R. Herzog, G. Müller, E. Herberg, M. Hashemi
Heidelberg University

Exercise 5

Date issued: 15th May 2023
Date due: 23rd May 2023

Homework Problem 5.1 (Affine Invariance of Newton's Method for Root Finding) 10 Points

Prove the statement in remark 5.29(*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 5.23 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

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

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

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

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

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

Homework Problem 5.2 (Newton Fractals in Root Finding) 10 Points

The convergence of Newton's method for varying initial values can be quite chaotic, depending on the initial value. A nice visualization of its behavior are so called fractal plots for root finding problems, which color each starting point according to the root that the method converged to when started at that point. Fractal plots are typically created for Newton's method in the complex numbers, but this is equivalent to working in $\mathbb{R}^2$, as we will see.

(*i*) Let $\phi \colon \mathbb{R}^2 \to \mathbb{C}$ denote the canonical isomorphism $\phi(x, y) = x + yi$ as well as $F_\mathbb{C} \colon \mathbb{C} \to \mathbb{C}$ be continuously differentiable and $F \coloneqq \phi^{-1} \circ F_\mathbb{C} \circ \phi \colon \mathbb{R}^2 \to \mathbb{R}^2$.

Further let $z^{(0)} = x^{(0)} + y^{(0)} i = \phi(x^{(0)}, y^{(0)}) \in \mathbb{C}$ be an initial value and let $z^{(k)}$ be a sequence of iterates of the local Newton's method in the complex numbers, defined as

$$z^{(k+1)} = z^{(k)} - F_\mathbb{C}'(z^{(k)})^{-1} F_\mathbb{C}(z^{(k)}),$$

started from $z^{(0)}$.

Show that Newton's method for F started at $(x^{(0)}, y^{(0)})$ is well defined and yields the sequence $(x^{(k)}, y^{(k)})^\intercal$ with $\phi(x^{(k)}, y^{(k)}) = z^{(k)}$ for all $k \in \mathbb{N}$.

(*ii*) Implement the local Newton's method for solving problems of the type

$$F(x) = 0$$

for $F \colon \mathbb{R}^n \to \mathbb{R}^n$ (Algorithm 5.23 of the lecture notes).

(*iii*) Reformulate the root finding problem $F_\mathbb{C}(z) = 0$, $z \in \mathbb{C}$ for

$$F_\mathbb{C} \colon \mathbb{C} \to \mathbb{C}, \quad F_\mathbb{C}(z) = z^3 - 1$$

into an equivalent problem in $\mathbb{R}^2$, solve the problem numerically for a grid of initial values and create a fractal plot by coloring each starting value corresponding to the root that the local Newton's method converged to starting from that point.

Homework Problem 5.3 (On the Restriction $\sigma \in (0, \frac{1}{2})$ in Globalized Newton) 7 Points

In the globalized Newton's method for optimization (Algorithm 5.30 of the lecture notes), the Armijo-parameter, which is typically chosen as $\sigma \in (0, 1)$, is restricted to the interval $(0, \frac{1}{2})$ so that the full Newton step size $\alpha^{(k)} = 1$ can in fact be accepted by the Armijo condition for $k \geq k_0$ and some $k_0 > 0$, in order to facilitate quadratic convergence in the final stages of the algorithm. We will investigate why that is:

(i) Show that the step length $\alpha^{(k)} = 1$ satisfies the Armijo condition for the Newton direction $d^{(k)} \neq 0$ for the quadratic function

$$f(x) = \frac{1}{2} x^\intercal A x + b^\intercal x + c$$

with s. p. d. $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, und $c \in \mathbb{R}$ if and only if $\sigma \leq \frac{1}{2}$.

(ii) Explain why we need to restrict ourselves to $\sigma < \frac{1}{2}$ for general nonquadratic problems.

Homework Problem 5.4 (Globalized Newton's Method in Optimization) 8 Points

Implement the globalized Newton's method for optimization (Algorithm 5.30 of the lecture notes), run it for the Rosenbrock's and/or Himmelblau's functions and compare its performance to that of your gradient descent implementation.

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