HW 1

1. Prove: f is convex $\Leftrightarrow$ $\forall x_1, \dots, x_k \in D$, $\alpha_1, \dots, \alpha_k \in [0,1]$, $\sum_{i=1}^{k} \alpha_i = 1$:

if D is convex and $\forall x, y \in D$, $\forall t \in [0,1]$: $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i)$$

($\Leftarrow$) If right side holds for all $x_i \in D$, we can choose α_i such that two of them are not 0: let's take α_1 and $\alpha_2 = 1 - \alpha_1$.
Then:

$$f(\alpha_1 x_1 + (1-\alpha_1) x_2) \leq \alpha_1 f(x_1) + (1-\alpha_1) f(x_2)$$

holds for every x_1, x_2, α_1, which is exactly the definition for convex functions. $\alpha_1 + 1 - \alpha_1 = 1$.

($\Rightarrow$) We will prove this by induction

$k=2$: holds by definition: $f(\underset{\alpha_1}{t} \underset{x_1}{x} + \underset{\alpha_2}{(1-t)} \underset{x_2}{y}) \leq tf(x) + (1-t)f(y)$ ✓

$k \to k+1$ $\quad f\left(\sum_{i=1}^{k+1} \alpha_i x_i\right) \leq \sum_{i=1}^{k+1} \alpha_i f(x_i)$

We know: $f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i) \quad \forall \alpha_i \in [0,1], \sum_{i=1}^{k} \alpha_i = 1, \quad \forall x_i \in D$.

$$f\left(t \cdot \sum_{i=1}^{k} \alpha_i x_i + (1-t) x_{k+1}\right) \underset{f \text{ convex}}{\leq} t \cdot f\left(\sum_{i=1}^{k} \alpha_i x_i\right) + (1-t) f(x_{k+1}) \underset{k\text{th step of induction}}{\leq}$$

$$\leq t\left(\sum_{i=1}^{k} \alpha_i f(x_i)\right) + (1-t) f(x_{k+1}) = \sum_{i=1}^{k} t\alpha_i f(x_i) + (1-t) f(x_{k+1})$$

$$\sum_{i=1}^{k} t\alpha_i + 1 - t = t \cdot \underbrace{\sum \alpha_i}_{=1} + 1 - t = 1$$

$$= \sum_{i=1}^{k+1} \beta_i f(x_i) \quad \text{with } \beta_i = t \cdot \alpha_i \text{ for } i = 1, \dots, k \text{ and } 1 - t = \beta_{k+1}$$

Since this holds for every t and every α_i (where $\sum \alpha_i = 1$), it holds for every β_i; $i = 1, \dots, k+1$, if $\sum_{i=1}^{k+1} \beta_i = 1$.
We can also write: $f\left(\sum_{i=1}^{k+1} \beta_i x_i\right) \leq \sum_{i=1}^{k+1} \beta_i f(x_i)$ ∎

2. f is L-Lip. if $\forall x, y \in D: \quad |f(x) - f(y)| \leq L \|x - y\|$

Prop. 26: f is L-Lip. iff. $\|\nabla f\| \leq L$.

$\Rightarrow$ The smallest possible L is $L = \max\{\|\nabla f\|\}$

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

$\|\nabla f\| = \sqrt{(2x + e^x - y)^2 + (2y - x)^2}$

$\operatorname{argmax} \|\nabla f\| = \operatorname{argmax} \|\nabla f\|^2$:

We are searching for max of fn $g(x,y) = (2x + e^x - y)^2 + (2y - x)^2$

$g_x(x,y) = \cancel{2}(2x + e^x - y) \cdot (2 + e^x) + \cancel{2}(2y - x) \cdot (-1) = 0$

$g_y(x,y) = \cancel{2}(2x + e^x - y) \cdot (-1) + \cancel{2}(2y - x) \cdot 2 = 0$

$-4x - 2e^x + 2y + 8y - 4x = 0$

$y = \frac{1}{5}(4x + e^x)$

$\cancel{8}x + \cancel{4}e^x - \cancel{2}y + \cancel{4}xe^x + \cancel{2}e^{2x} - \cancel{2}ye^x + \cancel{2}y + \cancel{2}x = 0$

$5x - 4/5(4x + e^x) + 2e^x + e^{2x} + 2xe^x - \frac{1}{5}(4x + e^x)e^x = 0$

$9/5x + 6/5 e^x + 4/5 e^{2x} + \frac{6}{5}xe^x = 0$

$9x + 6e^x + 4e^{2x} + 6xe^x = 0$

$3x(3 + 2e^x) + 2e^x(3 + 2e^x) = 0$

$(3 + 2e^x)(3x + 2e^x) = 0$

1. $3 + 2e^x = 0$
   $e^x = -\frac{3}{2}$ $\times$

2. $3x + 2e^x = 0$

$\Downarrow$ minimum ~~not~~ interested

$\|\nabla f\| = \sqrt{(2x + e^x - \frac{4}{5}x - \frac{1}{5}e^x)^2 + (\frac{8}{5}x + \frac{2}{5}e^x - x)^2} =$

$= \sqrt{(\frac{6}{5}x + \frac{4}{5}e^x)^2 + (\frac{3}{5}x + \frac{2}{5}e^x)^2} = \sqrt{(\frac{2}{5} \cdot 0)^2 + (\frac{1}{5} \cdot 0)^2} = 0$

We need to look at border values:

$x = -2: \|\nabla f\| = \sqrt{(\underbrace{-4 + e^{-2}}_{<0} - y)^2 + (\underbrace{2y + 2}_{\geq 0})^2}$ $\rightarrow$ is maximal when y is the biggest: $y = +2$

But it is smaller than (because we add e^{+2} to the abs. value)

$x = 2: \|\nabla f\| = \sqrt{(4 + e^{+2} - y)^2 + (2y - 2)^2} \rightarrow$ when $y = -2$

$= \sqrt{(6 + e^{+2})^2 + (-6)^2} = \sqrt{36 + 12e^{+2} + e^{+4} + 36} = \boxed{\sqrt{72 + 12e^{+2} + e^{+4}} = L}$

$\approx$ 14,672

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

Def: f is β-smooth if $\forall x,y \in D: \|\nabla f(x) - \nabla f(y)\| \leq \beta \|x-y\|$

Prop. 2.8.: f is β-smooth $\Leftrightarrow \|\nabla^2 f\| \leq \beta$. $\Leftrightarrow$ all eigenvalues are on $[0, \beta]$

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

eigenvalues: $(2+e^x - \lambda)(2-\lambda) - 1 = 0$

$$\lambda^2 + 4 + 2e^x - \lambda(2+e^x+2) - 1 = 0$$

$$\lambda^2 - (4 + e^x)\lambda + 2e^x + 3 = 0$$

$$\lambda_{1,2} = \frac{4+e^x \pm \sqrt{16 + 8e^x + e^{2x} - 4 \cdot 2e^x - 12}}{2} = \frac{4+e^x \pm \sqrt{4+e^{2x}}}{2}$$

$\Rightarrow$ max eigenvalue is at $x = 2$: $\boxed{\lambda_1 = \frac{4 + e^2 + \sqrt{4+e^4}}{2} = \beta \approx 9{,}522}$

f is α-strongly convex if $f(x) - \frac{\alpha}{2}\|x\|^2$ is convex

Prop. 2.11.: $\hookrightarrow$ iff each eigenvalue of $\nabla^2 f$ is $\geq \alpha$.

The smallest eigenvalue of $\nabla^2 f$ is at $x = -2$:

$$\boxed{\lambda_2 = \frac{4 + e^{-2} - \sqrt{4+e^{-4}}}{2} \approx 1{,}07 = \alpha}$$

f is convex if $\nabla^2 f$ is PSD. (for all $x, y \in K$)

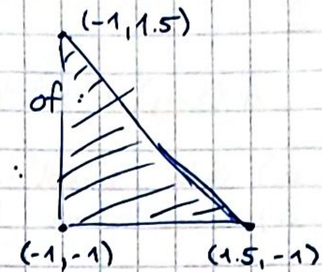
$\det(\nabla^2 f) = 4 + 2e^x - 1 = 3 + 2e^x > 0 \quad \forall x\ (,y)$

$2 + e^x > 0 \quad \forall x\ (,y)$

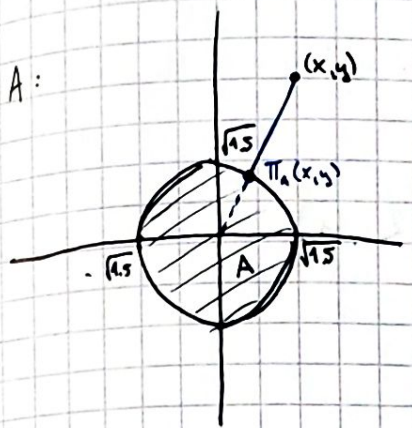
$\rightarrow \boxed{\begin{smallmatrix} + & \\ & + \end{smallmatrix}} \rightarrow$ & $H_f = \nabla^2 f$ PD $\Rightarrow$ f is convex everywhere.

3. Projections $\mathbb{R}^2 \to K$ to the closest points of

A: $x^2 + y^2 \le 1.5$ B: $[-1,1] \times [-1,1]$ C:



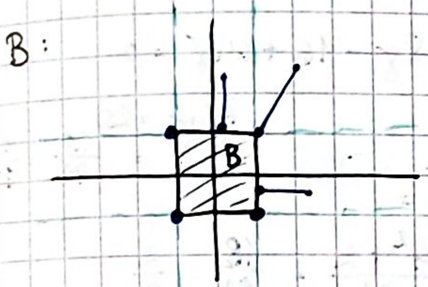
A:



$$\Pi_A(x,y) = \begin{cases} (x,y) & \text{if } (x,y) \in A \\ \frac{\sqrt{1.5}}{\sqrt{x^2+y^2}} \cdot (x,y) & \text{otherwise} \end{cases}$$

Π_A shortens each vector to lenght of $\sqrt{1.5}$, keeps its direction

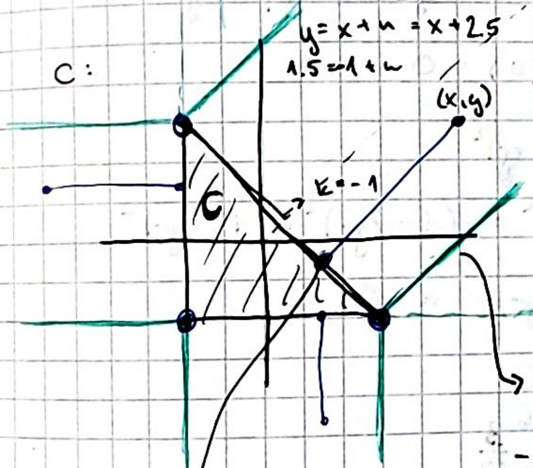
B:



$$\Pi_B(x,y) = (\Pi_{B1}(x), \Pi_{B2}(y))$$

$$\Pi_{B1}(x) = \begin{cases} x & ; -1 \le x \le 1 \\ -1 & ; x < -1 \\ 1 & ; x > 1 \end{cases}$$

C:



$$\Pi_C(x,y) = \begin{cases} (x,y) & ; (x,y) \in C \\ (-1,y) & ; x < -1,\ y \in (-1, 1.5) \\ (-1, 1.5) & ; y \ge \max(1.5, 2.5 + x) \\ (-1,-1) & ; x < -1,\ y < -1 \\ (x,-1) & ; x \in (-1, 1.5),\ y < -1 \\ (1.5, -1) & ; x \ge \max(1.5, y + 2.5) \\ \frac{1}{2}(0.5 - y + x,\ 0.5 + y - x) & ; \text{otherwise} \end{cases}$$

$k = 1$
$y = x + n$
$-1 = 1.5 + n$
$n = -2.5$

$(x,y) = \Pi(x,y) + ($
$\hookrightarrow y = x + n \to n_1 = y_1 - x_1$

$: y = -x + 0.5$

$$\left.\begin{aligned} -x + 0.5 &= x + (y_1 - x_1) \\ \tfrac{1}{2}(0.5 - y_1 + x_1) &= x \end{aligned}\right\}$$

$$y = \frac{1}{2}(0.5 - y_1 + x_1) + y_1 - x_1 = \frac{1}{2}(0.5 + y_1 - x_1)$$

4. $f(x,y) = x^2 + 2y^2$. $x_1 = (1,1)$

a) GD: find minimum of $f(x_2)$

b) How close to actual min of f: x^* can we get? $\min \|x^* - x_2\|_2$

GD: $X_{k+1} = x_k - \gamma \nabla f(x_k)$

$x_2 = x_1 - \gamma \nabla f(x_1)$

a) $\nabla f = [2x, 4y] \longrightarrow \nabla f(x_1) = [2, 4]$

$x_2 = [1 - 2\gamma, 1 - 4\gamma]$

$f(x_2) = (1-2\gamma)^2 + 2(1-4\gamma)^2$

$\frac{\partial f(x_2)}{\partial \gamma} = 2(1-2\gamma)\cdot(-2) + 4(1-4\gamma)\cdot(-4) = -4 + 8\gamma - 16 + 64\gamma = 0$

$72\gamma = 20$

$\gamma = \frac{20}{72} = \frac{5}{18}$

$\min f(x_2) = f(x_2)\Big|_{\gamma = \frac{20}{72}} = \left(\frac{32}{72}\right)^2 + 2\cdot\left(-\frac{8}{72}\right)^2 =$

$= \frac{32^2 + 2\cdot 8^2}{72^2} = \frac{4^2+2}{9^2} = \frac{18}{81} = \frac{2}{9}$

b) Actual minimum of $f(x,y)$ is $f(0,0) = 0$

$\|x^* - x_2\| = \|x_2\| = \sqrt{(1-2\gamma)^2 + (1-4\gamma)^2}$

γ at min is equal to γ at min of $\|x_2\|^2 = (1-2\gamma)^2 + (1-4\gamma)^2$

$\frac{\partial \|x_2\|^2}{\partial \gamma} = -2\cdot 2\cdot(1-2\gamma) - 4\cdot 2\cdot(1-4\gamma) = 0$

$-2 + 4\gamma - 4 + 16\gamma = 0$

$20\gamma = 6$

$\gamma = \frac{3}{10}$

$\|x_2\|\Big|_{\gamma=\frac{3}{10}} = \sqrt{\left(\frac{4}{10}\right)^2 + \left(-\frac{2}{10}\right)^2} = \sqrt{\frac{4+1}{25}} = \frac{\sqrt{5}}{5}$