5 Newtonova metóda

Na stránke v súbore `Kody5` nájdete všetky predprogramované kódy potrebné k vypracovaniu tohto zadania. Vyplňte odpoveďový hárok vo vašom priečinku `Cvičenie5` a nahrajte sem aj svoje kódy.

5.1 Newtonova metóda. Doplňte funkciu `newton(f,gr,He,x0,eps,n,m,a)` tak, aby obsahovala Newtonovu metódu. Vstupmi tejto funkcie sú:

`f`	–	predpis účelovej funkcie f,
`gr`	–	gradient účelovej funkcie f,
`He`	–	Hessova matica účelovej funkcie f,
`x0`	–	štartovací bod Newtonovej metódy,
`eps`	–	tolerancia pre hodnotu derivácie,
`n`	–	maximálny počet povolených iterácií,
`m`	–	rozmer úlohy,
`a`	–	hranica intervalu $[-a, a]$ pre vykresľovanie funkcie.

Výstupom funkcie je iteračná postupnosť bodov aproximujúca minimum x_iter a postupnosť funkčných hodnôt v týchto bodoch f_iter.

Fungovanie tejto funkcie otestujte na funkciách v nasledujúcich úlohách.

5.2 Newtonova metóda pre bikvadratickú funkciu a experimenty s podmienenosťou matíc. V súbore `Pr2.m` otestujte funkciu `newton` z predošlého príkladu na bikvadratickej funkcii

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

kde $x \in \mathbb{R}^n$ a A a Q sú symetrické matice. Gradient $\nabla f(x)$ a Hessovu maticu $\nabla^2 f(x)$ definujte ako funkciu x pomocou A, Q a b. (Ak je to potrebné, pravidlá vektorového derivovania nájdete na poslednej strane zadania.) Za štartovací bod zvoľte bod $x_0 = (-2, -5)^T$.

Experimentujte s podmienenosťou matíc a vaše závery spíšte do odpoveďového hárku. Všímajte si aj grafický výstup funkcie. Postupne voľte za parametre bikvadratickej funkcie f tieto dáta:

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

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

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

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

5.3 Numerické derivácie. Upravte funkciu `newton_num(f,x0,eps,n,m,delta,a)` tak, aby sa pri výpočte gradientu a Hessovej matice v Newtonovej metóde namiesto parciálnych derivácií používali ich numerické aproximácie

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

kde $\delta > 0$ je parameter a e_i je vektor, ktorého i-ta zložka je jednotka a ostatné sú nuly.

Vstupmi funkcie `newton_num` sú:

`f`	–	predpis účelovej funkcie f,
`x0`	–	štartovací bod Newtonovej metódy,
`eps`	–	tolerancia pre hodnotu derivácie,
`n`	–	maximálny počet povolených iterácií,
`m`	–	rozmer úlohy,
`delta`	–	parameter vystupúci v numerickej aproximácii,
`a`	–	hranica intervalu $[-a, a]$ pre vykresľovanie funkcie.

5.4 Newtonova metóda pre všeobecnú konvexnú funkciu. V súbore `Pr4.m` otestujte fungovanie Newtonovej metódy na konvexnej funkcii

$$f_1(x_1, x_2) = x_2^4 + (x_1-2)^2(5+x_2^2) - 2x_1x_2 + 9x_1 - 5x_2, \quad x_0 = (-3,3)^T.$$

Zároveň na tejto funkcii porovnajte štandardnú Newtonovu metódu (`newton`) a Newtonovu metódu s numerickými aproximáciami parciálnych derivácií (`newton_num`).

5.5 Zložitejšia funkcia. V súbore `Pr5.m` pomocou Newtonovej metódy nájdite minimum funkcie

$$f_2(x_1, x_2) = e^{x_1^2 - 2x_2 + 5} + e^{-2x_1 + x_2^2 + 2} + e^{10 \arctan(x_2)}$$

so štartovacím bodom $x_0 = (-2, 2)^T$. Rozhodnite, či je pre vás výhodnejšie využiť funkciu `newton` alebo `newton_num`.

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

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

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

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

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

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

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