Homework1

EE688. Optimal Control Theory

Hany Hamed

September 25, 2022

Contents

1 Solution of Problem 1: Linearization and Discretization

1.1 (a)

Restating the task: we want to show that $\forall\ \bar{x}_1 = \bar{h} > 0$ and $\forall\ \bar{x}_2 = \bar{T}_T$ such that $T_C \leq \bar{T}_T \leq T_H$ is a possible equilibrium point.

That means that we want to show that: $\exists \bar{u} \in \mathbb{R}^2$ satisfies "$\forall\ \bar{x}_1 = \bar{h} > 0$ and $\forall\ \bar{x}_2 = \bar{T}_T$ such that $T_C \leq \bar{T}_T \leq T_H$" such that $\dot{h}(t) = 0$ and $\dot{T}_T(t) = 0$

$$\begin{aligned}\dot{h}(t) &= \frac{1}{A_T}(q_C(t) + q_H(t) - c_D A_\circ \sqrt{2gh(t)} \\ &= \frac{1}{A_T}(\bar{u}_1 + \bar{u}_2 - c_D A_\circ \sqrt{2g\bar{x}_1}) = 0\end{aligned}$$

$$\begin{aligned}\frac{1}{A_T}(\bar{u}_1 + \bar{u}_2 - c_D A_\circ \sqrt{2g\bar{x}_1}) &= 0 \\ \bar{u}_1 + \bar{u}_2 - c_D A_\circ \sqrt{2g\bar{x}_1} &= 0 \\ c_D A_\circ \sqrt{2g\bar{x}_1} - \bar{u}_2 &= \bar{u}_1\end{aligned} \tag{1}$$

$$\begin{aligned}\dot{T}_T(t) &= \frac{1}{h(t)A_T}(q_C(t)[T_C - T_T(t)] + q_H(t)[T_H - T_T(t)] \\ &= \frac{1}{\bar{x}_1 A_T}(\bar{u}_1[T_C - \bar{x}_2] + \bar{u}_2[T_H - \bar{x}_2]) = 0\end{aligned}$$

$$\begin{aligned}\frac{1}{h(t)A_T}(q_C(t)[T_C - T_T(t)] + q_H(t)[T_H - T_T(t)]) &= 0 \\ \bar{u}_1[T_C - \bar{x}_2] + \bar{u}_2[T_H - \bar{x}_2] &= 0 \\ -\frac{\bar{u}_2[T_H - \bar{x}_2]}{[T_C - \bar{x}_2]} &= \bar{u}_1\end{aligned} \tag{2}$$

From Eq. 1 and Eq. 2:

$$\begin{aligned}-\bar{u}_2\frac{[T_H - \bar{x}_2]}{[T_C - \bar{x}_2]} &= c_D A_\circ \sqrt{2g\bar{x}_1} - \bar{u}_2 \\ \bar{u}_2(1 - \frac{T_H - \bar{x}_2}{T_C - \bar{x}_2}) &= c_D A_\circ \sqrt{2g\bar{x}_1} \\ \bar{u}_2 &= \frac{c_D A_\circ \sqrt{2g\bar{x}_1}}{(1 - \frac{T_H - \bar{x}_2}{T_C - \bar{x}_2})} \\ \bar{u}_1 &= c_D A_\circ \sqrt{2g\bar{x}_1} - \frac{c_D A_\circ \sqrt{2g\bar{x}_1}}{(1 - \frac{T_H - \bar{x}_2}{T_C - \bar{x}_2})}\end{aligned} \tag{3}$$

Therefore, $\forall\ \bar{x}_1 = \bar{h} > 0$ and $\forall\ \bar{x}_2 = \bar{T}_T$ such that $T_C \leq \bar{T}_T \leq T_H$ such that $\dot{h}(t) = 0$ and $\dot{T}_T(t) = 0$, $\exists \bar{u} \in \mathbb{R}^2$

1.2 (b)

$$f(x,u,t)=\dot{x}(t)=\begin{bmatrix}\dot{h}(t)\\ \dot{T}_T(t)\end{bmatrix}=\begin{bmatrix}\frac{1}{A_T}(q_C(t)+q_H(t)-c_DA_\circ\sqrt{2gh(t)})\\ \frac{1}{h(t)A_T}(q_C(t)[T_C-T_T(t)]+q_H(t)[T_H-T_T(t)]\end{bmatrix}=\begin{bmatrix}f_1(x,u,t)\\ f_2(x,u,t)\end{bmatrix}$$

$$x(t)=\begin{bmatrix}h(t)\\ T_T(t)\end{bmatrix}=\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

$$u(t)=\begin{bmatrix}q_C(t)\\ q_H(t)\end{bmatrix}=\begin{bmatrix}u_1(t)\\ u_2(t)\end{bmatrix}$$

$$\tilde{A}=\frac{\partial f}{\partial x}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}\end{bmatrix}$$

$$\tilde{B}=\frac{\partial f}{\partial u}=\begin{bmatrix}\frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u}\\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2}\end{bmatrix}$$

The linearized model is in the following form.

$$\begin{aligned}\dot{\tilde{x}}&=\tilde{A}(t)\tilde{x}(t)+\tilde{B}(t)\tilde{u}(t)\\ \dot{\tilde{x}}&=\tilde{A}(t)(x(t)-\bar{x})+\tilde{B}(t)(u(t)-\bar{u})\end{aligned}$$

such that $\bar{x},\bar{u}$ are the equilibrium point with the corresponding equilibrium input.

$$\frac{\partial f}{\partial x}=\begin{bmatrix}-\frac{1}{A_T}c_DA_\circ\frac{g}{\sqrt{2gh(t)}} & 0\\ -\frac{1}{h(t)^2A_T}(q_C(t)[T_C-T_T(t)]+q_H(t)[T_H-T_T(t)]) & -\frac{1}{h(t)A_T}(q_H(t)+q_C(t))\end{bmatrix} \tag{4}$$

$$\frac{\partial f}{\partial u}=\begin{bmatrix}\frac{1}{A_T} & \frac{1}{A_T}\\ \frac{1}{h(t)A_T}(T_C(t)-T_T(t)) & \frac{1}{h(t)A_T}(T_H(t)-T_T(t))\end{bmatrix} \tag{5}$$

Evaluating the Jacobians at equilibrium, we get the following.

$$\bar{x}=\begin{bmatrix}\tilde{x}_1\\ \tilde{x}_2\end{bmatrix},\ \bar{u}=\begin{bmatrix}\tilde{u}_1\\ \tilde{u}_2\end{bmatrix}$$

$$\tilde{A}(t)=\frac{\partial f}{\partial x}=\begin{bmatrix}-\frac{1}{A_T}c_DA_\circ\frac{g}{\sqrt{2g\bar{x}_1}} & 0\\ -\frac{1}{\bar{x}_1^2A_T}(\bar{u}_1[T_C-\bar{x}_2]+\bar{u}_2[T_H-\bar{x}_2]) & -\frac{1}{\bar{x}_1A_T}(\bar{u}_2+\bar{u}_1)\end{bmatrix} \tag{6}$$

$$\tilde{B}(t) = \frac{\partial f}{\partial u} = \begin{bmatrix} \frac{1}{A_T} & \frac{1}{A_T} \\ \frac{1}{\bar{x}_1 A_T}(T_C(t) - \bar{x}_2) & \frac{1}{\bar{x}_1 A_T}(T_H(t) - \bar{x}_2) \end{bmatrix} \tag{7}$$

Thus, the full linearized model is

$$\begin{aligned}
\dot{\tilde{x}} &= A(t)\tilde{x}(t) + B(t)\tilde{u}(t) \\
\dot{\tilde{x}} &= A(t)(x(t) - \bar{x}) + B(t)(u(t) - \bar{u}) \\
\begin{bmatrix} \dot{h}(t) \\ \dot{T}_T(t) \end{bmatrix} &= \begin{bmatrix} -\frac{1}{A_T} c_D A_\circ \frac{g}{\sqrt{2g\bar{x}_1}} & 0 \\ -\frac{1}{\bar{x}_1^2 A_T}(\bar{u}_1[T_C - \bar{x}_2] + \bar{u}_2[T_H - \bar{x}_2]) & -\frac{1}{\bar{x}_1 A_T}(\bar{u}_2 + \bar{u}_1) \end{bmatrix}(x(t) - \bar{x}) \\
&+ \begin{bmatrix} \frac{1}{A_T} & \frac{1}{A_T} \\ \frac{1}{\bar{x}_1 A_T}(T_C(t) - \bar{x}_2) & \frac{1}{\bar{x}_1 A_T}(T_H(t) - \bar{x}_2) \end{bmatrix}(u(t) - \bar{u})
\end{aligned} \tag{8}$$

1.3 (c)

We can find the linearized model such that $T_C = 10, T_H = 90, A_T = 3, A_\circ = 0.05, c_D = 0.7, g = 10, (\bar{h} = \bar{x}_1, \bar{T}_T = \bar{x}_2) = (1, 25)$

And from Eq. 3, we an find $\bar{q}_C = \bar{u}_1 = 0.1272$ and $\bar{q}_H = \bar{u}_2 = 0.0293$

$$\bar{x} = \begin{bmatrix} 1 \\ 25 \end{bmatrix}, \ \bar{u} = \begin{bmatrix} 0.1272 \\ 0.0293 \end{bmatrix}$$

By substituting into Eq. 8, we can obtain the linearized model with specific parameters and a specific equilibrium point as follow.

$$\begin{aligned}
\dot{\tilde{x}} &= A(t)\tilde{x}(t) + B(t)\tilde{u}(t) \\
\dot{\tilde{x}} &= A(t)(x(t) - \bar{x}) + B(t)(u(t) - \bar{u}) \\
\begin{bmatrix} \dot{h}(t) \\ \dot{T}_T(t) \end{bmatrix} &= \begin{bmatrix} -0.026 & 0 \\ 0 & -0.0522 \end{bmatrix}(x(t) - \bar{x}) \\
&+ \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -5 & \frac{65}{3} \end{bmatrix}(u(t) - \bar{u})
\end{aligned} \tag{9}$$

1.4 (d)

[1] Using the Zero-order hold method (ZOH), we can obtain the discretized model of Eq. 9 as follow:

$$x_{k+1} = A_d x_k + B_d u_k$$

such that $A_d = e^{\tilde{A}T_S}$, and $B_D = \tilde{A}^{-1}(e^{\tilde{A}T_S} - I)\tilde{B}$, I is the identity matrix and T_S is the sampling time = 0.05s

[1] Numeric calculation was done using Symbolab

$$A_d = \begin{bmatrix} 0.9987 & 0 \\ 0 & 0.99739 \end{bmatrix} \tag{10}$$

$$B_d = \begin{bmatrix} 0.01667 & 0.01667 \\ -0.25 & 1.08338 \end{bmatrix} \tag{11}$$

2 Solution of Problem 2: Stability

2.1 (a)

Proof:

We have the following to hold to be used in the proof.

- P is symmetric and positive definite ($P = P^T \succ 0$)
- $A^T PA - P = -Q$
- Q is symmetric and positive definite ($Q = Q^T \succ 0$)

Then $A^T PA - P = -Q$ is negative definite ($A^T PA - P = -Q \prec 0$)

In order to prove that the linear system $x_{k+1} = Ax_k$ is asymptotically stable, we need to prove that $\exists V : \mathbb{R}^n \to \mathbb{R}$ (a Lyapunov function) satisfies certain conditions, we will choose the Lyapunov function in the form of quadratic function $V = x^T Px$.

1. $V(0) = 0$
   As the function is quadratic, it holds
2. $V(x) > 0 \quad \forall\, x \in \mathbb{R}^n \backslash \{0\}$
   As the function is quadratic and $P \succ 0$, it holds
3. $V(x_{k+1}) - V(x_k) < 0$

$$\begin{aligned}
V(x_{k+1}) - V(x_k) &= x_{k+1}^T Px_{k+1} - x_k^T Px_k \\
&= (Ax_k)^T P(Ax_k) - x_k^T Px_k \\
&= x_k^T A^T PAx_k - x_k^T Px_k \\
&= x_k^T (A^T PA - P)x_k \\
&= -x_k^T Qx_k \prec 0 \quad (\text{As } Q = Q^T \succ 0)
\end{aligned}$$

4. $||x|| \to \infty, V(x) \to infty$: As $V(x) = x^T Px$ is quadratic and goes to infinity when x goes to infinity (Radially unbounded)

Then, the conditions are satisfied; then the system is asymptotically stable.

□

2.2 (b)

Solution 1: Reference: link1link2 and link3

First, let us denote the following operators:

- I is the identity matrix

- $vec(A)$ is the vector operator that stacks all the columns of matrix A with size (n× n) on top of each other to result a vector of size (n^2× 1)

$$A = \begin{bmatrix} \vdots & \vdots & \dots & \vdots \\ Col_1 & Col_2 & \cdots & Col_n \\ \vdots & \vdots & \dots & \vdots \end{bmatrix}$$

$$vec(A) = \begin{bmatrix} \vdots \\ Col_1 \\ \vdots \\ \vdots \\ Col_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ Col_n \\ \vdots \end{bmatrix}$$

- $\otimes$ is the Kronecker product between two matrices as following:

$$A \otimes B = \begin{bmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,m_A}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,m_A}B \\ \vdots & \vdots & \dots & \vdots \\ a_{n_A,1}B & a_{n_A,2}B & \cdots & a_{n_A,m_A}B \end{bmatrix}$$

  such that A with size ($n_A \times m_A$) and B with size ($n_B \times m_B$) and the result is a matrix with size ($n_a n_b \times m_a m_b$)

- Moreover, the following property holds according to the reference[2]:

$$vec(AXB) = (B^T \otimes A)vec(X)$$

  such that A, B, X are matrices.

- Furthermore, $eigenvalue(A \otimes B) = eigenvalue(A) \cdot eigenvalue(B)$ for all eigenvalues of the Kronecker product[3]

Therefore, we can rewrite Discrete-time Lyapunov equation (3) as following:

[2]Reference: Theorem 1 in link
[3]Reference: Theorem 2.3 in link

$$
\begin{aligned}
A^T PA - P &= -Q \\
vec(A^T PA) - vec(P) &= -vec(Q) \\
-vec(A^T PA) + vec(P) &= vec(Q) \\
vec(P) - (A^T \otimes A^T)vec(P) &= vec(Q) \\
[\mathbb{I} - A^T \otimes A^T]vec(P) &= vec(Q) \\
Cz &= b
\end{aligned}
$$

Therefore, it is similar to solving system of linear equations ($Cz = b$) and its solution is unique iff C is invertible. $C = \mathbb{I} - A^T \otimes A^T$ is invertible when the dimension of the null space of $C : (Dim(\text{null}(\mathbf{C}))) = 0$, which means that C matrix has no zero eigenvalues.

C has n^2 eigenvalues and to be all of them non-zeros the following should hold:

$$
\begin{aligned}
1 - \lambda_i \lambda_j &\neq 0 \ \forall i, j \in 1, ..., n \\
\lambda_i \lambda_j &\neq 1 \ \forall i, j \in 1, ..., n
\end{aligned}
\tag{12}
$$

such that $\lambda_k \ \forall k \in 1, ..., n$ are the eigenvalues for A^T

This holds iff the system (2) is the discrete-time linear system asymptotically stable, then the absolute value of eigenvalues of A is less than 1 ($|\lambda_i| < 1$) [4]. Thus, the absolute value eigenvalues of $A^T \otimes A^T$ are not equal to 1, and $[\mathbb{I} - A^T \otimes A^T]vec(P)$ is invertible and there is a unique solution for $vec(P)$, hence a unique solution for P

□

2.3 (c)

The solution is implied in 2.b, since the system is assumed to be asymptotically stable; therefore, we were able to find the unique existence of the matrix P. However, we did not prove that P is a symmetric and positive definite matrix.

To prove that there exists a symmetric and postive definte matrix P satisfies $A^T PA - P = -Q$. using the hint in (b), $P = \Sigma_{t=0}^{\infty}(A^T)^t QA^t$, then we can see that the system is stable, the absolute eigenvalues of A are less than 1, thus the summation converges while the exponent goes to infinity and the relation is quadratic. Since, Q is positive definite, P will be positive definite as well.

□

[4] Reference: ref1: Theorem 1, ref2, ref3: Page 21

2.4 (d)

From the Lyapunov stability theorem, given $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ we can find the solution of the DT-lyapunov equation (Matrix P) as following:

Using Matlab code, we can find the solution of P matrix

```
Q = eye(2)
A = [0, 1; -0.5, -0.5]
C = eye(4) - kron(A.', A.')
vec_Q = reshape(Q,1,[]).'
vec_P = inv(C)*vec_Q
P = reshape(vec_P, 2,2)

% or directly through P = dylap(A,Q)
```

$P = \begin{bmatrix} 1.75 & 0.5 \\ 0.5 & 3 \end{bmatrix}$

Therefore, according to the Lyapunov stability theorem and (a,b,c) proofs, we found P and proved that the system is asymptotically stable

$\square$

3 Solution of Problem 3: Convexity Preserving Operations

3.1 (a)

Show that the sub-level set $S = \{z \in \mathcal{R}^n : J_1(z) \leq \alpha\}$ is a convex set, such that $J_1 : \mathcal{R}^n \to \mathcal{R}$ is a covex function

Proof:

Let us pick two points $x_1, x_2 \in S$.

We need to show

$$(\lambda x_1 + (1-\lambda)x_2) \overset{?}{\in} S \;\forall \lambda \in [0,1]$$

Which means, we need to show

$$J_1(\lambda x_1 + (1-\lambda)x_2) \overset{?}{\leq} \alpha \;\forall \lambda \in [0,1]$$

$$\begin{aligned}
J_1(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda J_1(x_1) + (1-\lambda)J_1(x_2) && \text{(Convexity of } J_1\text{)} \\
&\leq \lambda\alpha + (1-\lambda)\alpha && \text{(The set definition)} \\
&= \alpha
\end{aligned}$$

Thus,

$J_1(\lambda x_1 + (1-\lambda)x_2) \leq \alpha \implies (\lambda x_1 + (1-\lambda)x_2) \in S \;\forall \lambda \in [0,1]$

$\implies$ The sub-level set S is a convex set

□

3.2 (b)

Show that $J_3 = J_1 + J_2$ is a convex function

Proof:

Let us pick two points $x_1, x_2 \in \mathcal{R}^n$

We need to show:

$$J_3(\lambda x_1 + (1-\lambda)x_2) \overset{?}{\leq} \lambda J_3(x_1) + (1-\lambda)J_3(x_2) \;\forall \lambda \in [0,1]$$

$$\begin{aligned}
J_3(\lambda x_1 + (1-\lambda)x_2) &= J_1(\lambda x_1 + (1-\lambda)x_2) + J_2(\lambda x_1 + (1-\lambda)x_2) && \text{(Definition of } J_3\text{)} \\
J_1(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda J_1(x_1) + (1-\lambda)J_1(x_2) && \text{(Convexity of } J_1\text{)} \\
J_2(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda J_2(x_1) + (1-\lambda)J_2(x_2) && \text{(Convexity of } J_2\text{)} \\
J_3(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda J_1(x_1) + (1-\lambda)J_1(x_2) + \lambda J_2(x_1) + (1-\lambda)J_2(x_2) \\
&= \lambda J_3(x_1) + (1-\lambda)J_3(x_2) && \text{(Common factors)}
\end{aligned}$$

Therefore,

$$J_3(\lambda x_1 + (1-\lambda)x_2) \leq \lambda J_3(x_1) + (1-\lambda)J_3(x_2)$$

and $J_3 = J_1 + J_2$ is a convex function

$\square$

3.3 (c)

Show that $J = J_1(Az + b)$ is a convex function

Proof:

Let us pick two points $x_1, x_2 \in \mathcal{R}^n$

We need to show:

$$J(\lambda x_1 + (1-\lambda)x_2) \overset{?}{\leq} \lambda J(x_1) + (1-\lambda)J(x_2)\ \forall \lambda \in [0,1]$$

As $f(z) = Az + b$ is an affine map, it means that it is a convex function. ("They are convex, but not strictly convex; they are also concave"[5]). Therefore, $f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda)f(x_2)$ holds.

$$\begin{aligned}
J(\lambda x_1 + (1-\lambda)x_2) &= J_1(A(\lambda x_1 + (1-\lambda)x_2) + b) && \text{(Definition of } J\text{)} \\
&= J_1(f(\lambda x_1 + (1-\lambda)x_2) \\
&= J_1(\lambda f(x_1) + (1-\lambda)f(x_2)) \\
&\leq \lambda J_1(f(x_1)) + (1-\lambda)J_1(f(x_2)) && \text{(Convexity of } J_1\text{)} \\
&= \lambda J_3(x_1) + (1-\lambda)J_3(x_2)
\end{aligned}$$

Therefore,

$$J(\lambda x_1 + (1-\lambda)x_2) \leq \lambda J(x_1) + (1-\lambda)J(x_2)\ \forall \lambda \in [0,1]$$

and $J = J_1(Az + b)$ is a convex function

$\square$

3.4 (d)

Show that $J(z) = \max(J_1(z), J_2(z))$ is a convex function

Proof

Let us first state the following two properties of the maximum operator such that $a, b, c, d \in \mathbb{R}$.

1. $\max(ca, cb) = c \cdot \max(a, b)$ (Scaling does not change in the comparison between the items)

[5] Reference: Section 1.4

2. $\max(a+b, c+d) \leq \max(a,c) + \max(b,d)$

   Proof[6]:

   Note:

   1)

   $\begin{cases} a \leq \max(a,c) \\ b \leq \max(b,d) \end{cases}$

   Therefore, $a+b \leq \max(a,c) + \max(b,d)$

   2)

   $\begin{cases} c \leq \max(a,c) \\ d \leq \max(b,d) \end{cases}$

   Therefore, $c+d \leq \max(a,c) + \max(b,d)$

   1), 2) have the same upper bounds. Thus,

$$\max(a+b, c+d) \leq \max(a,c) + \max(b,d)$$

□

Then, let us proceed to the main proof:
Let us pick two points $x_1, x_2 \in \mathcal{R}^n$
We need to show:

$$J(\lambda x_1 + (1-\lambda)x_2) \overset{?}{\leq} \lambda J(x_1) + (1-\lambda)J(x_2) \ \forall \lambda \in [0,1]$$

$$\max(J_1(\lambda x_1 + (1-\lambda)x_2), J_2(\lambda x_1 + (1-\lambda)x_2)) \overset{?}{\leq} \lambda \max(J_1(x_1), J_2(x_1)) + (1-\lambda)\max(J_1(x_2), J_2(x_2)) \ \forall \lambda \in [0,1]$$

$$\begin{aligned}
J(\lambda x_1 + (1-\lambda)x_2) &= \max(J_1(\lambda x_1 + (1-\lambda)x_2), J_2(\lambda x_1 + (1-\lambda)x_2) && \text{(Definition)} \\
&= \max(\lambda J_1(x_1) + (1-\lambda)J_1(x_2), \lambda J_2(x_1) + (1-\lambda)J_2(x_2)) && \text{(Convexity of } J_1, J_2\text{)} \\
&\leq \max(\lambda J_1(x_1), \lambda J_2(x_1)) + \max((1-\lambda)J_1(x_2), (1-\lambda)J_2(x_2)) && \text{(Property (2))} \\
&= \lambda \max(J_1(x_1), J_2(x_1)) + (1-\lambda)\max(J_1(x_2), J_2(x_2)) && \text{(Property (1))} \\
&= \lambda J(x_1) + (1-\lambda)J(x_2)
\end{aligned}$$

Therefore,

$$J(\lambda x_1 + (1-\lambda)x_2) \leq \lambda J(x_1) + (1-\lambda)J(x_2) \ \forall \lambda \in [0,1]$$

and $J = \max(J_1, J_2)$ is a convex function

□

[6] Reference: link

4 Solution of Problem 4: Lagrange Duality

Note: $x = [x_1 \ x_2]^T$, $u = [u_1 \ u_2 \ ... \ u_n]^T$

Primal problem:

$$\begin{aligned} J^* = \min_z \quad & J(z) \\ \text{subject to} \quad & g_i(z) \leq 0 \\ & h_i(z) = 0 \end{aligned}$$

Lagrangian:

$\mathcal{L}(z, u, v) = J(z) + u^T \cdot g(z) + v^T \cdot h(z)$, such that $u^T \geq 0$

Dual function:

$$d(u, v) = \min_z \mathcal{L}(z, u, v)$$

Dual problem:

$$\begin{aligned} d^* = \max_{u,v} \quad & d(u, v) \\ \text{subject to} \quad & u \geq 0 \\ & h_i(z) = 0 \end{aligned}$$

equivalent to:

$$\begin{aligned} d^* = \min_{u,v} \quad & -d(u, v) \\ \text{subject to} \quad & u \geq 0 \\ & h_i(z) = 0 \end{aligned}$$

4.1 P1

$$\begin{aligned} \min_x \quad & 0.5(x_1^2 + x_2^2) \\ \text{subject to} \quad & 1 - x_1 \leq 0 \end{aligned}$$

4.1.1 (a)

Yes, convex optimization problem.

Reasons:

- $J(x) = 0.5(x_1^2 + x_1^2)$ is a quadratic function (convex function).
- $g(x) = 1 - x_1$ is a hyperplane (convex function).

4.1.2 (b)

Lagrangian:

$$\mathcal{L}(x, u) = 0.5(x_1^2 + x_2^2) + u(1 - x_1)$$

Dual function:

$$d(u) = \min_x 0.5(x_1^2 + x_2^2) + u(1 - x_1)$$

Taking the gradient with respect to x of $\nabla_x \mathcal{L}(x, u) = 0$
$\begin{bmatrix} \nabla_{x_1} \mathcal{L}(x, u) \\ \nabla_{x_2} \mathcal{L}(x, u) \end{bmatrix} = \begin{bmatrix} x_1 - u \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
Then, $x_2 = 0, x_1 = u$ and $d(u) = -0.5u^2 + u$
Dual problem:

$$\begin{aligned} d^* = \max_u \quad & -0.5u^2 + u \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

equivalent to:

$$\begin{aligned} d^* = \min_u \quad & +0.5u^2 - u \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

4.1.3 (c)

Yes, Dual problem is convex optimization problem. It is a concave that can be written in convex form
Reasons:

- $0.5u^2 - u$ is a quadratic function (convex function).
- $g(u) = u$ is a hyperplane (convex function).

4.1.4 (d)

Taking the gradient with respect to u of $\nabla_u(0.5u^2 - u) = 0$, Then, $u = 1$, $d^* = d(u = 1) = 0.5$

4.1.5 (e)

The duality gap $J^* - d^* = 0$

4.1.6 (f)

- Primal and dual problems are convex (a,c)
- $\exists z^\circ$ such that $z^\circ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $g(z^\circ) \leq 0$. Thus, Slater's condition is satisfied (Weaker version)
- Duality gap = 0 (e), strong duality

4.2 P2

$$\begin{aligned} \min_x \quad & x_1 - x_2 \\ \text{subject to} \quad & x_1 + x_2 - 1 \leq 0 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{aligned}$$

4.2.1 (a)

Yes, convex optimization problem.

Reasons:

- $J(x) = x_1 - x_2$ is a hyperplane $a^T z + b$ (onvex function).
- $g(x)_1 = x_1 + x_2 - 1$ is a hyperplane (convex function).
- $g(x)_2 = -x_1$ is a hyperplane (convex function).
- $g(x)_3 = -x_2$ is a hyperplane (convex function).

4.2.2 (b)

Lagrangian:

$$\mathcal{L}(x,u) = x_1 - x_2 + u_1(x_1 + x_2 - 1) - u_2(x_1) - u_3(x_2)$$

Dual function:

$$\begin{aligned} d(u) = \min_x \quad & x_1 - x_2 + u_1(x_1 + x_2 - 1) - u_2(x_1) - u_3(x_2) \\ \text{subject to} \quad & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$

Taking the gradient with respect to x of $\nabla_x \mathcal{L}(x,u) = 0$
$\begin{bmatrix} \nabla_{x_1}\mathcal{L}(x,u) \\ \nabla_{x_2}\mathcal{L}(x,u) \end{bmatrix} = \begin{bmatrix} 1 - u_1 - u_2 \\ 1 - u_1 - u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
Then, $u_1 = 1 - u_2$, $u_1 = 1 - u_3$, then, $u_3 = u_2$.
Dual problem:

$$\begin{aligned} d^* = \max_u \quad & x_1 - x_2 + (1 - u_2)(x_1 + x_2 - 1) - u_2(x_1) - u_2(x_2) \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

$$\begin{aligned} d^* = \max_u \quad & (2 - 2u_2)x_1 - 2u_2 x_2 - 1 \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

equivalent to:

$$\begin{aligned} d^* = \min_{u} \quad & (2u_2 - 2)x_1 + 2u_2x_2 + 1 \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

4.2.3 (c)

Yes, Dual problem is convex optimization problem. It is a concave that can be written in convex form

Reasons:

- $(2u_2 - 2)x_1 + 2u_2x_2 + 1$ is in the form of $Au + b$ (convex function).
- $g(u) = u$ is a hyperplane (convex function).

4.2.4 (d)

Taking the gradient with respect to u of $\nabla_u((2u_2 - 2)x_1 + 2u_2x_2 + 1) = 0$, Then, $x_1 = -x_2$, and as $x_1 \geq 0, x_2 \geq 0$, then $x_1 = x_2 = 0$, $d^* = -1$

4.2.5 (e)

The duality gap $J^* - d^* = -1 + 1 = 0$

4.2.6 (f)

- Primal and dual problems are convex (a,c)
- $\exists z^\circ$ such that $z^\circ = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $g(z^\circ) \leq 0$. Thus, Slater's condition is satisfied (Weaker version)
- Duality gap = 0 (e), strong duality

4.3 P3

4.3.1 (a)

Not, it is not convex optimization problem.

Reasons:

- $g(x)_2 : x \in \{0,1\}$ is not a convex set.

  Proof:

  $\lambda \cdot 0 + (1-\lambda) \cdot 1$ not in $\{0,1\}$ for $0 < \lambda < 1$

4.3.2 (b)

Lagrangian:

$$\mathcal{L}(x,u) = -x + u(x - 0.5)$$

Dual function:

$$\begin{aligned} d(u) = \min_{x} & \quad -x + u(x-0.5) \\ \text{subject to} & \quad x \in \{0,1\} \end{aligned}$$

Taking the gradient with respect to x of $\nabla_x \mathcal{L}(x,u) == 0$, $-1 + u = 0$, then $u = 1$

Dual problem:

$$\begin{aligned} d^* = \max_{u} & \quad -x + (x-0.5) \\ \text{subject to} & \quad u \geq 0 \end{aligned}$$

4.3.3 (c)

The Dual problem is convex optimization problem. It is a concave that can be written in convex form

Reasons:

- $-x + (x-0.5) = -0.5$ is just a line (convex and concave).
- $g(u) = u$ is a hyperplane (convex function).

4.3.4 (d)

By the simplification, we can see that $d^* = \max_u -0.5 = 0.5$

4.3.5 (e)

The duality gap $J^* - d^* = 0 + 0.5 = 0.5$

4.3.6 (f)

- Primal is not convex
- Duality gap > 0
- No strong duality