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 $x_0 = (0, a)^T, (a, 0)^T, (1, a)^T, (a, 1)^T, (a, a)^T$.

4.2 Zložitejšia 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 príkladu 4.1 pre všeobecnú konvexnú funkciu, otestujte na niektorej z funkcií

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

(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 vo funkcii `zrez(f,a,b)`.

Domáca úloha

4.4 Gradient. ***[1 bod]*** Vyjadrite gradient $\nabla J(x)$ účelovej funkcie

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

4.5 Riešenie pomocou gradientnej metódy. ***[4 body]*** Riešte úlohu

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

pomocou grandientnej metódy

(a) s konštantným krokom,

(b) s približne optimálnym krokom, nájdeným pomocou *backtrackingu*,

(c) s optimálnym krokom, t.j. Cauchyho metódou.

Ako kritérium optimality použite $\left\| \nabla J(x^k) \right\| \leq \varepsilon = 10^{-3}$.

4.6 Porovnanie riešení. ***[1 bod]*** Porovnajte v tabuľke riešenia nájdené jednotlivými verziami gradientnej metódy a v porovnaní uveďte aj riešenie nájdené matlabovskou funkciou `fminsearch`, ktoré ste našli v prvej domácej úlohe.

4.7 Porovnanie rýchlosti. ***[1 bod]*** Porovnajte, aký dlhý čas trvalo nájdenie riešenia jednotlivým verziám gradientnej metódy a tiež, koľko iterácií potrebovali.

4.8 Grafické porovnanie konvergencie. ***[2 body]*** Označme J^* nájdené ε-presné riešenie. Pre tri verzie gradientnej metódy vykreslite do jedného obrázku grafy znázorňujúce vývoj hodnoty $J\left(x^k\right) - J^*$ s rastúcim číslom iterácie k. Na osi y použite logaritmickú mierku.

4.9 Zhodnotenie. ***[1 bod]*** Zhodnoťte výsledky porovnaní rôznych verzií gradientnej metódy.