ANLY561 Assignment4

Hongyang Zheng

In [1]:

```
'''
This code imports numpy packages and allows us to pass data from python to globa
l javascript
objects. It was developed by znah@github
'''

import json
import numpy as np
import numpy.random as rd
from ipywidgets import widgets
from IPython.display import HTML, Javascript, display

def json_numpy_serializer(o):
    if isinstance(o, np.ndarray):
        return o.tolist()
    raise TypeError("{} of type {} is not JSON serializable".format(repr(o), typ
e(o)))

def jsglobal(**params):
    code = [];
    for name, value in params.items():
        jsdata = json.dumps(value, default=json_numpy_serializer)
        code.append("window.{}={};".format(name, jsdata))
    display(Javascript("\n".join(code)))
```

In [2]:

```
%%javascript

// Loading the compiled MathBox bundle.
require.config({
    //baseUrl:'', paths: {mathBox: '../../tree/static/mathbox/build/mathbox-bund
le'}
    // online compilation
    baseUrl: '', paths: {mathBox: '../static/mathbox/build/mathbox-bundle'}
    // online compilation without local library-- remove baseUrl
    //paths: {mathBox: '//cdn.rawgit.com/unconed/mathbox/eaeb8e15/build/mathbox-
bundle'}
});

// Minified graphing functions

window.with_mathbox=function(element,func){require(['mathBox'],function(){var ma
thbox=mathBox({plugins:['core','controls','cursor','mathbox'],controls:{klass:TH
REE.OrbitControls},mathbox:{inspect:!1},element:element[0],loop:{start:!1},});va
r three=mathbox.three;three.renderer.setClearColor(new THREE.Color(0xFFFFFF),1.0
);three.camera.position.set(-1,1,2);three.controls.noKeys=!0;three.element.style
.height="400px";three.element.style.width="100%";function isInViewport(element){
var rect=element.getBoundingClientRect();var html=document.documentElement;var w
=window.innerWidth||html.clientWidth;var h=window.innerHeight||html.clientHeight
;return rect.top<h&&rect.left<w&&rect.bottom>0&&rect.right>0}
var intervalId=setInterval(function(){if(three.element.offsetParent===null){clea
rInterval(intervalId);three.destroy();return}
var visible=isInViewport(three.canvas);if(three.Loop.running!=visible){visible?t
hree.Loop.start():three.Loop.stop()}},100);func(mathbox);window.dispatchEvent(ne
w Event('resize'))})};window.plotGraph=function(mathbox,f,xlabel='x',ylabel='y',
zlabel='f(x,y)',rng=[[-3,3],[-5,5],[-3,3]]){var view=mathbox.cartesian({range:rn
g,scale:[1,1,1]},{rotation:(t)=>[0,t*0.02,0]}).grid({axes:[1,3]})
view.area({id:'yaxis',width:1,height:1,axes:[1,3],expr:function(emit,x,y,i,j){em
it(4,0,0);emit(0,0,0)},items:2,channels:3,}).text({font:'Helvetica',style:'bold'
,width:16,height:5,depth:2,expr:function(emit,i,j,k,time){emit(ylabel)},}).label
({color:'#000000',snap:!1,outline:2,size:24,offset:[0,-32],depth:.5,zIndex:1});v
iew.vector({points:'#yaxis',color:0x000000,width:9,start:!0});view.area({id:'xax
is',width:1,height:1,axes:[1,3],expr:function(emit,x,y,i,j){emit(0,0,4);emit(0,0
,0)},items:2,channels:3,}).text({font:'Helvetica',style:'bold',width:16,height:5
,depth:2,expr:function(emit,i,j,k,time){emit(xlabel)},}).label({color:'#000000',
snap:!1,outline:2,size:24,offset:[0,-32],depth:.5,zIndex:1,});view.vector({point
s:'#xaxis',color:0x000000,width:9,start:!0,});view.area({id:'zaxis',width:1,heig
ht:1,axes:[1,3],expr:function(emit,x,y,i,j){emit(0,4,0);emit(0,0,0)},items:2,cha
nnels:3,}).text({font:'Helvetica',style:'bold',width:16,height:5,depth:2,expr:fu
nction(emit,i,j,k,time){emit(zlabel)},}).label({color:'#000000',snap:!1,outline:
2,size:24,offset:[0,-32],depth:.5,zIndex:1,});view.vector({points:'#zaxis',color
:0x000000,width:9,start:!0,});var graph=view.area({id:'graph',width:64,height:64
,axes:[1,3],expr:function(emit,y,x,i,j){emit(y,f(x,y),x)},items:1,channels:3,});
view.surface({shaded:!0,lineX:!0,lineY:!0,points:graph,color:0x0000FF,width:1
,});return view};window.addSegment=function(view,p0,p1,col){view.array({width:12
8,expr:function(emit,i,time){var b=i/128;var a=1-b;emit(a*p0[1]+b*p1[1],a*p0[2]+
b*p1[2],a*p0[0]+b*p1[0])},channels:3,});view.line({color:col,width:10,size:2.5,s
troke:'dotted',start:!1,end:!1,})};window.addPoint=function(view,p,col,label){vi
ew.array({width:4,items:2,channels:3,expr:function(emit,i,t){emit(p[1],p[2],p[0
])},}).point({color:col,points:'<',size:15,depth:.5,zBias:50,}).text({font:'Helv
etica',style:'bold',width:16,height:5,depth:2,expr:function(emit,i,j,k,time){emi
t(label)},}).label({color:col,snap:!1,outline:2,size:24,offset:[0,-32],depth:.5,
zIndex:1,})};window.addCurve=function(view,ab,x,y,z,col){view.array({width:128,e
xpr:function(emit,i,time){var t=(ab[1]-ab[0])*(i/128)+ab[0];emit(y(t),z(t),x(t
))},channels:3,});view.line({color:col,width:20,size:2.5,start:!0,end:!0,})};win
```

```
dow.addClosedCurve=function(view,ab,x,y,z,col){view.array({width:128,expr:functi
on(emit,i,time){var t=(ab[1]-ab[0])*(i/128)+ab[0];emit(y(t),z(t),x(t))},channels
:3,});view.line({color:col,width:20,size:2.5,start:!1,end:!1,})};window.addSurfa
ce=function(view,ab,cd,x,y,z,col,opa){view.matrix({width:64,height:64,expr:funct
ion(emit,i,j,time){var p=(ab[1]-ab[0])*(i/64)+ab[0];var q=(cd[1]-cd[0])*(j/64)+c
d[0];emit(y(p,q),z(p,q),x(p,q))},items:1,channels:3}).surface({shaded:!0,lineX:!
1,lineY:!1,color:col,width:1,opacity:opa})}
window.addSequence=function(view,seq,col){var idx=0;var d=new Date();var start=d
.getTime();view.array({width:1,expr:function(emit,i,time){var nd=new Date();var
now=nd.getTime();if(1000<now-start){idx=idx+1;if(seq.length<=idx){idx=0}
start=now}
emit(seq[idx][1],seq[idx][2],seq[idx][0])},items:1,channels:3}).point({color:col
,points:'<',size:15,depth:.5,zBias:50,})}
```

Question 1

(1)

$f(x, y) = x^2 + y^2$ is strictly convex.

(2)

$f(x, y) = x^2$ is strictly convex.

(3)

$f(x, y) = x^2 - y^2$ is not convex.

(4)

$f(x, y) = -x^2$ is not convex.

(5)

$f(x, y) = -x^2 - y^2$ is not convex.

In [3]:

```
%%javascript

with_mathbox(element, function(mathbox) {

    var fcn = function(x, y) {
      return Math.sqrt(x*x + y*y);
    };

    var view = plotGraph(mathbox, fcn);
})
```

In [4]:

```
%%javascript

with_mathbox(element, function(mathbox) {

    var fcn = function(x, y) {
        return (x*x);
    }

    var view = plotGraph(mathbox, fcn)
})
```

In [5]:

```
%%javascript

with_mathbox(element, function(mathbox) {

    var fcn = function(x, y) {
        return (x*x-y*y);
    }

    var view = plotGraph(mathbox, fcn)
})
```

In [6]:

```
%%javascript

with_mathbox(element, function(mathbox) {

    var fcn = function(x, y) {
        return (-x*x);
    }

    var view = plotGraph(mathbox, fcn)
})
```

In [7]:

```
%%javascript

with_mathbox(element, function(mathbox) {

    var fcn = function(x, y) {
        return (-x*x-y*y);
    }

    var view = plotGraph(mathbox, fcn)
})
```

Question 2

(1)

$\partial_1 f(x, y) = 2x$, $\partial_2 f(x, y) = 2y$

choose any x_1, x_2, y_1, y_2 such that $(x_1, x_2) \neq (y_1, y_2)$, then :

$$\begin{gathered} f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\ = {x_1}^2 + {x_2}^2 - {y_1}^2 - {y_2}^2 - 2y_1(x_1 - y_1) - 2y_2(x_2 - y_2) \\ = {x_1}^2 + {x_2}^2 + {y_1}^2 + {y_2}^2 - 2y_1x_1 - 2y_2x_2 \\ = (x_1 - y_1)^2 + (x_2 - y_2)^2 \\ > 0 \end{gathered}$$

Therefore, $f(x_1, x_2) > f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2)$, the first order conditions for strict convexity holds. $f(x, y)$ is strictly convex.

(2)

$\partial_1 f(x, y) = 2x$, $\partial_2 f(x, y) = 0$

choose any x_1, x_2, y_1, y_2 such that $(x_1, x_2) \neq (y_1, y_2)$, then :

$$\begin{aligned} & f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\ =& {x_1}^2 + 0 - {y_1}^2 - 0 - 2y_1(x_1 - y_1) - 0 \\ =& {x_1}^2 + {y_1}^2 - 2y_1x_1 \\ =& (x_1 - y_1)^2 \\ \geqslant& 0 \end{aligned}$$

Therefore, $f(x_1, x_2) \geqslant f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2)$, the first order conditions for convexity holds. $f(x, y)$ is convex.

(3)

$\partial_1 f(x, y) = 2x$, $\partial_2 f(x, y) = -2y$

choose any x_1, x_2, y_1, y_2 such that $(x_1, x_2) \neq (y_1, y_2)$, then:

$$\begin{aligned} & f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\ =& {x_1}^2 - {x_2}^2 - {y_1}^2 + {y_2}^2 - 2y_1(x_1 - y_1) + 2y_2(x_2 - y_2) \\ =& {x_1}^2 - {x_2}^2 + {y_1}^2 - {y_2}^2 - 2y_1x_1 + 2y_2x_2 \\ =& (x_1 - y_1)^2 - (x_2 - y_2)^2 \end{aligned}$$

Therefore, $f(x_1, x_2) \geqslant f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2)$ does not hold for any x_1, x_2, y_1, y_2, the first order conditions for convexity fails. $f(x, y)$ is not convex.

(4)

$\partial_1 f(x, y) = -2x$, $\partial_2 f(x, y) = 0$

choose any x_1, x_2, y_1, y_2 such that $(x_1, x_2) \neq (y_1, y_2)$, then:

$$
\begin{aligned}
&f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\
=& - {x_1}^2 + 0 + {y_1}^2 - 0 + 2y_1(x_1 - y_1) - 0 \\
=& - {x_1}^2 - {y_1}^2 + 2y_1 x_1 \\
=& - (x_1 - y_1)^2 \\
\leqslant& 0
\end{aligned}
$$

Therefore, the first order conditions for convexity fails. $f(x, y)$ is not convex.

(5)

$\partial_1 f(x, y) = -2x$, $\partial_2 f(x, y) = -2y$

choose any x_1, x_2, y_1, y_2 such that $(x_1, x_2) \neq (y_1, y_2)$, then :

$$
\begin{aligned}
&f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\
=& - ({x_1}^2 + {x_2}^2 - {y_1}^2 - {y_2}^2 - 2y_1(x_1 - y_1) - 2y_2(x_2 - y_2)) \\
=& - ({x_1}^2 + {x_2}^2 + {y_1}^2 + {y_2}^2 - 2y_1 x_1 - 2y_2 x_2) \\
=& - ((x_1 - y_1)^2 + (x_2 - y_2)^2) \\
\leqslant& 0
\end{aligned}
$$

Therefore, the first order conditions for convexity fails. $f(x, y)$ is not convex.

Question 3

(a)

First, we know that $f(x, y)$ is defined at $(0, 0)$, since $f(0, 0) = 0$

Let $f_1 = y^2$, $f_2 = x^2 + y^2$, and $f_3 = \sqrt{x}$, and f_1, f_2, f_3 are continuous, so $f_3(f_2)$ is continuous.

Therefore, f is continuous because it is a ratio of two continuous functions $f_1, f_3(f_2)$.

Then, compute

$$
\lim_{(x,y)\to(0,0)} \frac{y^2}{\sqrt{(x^2 + y^2)}}
$$

.

Since

$$
0 < \frac{y^2}{\sqrt{(x^2 + y^2)}} \leqslant \frac{y^2}{\sqrt{y^2}} \leqslant |y|
$$

When $(x, y) \to (0, 0), y \to 0$, $|y| \to 0$, therefore, $\frac{x^2}{\sqrt{(x^2+y^2)}} \to 0$, when $(x, y) \to (0, 0)$

As a result, $f(x, y)$ is continuous at $(0, 0)$

(b)

$$
\begin{aligned}
g(t) =& f(at, bt) = \frac{(bt)^2}{\sqrt{(at)^2 + (bt)^2}} \\
=& \frac{b^2 t^2}{\sqrt{(a^2 + b^2)t^2}} \\
=& \frac{b^2 t^2}{|t|\sqrt{(a^2 + b^2)}} \\
=& \frac{b^2}{\sqrt{(a^2 + b^2)}}|t|
\end{aligned}
$$

Let $x, y \in \mathbb{R}$ and $m \in [0, 1]$, then we have

$$
\begin{aligned}
& g((1-m)x + my) \\
=& \frac{b^2}{\sqrt{(a^2 + b^2)}}|(1-m)x + my| \\
\leqslant& \frac{b^2}{\sqrt{(a^2 + b^2)}}(|(1-m)x| + |my|) \\
=& \frac{b^2}{\sqrt{(a^2 + b^2)}}((1-m)|x| + m|y|) \\
=& (1-m)g(x) + mg(y)
\end{aligned}
$$

Therefore, $g(t) = \frac{b^2}{\sqrt{(a^2+b^2)}}|t|$ is convex according to the definition of convex.

(c)

Let$(x_1, x_2) = (0, -5)$ and $(y_1, y_2) = (3, -4)$,$\partial_1 f(x, y) = \frac{-y^2 x}{(x^2+y^2)^{\frac{3}{2}}}$, $\partial_2 f(x, y) = \frac{2yx^2+y^3}{(x^2+y^2)^{\frac{3}{2}}}$

Then calculate:

$$
\begin{aligned}
& f(x_1, x_2) - f(y_1, y_2) - \partial_1 f(y_1, y_2)(x_1 - y_1) - \partial_2 f(y_1, y_2)(x_2 - y_2) \\
=& 5 - \frac{16}{5} - \frac{-48}{125} * (-3) - \frac{-136}{125} * (-1) \\
=& 5 - \frac{400}{125} - \frac{144}{125} - \frac{136}{125} \\
=& -\frac{55}{125} \\
<& 0
\end{aligned}
$$

Since the first order condition is not hold for $(x_1, x_2) = (0, -5)$ and $(y_1, y_2) = (3, -4)$, $f(x, y)$ is not convex on $\mathbb{R}$.

Question 4

First, calculate

$$det\begin{pmatrix} 1+x_1^2 & x_1x_2 \\ x_1x_2 & 1+x_2^2 \end{pmatrix}$$
$$= 1+x_1^2+x_2^2+x_1^2x_2^2-x_1^2x_2^2$$
$$= 1+x_1^2+x_2^2$$

Then, $f(x_1, x_2) = -log(1+x_1^2+x_2^2)$

$$\partial_1 f(x_1, x_2) = \frac{-2x_1}{1+x_1^2+x_2^2}$$

,

$$\partial_2 f(x_1, x_2) = \frac{-2x_2}{1+x_1^2+x_2^2}$$

$$\partial_{1,1} f(x_1, x_2) = \frac{-2+2x_1^2-2x_2^2}{(1+x_1^2+x_2^2)^2}$$

,

$$\partial_{1,2} f(x_1, x_2) = \frac{4x_1x_2}{(1+x_1^2+x_2^2)^2}$$

$$\partial_{2,1} f(x_1, x_2) = \frac{4x_1x_2}{(1+x_1^2+x_2^2)^2}$$

,

$$\partial_{2,2} f(x_1, x_2) = \frac{-2-2x_1^2+2x_2^2}{(1+x_1^2+x_2^2)^2}$$

$f(1,1) = ln3$, $\partial_1 f(0,0) = \partial_2 f(0,0) = \frac{-2}{3}$, $\partial_{1,1} f(0,0) = \partial_{2,2} f(0,0) = \frac{-2}{9}$, $\partial_{1,2} f(0,0) = \partial_{2,1} f(0,0) = \frac{4}{9}$

Thus the second order Taylor approximation to f at (0,0) is:

$$ln3 + \frac{-2}{3}(x_1-1) + \frac{-2}{3}(x_2-1) + \frac{1}{2}(\frac{-2}{9}(x_1-1)^2 + 2*\frac{4}{9}(x_1-1)(x_2-1) + \frac{-2}{9}(x_2-1)^2)$$
$$ln3 - \frac{2}{3}(x_1-1) - \frac{2}{3}(x_2-1) - \frac{1}{9}(x_1-1)^2 + \frac{4}{9}(x_1-1)(x_2-1) - \frac{1}{9}(x_2-1)^2$$

Question 5

(a)

Since A, B are both symmetric matrix, $A+B$ is also symmetric.

$$\boldsymbol{x^T}(A+B)\boldsymbol{x}$$
$$= \boldsymbol{x^T}(A\boldsymbol{x}+B\boldsymbol{x})$$
$$= \boldsymbol{x^T}A\boldsymbol{x}+\boldsymbol{x^T}B\boldsymbol{x}$$

Since $A, B \in SPD(2)$, $\boldsymbol{x^T}A\boldsymbol{x}$ and $\boldsymbol{x^T}B\boldsymbol{x}$ are greater or equal to 0. Then the sum of them is greater or equal to 0, which implies that $A+B \in SPD(2)$.

(b)

Choose $C, D \in SPD(2)$ and $t \in [0,1]$. Let $tC = M$ and $(1-t)D = N$, M, N are also symmetric matrix.

$$\begin{aligned}
&\boldsymbol{x}^{\boldsymbol{T}}(tC)\boldsymbol{x} \\
&= t\boldsymbol{x}^{\boldsymbol{T}}C\boldsymbol{x} \\
&\geqslant 0
\end{aligned}$$

$$\begin{aligned}
&\boldsymbol{x}^{\boldsymbol{T}}((1-t)D)\boldsymbol{x} \\
&= (1-t)\boldsymbol{x}^{\boldsymbol{T}}D\boldsymbol{x} \\
&\geqslant 0
\end{aligned}$$

Therefore, $M, N \in SPD(2)$.

According to part (a), since $M, N \in SPD(2)$, $M + N \in SPD(2)$. That is $tC + (1-t)D \in SPD(2)$.

Thus, $SPD(2)$ is a convex subset of $M_{2,2}$

(c)

Let $a, b, c, d \in \mathbb{R}$

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then

$$X^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

$$X^T X = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}$$

Since $a^2 + c^2, ab + cd, b^2 + d^2 \in \mathbb{R}$, $X^T X \in M_{2,2}$. And $X^T X$ is a symmetric matrix.

$$\begin{aligned}
&\boldsymbol{x}^{\boldsymbol{T}}(X^T X)\boldsymbol{x} \\
&= (x_1 x_2)\begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= (a^2 + c^2)x_1^2 + 2(ab + cd)x_1 x_2 + (b^2 + d^2)x_2^2 \\
&= (ax_1 + bx_2)^2 + (cx_1 + dx_2)^2 \\
&\geqslant 0
\end{aligned}$$

Therefore, $X^T X \in SPD(2)$

(d)

Since A is positive semidefinite and B is positive definite, A, B are both symmetric matrix, $A + B$ is also symmetric.

$$\begin{aligned}
&\boldsymbol{x}^{\boldsymbol{T}}(A + B)\boldsymbol{x} \\
&= \boldsymbol{x}^{\boldsymbol{T}}(A\boldsymbol{x} + B\boldsymbol{x}) \\
&= \boldsymbol{x}^{\boldsymbol{T}}A\boldsymbol{x} + \boldsymbol{x}^{\boldsymbol{T}}B\boldsymbol{x}
\end{aligned}$$

Since $\boldsymbol{x}^{\boldsymbol{T}}A\boldsymbol{x} \geqslant 0$ and $\boldsymbol{x}^{\boldsymbol{T}}B\boldsymbol{x} > 0$, $\boldsymbol{x}^{\boldsymbol{T}}A\boldsymbol{x} + \boldsymbol{x}^{\boldsymbol{T}}B\boldsymbol{x} > 0$, which implies that $A + B$ is positive definite.

(e)

Since A is positive definite, according to Sylvester's Criterion, $det(A) > 0$.

Since $det(A) \neq 0$, A is invertible.