}{\| g_1 \|} \frac{\| g_1 \|}{\| B_1 \|}$$

takže $\| s_1 \| \geq (\| s_1 \| \| B_1 \| / \| g_1 \|) m_1/M_1$
(d) Nechť $i \notin N_1$, $i \in N_3$ a $i \neq 1$. Nechť $k < i$ je index pro který neplatí současně $k \notin N_1$, $k \in N_3$ a $k \neq 1$. Pak podle (T3) a (T1a) platí

$$\| s_i \| \geq \Delta_i \geq \Delta_{k+1} \geq \min(\Delta_k, \underline{\beta} \| s_k \|) \geq \min(\| s_k \|, \underline{\beta} \| s_k \|) \geq \underline{\beta} \| s_k \|$$

takže podle (a)-(c) platí

$$\| s_i \| \geq \underline{\beta} \| s_k \| \geq \underline{c} m_i / M_i$$

kde

$$\underline{c} = \underline{\beta} \min\left((1-\overline{\omega}), \frac{2\underline{\sigma}(1-\underline{\rho}) \| B_1 \|}{\overline{G} + \| B_1 \|}, \frac{\| s_1 \| \| B_1 \|}{\| g_1 \|}\right)$$

Věta 31 Nechť $x_i \in R^n$, $i \in N$ je posloupnost generovaná metodou s lokálně omezeným krokem taková, že $\| B_i \| \leq \overline{B}$, $\forall i \in N$. Nechť funkce $F : R^n \to R$ splňuje podmínky (F1) a (F3). Pak platí

$$\liminf_{i\to\infty} \| g_i \| = 0$$

Důkaz Předpokládejme, že existuje číslo $\underline{\varepsilon} > 0$ tak, že $\| g_i \| \geq \underline{\varepsilon}$ $\forall i \in N$. Pak podle lemmatu 30 platí

$$\| s_i \| \geq \frac{\underline{c\varepsilon}}{\overline{B}}$$

Předpokládejme nejprve, že množina N_3 je nekonečná. Protože $N_3 \subset N_2$, můžeme psát

$$F_i - F_{i+1} = F(x_i) - F(x_i + s_i) \geq -\underline{\rho} Q_i(s_i) \geq \underline{\rho}\underline{\sigma}\underline{\varepsilon} \min\left(\| s_i \|, \frac{\underline{\varepsilon}}{\overline{B}}\right) \geq \underline{\rho}\underline{\sigma}\underline{\varepsilon}^2 \underline{c}/\overline{B}$$

$\forall i \in N_3$, takže

$$F_1 - \underline{F} \geq \lim_{i\to\infty}(F_1 - F_{i+1}) = \sum_{i=1}^{\infty}(F_i - F_{i+1}) \geq \sum_{i\in N_3}(F_i - F_{i+1}) \geq \sum_{i\in N_3} \underline{\rho}\underline{\sigma}\underline{\varepsilon}^2 \underline{c}/\overline{B} = \infty$$

což je spor. Předpokládejme nyní, že množina N_3 není nekonečná. Pak podle (T3a) platí $\Delta_i \to 0$ což spolu s (T1a) dává $\| s_i \| \to 0$. To je ale ve sporu s předpokladem že $\| s_i \| \geq \underline{c\varepsilon}/\overline{B}$.

Poznámka 33 Podmínku $\| B_i \| \leq \overline{B}$, $i \in N$, můžeme nahradit značně slabší podmínkou

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

Metody s lokálně omezeným krokem se používají hlavně ve spojení s Newtonovou nebo Gaussovou-Newtonovou metodou, kde podmínka $\| B_i \| \leq \overline{B}$, $i \in N$, bývá splněna (vyplývá z předpokladu $\| G_i \| \leq \overline{G}$).

Věta 32 (superlineární konvergence). Nechť $x_i \in R^n$, $i \in N$, je posloupnost generovaná metodou s lokálně omezeným krokem taková, že $x_i \to x^*$. Nechť funkce $F : R^n \to R$ splňuje podmínky (F3) a (F4). Nechť

$$\lim_{i\to\infty} \omega_i(s_i) = 0 \tag{A}$$

a

$$\lim_{i\to\infty} \frac{\| (G^* - B_i)s_i \|}{\| s_i \|} = 0 \tag{B}$$

Pak posloupnost x_i, $i \in N$, konverguje Q-superlineárně k bodu $x^* \in R^n$.

Důkaz (a) Naprosto stejným způsobem jako v důkazu věty 13 se ukáže, že existuje index $k_1 \in N$ takový, že

$$\| g_i \| / \overline{G} \leq \| s_i \| \leq \| g_i \| / \underline{G}$$

$\forall i \geq k_1$ pokud $\underline{G} < \underline{\lambda}(G^*)$ a $\overline{G} > \overline{\lambda}(G^*)$.
(b) Ukážeme, že existuje index $k_2 \geq k$ takový, že

$$-Q_i(s_i) \geq \frac{1}{2}\underline{G} \| s_i \|^2$$

$\forall i \geq k_2$ pokud $\underline{G} < \lambda(G^*)$. Označíme-li tak jako v důkazu věty 13 $u_i = (B_i s_i + g_i)/ \| g_i \|$ a $v_i = (G^* - B_i)s_i / \| s_i \|$, můžeme psát

$$\begin{aligned} -Q_i(s_i) &= -s_i^T g_i - \frac{1}{2} s_i^T B_i s_i = -s_i^T u_i \| g_i \| + \frac{1}{2} s_i^T B_i s_i = \\ &= -s_i^T u_i \| g_i \| + \frac{1}{2} s_i^T G^* s_i - \frac{1}{2} s_i^T v_i \| s_i \| \geq \\ &\geq \frac{1}{2}(\underline{\lambda}(G^*) - 2\overline{G} \| u_i \| - \| v_i \|) \| s_i \|^2 \end{aligned}$$

Jelikož $\| u_i \| \to 0$, $\| v_i \| \to 0$ a $\underline{G} < \underline{\lambda}(G^*)$, existuje index $k_2 \geq k_1$ tak, že $-Q_i(s_i) \geq (1/2)\underline{G} \| s_i \|^2$ $\forall i \geq k_2$
(c) Ukážeme, že existuje index $k_3 \geq k_2$ tak, že $i \in N_3$ $\forall i \geq k_3$. Podle věty o střední hodnotě platí

$$F(x_i + s_i) - F(x_i) = s_i^T g_i + \frac{1}{2} s_i^T G^* s_i + \frac{1}{2}\varepsilon_i \| s_i \|^2$$

kde $\varepsilon_i \to 0$ neboť $x_i \to x^*$ a $s_i \to 0$ (podle (a)). Z definice funkce $Q_i(s_i)$ dostaneme $s_i^T g_i = Q_i(s_i) - s_i^T B_i s_i/2$, takže

$$\rho_i(s_i) = \frac{F(x_i+s_i)-F(x_i)}{Q_i(s_i)} = 1 + \frac{1}{2}\frac{s_i^T(G^*-B_i)s_i + \varepsilon_i \parallel s_i \parallel^2}{Q_i(s_i)}$$

Podle (b) však platí

$$\left|\frac{s_i^T(G^*-B_i)s_i + \varepsilon_i \parallel s_i \parallel^2}{Q_i(s_i)}\right| \le \frac{(\parallel v_i \parallel + |\varepsilon_i|) \parallel s_i \parallel^2}{\underline{G} \parallel s_i \parallel^2 /2} \to 0$$

neboť $\parallel v_i \parallel \to 0$ a $|\varepsilon_i| \to 0$. Platí tedy $\rho_i(s_i) \to 1$ a jelikož $\underline{\rho} < 1$, existuje index $k_3 \ge k_2$ takový, že $\rho_i(s_i) \ge \underline{\rho}$ $\forall i \ge k_3$.
(d) Ukážeme, že existuje index $k \ge k_3$ takový, že $i \in N_1$ $\forall i \ge k$. Podle věty o střední hodnotě platí (pro $i \in N_2 \subset N_3$)

$$g_{i+1} = g(x_i+s_i) = g_i + G^* s_i + \varepsilon_i \parallel s_i \parallel$$

kde $\parallel \varepsilon_i \parallel \to 0$ neboť $x_i \to x^*$ a $\parallel s_i \parallel \to 0$. Označme

$$\lambda_i = \frac{\parallel g_{i+1} - g_i - B_i s_i \parallel}{\parallel s_i \parallel}$$

Pak podle předchozí úvahy platí $\lambda_i \le \parallel v_i \parallel + \parallel \varepsilon_i \parallel \to 0$. Jelikož zároveň $\omega_i \to 0$ existuje index $k_4 \ge k_3$ takový, že $\lambda_i < \underline{G}/2$ a $\omega_i < (\underline{G}/\overline{G})/2$ $\forall i \ge k_4$. Jestliže $i \notin N_1$ $\forall i \ge k_4$ pak $\parallel s_i \parallel \ge \Delta_i$ $\forall i \ge k_4$ a jelikož $i \in N_3$ $\forall i \ge k_4$, dostaneme $\parallel s_i \parallel \ge \Delta_{k_4}$ $\forall i \ge k_4$, což je spor s tím, že $\parallel s_i \parallel \to 0$. Existuje tedy index $k \ge k_4$ takový, že $k \in N_1$. Předpokládejme, že $i \in N_1$ pro nějaký index $i \ge k$ (platí to jistě pro $i = k$). Pak můžeme psát

$$\begin{aligned}
\parallel s_{i+1} \parallel &\le \frac{1}{\underline{G}} \parallel g_{i+1} \parallel \le \frac{1}{\underline{G}}(\parallel g_{i+1} - g_i - B_i s_i \parallel + \parallel B_i s_i + g_i \parallel) \le \\
&\le \frac{1}{\underline{G}}(\lambda_i + \overline{G}\omega_i) \parallel s_i \parallel < \left(\frac{1}{2}+\frac{1}{2}\right) \parallel s_i \parallel = \parallel s_i \parallel
\end{aligned}$$

Jelikož $i \in N_3$ podle (c), platí $\Delta_{i+1} \ge \Delta_i$, což dává $\parallel s_{i+1} \parallel < \parallel s_i \parallel \le \Delta_i \le \Delta_{i+1}$, takže $i+1 \in N_1$. Indukcí dostaneme $i \in N_1$ $\forall i \ge k$.
(e) Superlineární konvergence. Platí

$$\frac{\parallel g_{i+1} \parallel}{\parallel g_i \parallel} \le \frac{\parallel g_{i+1} - g_i - B_i s_i \parallel + \parallel B_i s_i + g_i \parallel}{\parallel g_i \parallel} \le \lambda_i/\underline{G} + \omega_i$$

což spolu s $\lambda_i \to 0$ a $\omega_i \to 0$ dává

$$\lim_{i\to\infty} \frac{\parallel x_{i+1} - x^* \parallel}{\parallel x_i - x^* \parallel} \le \frac{\overline{G}}{\underline{G}} \lim_{i\to\infty} \frac{\parallel g_{i+1} \parallel}{\parallel g_i \parallel} = 0$$

3.2. Metody s optimálním lokálně omezeným krokem

Definice 23 Metody s optimálním lokálně omezeným krokem používají směrové vektory $s_i \in R^n$, $i \in N$, takové, že

$$s_i = \arg \min_{\|s\| \leq \Delta_i} Q_i(s) \qquad (\overline{\mathrm{T}}1)$$

Věta 33 Směrový vektor určený podle ($\overline{\mathrm{T}}1$) vyhovuje podmínkám (T1) s $\overline{\omega} = 0$ a $\underline{\sigma} = 1/2$.

Důkaz Podmínka (T1a) je přímo součástí podmínky ($\overline{\mathrm{T}}1$).
(a) Předpokládejme, že $s_i \in R^n$ je řešením úlohy ($\overline{\mathrm{T}}1$), přičemž $\| s_i \| < \Delta_i$. Pak nutně $Q_i(s)$ je konvexní funkce a $B_i s_i + g_i = 0$, takže $\omega_i(s_i) = 0$ a

$$-Q_i(s_i) = g_i^T B_i^{-1} g_i - \frac{1}{2} g_i^T B_i^{-1} g_i = \frac{1}{2} g_i^T B_i^{-1} g_i \geq \frac{1}{2} \| g_i \|^2 / \| B_i \|$$

(b) Nechť $\| s_i \| = \Delta_i$. Položme

$$s = -\frac{g_i^T g_i}{g_i^T B_i g_i} g_i$$

a předpokládejme, že $\| s \| \leq \Delta_i$ a $g_i^T B_i g_i > 0$. Pak platí

$$-Q_i(s) = \frac{(g_i^T g_i)^2}{g_i^T B_i g_i} - \frac{1}{2} \frac{(g_i^T g_i)^2 g_i^T B_i g_i}{(g_i^T B_i g_i)^2} = \frac{1}{2} \frac{(g_i^T g_i)^2}{g_i^T B_i g_i} \geq \frac{1}{2} \| g_i \|^2 / \| B_i \|$$

Podle ($\overline{\mathrm{T}}1$) musí být $Q_i(s_i) \leq Q_i(s)$ takže nutně

$$-Q_i(s_i) \geq -Q_i(s) \geq \frac{1}{2} \| g_i \|^2 / \| B_i \|$$

(c) Nechť $\| s_i \| = \Delta_i$ a buď $\| s \| > \Delta_i$ nebo $g_i^T B_i g_i \leq 0$, kde $s \in R^n$ je vektor definovaný v (b). Jestliže $\| s \| > \Delta_i$, pak $\| g_i \|^3 / g_i^T B_i g_i > \Delta_i$ neboli

$$g_i^T B_i g_i < \| g_i \|^3 / \Delta_i$$

Stejná nerovnost platí pro $g_i^T B_i g_i < 0$ Položme $\| \tilde{s} \| = -(\Delta_i / \| g_i \|) g_i$ takže $\| \tilde{s} \| \leq \Delta_i$. Pak platí

$$-Q_i(\tilde{s}) = \Delta_i \| g_i \| - \frac{1}{2} \frac{\Delta_i^2}{\| g_i \|^2} g_i^T B_i g_i > \Delta_i \| g_i \| - \frac{1}{2} \Delta_i \| g_i \| = \frac{1}{2} \Delta_i \| g_i \| = \frac{1}{2} \| s_i \| \| g_i \|$$

Podle ($\overline{\mathrm{T}}1$) musí být $Q_i(s_i) \leq Q_i(\tilde{s})$ takže nutně

$$-Q_i(s_i) \geq -Q_i(\tilde{s}) \geq \frac{1}{2} \| g_i \| \| s_i \|$$

Věta 34 Vektor $s_i \in R^n$ je řešením úlohy ($\overline{\mathrm{T}}1$) 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 $(\| s_i \| - \Delta_i)\lambda_i = 0$.

Důkaz Dokážeme nejprve nutnost. Jestliže $\| s_i \| < \Delta_i$, pak nutně $B_i s_i + g_i = 0$ a $(\| s_i \| - \Delta_i) \neq 0$ a funkce $Q_i(s_i)$ je konvexní, takže matice B_i je pozitivně semidefinitní. Jsou tedy splněny dokazované podmínky s $\lambda_i = 0$. Jestliže $\| s_i \| = \Delta_i$ musí být splněny Kuhnovy-Tuckerovy podmínky $(B_i + \lambda_i I)s_i + g_i = 0$ a $(\| s_i \| - \Delta_i)\lambda_i = 0$ kde $\lambda_i \geq 0$. Zbývá dokázat pozitivní semidefinitnost matice $B_i + \lambda_i I$. Pro libovolný vektor $s \in R^n$ takový, že $\| s \| = \Delta_i$ platí

$$
\begin{aligned}
Q_i(s) - Q_i(s_i) &= (s - s_i)^T g_i + \frac{1}{2} s^T B_i s - \frac{1}{2} s_i^T B_i s_i = \\
&= (s_i - s)^T (B_i + \lambda_i I) s_i + \frac{1}{2} s^T B_i s - \frac{1}{2} s_i^T B_i s_i = \\
&= \frac{1}{2}(s_i - s)^T (B_i + \lambda_i I)(s_i - s) + \frac{1}{2}\lambda_i (s_i^T s_i - s^T s) = \\
&= \frac{1}{2}(s_i - s)^T (B_i + \lambda_i I)(s_i - s) \geq 0
\end{aligned}
$$

takže matice $B_i + \lambda_i I$ musí být pozitivně semidefinitní. Nyní dokážeme postačitelnost. Jestliže $\| s_i \| < \Delta_i$, je funkce $Q_i(s_i)$ konvexní (matice $B_i + \lambda_i I$ je pro $\lambda_i = 0$ pozitivně semidefinitní) , takže nutné podmínky jsou zároveň postačujícími podmínkami. Jestliže $\| s_i \| = \Delta_i$, pak dokazované podmínky implikují (tak jako dříve), že

$$
\begin{aligned}
Q_i(s) - Q_i(s_i) &= (s - s_i)^T g_i + \frac{1}{2} s^T B_i s - \frac{1}{2} s_i^T B_i s_i = \\
&= \frac{1}{2}(s_i - s)^T (B_i + \lambda_i I)(s_i - s) + \frac{1}{2}\lambda_i (s_i^T s_i - s^T s) \geq \\
&\geq \frac{1}{2}(s_i - s)^T (B_i + \lambda_i I)(s_i - s) \geq 0
\end{aligned}
$$

pro všechny vektory $s \in R^n$ takové, že $\| s \| \leq \| s_i \| = \Delta_i$.

Algoritmus Data $0 < \underline{\delta} < 1 < \overline{\delta}$ (obvykle $\underline{\delta} = 0.9$ a $\overline{\delta} = 1.1$).

Krok 1 Určíme $\underline{\gamma}$ jako maximální diagonální prvek matice $-B$. Položíme $\underline{\lambda} = 0$ a $\overline{\lambda} = \| g \| / \Delta + \| B \|$. Položíme $\lambda = \max(\underline{\gamma}, \underline{\lambda})$.

Krok 2 Položíme $\underline{\lambda} = \max(\underline{\gamma}, \underline{\lambda})$. Jestliže $\lambda < \underline{\gamma}$ položíme $\lambda = \sqrt{\underline{\lambda}, \overline{\lambda}}$.

Krok 3 Je-li matice $B + \lambda I$ SPD, určíme rozklad $R^T R = B + \lambda I$ a přejdeme na krok 4. V opačném případě určíme vektor $v \in R^n$ takový, že $\| v \| = 1$ a $v^T (B + \lambda I) v < 0$, položíme $\underline{\gamma} = \lambda - v^T (B + \lambda I) v$ a přejdeme na krok 2.

Krok 4 Určíme vektor $s \in R^n$ řešením rovnice $R^T Rs + g = 0$. Jestliže $\| s \| > \overline{\delta}\Delta$, položíme $\underline{\lambda} = \lambda$ a přejdeme na krok 6. Jestliže $\underline{\delta}\Delta \leq \| s \| \leq \overline{\delta}\Delta$ ukončíme výpočet. Jestliže $\| s \| < \underline{\delta}\Delta$ a $\lambda = 0$ ukončíme výpočet. Jestliže $\| s \| < \underline{\delta}\Delta$ a $\lambda \neq 0$ položíme $\overline{\lambda} = \lambda$ a přejdeme na krok 5.

Krok 5 Určíme vektor $v \in R^n$ tak, aby tento vektor byl dobrou aproximací vlastního vektoru matice B příslušného vlastnímu číslu $\underline{\lambda}(B)$ a aby platilo $\| v \| = 1$ a $v^T s \geq 0$. Určíme číslo $\alpha \geq 0$ tak, aby platilo $\| s + \alpha v \| = \Delta$. Jestliže $\alpha^2 \| Rv \|^2 \leq (1 - \underline{\delta}^2)(\| Rs \|^2 + \lambda\Delta^2)$, položíme $s := s + \alpha v$ a ukončíme výpočet. V opačném případě položíme $\underline{\gamma} = \lambda - \| Rv \|^2$ a přejdeme na krok 6.

Krok 6 Určíme vektor $v \in R^n$ řešením rovnice $R^T v = s$ a položíme

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

Pokud $\lambda < \underline{\lambda}$ položíme $\lambda = \underline{\lambda}$. Pokud $\lambda > \overline{\lambda}$ položíme $\lambda = \overline{\lambda}$. Přejdeme na krok 2

3.4. Nepřesné metody s lokálně omezeným krokem

K určení lokálně omezeného kroku můžeme velmi efektivně použít metodu sdružených gradientů aplikovanou na minimalizaci kvadratické funkce

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

(vynecháváme index i). Metoda sdružených gradientů používá rekurentní vztahy

$$s_1 = 0, \quad g_1 = g \quad p_1 = -g$$

a

$$q_i = Bp_i, \quad \alpha_i = \| g_i \|^2 / p_i^T q_i \tag{CG}$$
$$s_{i+1} = s_i + \alpha_i p_i$$
$$g_{i+1} = g_i + \alpha_i q_i, \quad \beta_i = \| g_{i+1} \|^2 / \| g_i \|^2$$
$$p_{i+1} = -g_{i+1} + \beta_i p_i$$

pro $1 \leq i \leq n$.

Poznámka 34 Platí $g_i = Bs_i + g$.

Poznámka 35 Hodnota $\alpha_i = \| g_i \|^2 / p_i^T q_i$ realizuje přesný výběr délky kroku, neboť platí

$$p_i^T g_{i+1} = p_i^T g_i + \alpha_i p_i^T q_i = - \| g_i \|^2 + (\| g_i \|^2 / p_i^T g_i) p_i^T q_i = 0$$

Použili jsme indukční krok

$$p_i^T g_i = - \| g_i \|^2 + \beta_{i-1} p_{i-1}^T g_i = - \| g_i \|^2$$

Věta 35 Aplikujeme-li metodu sdružených gradientů na kvadratickou funkci $Q(s)$ a platí-li $p_i^T B p_i > 0$ pro $1 \le i \le m$, pak

$$\begin{aligned} Q(s_{i+1}) &\le -\frac{1}{2} \| g \|^2 / \| B \| \\ Q(s_{i+1}) &< Q(s_i) \\ \| s_{i+1} \| &> \| s_i \| \end{aligned}$$

pro $1 \le i \le m$.

Důkaz Z důkazu věty 15 plyne, že

$$p_j^T B p_i = 0 \qquad \forall 1 \le j < i \le m$$

$$\begin{aligned} g_j^T g_i = 0 \qquad & \forall 1 \le j < i \le m+1 \\ p_j^T g_i = 0 \qquad & \forall 1 \le j < i \le m+1 \end{aligned}$$

použijeme-li (CG) dostaneme

$$\begin{aligned} Q(s_{i+1}) &= g^T(s_i + \alpha_i p_i) + \frac{1}{2}(s_i + \alpha_i p_i)^T B (s_i + \alpha_i p_i) = \\ &= Q(s_i) + \alpha_i g^T p_i + \alpha_i s_i^T B p_i + \frac{1}{2}\alpha_i^2 p_i^T B p_i = \\ &= Q(s_i) + \alpha_i g_i^T p_i + \frac{1}{2}\alpha_i^2 p_i^T B p_i = \\ &= Q(s_i) - \frac{\| g_i \|^4}{p_i^T B p_i} + \frac{1}{2}\frac{\| g_i \|^4}{p_i^T B p_i} = \\ &= Q(s_i) - \frac{1}{2}\frac{\| g_i \|^4}{p_i^T B p_i} < Q(s_i) \end{aligned}$$

Dále platí

$$\begin{aligned} \| s_{i+1} \|^2 &= (s_i + \alpha_i p_i)^T (s_i + \alpha_i p_i) = \| s_i \|^2 + \alpha_i^2 \| p_i \|^2 + 2\alpha_i s_i^T p_i = \\ &= \| s_i \|^2 + \alpha_i^2 \| p_i \|^2 + 2\alpha_i \sum_{j=1}^{i-1} \alpha_j p_j^T p_i = \\ &= \| s_i \|^2 + \alpha_i^2 \| p_i \|^2 + 2\alpha_i \sum_{j=1}^{i-1} \alpha_j \frac{\| p_j \|^2}{\| g_j \|^2} \| g_i \|^2 > \| s_i \| \end{aligned}$$

neboť

$$\begin{aligned} p_j^T p_i &= p_j^T(-g_i + \beta_{i-1} p_{i-1}) = \beta_{i-1} p_j^T p_{i-1} = \\ &= \left(\prod_{k=j}^{i-1} \beta_k \right) \| p_j \|^2 = (\| g_i \|^2 / \| g_j \|^2) \| p_j \|^2 \end{aligned}$$

Protože

$$s_2 = s_1 + \frac{\| g_1 \|^2}{p_1^T B p_1} p_1 = -\frac{\| g \|^2}{g^T B g} g$$

platí podle části (b) důkazu věty 33

$$-Q(s_2) \geq \frac{1}{2} \| g \|^2 / B$$

což spolu s $Q(s_{i+1}) < Q(s_i)$ pro $1 \leq i \leq m$ dává $Q(s_{i+1}) < -(1/2) \| g \|^2 / \| B \|$.

Poznámka 36 Jestliže $p_i^T B p_i \leq 0$ pak

$$Q(s_i + \alpha_i p_i) = Q(s_i) + \alpha_i g_i^T p_i + \frac{1}{2}\alpha_i^2 p_i^T B p_i \leq Q(s_i) - \alpha_i \| g_i \|^2 < Q(s_i)$$

pro libovolnou hodnotu $\alpha_i \geq 0$. Jestliže $\| s_i \| \leq \Delta$ a $p_i^T B p_i \leq 0$, určíme číslo $\alpha_i \geq 0$ tak, aby platilo $\| s_i + \alpha_i p_i \| = \Delta$ a položíme $s = s_i + \alpha_i p_i$. Podle věty 35 platí $Q(s) \leq -(1/2) \| g \|^2 / \| B \|$ pro $i \geq 2$. Podle části (c) důkazu věty 33 to platí i pro $i = 1$.

Poznámka 37 Číslo $\alpha_i \geq 0$, pro které platí $\| s_i + \alpha_i p_i \| = \Delta_i$, určujeme podle vzorce

$$\alpha_i = -p_i^T s_i + \sqrt{(p_i^T s_i)^2 + \Delta^2 - \| s_i \|^2}$$

Poznámka 38 Metoda sdružených gradientů generuje cestu v přístupné oblasti, to znamená křivku $s(t) \in R^n$ takovou, že

$$\frac{d \| s(t) \|}{dt} > 0$$

$$\frac{dQ(s(t))}{dt} < 0$$

Algoritmus 5 Data $0 < \omega < 1$, $0 < \Delta$, $m \geq n$

Krok 1 Položíme $s = 0$, $r = -g$, $\sigma = \| r \|^2$, $p = r$ a $k = 1$.

Krok 2 Položíme $\rho = \sigma$, vypočteme vektor $q = Bp$ a číslo $\tau = p^T q$. Jestliže $\tau \leq 0$, určíme číslo $\alpha \geq 0$ tak, aby platilo $\| s + \alpha p \| = \Delta$, položíme $s := s + \alpha p$ a ukončíme výpočet

Krok 3 Položíme $\alpha = \rho / \tau$. Jestliže $\| s + \alpha p \| \geq \Delta$, určíme číslo $\alpha \geq 0$ tak, aby platilo $\| s + \alpha p \| = \Delta$, položíme $s := s + \alpha p$ a ukončíme výpočet

Krok 4 Položíme $s := s + \alpha p$, $r := r - \alpha q$ a $\sigma = \| r \|^2$. Jestliže $\sigma \leq \omega^2 \| g \|^2$ nebo $k \geq m$, ukončíme výpočet

Krok 5 Položíme $\beta = \sigma / \rho$, $p := r + \beta p$, $k := k + 1$ a přejdeme na krok 2

(Obvykle volíme $m = n + 3$).

3.5. Využití směru největšího spádu (metody psí nohy)

Věta 36 Nechť jsou splněny předpoklady věty 35 pro $m \geq 1$, přičemž $\| s_{m+1} \| < \Delta$ a $Bs_{m+1} + g \neq 0$. Nechť $s_n \in R^n$ je vektor takový, že $Bs_n + g = 0$. Pak platí

$$\frac{dQ(s_{m+1} + \alpha(s_n - s_{m+1}))}{d\alpha} = (1 - \alpha)(s_n - s_{m+1})^T g_{m+1}$$

Důkaz Jelikož

$$\begin{aligned} Q(s_{m+1} + \alpha(s_n - s_{m+1})) &= g^T(s_{m+1} + \alpha(s_n - s_{m+1})) + \\ &+\frac{1}{2}(s_{m+1} + \alpha(s_n - s_{m+1}))^T B(s_{m+1} + \alpha(s_n - s_{m+1})) \end{aligned}$$

platí

$$\begin{aligned} \frac{dQ(s_{m+1} + \alpha(s_n - s_{m+1}))}{d\alpha} &= (s_n - s_{m+1})^T g + (s_n - s_{m+1})^T B(s_{m+1} + \alpha(s_n - s_{m+1})) = \\ &= (s_n - s_{m+1})^T B(s_{m+1} - s_n + \alpha(s_n - s_{m+1})) = \\ &= (1 - \alpha)(s_n - s_{m+1})^T B(s_{m+1} - s_n) = \\ &= (1 - \alpha)(s_n - s_{m+1})^T (Bs_{m+1} + g) = \\ &= (1 - \alpha)(s_n - s_{m+1})^T g_{m+1} \end{aligned}$$

Algoritmus 6 Data $0 < \Delta$, $m < n$

Krok 1 Jako v algoritmu 5.

Krok 2 Jako v algoritmu 5.

Krok 3 Jako v algoritmu 5.

Krok 4 Položíme $s := s + \alpha p$, $r := r - \alpha q$ a $\sigma = \| r \|^2$. Jestliže $k < m$ položíme $\beta = \sigma/\rho$, $p := r + \beta p$, $k := k + 1$ a přejdeme na krok 2.

Krok 5 Řešíme soustavu rovnic $Bs^* + g = 0$. Pokud $(s^* - s)^T r > 0$, určíme číslo $\alpha \geq 0$ tak, aby platilo $\| s + \alpha(s^* - s) \| = \Delta$. Jestliže $\alpha \geq 1$ položíme $s := s^*$ a ukončíme výpočet. Jestliže $\alpha < 1$ položíme $s := s + \alpha(s^* - s)$ a ukončíme výpočet. Pokud $(s^* - s)^T r \leq 0$, určíme číslo $\alpha \geq 0$ tak, aby platilo $\| s - \alpha(s^* - s) \| = \Delta$, položíme $s := s - \alpha(s^* - s)$ a ukončíme výpočet.

(Obvykle volíme $m \leq 3$. Pro $m = 1$ dostaneme jednoduchou metodu psí nohy).

3.6. Maticové rozklady (indefinitní matice B)

1) Gillův-Murrayův rozklad:

Gillův-Murrayův rozklad má tvar

$$R^T R = B + E$$

kde R je regulární horní trojúhelníková matice a E je pozitivně semidefinitní diagonální matice (může být $E = 0$). Na začátku eliminačního kroku máme matici

$$\begin{bmatrix} R_{i-1,i-1}, & R_{i-1,i}, & R_{i-1,n-i} \\ *, & B_{ii}^{i-1}, & B_{i,n-i}^{i-1} \\ *, & *, & B_{n-i,n-i}^{i-1} \end{bmatrix}$$

Eliminační krok (δ-malé, $\beta > \sqrt{\| B \|}$):

$$\gamma_i = \max_{i<j\leq n}(|B_{i,n-i}^{(i-1)}|)$$

$$\rho_i^2 = \max\left(|B_{ii}^{(i-1)}|, \frac{\gamma_i^2}{\beta^2}, \delta^2\right)$$

$$R_{ii} = \rho_i$$

$$R_{i,n-i} = B_{i,n-i}^{(i-1)}/R_{ii}$$

$$B_{n-i,n-i}^{(i)} = B_{n-i,n-i}^{(i-1)} - R_{i,n-i}^T R_{i,n-i}$$

Pro prvky matice E platí

$$E_{ii} = \rho_i^2 - B_{ii}^{(i-1)} = \rho_i^2 + R_{i,n-i} R_{i,n-i}^T - B_{ii}$$

Věta 37 Nechť $R^T R = B + E$ je Gillův-Murrayův rozklad s $\delta = 0$ a $\beta > \sqrt{\| B \|}$. Nechť

$$B_{kk}^{k-1} = \min_{1\leq i\leq n} B_{ii}^{(i-1)}$$

a nechť $v \in R^n$ je vektor určený řešením rovnice $Rv = e_k$ (e_k ke k-tý sloupec jednotkové matice). Není-li matice B pozitivně semidefinitní, platí

$$v^T B v = \frac{B_{kk}^{(k-1)}}{\rho_k^2} < 0$$

Důkaz Z rovnice $Rv = e_k$ plyne, že $v_k = 1/\rho_k$. Platí tedy

$$\begin{aligned} v^T B v &= v^T (B+E) v - v^T E v \leq v^T R^T R v - v_k^2 E_{kk} = \\ &= e_k^T e_k - E_{kk}/\rho_k^2 = \frac{\rho_k^2 - E_{kk}}{\rho_k^2} = \frac{B_{kk}^{(k-1)}}{\rho_k^2} \end{aligned}$$

Není-li matice B pozitivně semidefinitní, musí existovat index $1 \leq i \leq n$ tak, že $E_{ii} \neq 0$, neboli $\rho_i^2 \neq B_{ii}^{(i-1)}$. Mohou nastat dva případy. Buď $\rho_i^2 = |B_{ii}^{(i-1)}| \neq B_{ii}^{(i-1)}$, takže $B_{ii}^{(i-1)} < 0$ a tedy i $B_{kk}^{(k-1)} < 0$, nebo $\rho_i^2 = \gamma_i^2/\beta^2$. Ve druhém případě musí existovat index $i < j \leq n$ tak, že $\gamma_i = |B_{ij}^{(i-1)}|$, takže

$$|R_{ij}| = \frac{|B_{ij}^{(i-1)}|}{\rho_i} = \frac{\gamma_i}{\gamma_i/\beta} = \beta$$

což dává

$$B_{ii}^{(i-1)} = \rho_i^2 - E_{ii} = B_{ii} - R_{i,n-i} R_{i,n-1}^T \leq B_{ii} - \beta^2 < \| B \| - \| B \| = 0$$

2) Bunchův-Parlettův rozklad:

Bunchův-Parlettův rozklad má tvar

$$LDL^T = PBP^T$$

kde

$$L = \begin{bmatrix} I, & 0, & \dots, & 0 \\ L_{21}, & I, & \dots, & 0 \\ \vdots & \vdots & \vdots & \vdots \\ L_{n1}, & L_{n2}, & \dots, & I \end{bmatrix}, \quad D = \begin{bmatrix} D_{11}, & 0, & \dots, & 0 \\ 0, & D_{22}, & \dots, & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0, & 0, & \dots, & D_{nn} \end{bmatrix}$$

Tedy L je dolní trojúhelníková matice s jednotkovými bloky na diagonále a D je blokově diagonální matice (bloky mají rozměr 1×1 nebo 2×2). Na začátku eliminačního kroku máme matici

$$\begin{bmatrix} D_{11}, & L_{12}, & \dots, & L_{1,i-1}, & L_{1,m-i+1} \\ *, & D_{22}, & \dots, & L_{2,i-1}, & L_{2,m-i+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ *, & *, & \dots, & D_{i-1,i-1}, & L_{i-1,m-i+1} \\ *, & *, & \dots, & *, & B^{(i-1)} \end{bmatrix}$$

Eliminační krok má tvar:

$$\beta_i = \max_k |B_{kk}^{(i-1)}|$$

$$\gamma_i = \max_{k,l} |B_{kl}^{(i-1)}|$$

$$\alpha_i = \beta_i / \gamma_i$$

Jestliže $\alpha_i \geq \left(\sqrt{17}+1\right)/8$ volíme v i-tém kroku blok 1×1, jinak volíme blok 2×2. Je třeba provádět permutace (pivotový blok s indexy k a l se přenese do levého horního rohu matice $B^{(i-1)}$).Pak se provede transformace

$$B^{(i-1)} \rightarrow \begin{bmatrix} D_{ii}, & L_{i,m-i} \\ *, & B^{(i)} \end{bmatrix}$$

kde

$$D_{ii} = B_{ii}^{(i-1)}$$

$$L_{i,m-i} = D_{ii}^{-1} B_{i,m-i}^{(i-1)}$$

$$B^{(i)} = B_{m-i,m-i}^{(i-1)} - L_{i,m-i}^T B_{i,m-i}^{(i-1)}$$

Věta 38 Nechť $LDL^T = PBP^T$ je Bunchův-Parlettův rozklad. Nechť $u_i = 0$, pokud $\underline{\lambda}(D_{ii}) \geq 0$, a u_i je normalizovaný vlastní vektor příslušný $\underline{\lambda}(B_{ii})$, pokud $\underline{\lambda}(B_{ii}) < 0$. Nechť $L^T P v = u$, kde $u^T = [u_1, \ldots, u_m]$. Není-li matice B pozitivně semidefinitní, platí

$$v^T B v = \sum_{\underline{\lambda}(D_{ii})<0} \underline{\lambda}(D_{ii}) < 0$$

Důkaz Z rovnice $L^T P v = u$ dostaneme

$$v^T B v = v^T P^T L D L^T P v = u^T D u = \sum_{i=1}^{m} u_i^T D_{ii} u_i = \sum_{\underline{\lambda}(D_{ii})<0} \underline{\lambda}(D_{ii})$$

Není-li matice B pozitivně semidefinitní, existuje alespoň jeden blok D_{kk} matice D, který není pozitivně semidefinitní, takže $\underline{\lambda}(D_{kk}) < 0$. Platí tedy

$$v^T B v = \sum_{\underline{\lambda}(D_{ii})<0} \underline{\lambda}(D_{ii}) \leq \underline{\lambda}(D_{kk}) < 0$$

3.7. Newtonova metoda

Newtonova metoda používá matici $B_i = G(x_i)$, takže z (F3) plyne $\| B_i \| \leq \overline{G}$ a Newtonova metoda realizovaná jako metoda s lokálně omezeným krokem je globálně konvergentní pokud funkce $F : R^n \to R$ splňuje podmínky (F1) a (F3). Nejpoužívanější realizace Newtonovy metody:

1. Nepřesná Newtonova metoda ($\omega_i(s_i) > 0$). Jestliže platí (F4) a $\omega_i(s_i) \to 0$, je tato metoda Q-superlineárně konvergentní (soustava $B_i s_i + g_i = 0$ se řeší nepřesně metodou sdružených gradientů $\Rightarrow$ méně než $O(n^3)$ operací na iteraci, což je výhodné pro rozsáhlé úlohy).

2. Newtonova metoda s optimálním lokálně omezeným krokem. Pro tuto metodu platí velmi silná tvrzení.

Tvrzení 39 Nechť jsou splněny předpoklady (F1)-(F3). Nechť x_i, $i \in N$ je posloupnost určená Newtonovou metodou s optimálním lokálně omezeným krokem. Pak existuje hromadný bod $x^* \in R^n$ posloupnosti x_i, $i \in N$ takový, že $g(x^*) = 0$ a $G(x^*) \geq 0$.

Tvrzení 40 Nechť jsou splněny předpoklady tvrzení 39, přičemž bod $x^* \in R^n$ vyhovuje postačujícím podmínkám pro extrém ($g(x^*) = 0$ a $G(x^*) > 0$). Pak $x^* \in R^n$ je jediným hromadným bodem posloupnosti x_i, $i \in N$, a posloupnost x_i, $i \in N$, konverguje Q-superlineárně k bodu $x^* \in R^n$.

Numerické porovnání: 15 testovacích problémů (20 proměnných)

Metoda	IT - IF - IG	selhání	čas
S - (GM)	312 - 508 - 508	1	7.63
S - (BP)	342 - 488 - 488	1	16.03
T - optim. (GM)	281 - 321 - 296	-	4.61
T - psí n. (BP)	292 - 325 - 307	-	10.38

(S = metoda spádových směrů, T = metoda s lokálně omezeným krokem, GM = Gillův-Murrayův rozklad, BP = Bunchův-Parlettův rozklad)

3.8. Gaussova-Newtonova metoda pro součet čtverců

Nechť

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

Pak platí

$$g(x) = J^T(x)f(x) = \sum_{k=1}^{m} f_k(x)g_k(x)$$

$$G(x) = J^T(x)J(x) + C(x) = \sum_{k=1}^{m} g_k(x)g_k^T(x) + \sum_{k=1}^{m} f_k(x)G_k^T(x)$$

V Gaussově-Newtonově metodě zanedbáváme člen $C(x_i)$, takže

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

Zdůvodnění:

1. Úlohy s nulovým reziduem ($F(x^*) = 0$). Z $x_i \to x^*$ plyne $F(x_i) \to F(x^*) = 0$ a tedy $f_k(x_i) \to 0$ $\forall 1 \leq k \leq m$. Jestliže $\| G_k(x) \| \leq \overline{G}$, pak i

$$\| C(x_i) \| = \| \sum_{k=1}^{m} f_k(x_i)G_k(x_i) \| \leq \overline{G} \sum_{k=1}^{m} |f_k(x_i)| \to 0$$

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

2. Linearizace

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

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 $B_i = J_i^T J_i$.

Podmínky kladené na funkci $F : R^n \to R$. Podmínka (F1) je splněna vždy, neboť $F(x) \geq 0$ $\forall x \in R^n$. Podmínku (F3) nahradíme podmínkou

$$\| G_k(x) \| \leq \overline{G} \tag{$\overline{\text{F3}}$}$$

$\forall x \in R^n$, $\forall 1 \leq k \leq m$. Z (F2) a ($\overline{\text{F3}}$) plyne omezenost gradientů i funkčních hodnot

$$\| g_k(x) \| \leq \overline{g}$$

$$|f_k(x)| \leq \overline{f}$$

$\forall x \in \mathcal{L}(F(x_1))$, $\forall 1 \leq k \leq m$, a tudíž i (F3).

Věta 41 Nechť jsou splněny podmínky (F2) a ($\overline{\text{F3}}$). Pak Gaussova-Newtonova metoda realizovaná jako metoda s lokálně omezeným krokem je globálně konvergentní. Jestliže navíc $F(x_i) \to 0$ a $\omega_i(s_i) \to 0$ je rychlost konvergence Q-superlineární.

Důkaz Z (F2) a ($\overline{\text{F3}}$) plyne $\| G_k(x) \| \leq \overline{G}$, $\| g_k(x) \| \leq \overline{g}$, $|f_k(x)| \leq \overline{f}$ $\forall x \in R^n$, $\forall 1 \leq k \leq m$. Platí tedy jednak

$$\| G(x) \| \leq \sum_{k=1}^{m} \| g_k(x) \|^2 + \sum_{k=1}^{m} |f_k(x)| \, \| G_k(x) \| \leq m\overline{g}^2 + m\overline{fG}$$

(podmínka (F3)) a jednak

$$\| B_i \| = \| \sum_{k=1}^{m} g_k(x_i) g_k^T(x_i) \| \leq \sum_{k=1}^{m} \| g_k(x_i) \|^2 \leq m\overline{g}^2$$

takže podle věty 31 je metoda globálně konvergentní. Jak již bylo ukázáno z $F(x_i) \to 0$ plyne $B_i \to G(x_i)$, neboli

$$\frac{\| (B_i - G_i)s_i \|}{\| s_i \|} \leq \| B_i - G_i \| \to 0$$

což spolu s $\omega_i(s_i) \to 0$ implikuje Q-superlineární konvergenci (věta 32).

Určení směrového vektoru:

1. Normální soustava rovnic. Rovnice $B_i s_i + g_i = 0$ má tvar

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

2. Řešení linearizované úlohy ve smyslu nejmenších čtverců

$$J_i s_i + f_i \approx 0$$

   Používá se QR-rozklad $J_i = Q_i R_i$ kde $Q_i^T Q_i = I$, takže

$$Q_i^T J_i s_i = R_i s_i = Q_i^T f_i$$

   R_i - horní trojúhelníková matice. QR-rozklad je stabilní a je možné určit pseudohodnost matice J_i a následně snížit dimenzi soustavy. Při realizaci s lokálně omezeným krokem můžeme soustavu

$$(J_i^T J_i + \lambda I)s + J_i^T f_i = 0$$

nahradit linearizovanou úlohou

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

3. Systémové rovnice. Označme $r_i = -(J_i s_i + f_i)$. Směrový vektor hledáme tak, aby platilo $J_i r_i = 0$. To dohromady dává

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

což je soustava $m + n$ rovnic se symetrickou indefinitní maticí. Je to vhodné pro řídké úlohy nebo pro vážené úlohy. Jestliže

$$F(x) = \frac{1}{2} f^T(x) W f(x)$$

kde W je váhová matice, pak normální soustava má tvar

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

a označíme-li $r_i = -W(J_i s_i + f_i)$, dostaneme

$$\begin{bmatrix} W^{-1} & J_i \\ J_i^T & 0 \end{bmatrix} \begin{bmatrix} r_i \\ s_i \end{bmatrix} + \begin{bmatrix} f_i \\ 0 \end{bmatrix} = 0$$

takže některé váhy mohou být i nekonečné (úlohy s omezeními).

3.9. Hybridní metody

Gaussova-Newtonova metoda je velmi efektivní pro úlohy s nulovými rezidui, může však selhávat v případě úloh s velkými rezidui. Proto se nabízí tato strategie:

$F_i \to F^* = 0 \Rightarrow$ Gaussova-Newtonova metoda

$F_i \to F^* > 0 \Rightarrow$ Metoda BFGS (s proměnnou metrikou)

Lemma 42 Nechť $F_i \to F^* = 0$ Q-superlineárně. Pak

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

Nechť $F_i \to F^* > 0$. Pak

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

Důkaz Jestliže $F_i \to F^* = 0$ Q-superlineárně, pak platí

$$\lim_{i\to\infty} \frac{F_i - F_{i+1}}{F_i} = 1 - \lim_{i\to\infty} \frac{F_{i+1} - F^*}{F_i - F^*} = 1 - 0 = 1$$

Jestliže $F_i \to F^* > 0$ pak

$$\lim_{i\to\infty} \frac{F_i - F_{i+1}}{F_i} = \frac{1}{F^*} \lim_{i\to\infty} (F_i - F_{i+1}) = 0$$

Hybridní metoda: Nechť $B_1 = J_1^T J_1$. Jestliže $i > 1$ a $(F_i - F_{i+1})/F_i > \eta_0$, položíme

$$B_{i+1} = J_{i+1}^T J_{i+1}$$

Jestliže $i > 1$ a $(F_i - F_{i+1})/F_i \le \eta_0$, položíme

$$B_{i+1} = B_i + \frac{y_i y_i^T}{y_i^T d_i} - \frac{B_i d_i (B_i d_i)^T}{d_i^T B_i d_i}$$

kde $y_i = g_{i+1} - g_i$ a $d_i = x_{i+1} - x_i$. Obvykle $\eta_0 = 0.01$ pro metody spádových směrů a $\eta_0 = 0.0001$ pro metody s lokálně omezeným krokem.

Numerické porovnání: 30 testovacích problémů (2-12 proměnných)

Metoda	IT - IF - IG	selhání	čas
S - GN	1917 - 2974 - 2974	3	9.67
S - VM	1543 - 3256 - 3256	2	5.71
S - GN+VM	635 - 1037 - 1037	-	4.12
T - GN	605 - 748 - 634	-	2.64
T - VM	2155 - 2542 - 2183	2	7.58
T - GN+VM	536 - 664 - 565	-	2.58

4. Metody pro rozsáhlé řídké úlohy

Rozsáhlé úlohy nemůžeme řešit metodami, které vyžadují uchovávání velkých hustých matic (N, VM). Nejčastěji se pro tento účel používají některé speciální metody:

1. Metody s proměnnou metrikou s omezenou pamětí.
2. Diferenční verze nepřesné Newtonovy metody.

3. Metody pro řídké úlohy (N, VM)

4. Metody pro separovatelné úlohy (N, VM)

4.1. Metody s proměnnou metrikou s omezenou pamětí

Metody s proměnnou metrikou s omezenou pamětí jsou založeny na použití omezeného počtu kroků metody BFGS.

Lemma 43 Aktualizace získaná metodou BFGS se dá zapsat ve tvaru

$$H_+ = \gamma V^T H V + \frac{\rho}{b} d d^T$$

kde

$$V = I - \frac{1}{b} y d^T$$

přičemž $y = g_+ - g$, $d = x_+ - x$ a $a = y^T H y$, $b = y^T d$.

Důkaz Roznásobením dokazovaného vztahu dostaneme

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

Tentýž výsledek získáme úpravou vztahu

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

odpovídajícího metodě BFGS.

Definice 24 Řekneme, že základní optimalizační metoda je m-krokovou metodou BFGS s omezenou pamětí, jestliže

$$s_i = -H_i^i g_i$$

kde $H_{i-m}^i = I$, a

$$H_{j+1}^i = \gamma_j^i V_j^T H_j^i V_j + \frac{\rho_j}{b_j} d_j d_j^T$$

pro $i - m \le j \le i - 1$. Přitom $\gamma_{i-m}^i = b_i / a_i$ a $\gamma_j^i = 1$ pro $i - m < j \le i - 1$.

Lemma 44 Pro m-krokovou metodu BFGS s omezenou pamětí platí

$$H_{j+1}^i = \frac{b_i}{a_i} \left(\prod_{k=i-m}^{j} V_k \right)^T \left(\prod_{k=i-m}^{j} V_k \right) + \sum_{l=i-m}^{j} \frac{\rho_l}{b_l} \left(\prod_{k=l+1}^{j} V_k \right)^T d_l d_l^T \left(\prod_{k=l+1}^{j} V_k \right)$$

Důkaz (Indukcí) pro $j = i - m$ to platí (stačí dosadit $\gamma_{i-m}^i = b_i/a_i$ a $H_{i-m}^i = I$ do vztahu pro BFGS). Indukční krok:

$$\begin{aligned} H_{j+1}^i &= V_j^T H_j^i V_j + \frac{\rho_j}{b_j} d_j d_j^T = \frac{b_i}{a_i} V_j^T \left(\prod_{k=i-m}^{j-1} V_k \right)^T \left(\prod_{k=i-m}^{j-1} V_k \right) + \\ &+ \sum_{l=i-m}^{j-1} \frac{\rho_l}{b_l} V_j^T \left(\prod_{k=l+1}^{j-1} V_k \right) d_l d_l^T \left(\prod_{k=l+1}^{j-1} V_k \right) V_j + \frac{\rho_j}{b_j} d_j d_j^T = \\ &= \frac{b_i}{a_i} \left(\prod_{k=i-m}^{j} V_k \right)^T \left(\prod_{k=i-m}^{j} V_k \right) + \sum_{l=i-m}^{j} \frac{\rho_l}{b_l} \left(\prod_{k=l+1}^{j} V_k \right)^T d_l d_l^T \left(\prod_{k=l+1}^{j} V_k \right) \end{aligned}$$

Strangova formule. Vektor $s_i = -H_i^i g_i$ lze spočítat takto. Nejprve počítáme zpětnou rekurzí vektory

$$u_j = - \left(\prod_{k=j}^{i-1} V_k \right) g_i$$

pro $i - m \le j \le i - 1$. Protože

$$u_j = V_j u_{j+1} = \left(I - \frac{1}{b_j} y_j d_j^T \right) u_{j+1} = u_{j+1} - \frac{d_j^T u_{j+1}}{b_j} y_j$$

pro $i - m \le j \le i - 1$, kde $u_i = -g_i$, můžeme psát

$$u_i = -g_i$$

a

$$\sigma_j = d_j^T u_{j+1} / b_j$$

$$u_j = u_{j+1} - \sigma_j y_j$$

pro $i - m \le j \le i - 1$. Potom počítáme přímou rekurzí vektory

$$v_{j+1} = \frac{b_i}{a_i} \left(\prod_{k=i-m}^{j} V_k \right)^T u_{i-m} + \sum_{l=i-m}^{j} \frac{\rho_l}{b_l} \left(\prod_{k=l+1}^{j} V_k \right)^T d_l d_l^T u_{l+1}$$

pro $i - m \le j \le i - 1$. Protože

$$v_{j+1} = V_j^T v_j + \frac{\rho_j}{b_j} d_j d_j^T u_{j+1} = \left(I - \frac{1}{b_j} d_j y_j^T \right) v_j + \rho_j \sigma_j d_j = v_j + (\rho_j \sigma_j - y_j^T v_j / b_j) d_j$$

kde $v_{i-m} = (b_i/a_i)u_{i-m}$, můžeme psát

$$v_{i-m} = (b_i/a_i)u_{i-m}$$

a

$$v_{j+1} = v_j + (\rho_j\sigma_j - y_j^T v_j)d_j$$

pro $i - m \le j \le i - 1$. Nakonec položíme $s_i = v_i$.

Poznámka 39 Je třeba uchovávat pouze čísla σ_j, $i - m \le j \le i - 1$. Všechny vektory u_j, v_j, $i - m \le j \le i$ mohou být uloženy na témže místě v paměti počítače. Celkem tedy potřebujeme uložit $2m + 3$ vektorů (d_j, y_j, $i - m \le j \le i - 1$ a 3 vektory pro základní optimalizační metodu) a použijeme $O(mn)$ numerických operací.

Tvrzení 45 (Kvadratické ukončení). Nechť x_i, $i \in N$ je posloupnost generovaná m-krokovou metodou BFGS s omezenou pamětí s přesným výběrem délky kroku (platí $s_i^T g_{i+1} = 0\ \forall i \in N$) aplikovaná na ryze konvexní kvadratickou funkci (Q). Pak existuje index $k \le n$ tak, že $g_{k+1} = 0$ a $x_{k+1} = x^*$.

4.2. Diferenční verze nepřesné Newtonovy metody

Diferenční verze nepřesné Newtonovy metody jsou v podstatě nepřesné metody s lokálně omezeným krokem (algoritmus 5), kde se nepoužívá matice $B = G$ a násobení $q = Bp = Gp$ se nahražuje numerickým derivováním

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

kde δ je malá diference ($\delta = \sqrt{\varepsilon_M}$). Jinak se algoritmus 5 nemění. Jestliže výpočet gradientu vyžaduje $O(n)$ operací, je tento způsob úspornější než násobení matice vektorem (obecně $O(n^2)$ operací). Navíc není třeba počítat druhé derivace.

4.3. Diferenční verze Newtonovy metody pro řídké úlohy

Je-li Hessova matice řídká, není obvykle k jejímu výpočtu zapotřebí n diferencí (n gradientů).

Ukázka: (pásová struktura)

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

Nechť

$$v_1 = [1, 0, 0, 1, 0]^T$$
$$v_2 = [0, 1, 0, 0, 1]^T$$
$$v_3 = [0, 0, 1, 0, 0]^T$$

Pak platí

$$Gv_1 = [G_{11}, G_{21}, G_{34}, G_{44}, G_{54}]^T$$
$$Gv_2 = [G_{12}, G_{22}, G_{32}, G_{45}, G_{55}]^T$$
$$Gv_3 = [0, G_{23}, G_{33}, G_{43}, 0]^T$$

a můžeme použít diferenční vzorce

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

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

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

k přibližnému výpočtu všech prvků matice G. Je-li $G(x)$ pasová s m pásy, stačí použít m diferencí. Má-li matice G obecnou strukturu dostaneme kombinatorický problém ekvivalentní barvení jistého grafu. Známe-li matici G, můžeme použít nepřesnou metodu s lokálně omezeným krokem (algoritmus 5) nebo metodu s optimálním lokálně omezeným krokem s řídkým Gillovým-Murrayovým rozkladem.

4.4. Metody s proměnnou metrikou pro řídké úlohy

Je třeba zachovat řídkou strukturu a symetrii a splnit kvazinewtonovskou podmínku. Těmto požadavkům vyhovuje Tointova metoda. Tointova metoda je v hustém případě ekvivalentní metodě PSB, která není příliš efektivní. Navíc je třeba řešit dodatečnou soustavu lineárních rovnic (Tointův systém). Vhodnější způsob je zachovat řídkou strukturu bez symetrie a splnit kvazinewtonovskou podmínku a výsledek symetrizovat. Tím dostaneme Marwillovu metodu. Označme P_i, $1 \leq i \leq n$, diagonální matice takové, že

$$e_j^T P_i e_j = 0 \quad \text{pokud} \quad B_{ij} = 0$$
$$e_j^T P_i e_j \neq 0 \quad \text{pokud} \quad B_{ij} \neq 0$$

Definujme

$$U_+ = B + \sum_{i=1}^{n} (P_i d)^T P_i d)^- e_i^T (y - Bd) e_i (P_i d)^T$$

(symbol $-$ značí pseudoinverzi). Matice U_+ je obecně nesymetrická, má stejnou řídkou strukturu jako matice B a platí

$$U_+d = Bd+\sum_{i=1}^{n}(P_id)^TP_id)^-e_i^T(y-Bd)e_i(P_id)^TP_id = Bd+\sum_{i=1}^{n}e_i^T(y-Bd)e_i = Bd+y-Bd = y$$

(kvazinewtonovská podmínka). Matice B_+ se pak získá symetrizací

$$B_+ = \frac{1}{2}(U_+ + (U_+)^T)$$

(tím se může opět porušit kvazinewtonovská podmínka). Marwilova metoda je mnohem jednodušší než Tointova metoda. Není třeba řešit dodatečnou soustavu rovnic a numericky je toto metoda poměrně efektivní, zejména ve spojení s nepřesnými metodami (spádových směrů nebo s lokálně omezeným krokem).

4.5. Diferenční verze Newtonovy metody pro separovatelné úlohy

Předpokládejme, že platí

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

kde $m = O(n)$ a kde každá z funkcí $f_k : R^n \to R$, $1 \le k \le m$ závisí na $n_k = O(1)$ proměnných. Označme $\hat{g}_k(x) \in R^{n_k}$ spakovaný gradient funkce f_k (obsahuje pouze nenulové prvky z $g_k(x) \in R^n$) a $\hat{B}_k(x) \in R^{n_k \times n_k}$ spakovanou Hessovu matici funkce f_k). Pak k sestrojení gradientu a Hessovy matice

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

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

stačí určovat a uchovávat pouze spakované gradienty a spakované Hessovy matice. K výpočtu spakované Hessovy matice $G_k(x)$ můžeme použít diferenční vzorce

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

kde $\hat{e}_j$, $1 \le j \le n_k$ jsou sloupce jednotkové matice řádu n_k.

K určení spakované Hessovy matice $\hat{G}_k(x)$ je tedy zapotřebí $n_k^2 = O(1)$ operací takže k určení $G(x)$ je zapotřebí $mO(1) = O(n)$ operací.

4.6. Metody s proměnnou metrikou pro separovatelné úlohy

Spakované Hessovy matice $\hat{G}_k(x)$ se aproximují maticemi $\hat{B}_k$, které se určují pomocí metod s proměnnou metrikou

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

kde $\hat{y}_k = \hat{g}_k^+ - \hat{g}_k$ a $\hat{d}_k$ je spakovaný vektor, který obsahuje pouze ty prvky vektoru d, které odpovídají prvkům spakovaného gradientu $\hat{g}_k$. Dále $\hat{b}_k = \hat{y}_k^T\hat{d}_k$, $\hat{c}_k = \hat{d}_k^T\hat{B}_k\hat{d}_k$ a $\hat{\gamma}_k$, $\hat{\beta}_k$ jsou volné parametry (věta 26). Parametr $\hat{\beta}_k$ slouží ke škálování podobně jako u klasických metod s proměnnou metrikou.

Numerické výsledky: 10 testovacích funkcí se 100 proměnnými:

1) Řídké úlohy:

Metoda	NIT - NFV - NFG	čas
CG	2523 - 5071 - 5071	11.97
5-BFGS	2080 - 2257 - 2257	14.12
Nepřesná Newtonova (dif. verze)	795 - 913 - 6494	12.97
Řídká Newtonova (dif. verze) + CG	547 - 596 - 2548	13.57
Řídká Newtonova (dif. verze) + GM	417 - 445 - 1892	11.98
Řídká BFGS + CG	1352 - 2020 - 2020	21.86
Řídká BFGS + GM	2457 - 4862 - 4862	62.39
Hustá BFGS	1574 - 1716 - 1716	39.22

2) Separovatelné úlohy:

Metoda	NIT - NFV - NFG	čas
CG	2796 - 5638 - 5638	26.42
5-BFGS	2389 - 2586 - 2586	21.81
Nepřesná Newtonova (dif. verze)	774 - 903 - 6359	22.46
Separovatelná Newtonova (dif. verze) + CG	538 - 599 - 2213	31.25
Separovatelná Newtonova (dif. verze) + GM	416 - 446 - 1635	25.71
Separovatelná BFGS + CG	936 - 1048 - 1048	37.97
Separovatelná BFGS + GM	765 - 892 - 892	21.36

5. Optimalizace dynamických systémů

Uvažujeme úlohu s účelovou funkcí

$$F(x) = \int_{t_0}^{t_1} f_A(y(x,t),t)\,dt + f_T(y(x,t_1)) \tag{O}$$

kde

$$\frac{dy(x,t)}{dt} = f_S(x,y(x,t),t), \qquad y(x,t_0) = f_I(x)\;. \tag{D}$$

Přitom $x \in R^n$, $y : R^n \times [t_0,t_1] \to R^{n_S}$, $F : R^n \to R$, $f_A : R^{n_S} \times [t_0,t_1] \to R$, $f_T : R^{n_S} \to R$, $f_S : R^n \times R^{n_S} \times [t_0,t_1] \to R^{n_S}$, $f_I : R^n \to R^{n_S}$. Odstranění integrálu:

$$F(x) = F_A(x,t_1) + f_T(y(x,t_1)) \tag{$\overline{\mathrm{O}}$}$$

kde

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

$$\frac{dF_A(x,t)}{dt} = f_A(y,t), \qquad F_A(x,t_0) = 0 \tag{$\overline{\mathrm{D}}$}$$

Celkem se řeší $n_S + 1$ diferenciálních rovnic v přímém směru. Stačí spočítat hodnoty na konci intervalu. Úloha $(\overline{\mathrm{O}})+(\overline{\mathrm{D}})$ se řeší pomocí gradientních optimalizačních metod (CG, VM, N) proto je třeba počítat derivace účelové funkce. Předpoklady:

(A1) Existuje spojité řešení systému (D) na intervalu $[t_0,t_1]$ kdykoliv $x \in X \subset R^n$.

(A2) Funkce f_A, f_T, f_S, f_I jsou dvakrát spojitě diferencovatelné na $X \subset R^n$.

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

5.1. Přímý výpočet gradientu

Označme $u(x,t) = dy(x,t)/dx$, takže $u : R^n \times [t_0,t_1] \to R^{n_S \times n}$. Derivováním $(\overline{\mathrm{O}})$ a $(\overline{\mathrm{D}})$ dostaneme

$$g^T(x) = g_A^T(x,t_1) + \frac{\partial f_T(y(x,t_1))}{\partial y} u(x,t_1) \tag{$\overline{\mathrm{O1}}$}$$

kde

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

$$\frac{dg_A^T(x,t)}{dt} = \frac{\partial f_A(y,t)}{\partial y} u(x,t), \qquad g_A^T(x,t_0) = 0 \tag{$\overline{\mathrm{D1}}$}$$

Přitom $g^T(x) = dF(x)/dx$, $g_A^T(x,t) = dF_A(x,t)/dx$. Celkem se řeší $(n_S+1)(n+1)$ diferenciálních rovnic v přímém směru.

5.2. Zpětný výpočet gradientů

Nechť $p(t)$ je libovolná funkce taková, že $p : [t_0, t_1] \to R^{n_S}$ a nechť $y(x,t)$ je řešení systému (D), takže $f_S(x,y,t) - dy(x,t)/dt = 0$ pro $t \in [t_0, t_1]$. Použijeme-li (O), můžeme psát

$$F(x) = \int_{t_0}^{t_1} \{f_A(y,t) + p^T(t)(f_S(x,y,t) - \frac{dy(x,t)}{dt})\}\, dt + f_T(y(x,t_1))$$

a použitím pravidla integrování per partes dostaneme

$$\begin{aligned} F(x) &= \int_{t_0}^{t_1} \{f_A(y,t) + p^T(t) f_S(x,y,t) + \frac{dp^T(t)}{dt} y(x,t)\}\, dt \\ &\quad + p^T(t_0) y(x,t_0) - p^T(t_1) y(x,t_1) + f_T(y(x,t_1)) \,. \end{aligned}$$

Nyní můžeme $F(x)$ derivovat podle x, takže

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

Zvolíme-li funkci $p(t)$ tak, aby vypadly všechny členy s $dy(x,t)/dt$, čili tak, že

$$-\frac{dp(x,t)}{dt} = (\frac{\partial f_S(x,y,t)}{\partial y})^T p(x,t) + (\frac{\partial f_A(y,t)}{\partial y})^T, \qquad p(x,t_1) = (\frac{\partial f_T(y(x,t_1))}{\partial y})^T$$

pak platí

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

Dohromady to lze zapsat takto

$$g(x) = \tilde{g}_A(x,t_0) + (\frac{df_I(x)}{dx})^T p(x,t_0) \tag{$\overline{\mathrm{O2}}$}$$

kde

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

a

$$-\frac{d\tilde{g}_A(x,t)}{dt} = (\frac{\partial f_S(x,y,t)}{\partial x})^T p(t), \quad \tilde{g}_A(x,t_1) = 0 \tag{$\overline{\mathrm{D2}}$}$$

Celkem se řeší $n_S + 1$ diferenciálních rovnic v přímém směru a $2n_S + n$ diferenciálních rovnic ve zpětném směru.

5.3. Přímý výpočet Hessovy matice

Označme $v(x,t) = du(x,t)/dx = d^2y(x,t)/dx^2$, takže $v : R^n \times [t_0, t_1] \to R^{n_S \times n \times n}$. Derivováním $(\overline{O1})$ a $(\overline{D1})$ dostaneme

$$G(x) = G_A(x,t_1) + u^T(x,t_1)\frac{\partial^2 f_T(y(x,t_1))}{\partial y^2}u(x,t_1) + \frac{\partial f_T(y(x,t_1))}{\partial y}v(x,t_1) \tag{$\overline{O3}$}$$

kde

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

$$v(x,t_0) = \frac{d^2 f_I(x)}{dx^2}$$

$$\frac{dG_A(x,t)}{dt} = u^T(x,t)\frac{\partial^2 f_A(y,t)}{\partial y^2}u(x,t) + \frac{\partial f_A(y,t)}{\partial y}v(x,t), \quad G_A(x,t_0) = 0 \tag{$\overline{D3}$}$$

Přitom $G(x) = d^2F(x)/dx^2$ a $G_A(x) = d^2 f_A(x,t)dx^2$. Celkem se řeší $(n_S+1)(n^2+n+1)$ diferenciálních rovnic v přímém směru.

5.4. Přímá aproximace Hessovy matice (součet čtverců)

$$f_A(y,t) = \frac{1}{2}(y(x,t) - z(t))^T W(t)(y(x,t) - z(t))$$

$$\frac{\partial f_A(y,t)}{\partial y} = W(t)(y(x,t) - z(t)), \qquad \frac{\partial^2 f_A(y,t)}{\partial y^2} = W(t)$$

a podobně

$$f_T(y(x,t_1)) = \frac{1}{2}(y(x,t_1) - z(t_1))^T W_1(y(x,t_1) - z(t_1))$$

$$\frac{\partial f_T(y(x,t_1))}{\partial y} = W_1(y(x,t_1) - z(t_1)), \qquad \frac{\partial^2 f_T(y(x,t_1))}{\partial y^2} = W_1$$

Přitom $z : [t_0, t_1] \to R^{n_S}$, $W : [t_0, t_1] \to R^{n_S \times n_S}$ (SPD) (obecně $W_1 \neq W(t_1)$). Jestliže $F(x) \to 0$, pak nutně $y(x,t) \to z(t)$ takže $\partial f_A(y(x,t),t)/\partial y \to 0$ a $\partial f_T(y(x,t_1))/\partial y \to 0$. Můžeme tedy zanedbat tyto členy v $(\overline{O3})$ a $(\overline{\mathrm{D}3})$. Dostaneme tak

$$G(x) \approx B(x) = B_A(x,t_1) + u^T(x,t_1) W_1 u(x,t_1) \tag{$\overline{O4}$}$$

kde

$$\frac{dB_A(x,t)}{dt} = u^T(x,t) W(t) u(x,t), \qquad B_A(x,t_0) = 0 \tag{$\overline{\mathrm{D}4}$}$$

Celkem se řeší $(n_S + 1)(n + 1) + n^2$ diferenciálních rovnic v přímém směru.