Exercise 3

January 28, 2025

Problem 1.

Assume that $B \in \mathbb{R}^{d \times d}$ is a non-singular matrix (not necessarily orthogonal), $\hat{c} \in \mathbb{R}^d$, and $f : \mathbb{R}^d \to \mathbb{R}$ is a real-valued function. Define the function $g(x) = f(Bx + \hat{c})$.

a) Find expressions for $\nabla g(x)$ and $H_g(x)$ in the terms of f, B, and $\hat{c}$. (*You can (and maybe should) use the chain rule for the computation of ∇g and H_g.*)

b) Let $x_0 \in \mathbb{R}^d$ and denote by x_1 the result of one Newton step starting at x_0 for the minimisation of g with (Armijo) backtracking line search using the parameters $0 < c < 1$ (sufficient decrease parameter), $0 < \rho < 1$ (contraction factor), and $\hat{\alpha} = 1$ (initial step length).

   Moreover, let $y_0 = Bx_0 + \hat{c}$ and denote by y_1 the result of one Newton step starting at y for the minimisation of f with the (Armijo) backtracking line search and the same parameters $0 < c < 1$, $0 < \rho < 1$, and $\hat{\alpha} = 1$. Show that

$$y_1 = Bx_1 + \hat{c}. \tag{1}$$

c) Show that the relation (1) in general does not hold for the gradient descent method unless B is an orthogonal matrix.

Problem 2.

Implement both the gradient descent method and Newton's method with backtracking line search. Apply your method to the minimisation of the *Rosenbrock function* $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = 100(y - x^2)^2 + (1 - x)^2$.

The Newton direction is not necessarily a descent direction for this function, as f is not convex, and thus it might be necessary to modify the search directions in the Newton method. Do this by switching to the negative gradient direction, whenever the inequality

$$-\langle \nabla f(x_k), p_k^{\text{Newton}} \rangle \leq \epsilon \|\nabla f(x_k)\| \|p_k^{\text{Newton}}\|$$

holds (here $\epsilon > 0$ is some fixed small parameter).

Problem 3.

Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ (see Exercise 1, Problem 3a)

$$f(x,y) = \frac{x^2}{2} + x\cos y.$$

We want to apply a line search method for the minimisation of f, starting at the point $x_0 = (1, \frac{\pi}{4})$ and with the search direction $p_0 = (-1, 0)$. Formulate the (weak) Wolfe conditions for this situation! If the parameters are chosen as $c_1 = 0.1$ and $c_2 = 0.8$, what is the range of admissible values for the step length α?

Problem 4.

Consider the optimisation problem

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

where the objective function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is defined as

$$f(x,y) = 2x^2 - 2xy^2 - 12x + y^4 + 2y^2 + 36.$$

a) Compute all stationary (critical) points of the optimisation problem (2) and also find all local and global minima.

b) Starting at the point $x_0 = (1,2)$, compute one step of Newton's method with backtracking (Armijo) line search. Start with an initial step length $\alpha_0 = 1$, and use the parameters $c = 1/8$ (sufficient decrease parameter) and $\rho = 1/2$ (contraction factor).

Problem 5.

Consider the optimization problem

$$\min_{x\in\mathbb{R}^2} f(x), \tag{3}$$

where the objective function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is given by

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

a) Compute all stationary (critical) points of the optimisation problem (3) and also find all its local and global minima.

b) Starting at the point $x_0 = (0,0)$ compute one step of the gradient (steepest) descent method. Ensure that the step length satisfies the (weak) Wolfe conditions with $c_1 = \frac{1}{4}$ and $c_2 = \frac{3}{4}$.

c) In part b), can you also use the Newton search direction?