5 Newtonova metóda

5.1 Bikvadratickú funkcia. Pre $x \in \mathbb{R}^n$ a symetrické matice A a Q definujeme bikvadratickú funkciu ako

$$f(x) = \frac{1}{4}\left(x^T Q x\right)^2 + \frac{1}{2}x^T A x + b^T x.$$

Vyjadrite gradient $\nabla f(x)$ a Hessovu maticu $\nabla^2 f(x)$.

5.2 Newtonova metóda pre bikvadratickú funkciu. V súbore `newton.m` naprogramujte Newtonovu metódu pre bikvadratickú funkciu. Využite vzorce pre gradient a Hessovu maticu z predchádzajúceho príkladu. Otestujte na funkcii s

$$Q = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad A = \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -5 \end{pmatrix}$$

a so štartovacím bodom $x^0 = (-2, -5)^T$.

5.3 Experimenty s podmienenosťou matíc. Otestujte program z predchádzajúceho príkladu na bikvadratickej funkcii s parametrami

(a) $Q = \begin{pmatrix} 2 & 2 \\ 2 & 1.9 \end{pmatrix}, \quad A = \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -5 \end{pmatrix},$

(b) $Q = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad A = \begin{pmatrix} 7 & 6.9 \\ 6.9 & 7 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -5 \end{pmatrix},$

(c) $Q = \begin{pmatrix} 2 & 2 \\ 2 & 1.9 \end{pmatrix}, \quad A = \begin{pmatrix} 7 & 6.9 \\ 6.9 & 7 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -5 \end{pmatrix}.$

5.4 Newtonova metóda pre všeobecnú funkciu. Upravte program z predchádzajúceho príkladu pre všeobecnú konvexnú funkciu s explicitným zadaním potrebných derivácií. Otestujte na niektorej z funkcií

$$\begin{aligned} f_1(x_1, x_2) &= x_2^4 + (x_1 - 2)^2(5 + x_2^2) - 2x_1x_2 + 9x_1 - 5x_2, & x^0 &= (-3, 3)^T, \\ f_2(x_1, x_2) &= e^{x_1 + 3x_2 - 0.1} + e^{x_1 - 3x_2 - 0.1} + e^{-x_1 - 0.1}, & x^0 &= (-1, 1)^T. \end{aligned}$$

5.5 Numerické derivácie. Upravte program z predchádzajúceho príkladu tak, aby miesto parciálnych derivácií používal ich numerické aproximácie

$$\begin{aligned} \frac{\partial f(x)}{\partial x_i} &\approx \frac{f(x + \delta e_i) - f(x - \delta e_i)}{2\delta}, \\ \frac{\partial^2 f(x)}{\partial x_i \partial x_j} &\approx \frac{f(x + \delta e_i + \delta e_j) - f(x + \delta e_i - \delta e_j) - f(x - \delta e_i + \delta e_j) + f(x - \delta e_i - \delta e_j)}{4\delta^2}, \end{aligned}$$

kde $\delta > 0$ je parameter a e_i je vektor, ktorého i-ta zložka je jednotka a ostatné sú nuly. Porovnajte nájdené riešenia s riešeniami z predchádzajúceho príkladu.

5.6 Škaredšia funkcia. Otestujte program z predchádzajúceho príkladu na funkcii

$$f_3(x_1, x_2) = e^{x_1^2 - 2x_2 + 5} + e^{-2x_1 + x_2^2 + 2} + e^{10 \arctan(x_2)}$$

so štartovacím bodom $x^0 = (-2, 2)^T$.

Pravidlá vektorového derivovania

Vektorovému derivovaniu je venovaný Dodatok B v knihe (strana 273). Tu uvádzame len vybrané pravidlá derivovania.

Nech $f : \mathbb{R}^n \to \mathbb{R}$. Symbolom $\frac{df(x)}{dx}$ označujeme **riadkový** vektor

$$\frac{df(x)}{dx} = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}\right)$$

a symbolom $\nabla f(x) = \left(\frac{df(x)}{dx}\right)^T$ označujeme **stĺpcový** vektor gradientu funkcie f.

Derivácia skalárneho súčinu

Nech

$$u : \mathbb{R}^n \to \mathbb{R}^m, \quad v : \mathbb{R}^n \to \mathbb{R}^m, \quad f : \mathbb{R}^n \to \mathbb{R},$$

pričom $f(x) = u(x)^T v(x)$. Potom

$$\frac{df(x)}{dx} = \frac{d}{dx}\Big[u(x)^T v(x)\Big] = u(x)^T \frac{dv(x)}{dx} + v(x)^T \frac{du(x)}{dx},$$

$$\nabla f(x) = \left(\frac{df(x)}{dx}\right)^T = \nabla v(x)\ u(x) + \nabla u(x)\ v(x).$$

Derivácia zloženej funkcie

Nech

$$u : \mathbb{R}^m \to \mathbb{R}, \quad v : \mathbb{R}^n \to \mathbb{R}^m, \quad f : \mathbb{R}^n \to \mathbb{R},$$

pričom $f(x) = u\big(v(x)\big)$. Potom

$$\frac{df(x)}{dx} = \frac{du\big(v(x)\big)}{dx} = \frac{du\big(v(x)\big)}{dv}\frac{dv(x)}{dx},$$

$$\nabla f(x) = \left(\frac{df(x)}{dx}\right)^T = \nabla v(x)\ \nabla u\big(v(x)\big).$$

Domáca úloha

5.7 Hessova matica. ***[1 bod]*** Vyjadrite prvky Hessovej matice $\nabla^2 J(x)$ účelovej funkcie

$$J(x) = \sum_{i=1}^{m} (1 - v^i)\, x^T u^i + \ln\left(1 + e^{-x^T u^i}\right).$$

5.8 Newtonova metóda. ***[4 body]*** Riešte úlohu

$$\text{Min}\left\{ J(x) = \sum_{i=1}^{m} (1 - v^i)\, x^T u^i + \ln\left(1 + e^{-x^T u^i}\right) \;\middle|\; x \in \mathbb{R}^3 \right\}.$$

pomocou Newtonovej metódy

(a) s explicitne zadanými potrebnými deriváciami,

(b) s numerickými aproximáciami potrebných derivácií.

Ako kritérium optimality použite $\left\|\nabla J(x^k)\right\| \le \varepsilon = 10^{-5}$.

5.9 Približná Newtonova metóda. ***[1 bod]*** Upravte obe verzie Newtonovej metódy z predchádzajúceho príkladu tak, aby sa Hessova matica vyčísľovala len v každej N-tej iterácii. V iteráciách medzi tým sa použije posledná známa Hessova matica. Experimentujte s hodnotou N.

5.10 Porovnanie riešení. ***[1 bod]*** Porovnajte nájdené riešenia s tými, ktoré ste našli v predchádzajúcich domácich úlohách.

5.11 Porovnanie konvergencie. ***[3 body]*** Porovnajte konvergenciu rôznych verzií Newtonovej metódy (počet potrebných iterácií, trvanie výpočtu, grafické porovnanie). Zhodnoťte výsledky porovnania.