C6.2/B2. Continuous Optimization

Problem Sheet 2

Please hand-in for marking Problems 3, 4 and 5. The other problems are optional.

1. Consider that on each iteration of a generic linesearch method (GLM) for minimizing a function f, the stepsize is required to satisfy the Armijo condition of sufficient decrease

$$f(x^k + \alpha^k s^k) \leq f(x^k) + \beta \alpha^k \nabla f(x^k)^\top s^k, \tag{1}$$

for some constant $\beta \in (0, 1)$ independent of k. Show that there exists $\overline{\alpha}^k > 0$ such that (1) is satisfied for all $\alpha \in [0, \overline{\alpha}^k]$. Furthermore, show that the backtracking-Armijo (bArmijo) linesearch algorithm terminates in a finite number of iterations $i \geq 0$ with $\alpha^k := \alpha_{(i)} \geq \min\{\alpha_{(0)}, \tau\overline{\alpha}^k\}$.

2. Consider that on each iteration of a linesearch method for minimizing a function f, the stepsize is required to satisfy the Armijo condition of sufficient decrease (1) above. Let us assume that we want to find α^k to satisfy the Armijo condition (not by backtracking as described in lectures), but by interpolation (similar to Problem 5, Problem Sheet 1). In particular, assume that $\alpha_0 > 0$ is the initial guess of α^k and that it does not satisfy (1).

   (i) Using that s^k is a descent direction, justify why the interval $(0, \alpha_0)$ must contain steps that satisfy (1).

   (ii) Construct a quadratic polynomial $q(\alpha)$ that interpolates $\Phi(\alpha) := f(x^k + \alpha s^k)$ at $\Phi(0)$, $\Phi'(0)$ and $\Phi(\alpha_0)$. Find the expression of q based on the interpolation conditions.

   (iii) Let the new trial stepsize α_1 be the (global) minimizer of $q(\alpha)$ over $\alpha \in R$. Show that this minimizer is well-defined and it satisfies $\alpha_1 < \alpha_0/[2(1-\beta)]$. (*Comment: Since β is chosen to be small in practice, the previous strict inequality indicates that α_1 cannot be much greater than $0.5\alpha_0$ (though it can be smaller). This gives us an idea about the size of the new stepsize.*)

3. Consider the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^2} \; f(x), \tag{2}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is the function

$$f(x) = \frac{1}{2}(a x_1^2 + x_2^2), \quad \text{where } x = (x_1 \;\; x_2)^T \in \mathbb{R}^2, \tag{3}$$

and $a > 0$ is a constant.

   (i) Show that the Hessian $\nabla^2 f(x)$ is positive definite for all $x \in \mathbb{R}^2$. Find the global minimizer x^* of $f(x)$ over $x \in \mathbb{R}^2$ and show that it is the unique (local and global) minimizer of f over $\mathbb{R}^2$.

   (ii) Assume the steepest descent method with exact linesearches is used to solve (2), starting at $x^0 = (1 \;\; a)^T$. Let $\{x^k\}$ be the generated iterates. Show that

$$x^k = \left(\frac{a-1}{a+1}\right)^k \begin{pmatrix} (-1)^k \\ a \end{pmatrix}, \quad \forall k \geq 0. \tag{4}$$

   (iii) Show that the sequence $\{x^k\}$ converges linearly to x^* as $k \to \infty$, where x^k is given in (4) and x^*, in (i). Find the (precise) convergence factor. For any given $\epsilon \in (0, 1)$, find the number of iterations required to achieve $\|x^k - x^*\| \leq \epsilon$.

(iv) Assume that an inexact linesearch is used with the steepest descent method, so that the stepsize $\alpha^{k,ine}$ is now chosen as the largest step $\alpha > 0$ such that the following sufficient decrease (Armijo) condition holds

$$f(x^k + \alpha s^k) \leq f(x^k) + \beta\alpha \nabla f(x^k)^T s^k,$$

where s^k is the steepest descent direction, $\nabla f(x^k)$ denotes the gradient of f at x^k and $\beta \in (0,1)$. Show that, if $\beta \leq \frac{1}{2}$, then $\alpha^k \leq \alpha^{k,ine}$, where α^k is the stepsize determined by exact linesearch from x^k along s^k (see (ii)). Is this property true when f is a general convex quadratic function? Justify your answer.

(v) Assume that we make the following change of coordinates

$$y = \begin{pmatrix} a^{\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix} x.$$

Write the function f in the new coordinates, i.e., find $\overline{f}(y) = f(x(y))$. Describe what you observe if the problem is solved in this new coordinates. In other words, let the steepest descent method with exact linesearches be applied to $\overline{f}(y)$, starting from y^0 corresponding to x^0 under the change of variables, and deduce how many iterations it takes to solve the problem in the new coordinates. What property of the steepest descent method does this example illustrate?

4. Apply Newton's method (without linesearch) to minimizing the function of one variable,

$$f(x) = \frac{11}{546}x^6 - \frac{38}{364}x^4 + \frac{1}{2}x^2, \quad x \in \Re,$$

starting from $x^0 = 1.01$ and let $\{x^k\}$ be the generated sequence of iterates.

Show that the Hessian matrix $\nabla^2 f(x)$ is positive definite for all x, and that the sequence $\{f(x^k)\}$ is monotonically decreasing. Prove that the limit points of the sequence of iterates $\{x^k\}$ are $+1$ and -1 as $k \to \infty$, and that $\nabla f(\pm 1) \neq 0$. (*hint*: think of the Newton iterate x^{k+1} as a continuous function of x^k; alternatively, you may consider showing numerical results.)

5. Newton's method (without linesearch) is used to minimize the function of one variable

$$f(x) = x^6 - 14x^4 + 49x^2 - 36, \quad x \in \Re.$$

What is the order of convergence in the two cases when (a) the starting point x^0 is close to $\sqrt{7}$ and (b) x^0 is close to 0. Find a nonzero value of x^0 such that $x^{k+1} = -x^k$ on every iteration.

6. **(to be covered in problem class/revision)** Given a symmetric $n \times n$ matrix M, show that

$$\lambda_{\min}(M) \leq \frac{s^T M s}{\|s\|^2} \leq \lambda_{\max}(M), \quad \text{for all } s \in \Re^n,$$

where $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ are the smallest and largest eigenvalues of M, respectively.

Let $f : \Re^n \longrightarrow \Re$ be twice continuously differentiable. Then the gradient ∇f is Lipschitz continuous (namely, $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ for all x, $y \in \Re^n$ and some $L > 0$) if and only if the Hessian $\nabla^2 f$ is uniformly bounded above (namely, $\|\nabla^2 f(x)\| \leq L$ for all $x \in \Re^n$).

7. (**non-examinable**) Let $f \in \mathcal{C}^2(\Re^n)$ be Lipschitz continuous. Apply damped Newton's method with bArmijo linesearch (ie, GLM with Newton's direction on each iteration and bArmijo linesearch) to minimizing f. Assume that $\beta < 0.5$ and $\alpha_{(0)} = 1$ in bArmijo linesearch. Assume that $\nabla^2 f(x^k)$ is positive definite for all k and that $x^k \to x^*$ as $k \to \infty$ with $\nabla^2 f(x^*)$ positive definite also. Then, $\alpha^k = 1$ for all k sufficiently large and $x^k \to x^*$ quadratically.