Practical Assignment № 5

Optimal Control

variant number : 6

Student Name： Xu Miao, Zhou Haojie

HDU Number： 19322103, 19322233

ITMO number : 293687, 293806

Optimal control design for LTI plant (Bellman's dynamic programming)

1. Symbol description and experimental values（Group 6）

- **experimental values**

№	A	b	Q	r
6	$\begin{bmatrix} 0 & 1 \\ 9 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$	1

- **symbol description:**

To facilitate the description of subsequent calculations, we named the elements in the matrix as follows

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix},$$

2. find the parameters of optimal controller using Bellman's dynamic programming and simulate the control system work on a given time interval.

(Use the parameters for A, b, Q, r from work №2. Choose the initial conditions $x(0) = [1, 0]^T$. Plot the variables x_1, x_2, u and J.)

- **Description of the dynamic programming problem**

- **Plant:**

$$\dot{x}(t) = f(x(t), u(t)) = Ax + Bu \qquad (1)$$

- **Cost function (performance index, optimization criterion):**

$$J = \int_0^\infty x^T(\tau)Qx(\tau) + ru^2(\tau)d\tau = \int_0^\infty f_0(x(t), u(t))dt \quad (2)$$

- **Functional equation of dynamic programming:**

$$S(x)=\min_{u(t)}\left[\int_{0}^{\infty} f_0(x(t),u(t))dt\right] \tag{3}$$

- **Hamilton-Jacobi-Bellman (HJB) equation:**

$$\begin{aligned}&\min_{u(t)}\left[f_0(x(t),u(t))+\frac{\partial S}{\partial x}\dot{x}+\frac{\partial S}{\partial t}\right]=0\\&\min_{u(t)}\left[f_0(x(t),u(t))+\frac{\partial S}{\partial x}f(x(t),u(t))\right]=-\frac{\partial S}{\partial t}=0\end{aligned} \tag{4}$$

- **Optimal control law:**

$$\begin{aligned}&f_0(x(t),u(t))+\frac{\partial S}{\partial x}f(x(t),u(t))=0\\&\frac{\partial}{\partial u}\left[f_0(x(t),u(t))+\frac{\partial S}{\partial x}f(x(t),u(t))\right]=0\end{aligned} \tag{5}$$

- **calculation process**

- **Plant:**

$$\begin{aligned}\dot{x}_1(t)&=a_{11}x_1(t)+a_{12}x_2(t)+b_1u(t)\\\dot{x}_2(t)&=a_{21}x_1(t)+a_{22}x_2(t)+b_2u(t)\end{aligned} \tag{6}$$

- **Cost function:**

$$J=\int_{0}^{t_f} q_1x_1^2(t)+q_2x_2^2(t)+ru^2(t)dt \tag{7}$$

- **Hamilton-Jacobi-Bellman (HJB) equation:**

$$\min_{u(t)}\left[q_1x_1^2+q_2x_2^2+ru^2+\frac{\partial S}{\partial x}\dot{x}\right]=\frac{\partial S}{\partial t}=0 \tag{8}$$

$$q_1x_1^2+q_2x_2^2+ru^2+\frac{\partial S}{\partial x_1}(a_{11}x_1+a_{12}x_2+b_1u)+\frac{\partial S}{\partial x_2}(a_{21}x_1+a_{22}x_2+b_2u)=-\frac{\partial S}{\partial t}=0 \tag{9}$$

- **Optimal control law:**

$$q_1x_1^2+q_2x_2^2+ru^2+\frac{\partial S}{\partial x_1}(a_{11}x_1+a_{12}x_2+b_1u)+\frac{\partial S}{\partial x_2}(a_{21}x_1+a_{22}x_2+b_2u)=-\frac{\partial S}{\partial t}=0 \tag{10}$$

$$\frac{\partial}{\partial u}\left[-\frac{\partial S}{\partial t}\right]=2ru+\frac{\partial S}{\partial x_1}b_1+\frac{\partial S}{\partial x_2}b_2=0\Rightarrow u=-\frac{1}{2r}\left(\frac{\partial S}{\partial x_1}b_1+\frac{\partial S}{\partial x_2}b_2\right) \tag{11}$$

Bringing equation (11) into equation (10), we get:

$$
q_1x_1^2+q_2x_2^2+r\left(\frac{1}{2r}\left(\frac{\partial S}{\partial x_1}b_1+\frac{\partial S}{\partial x_2}b_2\right)\right)^2+\frac{\partial S}{\partial x_1}\left(a_{11}x_1+a_{12}x_2-\frac{b_1}{2r}\left(\frac{\partial S}{\partial x_1}b_1+\frac{\partial S}{\partial x_2}b_2\right)\right) \\
+\frac{\partial S}{\partial x_2}\left(a_{21}x_1+a_{22}x_2-\frac{b_2}{2r}\left(\frac{\partial S}{\partial x_1}b_1+\frac{\partial S}{\partial x_2}b_2\right)\right)=-\frac{\partial S}{\partial t}=0 \tag{12}
$$

Expanding equation (12) and then simplifying it we get:

$$
q_1x_1^2+q_2x_2^2-\frac{1}{4r}\left(\frac{\partial S}{\partial x_1}b_1\right)^2-\frac{b_2}{2r}\frac{\partial S}{\partial x_2}\frac{\partial S}{\partial x_1}b_1-\frac{1}{4r}\left(\frac{\partial S}{\partial x_2}b_2\right)^2+ \\
+\frac{\partial S}{\partial x_1}a_{11}x_1+\frac{\partial S}{\partial x_1}a_{12}x_2+\frac{\partial S}{\partial x_2}a_{21}x_1+\frac{\partial S}{\partial x_2}a_{22}x_2=-\frac{\partial S}{\partial t}=0 \tag{13}
$$

Organizing equation (13) we obtain:

$$
q_1x_1^2+q_2x_2^2-\left(\frac{1}{2\sqrt{r}}\left(\frac{\partial S}{\partial x_1}b_1\right)+\frac{1}{2\sqrt{r}}\left(\frac{\partial S}{\partial x_2}b_2\right)\right)^2+\frac{\partial S}{\partial x_1}(a_{11}x_1+a_{12}x_2)+\frac{\partial S}{\partial x_2}(a_{21}x_1+a_{22}x_2)=-\frac{\partial S}{\partial t}=0 \tag{14}
$$

- **Selection of functional equation of dynamic programming:**

Assume that the functional equation of dynamic programming is

$$
S=\psi_1x_1^2+\psi_{12}x_1x_2+\psi_2x_2^2 \tag{15}
$$

Bringing (15) into (14) we get:

$$
q_1x_1^2+q_2x_2^2-\frac{1}{4r}(b_1(2\psi_1x_1+\psi_{12}x_2)+b_2(\psi_{12}x_1+2\psi_2x_2))^2+ \\
+(2\psi_1x_1+\psi_{12}x_2)(a_{11}x_1+a_{12}x_2)+(\psi_{12}x_1+2\psi_2x_2)(a_{21}x_1+a_{22}x_2)=0 \tag{16}
$$

Organizing equation (16) we obtain:

$$
\left(q_1+2a_{11}\psi_1+a_{21}\psi_{12}-\left(\frac{b_1\psi_1}{\sqrt{r}}\right)^2-\left(\frac{b_2\psi_{12}}{2\sqrt{r}}\right)^2-\frac{b_1b_2\psi_1\psi_{12}}{r}\right)x_1^2+ \\
+\left(2a_{12}\psi_1+2a_{21}\psi_2+a_{11}\psi_{12}+a_{22}\psi_{12}-\frac{b_1b_2\psi_{12}^2}{2r}-\frac{b_1^2\psi_1\psi_{12}}{r}-\frac{b_2^2\psi_2\psi_{12}}{r}-\frac{2b_1b_2\psi_1\psi_2}{r}\right)x_1x_2 \\
+\left(q_{22}+2a_{22}\psi_2+a_{12}\psi_{12}-\left(\frac{b_2\psi_2}{\sqrt{r}}\right)^2-\left(\frac{b_1\psi_{12}}{2\sqrt{r}}\right)^2-\frac{b_1b_2\psi_2\psi_{12}}{r}\right)x_2^2 \\
=0 \tag{17}
$$

Let the left and right sides of equation (17) be equal to solve for the optimal control law:

- **Optimal control law:**

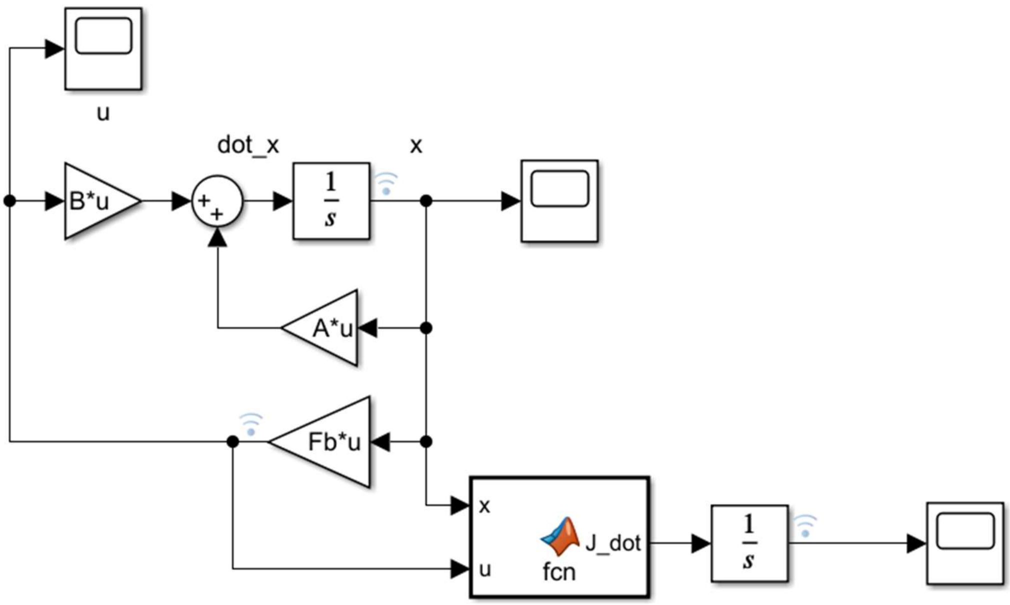
$$\begin{cases} 0 = q_1 + 2a_{11}\psi_1 + a_{21}\psi_{12} - \left(\frac{b_1\psi_1}{\sqrt{r}}\right)^2 - \left(\frac{b_2\psi_{12}}{2\sqrt{r}}\right)^2 - \frac{b_1b_2\psi_1\psi_{12}}{r} \\ 0 = 2a_{12}\psi_1 + 2a_{21}\psi_2 + a_{11}\psi_{12} + a_{22}\psi_{12} - \frac{b_1b_2\psi_{12}^2}{2r} - \frac{b_1^2\psi_1\psi_{12}}{r} - \frac{b_2^2\psi_2\psi_{12}}{r} - \frac{2b_1b_2\psi_1\psi_2}{r} \\ 0 = q_2 + 2a_{22}\psi_2 + a_{12}\psi_{12} - \left(\frac{b_2\psi_2}{\sqrt{r}}\right)^2 - \left(\frac{b_1\psi_{12}}{2\sqrt{r}}\right)^2 - \frac{b_1b_2\psi_2\psi_{12}}{r} \end{cases} \quad (18)$$

$$u = -\frac{1}{2r}((2\psi_1x_1 + \psi_{12}x_2)b_1 + (\psi_{12}x_1 + 2\psi_2x_2)b_2) \quad (19)$$

- **Code**

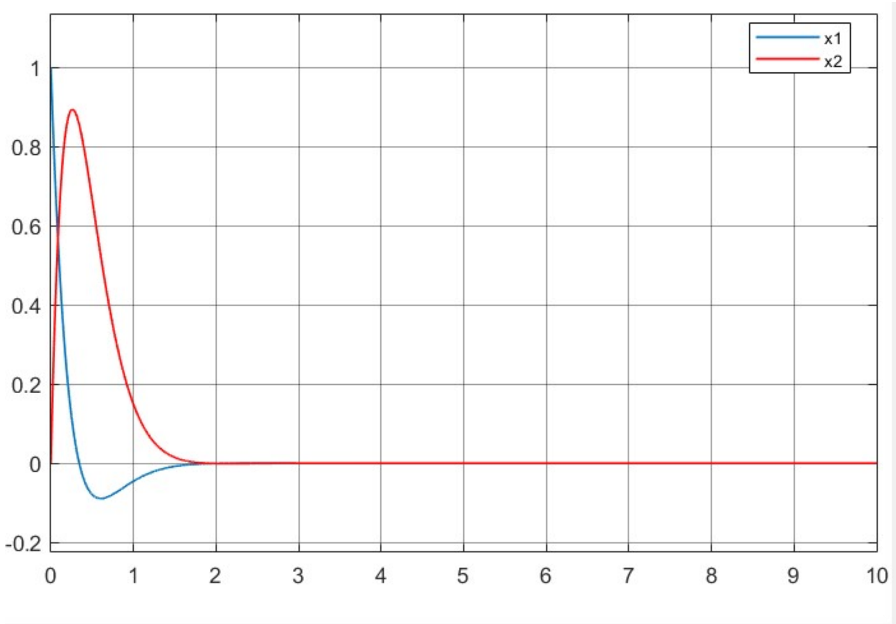
```
A = [0 1;9 -1];B = [1;0];Q = [1 0; 0 2];r = 1;

a11 = A(1,1);
a12 = A(1,2);
a21 = A(2,1);
a22 = A(2,2);

b1 = B(1,1);b2 = B(2,1);q1 = 1;q12 = 0;q22 = 2;
syms psi1 psi12 psi2
eqn1 = q1 + 2*a11*psi1 + a21*psi12 - (b1*psi1/sqrt(r))^2 - (b2*psi12/(2*sqrt(r)))^2 -
b1*b2*psi1*psi12/r == 0;
eqn2 = q12 + 2*a12*psi1 +2*a21*psi2 + a11*psi12 + a22*psi12- b1*b2*psi12^2/(2*r) -
b1^2*psi1*psi12/r - b2^2*psi2*psi12/r - 2*b1*b2*psi1*psi2/r == 0;
eqn3 = q22 +2*a22*psi2 + a12*psi12 - (b2*psi2/sqrt(r))^2 - (b1*psi12/(2*sqrt(r)))^2 -
b1*b2*psi2*psi12/r == 0;
[p1,p2,p3] = solve(eqn1,eqn2,eqn3,psi1,psi12,psi2);
vpa(p1)
p1 = double([ans(3,1)]); % psi1
vpa(p2)
p12 = double([ans(3,1)]); % psi12
vpa(p3)
p2 = double([ans(3,1)]); % psi2

Fb = [-1/(2*r) * (2*p1*b1 + p12*b2), -1/(2*r) * (2*p2*b2 + p12*b1)];
```

- **Simulink model**

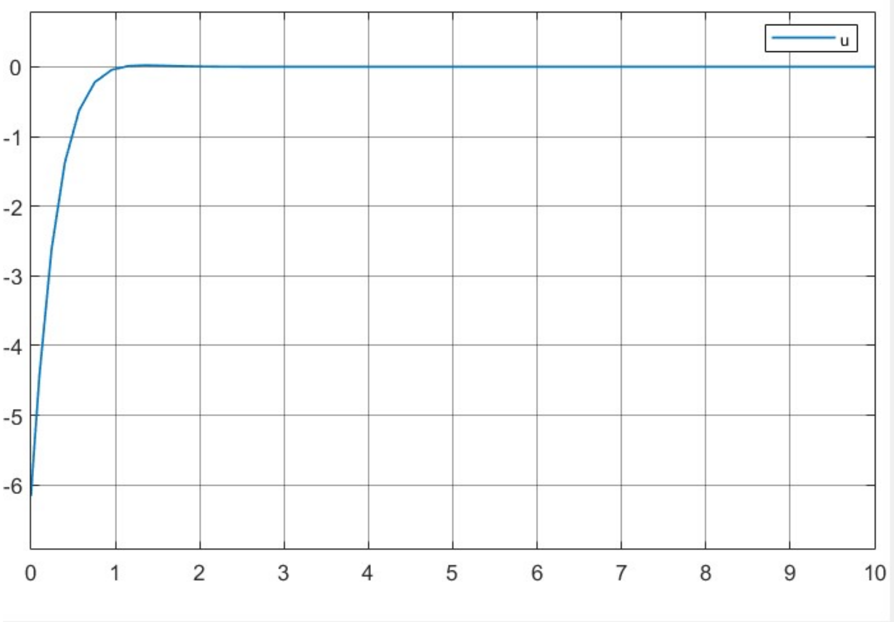


- **Simulink result**

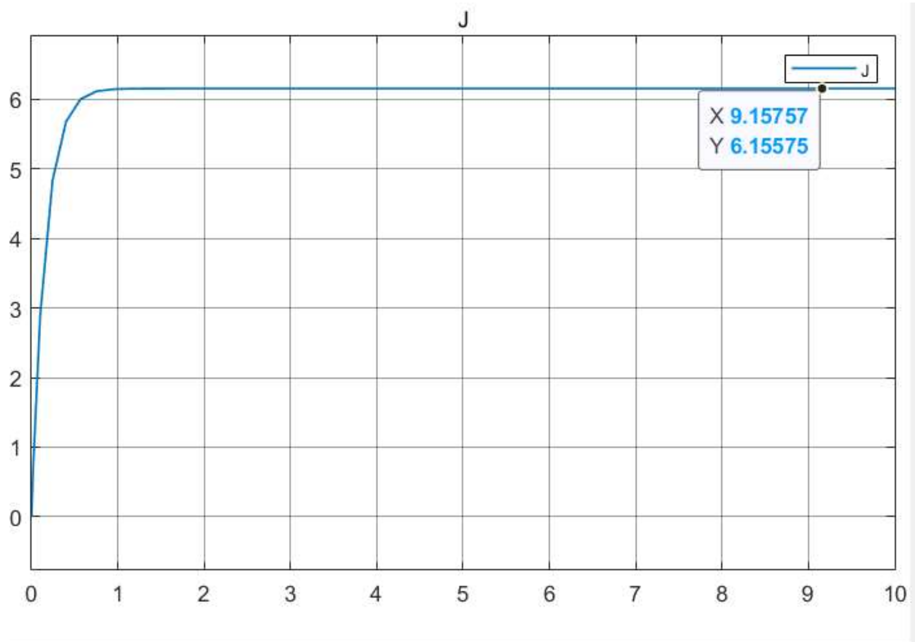
- State variables x(t)



- Control variables u(t)



- cost function J



3. **Plot the cost function for changed controller parameters from the optimal ones.**

   optimal ones:

$$u=-\frac{1}{2r}((2\psi_1x_1+\psi_{12}x_2)b_1+(\psi_{12}x_1+2\psi_2x_2)b_2)=-6.158x_1-2.051x_2 \quad (20)$$

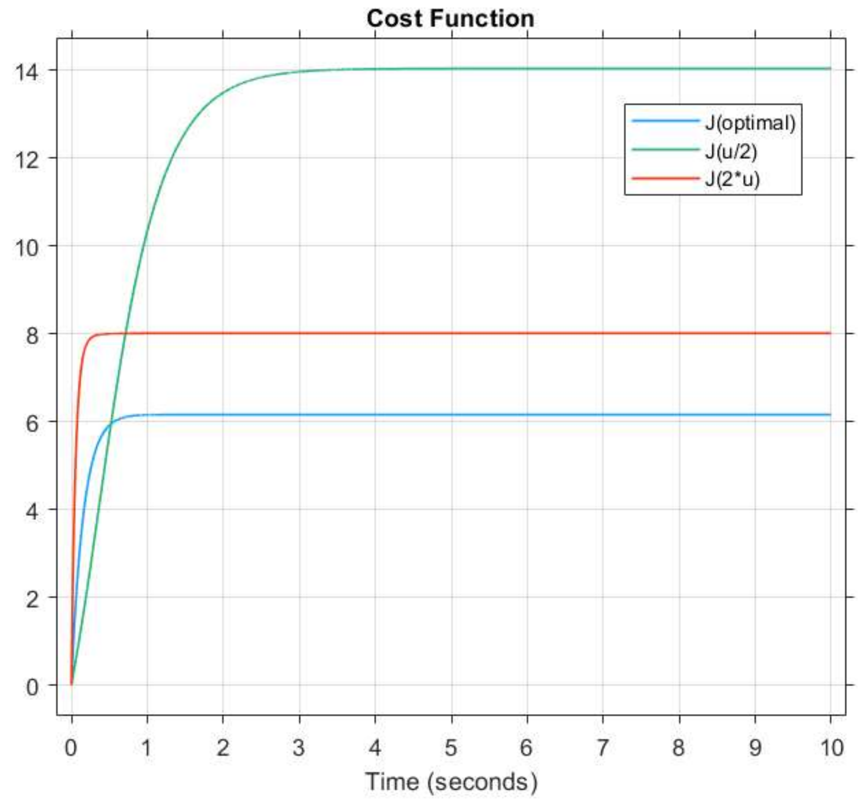
changed ones:

$$u(t)=-3x_1-x_2 \quad (21.1)$$

$$u(t)=-12x_1-4x_2 \quad (21.2)$$

The original control scheme is changed to Equation (5), and the simulation results are shown as follows:

- Cost Function



- **Conclusion:** By comparing the cost function of the optimal control and changed ones, we can find that no matter the parameters become larger or smaller, the cost alawys increase.

4. Comparison with the optimal control calculated by the LQR method (practice 2)

In Practical Assignment № 5 we used the LQR method to calculate the negative feedback coefficient as:

$$K = [6.1583 \qquad 2.0514] \tag{21}$$

Comparing with equation (20) we can find that the optimal negative feedback **control law of the system is the same** using Bellman dynamic programming method and LQR method.

- **Conclusion:**

In summary, we can see that we can use **different methods to obtain the optimal control law** (LOR, Bellman dynamic programming), and for the system in the exercise the two methods obtain the same control law.

However, during the experiment we also found that **these two methods have their own advantages and disadvantages** and are more suitable for different situations respectively.

LOR:

Advantages: Simpler, less computational process,

Disadvantages: Can only be used to calculate control rate coefficients for simple control forms as LSFB

Bellman dynamic programming:

Disadvantages: More complex, with more computational processes,

Advantages: It can be applied to the calculation of the control law in more complex cases (e.g. when F(x) in the loss function is not 0, when S(x,t) is time-dependent, i.e. the partial derivative with respect to t is not 0. Then the feedback coefficient of the optimal control law should be a function of time at this point. (The Bellman dynamic programming method can still be calculated in this case, but the LQR cannot obtain the optimal control law)