Prof. Dr. Karsten Urban
M. Sc. Lewin Ernst

Winter term 2024/2025
Sheet 04, 34 Points

Numerical Optimization
exercise sheet

review on 12.11.2024 during the exercise class

1. *(Step size too small or too big?)*
The following example shows that the step size has to be choosen very wisely. Consider the function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = x^2$ and the iteration

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} p^{(k)}, \qquad k \in \mathbb{N}_0,$$

where the initial value is given by $x^{(0)} = 1$.

a) Choose the search direction $p^{(k)} = -1$ and the step size $\alpha^{(k)} = 2^{-k-2}$ for $k \in \mathbb{N}_0$. Prove that $\lim_{k\to\infty} x^{(k)} = \frac{1}{2}$ and $\lim_{k\to\infty} f(x^{(k)}) = \frac{1}{4}$. Discuss this result.

b) Choose the search direction $p^{(k)} = (-1)^{k+1}$ and the step size $\alpha^{(k)} = 1 + \frac{3}{2^{k+2}}$ for $k \in \mathbb{N}_0$. Calculate $x^{(k)}$. Discuss your result.

(4 + 4 = 8 Points)

2. *(Wolfe conditions)*
Typical line search algorithms try out a sequence of candidate values for the step size α, stopping to accept one of these values when certain conditions are satisfied. We want to take a closer look at the Wolfe conditions which consist of

1.) a sufficient decrease condition of the objective function f (also known as *Armijo condition*)

$$f(x + \alpha p) \le f(x) + c_1 \alpha \nabla f(x)^T p, \qquad c_1 \in (0, 1) \tag{1}$$

and

2.) the curvature condition

$$\nabla f(x + \alpha p)^T p \ge c_2 \nabla f(x)^T p, \qquad c_2 \in (c_1, 1) \tag{2}$$

where f is the objective function, p is a descent direction and α is the step size.

a) Show that if $0 < c_2 < c_1 < 1$, there may be no step lengths that satisfy the Wolfe conditions.

b) Prove Lemma 2.3.10.

c) Prove that the quadratic function $\Phi_1^{(k)}$ that interpolates

$$\Phi_1^{(k)}(0) = \Phi^{(k)}(0), \quad \left(\Phi_1^{(k)}\right)'(0) = \left(\Phi^{(k)}\right)'(0) \quad \text{and} \quad \Phi_1^{(k)}(\alpha_0^{(k)}) = \Phi^{(k)}(\alpha_0^{(k)})$$

is given by $\Phi_1^{(k)}$ in Algorithm 2.3.3 Line 5. Then, make use of the fact that the Armijo's condition is not satisfied at α_0 to show that the quadratic $\Phi_1^{(k)}$ has positive curvature and the minimizer satisfies

$$\alpha_1 < \frac{\alpha_0}{2(1-c_1)}.$$

Remark: Since $c_1 \approx 10^{-4}$, this inequality gives us an idea of the new step length, in case of acceptance.

(4 + 5 + 5 = 14 Points)

3. *(Newton with Armijo stepsize control,* MATLAB*)*
Let the Himmelblau function be given. We shall visualize the Armijo step size control Algorithm 2.3.3.

a) Implement the MATLAB-function

```
alpha = armijo_step_size(f, gradf, x, p, alpha0, c1,...
                         delta_min, delta_max, maxIt, plot_flag)
```

which realizes Algorithm 2.3.3. The input parameters should be clear, except for `maxIt` and the `plot_flag`. The `maxIt` is used to break the for-loop in Line 15. Ignore the `plot_flag` input parameter for the moment.

b) In the material you will find the MATLAB-function `newton_armijo`. Write a script which applies Newton's method with your step size control algorithm to the Himmelblau-function. Use as an initial value `x0 = [2.6, -3.9]'` and `x0 = [2, -3.5]'`. Plot the path of the iterates in the same plot as the contour lines of the Himmelblau function.

c) Adjust your MATLAB-function `armijo_step_size.m`, such that if the `plot_flag` is true, then the functions $\Phi(\alpha)$, $\Phi_1(\alpha)$, $\Phi_2(\alpha)$[1] and $\Phi_{\text{lin}}(\alpha) := \Phi(0) + c_1\Phi'(0)\alpha$ as well as the interpolation points and minimizer are plotted. Choose the same initial points as in b) and use `maxIt = 1` for Newton's method as well as for the step size control. What do you observe?

(4 + 4 + 4 = 12 Points)

[1] This is the cubic interpolant. Plot it only within the loop of Line 15.