

The Hamilton-Jacobi-Bellman Equation

Contents

1	The Hamilton-Jacobi-Bellman Equation	2
1.1	The optimal control problem	2
1.2	Deriving the HJB equation	2
1.2.1	Summary for the analytical solution of the HJB equation	5
1.3	Solution of the HJB equation by examples	6
1.4	Exercises	14

The Hamilton-Jacobi-Bellman Equation

1.1 The optimal control problem

In the previous description of the dynamic programming method, we approximated continuous systems to discrete systems. This approach leads to a recurrence relation that is suited for digital implementation. In this chapter we shall consider another approach, for continuous systems, that leads to a nonlinear partial differential equation - the Hamilton-Jacobi-Bellman (HJB) equation. We shall show how the optimal performance measure, if it satisfies the Hamilton-Jacobi equation, determines an optimal control.

Consider the following continuous optimal control problem.

The process described by the state equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \quad (1.1)$$

is to be controlled to minimize the performance measure

$$J = h(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \quad (1.2)$$

where g and h are scalar functions, t_0 and t_f are fixed, τ is a variable of integration.

1.2 Deriving the HJB equation

The HJB equation will be obtained using the principle of optimality and the results obtained for the dynamic programming.

The continuous optimal control problem will be included in a larger class of problems by considering the performance measure:

$$J(\mathbf{x}(t), t, \mathbf{u}(\tau)_{t \leq \tau \leq t_f}) = h(\mathbf{x}(t_f), t_f) + \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \quad (1.3)$$

where t can be any value less than or equal to t_f and $\mathbf{x}(t)$ can be any admissible state value. Notice that the performance measure will depend on the numerical values for $\mathbf{x}(t)$ and t and on the optimal control history in the interval $[t, t_f]$, (Kirk, 2004).

Let us now attempt to determine the controls that minimize (1.3) for all admissible $\mathbf{x}(t)$ and all $t \leq t_f$. The minimum cost function is then:

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(\tau)} \left\{ h(\mathbf{x}(t_f), t_f) + \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \right\} \quad (1.4)$$

We shall split the time interval $[t, t_f]$ into $[t, t + \Delta t]$ and $[t + \Delta t, t_f]$, and are specifically interested in the case where $\Delta t \rightarrow 0$. By subdividing the interval we obtain:

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(\tau)} \left\{ h(\mathbf{x}(t_f), t_f) + \int_t^{t+\Delta t} g(\mathbf{x}, \mathbf{u}, \tau) d\tau + \int_{t+\Delta t}^{t_f} g(\mathbf{x}, \mathbf{u}, \tau) d\tau \right\} \quad (1.5)$$

The principle of optimality requires that:

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}} \left\{ \int_t^{t+\Delta t} g(\mathbf{x}, \mathbf{u}, \tau) d\tau + J^*(\mathbf{x}(t + \Delta t), t + \Delta t) \right\} \quad (1.6)$$

where $J^*(\mathbf{x}(t + \Delta t), t + \Delta t)$ is the minimum cost of the process for the time interval $t + \Delta t \leq \tau \leq t_f$ with 'initial' state $\mathbf{x}(t + \Delta t)$.

Assuming J^* has bounded second derivatives in both arguments, we can expand $J^*(\mathbf{x}(t + \Delta t), t + \Delta t)$ in a Taylor series about the point $(\mathbf{x}(t), t)$ (truncated after the first order terms) to obtain:

$$\begin{aligned} J^*(\mathbf{x}(t), t) &= \min_{\mathbf{u}} \left\{ \int_t^{t+\Delta t} g(\mathbf{x}, \mathbf{u}, \tau) d\tau + J^*(\mathbf{x}(t), t) + \right. \\ &\quad \left. + \left[\frac{\partial J^*}{\partial t}(\mathbf{x}(t), t) \right] \Delta t + \left[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x}(t), t) \right]^T [\mathbf{x}(t + \Delta t) - \mathbf{x}(t)] \right\} \end{aligned} \quad (1.7)$$

For small Δt ,

$$\begin{aligned} \mathbf{x}(t + \Delta t) - \mathbf{x}(t) &= \dot{\mathbf{x}}(t) \cdot \Delta t \\ \int_t^{t+\Delta t} g(\mathbf{x}, \mathbf{u}, \tau) d\tau &= g(\mathbf{x}, \mathbf{u}, t) \Delta t \end{aligned}$$

Because of $J^*(\mathbf{x}(t), t)$ is independent of $\mathbf{u}$ we cancel the terms from the right and left hand side of (1.7). $\dot{\mathbf{x}}$ will be replaced using the state equation (1.1):

$$0 = \min_{\mathbf{u}(\tau)} \left\{ g(\mathbf{x}, \mathbf{u}, t) \Delta t + \frac{\partial J^*(\mathbf{x}(t), t)}{\partial t} \Delta t + \left[\frac{\partial J^*(\mathbf{x}(t), t)}{\partial \mathbf{x}} \right]^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \Delta t \right\} \quad (1.8)$$

Dividing the above equation by Δt yields:

$$0 = \frac{\partial J^*(\mathbf{x}(t), t)}{\partial t} + \min_{\mathbf{u}(\tau)} \left\{ g(\mathbf{x}, \mathbf{u}, t) + \left[\frac{\partial J^*(\mathbf{x}(t), t)}{\partial \mathbf{x}} \right]^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\} \quad (1.9)$$

Define the Hamiltonian as:

$$H = g(\mathbf{x}, \mathbf{u}, t) + \left[\frac{\partial J^*(\mathbf{x}(t), t)}{\partial \mathbf{x}} \right]^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad (1.10)$$

Then we write the partial differential equation (1.9) as:

$$0 = \frac{\partial J^*(\mathbf{x}(t), t)}{\partial t} + \min_{\mathbf{u}} H \quad (1.11)$$

To find the boundary value for this partial differential equation set $t = t_f$; from (1.4) we have:

$$J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f) \quad (1.12)$$

We have obtained the **Hamilton-Jacobi-Bellman equation - HJB** (1.11) subject to the boundary conditions (1.12). It provides the solution to the optimal control problems for general nonlinear dynamical systems. However, analytical solution to the HJB equation is difficult to obtain in most cases.

A solution could be obtained analytically by "guessing" a form of the minimum cost function $J^*(x(t), t)$. In general, the HJB equation must be solved by numerical techniques. Actually, a numerical solution involves some sort of a discrete approximation to the exact optimization relationship (1.11), (Kirk, 2004).

To simplify the notations, let:

$$J_x^* = \frac{\partial J^*(\mathbf{x}(t), t)}{\partial x} \quad \text{and} \quad J_t^* = \frac{\partial J^*(\mathbf{x}(t), t)}{\partial t}$$

The HJB equation is then written as:

$$0 = J_t^* + \min_{\mathbf{u}} H$$

with the Hamiltonian:

$$H = g(\mathbf{x}, \mathbf{u}, t) + [J_x^*]^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$$

and the boundary condition:

$$J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f)$$

1.2.1 Summary for the analytical solution of the HJB equation

The summary will be written for a first-order system with one control input.

A. Statement of the optimal control problem

For a system described by the state equation:

$$\dot{x}(t) = f(x(t), u(t), t) \quad (1.13)$$

with $x(0) = x_0$, $t \in [t_0, t_f]$, determine the optimal control law $u^*(t)$ that minimizes the performance measure:

$$J = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t) dt \quad (1.14)$$

B. Solution

1. Write the Hamiltonian:

$$H(x, u, t, J_x^*) = g(x, u, t) + J_x^* \cdot f(x, u, t) \quad (1.15)$$

2. Minimize the Hamiltonian with respect to $u(t)$. Compute the necessary and sufficient conditions for minimum. Set:

$$\frac{\partial H}{\partial u} = 0$$

and obtain $u^*(t)$ as a function of (x, t, J_x^*) :

$$u^*(t) = U(x, t, J_x^*) \quad (1.16)$$

3. Obtain the minimum value of the Hamiltonian by replacing (1.16) into (1.15):

$$\min_u H = H(x, u^*, t, J_x^*)|_{u^*=U(x, t, J_x^*)}$$

4. Write the HJB equation

$$0 = J_t^* + \min_{\mathbf{u}} H$$

and the boundary condition

$$J^*(x(t_f), t_f) = h(x(t_f), t_f)$$

5. Solve the HJB equation and obtain $J^*(x(t), t)$.

6. Use the solution from step 5 to evaluate J_x^* and substitute into the expression for $u^*(t)$ from (1.16) to obtain the optimal control

Obs. Obtaining a solution of the HJB equation can be a difficult task. A usual approach could be "guessing" or selecting a form of the solution $J^*(x(t), t)$ that will transform the partial differential equation into an ordinary differential equation.

1.3 Solution of the HJB equation by examples

Example 1.1 A first order system is described by the differential equation

$$\dot{x}(t) = x(t) + u(t)$$

It is desired to find the control law that minimizes the cost function

$$J = \frac{1}{4}x^2(T) + \int_0^T \frac{1}{4}u^2(t)dt$$

The final time T is specified and the admissible state and control values are not constrained by any boundaries.

$$g(x(t), u(t), t) = \frac{1}{4}u^2(t)$$

$$\mathbf{f}(x(t), u(t), t) = x(t) + u(t)$$

The hamiltonian:

$$H(x(t), u(t), J_x^*) = \frac{1}{4}u^2(t) + J_x^* \cdot (x(t) + u(t))$$

and since the control is unconstrained, a necessary condition that optimal control must satisfy is:

$$\frac{\partial H}{\partial u} = \frac{1}{2}u(t) + J_x^* = 0$$

The control indeed minimizes the hamiltonian function because

$$\frac{\partial^2 H}{\partial u^2} = \frac{1}{2} > 0$$

The optimal control:

$$u^*(t) = -2J_x^*$$

when substituted into the HJB equation:

$$0 = J_t^* + \min_u H$$

gives:

$$\begin{aligned} 0 &= J_t^* + \frac{1}{4}(-2J_x^*)^2 + J_x^*(x(t) - 2J_x^*) \\ 0 &= J_t^* - (J_x^*)^2 + J_x^*x(t) \end{aligned} \quad (1.17)$$

The boundary value for J^* is:

$$J^*(x(T), T) = \frac{1}{4}x^2(T) \quad (1.18)$$

One way to solve the HJB equation is to guess a form for the solution and see if it can be made to satisfy the differential equation and the boundary conditions. Since $J^*(x(T), T)$ is quadratic in $x(T)$ guess

$$J^*(x(t), t) = \frac{1}{2}p(t)x^2(t) \quad (1.19)$$

where $p(t)$ represents an unknown scalar function of t that is to be determined. Notice that

$$J_x^* = \frac{\partial J^*}{\partial x} = p(t)x(t) \quad (1.20)$$

which together with the expression determined for $u^*(t)$ implies that:

$$u^*(t) = -2p(t)x(t)$$

Thus, if $p(t)$ can be found such that (1.17) and (1.18) are satisfied, the optimal control is linear feedback of the state - indeed this was the motivation for selecting the form (1.19).

Substituting (1.20) and

$$J_t^* = \frac{\partial J^*}{\partial t} = \frac{1}{2}\dot{p}(t)x^2(t) \quad (1.21)$$

into (1.17) gives:

$$0 = \frac{1}{2}\dot{p}(t)x^2(t) - p^2(t)x^2(t) + p(t)x^2(t)$$

Since this equation must be satisfied for all $x(t)$, $p(t)$ may be calculated from the differential equation:

$$\frac{1}{2}\dot{p}(t) - p^2(t) + p(t) = 0 \quad (1.22)$$

with the final condition: $p(T) = 1/2$. The final value $p(T)$ was obtained by replacing $t = T$ into (1.19) and setting it equal to the boundary condition of the HJB equation (1.18):

$$J^*(x(T), T) = \frac{1}{2}p(T)x^2(T) = \frac{1}{4}x^2(T), \Rightarrow p(T) = \frac{1}{2}$$

$p(t)$ is a scalar function of t ; therefore the solution can be obtained using the transformation $z(t) = 1/p(t)$, with the result:

$$p(t) = \frac{1}{1 + e^{2(t-T)}} \quad (1.23)$$

Obs. The solution of (1.22) is obtained as follows:

$$\begin{aligned} z(t) &= \frac{1}{p(t)}; \quad z(T) = \frac{1}{p(T)} = 2 \\ (1.22) &\Rightarrow -\frac{\dot{z}}{z^2} - \frac{2}{z^2} + \frac{2}{z} = 0 \\ &\quad \dot{z} - 2z + 2 = 0 \end{aligned}$$

The solution of homogeneous part of the above equation:

$$\dot{z} - 2z = 0, \text{ is } z(t) = C_1 e^{2t}$$

and the general solution:

$$z(t) = C_1 e^{2t} + C_2$$

The constants are calculated by replacing the general solution into the equation and using the final condition:

$$2C_1 e^{2t} - 2(C_1 e^{2t} + C_2) + 2 = 0 \Rightarrow -2C_2 + 2 = 0 \Rightarrow C_2 = 1$$

$$z(t) = C_1 e^{2t} + 1; \quad z(T) = C_1 e^{2T} + 1 = 2; \quad C_1 = e^{-2T}$$

Then, the general solution yields:

$$z(t) = e^{2(t-T)} + 1 \Rightarrow p(t) = \frac{1}{e^{2(t-T)} + 1}$$

The optimal control law is then:

$$u^*(t) = -2 \frac{\partial J^*(x, t)}{\partial x} = -2p(t)x(t)$$

$$u^*(t) = -\frac{2}{1 + e^{2(t-T)}}x(t)$$

The controller is a time-varying one and the block diagram is shown in Figure 1.1.

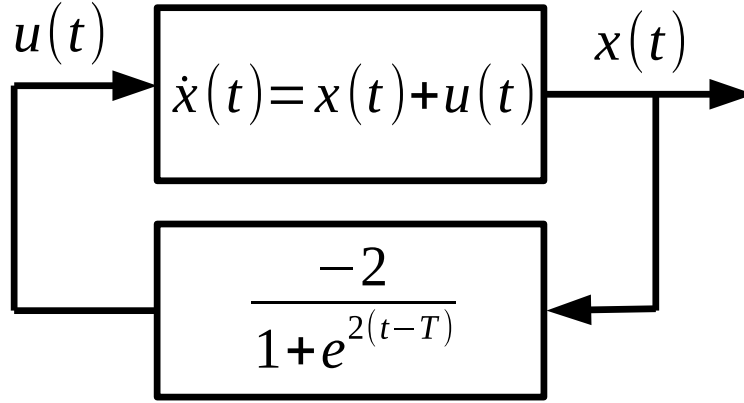


Figure 1.1: Optimal control for example 1.1

Notice that as $T \rightarrow \infty$, the linear time-varying feedback approaches constant feedback $p(t) = 1$ and the controlled system

$$\dot{x}(t) = x(t) - 2x(t) = -x(t)$$

is stable.

Example 1.2 A first order system is described by the differential equation

$$\dot{x}(t) = u(t), \quad x(0) = x_0, \quad t \in [0, T]$$

It is desired to find the control law that minimizes the cost function

$$J = \frac{1}{2}x^2(T) + \int_0^T \frac{1}{2}u^2(t)dt$$

The final time T is specified and the admissible state and control values are not constrained by any boundaries.

The functions involved in the Hamiltonian are:

$$g(x(t), u(t), t) = \frac{1}{2}u^2(t)$$

and

$$f(x(t), u(t), t) = u(t)$$

Therefore, the hamiltonian is:

$$H(x(t), u(t), J_x^*) = \frac{1}{2}u^2(t) + J_x^* \cdot u(t)$$

and since the control is unconstrained, a necessary condition that optimal control must satisfy is:

$$\frac{\partial H}{\partial u} = u(t) + J_x^* = 0 \quad (1.24)$$

The control indeed minimizes the hamiltonian function because

$$\frac{\partial^2 H}{\partial u^2} = 1 > 0$$

For the optimal control resulted from (1.24):

$$u^*(t) = -J_x^*$$

the minimum value of the hamiltonian is:

$$\min_u H = \frac{1}{2}J_x^{*2} + J_x^*(-J_x^*) = -\frac{1}{2}J_x^{*2}$$

When substituted into the HJB equation:

$$0 = J_t^* + \min_u H$$

it gives:

$$0 = J_t^* - \frac{1}{2}J_x^{*2} \quad (1.25)$$

The boundary value for J^* is:

$$J^*(x(T), T) = \frac{1}{2}x^2(T) \quad (1.26)$$

We will select a form for the solution and see if it can be made to satisfy the differential equation and the boundary conditions. Since $J^*(x(T), T)$ is quadratic in $x(T)$, the form of the solution can be:

$$J^*(x(t), t) = \frac{1}{2}p(t)x^2(t) \quad (1.27)$$

where $p(t)$ is an unknown scalar function of t that is to be determined. Notice that

$$J_x^* = \frac{\partial J^*}{\partial x} = p(t)x(t) \quad (1.28)$$

which together with the expression determined for $u^*(t)$ implies that:

$$u^*(t) = -p(t)x(t)$$

Thus, if $p(t)$ can be found such that (1.25) and (1.26) are satisfied, the optimal control is linear feedback of the state.

Substituting J_x^* from (1.28) and

$$J_t^* = \frac{\partial J^*}{\partial t} = \frac{1}{2}\dot{p}(t)x^2(t) \quad (1.29)$$

into (1.25) gives:

$$0 = \frac{1}{2}\dot{p}(t)x^2(t) - \frac{1}{2}p^2(t)x^2(t)$$

This equation can be divided by $x^2(t) \neq 0$ and then, $p(t)$ may be calculated from the ordinary differential equation:

$$\dot{p}(t) - p^2(t) = 0 \quad (1.30)$$

with the final condition: $p(T) = 1$, as resulted from the boundary condition of the HJB equation:

$$J^*(x(T), T) = \frac{1}{2}p(T)x^2(T) = \frac{1}{2}x^2(T), \Rightarrow p(T) = 1$$

The equation (1.30) can be solved analytically by changing the variable:

$$p(t) = \frac{1}{z(t)}, \Rightarrow \dot{p}(t) = -\frac{\dot{z}(t)}{z^2(t)}$$

The boundary condition in terms of the new variable $z(t)$ is:

$$z(T) = \frac{1}{p(T)} = 1$$

The ordinary differential equation becomes:

$$-\frac{\dot{z}(t)}{z^2(t)} - \frac{1}{z^2(t)} = 0 \quad \text{or} \quad \dot{z}(t) + 1 = 0$$

It has a simple solution: $z(t) = -t + C$ and the constant is computed using the final condition for z :

$$z(T) = -T + C = 1, \Rightarrow C = T + 1, \Rightarrow z(t) = -t + T + 1$$

Then, the function $p(t)$ is:

$$p(t) = \frac{1}{-t + T + 1} \quad (1.31)$$

The optimal control law is then:

$$u^*(t) = -J_x^* = -p(t)x(t) = \frac{1}{-t + T + 1}x(t)$$

It is also a time-varying state-feedback control law.

Example 1.3 , (Weber, 2000). A man is considering his lifetime plan of investment and expenditure. He has an initial level of savings $x(0)$ and no other income other than that which he obtains from investment at a fixed interest rate β . His total capital is therefore governed by the equation:

$$\dot{x}(t) = \beta x(t) - u(t),$$

where $\beta > 0$ and $u(t)$ is his rate of expenditure. He wishes to maximize

$$J = \int_0^T e^{-at} \sqrt{u(t)} dt$$

for a given T . Find his optimal policy $u^*(t)$.

The problem can be transformed into a minimum one by changing the sign of the performance measure:

$$J^*(x(t), t) = \min_u \int_0^T -e^{-at} \sqrt{u(t)} dt$$

Let $a = 1$; $\beta = 4$.

We solve this problem using the Hamilton-Jacobi-Bellman equation:

$$0 = \frac{\partial J^*}{\partial t} + \min_u H \Rightarrow J_t^* + H^* = 0$$

Then:

$$f(x, u, t) = 4x(t) - u(t); \quad g(x, u, t) = -e^{-t}\sqrt{u(t)}$$

and the Hamiltonian is:

$$H = g(x, u, t) + \frac{\partial J^*}{\partial x} \cdot f(x, u, t) = -e^{-t}\sqrt{u} + J_x^* \cdot (4x - u)$$

The boundary condition is: $J^*(x(T), T) = 0$ since $h(x(T), T) = 0$.

We calculate u^* that minimizes the hamiltonian from the necessary condition:

$$\begin{aligned} \frac{\partial H}{\partial u} &= 0 = -\frac{1}{2} \cdot \frac{e^{-t}}{\sqrt{u}} - J_x^* \\ u^* &= \frac{1}{4} \frac{1}{e^{2t} J_x^{*2}} = \frac{1}{4} e^{-2t} J_x^{*-2} \end{aligned}$$

Substituting u^* into the hamiltonian we obtain:

$$H^* = \frac{1}{4} \frac{(-3e^{-2t} + 16x J_x^{*2})}{J_x^*}$$

Suppose we try a solution of the form:

$$J^*(x(t), t) = p(t)\sqrt{x(t)}$$

Then:

$$\begin{aligned} J_t^* &= \frac{\partial J^*}{\partial t} = \dot{p}(t)\sqrt{x(t)} \\ J_x^* &= \frac{\partial J^*}{\partial x} = \frac{1}{2} \frac{p(t)}{\sqrt{x(t)}} \end{aligned}$$

By substituting the above, the HJB equation becomes:

$$\left(\dot{p} + \frac{1}{2} \frac{-3e^{-2t} + 4p^2}{p} \right) \sqrt{x} = 0$$

It has to be satisfied for any $x(t)$, thus we can obtain $p(t)$ from:

$$\dot{p} + \frac{1}{2} \frac{-3e^{-2t} + 4p^2}{p} = 0$$

subject to the final condition $p(T) = 0$.

The solution of the above equation is:

$$p(t) = \frac{1}{2} \sqrt{6e^{-2t} + 4e^{-4t} \cdot C}$$

The constant C is calculated from the final condition and has the value:

$$C = -\frac{3}{2e^{-2T}}$$

The unknown function $p(t)$ is:

$$p(t) = \frac{1}{2}\sqrt{6e^{-2t} - 6e^{-4t+2T}}$$

Then the optimal control law is given by:

$$u^*(t) = \frac{1}{4}e^{-2t}J_x^{*-2}, \text{ where } J_x^* = \frac{1}{2}\frac{p(t)}{\sqrt{x(t)}}$$

thus:

$$u^*(t) = \frac{4x(t)}{e^{2t}(e^{-2t} - 6e^{-4t+2T})} = \frac{4x(t)}{1 - 6e^{2(T-t)}}$$

1.4 Exercises

Exercise 1.1 Consider a general class of dynamical systems modeled by the differential equation:

$$\dot{x}(t) = ax(t) + bu(t)$$

with the associated cost index:

$$J = \frac{1}{2}fx^2(t_f) + \frac{1}{2}\int_{t_0}^{t_f} (qx^2(t) + ru^2(t)dt)$$

where the initial time $t_0 = 0$, the final time $t_f < \infty$ is fixed, and the final state $x(t_f)$ is free. Find a control $u^*(t)$ that minimizes J .

Exercise 1.2 Consider a first-order system described by the state equation

$$\dot{x}(t) = -x(t) + u(t)$$

Determine the optimal control $u^*(t)$ that minimizes the performance measure:

$$J = \frac{1}{2}x^2(1) + \int_0^1 u^2(t)dt$$

Hint. Consider as a candidate solution for the HJB equation $J^*(x(t), t) = x^2(t)p(t)$.

Exercise 1.3 Consider a first-order system described by the state equation

$$\dot{x}(t) = x(t) \cdot u(t)$$

Determine the optimal control $u^*(t)$ that minimizes the performance measure:

$$J = x^2(1) + \int_0^1 (x(t) \cdot u(t))^2 dt$$

Hint. Consider as a candidate solution for the HJB equation $J^*(x(t), t) = x^2(t)p(t)$.

Bibliography

Kirk, D. E. (2004). *Optimal Control Theory. An Introduction*. Dover Publications, Inc.

Weber, R. (2000). Optimization and control. online at www.statslab.cam.ac.uk.