

IE510 Applied Nonlinear Programming Homework 1

Instructor name: Ruoyu Sun

Student name: Wenxiao Yang

Warning: The work you are submitting for this homework assignment must be your own, and suspiciously similar homework submitted by multiple individuals may be reported to the University for investigation.

1 Reading

- Reading: Textbook Section 1.1
- Appendix A.

2 Problems

Note: You need to justify your answer for all questions (you cannot just answer yes or no for the question without a proof or analysis).

1. (15 + 5 points) Consider the sequence defined as

$$\mathbf{x}^{k+1} = A\mathbf{x}^k, k = 1, 2, \dots,$$

where $A \in \mathbb{R}^{2 \times 2}$, $\mathbf{x}^k \in \mathbb{R}^2$. In questions a) to c), The contraction/non-expansive mappings are defined for the Euclidean norm ¹.

a) Let the matrix $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where $a, b \in \mathbb{R}$. Find the sufficient and necessary condition C1 on a, b so that $f(x) = Ax$ is a contraction mapping. Find the sufficient and necessary condition C2 on a, b so that $f(x) = Ax$ is a nonexpansive mapping.

b) For the matrix defined in a), find the sufficient and necessary condition under which the sequence $\{\mathbf{x}^k\}$ converges for any initial point \mathbf{x}^k . Is it the same as C1 or C2 or none of them?

c) Now consider another matrix $A = \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix}$. Is $f(\mathbf{x}) = A\mathbf{x}$ a non-expansive mapping (under 2-norm)? Does the sequence $\{\mathbf{x}^k\}$ converge?

d) (bonus 5 points) For the matrix in c), can you find a norm of \mathbb{R}^2 so that $f(\mathbf{x}) = A\mathbf{x}$ is a contraction mapping with respect to this norm?

Hint 1: For any positive-definite matrix B , $\|\mathbf{x}\|_B \triangleq \sqrt{\mathbf{x}^T B \mathbf{x}}$ is a norm.

Hint 2: You may use computer to help the search.

¹Note: if the inequality $\|f(x) - f(y)\|_* \leq \gamma \|x - y\|_*$ where the constant $\gamma < 1$ holds for some norm $\|\cdot\|_*$ such as ℓ_1 norm or ℓ_2 norm, f is called a contraction mapping with respect to the norm $\|\cdot\|_*$.

1. a) $C_1: \max\{|a|, |b|\} < 1$

$f(x) = Ax$ is a contraction mapping

$$\Leftrightarrow \exists \gamma < 1. \text{ s.t. } \|f(x) - f(y)\| \leq \gamma \|x - y\| \quad \forall x, y$$

$$\text{Since } \|f(x) - f(y)\| = \|A(x-y)\| = \left\| \begin{bmatrix} a(x-y_1) \\ b(x-y_2) \end{bmatrix} \right\| = \sqrt{a^2(x_1-y_1)^2 + b^2(x_2-y_2)^2}$$

$$\exists \gamma < 1, \|f(x) - f(y)\| \leq \gamma \|x - y\| = \gamma \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}, \quad \forall x, y$$

$$\Leftrightarrow \exists \gamma < 1, \sqrt{a^2(x_1-y_1)^2 + b^2(x_2-y_2)^2} \leq \gamma \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2} \quad \forall x, y \Leftrightarrow \max\{|a|, |b|\} < 1.$$

$C_2: \max\{|a|, |b|\} \leq 1$.

$f(x) = Ax$ is a nonexpansive mapping

$$\Leftrightarrow \exists \gamma \leq 1. \text{ s.t. } \|f(x) - f(y)\| \leq \gamma \|x - y\| \quad \forall x, y$$

$$\text{Since } \|f(x) - f(y)\| = \sqrt{a^2(x_1-y_1)^2 + b^2(x_2-y_2)^2}$$

$$\exists \gamma \leq 1, \|f(x) - f(y)\| \leq \gamma \|x - y\|, \quad \forall x, y$$

$$\Leftrightarrow \exists \gamma \leq 1, \sqrt{a^2(x_1-y_1)^2 + b^2(x_2-y_2)^2} \leq \gamma \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2} \quad \forall x, y \Leftrightarrow \max\{|a|, |b|\} \leq 1.$$

(b) Sequence $\{x^k\}$ converges for any initial point x' .

$\Leftrightarrow \{x^k\}$ is Cauchy Sequence $\forall x' \in \mathbb{R}^2$.

\Leftrightarrow Given $\varepsilon > 0$, $\exists N_\varepsilon$ s.t. $\|x^{k+1} - x^m\| = \|A^k x' - A^m x'\| < \varepsilon$, $\forall k \geq m \geq N_\varepsilon$

$$\|A^k x' - A^m x'\| = \|(A^k - A^m)x'\| = \left\| \begin{bmatrix} a^k - a^m & 0 \\ 0 & b^k - b^m \end{bmatrix} x' \right\|$$

$$= \sqrt{(a^k - a^m)^2(x'_1)^2 + (b^k - b^m)^2(x'_2)^2} < \varepsilon \quad \forall x' \in \mathbb{R}^2, k \geq m \geq N_\varepsilon.$$

$$\Leftrightarrow \sqrt{2 \max\{(a^k - a^m)^2, (b^k - b^m)^2\}} x'^2 < \varepsilon \quad \forall x' \in \mathbb{R}$$

$\Leftrightarrow \forall \delta > 0$, $\exists N_8$ s.t. $\max\{|a^k - a^m|, |b^k - b^m|\} < \delta$, $\forall k \geq m \geq N_8$

$\Leftrightarrow \max\{|a|, |b|\} < 1$. the same as C_1 .

(c) $\|f(x) - f(y)\| = \|Ax - Ay\|$

$$\text{Let } x - y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \|f(x) - f(y)\| = \left\| \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \right\| = \sqrt{1.25} > 1 = \|x - y\|.$$

$\Rightarrow f(x)$ is not a non-expansive mapping.

We set $A_k = A^k = \begin{bmatrix} a_{11,k} & a_{12,k} \\ a_{21,k} & a_{22,k} \end{bmatrix}$, $k \geq 1$.

$$A_{k+1} = A \cdot A^k = \begin{bmatrix} 0.5 a_{11,k} + a_{21,k} & 0.5 a_{12,k} + a_{22,k} \\ 0.5 a_{21,k} & 0.5 a_{22,k} \end{bmatrix} = \begin{bmatrix} a_{11,k+1} & a_{12,k+1} \\ a_{21,k+1} & a_{22,k+1} \end{bmatrix}$$

$$\Rightarrow a_{21,k} = (0.5)^{k-1} \cdot 0 = 0 \quad a_{22,k} = (0.5)^{k-1} \cdot 0.5 = (0.5)^k$$

$$\Rightarrow a_{11,k} = (0.5)^k,$$

$$a_{12,k+1} = 0.5 a_{12,k} + (0.5)^k = \frac{a_{12,k}}{2^k} + \frac{k}{2^k} = \frac{k+1}{2^k}$$

$$x^{k+1} = A^k x^1 = \begin{bmatrix} (0.5)^k x_1 + \frac{k+1}{2^k} x_2 \\ (0.5)^k x_2 \end{bmatrix}$$

$$\lim_{k \rightarrow \infty} x^{k+1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Hence $\{x^k\}$ converges to 0

(d) Test $\|x\|_B \triangleq x^T B x$ firstly.

Contraction mapping: $\|Ax - Ay\|_B = \|A(x-y)\|_B < \|x-y\|_B$

$$\text{Let } z = x-y = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{R}^2, \quad Az = \begin{bmatrix} 0.5 z_1 + z_2 \\ 0.5 z_2 \end{bmatrix}$$

$$\text{Set the PI matrix } B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

$$\|Az\|_B = b_1 (0.5 z_1 + z_2)^2 + b_2 (0.5 z_2)^2$$

$$\|z\|_B = b_1 z_1^2 + b_2 z_2^2$$

$\|A\mathbf{z}\|_B < \|\mathbf{z}\|_B$ holds for all $\mathbf{z} \in \mathbb{R}^2$.

$$b_1(0.25z_1^2 + z_2^2 + z_1z_2 - z_1^2) - b_2(0.75z_2^2) < 0.$$

$$b_1 \frac{-0.75z_1^2 + z_2^2 + z_1z_2}{0.75z_1^2} - b_2 < 0.$$

$$b_1 \left(-\left(\frac{z_1}{z_2}\right)^2 + \frac{4}{3} + \frac{4}{3}\left(\frac{z_1}{z_2}\right) \right) - b_2 < 0.$$

Since $-\left(\frac{z_1}{z_2}\right)^2 + \frac{4}{3} + \frac{4}{3}\left(\frac{z_1}{z_2}\right) < \frac{16}{9}$

Hence $b_1 = 1$ $b_2 = 3$ can make the inequality holds for all $\mathbf{z} \in \mathbb{R}^2$

Hence $B = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ $\|\cdot\|_B$ is the norm we want.

2. $f(x, y) = (x+1)^2 + x^2y^2$

$$(1) \nabla f(x, y) = [2x+2+2xy^2, 2x^2y]^T$$

$$\nabla^2 f(x, y) = \begin{bmatrix} 2+2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}$$

$$\det([2+2y^2]) > 0$$

$\det(\nabla^2 f(x, y)) = 4x^2 - 12x^2y^2$ its sign depends on the value of x, y , so the function is neither convex nor concave.

$$(2) \nabla f(x, y) = 0 \Rightarrow \begin{cases} 2x+2+2xy^2 = 0 \\ 2x^2y = 0 \end{cases} \Rightarrow \begin{cases} x = -1 \\ y = 0 \end{cases}$$

$(-1, 0)$ is the unique stationary point.

(3) $(-1, 0)$ is the global minimum

$$f(x,y) = (x+1)^2 + x^2y^2 \geq 0 \quad \forall x,y \in \mathbb{R}$$

$f(-1, 0) = 0$, $(-1, 0)$ is the only stationary point.

$\Rightarrow (-1, 0)$ is the global-min

(4) No, $f(x,y) = (x+1)^2 + x^2y^2 \rightarrow +\infty$ as $x \rightarrow +\infty$

$$3. \nabla f(x,y) = [2x + \beta y + 1, 2y + \beta x + 2]^T$$

$$\nabla f(x, y) = 0 \Rightarrow \begin{cases} 2x + \beta y + 1 = 0 \\ 2y + \beta x + 2 = 0 \end{cases}$$

$$\text{if } \beta \neq \pm 2 \quad \begin{cases} x = \frac{2\beta - 2}{4 - \beta^2} \\ y = \frac{\beta - 4}{4 - \beta^2} \end{cases} \quad \text{if } \beta = \pm 2, \text{ there is no stationary point}$$

Set of all stationary points: $S = \left\{ \left(\frac{2\beta-2}{4-\beta^2}, \frac{\beta-4}{4-\beta^2} \right) \right\}$, $\beta \neq \pm 2$

$$\nabla^2 f(x,y) = \begin{bmatrix} 2 & \beta \\ \beta & 2 \end{bmatrix}$$

Eigenvalues of $\nabla^2 f(x,y)$: $(2-\lambda)^2 - \beta^2 = 0 \Rightarrow \lambda_1 = 2-\beta, \lambda_2 = 2+\beta$

If $\beta \in (-2, 2)$, $D^2f(x, y)$ is PD

the stationary point $(\frac{2\beta-2}{4-\beta^2}, \frac{\beta-4}{4-\beta^2})$ is global-min

Otherwise there is no global-min.

$$4.(a) \nabla f(x,y) = [4x^3 - 16x, 2y]^T$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 12x^2 - 16 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\nabla^2 f(x,y) \succeq 0 \iff x \geq \frac{2}{\sqrt{3}}, x \leq -\frac{2}{\sqrt{3}}$$

$$\nabla f(x,y) = 0 \Rightarrow \begin{cases} 4x^3 - 16x = 0 \\ 2y = 0 \end{cases} \Rightarrow \begin{cases} x=0 \\ y=0 \end{cases} \quad \begin{cases} x=2 \\ y=0 \end{cases} \quad \begin{cases} x=-2 \\ y=0 \end{cases}$$

Since $f(x,y) = (x^2 - 4)^2 + y^2 \geq 0 \forall x,y$

$f(2,0) = f(-2,0) = 0$, hence $(2,0), (-2,0)$ are global-min

For $(0,0)$, $\nabla^2 f(0,0) = \begin{bmatrix} -6 & 0 \\ 0 & 2 \end{bmatrix}$ which is neither PD nor ND, so $(0,0)$ is neither a local-max nor a local-min.

$$(b) f(x,y) = \frac{1}{2}x^2 + x \cos y.$$

$$\nabla f(x,y) = [x + \cos y, -x \sin y]^T$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 1 & -\sin y \\ -\sin y & -x \cos y \end{bmatrix}$$

$$\nabla f(x,y) = 0 \Rightarrow \begin{cases} x + \cos y = 0 \\ -x \sin y = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x=0 \\ y=k\pi + \frac{\pi}{2} \end{cases} \quad \begin{cases} x=-1 \\ y=2k\pi \end{cases} \quad \begin{cases} x=1 \\ y=(2k+1)\pi \end{cases} \quad k \in \mathbb{Z}$$

$$\det(1) = 1 > 0$$

$$\det \left(\begin{bmatrix} 1 & -\sin y \\ -\sin y & -x \cos y \end{bmatrix} \right) = -x \cos x - \sin^2 y$$

$$\nabla^2 f(x,y) \text{ is PD} \iff -x \cos x - \sin^2 y > 0.$$

When $\begin{cases} x=0 \\ y=k\pi + \frac{\pi}{2} \end{cases}$, $\nabla^2 f(x,y)$ is not PD.

When $\begin{cases} x=-1 \\ y=2k\pi \end{cases}$, $|-1| > 0$, $\nabla^2 f(x,y)$ is PD.

When $\begin{cases} x=1 \\ y=(2k+1)\pi \end{cases}$, $|1| > 0$, $\nabla^2 f(x,y)$ is PD

Hence the set of all local-min is

$$\{(-1, 2k\pi), (1, (2k+1)\pi), \forall k \in \mathbb{Z}\}.$$

$$(c) f(x,y) = \sin x + \sin y + \sin(x+y)$$

$$\nabla f(x,y) = [\cos x + \cos(x+y), \cos y + \cos(x+y)]^T$$

$$\nabla^2 f(x,y) = \begin{bmatrix} -\sin x - \sin(x+y) & -\sin(x+y) \\ -\sin(x+y) & -\sin y - \sin(x+y) \end{bmatrix}$$

$$\nabla f(x,y) = 0 \Rightarrow \begin{cases} \cos x + \cos(x+y) = 0 \\ \cos y + \cos(x+y) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x=\pi \\ y=\pi \end{cases}, \begin{cases} x=\frac{\pi}{3} \\ y=\frac{\pi}{3} \end{cases}, \begin{cases} x=\frac{5}{3}\pi \\ y=\frac{5}{3}\pi \end{cases}$$

$$\nabla^2 f(\pi, \pi) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0 \Rightarrow (\pi, \pi) \text{ can be local-min or local-max}$$

$$\nabla^2 f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \begin{bmatrix} -\sqrt{3} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\sqrt{3} \end{bmatrix} < 0 \Rightarrow \left(\frac{\pi}{3}, \frac{\pi}{3}\right) \text{ is local-max}$$

$$\nabla^2 f\left(\frac{5}{3}\pi, \frac{5}{3}\pi\right) = \begin{bmatrix} \sqrt{3} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \sqrt{3} \end{bmatrix} > 0 \Rightarrow \left(\frac{5}{3}\pi, \frac{5}{3}\pi\right) \text{ is local-min}$$

$$(d) f(x,y) = (y-x^2)^2 - x^2 = x^4 - 2yx^2 + y^2 - x^2$$

$$\nabla f(x,y) = [4x^3 - 4yx - 2x, -2x^2 + 2y]^T$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 12x^2 - 4y - 2 & -4x \\ -4x & 2 \end{bmatrix}$$

$$\nabla f(x,y) = 0 \Rightarrow \begin{cases} 4x^3 - 4yx - 2x = 0 \\ -2x^2 + 2y = 0 \end{cases} \Rightarrow \begin{cases} x=0 \\ y=0 \end{cases}$$

(0, 0) is the only stationary point

$$\nabla^2 f(0,0) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \text{ which is neither PD nor NI.}$$

$\Rightarrow (0,0)$ is neither local-max nor local-min.

(e)

We can find a C s.t. the $S = \{x \mid f(x,y) \leq C, \forall y \in [-1,1]\}$ is non-empty.

Since $y \in (-1,1)$, $f(x,y) \rightarrow +\infty$ as $x \rightarrow +\infty$,
the S is compact.

\Rightarrow then the global-min of $f(x,y), x \in \mathbb{R}, y \in [-1,1]$ exists.

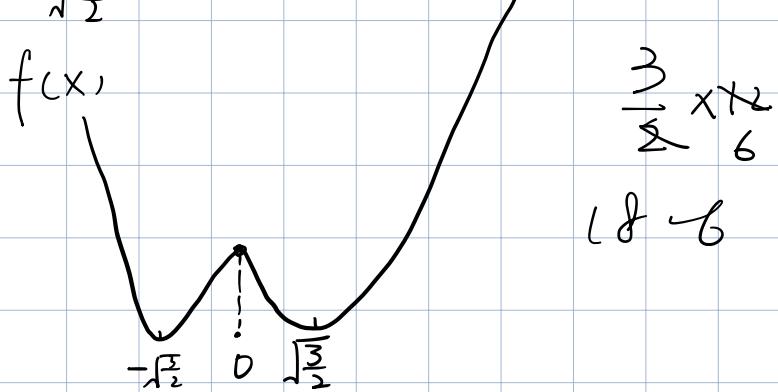
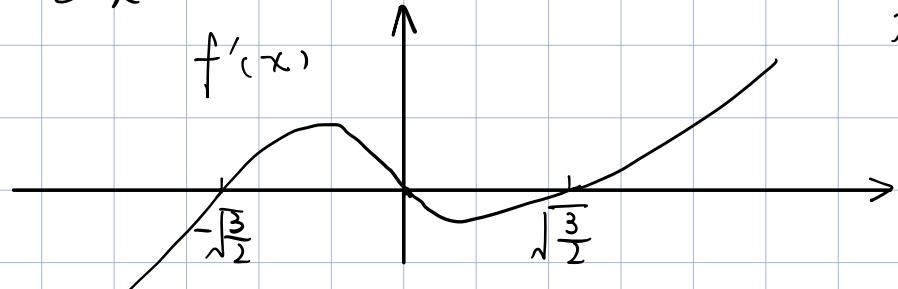
$$y^* = \begin{cases} x^2, & x \in [-1,1] \\ 1, & \text{otherwise} \end{cases}$$

When $x \in [-1,1]$, $f(x, y^*) = -x^2 \geq -1$. (equal when $(\pm 1, 1)$)

When $x < -1$ or $x > 1$, $f(x, y^*) = x^4 - 3x^2 + 1$

$$\frac{\partial f(x, y^*)}{\partial x} = 4x^3 - 6x \quad \frac{\partial^2 f(x, y^*)}{\partial x^2} = 12x^2 - 6$$

$$\frac{\partial f(x, y^*)}{\partial x} = 0, \quad x < -1 \text{ or } x > 1 \Rightarrow x_1 = \sqrt{\frac{3}{2}}, \quad x_2 = -\sqrt{\frac{3}{2}}, \quad x_3 = 0.$$



$$\frac{\partial^2 f(\sqrt{\frac{3}{2}}, y^*)}{\partial x^2} = \frac{\partial^2 f(-\sqrt{\frac{3}{2}}, y^*)}{\partial x^2}$$

$$= 12 > 0.$$

$$\frac{\partial^2 f(0, y^*)}{\partial x^2} = -6 < 0.$$

$$f\left(\sqrt{\frac{3}{2}}, 1\right) = f\left(-\sqrt{\frac{3}{2}}, 1\right) = -\frac{5}{4} < -1 = f(\pm 1, 1).$$

$\Rightarrow \left(\sqrt{\frac{3}{2}}, 1\right), \left(-\sqrt{\frac{3}{2}}, 1\right)$ are global-min.

5. (a) Since $g(\alpha) = f(x^* + \alpha d)$ is minimized at $\alpha = 0$. $\forall d \in \mathbb{R}^n$

$$g'(\alpha) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x^* + \alpha d) d_i = \nabla f(x^* + \alpha d)^T \cdot d$$

$\alpha = 0$ is global-min of $g(\alpha)$ $\forall d \in \mathbb{R}^n$

$$\Rightarrow g'(0) = \nabla f(x^*)^T \cdot d = 0 \quad \forall d \in \mathbb{R}^n$$

$$\Rightarrow \nabla f(x^*) = 0$$

(b). Consider func $f(y, z) = (z - py^2)(z - qy^2)$, $0 < p < q$.

$$\nabla f(y, z) = [4pqy^3 - 2pyz - 2qyz, 2z - qy^2 - py^2]^T$$

$$\nabla^2 f(y, z) = \begin{bmatrix} 12pqy^2 - 2pz - 2qz, & -2py - 2qy \\ -2qy - 2py, & 2 \end{bmatrix}$$

$$\nabla f(0, 0) = 0 \quad \nabla^2 f(0, 0) \succeq 0.$$

$\Rightarrow (0, 0)$ is a local-min of f .

if $p < m < q$, $f(y, my^2) = (m-p)(m-q)y^4 < 0$. if $y \neq 0$.

while $f(0, 0) = 0$. proved.

HW 1, P6 by Sagemode

```
In [1]: set_random_seed(0)
```

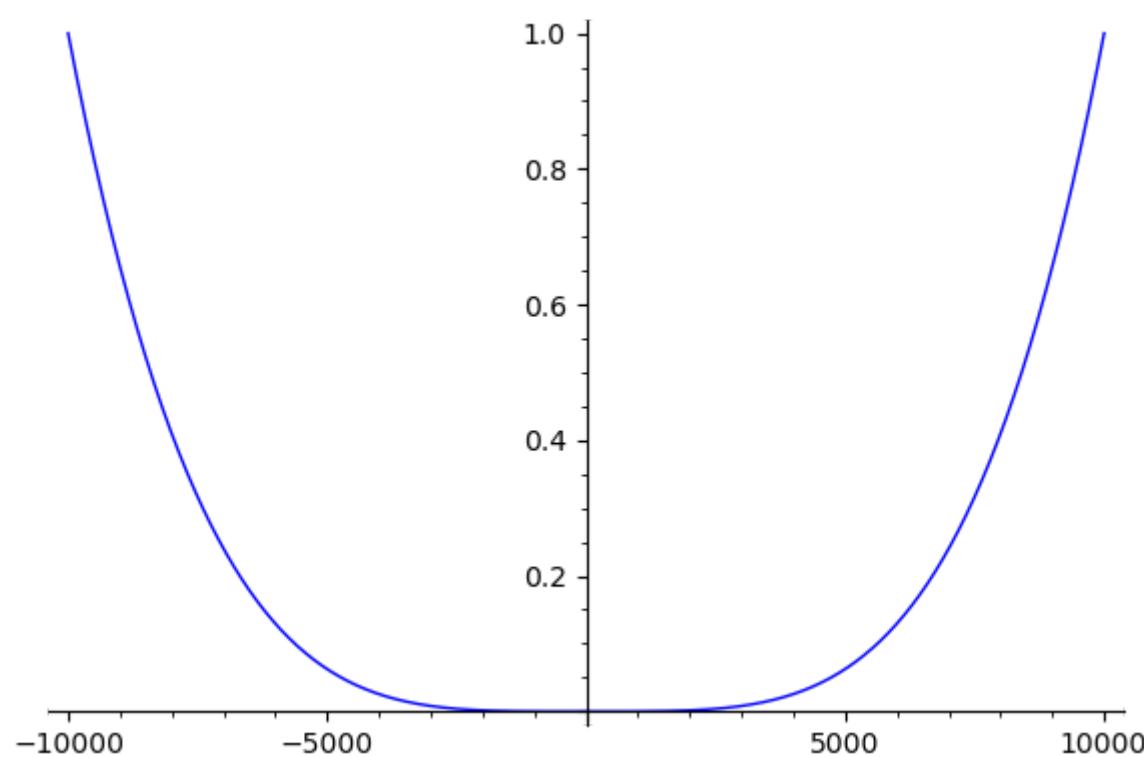
(a)

(i)

```
In [2]: x=var('x')
f=(1-x**2)**2
```

```
In [3]: plot(f, x,-10000,10000)
```

Out[3]:



(ii)

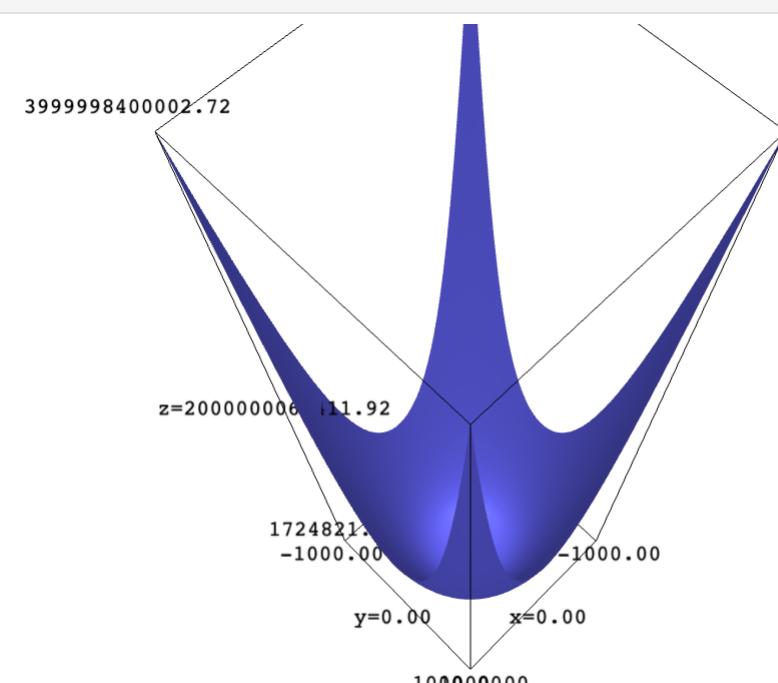
```
In [4]: A=matrix(2,2,[1,0.6,0.6,1])
var('x1','x2')
```

Out[4]: (x1, x2)

```
In [5]: f1=lambda x1,x2: ((A-vector([x1,x2]).column() * vector([x1,x2]).row())).norm('frob'))**2
```

```
In [6]: plot3d(f1, (x1,-1000,1000), (x2,-1000,1000), aspect_ratio=[1,1,1])
```

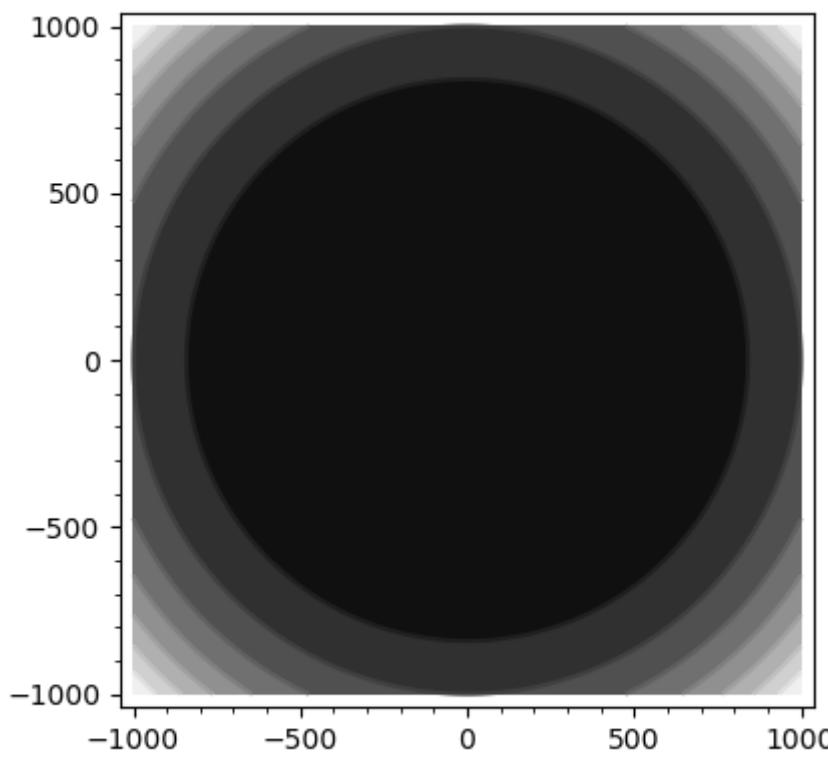
Out[6]:



(i)

```
In [7]: contour_plot(f1, (-1000,1000), (-1000,1000))
```

Out[7]:



(iii)

Pick a point $\hat{\theta} = 0_{200 \times 1} = [0, 0, \dots, 0]^T$ Pick two vectors $u = [1, 0, 0, \dots, 0]^T$ and $v = [0, 1, 0, \dots, 0]^T$

```
In [8]: u=vector([0]*200)
u[0]=1
v=vector([0]*200)
v[1]=1
```

```
In [9]: u.column()*v.row()
```

```
Out[9]: 200 x 200 dense matrix over Integer Ring (use the '.str()' method to see the entries)
```

```
In [10]: B=Matrix(RR, 200, lambda i,j: normalvariate(0, 1))
A=B*B.transpose()
```

$$(x_1 u)(x_2 v)^T = [t_{ij}]_{i,j}, t_{ij} = x_1 x_2 \text{ if } i = 1, j = 2, \text{ otherwise } t_{ij} = 0$$

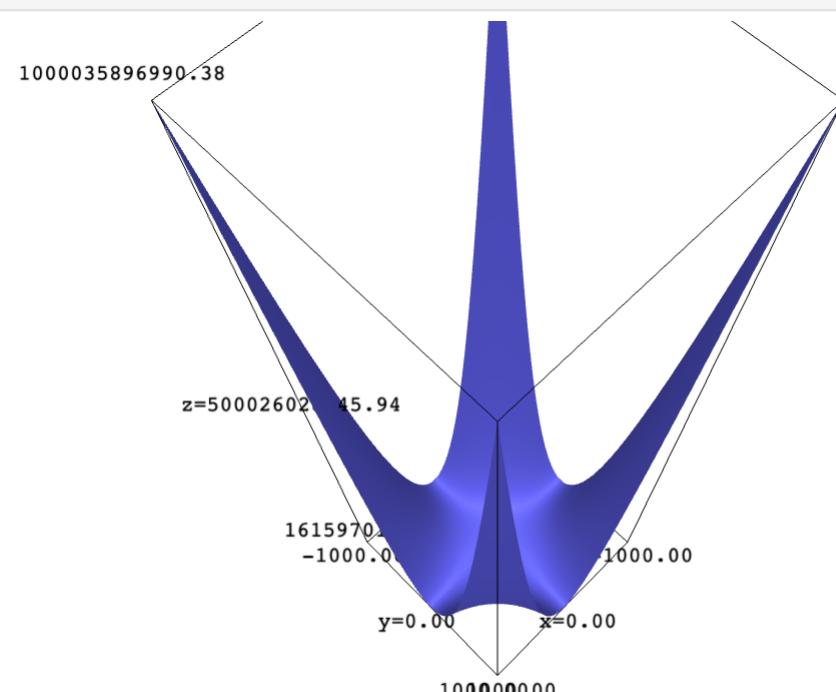
$$\|A - (x_1 u)(x_2 v)^T\|_F^2 = \|A\|_F^2 - a_{12}^2 + (a_{12} - x_1 x_2)^2 = \|A\|_F^2 + (x_1 x_2)^2 - 2a_{12}x_1 x_2$$

```
In [11]: a=A[1,2]
Af=A.norm('frob')**2
```

```
In [12]: f2=lambda x1,x2: Af+(x1*x2)**2-2*a*x1*x2
```

```
In [13]: plot3d(f2, (x1,-1000,1000), (x2,-1000,1000), aspect_ratio=[1,1,1])
```

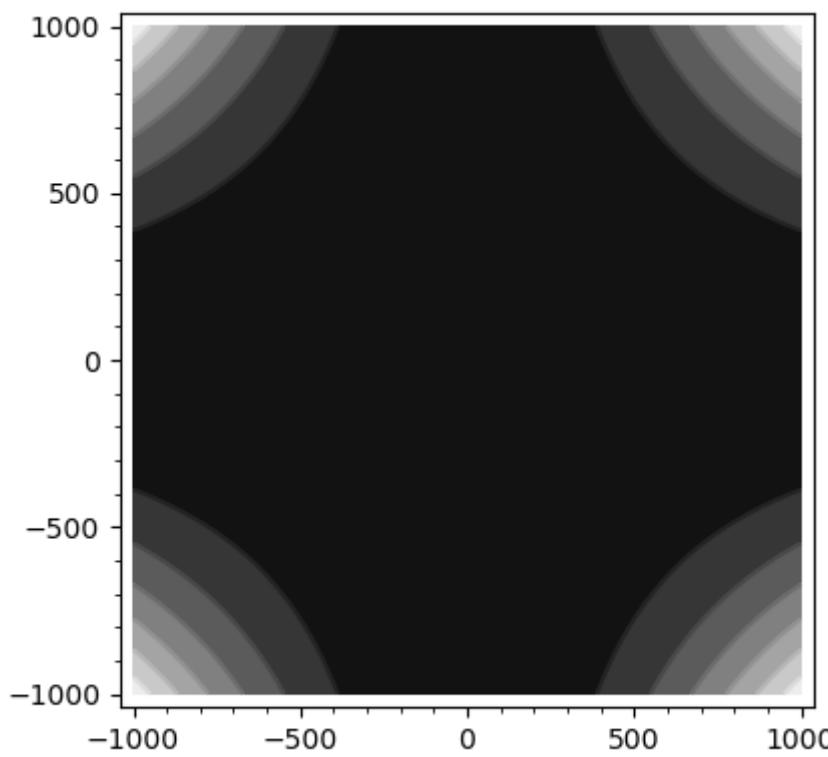
```
Out[13]:
```



(i)

```
In [14]: contour_plot(f2, (-1000,1000), (-1000,1000))
```

```
Out[14]:
```



(b)

$$(n, d, m) = (100, 30, 5)$$

```
In [15]: n=100
d=30
m=5
```

```
In [16]: def sigma(x):
    if x<0:
        return exp(x)-1
    else:
        return x
```

```
In [17]: Train=Matrix(n,d+1, lambda i,j: normalvariate(0, 1))
```

```
In [18]: x_train=Train[:, :-1]
y_train=Train[:, -1]
```

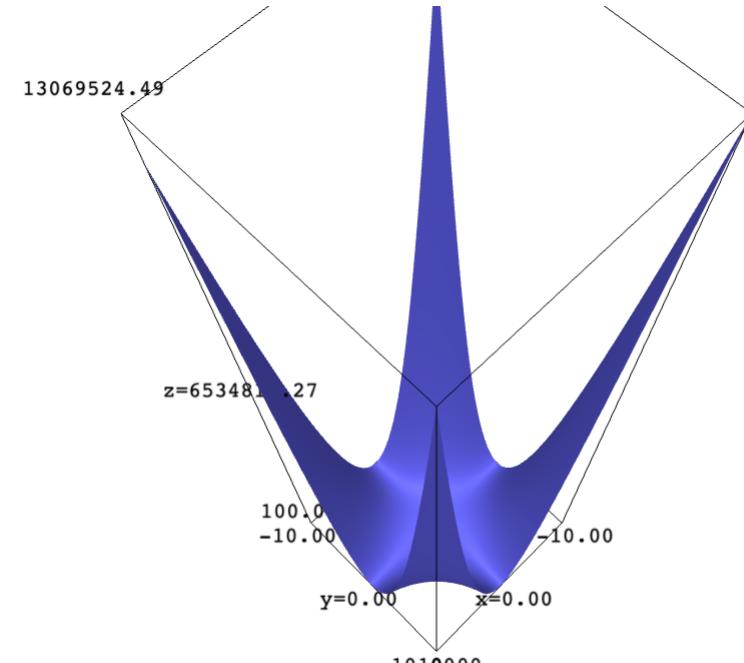
Projection onto a 2-dimensional space: $v = x_1[1, 0, 0, \dots, 0]_{m \times 1}^T$, $W = x_2[[1, 1, \dots, 1]_{1 \times 30}, 0_{1 \times 30}, 0_{1 \times 30}, 0_{1 \times 30}, 0_{1 \times 30}]_{m \times d}^T$

$$\sum_{i=1}^n (y_i - v^T \sigma(Wx_i))^2 = \sum_{i=1}^n (y_i - x_1 \sigma(x_2 \text{sum}(x_i)))^2$$

```
In [19]: def f3(x1,x2):
    fv=0
    for i in range(n):
        fv+=(y_train[i][0]-x1*sigma(x2*sum(x_train[i])))**2
    return fv
```

```
In [20]: plot3d(f3, (x1,-10,10), (x2,-10,10), aspect_ratio=[1,1,1])
```

Out[20]:

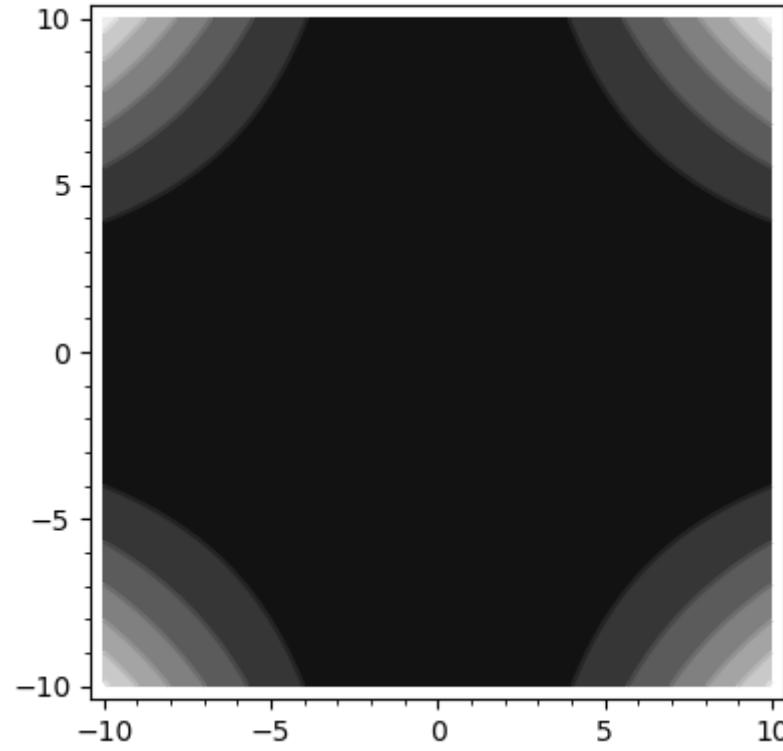


(i)

In [21]:

```
contour_plot(f3, (-10,10), (-10,10))
```

Out[21]:



$$(n, d, m) = (100, 10, 5)$$

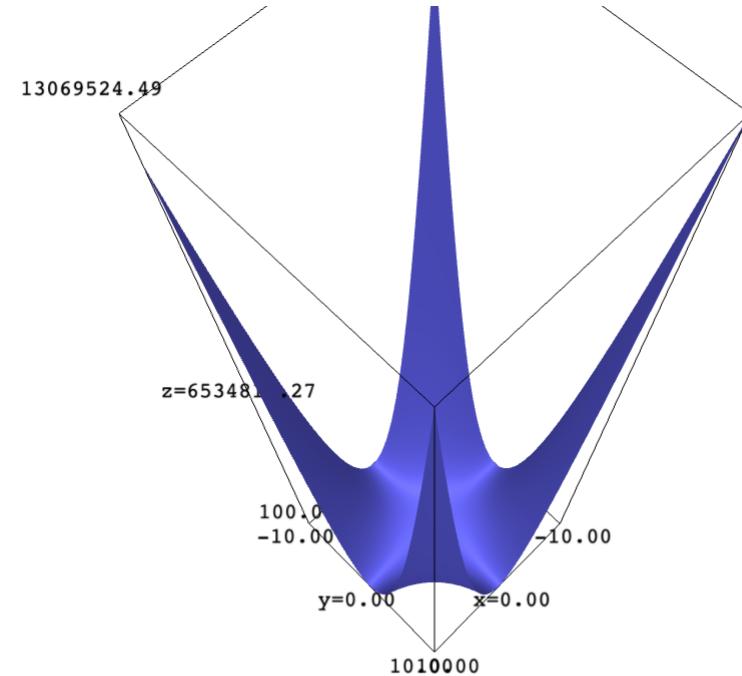
In [22]:

```
n=100
d=10
m=5
```

In [23]:

```
plot3d(f3, (x1,-10,10), (x2,-10,10), aspect_ratio=[1,1,1])
```

Out[23]:

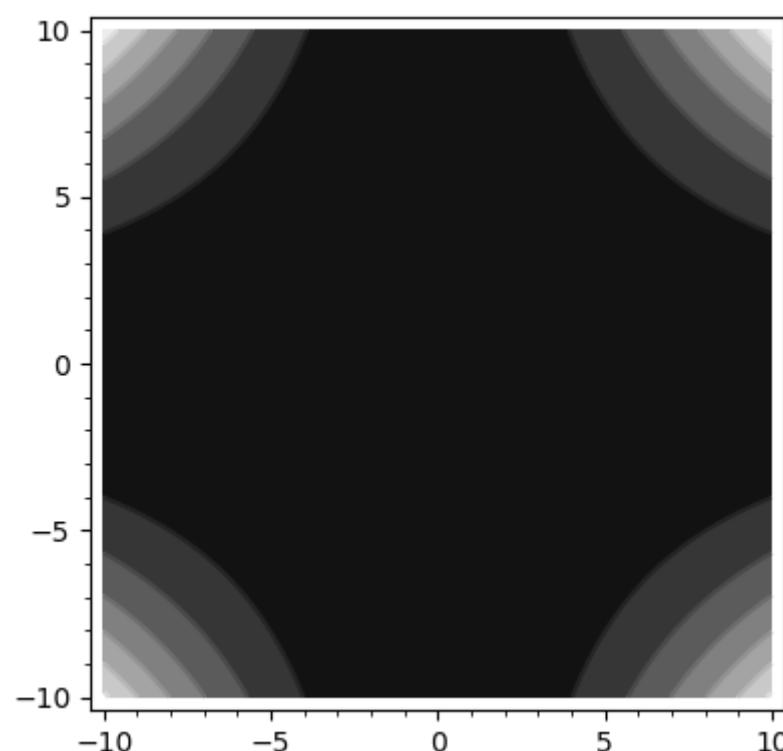


(i)

In [24]:

```
contour_plot(f3, (-10,10), (-10,10))
```

Out[24]:

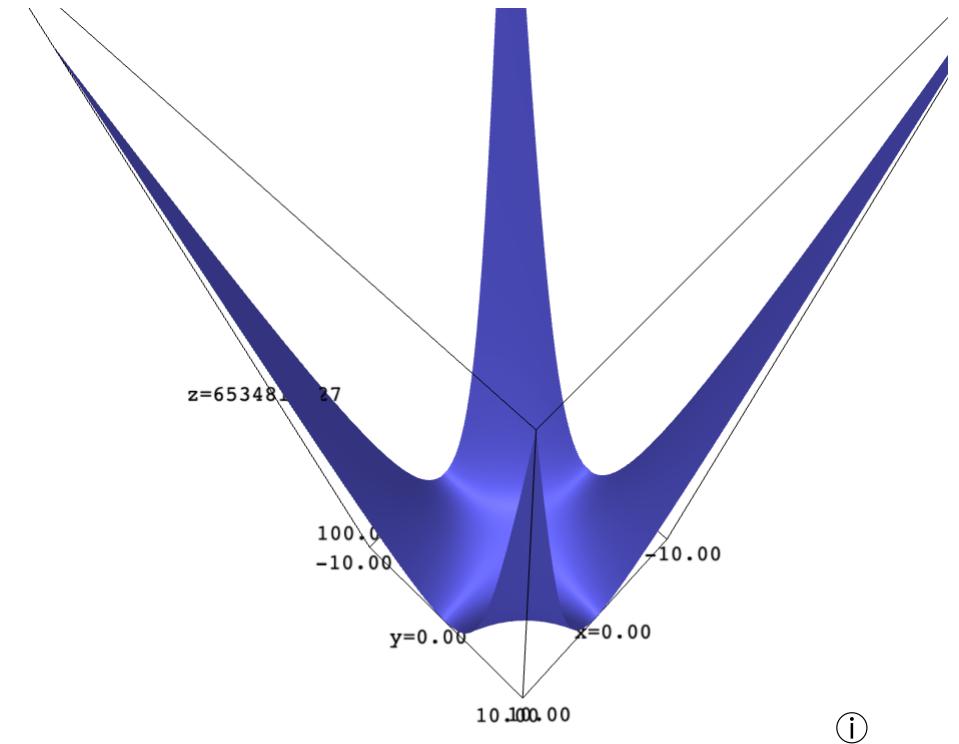


$$(n, d, m) = (100, 3, 5)$$

```
In [25]: n=100  
d=3  
m=5
```

```
In [26]: plot3d(f3, (x1,-10,10), (x2,-10,10), aspect_ratio=[1,1,1])
```

Out[26]:



```
In [27]: contour_plot(f3, (-10,10), (-10,10))
```

Out[27]:

