Nonlinear Optimization - Homework 1

Christian Segercrantz 481056

October 18, 2021

1.1

We will first examine weather the set is non-empty and convex. We can see that the set is non-empty by finding a number that is contained in the set. Let us examine the point (1,1). The set is subject to $x_2 \geq x_1^2$ and $x_2 \leq 4$ and by inserting the point (1,1) into the conditions we get $1 \geq 1$ and $1 \leq 4$ which both hold. The set is thus non-empty.

From lecture 3, slide 3 we know that polynomial functions are convex functions. Since we have the sum of two second degree polynomial functions in different dimensions we can conclude that the function is convex. The epigraph of a convex function is also convex and the set is thus convex.

The necessary optimality condition for our nonempty convex set is $\xi^\top(x-\bar{x}) \geq 0$. It is known that second degree functions are differentiable, and thus we know that $\xi = \nabla f(\bar{x})$. Thus the condition becomes

$$\nabla f(\bar{x})^\top (x-\bar{x}) \geq 0. \tag{1}$$

The gradient becomes

$$\nabla f(\bar{x}) = \begin{bmatrix} 2(\bar{x}_1-4) \\ 2(\bar{x}_2-6) \end{bmatrix}. \tag{2}$$

The gradient at (2,4) becomes

$$\nabla f(2,4) = \begin{bmatrix} 2(2-4) \\ 2(4-6) \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}. \tag{3}$$

The complete expression then becomes

$$\nabla f(\bar{x})^\top (x-\bar{x}) \geq 0 \tag{4}$$

$$\begin{bmatrix} -4 & -4 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right) \geq 0 \tag{5}$$

$$-4(x_1 - 2 + x_2 - 4) = -4(x_1 + x_2 - 6) \geq 0 \tag{6}$$

Due to our constraints of the function, choosing the largest possible values for x will result in the outcome $0 \geq 0$ and any other value will result in a the LHS being greater than 0 since the parenthesis will be negative and the product, thus, result in a non-negative value. We have thus proven that the point (2,4) is optimal. Based on Theorem 1 in lecture 3, a optimal point of a convex function is a globally optimal point.

1.2

The closest point theorem says that $(x-\bar{x})^\top(x'-\bar{x})$ for any point x not in the convex set S and any point x' in the set S the point $\bar{x}$ is the closest point in the set S to y. Let's choose our point x' in the S to be $\bar{y}$ for x and $\bar{x}$ for y. Our two closest-point equations thus become

$$(x-\bar{x})^\top(\bar{y}-\bar{x}) \geq 0 \tag{7}$$

and

$$(y-\bar{y})^\top(\bar{x}-\bar{y}) \geq 0. \tag{8}$$

Further we know that $||a-b||^2 = (a-b)^\top(a-b)$ for any $a, b \in \mathbb{R}^n$. We can thus write

$$||\bar{x}-\bar{y}||^2 = (\bar{x}-\bar{y})^\top(\bar{x}-\bar{y}) \tag{9}$$

and add both the distances distances to the RHS. We will negate both equations so that they are greater than or equal to zero.

$$||\bar{x}-\bar{y}||^2 \leq (\bar{x}-\bar{y})^\top(\bar{x}-\bar{y}) + (\bar{y}-y)^\top(\bar{x}-\bar{y}) + (x-\bar{x})^\top(\bar{y}-\bar{x}) \tag{10}$$

$$\leq ((\bar{x}+\bar{y})^\top + (\bar{y}-y)^\top + (x-\bar{x})^\top)(\bar{x}-\bar{y}). \tag{11}$$

We can now open up the transpose parenthesis and simplify the expression into

$$||\bar{x} - \bar{y}||^2 \leq (x^\top - y^\top)(\bar{x} - \bar{y}) = (x - y)^\top(\bar{x} - \bar{y}). \tag{12}$$

We can now use the Cauchy–Schwarz inequality to show that

$$\begin{aligned} (x - y)^\top(\bar{x} - \bar{y}) \leq& ||x - y|| \cdot ||\bar{x} - \bar{y}|| \qquad (13) \\ \implies ||\bar{x} - \bar{y}||^2 \leq& ||x - y|| \cdot ||\bar{x} - \bar{y}|| \qquad (14) \\ ||\bar{x} - \bar{y}|| \leq& ||x - y|| \qquad (15) \end{aligned}$$

which was what we wanted to show.

1.3

(a)

$a^\top x \geq \alpha$ and $a^\top x \geq \beta$ form half-spaces. Since $\alpha \leq \beta$ that are two options:

1. $\alpha = \beta \implies$ we have a hyperplane which is convex.
2. $\alpha < \beta$ which is the area between two hyperplanes i.e. the intersection of half-spaces which is convex.

The set is thus convex.

(b)

The set forms a hyper cube, i.e. intersections of hyperplanes. We know from exercise 2.1 we know that the intersection of convex sets are also convex. Hence, the set is convex

(c)

The sets $a_1^\top x \leq b_1$ and $a_2^\top x \leq b_2$ form two half-spaces. Similarly part (a), there are two outcomes

1. The intersection of the half-spaces are empty $S = \emptyset$
2. The intersection is non-empty $S \neq \emptyset$.

A empty set is by definition convex. We know from exercise 2.1 we know that the intersection of convex sets are also convex. Since half-spaces are convex by lecture 2 slide 6 the set S is convex.

(d)

By the definition of a convex set, the point x will belong to the set S as $x = \lambda x_1 + (1 - \lambda)x_2$ where $x_1, x_2 \in S$ and $0 \leq \lambda \leq 1$. In order for the set to be convex we show the following:

$$\begin{aligned} x &= A^\top y \qquad (16) \\ &= A^\top(\lambda y_1 + (1 - \lambda)y_2) \qquad (17) \\ &= \lambda A^\top y_1 + (1 - \lambda)A^\top y_2 \qquad (18) \\ x &= \lambda x_1 + (1 - \lambda)x_2. \qquad (19) \end{aligned}$$

As we can see, any two points y_1, y_2 belong to the set. The set is thus convex.

1.4

(a)

Let us call $h(x) = Ax + b$. By lecture 3, slide 4 we know that that $f(x)$ is convex if $f(x) = g(h(x))$ if $g(x)$ is convex and $h(x)$ is affine. We can thus conclude that $f(x) = g(Ax + b)$

(b)

Let $f(x) = \alpha h(x) + \beta$. By the definition of convex function we set

$$f(\lambda x + (1-\lambda)y) = \alpha h(\lambda x + (1-\lambda)y) + \beta. \tag{20}$$

As we know that $h(x)$ is convex we can set

$$
\begin{align}
&\leq \alpha(\lambda h(x) + (1-\lambda)h(y)) + \beta \tag{21}\\
&= \lambda \alpha h(x) + (1-\lambda)\alpha h(y) + \beta \tag{22}\\
&= \lambda \alpha h(x) + \lambda\beta + (1-\lambda)\alpha h(y) + (1-\lambda)\beta \tag{23}\\
&= \lambda(\alpha h(x) + \lambda) + (1-\lambda)(\alpha h(y) + \beta) \tag{24}\\
&= \lambda f(x) + (1-\lambda)f(x) \tag{25}
\end{align}
$$

Which proves that our function $f(x)$ is convex.

(c)

Let $f(x) = e^{\beta x^\top Ax}$ and A a positive semidefinite matrix $x^\top Ax \geq 0$. Let $g(x) = e^x$ and $h(x) = \beta x^\top Ax$. We know that $g(x)$ is a convex as per the lecture 3 slide 3. For a non-decreasing function it holds that $f'(x) \geq 0$. For $g(x)$ it means that $x \geq 0$.

We can show that $h(x)$ is convex by showing that the Hessian of it is positive semidefinite, as per Theorem 12, lecture 3

$$H(h(x)) = \nabla^2 \beta x^\top Ax = \beta 2A \tag{26}$$

We know that $\beta > 0$ and as stated before that the matrix A is positive semidefinite and thus the Hessian is positive semidefinite.

By lecture 3 slide 4, we know that $g(h(x))$ is convex as long as $g(x)$ is a convex non-decreasing function and $h(x)$ is convex and positive-semidefinite. We have thus shown that $f(x) = e^{\beta x^\top Ax}$ is convex.

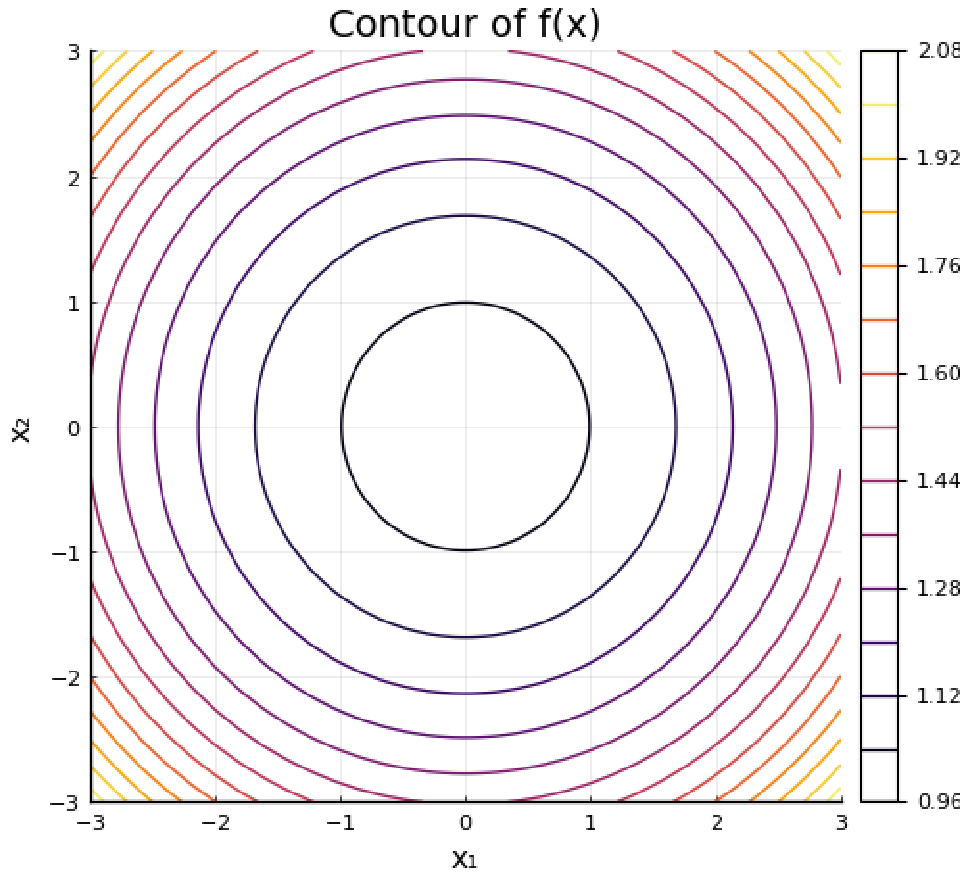

Figure 1: The contour plot of 1.4c