Problem 1

a) $g_1 : \mathbb{R}^{d_1} \to \mathbb{R}$ - convex

$g_2 : \mathbb{R} \to \mathbb{R}$ - convex

$h(x) = g_2(g_1(x))$ - convex?

Product rule

$$\boxed{(fg)' = f'g + fg'}$$

$$h'(x) = g_2'(g_1(x)) \cdot g_1'(x)$$

$$h''(x) = \underbrace{g_2''(g_1(x))}_{\geq 0} \cdot \underbrace{g_1'(x)}_{\geq 0}{}^2 + \underbrace{g_2'(g_1(x))}_{?} \cdot \underbrace{g_1''(x)}_{\geq 0}$$

From the expression above, we notice that if g_2 is decreasing and has a sufficiently high negative slope, then it could be the case that $h''(x) < 0$ and, hence, non-convex.

b) From the same expression, we notice that if g_2 is non-decreasing, then $h''(x) \geq 0$ and hence convex.

c) One observation is that this expression makes sense iff $d_1 = d_2 = \dots = d_n$.

We know that max is a convexity-preserving operation. So, $\max(g_1(x), g_2(x))$ is convex.

$$\text{Then } \overbrace{\max(\underbrace{\max(g_1(x), g_2(x))}_{\text{convex}}, \underbrace{g_3(x)}_{\text{convex}})}^{\text{convex}} =$$

$$= \max(g_1(x), g_2(x), g_3(x)) \text{ is also convex.}$$

By mathematical induction, we can then prove that $\max(g_1, \dots, g_n)$ is also convex.

Problem 2

a) $f(x_1, x_2) = 0.5x_1^2 + x_2^2 + 2x_1 + x_2 + \cos(\sin(\sqrt{\pi}))$

$$\frac{\partial f}{\partial x_1} = x_1 + 2 = 0 \Rightarrow x_1 = -2$$

$$\frac{\partial f}{\partial x_2} = 2x_2 + 1 = 0 \Rightarrow x_2 = -\frac{1}{2}$$

$$\Rightarrow x^* = \left[-2, -\frac{1}{2}\right]$$ – the minimizer

b) $\sigma = 1$, $x^{(0)} = (0, 0)$

First iteration:

• Compute the gradient in $(0,0)$:

$$\nabla f = \begin{bmatrix} x_1 + 2 & 2x_2 + 1 \end{bmatrix}$$

$$\nabla f(0,0) = \begin{bmatrix} 2 & 1 \end{bmatrix}$$

$$x^{(1)} = x^{(0)} - \sigma \nabla f(0,0) = (0,0) - (2,1)$$
$$= (-2, -1)$$

• Update x

Second iteration:

$$\nabla f(-2,-1) = [0,-1]$$

$$x^{(2)} = x^{(1)} - \sigma \nabla f(-2,-1) = (-2,-1) - (0,-1)$$
$$= (-2, 0)$$

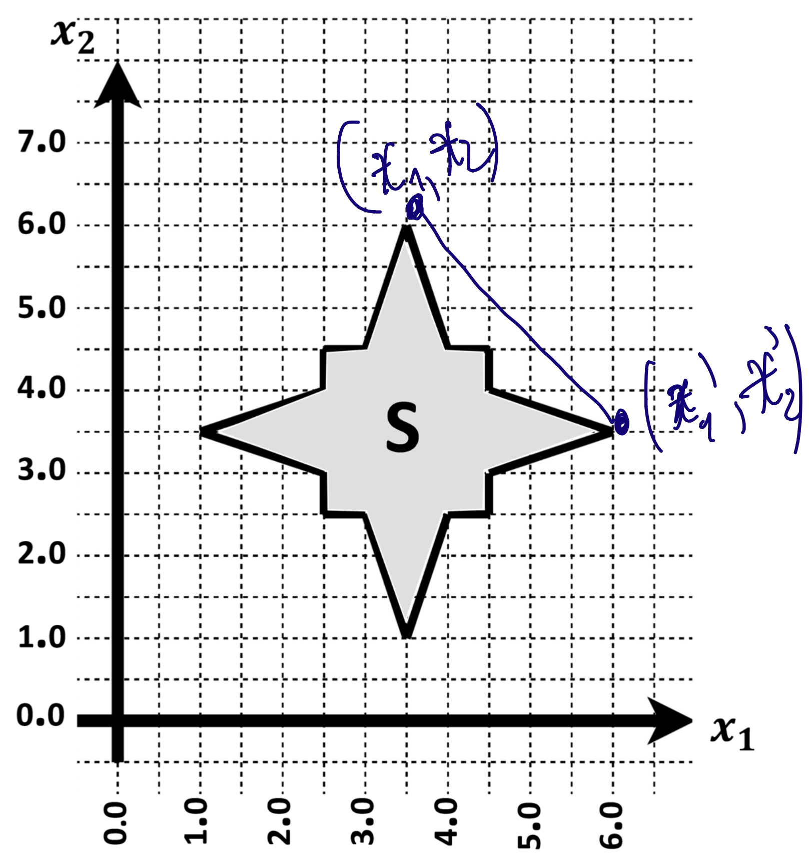
c) The gradient descent procedure from b) will never converge to the minimum, since, on the second coordinate, the gradient is too big (the function is steep) and the learning rate is also too large ($\overset{\text{"}\sigma\text{"}}{1} \cdot \overset{\text{"}\nabla f_y\text{"}}{1} = 1$, this is the least amount that will be added/subtracted from 0, being impossible to reach the min: $-\frac{1}{2}$).

One potential solution would be to make the learning rate smaller such that we could subtract/add a sufficiently small number in order to get $-\frac{1}{2}$.

Problem 4

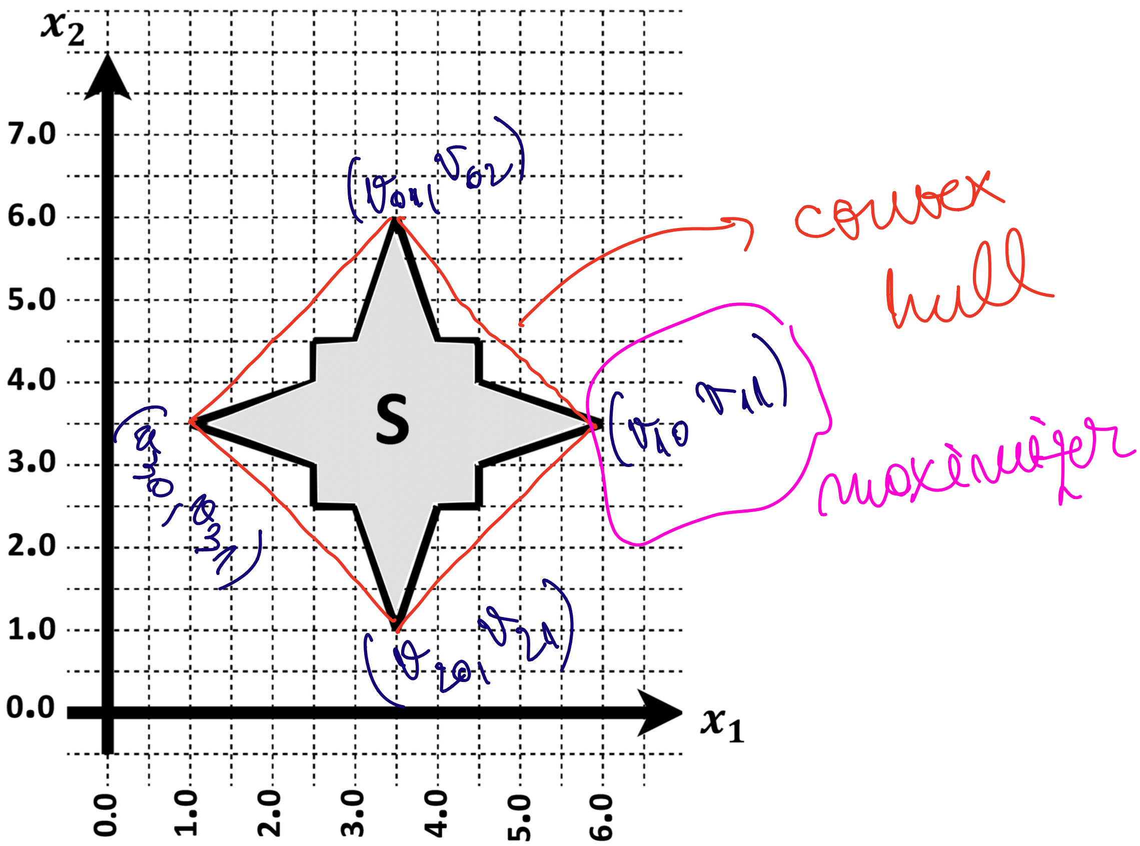
$$f(x_1, x_2) = e^{x_1 + x_2} - 5 \log x_2$$

a)



The region is not convex, since by choosing the 2 points shown above, there are points on that line that are not in the region and, hence, that violate the convexity definition $(\lambda x + (1-\lambda) y \in S)$ for $x, y \in S$

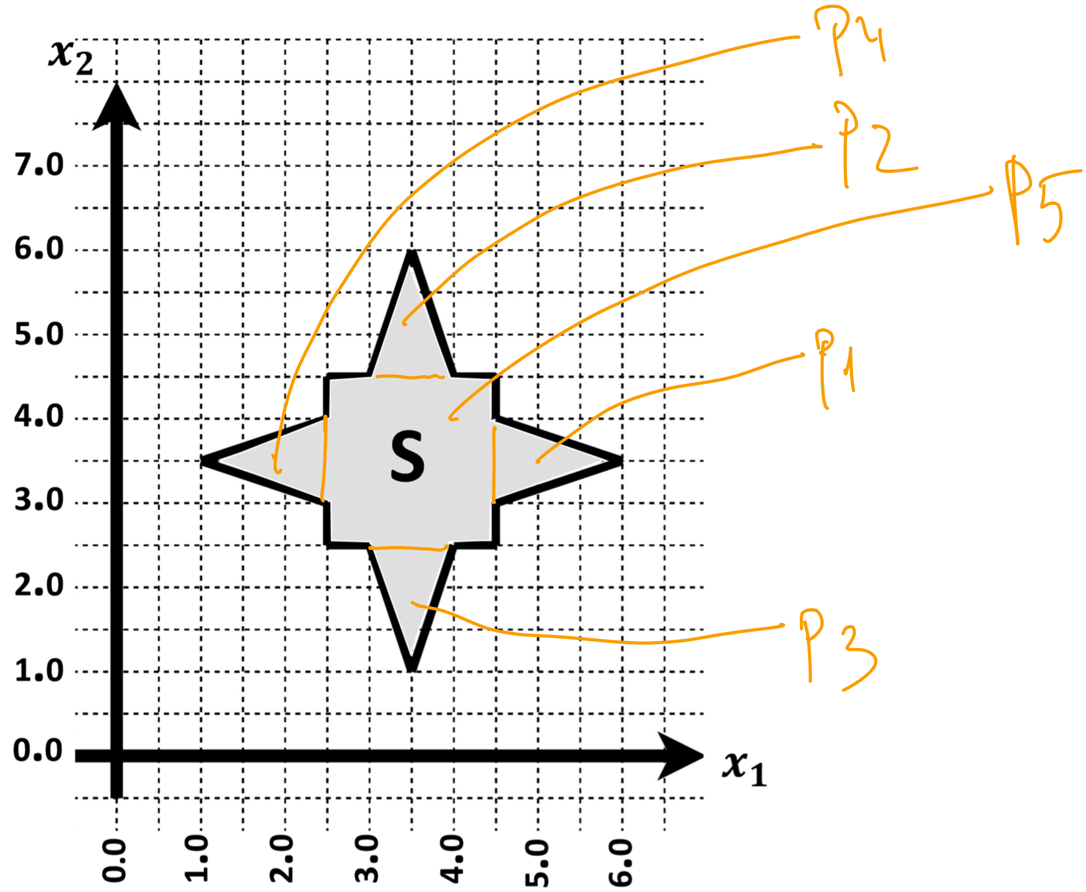
b) For getting the maximum, it is sufficient to check only the vertices of the convex hull.



There are 4 vertices on the convex hull that we need to check. However, intuitively, (v_{10}, v_{11}) is the maximum since $x_1 + x_2$ reaches its maximum in this point and x_2 is small enough such that $f(x_1, x_2)$ is max.

Therefore $(6.0, 3.5)$ is the maximizer.

d)



Since ConvOpt(f, D) works only on convex surfaces, one idea would be to split our surface into multiple convex sub-surfaces, compute the minimum on each one and then find the global minimum by comparing all of the previous minima.

One such potential split is shown in the figure above in which we get 5 convex polygons. Then it's easy to compute all the 5 minima and get the smallest one.