NTNU

Kunnskap for en bedre verden

TTK4135 - Optimization and Control

Exercise #8

Author:
Sondre Pedersen

March 13, 2025

Problem 1: The Rosenbrock function

Compute the gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ of the Rosenbrock function

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Show that $x^* = (1,1)^\top$ is the only local minimizer of this function, and that the Hessian matrix at that point is positive definite.

$$\nabla f(x) = \begin{bmatrix} 200(x_2 - x_1^2)(-2x_1) - 2(1 - x_1) \\ 200(x_2 - x_2^2) \end{bmatrix} = \begin{bmatrix} -400(x_1 x_2 - x_1^3) + 2x_1 - 2 \\ 200(x_2 - x_1^2) \end{bmatrix}.$$

$$\nabla^2 f(x) = \begin{bmatrix} -400x_2 + 1200x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}.$$

$\nabla f(x) = 0$ at the points where x is a local minimizer. We have that

- $200(x_2 - x_1^2) = 0 \implies x_2 = x_1^2$
- $-400(x_1 x_2 - x_1^3) + 2x_2 - 2 = 0 \implies -400(x_1^3 - x_1^3) + 2x_1 - 2 = 0 \implies x_1 = 1 = x_2$

So $x_1 = x_2 = 1$ is the only minimizer for this function.
The Hessian matrix at that point is $\begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$.
Since the top left element is positive and the determinant is positive $(802 * 200 - 400 * 400)$, the Hessian is positive definite.

Problem 2: The Newton Direction

Consider the model function m_k based on the second-order Taylor approximation

$$m_k(p) = f_k + p^\top \nabla f_k + \frac{1}{2} p^\top \nabla^2 f_k p \approx f(x_k + p).$$

(a) Derive the expression for the Newton direction

$$p_k^N = -(\nabla^2 f_k)^{-1} \nabla f_k.$$

using the model function m_k. The Newton direction is the direction that minimized $m_k(p)$.

(b) Assume that $\nabla^2 f_k$ is not positive definite. In this case, is the Newton-direction p_k^N a descent direction? Is it even defined?

(c) Given an unconstrained minimization problem with objective function

$$f(x) = \frac{1}{2} x^\top G x + c^\top x.$$

with $G = G^\top > 0$ and $x \in \mathbb{R}^n$. Show that an iteration algorithm based on Newton direction (i.e., $x_{k+1} = x_k + p_k^N$) always converges to the optimum in *one step*.

(a) Derive Newton direction

$$\begin{aligned} \frac{\partial m_k}{\partial p} &= 0 \\ &= \nabla f_k + \frac{1}{2}(\nabla^2 f_k)^\top p + \frac{1}{2} \nabla^2 f_k p \\ &= \nabla f_k + \nabla^2 f_k p \\ \implies p &= -(\nabla^2 f_k)^{-1} f_k \end{aligned}$$

Here we use that $(\nabla^2 f_k)^\top p = \nabla^2 f_k p$ (Symmetric Hessian).

(b) When Hessian is not pd

The Newton direction is a descent direction if it satisfies:

$$\nabla f_k^\top p_N^\top < 0.$$

Expanding with the expression for p_N^k:

$$\nabla f_k^\top (-(\nabla^2 f_k)^{-1} \nabla f_k) = -\nabla f_k^\top (\nabla^2 f_k)^{-1} \nabla f_k.$$

If $\nabla^2 f_k$ is positive definite, then $(\nabla^2 f_k)^{-1}$ is also positive defininte. This ensures that the quadratic form is strictly negative, making p_n^k a descent direction.

However, if $\nabla^2 f_k$ is not positive definite, then $(\nabla^2 f_k)^{-1}$ is not well defined, and could be a direction that is not descending.

p_N^k is defined if $\nabla^2 f_k$ has negative eigenvalues. Otherwise it is not defined.

(c) Newton convergence

The Newton direction is given by $p_k^N = -(\nabla^2 f_k)^{-1} \nabla f_k$. Using $\nabla f(x_k) = Gx_k + c$ and $\nabla^2 f_k = G$ gives

$$\begin{aligned} p_k^N &= -G^{-1}(Gx_k + c) \\ &= -x_k - G^{-1}c \end{aligned}$$

The next step is now $x_{k+1} = x_k + p_k^N = x_k - x_k - G^{-1}c = -G^{-1}c$.
This is the optimal solution. Setting $\nabla f(x) = 0$ gives $Gx^* + c = 0 \implies x^* = -G^{-1}c$.

Problem 3: Search direction

Consider the function

$$f(x) = (x_1 + x_2^2)^2.$$

At the point $x_k^\top = (1, 0)$ we consider the search direction $p_k^\top = (-1, 1)$. Show that p_k is a descent direction, and find all minimizers of

$$\min_{\alpha > 0} f(x_k + \alpha p_k).$$

$$\nabla f(x) = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix}.$$

$$\nabla f(1, 0) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Given $p_k^\top = (-1, 1)$, we get that

$$\nabla f(1, 0)^\top p_k = \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -2.$$

Since the result is negative, we have a descent direction.

Define $x(\alpha) = x_k + \alpha p_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 - \alpha \\ \alpha \end{bmatrix}$. Using this in f(x) gives $f(1 - \alpha, \alpha) = ((1 - \alpha) + \alpha^2)^2 := g(\alpha)$.

$$\frac{dg}{d\alpha} = 2(1 - \alpha + \alpha^2) \times (-1 + 2\alpha) = 0 \implies \alpha = \frac{1}{2}.$$