MA40050: Numerical Optimisation & Large–Scale Systems

Problem Sheet 5

Instructions: Submit solutions to Question 1 and 3 by **Thursday, 2nd April, 12.15pm**
As a single file through Moodle (either PDF or ZIP of photos/scans).
Name your file '(given name) (family name) PS5'.

1. **Inexact Newton Methods:** Suppose that in Newton's Method, instead of solving the linear system $DF(x_n)(x_{n+1} - x_n) = -F(x_n)$ exactly, we employ an approximate solution technique (for example, an iterative method) and require only that

$$|DF(x_n)(x_{n+1} - x_n) + F(x_n)| \leq \rho_n |F(x_n)|,$$

where $\rho_n > 0$. Review the proof of Theorem 3.2 and, under the same assumptions on F and DF, prove the following facts on this inexact Newton iteration:

   (a) There exist $\delta > 0$ and $\bar{\rho} > 0$ such that, if $|x_0 - x_*| \leq \delta$ and $\rho_n \leq \bar{\rho}$ for all n, then $x_n \to x_*$ q-linearly.

   (b) If, in addition, $\rho_n \to 0$ then $x_n \to x_*$ q-superlinearly.

   (c) If $\rho_n \leq |F(x_n)|$ then there exists $\delta > 0$ such that, if $|x_0 - x_*| \leq \delta$, then $x_n \to x_*$ q-quadratically.

2. Implement **Algorithm 5.2** (basic trust region method) with $H_n = \nabla^2 f(x_n)$ and $x' = x_n^c$, the Cauchy point, and apply it to the Rosenbrock function

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

Print the trust region radius and the actual step length used at each iteration. As before, try the initial points $x_0 = (1.2, 1.2)^T$ and $x_0 = (-1.2, 1)^T$, and choose $\Delta_0 = 0.5$. What do you observe? Is the method sensitive to the choice of the parameters Δ_0 and ρ_{ac}?

3. Suppose that $f \in \mathrm{C}^1(\mathbb{R}^N; \mathbb{R})$ and ∇f is Lipschitz continuous with Lipschitz constant L. Further, suppose that $\nabla f(x_n) \neq 0$ and $\|H_n\| < +\infty$. Prove the following results:

   (a) If $\Delta_n \leq |\nabla f(x_n)|/\|H_n\|$, then the Cauchy point lies on the trust region boundary, i.e.,

$$\alpha_n^c = \frac{\Delta_n}{|\nabla f(x_n)|} \quad \text{and} \quad m_n(x_n) - m_n(x_n^c) \geq \tfrac{1}{2}\Delta_n |\nabla f(x_n)|.$$

   (b) If

$$\Delta_n \leq \frac{3}{4}\frac{|\nabla f(x_n)|}{L + \|H_n\|} \quad \text{and} \quad m_n(x'_{n+1}) \leq m_n(x_n^c),$$

   then $\rho_n \geq 1/4$ and x'_{n+1} in Algorithm 5.2 is accepted.

4. Let $f(x) = (1 - x_1)^2 + (x_2 - x_1^2)^2$, $x_0 = (-1, 1)^T$ and $H_0 = \nabla^2 f(x_0)$. Apply one iteration of the dogleg method to this function to compute x_1' for $\Delta_0 = 0.2,\ 1$ and 5 . What is ρ_0 in each case and what values of Δ_1 and x_1 does Algorithm 5.2 produce with $\rho_{ac} = 1/8$.

You may assume the following result without proof: Let $\nabla f(x_n)^T H_n \nabla f(x_n) > 0$, let H_n be invertible and define $\gamma := (x_n^N - x_n^U) \cdot (x_n - x_n^U)$. If $\lvert x_n^U - x_n \rvert \le \Delta_n \le \lvert x_n^N - x_n \rvert$, then the dogleg path Γ_n intersects the boundary of the trust region at $x'_{n+1} = x_n^U + t_{\min}(x_n^N - x_n^U), \quad \text{where} \quad t_{\min} := \frac{\gamma + \sqrt{\gamma^2 + \lvert x_n^N - x_n^U \rvert^2 \left(\Delta_n^2 - \lvert x_n^U - x_n \rvert^2\right)}}{\lvert x_n^N - x_n^U \rvert^2}$ and this is the candidate point chosen by the dogleg method in that case.