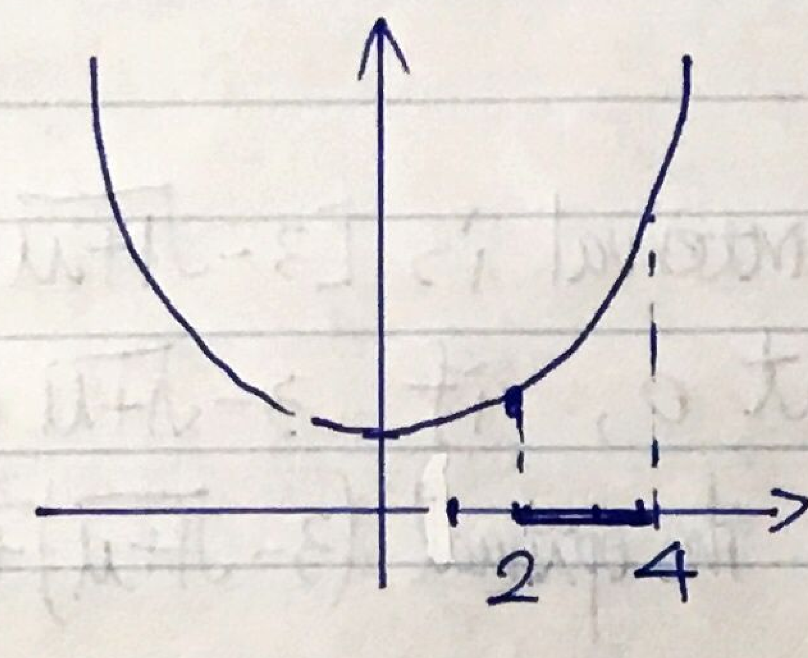
291A Homework 4 A53204772 Hao-en Sung

Exercise from textbook Chapter Five

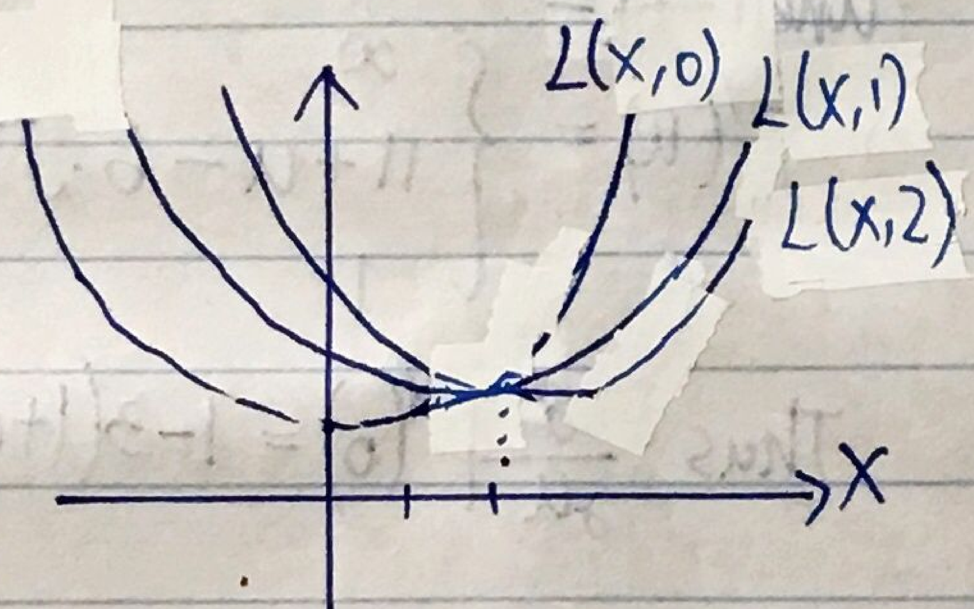
5.1

(a) Since $(x-2)(x-4) \le 0$, I know that $2 \le x \le 4$.



With the figure, it is obvious that the optimal happens when $x=2$, and the optimal value will be $2^2+1=5$. *

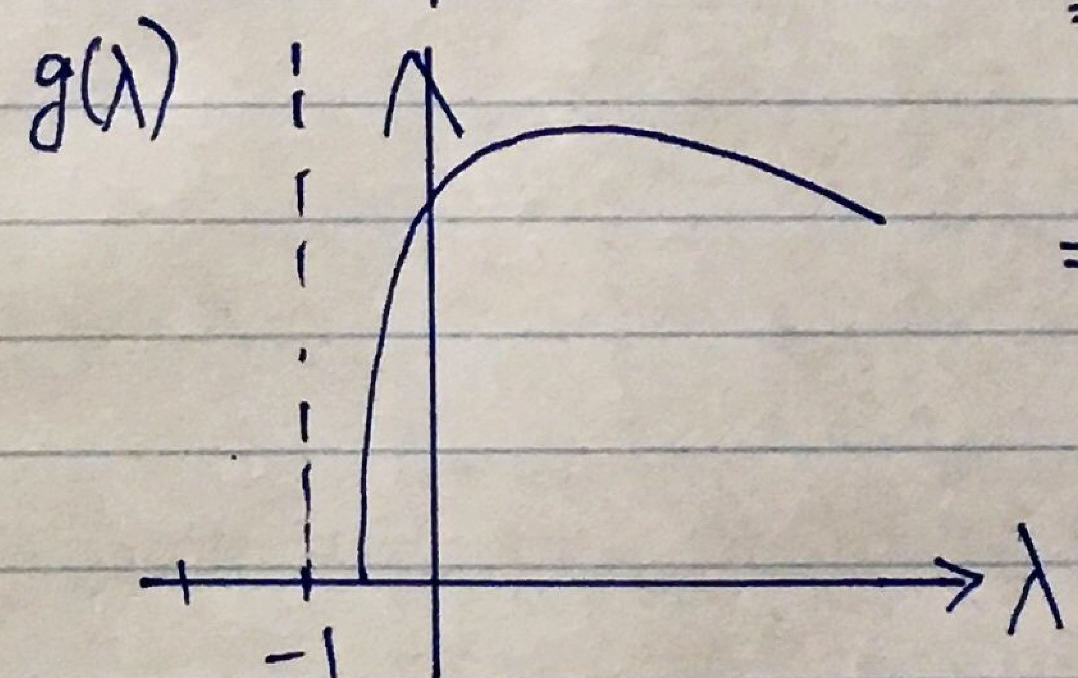
(b) $L(x,\lambda) = x^2+1+\lambda(x-2)(x-4) = (\lambda+1)x^2 - 6\lambda x + (1+8\lambda)$



$\frac{\partial}{\partial x} L(x,\lambda) = 2(\lambda+1)\cdot x - 6\lambda = 0$

$\Rightarrow x = \frac{3\lambda}{\lambda+1}$

$\Rightarrow \inf_x L(x,\lambda) = \begin{cases} -\infty & ; \lambda < -1 \\ \frac{-9\lambda^2}{\lambda+1} + 1 + 8\lambda & ; \lambda \ge -1 \end{cases}$



$\Rightarrow g(\lambda) = \inf_x L(x,\lambda)$

From the figure, we know that when $\lambda=2$, $g(\lambda)$ has the optimal value $p^* = 5$

*

(c) Dual problem

$$\max_\lambda \; g(\lambda) = \frac{-9\lambda^2}{\lambda+1} + 1 + 8\lambda,$$

s.t. $\lambda \ge 0$

Since the optimal $d^*=5$ happens when $\lambda=2$, we can directly infer the strong duality holds for the problem. *

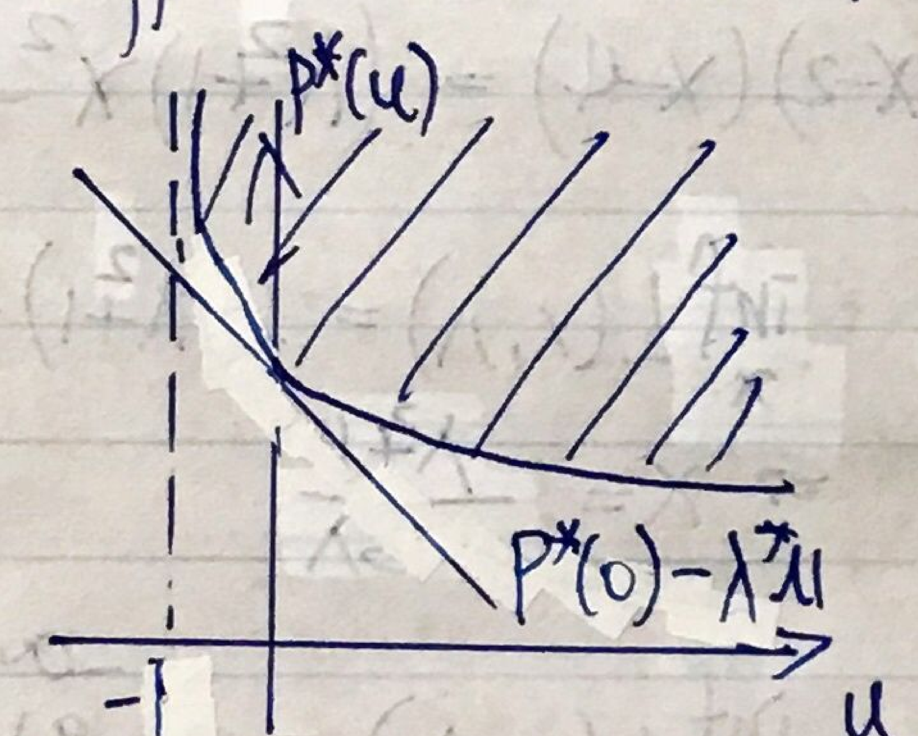
(d) From the constraint $(x-2)(x-4) \leq u$, I have $x^2-6x+(8-u) \leq 0$.

Then I know $6^2-4\cdot(8-u) \geq 0$

$\Rightarrow 8-u \leq 9$

$\Rightarrow u \geq -1$

For $u \geq -1$, the feasible interval is $[3-\sqrt{1+u}, 3+\sqrt{1+u}]$.
Since the optimal happens at 0, if $3-\sqrt{1+u} \leq 0$, i.e. $u \geq 8$, the optimal value is 0; or the optimal $(3-\sqrt{1+u})^2+1 = 11+u-6\sqrt{1+u}$ happens when $x = 3-\sqrt{1+u}$.



In a clearer statement, I have

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

Thus $\frac{\partial}{\partial u}p^*(0) = 1-3(1+0)^{-\frac{1}{2}} = -2 = -\lambda^*$ *

5.9

(a) From hint, I need to first prove that $\begin{bmatrix} \sum_{k=1}^{m} a_k a_k^T & a_i \\ a_i^T & 1 \end{bmatrix} \succeq 0$.

It is pretty straightforward to tell so if we decompose it into summation over two matrix, i.e.

$$\begin{bmatrix} \sum_{k=1}^{m} a_k a_k^T & a_i^T \\ a_i^T & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \sum_{k\neq i}^{m} a_k a_k^T & 0 \\ 0 & 0 \end{bmatrix}}_{\succeq 0} + \underbrace{\begin{bmatrix} a_i \\ 1 \end{bmatrix}\begin{bmatrix} a_i \\ 1 \end{bmatrix}^T}_{\succeq 0} \succeq 0.$$

Thus, the Schur complement of the $(1,1)$ block in this matrix is PSD, i.e., $1 - a_i^T\left(\sum_{k=1}^{m} a_k a_k^T\right)^{-1} a_i \geq 0$, which is feasible.

(b) Firstly, if I consider $\lambda = t \cdot 1$, I have

$$g(\lambda) = \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log t - mt + n.$$

To get best t, I need to let $\frac{\partial g(\lambda)}{\partial t} = 0$, i.e.

$$\frac{\partial g(\lambda)}{\partial t} = \frac{n}{t} - m = 0 \Rightarrow t = \frac{n}{m},$$

and I have

$$g(\lambda) = \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log\left(\frac{n}{m}\right).$$

Since the primal objective value for $X = X_{sm}$ is $-\log\det\left(\sum_{i=1}^{m} a_i a_i^T\right)^{-1}$, the gap between X_{sm} and λ is $n\log\left(\frac{n}{m}\right)$.

In other words, X_{sm} is no more than $n\log\left(\frac{n}{m}\right)$ suboptimal.

In the view that the volumn V of a ellipsoid with matrix X is given by $V = \exp\left(\frac{-\theta}{2}\right)$, where $\theta = -\log\det X$ in this case, I know the bound will be bounded by $\left(\frac{m}{n}\right)^{\frac{n}{2}}$.

5.40

The first step is pretty straightforward. I just need to replace $\lambda_{max}\left(\sum_{i=1}^{p} x_i v_i v_i^T\right)$ with t, which indicates that

$$\lambda_{max}\left(\sum_{i=1}^{p} x_i v_i v_i^T\right) = t$$

$$\Rightarrow \sum_{i=1}^{p} x_i v_i v_i^T \succeq t \cdot I$$

After that, I can derive $L(t, x, Z, \lambda, \nu)$ as

$$L(t, x, Z, \lambda, \nu) = \frac{1}{t} - tr\left(Z \cdot \left(\sum_{i=1}^{p} x_i v_i v_i^T - t \cdot I\right)\right) - \lambda^T x + \nu(1^T x - 1)$$

$$= \frac{1}{t} + t \cdot tr(Z) + \sum_{i=1}^{p} x_i(-v_i^T Z v_i - \lambda_i + \nu) - \nu$$

To make sure x_i is not unbounded below, $(-v_i^T Z v_i - \lambda_i + \nu) = 0$, and we can simplify $L(t, x, Z, \lambda, \nu)$ as

$$L(t, x, Z, \lambda, \nu) = \frac{1}{t} + t \cdot tr(Z) - \nu$$

$$\frac{\partial L(t, x, Z, \lambda, \nu)}{\partial t} = \frac{-1}{t^2} + tr(Z) = 0 \Rightarrow t = \frac{1}{\sqrt{tr(Z)}}$$

$$\Rightarrow g(Z, \lambda, \nu) = \inf_{t,x} L(t, x, Z, \lambda, \nu) = \begin{cases} \sqrt{tr(Z)} + \sqrt{tr(Z)} - \nu = 2\sqrt{tr(Z)} - \nu & , \text{if } -v_i^T Z_i v_i - \lambda_i + \nu = 0 \\ -\infty & , \text{otherwise} \end{cases}$$

$\Rightarrow$ the dual problem becomes

$$\max_{Z,\nu} \ 2\sqrt{tr(Z)} - \nu,$$
$$\text{s.t.} \ v_i^T Z v_i \leq \nu, \ Z \succeq 0$$

$\Rightarrow$ If define $W = \frac{1}{\nu} Z$, I have

$$\max_{W,\nu} \ 2 \cdot \sqrt{\nu}\sqrt{tr(W)} - \nu,$$
$$\text{s.t.} \ v_i^T W v_i \leq 1, \ W \succeq 0$$

$\Rightarrow$ If $\nu = tr(W)$, when optimizing over ν, I have

$$\max_{W} \ tr(W)$$
$$\text{s.t.} \ v_i^T W v_i \leq 1, \ W \succeq 0$$

Assignments

A4.1

(a) After solving the problem with CVX (codes attached as another file),

I know the optimal primal variable values: $x_1^* = -2.3333$, $x_2^* = 0.1667$,

the optimal dual variable values: $\lambda_1^* = 1.4557$, $\lambda_2^* = 3.7735$, $\lambda_3^* = 0.1208$

the optimal objective value: 8.2222

To satisfy KKT conditions, there are four requirements.

(1) Stationarity:

$$\frac{\partial}{\partial x_1}: \; 2x_1^* - x_2^* - 1 + \lambda_1^* + \lambda_2^* + 5\lambda_3^* = 0$$

$$\frac{\partial}{\partial x_2}: \; 4x_2^* - x_1^* + 2\lambda_1^* - 4\lambda_2^* + 76\lambda_3^* = 0$$

(2) Primal Feasibility:

$$x_1^* + 2x_2^* + 2 = 0$$

$$x_1^* - 4x_2^* + 3 = 0$$

$$5x_1^* + 76x_2^* - 1 = 0$$

(3) Dual Feasibility:

$$\lambda_1^* \geq 0$$

$$\lambda_2^* \geq 0$$

$$\lambda_3^* \geq 0$$

(4) Complementary Slackness:

$$\lambda_1^*(x_1^* + 2x_2^* + 2) = \lambda_1^* \cdot 0 = 0$$

$$\lambda_2^*(x_1^* - 4x_2^* + 3) = \lambda_2^* \cdot 0 = 0$$

$$\lambda_3^*(5x_1^* + 76x_2^* - 1) = \lambda_3^* \cdot 0 = 0$$

From (1) to (4), I prove that all KKT requirements are held.

(b) Similar to (a), I implement the codes and attached it as another file. With CVX program, I can filled up the following tables:

δ_1	δ_2	P^*_{pred}	P^*_{exact}
0	0	8.2222	8.2222
0	-0.1	8.5996	8.7064
0	0.1	7.8449	7.9800
-0.1	0	8.3678	8.5650
-0.1	-0.1	8.7451	8.8156
-0.1	0.1	7.9904	8.3189
0.1	0	8.0767	8.2222
0.1	-0.1	8.4540	8.7064
0.1	0.1	7.6993	7.7515

It is known that P_{pred} can be calculated by $P^*(0) - \lambda^T u - \nu^T v$.
Besides, based on the results I obtain, I can show $P^*_{pred} \leq P^*_{exact}$.

A 16.9

(a) The problem can be rewritten as follows.

$$\min_c \; p^T(u+c)$$

$$\begin{aligned}
\text{s.t.}\quad & q_1 = q_T + C_T \\
& q_{i+1} = q_i + c_i, && i = 1, \ldots, T-1 \\
& -D \leq c_i \leq C, && i = 1, \ldots, T \\
& 0 \leq q_i \leq Q, && i = 1, \ldots, T
\end{aligned}$$

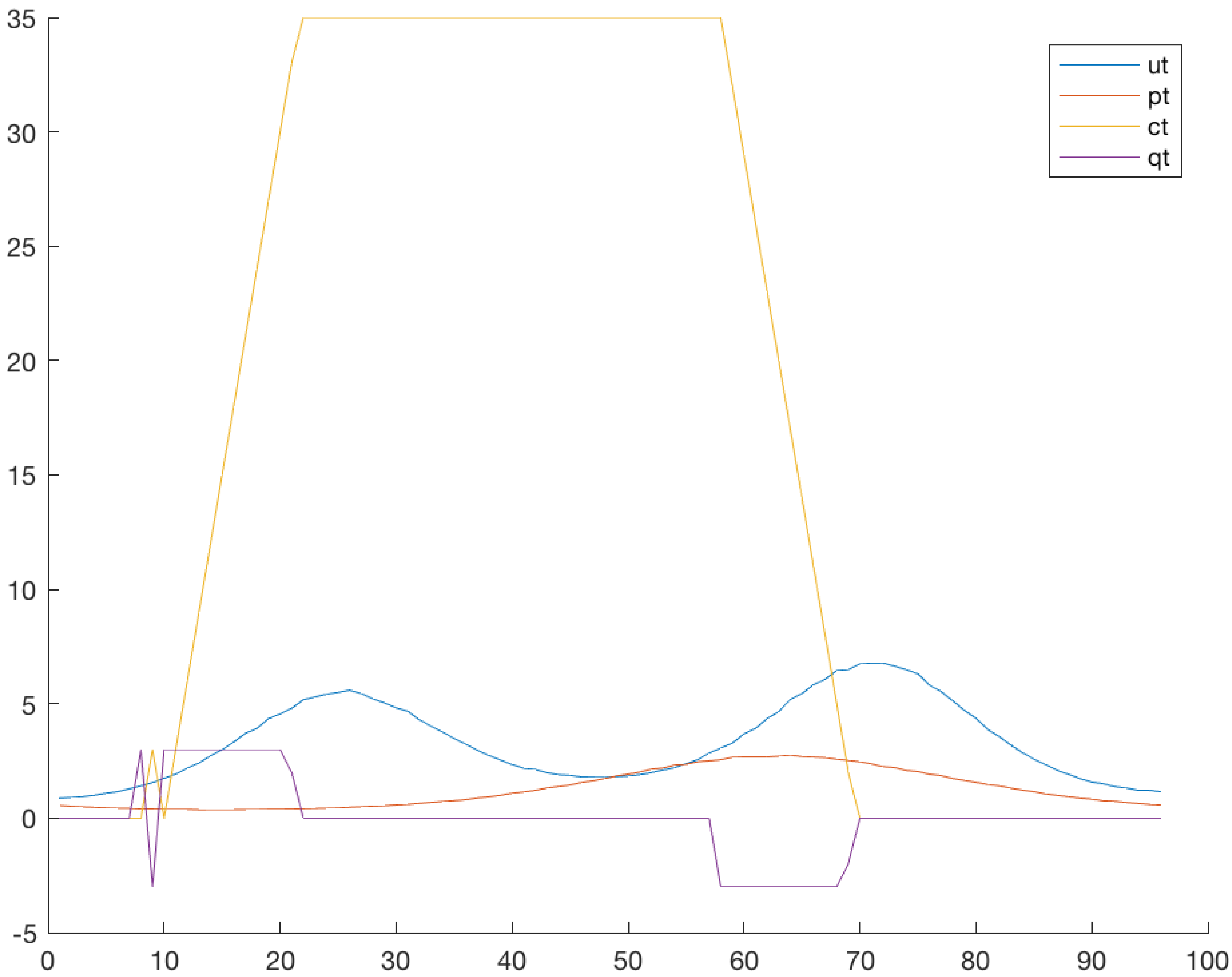
Since the objective function and constraints are all linear, it is a valid convex optimization problem.

(b) The optimal objective value is 378.6023 when $Q=35$, $C=D=3$. The plot can be found in another attached file.

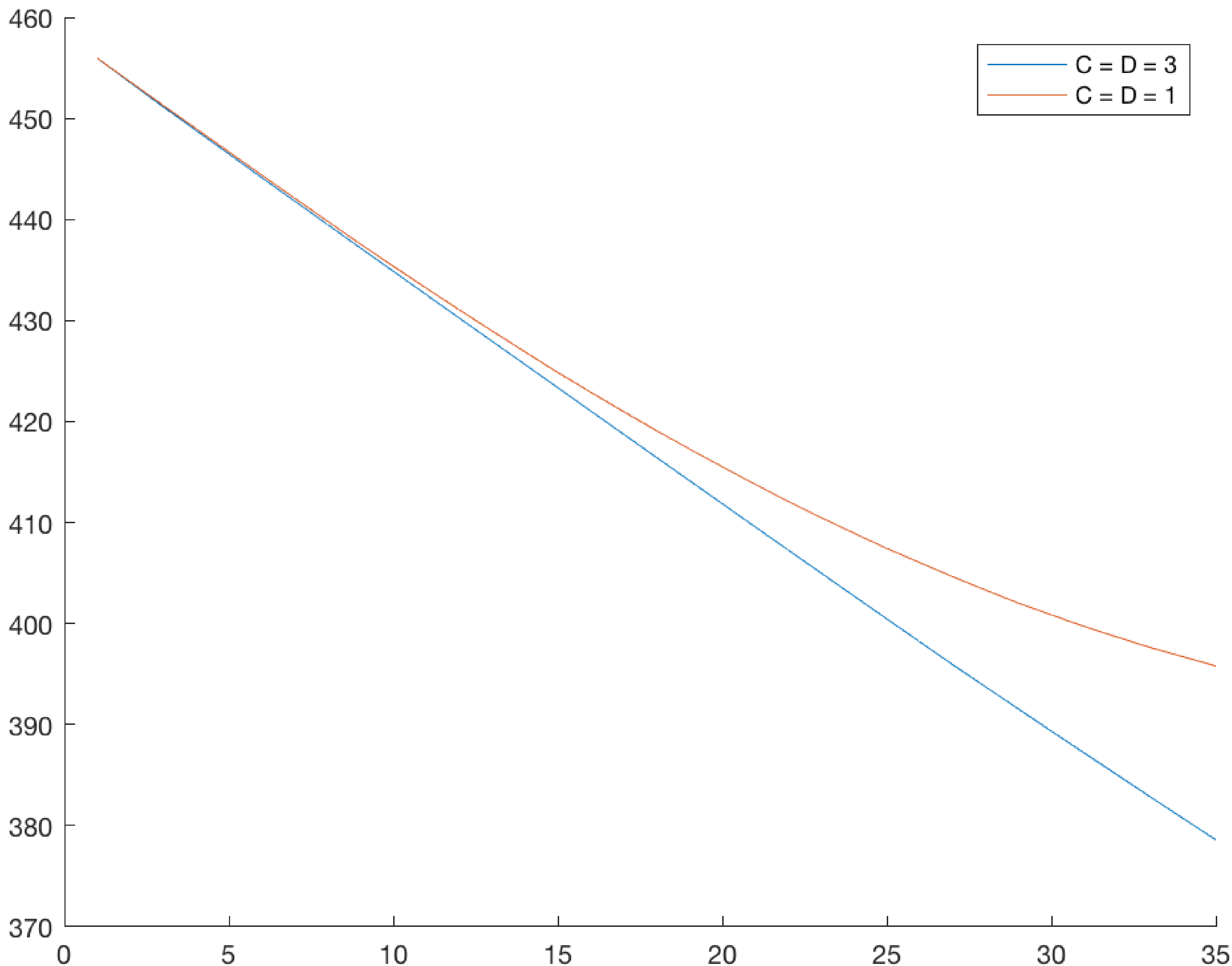
(c) The plot can be found in another attached file.

```matlab
function a1()
% (a)
cvx_begin quiet
    variables x1 x2;
    dual variables y1 y2 y3;
    minimize 0.5 * x1^2 + 1.5 * x2^2 + 0.5 * (x1 - x2)^2 - x1;
    subject to
        y1: x1 + 2 * x2 <= -2;
        y2: x1 - 4 * x2 <= -3;
        y3: 5 * x1 + 76 * x2 <= 1;
cvx_end
opt = cvx_optval;

% (b)
u1 = -2;
u2 = -3;
d = [0, -0.1, 0.1];
for i = 1:3
    for j = 1:3
        cvx_begin quiet
            variables x1 x2;
            minimize 0.5 * x1^2 + 1.5 * x2^2 + 0.5 * (x1 - x2)^2 - x1;
            subject to
                x1 + 2 * x2 <= u1 + d(i);
                x1 - 4 * x2 <= u2 + d(j);
                5 * x1 + 76 * x2 <= 1;
        cvx_end
        pp = opt - y1 * d(i) - y2 * d(j);
        pe = cvx_optval;
        fprintf('d1=%+f d2=%+f, pp=%+f, pe=%+f\n', d(i), d(j), pp, pe);
    end
end
end
```



35
30
25
20
15
10
5
0
-5
0
10
20
30
40
50
60
70
80
90
100
ut
pt
ct
qt


C = D = 3
C = D = 1
460
450
440
430
420
410
400
390
380
370
0
5
10
15
20
25
30
35

```matlab
function a2()
% (b)
randn('seed', 1);
T = 96;
t = (1:T)';
p = exp(-cos((t-15)*2*pi/T)+0.01*randn(T,1));
u = 2*exp(-0.6*cos((t+40)*pi/T) -0.7*cos(t*4*pi/T)+0.01*randn(T,1));

[optv, c, q] = optimize(T, p, u, 3, 3, 35);
disp(optv);

hold on
plot(t, u);
plot(t, p);
plot(t, c);
plot(t, q);
hold off

legend('ut', 'pt', 'ct', 'qt');


% (c)
Q = (1:35)';

cd3_v = zeros(size(Q));
for i = 1:size(Q,1)
    [optv, ~, ~] = optimize(T, p, u, 3, 3, Q(i));
    cd3_v(i) = optv;
end

cd1_v = zeros(size(Q));
for i = 1:size(Q,1)
    [optv, ~, ~] = optimize(T, p, u, 1, 1, Q(i));
    cd1_v(i) = optv;
end

hold on
plot(Q, cd3_v);
plot(Q, cd1_v);
hold off

legend('C = D = 3', 'C = D = 1')
end

function [cvx_optval, q, c] = optimize(T, p, u, C, D, Q)
    cvx_begin quiet
        variable q(T);
        variable c(T);
        minimize (p' * (u + c));

        q(1) == q(T) + c(T);
        for i = 1:(T-1)
            q(i+1) == q(i) + c(i);
        end

        for i = 1:T
            c(i) <= C;
            c(i) >= -D;
            q(i) >= 0;
            q(i) <= Q;
        end
    cvx_end
end
```