Nonlinear Optimization - Homework 1

Christian Segercrantz 481056

October 20, 2021

2.1

$$f(x) = x_1^3 - x_1 + x_2^3 - x_2 \tag{1}$$

(a)

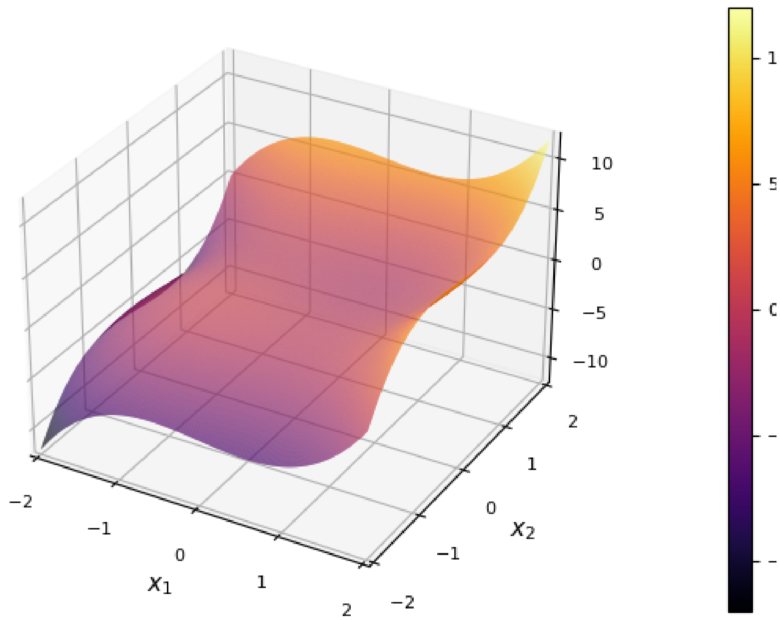

Figure 1: The surface plot of equation (1).

By examining Figure 1, I would conclude that the function is non-convex as we can clearly see areas that do not follow the definition of a convex function $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$.

(b)

In order for us to find the critical point, we will solve the first-order derivative of the function.

$$\nabla f(x_1, x_2) = \begin{bmatrix} 3x_1^2 - 1 \\ 3x_2^2 - 1 \end{bmatrix}. \tag{2}$$

The first-order necessary condition is $\nabla f(\bar{x}) = 0$ for our unconstrained problem . We can solve the equations to get the critical points:

$$2x^2 - 1 = 0 \iff x = \pm\sqrt{1/3}. \tag{3}$$

This gives us the four points $(\sqrt{1/3}, \sqrt{1/3})$, $(-\sqrt{1/3}, \sqrt{1/3})$, $(\sqrt{1/3}, -\sqrt{1/3})$, and $(-\sqrt{1/3}, -\sqrt{1/3})$.

(c)

We calculate the hessian from the gradient and receive

$$Hf(x) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 6x_2 \end{bmatrix} \tag{4}$$

In order for us to know how the function behaves, we need to examine the semi-definiteness of the Hessian. A positive semi-definite matrix has all of its eigenvalues non-negative and vice verse for a negative semi-definite matrix for all x. Additionally, a convex(/concave) functions hessian, if twice differentiable, is positive(/negative) semi-definite. The two eigen values of the function is $\lambda_1 = 6x_1$ and $\lambda_2 = 6x_2$. We can see that only for values where both x_1 and x_2 are non-negative(/-positive) is the hessian positive(/negative) semi-definite. In other words, the function is not convex(concave).

Using Julia we can solve the eigen values in order to get the curvature around the points. The point $(\sqrt{1/3}, \sqrt{1/3})$ has all positive eigen values, $(-\sqrt{1/3}, -\sqrt{1/3})$ has all negative eigen values and the other two one positive and one negative. From this we can conclude that $(\sqrt{1/3}, \sqrt{1/3})$ is a local minima, $(-\sqrt{1/3}, -\sqrt{1/3})$ a local maxima and the two others saddle points.

2.2

$$f(x_1, x_2) = 2x_1^2 - x_1x_2 + x_2^2 - 3x_1 + e^{2x_1+x_2} \tag{5}$$

(a)

By corollary 6 of lecture 4, we know that the first-order necessary condition for unconstrained problems is that *"Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\bar{x}$. If $\bar{x}$ is a local minimum, then $\nabla f(\bar{x}) = 0$"*. The equation to be satisfied is thus $\nabla f(\bar{x}) = 0$. We calculate ∇f:

$$\nabla f = \begin{bmatrix} 4x_1 - x_2 - 3 + 2e^{2x_1+x_2} \\ -x_1 + 2x_2 + e^{2x_1+x_2} \end{bmatrix} \tag{6}$$

For this to be a sufficient condition for optimality the function $f(x_1, x_2)$ has to be convex. Then, based on theorem 8 from lecture 4 can we have sufficient conditions. Since we know the following things:

1. Polynomials are convex,
2. The linear combination of convex functions are convex,

we can conclude that the function $f(x_1, x_2)$ is convex. We thus have the necessary and sufficient condition for optimality.

(b)

If $\bar{x} = (0, 0)$ is to be a optimal point, must it satisfy $\nabla f(0, 0) = \mathbf{0}$.

$$\begin{aligned} \nabla f &= \begin{bmatrix} 4x_1 - x_2 - 3 + 2e^{2x_1+x_2} \\ -x_1 + 2x_2 + e^{2x_1+x_2} \end{bmatrix} && (7) \\ &= \begin{bmatrix} 0 - 0 - 3 + 2e^0 \\ -0 + 0 + e^0 \end{bmatrix} && (8) \\ &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (9) \end{aligned}$$

We can see that the point $(0, 0)$ does not satisfy the condition.

The direction d that makes the function decrease must satisfy the condition $\nabla f(\bar{x})^\top d < 0$.

(c)

To find the minimum for $f(x)$ in the direction d= $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ we must calculate the step size $\bar{\lambda} = \text{argmin}_\lambda d^\top \nabla(x + \lambda d)$. We can