MA40050: Numerical Optimisation & Large–Scale Systems

Problem Sheet 3

Instructions: Hand in solutions to Questions 2 and 4 by **Thursday, 12 March, 12.15pm** (either in one of the lectures or to the course pigeon hole in 4W Level 1).

1. A function $f\colon \mathbb{R}^N \to \mathbb{R}$ is convex if

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y), \quad \text{for any } \alpha \in [0,1] \text{ and } x, y \in \mathbb{R}^N.$$

   Prove that any local minimiser of a convex function is a global minimiser.

2. An **exact line search algorithm** chooses

$$\alpha_n := \inf\{\alpha \ge 0 : \phi'(\alpha) = 0\} \quad \text{where} \quad \phi(\alpha) := f(x_n + \alpha s_n),$$

   i.e. the first critical point of f encountered along the half line $\{x_n + \alpha s_n : \alpha \ge 0\}$.

   A sequence $(x_k)_{k\ge 0} \subset \mathbb{R}^N$ is said to **converge r-linearly** to $x_* \in \mathbb{R}^N$ with **r-factor** σ, if there exists a sequence $(\xi_k)_{k\ge 0} \subset \mathbb{R}$ that converges q-linearly to 0 with q-factor σ and

$$|x_k - x_*| \le \xi_k\,, \quad \forall k \ge 0.$$

   Apply the method of steepest descent with exact line search to the function

$$f(x) = 2x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 2x_2, \quad x = (x_1, x_2)^T \in \mathbb{R}^2.$$

   (a) Let x_k denote the kth iterate with $x_0 = 0$. Show that $x_{2k} = (0, 1 - 5^{-k})^T$, $k \ge 0$. Sketch the successive steps taken by the method and deduce the minimiser x^*.

   (b) Show that the sequence $(x_k)_{k\ge 0}$ in Part (a) converges r-linearly to x_* and determine the r-factor σ.

3. Implement **Algorithm 4.1** (backtracking line search) and **Algorithm 4.2** (steepest descent) in `Matlab`. Use this code 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? What value of θ_{sd} gives the best convergence behaviour? Compare the behaviour against the behaviour of Newton's Method, which you implemented on **Problems Sheet 2**.

4. **The Gauss-Newton Iteration.** Let $R \in \mathrm{C}^2(\mathbb{R}^N; \mathbb{R}^M)$ where $M \geq N$. We rewrite the overdetermined nonlinear system $R(x) = 0$ as a *nonlinear least squares problem*, i.e., we aim to minimize

$$f(x) = \tfrac{1}{2}|R(x)|^2 = \tfrac{1}{2}\sum_{j=1}^{N} R_j(x)^2.$$

   (a) Assume that $R(x_*) = 0$ and that $DR(x_*)$ has full rank. Prove that x_* is a strict local minimizer of f.

   (b) (**harder**) Recall from *Problem Sheet 1, Q. 1(c)* that

$$\nabla f(x) = DR(x)^T R(x), \quad \text{and} \quad \nabla^2 f(x) = DR(x)^T DR(x) + \sum_{j=1}^{N} R^j(x) \nabla^2 R^j(x).$$

   The *Gauss-Newton Iteration* simply discards the (expensive) second term of the Hessian and uses the steps

$$x_{n+1} = x_n - (DR(x_n)^T DR(x_n))^{-1} \nabla f(x_n).$$

   Under the assumptions that both DR and $\nabla^2 f$ are Lipschitz continuous in a neighbourhood of x_*, prove that the Gauss–Newton Iteration converges q-quadratically to x_*, for any starting value x_0 sufficiently close to x_*.