2. Kvázinewtonovské metódy

2.1. Optimalizačné metódy vo všeobecnosti

Optimalizačné metódy na hľadanie minima funkcií majú iteračný charakter. Na začiatku zvolíme štartovací bod $\mathbf{x}_0$. Nové približné riešenia konštruujeme podľa vzťahu

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{s}_k \,,$$

kde $\mathbf{s}_k$ je smer posunu a λ_k je veľkosť kroku.

To znamená, že do iterácie vstupuje približné riešenie $\mathbf{x}_k$ z predchádzajúcej iterácie. V tomto bode zvolíme smer $\mathbf{s}_k$ ako n - rozmerný nenulový vektor a v tomto smere zvolíme veľkosť posunu – krok λ_k.

Sú rôzne možnosti voľby kroku. V tejto práci budem používať výpočtovo najnáročnejšiu voľbu kroku – *optimálny krok*. Optimálny krok dostaneme ako riešenie nasledovnej úlohy:

$$\text{Min}\left\{\varphi(\lambda) = f\left(\mathbf{x}_k + \lambda \mathbf{s}_k\right) \;\middle|\; \lambda \in R\right\} \tag{1}$$

Riešenie tejto úlohy je podstatne jednoduchšie ako riešenie pôvodného problému (1.1) z úvodu tejto práce, pretože $\lambda \in R$. To znamená, že ide o minimalizáciu jednorozmernej úlohy na voľný extrém. Na vyriešenie úlohy (1) môžeme použiť metódu zlatého rezu alebo iné.

Smery $\mathbf{s}_k$ v tejto práci budú generované nasledovne

$$\mathbf{s}_k = -\mathbf{H}_k \nabla f(\mathbf{x}_k)\,, \tag{2}$$

kde matica $\mathbf{H}_k$ je kladne definitná. Označme $\mathbf{H}_k > 0$.

Definícia 2.1. Majme funkciu $f(\mathbf{x})$ definovanú na R^n a bod $\mathbf{x}_k \in R^n$. Smer $\mathbf{s}_k \in R^n$ nazývame *spádovým smerom* funkcie $f(\mathbf{x})$ v bode $\mathbf{x}_k$, ak existuje kladné číslo λ^* tak, že pre všetky $0 < \lambda < \lambda^*$ platí

$$f(\mathbf{x}_k + \lambda \mathbf{s}_k) < f(\mathbf{x}_k) \tag{3}$$

Lema 2.1. Smer $\mathbf{s}_k = -\mathbf{H}_k \nabla f(\mathbf{x}_k)$ je spádový.

Dôkaz: Derivovaním funkcie $\varphi(\lambda) = f(\mathbf{x}_k + \lambda \mathbf{s}_k)$ v bode $\lambda = 0$ dostávame

$$\varphi'(0) = [\nabla f(\mathbf{x}_k)]^T \mathbf{s}_k. \tag{4}$$

Z vlastností derivácie smer $\mathbf{s}_k$ bude spádový, ak

$$[\nabla f(\mathbf{x}_k)]^T \mathbf{s}_k < 0. \tag{5}$$

Dosadením (2) do (5) a využitím $\mathbf{H}_k > 0$ je lema dokázaná. ■

2.2. Kvázinewtonovské metódy

Kvázinewtonovské metódy vychádzajú z Newtonovej metódy, pričom sa snažia odstrániť jej hlavné nedostatky. Newtonova metóda je založená na aproximácii minimalizovanej funkcie Taylorovým polynómom druhého stupňa. Minimum tejto kvadratickej aproximácie v bode $\mathbf{x}_k$ určuje ďalšie približné riešenie $\mathbf{x}_{k+1}$. Iteračný vzorec má tvar:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\nabla^2 f(\mathbf{x}_k)]^{-1} \nabla f(\mathbf{x}_k) \tag{6}$$

Kvôli jednoduchosti v ďalšom gradient $\nabla f(\mathbf{x}_k)$ označíme $\mathbf{g}(\mathbf{x}_k)$ a Hessovu maticu druhých parciálnych derivácii $\nabla^2 f(\mathbf{x}_k)$ ako $\mathbf{G}(\mathbf{x}_k)$. Rozdiel dvoch po sebe nasledujúcich približných riešení ako $\mathbf{p}_k$ a rozdiel dvoch po sebe nasledujúcich gradientov ako $\mathbf{y}_k$.

Teda:

$$\mathbf{g}(\mathbf{x}_k) = \nabla f(\mathbf{x}_k) \ , \ \mathbf{g}(\mathbf{x}_k) = \mathbf{g}_k , \tag{7.a}$$

$$\mathbf{G}(\mathbf{x}_k) = \nabla^2 f(\mathbf{x}_k) \ , \ \mathbf{G}(\mathbf{x}_k) = \mathbf{G}_k , \tag{7.b}$$

$$\mathbf{p}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \tag{7.c}$$

$$\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k . \tag{7.d}$$

Smery v Newtonovej metóde sú spádové vtedy, keď sú matice $\mathbf{G}_k$ kladne definitné ($\mathbf{G}_k > 0$). To však ešte nezaručuje aby $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$, pretože veľkosť kroku λ_k je konštantne rovná jednej. Ak $\lambda^* > 0$ je najväčšie λ, pre ktoré platí (4) a $\lambda^* < 1$, tak k poklesu funkčnej hodnoty nedôjde. Táto metóda nie je globálne konvergentná. Preto je vhodné regulovať veľkosť kroku.

Modifikovaná Newtonova metóda používa rovnaký smer $\mathbf{s}_k = \mathbf{p}_k = -\mathbf{G}_k^{-1}\mathbf{g}_k$ a optimálny krok . Za predpokladu kladnej definitnosti matíc $\mathbf{G}_k^{-1}$ je táto metóda globálne konvergentná.

Vidíme, že v algoritme modifikovanej Newtonovej metódy je potrebné počítať Hessovu maticu druhých parciálnych derivácii $\mathbf{G}$. V praktických úlohách býva veľmi ťažké získať analytické vyjadrenie pre prvky Hessovej matice. Navyše v každej iterácii sa rieši sústava n lineárnych rovníc $\mathbf{G}_k \cdot \mathbf{p}_k = \mathbf{g}_k$ ($\mathbf{p}_k$ nepoznáme). Toto je spojené s veľkou výpočtovou náročnosťou.

Kvázinewtonovské metódy sa snažia odstrániť tieto nedostatky. Nepoužívajú Hessovu maticu, ale pracujú len s aproximáciou inverznej Hessovej matice, ktorá je konštruovaná na základe znalosti gradientov v každom bode. Tým odpadá aj nutnosť riešenia lineárneho systému, pretože dostaneme už inverznú Hessovu maticu a teda ďalšie priblíženie dostaneme priamo zo vzťahu (6).

Majme kvadratickú funkciu

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}})^T \mathbf{G}(\mathbf{x} - \overline{\mathbf{x}}) + f(\overline{\mathbf{x}}), \qquad \text{kde } \mathbf{G} > 0. \tag{8}$$

zrejme $\overline{\mathbf{x}} \in R^n$ je jej bodom minima.

V prípade kvadratickej funkcie (8) pre dva body $\mathbf{x}_{k+1}$ a $\mathbf{x}_k$ platí:

$$(\mathbf{g}_{k+1} - \mathbf{g}_k) = \mathbf{G}\,(\mathbf{x}_{k+1} - \mathbf{x}_k), \tag{9a}$$

$$(\mathbf{x}_{k+1} - \mathbf{x}_k) = \mathbf{G}^{-1}(\mathbf{g}_{k+1} - \mathbf{g}_k). \tag{9b}$$

Použijúc označenie (7) dostávame vzťahy:

$$\mathbf{y}_k = \mathbf{G}\,\mathbf{p}_k \tag{10a}$$

respektíve

$$\mathbf{p}_k = \mathbf{G}^{-1}\mathbf{y}_k \tag{10b}$$

Uvedené vzťahy naznačujú, že maticu druhých parciálnych derivácii **G** kvadratickej funkcie možno vypočítať na základe znalosti prvých derivácii **g**. Samozrejme, že vzťah (10) neurčuje **G** jednoznačne, ale to nám umožní dostatočne veľký počet týchto podmienok. Pre kvadratickú funkciu znalosť n gradientov $\mathbf{g}_k$ v n rôznych bodoch $\mathbf{x}_k$ obsahuje dostatok informácii na určenie **G**.

Označme k-tu aproximáciu inverznej Hessovej matice $\mathbf{H}_k$, teda

$$\mathbf{H}_k \approx \mathbf{G}^{-1}. \tag{11}$$

V Newtonovej metóde išlo o kvadratickú aproximáciu a aj v kvázinewtonovských metódach sa zachovávajú niektoré vlastnosti z minimalizácie kvadratickej funkcie. Vzťahy (10.a) respektíve (10b) sa nazývajú *kvázinewtonovská podmienka.* Ak sme v iteračnom algoritme boli v bode $\mathbf{x}_k$ s gradientom $\mathbf{g}_k$ a maticou $\mathbf{H}_k$ a určili sme nový bod $\mathbf{x}_{k+1}$, tak dostávame novú informáciu o gradiente v tomto bode $\mathbf{g}_{k+1}$. Táto nová informácia sa cez kvázinewtonovskú podmienku premietne do požiadavky na k+1 aproximáciu inverznej Hessovej matice nasledovne:

$$\mathbf{H}_{k+1}\mathbf{y}_k = \mathbf{p}_k \tag{12}$$

Vystupuje tu aj požiadavka na zachovanie starej informácie - toho, čo sme už poznali, v podmienke:

$$\mathbf{H}_{k+1}.\mathbf{w} = \mathbf{H}_k.\mathbf{w}, \tag{13}$$

kde **w** sú vektory z podpriestoru V, ktorý neobsahuje vektor $\mathbf{y}_k$.

Matica $\mathbf{H}_{k+1}$ je na základe matice $\mathbf{H}_k$ konštruovaná pomocou aditívnej korekcie:

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \Delta\,\mathbf{H}_k\,. \tag{14}$$

Korekčná matica $\Delta\,\mathbf{H}_k$ je tvorená pomocou známych vektorov $\mathbf{p}_k$ a $\mathbf{y}_k$ a matice $\mathbf{H}_k$. Zo vzťahov (12), (13) a (14) pre korekčnú maticu vyplýva:

$$\Delta\,\mathbf{H}_k.\mathbf{y}_k = \mathbf{p}_k - \mathbf{H}_k.\mathbf{y}_k\,, \tag{15a}$$

$$\Delta\,\mathbf{H}_k.\mathbf{w} = 0 \qquad \text{pre } \forall\, \mathbf{w} \in V. \tag{15b}$$

Minimalizovaná funkcia f(.) je konvexná a dvakrát spojite diferencovateľná, teda jej Hessova matica je symetrická a kladne definitná. Preto prirodzené požiadavky na matice $\mathbf{H}_k$ sú, aby boli kladne definitné a symetrické. Taktiež z hľadiska konštrukcie by tieto matice mali byť jednoduché, hodnosti jedna alebo dva.

Modelový algoritmus na riešenie úlohy (1.1) kvázinewtonovskými metódami vyzerá nasledovne:

Algoritmus 1

Štart:	$\mathbf{x}_0 \in R^n$, $\mathbf{g}_0 = \nabla f(\mathbf{x}_0)$, $\mathbf{H}_0 > 0$.
	$\varepsilon > 0$ – konštanta pre kritérium presnosti.
	Počítadlo iterácii $k = 0$.
Cyklus:	$k = 0, 1, \ldots$
1. krok	Test dosiahnutej presnosti
	Ak $\|\nabla f(\mathbf{x}_k)\| < \varepsilon$, tak STOP.
2. krok	Vypočítame spádový smer $\mathbf{s}_k = -\mathbf{H}_k g_k$.
3. krok	Vypočítame optimálny krok $\lambda_k > 0$.
4. krok	Výpočet novej aproximácie $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{s}_k$
	a nového gradientu $\mathbf{g}_{k+1} = \nabla f(\mathbf{x}_{k+1})$.
5. krok	Vypočítame vektory
	$\mathbf{p}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$,
	$\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$,
6. krok	Pomocou vektorov $\mathbf{y}_k$, $\mathbf{p}_k$ a matice $\mathbf{H}_k$ konštruujeme korekčnú maticu $\Delta\mathbf{H}_k$ vyhovujúcu vzťahom (15a, b).
7. krok	Vypočítame novú maticu $\mathbf{H}_{k+1} = \mathbf{H}_k + \Delta\mathbf{H}_k$.
8. krok	Zväčšíme k o jeden a vrátime sa na 1.krok
Výstup:	$\mathbf{x}_k$ - aproximácia bodu minima $f(\mathbf{x})$

Matica $\mathbf{H}_{k+1}$ je konštruovaná tak, aby v každej iterácii spĺňala kvázinewtonovskú podmienku (12). Informáciu z tejto podmienky získanú v k-tej iterácii si zachováva vo všetkých ďalších iteráciach, ako je to formulované v podmienke (13). Teda po n iteráciach v prípade kvadratickej funkcie máme:

$$\mathbf{G}\mathbf{p}_k = \mathbf{y}_k \qquad \text{pre } k = 0, 1, \ldots, n-1.$$

Z (10) a (11) pre $\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{n-1}$ nezávislé dostávame

$$\mathbf{H}_n\mathbf{y}_k = \mathbf{p}_k \qquad \text{per } k = 0, 1, \ldots, n-1.$$

Spojením týchto dvoch vzťahov máme

$$\mathbf{H}_n\mathbf{G}\mathbf{P} = \mathbf{P},$$

kde **P** je matica, ktorej stĺpce tvoria vektory $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{n-1}$.
(Presný dôkaz bude uvedený neskôr.)

Bolo uvedené, že na určenie Hessovej matice **G** rozmeru nxn kvadratickej funkcie vo všeobecnosti postačuje n podmienok (10a). Na určenie všetkých členov matice $\mathbf{G}^{-1}$ v kvadratickom prípade stačí n podmienok (10b), pokiaľ sú vektory $\mathbf{p}_k$ lineárne nezávislé pre $k = 0, 1, \dots, n-1$. To znamená, že tento algoritmus po n iteráciach nájde minimum kvadratickej funkcie. Touto vlastnosťou sa vyznačujú metódy združených smerov.

2.3. Združené smery a ich vlastnosti

Definícia 2.2. Nech $\mathbf{G} = \mathbf{G}^T$ je kladne definitná matica typu nxn. Smery $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$, t. j. nenulové vektory $\mathbf{s}_i \in R^n$ sa nazývajú ***G** – združené* , ak pre každú dvojicu $\mathbf{s}_i, \mathbf{s}_j$ ($i \neq j$) platí

$$\mathbf{s}_i^T \mathbf{G}\, \mathbf{s}_j = 0 \qquad (16)$$

Lema 2.2. **G** – združené smery sú lineárne nezávislé.

Veta 2.1.

Nech je iteračný proces $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{s}_k$, $k \in N$ aplikovaný na kvadratickú funkciu (8). Nech smery $\{\mathbf{s}_0, \mathbf{s}_2, \dots, \mathbf{s}_{n-1}\}$ sú **G** združené a λ_k je optimálny krok. Potom platí

$$\mathbf{s}_i^T \mathbf{g}_{j+1} = 0 \qquad \text{pre} \quad \forall\, 0 \le i \le j \le n-1, \qquad (17)$$

$$\mathbf{g}_n = 0,$$

a teda $\mathbf{x}_n = \bar{\mathbf{x}}$, t.j. iteračný proces určí minimum kvadratickej funkcie za n krokov.

<u>Dôkaz:</u> Používame optimálny krok, teda λ_k dostaneme ako riešenie úlohy (1)

$$\text{Min} \left\{ \varphi(\lambda) = f(\mathbf{x}_k + \lambda \mathbf{s}_k) \;\middle|\; \lambda \in R \right\}.$$

Pre optimálny krok musí platiť $\varphi'(\lambda_k) = 0$.

Počítajme

$$\varphi'(\lambda_k) = [\mathbf{g}(\mathbf{x}_k + \lambda_k \mathbf{s}_k)]^T.\mathbf{s}_k = \mathbf{g}^T_{k+1} \mathbf{s}_k .$$

Teda $$\mathbf{g}_{k+1}^T \mathbf{s}_k = 0. \quad (18)$$

Ďalej

podľa (10a) $$\mathbf{y}_k = \mathbf{G}\mathbf{p}_k = \lambda_k \mathbf{G}\mathbf{s}_k, \quad (19)$$

podľa (16) a (19) $$\mathbf{s}_i^T \mathbf{y}_j = \mathbf{y}_i^T \mathbf{s}_j = 0 \quad \text{pre } \forall\, i, j \in [0, n-1] \text{ a } i \neq j. \quad (20)$$

Vzťah (17) dokážeme indukciou.

I. krok (i = 0, j = 0) $\mathbf{s}_0^T \mathbf{g}_1 \overset{18}{=} 0$

Predpokladajme, že platí $\mathbf{s}_i^T \mathbf{g}_j = 0$, kde $0 \le i < j$.

II. krok $j \to j + 1$

Podľa indukčného predpokladu a vzťahov (7.d) a (20) máme

$$\mathbf{s}_i^T \mathbf{g}_{j+1} = \mathbf{s}_i^T \mathbf{g}_j + \mathbf{s}_i^T \mathbf{y}_j = 0.$$

Navyše podľa (18) $\mathbf{g}_{j+1}^T \mathbf{s}_j = 0$, teda máme (17) dokázané pre $0 \le i \le j \le n-1$.

Vektory { $\mathbf{s}_0, \mathbf{s}_2, \dots, \mathbf{s}_{n-1}$} sú podľa lemy 2.2 lineárne nezávislé. Z toho vyplýva, že $\mathbf{g}_n$ je ortogonálny k n lineárne nezávislým vektorom v R^n, a teda $\mathbf{g}_n = 0$.

Derivovaním funkcie (8) dostávame $0 = \mathbf{g}_n = \mathbf{g}(\mathbf{x}_n) = \mathbf{G}(\mathbf{x}_n - \bar{\mathbf{x}})$. Keďže $\mathbf{G} > 0$, musí platiť $\mathbf{x}_{n+1} = \bar{\mathbf{x}}$. ■

2.4. Konkrétne kvázinewtonovské formuly

Prvou známou kvázinewtonovskou formulou bola DFP formula (Davidon, Fletcher, Powell) z roku 1963 s korekčnou maticou v tvare:

$$\Delta \mathbf{H}_k = \frac{1}{\mathbf{p}_k^T \mathbf{y}_k^T} \mathbf{p}_k \mathbf{p}_k^T - \frac{1}{\mathbf{y}_k^T \mathbf{H}_k \mathbf{y}_k} \mathbf{H}_k \mathbf{y}_k \mathbf{y}_k^T \mathbf{H}_k \quad (21)$$

Vlastnosti DFP formuly:

- Ak je matica $\mathbf{H}_0$ kladne definitná, DFP formula generuje kladne definitné matice $\mathbf{H}_k$.
- Ak f(.) je kvadratická funkcia (8) , tak algoritmus je identický metóde združených smerov, t. j. generuje **G** – združené smery a konverguje najviac za n krokov.
- Ak f je ako v b) a konvergencia k riešeniu vyžaduje plných n krokov, potom n-tá aproximácia $\mathbf{H}_n = \mathbf{G}^{-1}$.

(Dôkaz viď. [1])

DFP metóda je špeciálnou formulou z Broydenovej triedy kvázinewtonovských formúl. Korekčná matica Broydenovej triedy formúl má tvar:

$$\Delta\mathbf{H}_k = \frac{\mathbf{p}_k\mathbf{p}_k^T}{\mathbf{p}_k^T\mathbf{y}_k} + \frac{\mathbf{H}_k\mathbf{y}_k\mathbf{y}_k^T\mathbf{H}_k}{\mathbf{y}_k^T\mathbf{H}_k\mathbf{y}_k} + \varphi_k(\mathbf{y}_k^T\mathbf{H}_k\mathbf{y}_k)\mathbf{v}_k\mathbf{v}_k^T \quad , \tag{22a}$$

kde

$$\mathbf{v}_k = \frac{\mathbf{p}_k}{\mathbf{p}_k^T\mathbf{y}_k} - \frac{\mathbf{H}_k\mathbf{y}_k}{\mathbf{y}_k^T\mathbf{H}_k\mathbf{y}_k} \tag{22b}$$

a φ_k je Broydenov parameter.

(DFP zodpovedá voľbe hodnoty parametra $\varphi_k = 0$.)

Broydenova formula (aj DFP) má korekčnú maticu hodnosti dva.

Do triedy formúl hodnosti jedna patrí SR1 formula. Jej tvar je

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\mathbf{r}_k\mathbf{r}_k^T}{\mathbf{r}_k^T\mathbf{y}_k}, \tag{23a}$$

kde

$$\mathbf{r}_k = \mathbf{p}_k - \mathbf{H}_k\mathbf{y}_k\,. \tag{23b}$$

Zovšeobecnením všetkých spomenutých formúl je Orenova trieda formúl, ktorej je venovaná nasledujúca kapitola. Oren aproximáciu Hessovej matice v iterácii vylepšuje nielen aditívnou korekciou, ale pomocou nového parametra γ mení veľkosť členov aproximácie z predchádzajúcej iterácie. Maticu $\mathbf{H}_{k+1}$ konštruuje podľa vzťahu

$$\mathbf{H}_{k+1} = \gamma\mathbf{H}_k + \Delta\mathbf{H}_k\,.$$