

# Riccati eq., Linear Quadratic Regulator

## Control Theory, Lecture 7

by Sergei Savin

Spring 2024

## ■ Hamilton-Jacobi-Bellman equation

- ▶ Definitions
- ▶ Cost, optimal cost
- ▶ Differentiating optimal cost

## ■ Algebraic Riccati equation

- ▶ HJB for LTI
- ▶ Linear Quadratic Regulator
- ▶ Numerical methods

Let us define dynamics:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad (1)$$

with initial conditions  $\mathbf{x}(0)$ .

Additionally we define *control policy* as:

$$\mathbf{u} = \pi(\mathbf{x}, t) \quad (2)$$

To connect with the previous ways we talked about control, we can say that choosing different control gains and different feed-forward control amounts to choosing a different control policy.

Let  $J$  be an additive cost function:

$$J(\mathbf{x}_0, \pi(\mathbf{x}, t)) = \int_0^\infty g(\mathbf{x}, \mathbf{u}) dt \quad (3)$$

where  $g(\mathbf{x}, \mathbf{u})$  is instantaneous cost and  $\mathbf{x}_0 = \mathbf{x}(0)$  is the initial conditions. Notice that  $J$  depends on  $\mathbf{x}_0$  rather than  $\mathbf{x}(t)$ , since initial conditions and control policy completely define the trajectory of the system  $\mathbf{x}(t)$ .

Let  $J^*$  be the optimal (lowest possible) cost. In other words:

$$J^*(\mathbf{x}_0) = \inf_{\pi} J(\mathbf{x}_0, \pi(\mathbf{x}, t)) \quad (4)$$

Optimal cost is attained when optimal policy is attained:

$$\pi = \pi^*(\mathbf{x}, t)$$

# HAMILTON-JACOBI-BELLMAN EQUATION

With this, we can formulate *Hamilton-Jacobi-Bellman equation* (HJB):

$$\min_{\mathbf{u}} \left[ g(\mathbf{x}, \mathbf{u}) + \frac{\partial J^*}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}) \right] = 0 \quad (5)$$

We can find control that delivers minimum to the function (5):

$$\mathbf{u}^* = \operatorname{argmin}_{\mathbf{u}} \left[ g(\mathbf{x}, \mathbf{u}) + \frac{\partial J^*}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}) \right] \quad (6)$$

For LTI, dynamics is:

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \quad (7)$$

We can choose quadratic cost:

$$g(\mathbf{x}, \mathbf{u}) = \mathbf{x}^\top \mathbf{Qx} + \mathbf{u}^\top \mathbf{Ru} \quad (8)$$

Then HJB becomes:

$$\min_{\mathbf{u}} [\mathbf{x}^\top \mathbf{Qx} + \mathbf{u}^\top \mathbf{Ru} + \frac{\partial J^*}{\partial \mathbf{x}} (\mathbf{Ax} + \mathbf{Bu})] = 0 \quad (9)$$

where  $\mathbf{Q} = \mathbf{Q}^\top \geq 0$  and  $\mathbf{R} = \mathbf{R}^\top > 0$ .

## ALGEBRAIC RICCATI (LTI CASE), 2

There is a theorem that says that for LTI with quadratic cost,  $J^*$  has the form:

$$J^* = \mathbf{x}^\top \mathbf{S} \mathbf{x} \quad (10)$$

where  $\mathbf{S} = \mathbf{S}^\top \geq 0$ .

Then HJB becomes:

$$\min_{\mathbf{u}} \left[ \mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{u}^\top \mathbf{R} \mathbf{u} + \mathbf{x}^\top \mathbf{S} (\mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u}) + (\mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u})^\top \mathbf{S} \mathbf{x} \right] = 0$$

Simplifying, we get:

$$\min_{\mathbf{u}} \left[ \mathbf{u}^\top \mathbf{R} \mathbf{u} + \mathbf{x}^\top (\mathbf{Q} + \mathbf{S} \mathbf{A} + \mathbf{A}^\top \mathbf{S}) \mathbf{x} + \mathbf{x}^\top \mathbf{S} \mathbf{B} \mathbf{u} + \mathbf{u}^\top \mathbf{B}^\top \mathbf{S} \mathbf{x} \right] = 0$$

Finding partial derivative of the HJB with respect to  $\mathbf{u}$  and setting it to zero (as it is an extreme point) we get:

$$2\mathbf{u}^\top \mathbf{R} + 2\mathbf{x}^\top \mathbf{S}\mathbf{B} = 0 \quad (11)$$

This expression can be transposed and  $\mathbf{u}$  separated:

$$\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S}\mathbf{x} \quad (12)$$

This is the desired control law. We can see that it is *proportional*. We can re-write it as:

$$\mathbf{u} = -\mathbf{K}\mathbf{x} \quad (13)$$

where  $\mathbf{K} = \mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S}$  is the controller gain. This control law is called Linear Quadratic Regulator (LQR).



Substituting found control law into the HJB, we find:

$$\min_{\mathbf{u}} [\mathbf{x}^\top (\mathbf{Q} + \mathbf{S}\mathbf{A} + \mathbf{A}^\top \mathbf{S})\mathbf{x} + \mathbf{x}^\top \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{R}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S}\mathbf{x} - \mathbf{x}^\top \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S}\mathbf{x} - \mathbf{x}^\top \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S}\mathbf{x}] = 0 \quad (14)$$

Simplifying, we get:

$$\mathbf{x}^\top (\mathbf{Q} + \mathbf{S}\mathbf{A} + \mathbf{A}^\top \mathbf{S} - \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S})\mathbf{x} = 0 \quad (15)$$

which would hold for all  $\mathbf{x}$  iff:

$$\mathbf{Q} - \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{S} + \mathbf{S}\mathbf{A} + \mathbf{A}^\top \mathbf{S} = 0 \quad (16)$$

This is the *Algebraic Riccati equation*.

There are a number of ways to solve LQR:

- In MATLAB there is a function  $[K,S,P] = \text{lqr}(A,B,Q,R)$ , where  $P = \text{eig}(A-B*K)$
- In Python, there is  $S = \text{scipy.linalg.solve\_continuous\_are}(A,B,Q,R)$

# LQR AND POLE PLACEMENT

- Pole placement **upsides**: allows to design exactly how fast the control error decays to zero; allows to design control error oscillations.
- Pole placement **downsides**: may require unreasonably high control gains. Easy to ask for "unreasonable" performance.
- LQR **upsides**: easy to produce "reasonable" control gains.
- LQR **downsides**: may produce very slow decaying control error with oscillations.

Consider discrete dynamics:

$$\mathbf{x}_{i+1} = \mathbf{A}\mathbf{x}_i + \mathbf{B}\mathbf{u}_i \quad (17)$$

with a cost function:

$$J = \sum_{i=0}^N (\mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i + \mathbf{u}_i^\top \mathbf{R} \mathbf{u}_i) \quad (18)$$

Let us find the optimal control policy for this case.

Let us consider a linear control law  $\mathbf{u}_i = -\mathbf{K}_i \mathbf{x}_i$ . Given initial conditions  $\mathbf{x}_0$  we find  $\mathbf{x}_1$  as:

$$\mathbf{x}_1 = (\mathbf{A} - \mathbf{B}\mathbf{K}_i)\mathbf{x}_0 \quad (19)$$

let us define components of the cost associated with each iteration as  $J_1, J_2$ , etc:

$$J_i = \mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i + \mathbf{u}_i^\top \mathbf{R} \mathbf{u}_i \quad (20)$$

Since  $\mathbf{u}_0 = -\mathbf{K}_0 \mathbf{x}_0$  we can show that  $J_1$  is a quadratic form of initial conditions:

$$J_0 = \mathbf{x}_0^\top (\mathbf{Q} + \mathbf{K}_i^\top \mathbf{R} \mathbf{K}_i) \mathbf{x}_0 \quad (21)$$

Similarly, we can show that all  $J_i$  can be written as a quadratic form of initial conditions.

Let us define cost-to-go as optimal cost for given initial conditions:

$$V_0 = \min_{\mathbf{u}} \sum_{i=0}^N (\mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i + \mathbf{u}_i^\top \mathbf{R} \mathbf{u}_i) \quad (22)$$

If  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots$  is a sequence of states that form an optimal trajectory, let us define the cost-to-go starting from each of these states as:

$$V_i = \min_{\mathbf{u}} \sum_{k=i}^N (\mathbf{x}_k^\top \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^\top \mathbf{R} \mathbf{u}_k) \quad (23)$$

We can note that the optimal cost will take a form of a quadratic function:

$$V_i = \mathbf{x}_i^\top \mathbf{P}_i \mathbf{x}_i \quad (24)$$

Defining  $\mathbf{x}_0 = \mathbf{z}$  and  $\mathbf{u}_0 = \mathbf{w}$  we can write cost-to-go as:

$$V_0(\mathbf{z}) = \min_{\mathbf{w}} (\mathbf{z}^\top \mathbf{Q} \mathbf{z} + \mathbf{w}^\top \mathbf{R} \mathbf{w} + V_1(\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})) \quad (25)$$

where  $V_1(\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})$  is the optimal cost-to-go on the next step.

As the next step is closer to the goal, the optimal cost-to-go on the next step is both smaller than on the current step, and is contained in it.

The equation (25) is called *Bellman* equation.

Since  $V_0 = \mathbf{x}_0^\top \mathbf{P}_0 \mathbf{x}_0$  and  $V_1 = \mathbf{x}_1^\top \mathbf{P}_1 \mathbf{x}_1$  we can re-write Bellman equation as:

$$\mathbf{z}^\top \mathbf{P}_0 \mathbf{z} = \min_{\mathbf{w}} (\mathbf{z}^\top \mathbf{Q} \mathbf{z} + \mathbf{w}^\top \mathbf{R} \mathbf{w} + (\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})^\top \mathbf{P}_1 (\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})) \quad (26)$$

To find minimum over  $\mathbf{w}$  we take partial derivative and set it to zero:

$$\frac{\partial}{\partial \mathbf{w}} \left( \mathbf{z}^\top \mathbf{Q} \mathbf{z} + \mathbf{w}^\top \mathbf{R} \mathbf{w} + (\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})^\top \mathbf{P}_1 (\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w}) \right) = 0 \quad (27)$$

$$2\mathbf{w}^\top \mathbf{R} + 2(\mathbf{A} \mathbf{z} + \mathbf{B} \mathbf{w})^\top \mathbf{P}_1 \mathbf{B} = 0 \quad (28)$$

$$\mathbf{R} \mathbf{w} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{B} \mathbf{w} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{A} \mathbf{z} = 0 \quad (29)$$



From  $\mathbf{R}\mathbf{w} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{B}\mathbf{w} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{A}\mathbf{z} = 0$ , we can find expression for the optimal control law:

$$(\mathbf{R} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{B})\mathbf{w} = -\mathbf{B}^\top \mathbf{P}_1 \mathbf{A}\mathbf{z} \quad (30)$$

$$\mathbf{w} = -(\mathbf{R} + \mathbf{B}^\top \mathbf{P}_1 \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{P}_1 \mathbf{A}\mathbf{z} \quad (31)$$

In general, given optimal cost-to-go matrix  $\mathbf{P}_{i+1}$  we find optimal control law:

$$\mathbf{u}_i = -(\mathbf{R} + \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{A}\mathbf{x}_i \quad (32)$$

# BACK-PROPAGATION, 1

Let us define  $\mathbf{M} = (\mathbf{R} + \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{B})^{-1}$  and  $\mathbf{N} = \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{A}$  we can re-write the control law:

$$\mathbf{u}_i = -\mathbf{M}\mathbf{N}\mathbf{x}_i \quad (33)$$

We can substitute the control law into the Bellman eq.:

$$\mathbf{x}_i^\top \mathbf{P}_i \mathbf{x}_i = \min_{\mathbf{u}} (\mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i + \mathbf{u}_i^\top \mathbf{R} \mathbf{u}_i + (\mathbf{A} \mathbf{x}_i + \mathbf{B} \mathbf{u}_i)^\top \mathbf{P}_{i+1} (\mathbf{A} \mathbf{x}_i + \mathbf{B} \mathbf{u}_i))$$

$$\begin{aligned} \mathbf{x}_i^\top \mathbf{P}_i \mathbf{x}_i &= \mathbf{x}_i^\top \mathbf{Q} \mathbf{x}_i + \mathbf{x}_i^\top \mathbf{N}^\top \mathbf{M} \mathbf{R} \mathbf{M} \mathbf{N} \mathbf{x}_i + \\ &+ (\mathbf{A} \mathbf{x}_i - \mathbf{B} \mathbf{M} \mathbf{N} \mathbf{x}_i)^\top \mathbf{P}_{i+1} (\mathbf{A} \mathbf{x}_i - \mathbf{B} \mathbf{M} \mathbf{N} \mathbf{x}_i) \end{aligned}$$

$$\begin{aligned} \mathbf{P}_i &= \mathbf{Q} + \mathbf{N}^\top \mathbf{M} \mathbf{R} \mathbf{M} \mathbf{N} + \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{A} - \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{B} \mathbf{M} \mathbf{N} - \\ &\quad \mathbf{N}^\top \mathbf{M} \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{A} + \mathbf{N}^\top \mathbf{M} \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{B} \mathbf{M} \mathbf{N} \end{aligned}$$

$$\mathbf{P}_i = \mathbf{Q} + \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{A} + \mathbf{N}^\top \mathbf{M} (\mathbf{R} + \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{B}) \mathbf{M} \mathbf{N} - \mathbf{N}^\top \mathbf{M} \mathbf{N} - \mathbf{N}^\top \mathbf{M} \mathbf{N}$$

$$\mathbf{P}_i = \mathbf{Q} + \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{A} - \mathbf{N}^\top \mathbf{M} \mathbf{N}$$

From  $\mathbf{P}_i = \mathbf{Q} + \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{A} - \mathbf{N}^\top \mathbf{M} \mathbf{N}$  we obtain the final result:

$$\mathbf{P}_i = \mathbf{Q} + \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{A} - \mathbf{A}^\top \mathbf{P}_{i+1} \mathbf{B} (\mathbf{R} + \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{P}_{i+1} \mathbf{A}$$

This equation can be used to compute  $\mathbf{P}_i$  from known  $\mathbf{P}_{i+1}$ .

- Underactuated robotics. Linear Quadratic Regulators.
- Discrete LQR. Stanford, EE363.
- Discrete LQR (infinite horizon). Stanford, EE363.

Lecture slides are available via Github:

[github.com/SergeiSa/Control-Theory-2024](https://github.com/SergeiSa/Control-Theory-2024)



# Appendix A: Illustration of HJB

Consider the additive cost  $J(\mathbf{x}_0, \pi(\mathbf{x})) = \int_0^\infty g(\mathbf{x}, \mathbf{u}) dt$ , where  $\mathbf{u} = \pi(\mathbf{x})$  is a control policy. The function  $g(\mathbf{x}, \mathbf{u}) \geq 0$  can be interpreted as a rate of change of cost.

Let  $\pi^*(\mathbf{x})$  be the optimal control policy. Applying the optimal policy to the dynamics  $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u})$  we obtain optimal dynamics:

$$\dot{\mathbf{x}} = f^*(\mathbf{x}) = f(\mathbf{x}, \pi^*(\mathbf{x})) \quad (34)$$

Given initial conditions  $\mathbf{x}_0 = \mathbf{z}$  we generate an optimal trajectory  $\mathbf{x}^* = \mathbf{x}^*(t, \mathbf{z})$ . Given optimal trajectory and optimal control policy we find optimal cost:

$$J^*(\mathbf{z}) = J(\mathbf{z}, \pi^*(\mathbf{x})) \quad (35)$$

Equivalently, we find optimal instantaneous cost:

$$g^*(\mathbf{x}) = g(\mathbf{x}, \pi^*(\mathbf{x})) \quad (36)$$

Since optimal cost depends on initial conditions only, we can find a function  $J^* = J^*(\mathbf{z})$  defined over  $\mathbb{R}^n$ .

Lets us consider a trajectory  $\mathbf{x}^* = \mathbf{x}^*(t, \mathbf{z})$  and sequential points on this trajectory  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2$ , etc. associated with the time points  $t_0, t_1, t_2$ , etc. We can define incurred cost (incurred while moving from the initial state  $\mathbf{z} = \mathbf{x}_0$  to the given point) for each of these points:  $S_0, S_1, S_2$ , etc.:

$$S_i = \int_0^{t_i} g^*(\mathbf{x}) dt \quad (37)$$

Since  $g^*(\mathbf{x}) \geq 0$ , we observe that  $S_0 \leq S_1 \leq S_2 \leq \dots$ . Moving along a trajectory  $\mathbf{x}^*(t, \mathbf{z})$  we incur monotonically increasing cost. We can describe it as a time function  $S(t)$ . The rate of increase of this function is given by instantaneous cost  $g^*(\mathbf{x})$ .



We know that the optimal cost from the point  $\mathbf{z}$  is given as  $J^*(\mathbf{z})$ . For each sequential point on the trajectory we can define *cost-to-go*  $V_i$  as a difference between the optimal cost and the incurred cost:

$$V_i = J^*(\mathbf{z}) - S_i \quad (38)$$

For a given trajectory, we can describe cost-to-go as a time function  $V(t) = J^*(\mathbf{z}) - S(t)$ . Where as  $S(t)$  is monotonically increasing, the function  $V(t)$  is monotonically decreasing, with a rate of change given as  $-g^*(\mathbf{x})$ .

Note that the cost-to-go can be equivalently found as:

$$V(t) = J^*(\mathbf{x}^*(t)) \quad (39)$$

since the optimal cost we incur by starting from the point  $\mathbf{x}_i$  is equivalent to the cost "have left to incur" when we reach  $\mathbf{x}_i$  from  $\mathbf{x}_0$ .

With that, we can make an observation: for an optimal policy we see the rate of change of the cost-to-go function equal to  $-g^*(\mathbf{x})$ . But this rate of change can be found by taking a derivative of  $J^*(\mathbf{x})$  with respect to the vector field  $\dot{\mathbf{x}} = f^*(\mathbf{x})$ :

$$-g^*(\mathbf{x}) = \frac{\partial J^*}{\partial \mathbf{x}} f^*(\mathbf{x}) \quad (40)$$

For sub-optimal control policies, the incurred cost will outpace the decrease of the cost-to-go:

$$g(\mathbf{x}, \mathbf{u}) + \frac{\partial J^*}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u}) \geq 0 \quad (41)$$

Optimal policy recovers the sought equality:

$$\min_{\mathbf{u}} \left[ g(\mathbf{x}, \mathbf{u}) + \frac{\partial J^*}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}) \right] = 0 \quad (42)$$