Introduction to Optimization: Homework #1

Due on October 10, 2020

Professor Andre Milzarek

Peng Deng

Assignment A1.1

A local family-owned plastic cup factory wants to optimize its production mix in order to maximize its profit. The factory produces personalized beer mugs and champagne glasses. The profit on a box of beer mugs is 200 RMB while the profit on a box of champagne glasses is 150 RMB.

The cups are manufactured with a machine called a plastic extruder which feeds on plastic resins. Each box of beer mugs requires 9 kg of plastic resins to produce while champagne glasses require 5.5 kg per box. The daily supply of plastic resins is limited to at most 900 kg. About 15 boxes of either product can be produced per hour. At the moment the family wants to limit their work day to 8 hours.

Write an optimization problem to maximize the profit of this company.

Solution

Firstlt, we set the company produce x boxes beer mugs and y boxes champagne glasses. Thus we can get the follow definition.

◦ **Decisions variable:** the number of boxes of beer mugs x and the umber of boxes of beer mugs y.
◦ **Objective:** maximize the profit $f(x, y) = 200x + 150y$.
◦ **Constraints:**
1)The total use of plastic resins: $9x + 5.5y \leq 900$.
2)The total boxes of beer mugs and champagne glasses: $x + y \leq (15)(8) = 120$.
3)The boxes of beer mugs and champagne glasses should greater than or equal to zero: $x \geq 0, y \geq 0$, and must be integers.

Then we solve the optimization problem by drawing the linear programming graph as Figure 1. As we can see from Figure 1, the red dash line is the profit function with variable profits. And the points in the blue area are the available decisions. So we can find the maximum of profit near the point $C(x = 68.57, y = 51.43)$. Because the boxes should be integers, finally we can find the maximum of profit as $x = 68, y = 52$, and the maximum of profit is

$$f(x, y)_{max} = 200x + 150y = (200)(68) + (150)(52) = 21400 \tag{1}$$

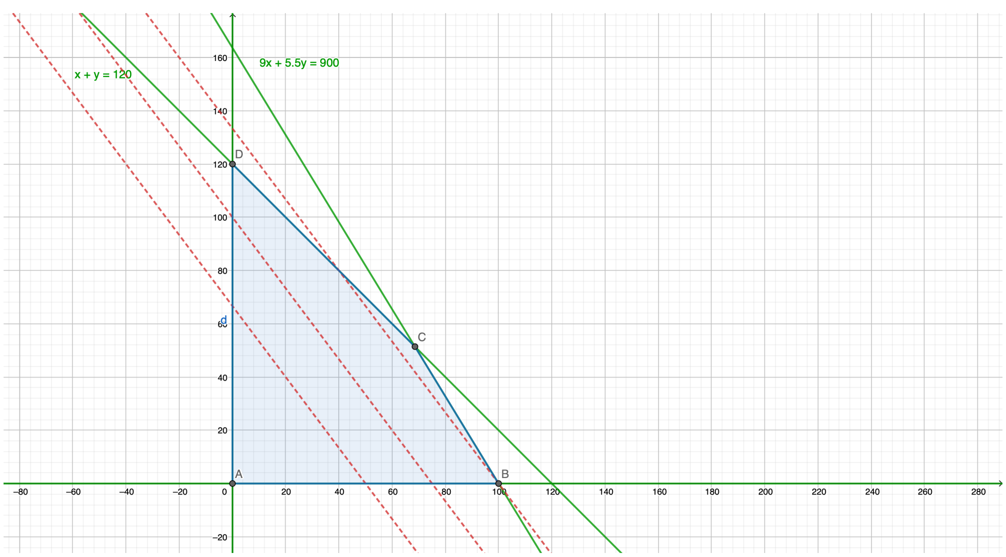


Figure 1 Linear programming of the factory profit

Assignment A1.2

A local brewery produces Ale and Beer. The production is currently limited by the available resources of corn, hops, and barley malt. To produce one box of Ale 5 kg of corn, 4 kg of hops and 35 kg of malt are required. To make one box of Beer 15 kg of corn, 4 kg of hops, and 20 kg of malt are required. Suppose that only 480 kg of corn, 160 kg of hops and 1190 kg of malt are available.

a) Formulate a linear program (= linear optimization problem) to maximize the profit of the brewery when one box of Ale achieves a profit of 130 RMB and one box of Beer achieves a profit of 230 RMB.

b) The brewery also wants to consider time constraints. The production of one box of Ale requires 1.5 hours while the production of one box of Beer requires 2.5 hours. The two production processes can not run at the same time and the brewery can only run for up to 12 hours each day.

   Revise the linear program from part a) and include these time constraints in an appropriate way to find an optimal production plan for one week of production.

Solution

Firstlt, we set the brewery produce x boxes Ale and y boxes Beer.

Subproblem a)

We can derive the definition as follow.

∘ **Decisions variable:** the number of boxes of Ale x and the umber of boxes of Beer y.
∘ **Objective:** maximize the profit $f(x, y) = 130x + 230y$.
∘ **Constraints:**
 1)The total use of corn: $5x + 15y \leq 480$.
 2)The total use of hops: $4x + 4y \leq 160$.
 3)The total use of malt: $35x + 20y \leq 1190$.
 4)The boxes of Ale and Beer should greater than or equal to zero: $x \geq 0, y \geq 0$, and must be integers.

Then we solve the optimization problem by drawing the linear programming graph as Figure 2. As we can see from Figure 2, the red dash line is the profit function with variable profits. And the points in the blue area are the available decisions. So we can find the maximum of profit with the point $D(x = 12, y = 28)$ as

$$f(x, y)_{max} = 130x + 230y = (130)(12) + (230)(28) = 8000 \tag{2}$$

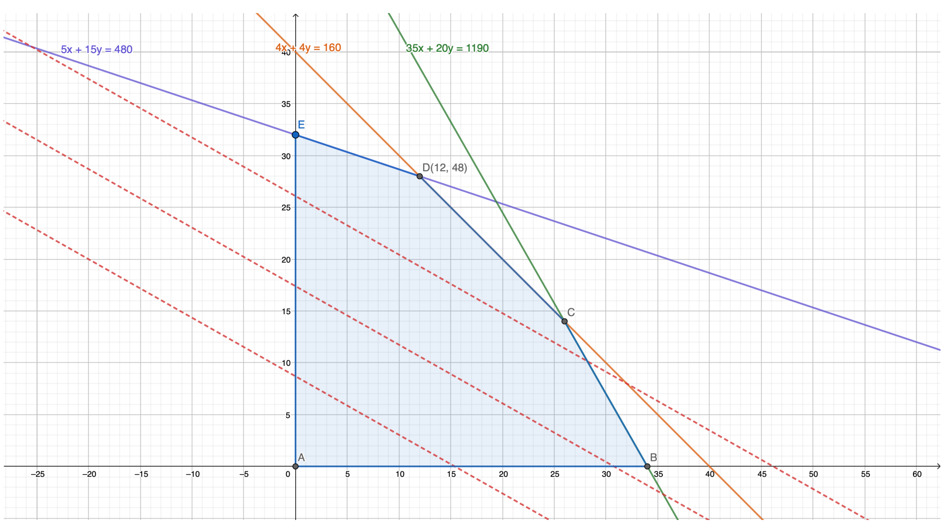


Figure 2 Linear programming of the brewery profit

Assignment A1.2 continued on next page...

Subproblem b)
With the consideration of time constraints, we should add the constraints as follow

$$1.5x + 2.5y \le (12)(7) = 84 \tag{3}$$

Then we solve the optimization problem by drawing the linear programming graph as Figure 3. As we can see from Figure 3, the red dash line is the profit function with variable profits. And the points in the blue area are the available decisions, and the pink line is the new adding time constraint. So we can find the maximum of profit with the point $E(x = 6, y = 30)$ as

$$f(x, y)_{max} = 130x + 230y = (130)(6) + (230)(30) = 7680 \tag{4}$$

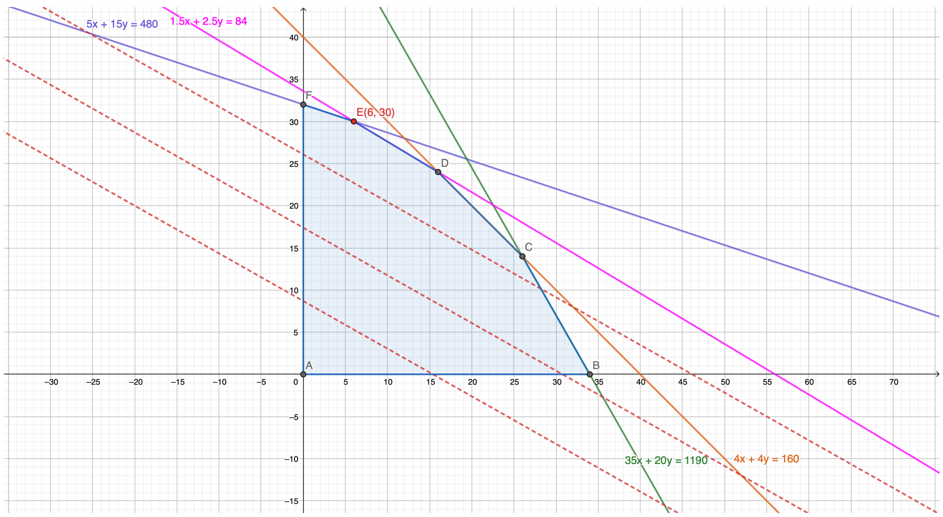


Figure 3 Linear programming with time constraint

Assignment A1.3

Classify the following matrices and verify whether they are positive definite, positive semidefinite or indefinite.

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

What are the eigenvalues of the matrices A_3 and A_4?

Solution

○ **For matrix A_1:** Because matrix A_1 is a diagonal matrix, we can get the eigenvalues of A_1 as $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$. Thus, two eigenvalues are positive, and one eigenvalue is negative, so matrix A_1 is **indefinite**.

○ **For matrix A_2:** Because matrix A_2 is a real symmetric matrix, we can use the eigenvalues of A_2 to judge the definiteness of matrix A_2.
The characteristic polynomial of A_2 can be derived as follow

$$\begin{aligned} \left|A_2 - \lambda I\right| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 2 \\ 1 & 2 & 5-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} 1-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} + (1)\begin{vmatrix} 0 & 1-\lambda \\ 1 & 2 \end{vmatrix} \\ &= \lambda(1-\lambda)(\lambda-6) \end{aligned} \tag{5}$$

Assignment A1.3 continued on next page...

Thus, the eigenvalues of A_2 can be derived as $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 6$. So $\forall \lambda_i \geq 0$, thus matrix A_2 is **positive semidefinite**.

∘ **For matrix A_3:** For $\forall \boldsymbol{x} \in \mathbb{R}^3$, we have

$$\begin{aligned}
\boldsymbol{x}^T A_3 \boldsymbol{x} &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} x_1 & x_2 + 2x_3 & x_1 - x_2 + 4x_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= x_1^2 + x_2^2 + 4x_3^2 + x_2x_3 + x_1x_3 \\
&= \left(x_1 + \frac{1}{2}x_3 \right)^2 + \left(x_2 + \frac{1}{2}x_3 \right)^2 + \frac{7}{2}x_3^2 \geq 0
\end{aligned} \tag{6}$$

Thus we can find $\boldsymbol{x}^T A_3 \boldsymbol{x} > 0$ for $\forall \boldsymbol{x} \in \mathbb{R}^3 \backslash \{\boldsymbol{0}\}$. So matrix A_3 is **positive definite**.

In order to calculate the eigenvalues of matrix A_3. We write down the characteristic polynomial of A_3 as follow

$$\begin{aligned}
\left| A_3 - \lambda I \right| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{vmatrix} \\
&= (1-\lambda) \begin{vmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{vmatrix} \\
&= (1-\lambda)(\lambda-2)(\lambda-3)
\end{aligned} \tag{7}$$

Thus, the eigenvalues of A_3 can be derived as $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$.

∘ **For matrix A_4:** For $\forall \boldsymbol{x} \in \mathbb{R}^3$, we have

$$\begin{aligned}
\boldsymbol{x}^T A_4 \boldsymbol{x} &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} x_1 - x_3 & -x_2 + x_3 & x_1 - x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= x_1^2 - x_2^2
\end{aligned} \tag{8}$$

So when $x_1 \geq x_2, \boldsymbol{x}^T A_4 \boldsymbol{x} \geq 0$; when $x_1 \leq x_2, \boldsymbol{x}^T A_4 \boldsymbol{x} \leq 0$. Thus matrix A_4 is **indefinite**.

In order to calculate the eigenvalues of matrix A_4. We write down the characteristic polynomial of A_4 as follow

$$\begin{aligned}
\left| A_4 - \lambda I \right| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & -1-\lambda & -1 \\ -1 & 1 & -\lambda \end{vmatrix} \\
&= (1-\lambda) \begin{vmatrix} -1-\lambda & -1 \\ 1 & -\lambda \end{vmatrix} + (1) \begin{vmatrix} 0 & -1-\lambda \\ -1 & 1 \end{vmatrix} \\
&= -\lambda(\lambda^2 + 1)
\end{aligned} \tag{9}$$

Thus, the eigenvalues of A_4 can be derived as $\lambda_1 = 0, \lambda_2 = i, \lambda_3 = -i$.

Assignment A1.4

Let $y^1, y^2, \ldots, y^k$ be k different and given points in $\mathbb{R}^2$, i.e., it holds that $y^i \in \mathbb{R}^2$ for all $i = 1, \ldots, k$. We want to find a circle in $\mathbb{R}^2$ with minimum radius that contains all of these points.

Formulate this problem as an optimization problem with the center and radius of the circle as optimization variables. Classify this type of optimization problem.

Solution

We set the center of the circle as $\boldsymbol{x} \in \mathbb{R}^2$, and the radius of the circle as $r \in \mathbb{R}$. Thus we can get the follow definition of the optimization problem.

- **Decisions variable:** the center of the circle $\boldsymbol{x}$, and the radius of the circle r.
- **Objective:** minimize the ridus r.
- **Constraints:** The circle should contains all the given points $y^1, y^2, \ldots, y^k$:

$$\forall i = 1, \ldots, k, \quad \left\| y^i - \boldsymbol{x} \right\| \leq r \tag{10}$$

Thus, the optimization problem can be classified as constrained, nonlinear, continuous optimization problem.

Assignment A1.5

In this exercise, we want to study and visualize different feasible sets and solve optimization problems via graphical considerations.

a) Sketch the following sets in $\mathbb{R}^2$:
   - $X_1 := \{\boldsymbol{x} \in \mathbb{R}^2 : |x_1| - |x_2| \geq 1\}$.
   - $X_2 := \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 \leq 0, (x_1 + 1)^2 + x_2^2 \geq 1\}$.
   - $X_3 := \{\boldsymbol{x} \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 \leq 1, x_1 - x_2 \leq 0, x_1 + x_2 \leq 0\}$.
   - $X_4 := \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 x_2 \geq 0, \|\boldsymbol{x}\|_\infty \geq 1\}$, where $\|\boldsymbol{x}\|_\infty := \max\{|x_1|, |x_2|\}$ denotes the maximum norm in $\mathbb{R}^2$.

b) Analyze which of the sets X_1, X_2, X_3, or X_4 is bounded and explain your answer.

c) Consider the nonlinear program

$$\min f(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in X_4 \tag{11}$$

and determine (graphically) all local and global minimizer of problem (11) for the two choices $f(\boldsymbol{x}) := x_1$ and $f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2)$. Are the minimizer strict?

d) In this part, we consider an optimization problem with equality constraints

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} f(\boldsymbol{x}) = \frac{1}{2} \|\boldsymbol{x}\|^2 \quad s.t. \quad 1 - x_1^2 + x_2 = 0 \tag{12}$$

Let us define $h : \mathbb{R}^2 \to \mathbb{R}$, $h(\boldsymbol{x}) = 1 - x_1^2 + x_2$ and let $\{\boldsymbol{x} \in \mathbb{R}^2 : h(\boldsymbol{x}) = 0\}$ denote the feasible set associated with problem (12).

   – Draw the sets $L_\alpha := \{\boldsymbol{x} \in \mathbb{R}^2 : f(\boldsymbol{x}) = \alpha\}$ for different values of α¿0 and the feasible set X and determine the solution $\boldsymbol{x}^* \in X$ of problem (12) graphically.

   Remark: The sets L_α are called *contours* or *contour lines* of f.

   – Modify your sketch and draw the gradients $\nabla f(\boldsymbol{x}^*)$and $\nabla h(\boldsymbol{x}^*)$(at the point $\boldsymbol{x}^*$). What kind of connection exists between those two vectors?

e) Give an example of a function f that has a strict local minimum but no global minimum.

Solution

Subproblem a)

The graph of X_1 is showed as Figure 4. The blue area is the feasible points of set X_1. The graph of X_2 is showed as Figure 5. The blue area is the feasible points of set X_2. The graph of X_3 is showed as Figure 6. red point A is the only feasible points] of set X_3. The graph of X_4 is showed as Figure 7. The blue area is the feasible points of set X_4.

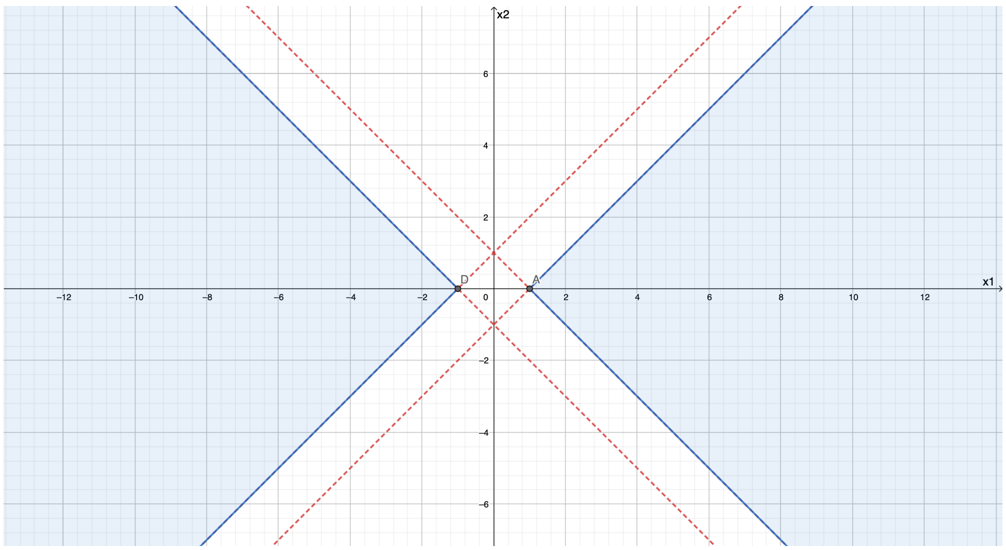


Figure 4 Graph of set X_1

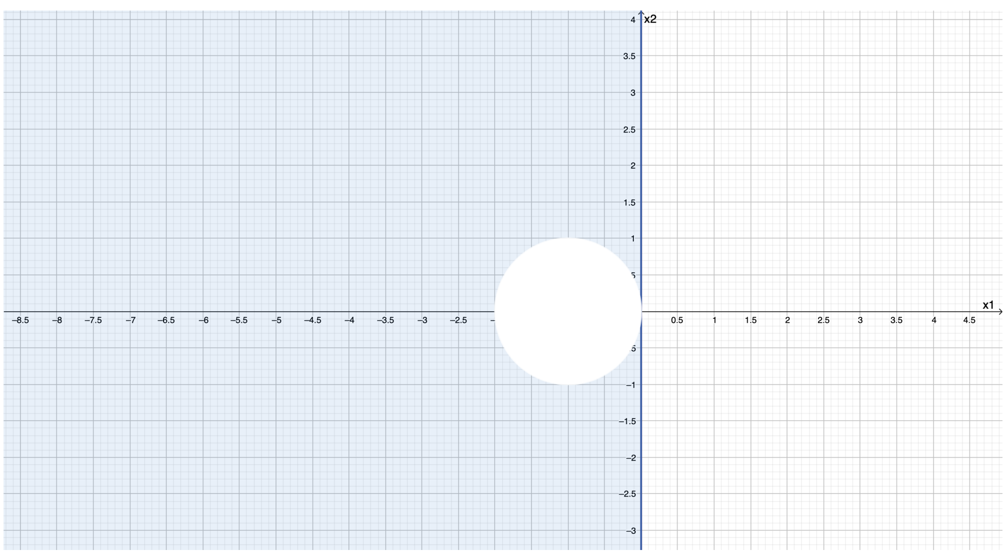


Figure 5 Graph of set X_2

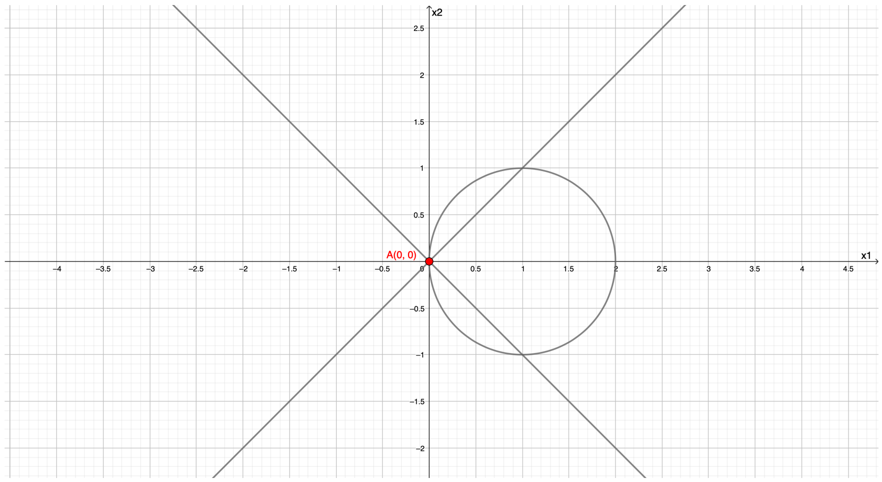


Figure 6 Graph of set $\boldsymbol{X}_3$

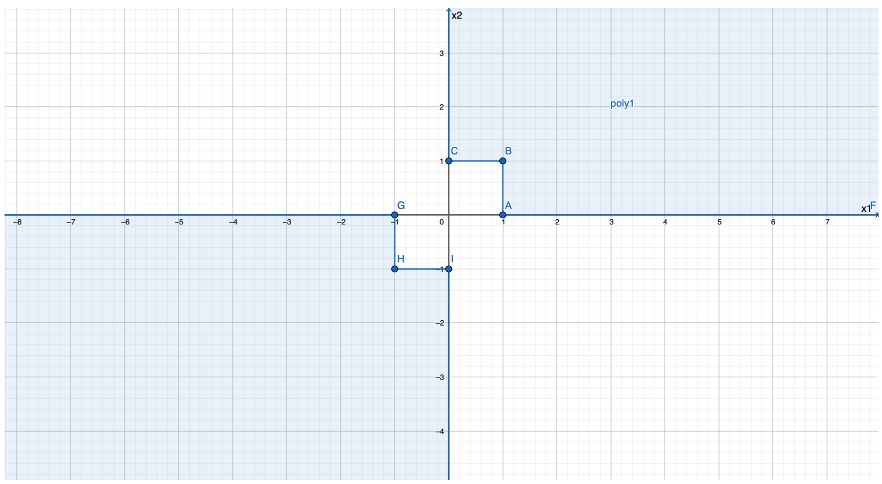


Figure 7 Graph of set $\boldsymbol{X}_4$

Subproblem b)

○ Set X_1 is not bounded, because both x_1 and x_2 can reach ∞, which means it is not bounded.
○ Set X_2 is not bounded, because x_1 can reach $-\infty$ and x_2 can reach ∞, which means it is not bounded.
○ Set X_3 is bounded, because x_1 can only be 0 and x_2 can only be 0, which means there is just one point in Set X_3, say $\boldsymbol{x}(0, 0)$. So Set X_3 is bounded.
○ Set X_4 is not bounded, because x_1 and x_2 can reach ∞, which means it is not bounded.

Subproblem c)

○ For the first choice, $f(\boldsymbol{x}) := x_1$. Because the variable x_1 can can reach all the real number continuously, $f(x)$ is not bounded (see Figure 8), which means there is not any local or global minimizer.

○ For the second choice, $f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2)$. Every time, we can set $f(\boldsymbol{x})$ as a constant, so we can draw the circle due to the equation. With the constant change, we can get different circles, the circle with the mimimum radius is what we want to find. As we can see in Figure 9, the red dash circle is drawed according to the eqaution, and the mimimum circle is the circle which goes through point A, C, G, I, so these points are global minimizer, but they are not strict global minimizer. However, they are strict local minimizer.

Thus, we can find four global minimizer $A(x_1 = 1, x_2 = 0)$, $C(x_1 = 0, x_2 = 1)$, $G(x_1 = -1, x_2 = 0)$, $I(x_1 = 0, x_2 = -1)$. None of them is strict global minimizer, but all of them are strict local minimizer.

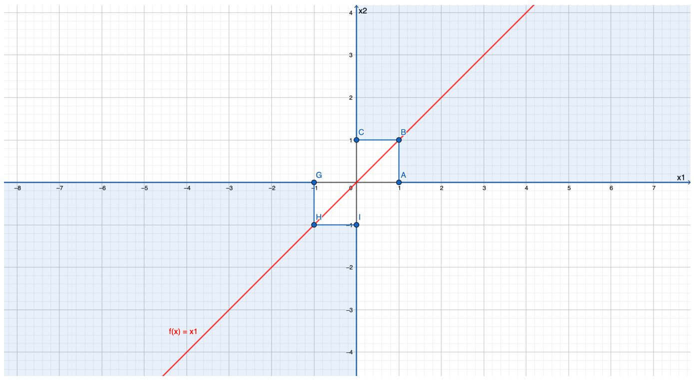


Figure 8 $f(\boldsymbol{x}) = x_1$

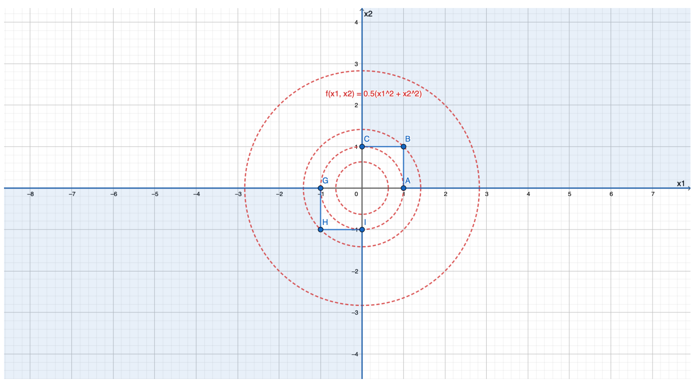


Figure 9 $f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2)$

In order to see the outcome more vividly, we draw the 3D graph of $f(\boldsymbol{x})$ as showing in Figure 10, it can help us to find the minimizer more intuitively.

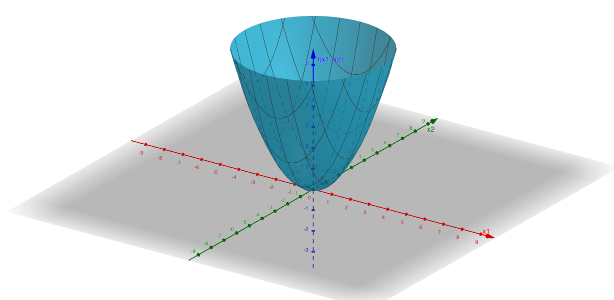


Figure 10 The 3D graph of $f(\boldsymbol{x})$

Subproblem d)

The test set L_α is showed by Figure 11, with differnt α, there are different circles (drawed by blue and green dash line) to denote the set L_α. The green circle($\alpha = \frac{3}{8}$) is tangent with the function $h(\boldsymbol{x}) = 0$(drawed by red color) at point $A\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$ and point $B\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$. Thus we can find the solution $\boldsymbol{x}^* \in X$ of problem (12) as Point $A\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$ and point $B\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$, which lead the value of $f(\boldsymbol{x})$ to be minimum.

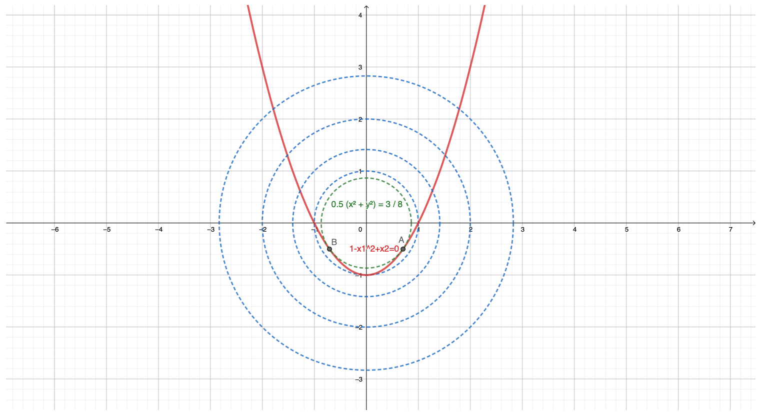


Figure 11 L_α **and** $h(x) = 0$

Because $f(\boldsymbol{x}) := \frac{1}{2}(x_1^2 + x_2^2)$ and $h(\boldsymbol{x}) = 1 - x_1^2 + x_2$, the gradient of $f(\boldsymbol{x})$ can $h(\boldsymbol{x})$can be derived as

$$\begin{aligned} \nabla f(x) &= \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ \nabla h(x) &= \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix} \end{aligned} \tag{13}$$

We have two points (A and B) as the $\boldsymbol{x}^*$.

◦ For point $A\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$, we get the gradient

$$\begin{aligned} \nabla f(A) &= \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{1}{2} \end{pmatrix} \\ \nabla h(A) &= \begin{pmatrix} -\sqrt{2} \\ 1 \end{pmatrix} \end{aligned} \tag{14}$$

We draw the two vectors in Figure 12, the yellow dash is the vector $\nabla f(A)$, and the pink dash is the vector $\nabla h(A)$. We can see that the two vectors are parallel and have opposite direction.

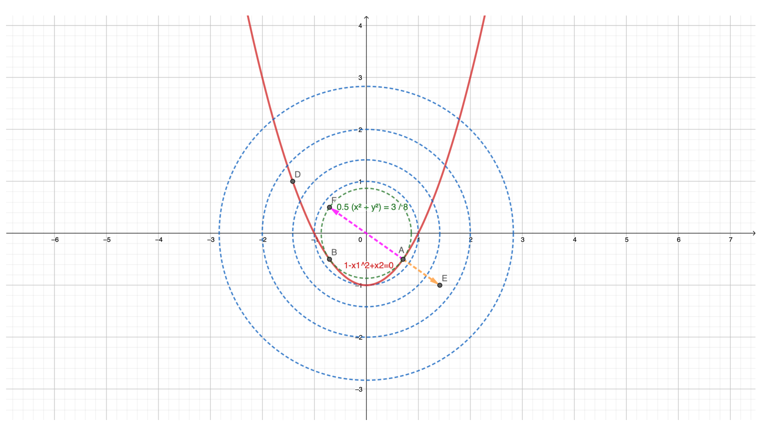


Figure 12 L_α **and** $h(x)$ **with gradients**

◦ For point $B\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$, we get the gradient

$$\begin{aligned} \nabla f(B) &= \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ -\frac{1}{2} \end{pmatrix} \\ \nabla h(B) &= \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \end{aligned} \tag{15}$$

We draw the two vectors in Figure 13, the orange dash is the vector $\nabla f(B)$, and the pink dash is the vector $\nabla h(B)$. We can see that the two vectors are parallel and have opposite direction.

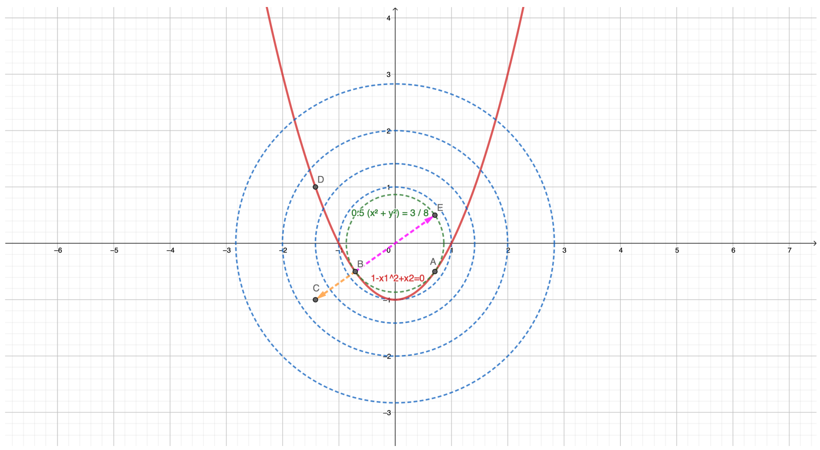


Figure 13 L_α **and** $h(x)$ **with gradients**

Thus, at the point $\boldsymbol{x}^*$, the connection between the two vectors $\nabla f(\boldsymbol{x}^*)$ and $\nabla h(\boldsymbol{x}^*)$ is that they are parallel and have opposite direction.

Subproblem e)

We give the function $f(x) = 2x^2 - \frac{1}{2}x^3, x \in \mathbb{R}$, and we draw the graph of this function. As we can see from Figure 14, there is a strict local minimizer $x = 0$, as well as the strict local minimum $f(0) = 0$. However, there is no global minimum, because when $x \to +\infty$, the value of $f(x)$ will be $-\infty$.

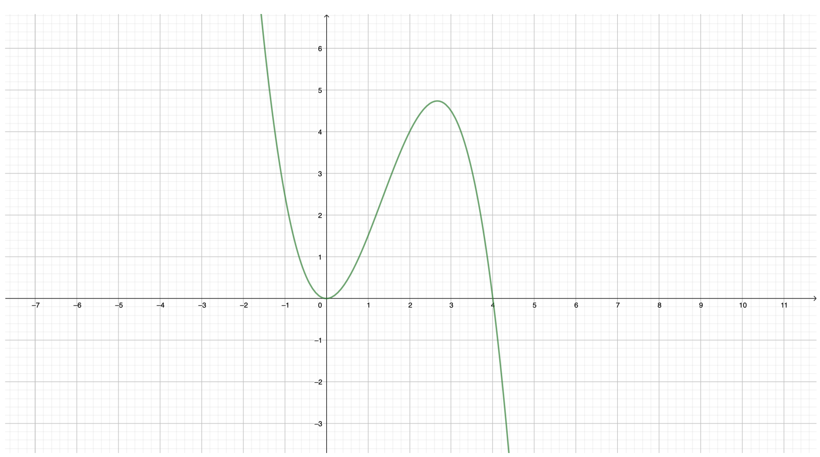


Figure 14 Graph of function $f(x) = 2x^2 - \frac{1}{2}x^3$

Assignment A1.6

In this exercise, we investigate differentiability properties of different functions.

a) Calculate the gradient and Hessian of the following mappings:

$$\begin{aligned}
&f_1 : \mathbb{R}^2 \to \mathbb{R},\ f_1(\boldsymbol{x}) = (x_1^2 - x_2)^2 + (x_1 - x_2^2)_2 - (x_1 - 1)^3 + (x_2 - 1)^3,\\
&f_2 : \mathbb{R}^2 \to \mathbb{R},\ f_2(\boldsymbol{x}) = \cos(x_1)\sin(x_2) - \frac{x_1}{1 + x_2^2},\\
&f_3 : \mathbb{R}^2 \to \mathbb{R},\ f_3(\boldsymbol{x}) = x_1^4 + 2(x_1 - x_2)x_1^2 + 4x_2^2
\end{aligned} \tag{16}$$

b) Find all points $\boldsymbol{x}^* \in \mathbb{R}^2$ such that $\nabla f_3(\boldsymbol{x}^*) = 0$. Calculate the Hessian $\nabla^2 f_3(\boldsymbol{x}^*)$ at those points and investigate the definiteness of the Hessians (i.e., decide whether $\nabla^2 f(\boldsymbol{x}^*)$ is positive (semi-)definite, negative (semi-)definite, or indefinite).

c) Write a MATLAB or Python code to create a 3D-plot (surface plot) and a (two-dimensional) contour plot of the function f_3. Include the points $\boldsymbol{x}^*$ that you have found in part b) in your plot.

Solution

Subproblem a)

○ The gradient of f_1, f_2, f_3 can be derived as follow

$$\begin{aligned}
\nabla f_1(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 4x_1^3 - 3x_1^2 - 2x_2^2 - 4x_1x_2 + 8x_1 - 3 \\ 4x_2^3 + 3x_2^2 - 2x_1^2 - 4x_1x_2 - 4x_2 + 3 \end{pmatrix}\\
\nabla f_2(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -\sin(x_1)\sin(x_2) - \frac{1}{1+x_2^2} \\ \cos(x_1)\cos(x_2) + \frac{2x_1x_2}{\left(1+x_2^2\right)^2} \end{pmatrix}\\
\nabla f_3(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial f_3}{\partial x_1} \\ \frac{\partial f_3}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 4x_1^3 + 6x_1^2 - 4x_1x_2 \\ -2x_1^2 + 8x_2 \end{pmatrix}
\end{aligned} \tag{17}$$

○ The Hessian of f_1, f_2, f_3 can be derived as follow

$$\begin{aligned}
\nabla^2 f_1(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial^2 f_1}{\partial x_1 \partial x_1} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_1}{\partial x_2 \partial x_1} & \frac{\partial^2 f_1}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} 12x_1^2 - 6x_1 - 4x_2 + 8 & -4x_2 - 4x_1 \\ -4x_1 - 4x_2 & 12x_2^2 + 6x_2 - 4x_1 - 4 \end{pmatrix}\\
\nabla^2 f_2(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial^2 f_2}{\partial x_1 \partial x_1} & \frac{\partial^2 f_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_2}{\partial x_2 \partial x_1} & \frac{\partial^2 f_2}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} -\sin(x_2)\cos(x_1) & -\sin(x_1)\cos(x_2) + \frac{2x_2}{\left(1+x_2^2\right)^2} \\ -\sin(x_1)\cos(x_2) + \frac{2x_2}{\left(1+x_2^2\right)^2} & -\cos(x_1)\sin(x_2) + \frac{2x_1\left[\left(1+x_2^2\right) - 4x_2^2\right]}{\left(1+x_2^2\right)^3} \end{pmatrix}\\
\nabla^2 f_3(\boldsymbol{x}) &= \begin{pmatrix} \frac{\partial^2 f_3}{\partial x_1 \partial x_1} & \frac{\partial^2 f_3}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_3}{\partial x_2 \partial x_1} & \frac{\partial^2 f_3}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} 12x_1^2 + 12x_1 - 4x_2 & -4x_1 \\ -4x_1 & 8 \end{pmatrix}
\end{aligned} \tag{18}$$

Subproblem b)

We set $\nabla f_3(\boldsymbol{x}^*) = \boldsymbol{0}$, so we can derive $\boldsymbol{x}^*$.

$$
\begin{aligned}
\nabla f_3(\boldsymbol{x}) &= \begin{pmatrix} 4x_1^3 + 6x_1^2 - 4x_1x_2 \\ -2x_1^2 + 8x_2 \end{pmatrix} = \mathbf{0} \\
\Rightarrow &\begin{cases} 4x_1^3 + 6x_1^2 - 4x_1x_2 = 0 \\ -2x_1^2 + 8x_2 = 0 \end{cases} \\
\Rightarrow &\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases} \text{or} \quad \begin{cases} x_1 = -2 \\ x_2 = 1 \end{cases}
\end{aligned} \tag{19}
$$

Thus, we find all $\boldsymbol{x}^*$, say point $A(0,0)$ and point $B(-2,1)$. So we can calculate the Hessian matrix at point A and B as follow

$$
\begin{aligned}
\nabla^2 f_3(A) &= \begin{pmatrix} \frac{\partial^2 f_3}{\partial x_1 \partial x_1} & \frac{\partial^2 f_3}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_3}{\partial x_2 \partial x_1} & \frac{\partial^2 f_3}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} 12x_1^2 + 12x_1 - 4x_2 & -4x_1 \\ -4x_1 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 8 \end{pmatrix} \\
\nabla^2 f_3(B) &= \begin{pmatrix} \frac{\partial^2 f_3}{\partial x_1 \partial x_1} & \frac{\partial^2 f_3}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_3}{\partial x_2 \partial x_1} & \frac{\partial^2 f_3}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} 12x_1^2 + 12x_1 - 4x_2 & -4x_1 \\ -4x_1 & 8 \end{pmatrix} = \begin{pmatrix} 20 & 8 \\ 8 & 8 \end{pmatrix}
\end{aligned} \tag{20}
$$

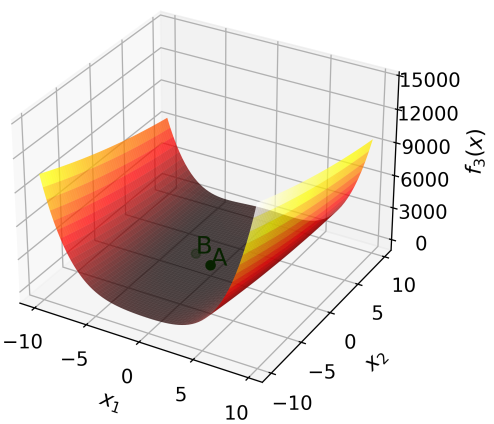
∘ **For point** A**:** because the Hessian matrix at point A is a diagonal matrix, we can get the eigenvalues of $\nabla^2 f_3(A)$ as $\lambda_1 = 0, \lambda_2 = 8$. So $\forall \lambda_i \geq 0$, thus matrix $\nabla^2 f_3(A)$ is **positive semidefinite**.

∘ **For point** B**:** because the Hessian matrix at point B, say $\nabla^2 f_3(B)$ is a real symmetric matrix, we can use the eigenvalues of $\nabla^2 f_3(B)$ to judge the definiteness of matrix $\nabla^2 f_3(B)$.
The characteristic polynomial of $\nabla^2 f_3(B)$ can be derived as follow

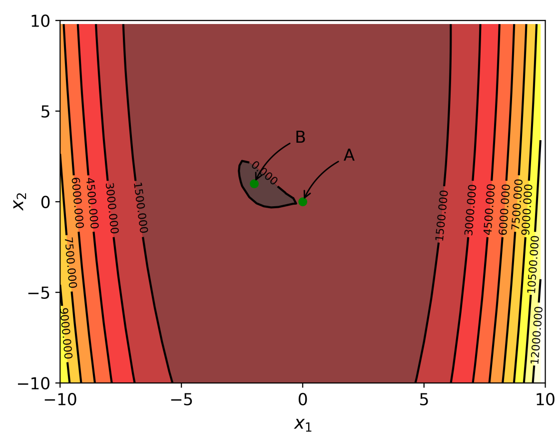
$$
\left| \nabla^2 f_3(B) - \lambda I \right| = \begin{vmatrix} 20 - \lambda & 8 \\ 8 & 8 - \lambda \end{vmatrix} = (\lambda - 4)(\lambda - 24) \tag{21}
$$

Thus, the eigenvalues of $\nabla^2 f_3(B)$ can be derived as $\lambda_1 = 4, \lambda_2 = 24$. So $\forall \lambda_i > 0$, thus matrix $\nabla^2 f_3(B)$ is **positive definite**.

Subproblem c)
By using python, we draw the 3D-plot (surface plot) and the contour plot of the function f_3 as Figure 15. Figure 15(a) is the 3D-plot and Figure 15(b) is the contour plot. Points A and B are labeled in the figure.



(a) 3D-plot



(b) Contour plot

Figure 15 3D and Contour plot of f_3