IE510 Applied Nonlinear Programming Homework 1

Instructor name: Ruoyu Sun Student name: Wenxiao Yang.

Warning: The work your 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

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

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, $\|x\|_B \triangleq x^T Bx$ is a norm.
Hint 2: You may use computer to help the search.

[1] Note: 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

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

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

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

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

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

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

$\Leftrightarrow \exists\, \gamma \le 1.$ s.t. $\|f(x) - f(y)\| \le \gamma \|x-y\| \quad \forall x, y$

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

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

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

(b) Sequence $\{x^k\}$ converges for any initial point x^1.

$\Leftrightarrow \{x^k\}$ is Cauchy Sequence $\forall x^1 \in \mathbb{R}^2$.

$\Leftrightarrow$ Given $\varepsilon > 0$, $\exists N_\varepsilon$ s.t. $\|x^{k+1} - x^{m+1}\| = \|A^k x^1 - A^m x^1\| < \varepsilon$, $\forall k \ge m \ge N_\varepsilon$

$\|A^k x^1 - A^m x^1\| = \|(A^k - A^m)x^1\| = \left\|\begin{bmatrix} a^k - a^m & 0 \\ 0 & b^k - b^m \end{bmatrix} x^1\right\|$

$= \sqrt{(a^k-a^m)^2(x_1^1)^2 + (b^k-b^m)^2(x_2^1)^2} < \varepsilon \quad \forall x^1 \in \mathbb{R}^2,\ k \ge m \ge N_\varepsilon.$

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

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

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

(c) $\|f(x) - f(y)\| = \|A(x-y)\|$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $z \in \mathbb{R}^2$

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, $f(x,y) = (x+1)^2 + x^2y^2 \to +\infty$ as $x \to +\infty$

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

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

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

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

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

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

if $\beta \in (-2,2)$ $\nabla^2 f(x,y)$ is PD

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

Otherwise there is no global-min.

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

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

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

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

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

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

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

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

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

$\det(1) = 1 > 0$

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

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

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

Hence the set of all local-min is

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

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

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

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

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

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

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

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

$\nabla^2 f(\frac{5}{3}\pi, \frac{5}{3}\pi) = \begin{bmatrix} \sqrt{3}, & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2}, & \sqrt{3} \end{bmatrix} \succ 0 \Rightarrow (\frac{5}{3}\pi,\frac{5}{3}\pi)$ 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

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

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

(e)

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

Since $y \in (-1,1)$, $f(x,y) \to +\infty$ as $x \to +\infty$,

the S is compact.

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

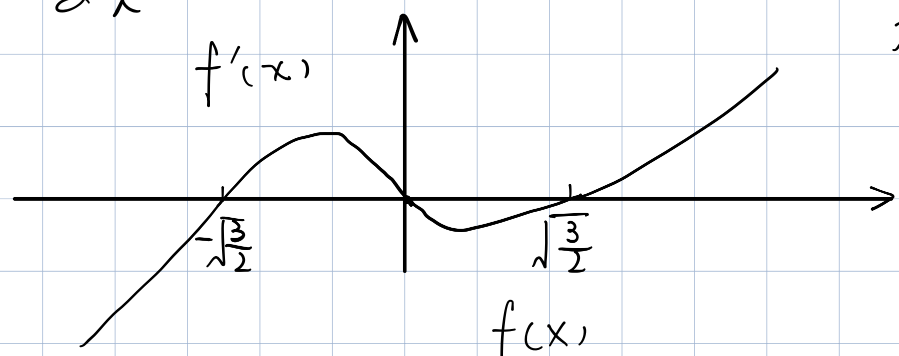
$$y^* = \begin{cases} x^2, & x \in [-1,1] \\ 1, & \text{otherwise} \end{cases}$$

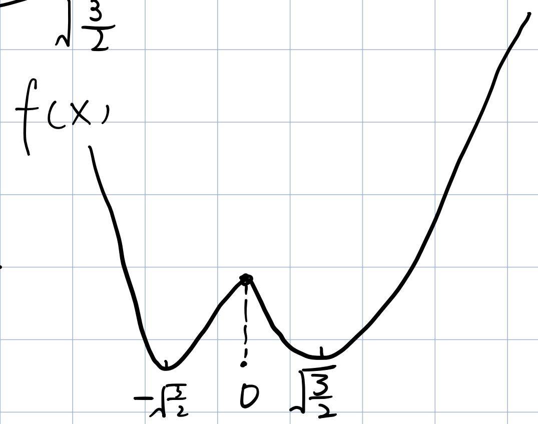
When $x \in [-1,1]$, $f(x,y^*) = -x^2 \ge -1$. (equal when $(\pm 1, 1)$)

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

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

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





$\frac{3}{\cancel{2}} \times \cancel{12}$ 6

18 - 6

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

$= 12 > 0.$

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

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

$\Rightarrow (\sqrt{\frac{3}{2}}, 1)\ (-\sqrt{\frac{3}{2}}, 1)$ 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.