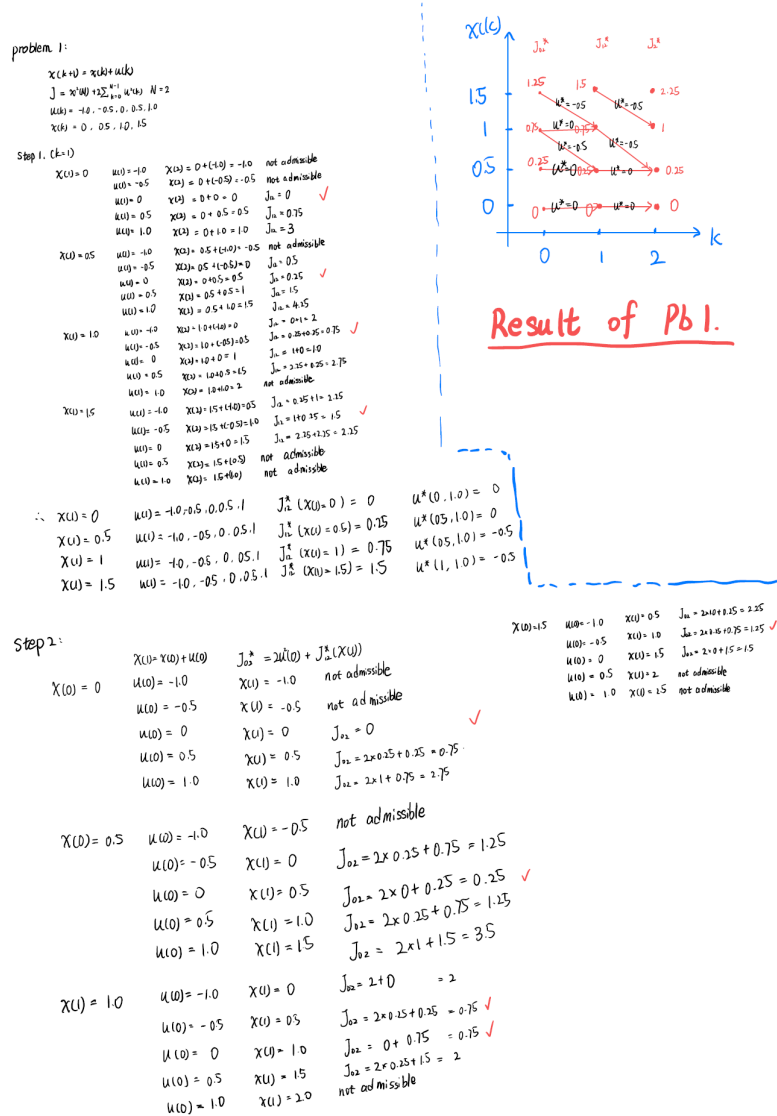


Assignment 2

Shuhao Bian

April 8, 2023

1 Problem 1



According to the Figure 1, the best input is:

- $x^*(0) = 1 \rightarrow x^*(1) = 1 \rightarrow x^*(2) = 0.5, (u^*(1) = 0, u^*(2) = -0.5)$

- $x^*(0) = 1 \rightarrow x^*(1) = 0.5 \rightarrow x^*(2) = 0.5, (u^*(1) = -0.5, u^*(2) = 0)$

2 Problem 2

2.1 No constraints input, $T_s = 0.1s$

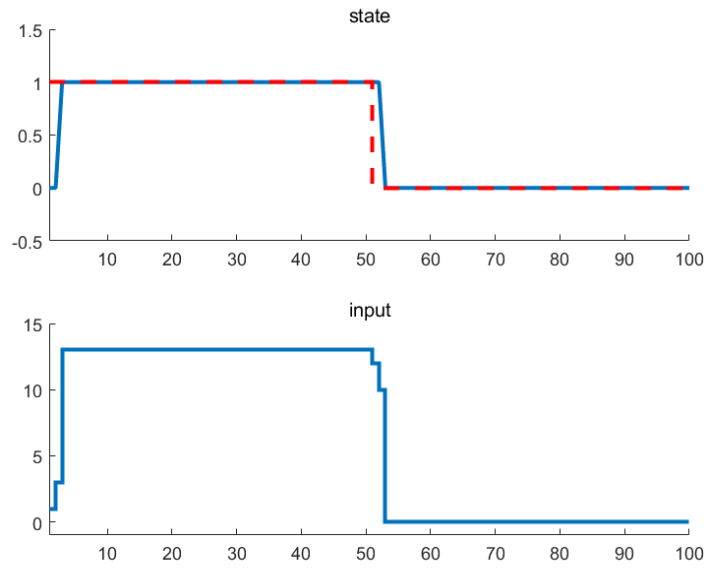


Figure 2: Graph q1

In the figure 2, the output of the system followed the reference well with no overshoot and no undershoot. The maximum input is bigger than 10, Δu is unlimited.

2.2 $|u(t)| \leq 10, T_s = 0.1s$

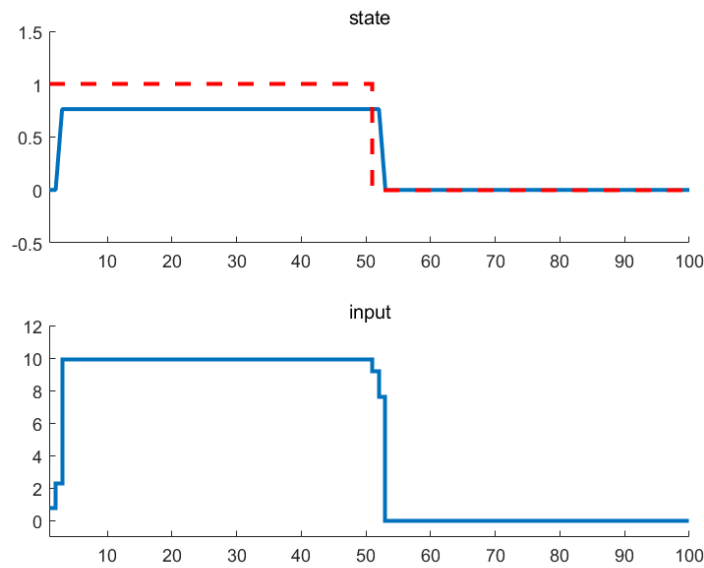


Figure 3: Graph q2

In the figure 3, the input signal is limited to 10, the output of the system cannot follow the reference well.

2.3 $|u(t)| \leq 10, |\Delta u(t)| \leq 1, T_s = 0.1s$

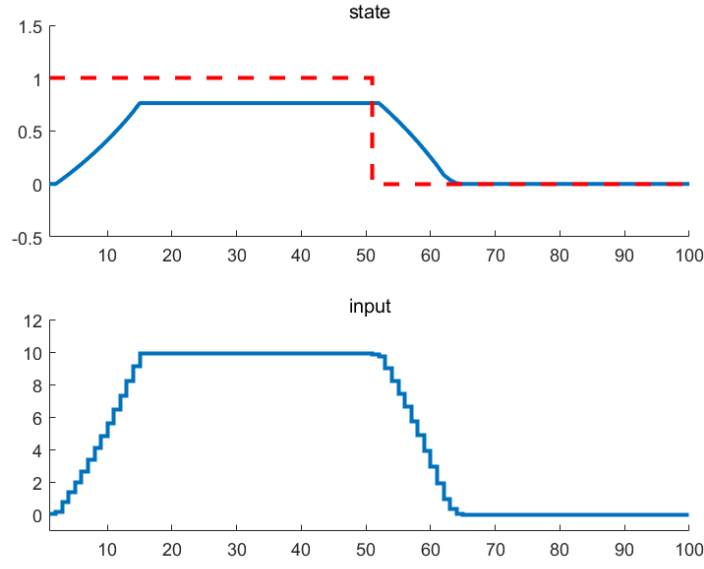


Figure 4: Graph q3

In the figure 4, the input signal is limited to 10, and Δu is limited to 1, the output of the system cannot follow the reference well.

2.4 $|u(t)| \leq 10, |\Delta u(t)| \leq 3, T_s = 0.1s$

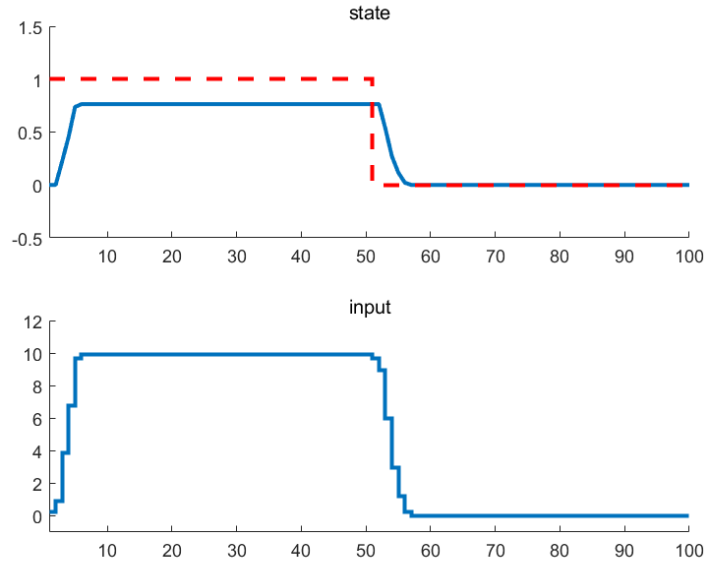


Figure 5: Graph q4

In the figure 5, the input signal is limited to 10, and Δu is limited to 3, the output of the system cannot follow the reference well.

2.5 $|u(t)| \leq 10, |\Delta u(t)| \leq 3, T_s = 1s$

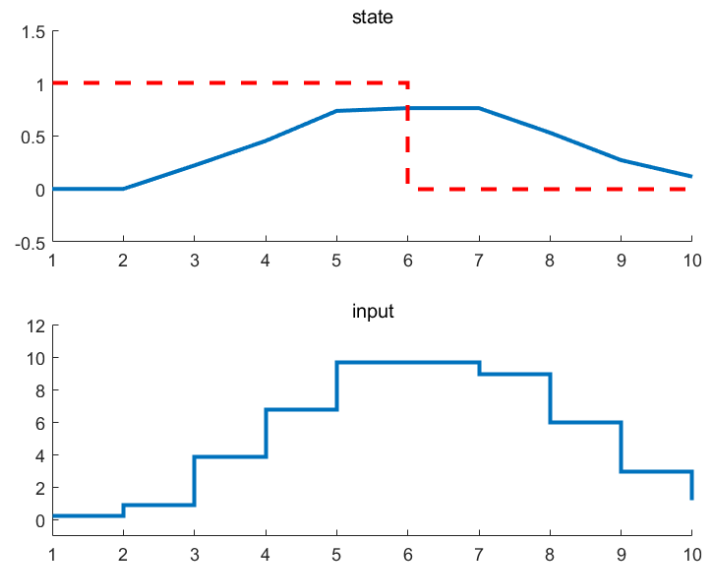


Figure 6: Graph q5

In the figure 6, the input signal is limited to 10, and Δu is limited to 3, the system badly followed the reference.

3 Problem 3

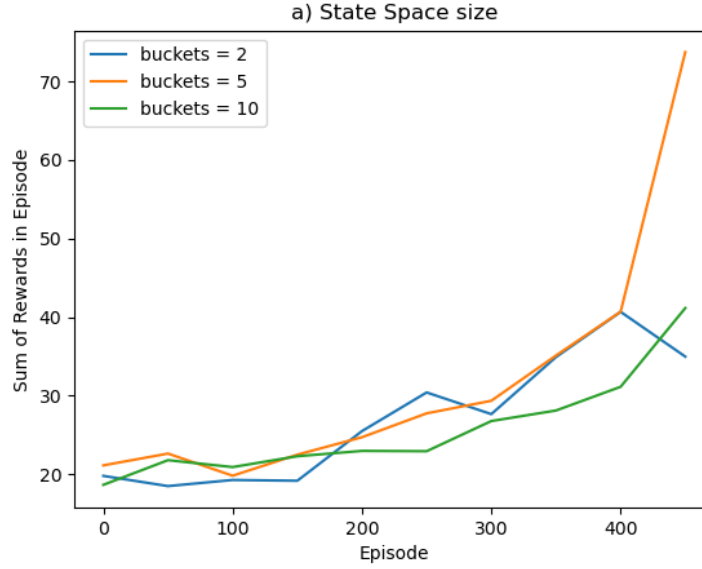


Figure 7: Graph q1

In the figure 7, the reward of the system is ascending, the reward of the system is ascending. In the last episode the reward of the system with buckets 5 is higher than the reward of the system with buckets 10 and 2.

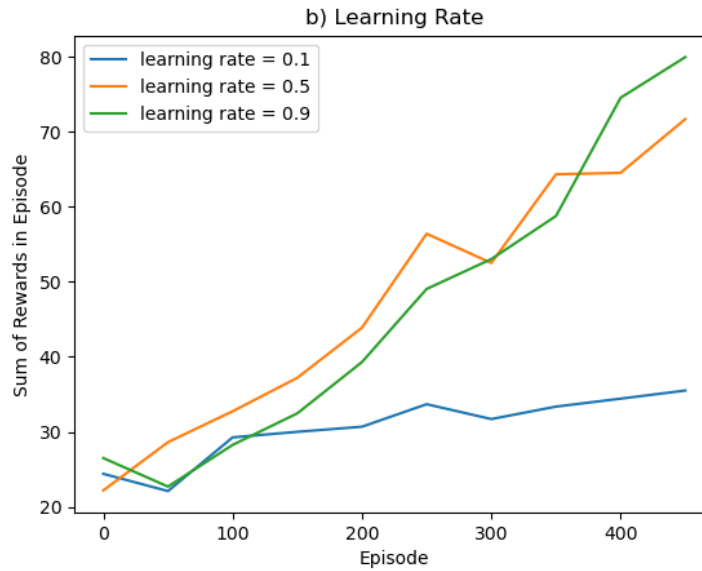


Figure 8: Graph q2

In the figure 8, the reward of the system is ascending, the system with learning rate 0.5 and 0.1 is ascending faster than the system with learning rate 0.9.

In the figure 9, the reward of the system is fluctuating and slightly ascending, the performances of the system with different discount factor are similar.

In the figure 10, the reward of the system is ascending, the system with epsilon 0.1 is ascending faster than the system with epsilon 0.9.

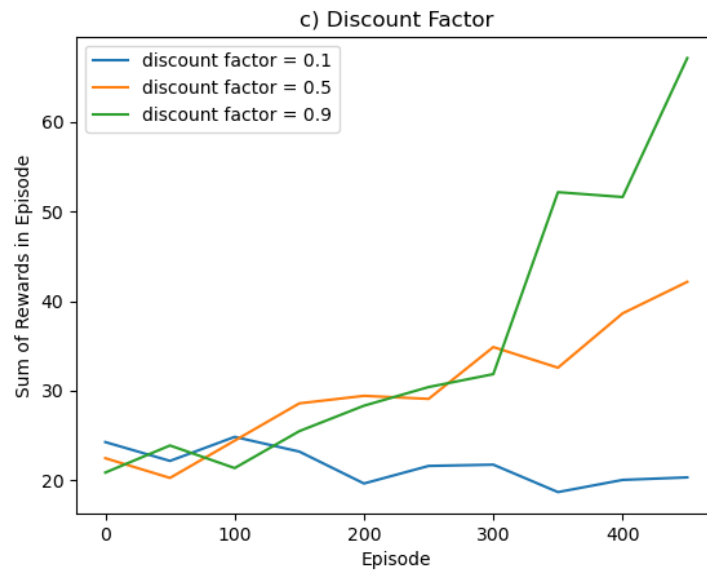


Figure 9: Graph q3

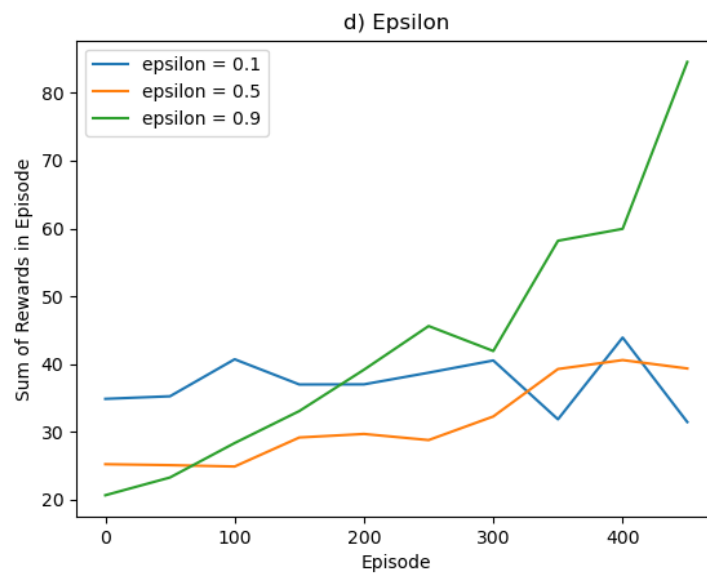


Figure 10: Graph q4