MA40050: Numerical Optimisation & Large–Scale Systems

Problem Sheet 4

Instructions: Hand in solutions to Questions 1 and 2 by **Thursday, 19th March, 12.15pm** (either in one of the lectures or to the pigeon hole in 4W Level 1).

1. (a) Prove that, for any choice of descent direction s_n, the $(n+1)$ th iterate obtained with exact line search satisfies $\nabla f(x_{n+1}) \cdot s_n = 0$.

   (b) Let us consider the quadratic function $f(x) = \frac{1}{2}x^T Ax + b^T x + c$, where $A \in \mathbb{R}^{N\times N}$ is spd, $b \in \mathbb{R}^N$ and $c \in \mathbb{R}$ and assume that s_n is a descent direction. Prove that

   $$\alpha_n = -\frac{\nabla f(x_n)^T s_n}{s_n^T A s_n}.$$

   (c) Let f be as in (b) and let $x_* - x_0$ be parallel to an eigenvector of A. Prove that Algorithm 4.2, i.e. steepest descent with $s_n = -\nabla f(x_n)$, with exact line search converges in one step.

2. **Slow Convergence of Steepest Descent.** Let $f(x) = \frac{1}{2}x^T Ax$ where

   $$A = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{with } \gamma \geq 1,$$

   and consider applying Algorithm 4.2 with **exact line searches** to this objective function.

   (a) Prove that the function $\phi(\alpha) = f(x_n - \alpha \nabla f(x_n))$ has a unique minimizer α_n and that this is the step-length taken by exact line search.

   (b) Compute α_n explicitly. Then show that, for $x_0 = (1, \gamma)^T$, we have

   $$x_n = \left(\frac{\gamma - 1}{\gamma + 1}\right)^n \begin{pmatrix} (-1)^n \\ \gamma \end{pmatrix}.$$

   (c) Deduce that $x_n \to x_*$ q-linearly with q-factor $1 - \dfrac{2}{1 + \kappa(A)}$ and that this is sharp.

3. Implement **Algorithm 4.3** with $B_n = \nabla^2 f(x_n)$ (Newton's Method with line search) in `Matlab`. Use this algorithm with backtracking line search (**Algorithm 4.1** implemented on **Problem Sheet 3**) to minimise the Rosenbrock function

   $$f(x) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$

   Print the step length used at each iteration. Try the initial point $x_0 = (1.2, 1.2)^T$ and then the more difficult point $x_0 = (-1.2, 1)^T$. What do you observe? Are there any advantages to using backtracking line search for Newton; why not fix $\alpha = 1$?