ECE-6500

ASSIGNMENT-3

- UTKARSH PRATAP SINGH JADON
500711257

Q2) Prove: Primal-Dual Optimality Theorem

According to necessary and sufficient conditions for the optimality of primal-dual pairs,

$$L(x^*, \mu, \lambda) \leq L(x^*, \mu^*, \lambda^*)$$

For some point on the Lagrangian funct' $\mu = 0$ & $\lambda = 0$, we have

$$L(x^*, 0, 0) \leq L(x^*, \mu^*, \lambda^*)$$

Above eqn can be re-written as,

$$\cancel{f(x^*)} \leq \cancel{f(x^*)} + \mu^{*T} g(x^*) + \lambda^{*T} \cancel{h(x^*)}^{\,0}$$

$$\Rightarrow \mu^{*T} g(x^*) \geq 0 \quad \text{——— (1)}$$

But from i) Primal feasibility : $g(x^*) \leq 0$
ii) Dual feasibility : $\mu^* \geq 0$

$$\hookrightarrow \mu^{*T} g(x^*) \leq 0 \quad \text{——— (2)}$$

Eqn (1) & (2) are satisfied only when equality is attained, i.e.

$$\boxed{\mu^* g(x^*) = 0}$$

Hence, Complementary Slackness Condition is proved.

Q5) $G = \{ (u,t) \mid \exists x \in D,\ f_0(x) = t,\ f_1(x) = u \}$
$A = \{ (u,t) \mid \exists x \in D,\ f_0(x) \leq t,\ f_1(x) \leq u \}$

a) Minimize x subject to $x^2 \leq 1$

G is a ~~a~~ curve
$G = \{ (u,t) \mid u \in D,\ u = f_1(t) \}$
A is set of points ~~at~~ north-east (above & to the right) of G.

$x^* = -1$, $\lambda = 1$, $p^* = -1$, $d^* = -1$
Problem is convex
Strong duality holds
Slater's condition holds

b) Minimize x subject to $x^2 \leq 0$

G is a curve
$G = \{ (u,t) \mid u \in D,\ u = f_1(t) \}$
A is set of points above & to the right of G.

$x^* = 0$, $p^* = 0$, $d^* = 0$. Dual optimum is not achieved.
Problem is convex.
Strong duality holds
Slater's condition does not hold

d) Minimize x subject to $f_1(x) \le 0$ where

$$f_1(x) = \begin{cases} -x+2 & x \geq 1 \\ x & -1 \le x \le 1 \\ -x-2 & x \le -1 \end{cases}$$

G is a curve, $G = \{(u,t) \mid u \in D,\ u = f_1(t)\}$
A is set of points above and to the right of G.

$x^* = -2$, $p^* = -2$, $\lambda^* = 1$, $d^* = -2$
Problem is not convex.
Strong duality holds

Q6) Minimize $-3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3)$
subject to $x_1^2 + x_2^2 + x_3^2 = 1$

a) KKT conditions:
$x_1^2 + x_2^2 + x_3^2 = 1$, $(-3+\nu)x_1 + 1 = 0$, $(1+\nu)x_2 + 1 = 0$, $(2+\nu)x_3 + 1 = 0$

b) KKT conditions imply that, $\nu \neq 3$, $\nu \neq -1$, $\nu \neq -2$
Therefore, by eliminating x, KKT conditions can be reduced to nonlinear eq^n in ν:

eq. $$\frac{1}{(-3+\nu)^2} + \frac{1}{(1+\nu)^2} + \frac{1}{(2+\nu)^2} = 1$$

Above eq^n gives four sol^n: $\nu = -3.15$, $\nu = 0.22$, $\nu = 1.89$, $\nu = 4.04$
Corresponding values of x are:
$x = (0.16, 0.47, -0.87)$, $x = (0.36, -0.82, 0.45)$
$x = (0.90, -0.35, 0.26)$, $x = (-0.97, -0.20, 0.17)$

c) Largest of the four values, $\nu^* = 4.04$
Simplest way to see this is by comparing objective values of four solutions x;

$f_0(x) = 1.17$, $f_0(x) = 0.67$, $f_0(x) = -0.56$, $f_0(x) = -4.70$

$$\nabla^2 f_0(x^*) + \nu^* \nabla^2 f_1^*(x^*) \succeq 0$$

because x^* is a minimizer of $L(x, \nu^*)$

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \nu^* \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \succeq 0$$

$\therefore \quad \nu^* \geq 3$

Q4) Minimize e^{-x_2}

Subject to $\|x\| \leq x_1,\ x_2 \geq 0$

a) Constraint set is $\|x\| \leq x_1$ & $x_2 \geq 0$

$\Rightarrow \sqrt{x_1^2 + x_2^2} \leq x_1$ & $x_2 \geq 0$

$\Rightarrow x_1^2 + x_2^2 \leq x_1^2$ & $x_2 \geq 0$

$\Rightarrow x_2^2 \leq 0$ & $x_2 \geq 0$

$\Rightarrow x_2 \leq 0$ & $x_2 \geq 0$

$\therefore x_2 = 0$

Optimal value $p^* = e^{-0} \Rightarrow \underline{\underline{p^* = 1}}$

e) Relax constraint $x_2 \geq 0$

$\therefore$ Constraint set, now is $x_2 \leq 0$ only

Here, $x_2^* = 0$

$\therefore \boxed{p^* = 1}$ = Optimum value

Now, $q^* = \max\{q(\mu)\}$
s.t. $\mu \geq 0$

And, $q(\mu) = \inf_{x \in X} L(x, \mu)$

where $L(x, \mu) = f(x) + \mu^T g(x)$

$\therefore L(x, \mu) = e^{-x_2} + \mu x_2$

$q(\mu) = \inf\{e^{-x_2} + \mu x_2\}$

$$\frac{\partial}{\partial x_2}(e^{-x_2} + \mu x_2) = -e^{-x_2} + \mu = 0$$
$$\Rightarrow \mu = e^{-x_2}$$

Substituting the value of μ in $q(\mu)$, we get:

$q(\mu) = \mu - \mu \ln \mu$

Now, $q^* = \max\{q(\mu)\}$

$$\frac{dq(\mu)}{d\mu} = 1 - \ln\mu - \cancel{\mu}\frac{1}{\cancel{\mu}} = -\ln\mu = 0$$
$$\Rightarrow \mu = 1$$

Substituting $\mu = 1$ in $q(\mu) \Rightarrow q(\mu) = 1 - 1.\ln 1 \Rightarrow \underline{\underline{q^* = 1}}$

Here, $p^* = q^* = 1$

$\therefore$ No duality gap exists

c) Relax costraint $|x_1| \leq x_1$ only

$\therefore$ Costraint set now is only $x_2 \geq 0$

We know that primal problem is given by

min $f(x)$

s.t. $g_i(x) \leq 0$

$\therefore$ Here, $g(x) = -x_2$

i.e. $x_2 \geq 0 \Rightarrow g(x) \leq 0$

Here, as $x_2 \to \infty$

$$p^* = e^{-x_2} = e^{-\infty} = \frac{1}{e^{\infty}} = \frac{1}{\infty} = 0$$

$\therefore$ $\boxed{p^* = 0}$

$$L(x, \mu) = f(x) + \mu^T g(x)$$
$$= e^{-x_2} - \mu x_2$$

$$q(\mu) = \inf (e^{-x_2} - \mu x_2)$$

$$\frac{\partial}{\partial x_2}(e^{-x_2} - \mu x_2) = -e^{-x_2} - \mu = 0$$
$$\Rightarrow \mu = -e^{-x_2}$$
$$\Rightarrow e^{-x_2} = -\mu$$

We know that $\mu \geq 0$

But exponential function can not output negative value,

$\therefore$ No μ attains q^*

d) In e), Optimum value is attained at $x_2^* = 0$, which is $p^* = 1$. And $q^* = 1$, $\therefore$ Strong duality holds.

In c), Optimum value is attained at $x_2^* \to \infty$, which is $p^* = 0$. And no μ attains q^*

Q7] Minimize $\sum_{i=1}^{n} \left[x_i \log\left(\frac{x_i}{c_i}\right) - x_i \right]$

Subject to $x \geq 0$ $\quad Ax = b \quad \lambda \in \mathbb{R}$

$\hookrightarrow \mu \in \mathbb{R}^k$

Where $c \geq 0$, $b \in \mathbb{R}^k$ and $A \in \mathbb{R}^{k \times n}$ matrix with i^{th} column denoted as $A_i \in \mathbb{R}^k$

We know that,

$$q(\mu, \lambda) = \inf_{x \in X} L(x, \mu, \lambda)$$

where $L(x, \mu, \lambda) = f(x) + \sum_i \mu^T g(x) + \sum \lambda^T h(x)$

Here $f(x) = \sum_i \left[x_i \log\left(\frac{x_i}{c_i}\right) - x_i \right]$

$g(x) = -x$

$h(x) = Ax - b$

$\therefore L(x, \mu, \lambda) = \sum_i \left[x_i \log\left(\frac{x_i}{c_i}\right) - x_i \right] - \sum \mu_i x_i + \lambda^T (Ax - b)$

$$q(\mu, \lambda) = \inf_{x \geq 0} \left\{ \sum_i \left[x_i \log\left(\frac{x_i}{c_i}\right) - x_i \right] - \sum_i \mu_i x_i + \lambda^T (Ax - b) \right\}$$

$$= \inf_{x \geq 0} \left\{ \sum_i x_i \left[\log x_i - \log c_i - 1 - \mu + \lambda^T A_i \right] \right\} - \lambda^T b$$

$$= \sum_i \inf_{x_i \geq 0} \left\{ x_i \left[\log x_i - \log c_i - 1 - \mu + \lambda^T A_i \right] \right\} - \lambda^T b$$

$$\frac{\partial q(\mu, \lambda)}{\partial x} = \left(\log x_i - \log c_i - 1 - \mu + \lambda^T A_i \right) + \left(x_i \cdot \frac{1}{x_i} \right) = 0$$

$$\Rightarrow \log x_i - \log c_i \cancel{-1} - \mu + \lambda^T A_i \cancel{+1} = 0$$

$$\Rightarrow \log \left(\frac{x_i}{c_i} \right) = \mu - \lambda^T A_i$$

$$\Rightarrow x^* = c_i e^{(\mu - \lambda^T A_i)}$$

Substituting the value of x^* in $q(\mu, \lambda)$, we get:

$$q(\mu, \lambda) = \sum_i c_i \left(\exp(\mu - \lambda^T A_i) \right) \left[\cancel{\log c_i} + \cancel{\mu} \cancel{- \lambda^T A} \cancel{- \log c} - 1 \cancel{- \mu} \cancel{+ \lambda^T A} \right] - \lambda^T b$$

$$q(\mu, \lambda = - \sum c_i e^{\mu - \lambda^T A_i} - \lambda^T b$$

Dual Problem:

$$q^* = \max_{\mu \geq 0 \quad \lambda \in \mathbb{R}^k} q(\mu, \lambda)$$

$$= \max \left\{ - \sum_i c_i e^{\mu - \lambda^T A_i} - \lambda^T b \right\}$$

Now, $\frac{dq}{d\mu} = -\sum c_i e^{\mu - \lambda^T A_i} = 0$

which is not possible as exponential funct° does not give value of zero for real valued domain

$\therefore$ No μ attains q^*
Hence dual problem does not exist

Q3) Minimize x^2+1
s.t. $(x-2)(x-4) \le 0$ with $x \in \mathbb{R}$

a) Feasible set is the interval $[2,4]$
Optimal point $x^* = 2$
Optimal value $p^* = 5$

b) $L(x,\lambda) = x^2+1 + \lambda(x-2)(x-4)$
$= (1+\lambda)x^2 - 6\lambda x + 1 + 8\lambda$

Plot of the graph is at the end of the question.
We can see from it that as λ increases from 0 to 2, it reaches max at $\lambda=2$ and then decreases again as λ increases above 2.

$\therefore$ $p^* = g(\lambda)$ where $\lambda = 2$

$\frac{\partial L(x,\lambda)}{\partial x} = 0$

$\bar{x} = \frac{3\lambda}{1+\lambda}$ gives lowest Lagrangian

$\therefore$ $g(\lambda) = \begin{cases} \frac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda, & \lambda > -1 \\ -\infty, & \lambda \le -1 \end{cases}$

From the graph, we can see that $p^* = 5$ for $\lambda = 2$.

c) Lagrange dual problem is

Maximize $-9\lambda^2/(1+\lambda) + 1 + 8\lambda$

Subject to $\lambda \geq 0$

The dual optimum occurs at $\lambda = 2$ and $d^* = 5$

From the graph, we can see that it is a concave maximization problem.

As $p^* = q^*$, we can say that strong duality holds

d) Perturbed problem is infeasible for $u < -1$, since

$\inf_x (x^2 - 6x + 8) = -1$

for $u \geq -1$, feasible set is the interval

$$[3 - \sqrt{1+u},\ 3 + \sqrt{1+u}]$$

given by the two roots of $x^2 - 6x + 8 = u$.

for $-1 \leq u \leq 8$ the optimum is

$x^*(u) = 3 - \sqrt{1+u}$

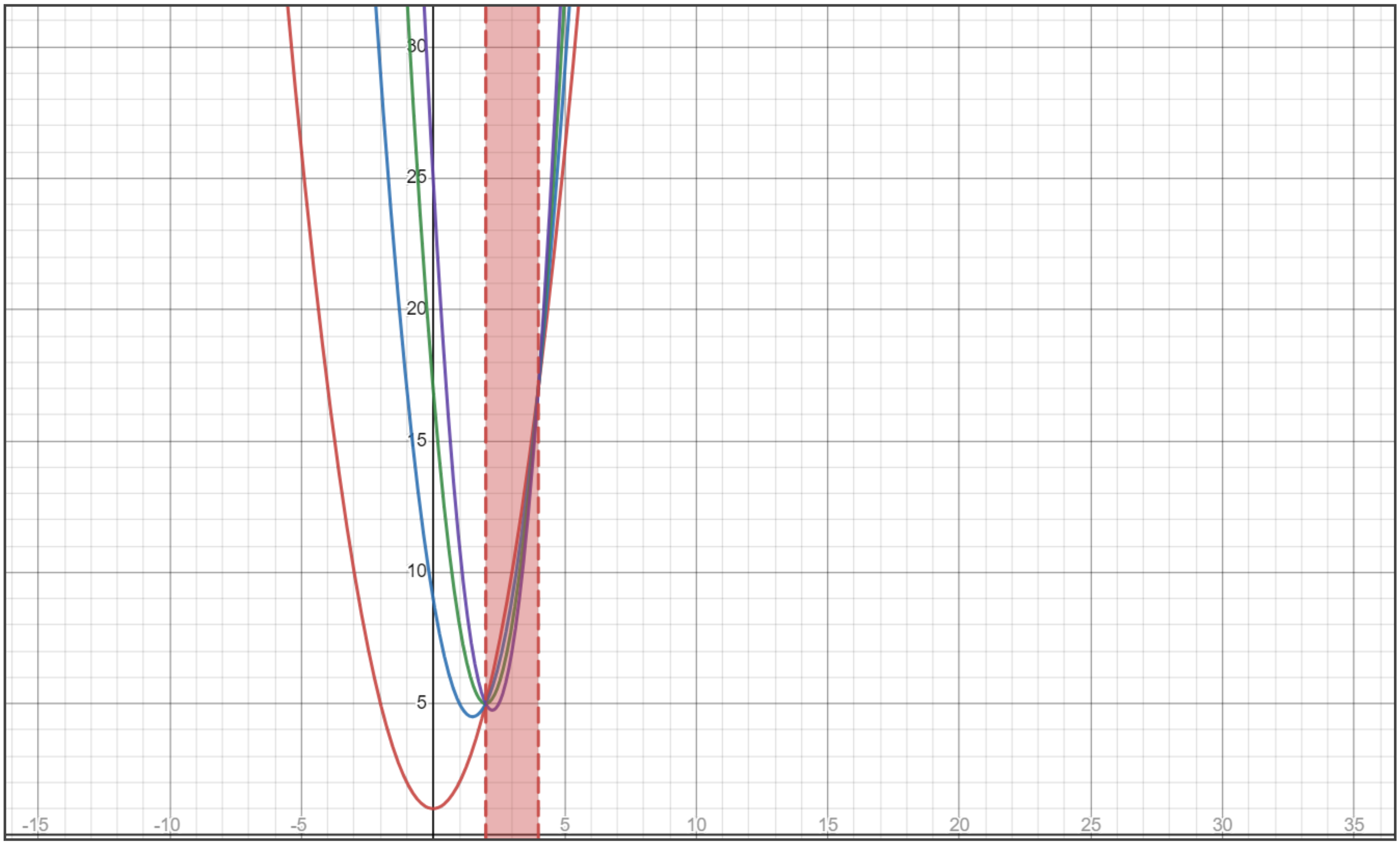
for $u \geq 8$, optimum is the unconstrained minimum of f_0

i.e. $x^*(u) = 0$.

$$\therefore \quad p^*(u) = \begin{cases} \infty, & u < -1 \\ 11 + u - 6\sqrt{1+u} & -1 \leq u \leq 8 \\ 1 & u \geq 8 \end{cases}$$

Finally, we note that $p^*(u)$ is a differentiable funct^n^ $\forall u$

$$\frac{dp^*(0)}{du} = -2 = -\lambda^*$$



1

$x^2 + 1$

2

$x^2 + 1 + (1 \cdot (x - 2) \cdot (x - 4))$

3

$x^2 + 1 + (2 \cdot (x - 2) \cdot (x - 4))$

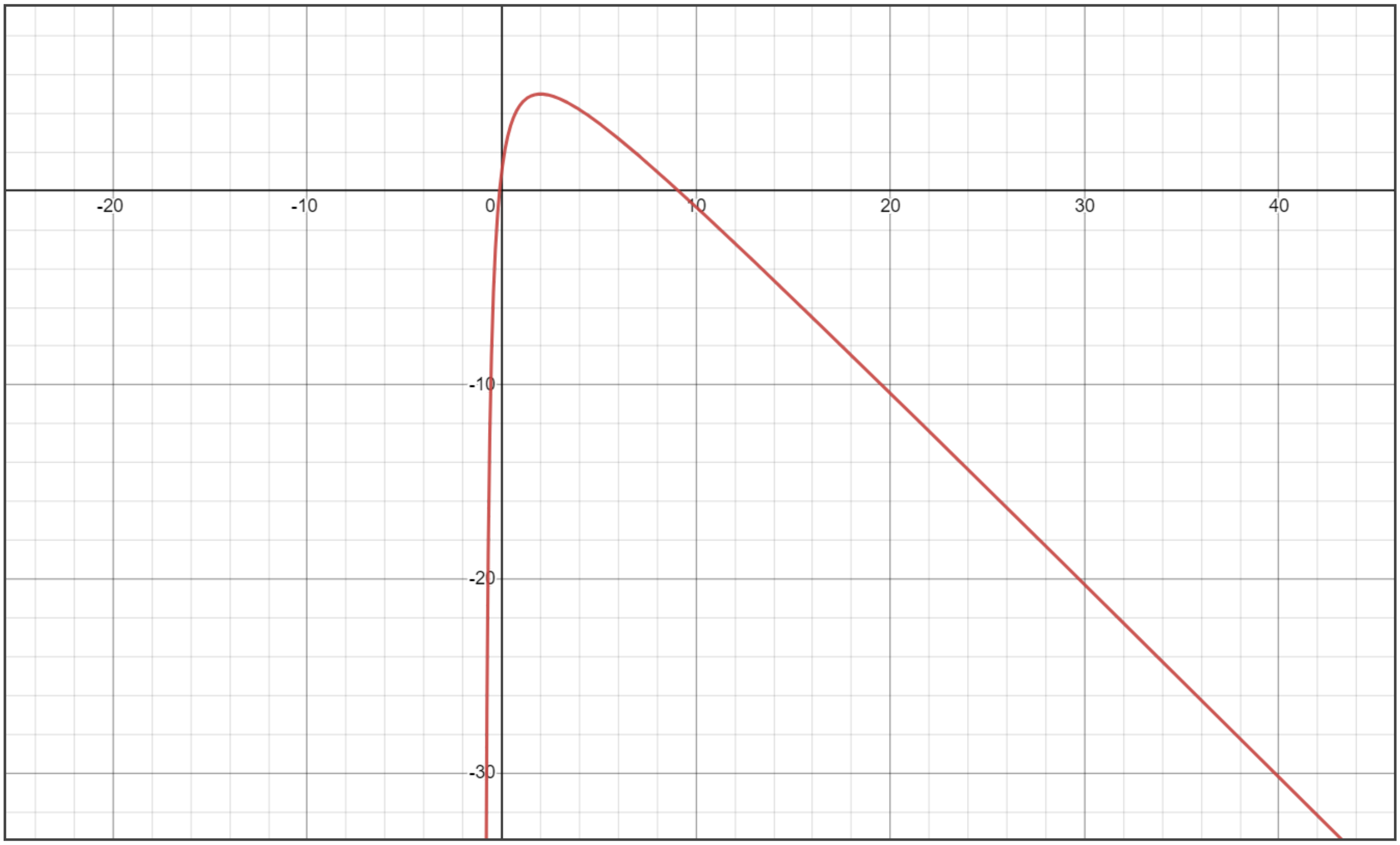
4

$x^2 + 1 + (3 \cdot (x - 2) \cdot (x - 4))$

5

$2 < x < 4$

6



1

$$-\frac{9x^2}{1+x}+1+8x$$

Q1) A non-vertical hyperplane with normalize $\begin{pmatrix} \tilde{\mu} \\ 1 \end{pmatrix}$ that achieves q^*

~~Then, From~~ We know that,

$$q\begin{pmatrix} \tilde{\mu} \\ 1 \end{pmatrix} = \inf_{x \in X} \left\{ f(x) + \tilde{\mu}^T g(x) \right\}$$

with $\tilde{\mu}^T \geq 0$

for $x^* \in C$

$$q(\tilde{\mu}) \leq f(x^*) + \tilde{\mu}^T g(x^*)$$

$$q^* \leq f(x^*) + \sum_{j=1}^{m} \mu_j g_j(x^*)$$

$$\sum_{j=1}^{m} \mu_j g_j(x^*) \geq q^* - f(x^*)$$

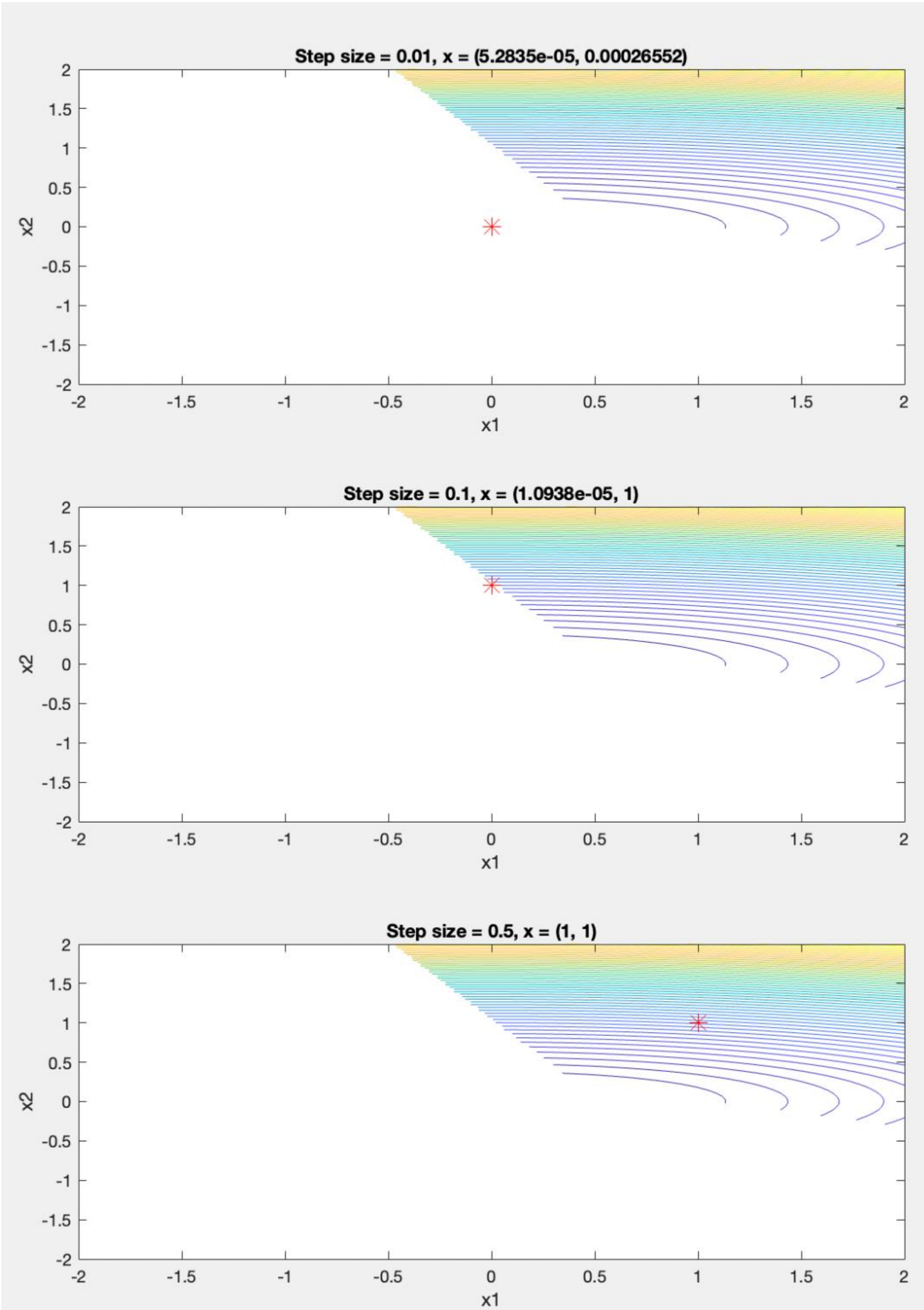
By looking at the eqⁿ to prove, multiply both sides with -1 to get

$$\sum_{j=1}^{m} \mu_j (-g_j(x)) \leq f(x^*) - q^*$$

But $x^* = y \Rightarrow$ Satisfy the Slater's Condition i.e. $g(y) < 0$

$$\sum_{j=1}^{m} \mu_j \leq \frac{f(y) - q^*}{-g_j(y)}$$

Question 8



Step size = 0.01, x = (5.2835e-05, 0.00026552)
Step size = 0.1, x = (1.0938e-05, 1)
Step size = 0.5, x = (1, 1)
x1
x2

The gradient projection method is a popular iterative optimization algorithm for solving constrained optimization problems. Unlike gradient descent, the gradient projection method does not require a hyperparameter for step size or learning rate. Instead, the step size is determined by projecting the gradient onto the feasible region, which is defined by the constraints of the optimization problem. The algorithm repeats this projection step to ensure that each iterate remains feasible. The convergence of the gradient projection method is less sensitive to the choice of step size because the projection step determines the step size. Consequently, the gradient projection method can converge with a large step size since the projection step limits the distance between the iterates and the feasible region.

```
% Define the objective function and its gradient
f = @(x) x(1)^2 + 9*x(2)^2;
grad_f = @(x) [2*x(1); 18*x(2)];

% Define the projection operator onto the feasible set
proj = @(x) max(x, 0);

% Define the constraint functions and their gradients
g1 = @(x) 2*x(1) + x(2) - 1;
g2 = @(x) x(1) + 3*x(2) - 1;
grad_g1 = @(x) [2; 1];
grad_g2 = @(x) [1; 3];

% Define the Lagrange multiplier function and its gradient
L = @(x,lambda) f(x) + lambda(1)*max(-g1(x), 0) + lambda(2)*max(-g2(x), 0);
grad_L = @(x,lambda) grad_f(x) + lambda(1)*grad_g1(x).*((g1(x) < 0)) + lambda(2)*grad_g2(x).*((g2(x) < 0));

% Set the initial guess, maximum number of iterations, and tolerance level
x0 = [1; 1];
maxIter = 1000;
tol = 1e-6;

% Set the step sizes to test
alphas = [0.01, 0.1, 0.5];

% Create a figure to show the contour plot and solution trajectory for each step size
figure;
for i = 1:length(alphas)
    % Initialize the variables for this step size
    alpha = alphas(i);
    x = x0;
    lambda = [0; 0];

    % Create a contour plot of the objective function and constraints
    subplot(length(alphas), 2, 2*i-1);
    x1_vals = linspace(-2, 2, 100);
    x2_vals = linspace(-2, 2, 100);
    [X1,X2] = meshgrid(x1_vals,x2_vals);
    Z = X1.^2 + 9*X2.^2;
    g1_vals = 2*X1 + X2 - 1;
    g2_vals = X1 + 3*X2 - 1;
    Z(g1_vals < 0 | g2_vals < 0) = NaN;
    contour(X1,X2,Z,50);
    hold on;
    xlabel('x1');
    ylabel('x2');
    title(['Step size = ', num2str(alpha)]);

    % Run the projection gradient method
    for iter = 1:maxIter
        % Compute the projected gradient
        proj_grad = proj(x - alpha*grad_L(x, lambda));

        % Update the Lagrange multipliers
        lambda = max(lambda - alpha*[max(-g1(x), 0); max(-g2(x), 0)], 0);

        % Update the solution
        x_old = x;
        x = proj(x - alpha*proj_grad);

        % Check for convergence
        if norm(x - x_old) < tol
            break;
```

```
        end
    end

    % Plot the solution trajectory on the contour plot
    plot(x(1), x(2), 'r*', 'MarkerSize', 10);
    xlabel('x1');
    ylabel('x2');
    title(['Step size = ', num2str(alpha), ', x = (', num2str(x(1)), ', ', num2str(x(2)), ')']);
end
```