

Exercise #3

January 31, 2023

Problem 1.

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

- Find the expressions for $\nabla g(x)$ and $\nabla^2 g(x)$ in the terms of f , B and c . (**Hint:** You can use the chain-rules for ∇ and ∇^2 .)
- Let $x \in \mathbb{R}^d$ and denote by x_1 the result of one Newton step starting at x for the minimization of g with the (Armijo) backtracking line search and the parameters $0 < c < 1$ (sufficient decrease parameter), $0 < \rho < 1$ (contraction factor), and $\hat{\alpha} = 1$ (initial step length). Moreover, let $y = Bx + c$ and denote by y_1 the result of one Newton step starting at y for the minimization of f with the (Armijo) backtracking line search and the same parameters $0 < c < 1$, $0 < \rho < 1$, and $\hat{\alpha} = 1$. Show that

$$y_1 = Bx_1 + c. \quad (1)$$

- 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 minimization of *Rosenbrock* function $f(x, y) = 100(y - x^2)^2 + (1 - x)^2$.

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 $-\nabla f(x_k)^T p_k^{\text{Newton}} \leq \epsilon \|\nabla f(x_k)\| \|p_k^{\text{Newton}}\|$ holds (here $\epsilon > 0$ is some fixed small parameter).

Problem 3.

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

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

We want to apply a line search method for the minimization of f , starting at the point $x_0 = (1, \frac{\pi}{4})$ and with the search direction $p_0 = (-1, 0)$. State the Wolfe conditions for this line search and explain their purpose! 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 optimization problem

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

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

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

- Compute all stationary (critical) points of the optimization problem (2) and also find all its local and global minima.
- 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

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

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

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

- Compute all stationary (critical) points of the optimization problem (3) and also find all its local and global minima.
- Starting at the point $x_0 = (0, 0)$ compute one step of the gradient (steepest) descent method. Ensure that the step length satisfies the Wolfe conditions with $c_1 = \frac{1}{4}$ and $c_2 = \frac{3}{4}$.
- In part b), can you also use the Newton search direction?