4 Gradientné metódy

4.1 Cauchyho metóda pre kvadratickú funkciu. Doplňte súbor `cauchy.m` tak, aby pomocou Cauchyho metódy hľadal minimum kvadratickej funkcie z prednášky

$$Q_1(x) = \frac{1}{2}x^T \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} x.$$

Experimentujte s parametrom $a > 0$ a so štartovacím bodom x_0. Pre hodnoty $a = 1, 3, 10$ vyskúšajte body $x_0 = (a, 0)^T, (a, 1)^T, (a, a)^T$.

4.2 Iná kvadratická funkcia. Vyskúšajte naprogramovanú metódu na inej kvadratickej funkcii, napríklad

$$Q_2(x) = \frac{1}{2}x^T \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} x + x^T \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

so štartovacím bodom $x_0 = (3, 2)^T$.

4.3 Gradientná metóda pre všeobecnú funkciu. Upravte program z predchádzajúceho príkladu pre všeobecnú konvexnú funkciu (nie nutne kvadratickú). Otestujte na niektorej z funkcií

$$\begin{aligned} f_1(x_1, x_2) &= 5x_1^2 + x_1^2x_2^2 + x_2^4 - 2x_1x_2 + 9x_1 - 5x_2, \quad & x_0 = (1, 0)^T, \\ f_2(x_1, x_2) &= e^{x_1+3x_2-0{,}1} + e^{x_1-3x_2-0{,}1} + e^{-x_1-0{,}1}, \quad & x_0 = (-1, 1)^T. \end{aligned}$$

(a) Gradientná metóda s konštantným krokom. Použite gradientnú metódu s konštantným krokom. V každej iterácii kontrolujte, či hodnota účelovej funkcie skutočne klesne. Skúste rôzne dĺžky kroku c.

(b) Gradientná metóda s približne optimálnym krokom. Použite gradientnú metódu s približne optimálnym krokom, ktorý nájdite pomocou *backtrackingu*.

(c) Cauchyho metóda. Použite gradientnú metódu s optimálnym krokom, t.j. Cauchyho metódu. Nájdenie optimálnej dĺžky kroku zodpovedá riešeniu jednorozmerného problému

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

Tento problém riešte metódou zlatého rezu, využite funkciu `zrez(f,a,b)`.

4.4 Konvergencia gradientnej metódy s konštantným krokom. Podľa prednášky gradientná metóda s konštantným krokom c konverguje pre $0 < c < 2/L$, kde L je Lipschitzovská konštanta pre gradient $\nabla f(x)$. Ukážte, že pre kvadratickú funkciu $Q(x) = \frac{1}{2}x^TGx + h^Tx$ možno za L zvoliť najväčšiu vlastnú hodnotu μ_{max} matice G, t.j. ukážte, že pre ľubovoľné x, y z $\mathbb{R}^n$ platí

$$\|\nabla Q(x) - \nabla Q(y)\| \le \mu_{max}\|x - y\|.$$

Overte na funkcii Q_2 z druhého príkladu, že metóda skutočne konverguje pre $c = 1.9/L$. Ďalej zistite, či konverguje pre $c = 2.1/L$. Hovorí o tom tvrdenie Vety 1 z prednášky?