1 Chapter 3

1.1 Exercise 1

The Rosenbrock function

$$f : \mathbb{R}^2 \to \mathbb{R}, x \mapsto 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$

also compare `http://en.wikipedia.org/wiki/Rosenbrock_function`, is frequently utilized to test optimization methods.

a. The absolute minimum of f, at the points $(1, 1)^T$, can be seen without any calculations. Show that it is the only extremal point.

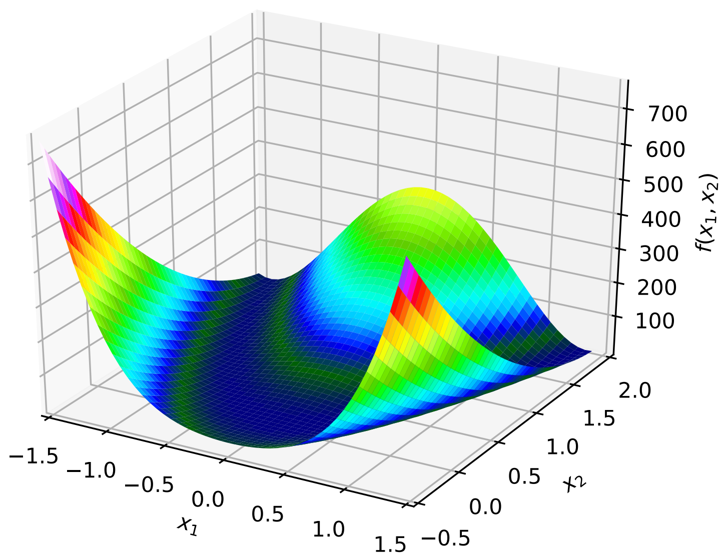
   All local (and thus global) extrama occur at critical points of f, therefore it is sufficient to show that the only critical point exists at $(1, 1)^T$.

$$\nabla f(x) = \begin{pmatrix} 200(x_2 - x_x^2)(-2x_1) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}$$

   The only solution to $\nabla f(x) = 0$ is $x = (1, 1)^T$, which is at the absolute minimum $(1, 1)^T$, which we already have. There are no other critical points, thefore this is the only extrema.

b. Graph the function f on $[-1.5, 1.5] \times [-0.5, 2]$ as a 3D plot or as a level curve plot to understand why f is reffered to as the *banana function* and the like.

Figure 1: The ROSENBROCK Function



c. Implement the NELDER-MEAD method. Visualize the level curves of the given function together with the polytop for each iteration. Finnaly visualize the trajectory of the ecenters of gravity of the polytopes!

d. Test the program with the starting polytope given by the vertices $(-1, 1)^T$, $(0, 1)^T$, $(-0.5, 2)^T$ and the parameters $(\alpha, \beta, \gamma) := (1, 2, 0.5)$ and $\varepsilon := 10^{-4}$ using the ROSENBROCK function. How many iterations are needed? What is the distance between the calculated solution and the exact minimizer $(1, 1)^T$?

e. Find $(\alpha, \beta, \gamma) \in [0.9, 1.1] \times [1.9, 2.1] \times [0.4, 0.6]$ such that with $\varepsilon := 10^{-4}$ the algorithm terminates after as few iterations as possible. What is the distance between the solution and the minimizer $(1, 1)^T$ in this case?

f. If the distance in *e* was greater than in *d*, reduce ε until the algorithm gives a result—with the (α, β, γ), found in e—which is not farther away from $(1, 1)^T$ and at the same time needs fewer iterations than the solution in d.