Numerical Optimization 2024 - Homework 4

Deadline: Wednesday, April 17, 15:30.

In this homework, we fully focus on the Rosenbrock function

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2. \tag{1}$$

Basically, your task is to check what does the theory from Section 3.2 (CONVERGENCE OF LINE SEARCH METHODS) say about convergence of the **steepest descent** applied to this function (the search direction p_k below corresponds to $-\nabla f(x_k)$). More precisely, you should check if the assumptions of the convergence theorems are satisfied and draw a conclusion based on that. Try to answer the best you can and do not worry if you are not sure about something. The purpose is that you try to think about the behaviour of these methods, not that you provide a perfect answer. You can also help yourself by programming the method and trying it.

Problem 1. Sketch approximately the contour lines of the Rosenbrock function, i.e. the curves where it attains a fixed value c (the set of all x such that $f(x) = c$). You do not have to really compute it from $f(x) = c$; try to guess it and helps yourself by trying some points (e.g. $(1,1)$, $(0,0)$, $(-1,1)$, $(0,1)$, etc.). Specifically, sketch the set of points where f attains values 0, 1, and 4. See some figures in the book for inspiration if you need.

Problem 2. Analyze convergence of the steepest descent algorithm. In this task, you can assume that the line search is exact, i.e. α_k is the global minimizer of the function $\phi(\alpha) = f(x_k + \alpha p_k)$. More precisely, with the aid of results from Sections 3.2 and 3.3, try to answer the following questions and justify/explain your answers:

(i) What convergence do you expect (local, global, none, to which point)?

(ii) If you expect convergence, then how fast it will be (linear, superlinear, quadratic)? If linear, try to estimate the corresponding constant.

(iii) Suppose you start at $x_0 = (-1,1)$: Can you guess the trajectory of the iterations? Will they go directly towards the limit? Use the contour lines and the fact that the search direction is a descent direction to answer.

Problem 3. Detail the line search process (selection of α_k), e.g. write explicitly the function $\alpha \to f(x_k + \alpha p_k)$, sketch its graph for $\alpha \in [0,1]$ and try to depict the backtracking process (see Algorithm 3.1 from the book). For x_k take points

(i) $x_k = (1.2, 1.2)^T$;

(ii) $x_k = (-1.2, 1)^T$.