E. Herberg, M. Marić, V. Stein
Heidelberg University

Exercise 5

Date issued: 13th May 2024
Date due: 22nd May 2024

Homework Problem 5.1 (Efficiency of Wolfe-Powell Step Sizes for $C^{1,1}$ Functions) 5 Points

Let $f \in C^1$ and let $x^{(0)} \in \mathbb{R}^n$ be an initial iterate of the generic descent scheme (Algorithm 4.2). Further assume that f' is Lipschitz continuous on the sublevel set $\mathcal{M}_f(x^{(0)}) \coloneqq \{x \in \mathbb{R}^n \mid f(x) \le f(x^{(0)})\}$.

Show that step sizes $\alpha^{(k)}$ that satisfy the Wolfe-Powell-conditions at $x^{(k)}$ for the descent direction $d^{(k)}$ for all k are efficient and that there is a $c > 0$ such that

$$f(x^{(k)} + \alpha^{(k)} d^{(k)}) - f(x^{(k)}) \le -c \left(\cos \measuredangle\big(-\nabla_M f(x^{(k)}), d^{(k)}\big) \|f'(x^{(k)})^\intercal\|_{M^{-1}} \right)^2$$

for all $k \ge 0$.

Homework Problem 5.2 (Scaling Invariance of Armijo- and Curvature Conditions) 5 Points

Show the statement of remark 4.21, i. e. that when a step length α satisfies any of the Armijo- or curvature conditions (4.12), (4.17) and (4.18) for $g(x) \coloneqq \gamma f(Ax + b) + \delta$ at $x \in \mathbb{R}^n$ with search direction $d \in \mathbb{R}^n$, where $A \in \mathbb{R}^{n \times n}$ is non-singular, $b \in \mathbb{R}^n$, $\gamma > 0$ and $\delta \in \mathbb{R}$, then it satisfies the respective conditions for f at $Ax + b$ with the search direction Ad.

Homework Problem 5.3 (Implementation of Nonlinear Steepest Descent and Armijo Backtracking) 8 Points

Implement the M-steepest descent method as outlined in Algorithm 4.22 with the original Armijo backtracking as outlined in Algorithm 4.11. You can also try to use the modified (interpolating) Armijo backtracking as outlined in Algorithm 4.15.

Visualize and examine the effect of the parameters of the step size strategy on the behavior of the algorithm when applied to quadratic, strongly convex functions, Rosenbrock's and/or Himmelblau's functions.

Homework Problem 5.4 (Affine Invariance of Newton's Method for Root Finding) 10 Points

Prove the statement in remark 4.29(iii) of the lecture notes concerning affine invariance of local Newton's method for solving the root finding problem $F(x) = 0$ with continuously differentiable $F\colon \mathbb{R}^n \to \mathbb{R}^n$ (Algorithm 4.23 of the lecture notes).

I. e., let $A \in \mathbb{R}^{n\times n}$ be regular and $b \in \mathbb{R}^n$ and consider a sequence $(x^{(k)})_{k\in\mathbb{N}_0}$ of iterates produced by Newton's method for F started from $x^{(0)} \in \mathbb{R}^n$. Prove that:

(i) Newton's method for the function

$$G\colon \mathbb{R}^n \mapsto \mathbb{R}^n, \quad G(y) := F(Ay + b)$$

with initial value $y^{(0)} \in \mathbb{R}^n$ such that $x^{(0)} = Ay^{(0)} + b$ is well defined and produces the sequence $(y^{(k)})_{k\in\mathbb{N}_0}$ of iterates with

$$x^{(k)} = Ay^{(k)} + b.$$

(ii) Newton's method for the function

$$H\colon \mathbb{R}^n \mapsto \mathbb{R}^n, \quad H(y) := AF(y)$$

with initial value $y^{(0)} \in \mathbb{R}^n$ such that $x^{(0)} = y^{(0)}$ is well defined and produces the sequence $(y^{(k)})_{k\in\mathbb{N}_0}$ of iterates with

$$x^{(k)} = y^{(k)}.$$

(iii) Explain why we can not expect a similar transformation result to hold for the iterates of Newton's method when we expand the transformation in Part (ii) by an additional constant shift, as in

$$H\colon \mathbb{R}^n \mapsto \mathbb{R}^n, \quad H(y) := AF(y) + b.$$

Please submit your solutions as a single pdf and an archive of programs via moodle.