CONVEX OPTIMISATION ASSIGNMENT

TADIPATRI UDAY KIRAN REDDY
EE19BTECH11038

May 3, 2022

Question 1

(a)

Problem (1) is convex.(Positive sum of norm functions is convex)
Problem (2) is convex.(Norm is convex and Norm ball is convex)
Problem (3) is convex.(Positive sum of norm function is convex)
Problem (4) is convex.(Norm is convex and Norm ball is convex)

(b)

$$\nabla Objective = \left(\partial \frac{\overline{x}^T(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})\overline{x} - \overline{y}^T\mathbf{A}\overline{x} + \overline{y}^T\overline{y}}{\partial \overline{x}}\right)^T$$
$$\implies \nabla Objective = 2(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})\overline{x} - 2\overline{y}^T\mathbf{A}$$

(c)

$$\Delta Objective = 0$$
$$\implies \overline{x}^* = (\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y}$$

(d)

Given, $||(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y}||_2 \leq 1$

$$\mathbf{A}^T\left(\overline{y} - \mathbf{A}\overline{x_1^*}\right) = \mathbf{A}^T\left(\overline{y} - \mathbf{A}(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y}\right) = \mathbf{A}^T\overline{y} - \left(\mathbf{A}^T\mathbf{A} + \alpha I\right)(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y} + \alpha(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y}$$
$$\implies \mathbf{A}^T\left(\overline{y} - \mathbf{A}\overline{x_1^*}\right) = \alpha(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\overline{y} \implies ||\mathbf{A}^T\left(\overline{y} - \mathbf{A}\overline{x_1^*}\right)||_2 \leq \alpha$$

Thus $\overline{x_1}^*$ is a feasible point in Problem (2). Therefore the cost with this point will always be less than optimal cost, $\implies ||\overline{x_2}^*||_2 \leq ||\overline{x_1}^*||_2$

(e)

LP

$$||\mathbf{A}^T(\overline{y} - \mathbf{A}\overline{x})||_\infty \leq \alpha \equiv \mathbf{A}^T(\overline{y} - \mathbf{A}\overline{x}) \leq \alpha$$

Let's take epigraph form of objective,

$$||\overline{x}|| \leq t \equiv \overline{x} \leq t; \overline{x} \geq -t$$

Final reformulation is,

$$\begin{aligned} \min \quad & t \\ \text{s.t} \quad & \mathbf{A}^T(\overline{y} - \mathbf{A}\overline{x}) \leq \alpha \\ & \overline{x} \leq t \\ & \overline{x} \geq -t \end{aligned} \equiv \begin{aligned} \min \quad & t \\ \text{s.t} \quad & \begin{bmatrix} -\mathbf{A}^T\mathbf{A} \\ 1 \\ -1 \end{bmatrix} \overline{x} \leq \begin{bmatrix} \alpha - \mathbf{A}^T\overline{y} \\ t \\ t \end{bmatrix} \end{aligned}$$

QP

$$||\overline{x}||_1 \leq t \implies \overline{x} \leq t; \overline{x} \geq -t$$

Now let $\overline{z} = \begin{bmatrix} \overline{x} \\ t \end{bmatrix}$,

$$\begin{gathered} \mathbf{A}\overline{x} - \overline{b} = \begin{bmatrix} \mathbf{A} & 0 \end{bmatrix} \overline{z} - \overline{b} = \mathbf{U}\overline{z} - \overline{b} \\ \alpha t = \begin{bmatrix} \overline{0} & 1 \end{bmatrix} \overline{z} \\ ||\mathbf{A}\overline{x} - \overline{b}||_2^2 + \alpha||\overline{x}||_1 \leq ||\mathbf{U}\overline{z} - \overline{b}||_2^2 + \begin{bmatrix} \overline{0} & 1 \end{bmatrix} \overline{z} \\ \mathbf{U}^T\mathbf{U} = \begin{bmatrix} \mathbf{A} & \overline{0} \\ \overline{0} & 0 \end{bmatrix} \geq 0 \end{gathered}$$

Final reformation is,

$$\begin{aligned} \min \quad & ||\mathbf{U}\overline{z} - \overline{b}||_2^2 + \begin{bmatrix} \overline{0} & 1 \end{bmatrix} \overline{z} \\ \text{s.t} \quad & \begin{bmatrix} \overline{1} & 0 \\ \overline{1} & 0 \end{bmatrix} \overline{z} \geq \begin{bmatrix} t \\ -t \end{bmatrix} \end{aligned}$$

Question 2

$f_2(\overline{x}) = \sum_i ||\mathbf{A_i}\overline{x} - \overline{b}||_2$

Apply epigraph trick on the objective,

$$||\mathbf{A_i}\overline{x} - \overline{b}||_2 \leq t_i$$

$$\begin{aligned} \min \quad & \overline{1}^T\overline{t} \\ \text{s.t} \quad & ||\mathbf{A_i}\overline{x} - \overline{b}||_2 \leq t_i \end{aligned}$$

The above is in form of *SOCP*.

$f_1(\overline{x}) = \sum_i ||\mathbf{A_i}\overline{x} - \overline{b}||_1$

$$\begin{gathered} ||\mathbf{A_i}\overline{x} - \overline{b}||_1 \leq \overline{1}^T\overline{t_i} \\ \implies \mathbf{A_i}\overline{x} - \overline{b} \leq \overline{t_i}; \mathbf{A_i}\overline{x} - \overline{b} \geq -\overline{t_i} \end{gathered}$$

$$\begin{aligned} \min \quad & \sum_i \overline{1}^T\overline{t_i} \equiv T\overline{1} \\ \text{s.t} \quad & \mathbf{A_i}\overline{x} - \overline{b} \leq \overline{t_i} \\ & \mathbf{A_i}\overline{x} - \overline{b} \geq -\overline{t_i} \end{aligned}$$

The above formulation is a *LP*.

$f_\infty(\overline{x}) = \sum_i ||\mathbf{A_i}\overline{x} - \overline{b}||_\infty$

$$\begin{aligned}
&||\mathbf{A_i}\overline{x} - \overline{b}||_\infty \leq t_i \\
\implies &\mathbf{A_i}\overline{x} - \overline{b} \leq \overline{1}t_i; \mathbf{A_i}\overline{x} - \overline{b} \geq -\overline{1}t_i \\
\min \quad &\overline{1}^T\overline{t} \\
\text{s.t} \quad &\mathbf{A_i}\overline{x} - \overline{b} \leq \overline{1}t_i \\
&\mathbf{A_i}\overline{x} - \overline{b} \geq -\overline{1}t_i
\end{aligned}$$

The above formulation is a *LP*.

Question 3

(a)

To prove that the objective function is quasi-concave we ned to show that all α level supersets are convex.

$$\frac{\overline{\mu}^T\overline{x}}{||\mathbf{V}\overline{x}||_2} \geq \alpha$$

$$\implies ||\mathbf{V}\overline{x}||_2 \leq \frac{1}{\alpha}\overline{\mu}^T\overline{x}$$

Given that $\mathbf{V}$ is symmentric which means quadratic part is convex and the above constarints look like SOCP constraints thus it is convex.

(b)

$$\overline{z} = \frac{\overline{x}}{\overline{\mu}^T\overline{x}}$$

$$\implies \frac{\overline{z}}{\overline{1}^T\overline{z}} = \frac{\frac{\overline{x}}{\overline{\mu}^T\overline{x}}}{\frac{\overline{1}^T\overline{x}}{\overline{\mu}^T\overline{x}}}$$

$$\implies \boxed{\overline{x} = \frac{\overline{z}}{\overline{1}^T\overline{z}}}$$

$$\frac{\overline{\mu}^T\overline{x}}{||\mathbf{V}\overline{x}||_2} = \frac{\overline{\mu}^T\frac{\overline{z}}{\overline{1}^T\overline{z}}}{||\mathbf{V}\frac{\overline{z}}{\overline{1}^T\overline{z}}||_2} = \frac{sgn\left(\overline{1}^T\overline{z}\right)}{||\mathbf{V}\overline{z}||_2}$$

Given $\overline{\mu}^T\overline{x} \geq 0$ and $\overline{1}^T\overline{x} = 1$ which means $\overline{1}^T\overline{z} = \frac{1}{\overline{\mu}^T\overline{x}} > 0$.

$$\implies \boxed{\frac{\overline{\mu}^T\overline{x}}{||\mathbf{V}\overline{x}||_2} = \frac{1}{||\mathbf{V}\overline{z}||_2}}$$

$$||\overline{x}||_1 \leq L \implies ||\frac{\overline{z}}{\overline{1}^T\overline{z}}||_1 \leq L$$

$$||\overline{z}||_1 \leq L\overline{1}^T\overline{z}$$

Now transformed problem is,

$$\begin{aligned}
\min \quad &||\mathbf{V}\overline{z}||_2 \\
\text{s.t} \quad &||\overline{z}||_1 \leq L\overline{1}^T\overline{z} \\
&\overline{1}^T\overline{z} \geq 0
\end{aligned}$$

The above transformed problem has both convex objective and constraints thus it is convex optimisation problem.

Question 4

(a)

Lemma: $(\mathbf{A}+\mathbf{B})^{-1} = \mathbf{A}^{-1} - (I + \mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$.
If $g(x)$ is convex then so is $\overline{a}^T g(x)\overline{a}$ because the map linear with respective to $g(x)$.
Now it is sufficient to prove the convexity of $\mathbf{X}^{-1}$. We do this by contradiction assume that the function is not convex which means,

$$(\alpha\mathbf{A})^{-1} + ((1-\alpha)\mathbf{B})^{-1} < (\alpha\mathbf{A} + (1-\alpha)\mathbf{B})^{-1}$$
$$\frac{1}{\alpha}\mathbf{A}^{-1} + \frac{1}{1-\alpha}\mathbf{B}^{-1} < \alpha\mathbf{A}^{-1} - \frac{1-\alpha}{\alpha^2}(I + \frac{1-\alpha}{\alpha}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$$

Since the matrices are positve semi-definite multiplication on inequality will not change the sign.

$$\frac{1}{1-\alpha}\mathbf{B}^{-1} < -\frac{1-\alpha}{\alpha^2}(I + \frac{1-\alpha}{\alpha}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$$
$$\implies \left(\frac{1-\alpha}{\alpha}\mathbf{A}^{-1}\mathbf{B}\right)^2 + \left(\frac{1-\alpha}{\alpha}\mathbf{A}^{-1}\mathbf{B}\right) + I < 0$$

Since $\mathbf{A} \geq 0$ and $\mathbf{B} \geq 0$ so is $\mathbf{A}^{-1}\mathbf{B} \geq 0 \implies \frac{1-\alpha}{\alpha}\mathbf{A}^{-1}\mathbf{B} \geq 0$.Therfore the above obtained sum is just sum of positive semi definite matrices which is positive semi definite but we got negative definite which is a contradiction. Thus our assumption is wrong. Therfore $\mathbf{X}^{-1}$ is convex and so is $\overline{a}^T\mathbf{X}^{-1}\overline{a}$.

(b)

Let $\overline{a_i}$ be ith column of identity matrix then from previous results $\overline{a_i}^T\mathbf{X}^{-1}\overline{a_i}$ is convex, this function just picks (i, i) element of $\mathbf{X}^{-1}$ which is a diagonal element. Thus diagonal elements are convex functions of $\mathbf{X}$.

(c)

$trace(\mathbf{X}^{-1})$ is just sum of diagonal elements of $\mathbf{X}^{-1}$ which are individually convex, since sum of convex functions are convex. $trace(\mathbf{X}^{-1})$ is convex.

(d)

Let transform this problem in epigraph problem.

$$\begin{aligned} \min_{t,\mathbf{X}} \quad & t \\ \text{s.t} \quad & t \geq \overline{a}^T\mathbf{X}^{-1}\overline{a} \equiv \begin{bmatrix} t & \overline{a}^T \\ \overline{a} & \mathbf{X} \end{bmatrix} \geq 0 \\ & \mathbf{A}\mathbf{X} = \mathbf{B} \\ & \mathbf{X} \geq 0 \end{aligned}$$

Now we say $\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \overline{u}^T \end{bmatrix}$, $\overline{u} = \begin{bmatrix} t \\ 0 \\ . \\ 0 \end{bmatrix}$

$$\begin{bmatrix} t & \overline{a}^T \\ \overline{a} & \mathbf{X} \end{bmatrix} \geq 0 \implies \mathbf{UZ} + \mathbf{V} \geq 0$$
$$\mathbf{U} = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 1 \\ 1 & 1 & 1 & . & . & . & 0 \end{bmatrix}$$
$$\mathbf{V} = \begin{bmatrix} 0 & \overline{a}^T \\ \overline{a} & \mathbf{0} \end{bmatrix}$$
$$\mathbf{AX} = \mathbf{B} \implies \mathbf{WZ} = \mathbf{B}$$
$$\mathbf{W} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \end{bmatrix}$$
$$\mathbf{X} \geq 0 \implies \mathbf{YZ} \geq 0$$
$$\mathbf{Y} = \begin{bmatrix} I & 0 \end{bmatrix}$$

Final SDP problem is,

$$\begin{aligned} \min_{t,\mathbf{X}} \quad & \begin{bmatrix} 0 & 1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Z} \\ \text{s.t} \quad & \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \\ -\mathbf{W} \\ \mathbf{Y} \end{bmatrix} \mathbf{Z} \geq - \begin{bmatrix} \mathbf{V} \\ \mathbf{B} \\ -\mathbf{B} \\ \mathbf{0} \end{bmatrix} \end{aligned}$$

(e)

This reformulation is similar to previous one we can rewrite trace as written in (c).

$$trace(\mathbf{X}^{-1}) = \sum_i \overline{a_i}^T \mathbf{X}^{-1} \overline{a_i}$$

Where $\overline{a_i}$ is ith column of Identity matrix. Now we apply the same epigraph trick,

$$\overline{a_i}^T \mathbf{X}^{-1} \overline{a_i} \leq t_i$$
$$\begin{bmatrix} t_i & \overline{a_i}^T \\ \overline{a_i} & \mathbf{X} \end{bmatrix} \geq 0$$

$$\begin{aligned} \min_{t,\mathbf{X}} \quad & \begin{bmatrix} \overline{0}^T & 1 \end{bmatrix} \mathbf{Z} \overline{1} \\ \text{s.t} \quad & \begin{bmatrix} \mathbf{U_1} \\ \mathbf{U_2} \\ \mathbf{U_3} \\ . \\ \mathbf{U_n} \\ \mathbf{W} \\ -\mathbf{W} \\ \mathbf{Y} \end{bmatrix} \mathbf{Z} \geq - \begin{bmatrix} \mathbf{V_1} \\ \mathbf{V_2} \\ \mathbf{V_3} \\ . \\ \mathbf{V_n} \\ \mathbf{B} \\ -\mathbf{B} \\ \mathbf{0} \end{bmatrix} \end{aligned}$$

Where, $\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \overline{t}^T \end{bmatrix}$

Question 5

(a)

$f_1(x)$

This function tries to find a hyperplane which lies above all the sample points. Objective tries to keep this hyperplane as nearer to the sample points as possible.

$f_2(x)$

This is just a convex hull of all samples points but we try to choose a convex hull such that the objective is maximised.

$f_3(x)$

This is also a convex hull of all samples points but with less amplitude of conmbination.

(b)

$f_1(x)$

Here both objective and constraints are affine. Thus the problem is both convex and concave optimisation problem.

$f_2(x)$

In this we are trying to find a surface such that convex combination is maximized. Technically this is also LP with all affine constraints. But in general we can say it is concave.

$f_3(x)$

Here also the problem is LP with affine constraints but in general since it is minization problem we can tell that it is convex problem.

(c)

Since f_1 is LP we can say that it satisfies strong duality.

$$\mathcal{L}(x, m, c, \overline{\lambda}) = mx + c + \sum_i \lambda_i (f(i) - mi - c)$$

$$g(\lambda) = \min_{m,c} m(x - \sum_i \lambda_i i) + c(1 - \sum_i \lambda_i) + \sum_i \lambda_i f(i)$$

This would go to infinity to avoid that, we make coefficients 0 which would yeild.

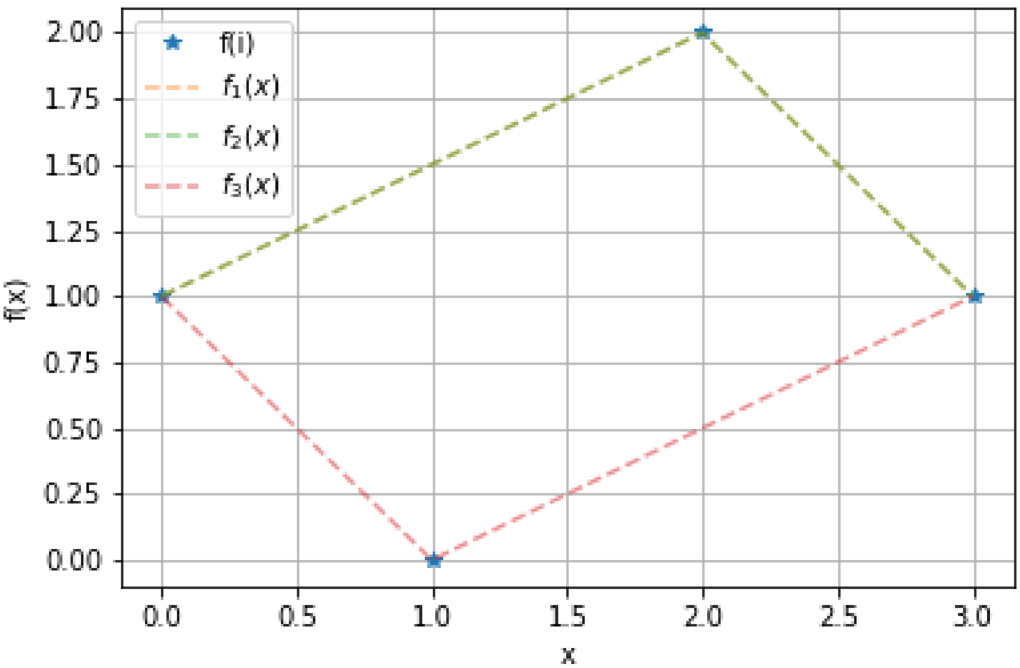
$$\begin{aligned} \min \quad & \sum_i \lambda_i f(i) \\ \text{s.t} \quad & \sum_i \lambda_i = 1 \\ & x = \sum_i \lambda_i i \\ & \lambda_i \geq 0 \end{aligned}$$

Observe that the obtained dual is same as f_2 thus both their optimal values are same.

(d)

As shown previously f_1 and f_2 are primal-dual pairs. If any dual variable is zero which means that constraint has slack,i.e,≤ 0 else constraint $= 0$ this implies that if f(i) has contributed to f_2 which means corresponding α_i is 0.
If three points are non colinear then at max line can pass through only 2 points which means for the third point there is slackness thus it's corresponding dual variable is 0, rest of them need not be zero. Thus in such a senario at most 2 of them can be non-zero.

(e)



Question 6

Primal is

$$\overline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\min \quad \overline{x}^T \begin{bmatrix} 1 & -0.5 \\ -0.5 & 2 \end{bmatrix} \overline{x} + \begin{bmatrix} -1 & 0 \end{bmatrix} \overline{x}$$

$$\text{s.t} \quad \begin{bmatrix} 1 & -2 \\ 1 & 4 \\ 5 & -76 \end{bmatrix} \overline{x} \leq \begin{bmatrix} u_1 \\ u_2 \\ 1 \end{bmatrix}$$

Dual is

(a)

The above objective is in quadratic form and eigen decomposition of the hessian is

$$\begin{bmatrix} 1 & -0.5 \\ -0.5 & 2 \end{bmatrix} = \begin{bmatrix} -0.92387953 & 0.38268343 \\ -0.38268343 & -0.92387953 \end{bmatrix} \begin{bmatrix} 0.79289322 & 0 \\ 0 & 2.20710678 \end{bmatrix} \begin{bmatrix} -0.92387953 & -0.38268343 \\ 0.38268343 & -0.92387953 \end{bmatrix}$$

Here both eigen values are positive, which implies that hessian is positive semidefnite. With linear constraints, The problem is convex and is a QP.

(b)

After solving the problem with $CVXPY$ we get,

$$x_1^* = -2.33; x_2^* = -0.167$$
$$\lambda_1^* = 2.864; \lambda_2^* = 2.298; \lambda_3^* = 0.067$$

(c)

KKT Conditions

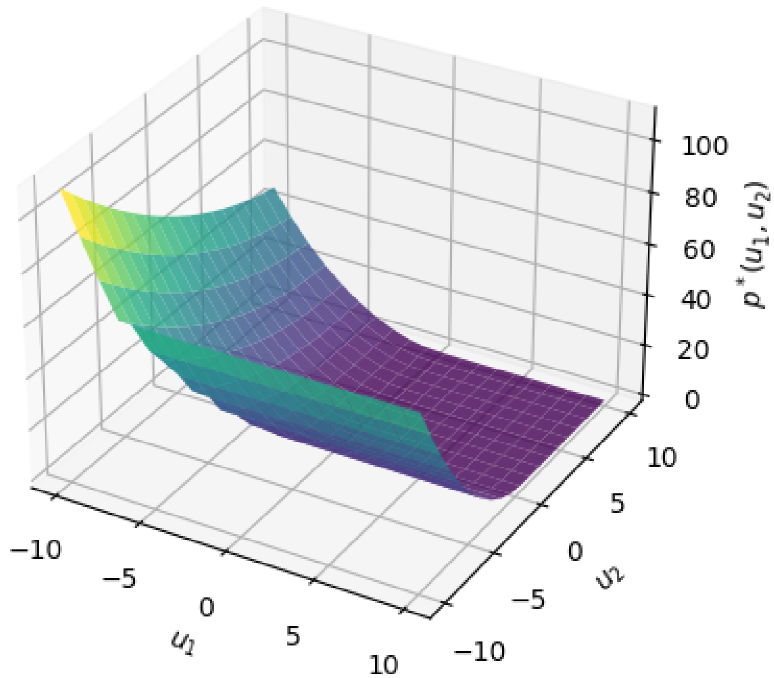
1. $f_i(x^*) \leq 0$, Satisfied.

$$\begin{bmatrix} 1 & -2 \\ 1 & 4 \\ 5 & -76 \end{bmatrix} \begin{bmatrix} -2.33 \\ -0.167 \end{bmatrix} \leq \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix} \implies \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix} \leq \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}$$

2. $\lambda_i^* \geq 0$, Satisfied.
3. $\lambda_i^* f_i(x^*) = 0$, Satisfied.

$$\begin{bmatrix} 5.167 & 1.834 & 0 \end{bmatrix} \left(\begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix} - \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix} \right) = 0$$

(d)

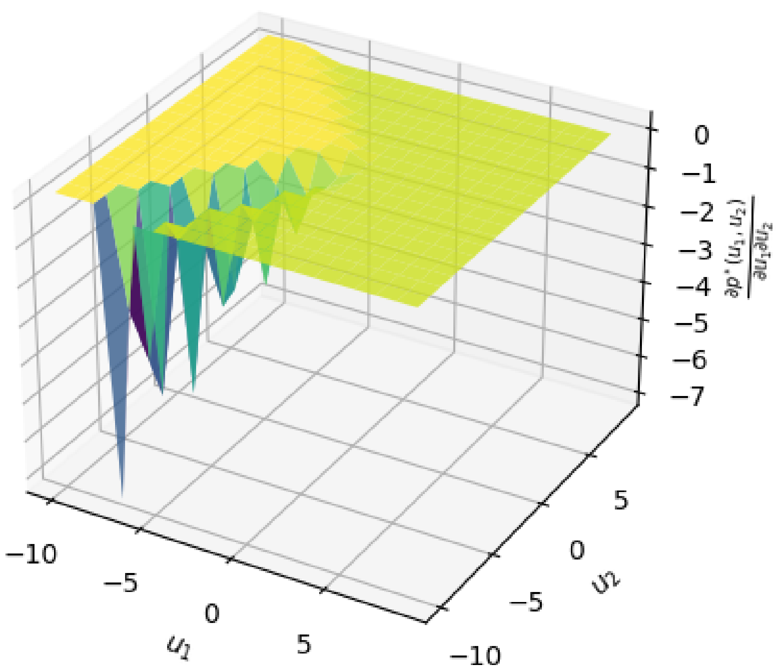


(e)

From above graph it seems like $p^*(u_1, u_2)$ is a convex function.

(f)

Partial numerical derivatives at $u_1 = -2$ is -2.8679 and $u_2 = -3$ is -2.294 and corresponding lambda are 2.864 and 2.298 we see that they are almost equal in magnitude.



0
−1
−2
−3
−4
−5
−6
−7
$\frac{\partial p^*(u_1,u_2)}{\partial u_1 \partial u_2}$
−10
−5
0
5
u_1
5
0
−5
−10
u_2