Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimisation I Spring 2020

Exercise set 3

1 Consider the quadratic function

$$f(x) = \frac{1}{2}x^T Q x - b^T x$$

with $Q \in \mathbb{R}^{d\times d}$ symmetric and positive definite and $b \in \mathbb{R}^d$.

a) Compute the gradient and the Hessian of the function f, verify that f is strictly convex, and find the unique global minimum of f.

b) Let $x \in \mathbb{R}^d$, and let $p \in \mathbb{R}^d$ be a direction satisfying the inequality $\nabla f(x)^T p < 0$. Compute analytically the step length $\alpha_{x,p}$ that solves the (exact) line search problem $\min_{\alpha>0} f(x + \alpha p)$.

c) Let x, $p \in \mathbb{R}^d$, and $\alpha_{x,p}$ be as in the previous question. Show that the step length $\alpha_{x,p}$ satisfies the strong Wolfe conditions if and only if $c_1 \le 1/2$.

2 *Cf. the exam 2016.*

A function $f\colon \mathbb{R}^d \to \mathbb{R}$ is called *strongly convex*,[1] if there exists $c > 0$ such that the function $x \mapsto f(x) - \frac{c}{2}\|x\|^2$ is convex.

a) Show that a twice differentiable function $f\colon \mathbb{R}^d \to \mathbb{R}$ is strongly convex, if and only if there exists $c > 0$ such that

$$p^T \nabla^2 f(x) p \ge c\|p\|^2$$

for all $p \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$.
Verify that this is equivalent to the condition that all the eigenvalues of all the matrices $\nabla^2 f(x)$, $x \in \mathbb{R}^d$, are larger or equal to c.

b) Assume that $f\colon \mathbb{R}^d \to \mathbb{R}$ is twice continuously differentiable and strongly convex, and assume that x^* is a minimiser of f. Show that there exists $c > 0$ such that

$$f(x) \ge f(x^*) + \frac{c}{2}\|x - x^*\|^2$$

for every $x \in \mathbb{R}^d$.

c) Show that every strongly convex and continuously differentiable function $f\colon \mathbb{R}^d \to \mathbb{R}$ has a minimizer $x^* \in \mathbb{R}^d$ and that that minimizer is unique.

[1] Not to be confused with *strictly convex*, which is a weaker notion!

d) Consider the *damped Newton method*

$$x_{k+1} = x_k + \alpha_k p_k \qquad \text{with} \qquad p_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k),$$

where the step length parameter α_k is chosen according to backtracking Armijo line search with parameters $\hat{\alpha} > 0$, $0 < c_1 < 1$, and $0 < \rho < 1$.

Show that p_k is a descent direction in each step, and that the sequence x_k converges to the unique minimizer x^* of f.

Hint: Use Theorem 1 in the note "Convergence of descent methods with backtracking (Armijo) linesearch..." by Anton Evgrafov.

3 Consider the function

$$f(x, y) = 2x^2 + y^2 - 2xy + 2x^3 + x^4.$$

a) Compute all stationary points of f and find all global or local minimisers of f.

b) Consider the gradient descent method with backtracking for the minimisation of f. Use the parameters $\rho = 1/2$ and $c_1 = 1/4$. Perform one step with starting value $(x_0, y_0) = (-1, 0)$. Does the method converge to a minimiser of f?

c) Consider Newton's method with backtracking for the minimisation of f. Use the parameters $\rho = 1/2$ and $c_1 = 1/4$. Perform one step with starting value $(x_0, y_0) = (-1, 0)$. Does the method converge to a minimiser of f?

4 Let

$$A := \begin{pmatrix} 2 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 2 \end{pmatrix} \qquad \text{and} \qquad b := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Use the CG-method with initialisation $x_0 = 0$ for solving the linear system $Ax = b$.

5 (Cf. Exercise 5.1 in Nocedal & Wright)

Implement the CG method for the solution of linear systems $Ax = b$ with symmetric and positive definite matrix $A \in \mathbb{R}^{n \times n}$.

Use your method in the case where A is the Hilbert matrix, the elements of which are

$$A_{i,j} = \frac{1}{i + j - 1}, \qquad 1 \le i, j \le n.$$

Use the right hand side $b = (1, 1, \ldots, 1)^T$ and the initialisation $x_0 = 0$. Test your code for dimensions $n = 5,\ 8,\ 12,\ 20$. How many iterations are required to reduce the residual below 10^{-6}? Why do your results not contradict the theoretical results concerning the CG method that were discussed in the lecture?

Hint: You might want to have a look at the condition number of the Hilbert matrix.

(Note that in MATLAB the Hilbert matrix can be produced with the command `hilb`, and in Python using `scipy.linalg.hilbert`.)