

Submission for HW 3  
CS: 427 Mathematics for Data Science  
Autumn 2020-21

K. Sai Anuroop, 170030035  
Sunday, October 18, 2020

Problem 1

Prove that a function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is convex if and only if its hessian is positive semi-definite.

Answer Assuming that  $f$  is twice differentiable and domain of  $f$  is convex.

Sketch of the proof:

① We first prove it for the case of  $f: \mathbb{R} \rightarrow \mathbb{R}$  and then

② We use the first-order condition for convexity, i.e.,  $f$  is convex if and only if  $\text{dom} f$  is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \text{ holds } \forall x, y \in \text{dom} f.$$

to prove the statement in problem.

① Let us assume  $n=1$ .

$\Rightarrow$  Suppose  $f: \mathbb{R} \rightarrow \mathbb{R}$  is convex. Let  $x, y \in \text{dom} f$  where  $y > x$ . By the first order condition

$$f(y) \geq f(x) + f'(x)(y-x) \text{ and } f(y) \leq f(x) + f'(y)(y-x)$$

$$\text{This means } f(x) + f'(x)(y-x) \leq f(y) \leq f(x) + f'(y)(y-x)$$

$$\Rightarrow f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x)$$

So we get  $f'(y)(y-x) \geq f'(x)(y-x)$

$$\Rightarrow f'(y)(y-x) - f'(x)(y-x) \geq 0$$

$$\Rightarrow \frac{f'(y) - f'(x)}{y-x} \geq 0 \quad (\text{note that } y > x).$$

Taking  $y \rightarrow x$ ,

$$\lim_{y \rightarrow x} \frac{f'(y) - f'(x)}{y-x} = f''(x)$$

$$\text{So, } \lim_{y \rightarrow x} \frac{f'(y) - f'(x)}{y-x} = f''(x) \geq 0.$$

Therefore  $f''(x) \geq 0$ .

$\Leftarrow$  Suppose  $f''(z) \geq 0 \quad \forall z \in \text{dom } f$ .

Consider any two points  $x, y \in \text{dom } f$  where  $x < y$ .

Now consider 
$$\int_x^y f''(z)(y-z) dz$$

Note that  $f''(z) \geq 0 \quad \forall z \in \text{dom } f$ .

and  $y-z \geq 0 \quad \forall z \in [x, y]$ .

$$\text{So, } 0 \leq \int_x^y f''(z)(y-z) dz$$

we now apply  $\boxed{\int u dv = uv - \int v du}$  rule to

solve the above integral.

$$\begin{aligned}
 0 &\leq \int_x^y f''(z) (y-z) dz \\
 &= \int_x^y (y-z) d(f'(z)) \\
 &= (y-z)f'(z) \Big|_x^y - \int_x^y f'(z) (-dz) \\
 &= -(y-x)f'(x) + \int_x^y f'(z) dz \\
 &= -(y-x)f'(x) + f(z) \Big|_x^y \\
 &= -(y-x)f'(x) + f(y) - f(x)
 \end{aligned}$$

$$\Rightarrow f(y) - f(x) - (y-x)f'(x) \geq 0$$

i.e.,  $f(y) \geq f(x) + f'(x)(y-x)$ .

Therefore  $f$  is convex.

□  
(end of ① proof).

② Consider that a function is convex if and only if it is convex on all lines. In other words, the function  $g(t) = f(x_0 + tv)$  is convex in  $t$  for all  $x_0 \in \text{dom } f$  and all  $v$ .

Therefore, for any  $n \geq 1$ ,  $f$  is convex if and only if  $g''(t) = v^T f''(x_0 + tv) v \geq 0$

$\forall x_0 \in \text{dom } f$ ,  $\forall \epsilon \in \mathbb{R}^n$  and  $t$  such that  $x_0 + t \epsilon \in \text{dom } f$   
 here, it is necessary and sufficient that

$$\nabla^2 f(x) \succeq 0 \quad \forall x \in \text{dom } f.$$

Therefore,  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is convex if and only if  
 if  $\nabla^2 f(x) \succeq 0$  (Hessian of  $f$  is positive  
 semi-definite). □  
(end of ② proof).

Problem 2 When is the epigraph of a function a halfspace?  
 When is the epigraph of a function a convex cone?  
 When is the epigraph of a function a polyhedron?

Answer An epigraph of a function, denoted by  $\text{epi } f$  is defined as follows:  $\{ (x, t) \mid f(x) \leq t \}$

$$\text{epi } f := \{ (x, t) : x \in \mathbb{R}^n, t \in \mathbb{R}, f(x) \leq t \}$$

strictly speaking

$$\boxed{\text{epi } f := \{ (x, t) \mid x \in \text{dom } f, f(x) \leq t \}}$$

A halfspace is a set of the form  $\{x \mid a^T x \leq b\}$   
 where  $x \in \mathbb{R}^n$ ,  $a (\neq 0) \in \mathbb{R}^n$ ,  $b \in \mathbb{R}$ .



Suppose epigraph is a cone.

$\Rightarrow \forall a > 0$ , if  $(x, t) \in \text{epi } f$  then  $(ax, at) \in \text{epi } f$

clearly,  $(x, f(x)) \in \text{epi } f$ .

So,  $(ax, af(x)) \in \text{epi } f$ .

$$\Rightarrow f(ax) \leq af(x) \quad \text{---(i)}$$

Similarly,  $(ax, f(ax)) \in \text{epi } f$ . So, if the epigraph

is a cone,  $(x, f(ax)/a) \in \text{epi } f$

$$\text{i.e., } f(x) \leq f(ax)/a \quad \text{---(ii)}$$

By combining (i) and (ii) we get

$$af(x) \leq f(ax) \leq af(x).$$

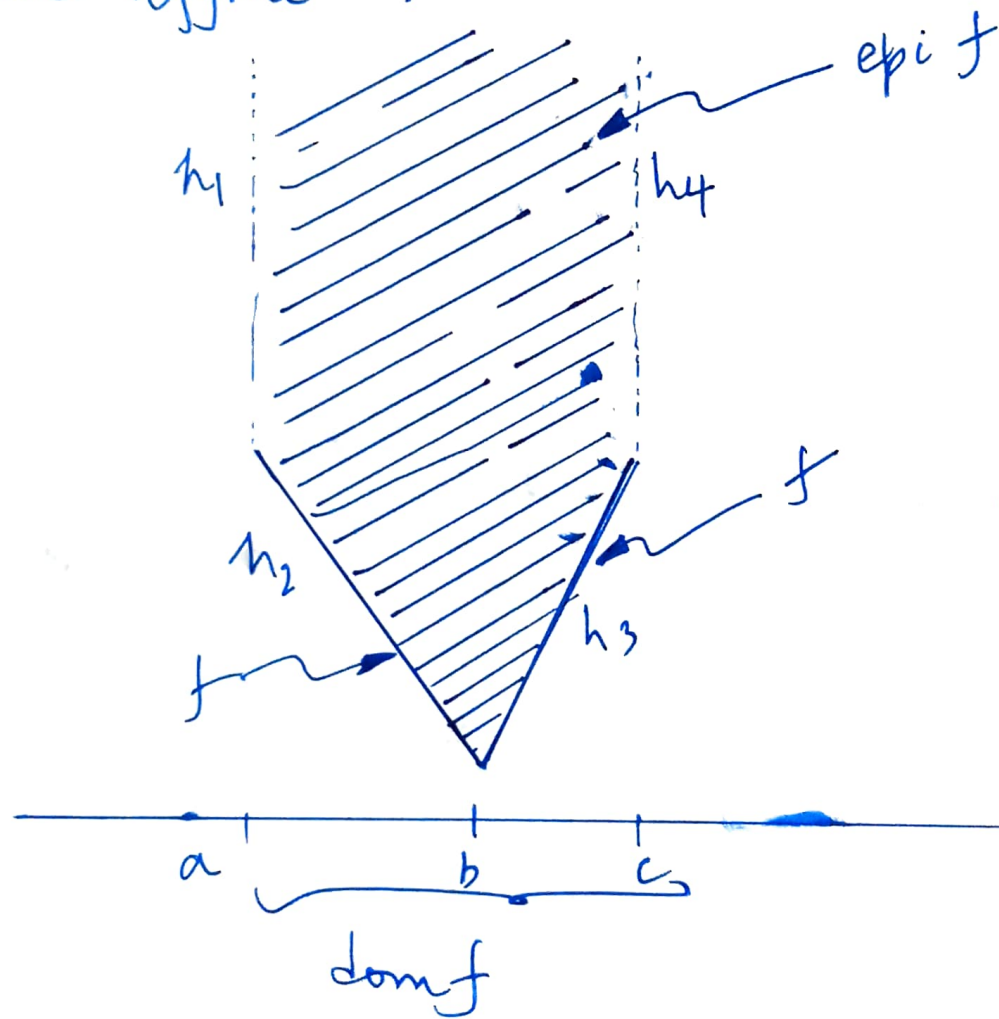
$\Rightarrow f(ax) = af(x)$ . So this says  $f$  is positive homogeneous.

$\therefore$  Epigraph of a function is a convex cone if  $f$  is positively homogeneous.

---

A polyhedron is the intersection of a finite number of halfspaces and hyperplanes.

We just proved that the epigraph of  $f$  is a halfspace if  $f$  is affine. Thus for the epigraph to be a polyhedron, we require that  $f$  be 'piecewise affine'.



Note in this figure that  $f$  is piecewise affine. The shaded region is  $\text{epi } f$  which is a polyhedron too. The halfspaces which form epigraph are:  $h_1, h_2, h_3$  &  $h_4$  in the figure.

Problem 3 A function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is monotone if  $\forall x, y \in \mathbb{R}^n$   
 $(f(x) - f(y))^T (x - y) \geq 0$ .

Show that gradient  $\nabla f$  of  $f$  is monotone.

Is every monotone mapping a gradient of some convex function?

Answer the question is slightly incorrect. It should have been as follows:

[ A function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is monotone if  $\forall x, y \in \mathbb{R}^n$   
 $(f(x) - f(y))^T (x - y) \geq 0$ .

Suppose  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  is a differentiable convex function.

Show that  $\nabla g$  is monotone. Is every monotone mapping the gradient of some convex function? ]

Answer

Since  $g$  is convex, we have

$$g(x) \geq g(y) + \nabla g(y)^T (x - y) \text{ and}$$

$$g(y) \geq g(x) + \nabla g(x)^T (y - x), \text{ for any } x, y \in \mathbb{R}^n$$

(by first order convexity).

Combining the above two inequalities we get

$$(\nabla g(x) - \nabla g(y))^T (x - y) \geq 0 \text{ which says}$$

$\nabla g$  is monotone.

for the other part of the question, consider

$$\phi(x) = \begin{bmatrix} x_1 \\ x_1/2 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

claim  
 $\phi(x)$  is monotone.

$$\phi(x) = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\phi(y) = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\phi(x) - \phi(y) = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 - y_1 \\ \frac{x_1 - y_1}{2} + x_2 - y_2 \end{bmatrix}$$

$$(\phi(x) - \phi(y))^T (x - y) = \begin{bmatrix} x_1 - y_1 & \frac{x_1 - y_1}{2} + x_2 - y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix} \begin{bmatrix} x_1 - y_1 & \frac{x_1 - y_1}{2} + x_2 - y_2 \end{bmatrix}$$
$$= (x_1 - y_1)^2 + (x_2 - y_2)^2 + \frac{(x_2 - y_2)(x_1 - y_1)}{2}$$

$$\geq \frac{(x_1 - y_1)^2}{4} + \frac{(x_2 - y_2)^2}{4} + \frac{(x_2 - y_2)(x_1 - y_1)}{2}$$

$$= \left[ \frac{(x_1 - y_1) + (x_2 - y_2)}{2} \right]^2$$

$$\geq 0.$$

□

(P.T.O)



† We know that for a given field  $F = (F_1, \dots, F_n)$  which is smooth ( $C^1$ ) to be gradient, it is necessary that  $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ ,  $1 \leq i < j \leq n$ .  
 Let us check if  $\phi$  satisfies this condition.

$$\frac{\partial \phi_1}{\partial x_2} = \frac{\partial x_1}{\partial x_2} = 0$$

$$\frac{\partial \phi_2}{\partial x_1} = \frac{\partial (x_1/2 + x_2)}{\partial x_1} = \frac{1}{2}$$

$$\Rightarrow \frac{\partial \phi_1}{\partial x_2} \neq \frac{\partial \phi_2}{\partial x_1}$$

$\therefore \phi$  cannot be a gradient.

So, we have shown a counter example where a monotone function is not the gradient of a convex function.

---

† This comes from the fact that curl of a gradient is 0

Problem 4 Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is convex,  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  is concave,  $\text{dom } f = \text{dom } g = \mathbb{R}^n$ , and for all  $x$ ,  $g(x) \leq f(x)$ . Show that there exists an affine function  $h$  such that for all  $x$ ,  $g(x) \leq h(x) \leq f(x)$ .  
In other words, if a concave function  $g$  is an under-estimator of a convex function  $f$ , then we can fit an affine function between  $f$  and  $g$ .

Answer Recall that  $\text{epi } f := \{(x, t) : x \in \text{dom } f, t \geq f(x)\}$   
 $\text{hypo } f := \{(x, t) : x \in \text{dom } f, t \leq f(x)\}$

Notice that  $\text{int}(\text{epi } f) \neq \emptyset$  (interior of  $\text{epi } f$  is not empty)

as  $\text{dom } f = \mathbb{R}^n$ .

Also,  $\text{int}(\text{epi } f) \cap \text{hypo } g = \emptyset$  as

$f(x) < t$  for  $(x, t) \in \text{int } \text{epi } f$  and

$t \geq g(x)$  for  $(x, t) \in \text{hypo } g$ .

Thus, these two sets ( $\text{int}(\text{epi } f)$  and  $\text{hypo } g$ )

can be separated by a hyperplane. This means

there exist  $a \in \mathbb{R}^n$ ,  $b \in \mathbb{R}$  ( ~~$a \neq 0$  or  $b \neq 0$~~ )

(where both  $a$  and  $b$  are not zero at the same time),  
and  $c \in \mathbb{R}$  such that

$$a^T x + b t \geq c \geq a^T y + b v \quad \text{where}$$

$$t \geq f(x) \text{ and } v \leq g(y).$$

Note that if  $b=0$ ,  $a^T x \geq a^T y \Rightarrow a=0$ .  
 $\forall x, y$

So,  $b \neq 0$ . Let  $x=y$ .

Then,  $b t \geq c \geq b v$ .

$$\Rightarrow b t \geq b v \quad - (i)$$

But we have  $t > f(x)$  and  $v \leq g(x)$   
and  $g(x) \leq f(x)$

$$\text{So, } v \leq g(x) \leq f(x) < t$$

$$\Rightarrow t > v \quad - (ii)$$

Plugging (ii) in (i) we get,

$$b(t-v) \geq 0$$

$$\Rightarrow \underline{\underline{b > 0}}$$

Let us now separate  $\text{int}(\text{epi } f)$  and  $\text{hypo } g$  by a hyperplane.

Consider a point  ~~$(x, t) \in \text{int}(\text{epi } f)$~~

$(x, t) \in \text{int}(\text{epi } f)$  and

$(y, v) = (x, g(x)) \in \text{hypo } g$ .

Then  $a^T x + b t \geq c \geq a^T x + b g(x)$ .

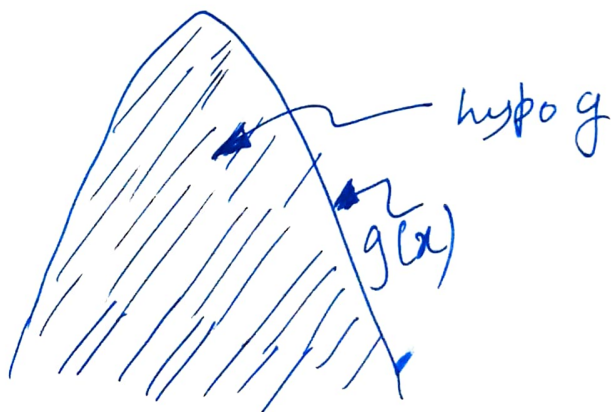
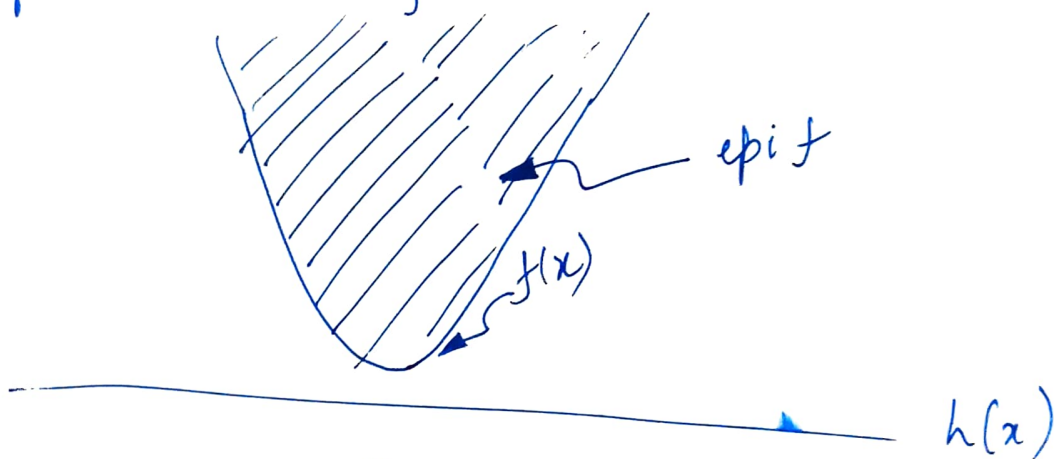
Dividing by  $b$ . (note that we proved  $b > 0$ )

$$\frac{a^T x}{b} + t \geq \frac{c}{b} \geq \frac{a^T x}{b} + g(x)$$

$$\Rightarrow \boxed{t \geq -\frac{a^T x + c}{b} \geq g(x) \quad \forall t \geq f(x)}$$

Thus, we have  $h(x) = \frac{c - a^T x}{b}$ , an affine function which lies between  $f$  and  $g$ .

A representation of the theorem is as follows:





# Submission for HW 3 (Programming Assignment)

*CS 427: Mathematics for Data Science, Autumn 2020-21*

K. Sai Anuroop, 170030035

October 18, 2020

## 1 Question 1

Plot a 3D graph and a contour map of  $f(x, y) = x^2 - y^2 \forall x, y \in [-5, 5]$

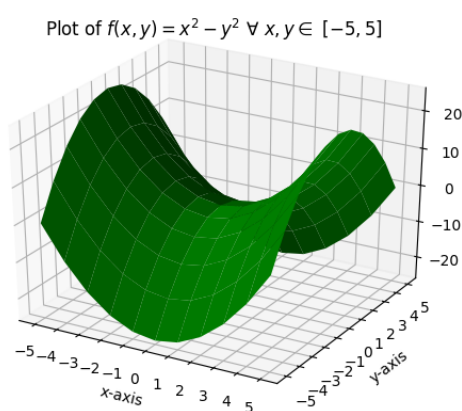


Figure 1: 3D graph

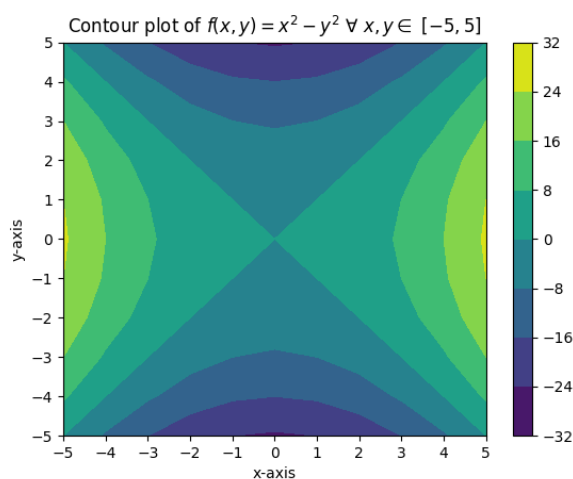


Figure 2: Contour map

## 2 Question 2

Randomly generate a set of 24 points that belong to the set  $\{(x, y) : x, y \in [-5, 5]\}$ . Create a scatter plot and outline the convex hull of the set you just created.

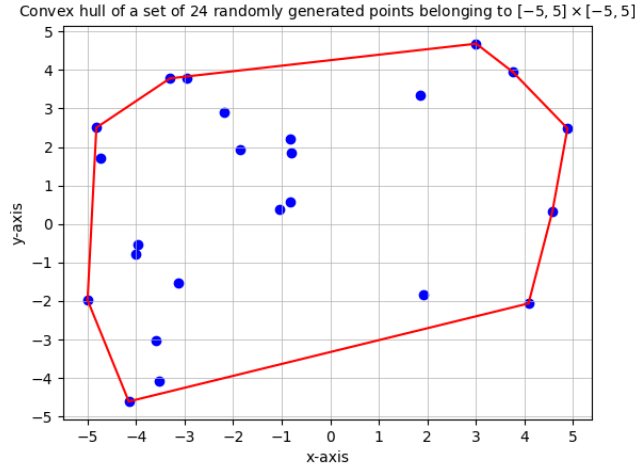


Figure 3: Scatter plot and convex hull

## 3 Question 3

Check if the function  $f(x) = x^T A x$  for  $A \in R^{2 \times 2}$  where all components of  $x$  are integers in  $[-10, 10]$ , is convex. Find 11 counter examples if it is not.

I take a random vector  $x$  where  $x_{ij} \in [-10, 10]$  for  $i, j = \{1, 2\}$ . I then rotate and scale it using the matrix  $A := \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ . Here  $f(x) = x^T A x = (x^2 + y^2) \cos \theta + 2xy \sin \theta$ . There is a result which states that

$f(x) = x^T A x$  is convex if and only if  $A$  is positive semidefinite. By the way I defined the matrix  $A$ , it is symmetric. However, it can be analysed that for some values of  $\theta$ ,  $f$  takes negative values. So I provide 11 such values of  $A$  for which  $f$  is not convex. For better visualisation, I plotted the rotated and scaled vector  $Ax$  for some values of  $A$ . I also plot the corresponding surfaces of  $f$ .

Here are the counter-examples in terms of the matrix  $A$ :

$$\begin{pmatrix} 0.27 & 0.96 \\ 0.96 & 0.27 \end{pmatrix} \begin{pmatrix} 0.17 & 0.99 \\ 0.99 & 0.17 \end{pmatrix} \begin{pmatrix} 0.07 & 1.0 \\ 1.0 & 0.07 \end{pmatrix} \begin{pmatrix} -0.03 & 1.0 \\ 1.0 & -0.03 \end{pmatrix} \begin{pmatrix} -0.13 & 0.99 \\ 0.99 & -0.13 \end{pmatrix} \begin{pmatrix} -0.23 & 0.97 \\ 0.97 & -0.23 \end{pmatrix} \begin{pmatrix} -0.32 & 0.95 \\ 0.95 & -0.32 \end{pmatrix} \\ \begin{pmatrix} -0.42 & 0.91 \\ 0.91 & -0.42 \end{pmatrix} \begin{pmatrix} -0.5 & 0.86 \\ 0.86 & -0.5 \end{pmatrix} \begin{pmatrix} -0.59 & 0.81 \\ -0.59 & 0.81 \end{pmatrix} \begin{pmatrix} -0.67 & 0.75 \\ 0.75 & -0.67 \end{pmatrix}$$

Clearly as seen in the below plots,  $f$  loses its convexity as we vary the value of  $\theta$  from 0 to  $\pi$ .

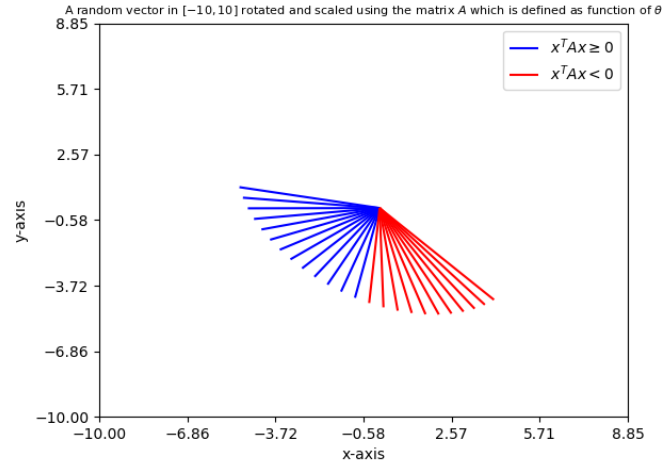


Figure 4: Rotated and scaled vectors  $Ax$  for different matrices  $A$

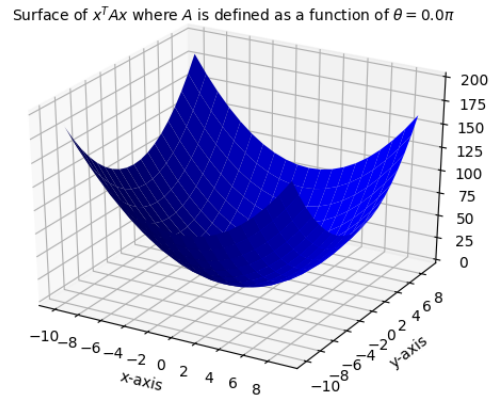


Figure 5: Surface of  $f$  for  $\theta = 0$

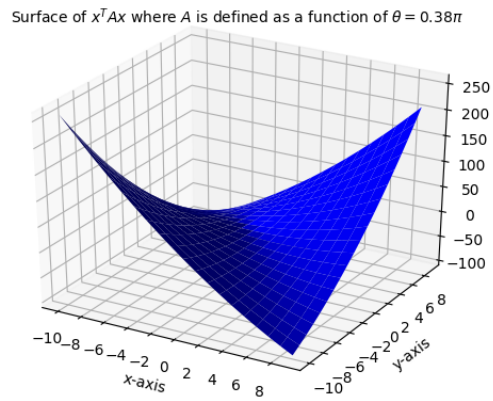


Figure 6: Surface of  $f$  for  $\theta = 0.38\pi$

Surface of  $x^T A x$  where  $A$  is defined as a function of  $\theta = 0.64\pi$

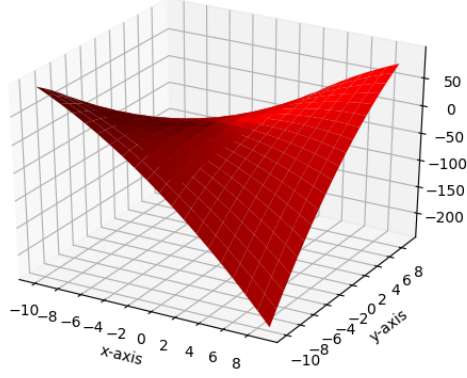


Figure 7: Surface of  $f$  for  $\theta = 0.64\pi$

Surface of  $x^T A x$  where  $A$  is defined as a function of  $\theta = 0.73\pi$

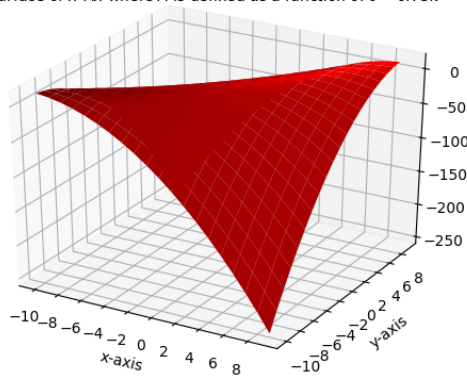


Figure 8: Surface of  $f$  for  $\theta = 0.73\pi$





Scan this QR code to access the GitHub repository of my homework solutions at  
[https://github.com/ksanu1998/MDS\\_HW\\_Solutions](https://github.com/ksanu1998/MDS_HW_Solutions)