THEORETICAL PART

GD HW1

[1] function f, its graph Γ.

f is convex $\Longleftrightarrow$ for each $x_1, x_2, \dots, x_k \in D$ and $\alpha_1, \dots, \alpha_k \in [0,1]$,
$\sum_{i=1}^{k} \alpha_i = 1$: $f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i)$ (1)

Proof (induction):

$\Rightarrow$: for $k=1$: $\alpha_1 = 1$ $\quad f(\alpha_1 x_1) = f(x_1) = \alpha_1 \cdot f(x_1) = f(x_1)$.

$k \to k+1$: We assume (1) holds

$$f\left(\sum_{i=1}^{k+1} \alpha_i x_i\right) = f\left(\sum_{i=1}^{k} \alpha_i x_i + \alpha_{k+1} x_{k+1}\right) = f\left(\beta\left(\sum_{i=1}^{k} \frac{\alpha_i}{\beta} x_i\right) + \alpha_{k+1} x_{k+1}\right) =$$

$\beta = \sum_{i=1}^{k} \alpha_i \Longrightarrow \boxed{\sum_{i=1}^{k} \alpha_i + \alpha_{k+1} = \beta + \alpha_{k+1} = 1}$

$$\Rightarrow f\left(\beta\left(\sum_{i=1}^{k} \frac{\alpha_i}{\beta} x_i\right) + \alpha_{k+1} x_{k+1}\right) \overset{\text{convexity def}}{\leq} \beta \cdot f\left(\sum_{i=1}^{k} \frac{\alpha_i}{\beta} x_i\right) + \alpha_{k+1} \cdot f(x_{k+1}) \leq$$

$$\overset{(1)}{\leq} \beta \cdot \sum_{i=1}^{k} \frac{\alpha_i}{\beta} f(x_i) + \alpha_{k+1} f(x_{k+1}) =$$

induction assumption

$$= \sum_{i=1}^{k+1} \alpha_i f(x_i) \quad \blacksquare$$

$\Leftarrow$: Assume f isn't convex.
Then there exist points x, y for which
$f(\alpha_1 x + \underbrace{(1-\alpha_1)}_{\alpha_2} y) > \alpha_1 f(x) + (1-\alpha_1) \cdot f(y)$ $\quad \rightarrow\leftarrow$

def of (non) convexity

$\blacksquare$

$\boxed{2}$ $f(x,y) = x^2 + e^x + y^2 - xy$, f restricted to $K = [-2,2] \times [-2,2]$

β-smooth, L-Lipschitz, strongly convex. Find L, α, β on K

~~Finding L = [illegible] gradient [illegible] f on K~~

$$\nabla f = [2x + e^x - y,\ 2y - x]$$

~~$\|\nabla f\|^2 = (2x+e^x-y)^2 + (2y-x)^2 = 4x^2 + e^{2x} + y^2 + 4xe^x - 4xy - 2e^x y + 4y^2 - 4xy + x^2 = \boxed{5x^2 + e^{2x} + 4xe^x - 8xy - 2e^x y + 5y^2}$~~

$$H_f = \begin{bmatrix} 2+e^x & -1 \\ -1 & 2 \end{bmatrix}$$

1) $2 + e^x > 0$ 2) $\det(H_f) = 4 + 2e^x - 1 = 3 + 2e^x > 0$

$\Rightarrow H_f$ is PD $\Rightarrow f$ is convex

Finding β:

$\|H_f\| = 2 + e^x$, maximal at $x = 2 \Rightarrow \|H_f\|_{max} = \boxed{\frac{4 + e^2 + \sqrt{e^4 + 4}}{2} = \beta}$

eigenvalues: $\begin{vmatrix} 2+e^x-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = (2+e^x-\lambda)(2-\lambda) - 1 = 4 + 2e^x - 2\lambda - 2\lambda - \lambda e^x + \lambda^2 - 1 =$

$= +\lambda^2 + \lambda(-4 - e^x) + 3 + 2e^x \Rightarrow \lambda_{1,2} = \frac{4 + e^x \pm \sqrt{16 + 8e^x + e^{2x} - 12 - 8e^x}}{2} =$

$= \frac{4 + e^x \pm \sqrt{e^{2x} + 4}}{2} \Rightarrow \lambda_{max}(x) = \frac{4 + e^x + \sqrt{e^{2x} + 4}}{2}$

$\lambda_{min}(x) = \frac{4 + e^x - \sqrt{e^{2x} + 4}}{2}$

$\lambda_{max} \Rightarrow$ since e^x, e^{2x} are increasing functions, λ_{max} is also increasing $\Rightarrow \lambda_{max,max} = \lambda_{max}(2) = \frac{4+e^2+\sqrt{e^4+4}}{2}$

$\frac{d\lambda_{min}}{dx} = \frac{1}{2}e^x - \frac{1}{\cancel{2}\sqrt{e^{2x}+4}} e^{2x} \cdot \cancel{2} = \frac{1}{2}e^x\left(1 - \underbrace{\frac{e^x}{\sqrt{e^{2x}+4}}}_{<1}\right) \geq 0 \Rightarrow \lambda_{min}(x)$ is always increasing

$\Rightarrow \lambda_{min,min} = \lambda_{min}(-2) = \frac{4+e^{-2}-\sqrt{e^{-4}+4}}{2}$

Finding α: minimal eigenvalue of H_f is $\frac{4+e^{-2}-\sqrt{e^{-4}+4}}{2}$

$\Rightarrow \boxed{\alpha = \frac{4+e^{-2}-\sqrt{e^{-4}+4}}{2}}$

Finding L:

$$\|\nabla f\|^2 = \underbrace{\boxed{5x^2 + e^{2x} + 4xe^x}}_{1} \underbrace{\boxed{-8xy - 2e^x y + 5y^2}}_{2}$$

Since it may be hard to find the maximum of this function, some common sense will be applied.
In 1 we can see that for $x=2$ we will get the largest value. If we pick $x=2$ then for 2 we can still get the largest possible value with $y=-2$ since the first 2 terms will become positive, and the last one is squared so it will remain positive.

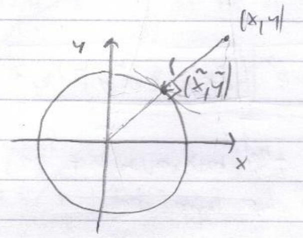
$\Rightarrow \|\nabla f\|^2_{max} = \|\nabla f\|^2(2,-2) = 20 + e^4 + 8e^2 + 32 + 4e^2 + 20 =$

$= \boxed{72 + 12e^2 + e^4 = L^2}$

$\boxed{L = \sqrt{72 + 12e^2 + e^4}}$

③

a) $D: x^2+y^2 \le 1.5$; $\quad \pi_D : \mathbb{R}^2 \to D$

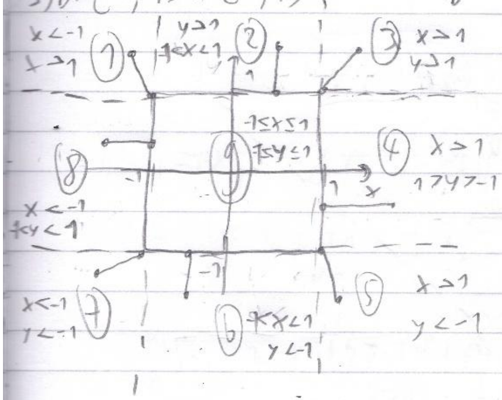


$$\pi_D(x,y) = \begin{cases} (x,y) & \text{if } x^2+y^2 \le 1.5 \\ (\tilde{x},\tilde{y}) & \text{if } x^2+y^2 > 1.5 \end{cases}$$

$(\tilde{x},\tilde{y}) \equiv$ projection of (x,y) onto the circle

$$\underbrace{\tilde{x} = x}_{\text{direction}} \cdot \underbrace{\frac{\sqrt{1.5}}{\sqrt{x^2+y^2}}}_{\text{scaling}} \qquad \tilde{y} = y \frac{\sqrt{1.5}}{\sqrt{x^2+y^2}}$$

b) $D: [-1,1] \times [-1,1]$, $\quad \pi_D : \mathbb{R}^2 \to D$



For each area:

① : $\pi_D(x,y) = (-1,1)$
② : $\pi_D(x,y) = (x,1)$
③ : $\pi_D(x,y) = (1,1)$
④ : $\pi_D(x,y) = (1,y)$
⑤ : $\pi_D(x,y) = (1,-1)$
⑥ : $\pi_D(x,y) = (x,-1)$
⑦ : $\pi_D(x,y) = (-1,-1)$
⑧ : $\pi_D(x,y) = (-1,y)$
⑨ : $\pi_D(x,y) = (x,y)$

Written more compactly:

①, ③, ⑤, ⑦ : $|x| > 1$, $|y| > 1$

$\pi_D(x,y) = \left(\frac{x}{|x|}, \frac{y}{|y|}\right)$

②, ⑥ : $|y| > 1$, $|x| < 1$

$\pi_D(x,y) = \left(x, \frac{y}{|y|}\right)$

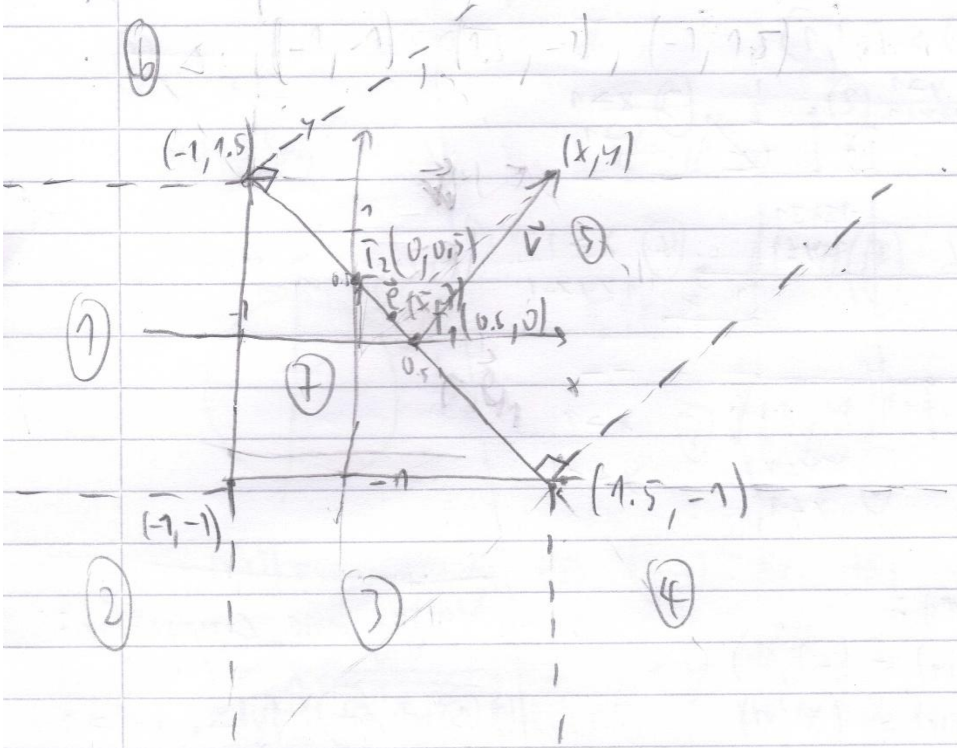
⑧, ④ : $|x| > 1$, $|y| < 1$

$\pi_D(x,y) = \left(\frac{x}{|x|}, y\right)$

⑨ : $|x| < 1$, $|y| < 1$

$\pi_D(x,y) = (x,y)$

d.



1 : $\pi_D(x,y) = (-1, y)$; $x < -1$, $-1 < y < 1.5$

2 : $\pi_D(x,y) = (-1,-1)$; $x < -1$, $y < -1$

3 : $\pi_D(x,y) = (x, -1)$; $-1 < x < 1.5$, $y \leq -1$

4 : $\pi_D(x,y) = (1.5, -1)$; $x > 1.5$, $y \leq x - 2.5$

5 : $\pi_D(x,y) = (\tilde{x}, \tilde{y})$; $y > -x + 0.5$, $y > x - 2.5$, $y < x + 2.5$

$\vec{v} = (x - 0.5, y)$, $\vec{e} = (-1, 1)$ $\quad t = \frac{\vec{v} \cdot \vec{e}}{|\vec{v}||\vec{e}|} = \cos\varphi$

$$\boxed{t = \frac{0.5 - x + y}{\sqrt{2} \cdot \sqrt{(x-0.5)^2 + y^2}} \quad (\tilde{x}, \tilde{y}) = (0.5, 0) + t \cdot \vec{e}}$$

6 $\pi_D(x,y) = (-1, 1.5)$; $y > 1.5$, $y > x + 2.5$

7 $\pi_D(x,y) = (x,y)$; $x \geq -1$, $y \geq -1$, $y \leq -x + 0.5$.

$F(x,y) = x^2 + 2y^2 \qquad x_1 = (1,1)$

$\nabla f = [2x, 4y] \quad \nabla f(x_1) = [2,4] \quad x_2 = x_1 - \gamma \cdot \nabla f(x_1)$

$f(x_2) = f(x_1 - \gamma \nabla f(x_1)) = f([1-2\gamma, 1-4\gamma]) =$

$= 1 - 4\gamma + 4\gamma^2 + 2 - 16\gamma + 32\gamma^2 =$

$= 36\gamma^2 - 20\gamma + 3 = 0$

$\frac{df(x_2)}{d\gamma} = 72\gamma - 20 = 0 \Rightarrow \boxed{\gamma = \frac{20}{72}}$

$\min(f(x_2)) = 36\left(\frac{20}{72}\right)^2 - 20\frac{20}{72} + 3 = \boxed{\frac{2}{9}}$

)

$x_2 = [1-2\gamma, 1-4\gamma] \qquad x^* = [0,0]$

$d^2 = (1-2\gamma)^2 + (1-4\gamma)^2 = 1 - 4\gamma + 4\gamma^2 + 1 - 8\gamma + 16\gamma^2 =$

$= 20\gamma^2 - 12\gamma + 2$

$\frac{\partial d^2}{\partial \gamma} = 40\gamma - 12 = 0 \Rightarrow \boxed{\gamma = \frac{12}{40} = \frac{3}{10}}$

$d^2_{min} = 20 \cdot \left(\frac{3}{10}\right)^2 - 12 \cdot \frac{3}{10} + 2 = \boxed{\frac{1}{5} = d^2_{min}}$

$\boxed{d_{min} = \sqrt{\frac{1}{5}}}$