MA40050: Numerical Optimisation & Large–Scale Systems

Problem Sheet 1

Instructions: Hand in solutions to Questions 2 & 3 by **Thursday, 13th February, 12.15pm** (either in one of the lectures or to the pigeon hole in 4W Level 1).

1. Let $A \in \mathbb{R}^{N\times N}$ be invertible, then its condition number is $\kappa(A) = \|A\|\|A^{-1}\|$. Suppose we wish to solve the linear system $Ax = b$, but we know only an erroneous right-hand side $\tilde{b}$. Let $A\tilde{x} = \tilde{b}$, i.e., $\tilde{x}$ is the erroneous solution. Prove the following estimate on the relative error:
$$\frac{\|x-\tilde{x}\|}{\|x\|} \leq \kappa(A)\frac{\|b-\tilde{b}\|}{\|b\|}$$

2. Find the gradients $\nabla f(x)$ and Hessians $\nabla^2 f(x)$ of the following functions $f\colon \mathbb{R}^N \to \mathbb{R}$:
   (a) $f(x) := b^T x$, where $b \in \mathbb{R}^N$;
   (b) $f(x) := (1-x_1)^2 + 100(x_2 - x_1^2)^2$, $N = 2$ (Rosenbrock function);
   (c) $f(x) := F(x)^T F(x)$, where $F(x) = [F_1(x), \ldots, F_M(x)]^T$. Use the matrix $J(x)$, the $M \times N$ Jacobian matrix of $F(x)$ with (i,j) th element $J_{ij} = \partial F_i/\partial x_j$.

3. Let U be an open convex set in $\mathbb{R}^N$. Prove the following Taylor formulae:
   (a) If $F \in \mathrm{C}^1(U;\mathbb{R}^M)$ then
$$F(x+h) = F(x) + \nabla F(x)^T h + \int_0^1 \big(\nabla F(x+th) - \nabla F(x)\big)^T h \,\mathrm{d}t.$$
   (b) If $f \in \mathrm{C}^2(U;\mathbb{R})$ then
$$f(x+h) = f(x) + \nabla f(x)\cdot h + \tfrac{1}{2}h^T\nabla^2 f(x)h + h^T\Big[\int_0^1 (1-t)(\nabla^2 f(x+th) - \nabla^2 f(x))\,\mathrm{d}t\Big]h.$$
   [**Hint:** Define $\varphi(t) = f(x+th)$ and prove and then use the one-dimensional analogues using the Fundamental Theorem of Calculus for φ and φ'.]

4. Prove **Theorem 2.7**, the **Contraction Mapping Theorem.**
   [**Hint:** You may follow the following steps (mostly revision):
   - Show that $\forall k > l \geq 0$, $\|x^k - x^l\| \leq \dfrac{\alpha^l}{1-\alpha}\|x^1 - x^0\|.$
   - Deduce that $\{x^k\}$ is a Cauchy sequence in $\bar{B}_r(x_0)$ and hence convergent to some $x_* \in \bar{B}_r(\boldsymbol{x}_0)$.
   - Show that $\|x_* - G(x_*)\| \leq \|x_* - x^{k+1}\| + \alpha\|x_* - x^k\|$ and deduce that $G(x_*) = x_*$.
   - Show that x_* is the unique fixed point of G in $\bar{B}_r(x_0)$.]