6 Kvázinewtonovské metódy

6.1 Kvázinewtonovské metódy pre náhodnú kvadratickú funkciu. Naprogramujte kvázinewtonovskú metódu pre náhodnú kvadratickú funkciu. To znamená, že v účelovej funkcii

$$Q(x) = \frac{1}{2}x^T G x + h^T x$$

budú kladne definitná, symetrická $n \times n$ matica G a n-rozmerný vektor h náhodne vygenerované. Štartovací bod x^0 určite tiež náhodne. Použite

(a) DFP metódu s korekčnou maticou

$$\Delta H_{DFP} = \frac{pp^T}{p^T y} - \frac{Hyy^T H}{y^T H y},$$

(b) BFGS metódu s korekčnou maticou

$$\Delta H_{BFGS} = \left(1 + \frac{y^T H y}{p^T y}\right)\frac{pp^T}{p^T y} - \frac{Hyp^T + py^T H}{p^T y},$$

(c) SR1 metódu s korekčnou maticou

$$\Delta H_{SR1} = \frac{(p - Hy)(p - Hy)^T}{y^T (p - Hy)}.$$

Pozorujte, ako závisí potrebný počet iterácií od rozmeru úlohy n.

Poznámka: Pre ľubovoľnú štvorcovú maticu A je matica $AA^T + I$ kladne definitná.

6.2 Kvázinewtonovské metódy s približne optimálnym krokom. Naprogramujte kvázinewtonovskú metódu pre niektorú z funkcií

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

Použite (a) DFP metódu, (b) BFGS metódu, (c) SR1 metódu.
Dĺžku kroku voľte približne optimálne pomocou *backtrackingu*. Rozhodnite sa, či programu zadáte potrebné derivácie, alebo použijete ich numerické aproximácie. Porovnajte rýchlosť konvergencie s Newtonovou metódou, ktorú ste pre rovnaké funkcie programovali na minulom cvičení.

6.3 Kvázinewtonovské metódy s optimálnym krokom. Upravte programy z príkladu 6.2 tak, aby volili optimálnu dĺžku kroku. Využite niektorú z metód minimalizácie funkcie jednej premennej, ktoré ste programovali na prvom cvičení, alebo použite metódu zlatého rezu vo funkcii `zrez(f,a,b)`. Porovnajte rýchlosť konvergencie s programom s približne optimálnym krokom z príkladu 6.2.

Domáca úloha - do nedele 28.5.

6.4 Kvázinewtonovské metódy. *[3 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 troch spomínaných Kvázinewtonovských metód

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

(b) s optimálnym krokom.

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

6.5 Porovnanie všetkých metód. *[3 body]* Spravte tabuľku s porovnaním všetkých použitých metód v rámci domácich úloh. V tabuľke uveďte názov metódy, použité ε, nájdené riešenie, počet iterácií a čas výpočtu.

6.6 Grafické porovnanie. *[2 body]* Graficky porovnajte vývoj hodnoty $J\left(x^k\right) - J^*$

(a) pre tri kvázinewtonovské metódy s približne optimálnym aj optimálnym krokom,

(b) pre najrýchlejšiu (v zmysle počtu iterácií) kvázinewtonovskú metódu s približne optimálnym krokom, najrýchlejšiu kvázinewtonovskú metódu s optimálnym krokom, Newtonovu metódu a približnú Newtonovu metódu.

6.7 Zhodnotenie metód. *[2 body]* Na základe predchádzajúcich porovnaní zhodnoťte, ktoré z metód sa javia byť vhodné na odhad parametrov logistickej regresie a ktoré nie.

6.8 Binárna klasifikácia. *[2 body]* Použite odhadnutý model na binárnu klasifikáciu, t.j. ak pravdepodobnosť prijatia študenta je aspoň 50 %, tak ho klasifikujeme ako prijatého, inak ako neprijatého. Na dátach zo súboru *data2.csv* (majú rovnakú štruktúru ako pôvodné dáta, z ktorých sme odhadovali parametre) otestujte, ako dobre náš model klasifikuje študentov.

(a) Zistite, akú časť študentov klasifikuje správne, koľko neprijatých klasifikuje ako prijatých a koľko prijatých klasifikuje ako neprijatých.

(b) Vykreslite obrázok, ktorý graficky ukáže úspešnosť klasifikácie. Pre jednotlivých študentov vyznačte, či boli prijatí a či ich váš model klasifikuje ako prijatých. Ukážka výstupu so štyrmi študentami: