Fundamentals of Optimal Control

Lecture “Introduction to Control and Machine Learning”

Zhengping Ji, Enrique Zuazua

05.03.2026

Friedrich-Alexander-Universität Erlangen-Nürnberg

Motivation

In applications, we need to find a control signal for a dynamical system over a period of time such that an objective function is optimized.

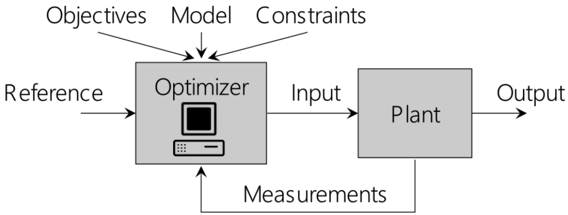

Figure 1: Optimal control system

Key character: infinite-dimensional function space; dynamic constraints.

What is Optimal Control

For a dynamical system, optimal control does not only find “a” solution, it finds the *best* solution.

Defining the optimality

"Best" is defined by minimizing a **cost function** $J(u, x)$.

Real-World Analogy: Driving a Car

- **State (x):** Your car's position and velocity.
- **Control (u):** The gas pedal and steering wheel.
- **Goal:** Get to your destination (regulate state to a target).
- **Cost (J):** Minimize fuel used (control effort) and time taken (state error).

Optimal control finds the *exact* steering and gas inputs to minimize this cost.

Optimal Control Problem: Mathematical Formulation

A typical (finite-horizon) optimal control problem can be formulated as

$$\min_u \int_0^T \ell\big(x(t), u(t)\big)\,\mathrm{d}t + g(x(T)) \tag{1}$$

$$\text{subject to} \quad \dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0 \in \mathbb{R}^n \tag{2}$$

where $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is called the *running cost*, $g : \mathbb{R}^n \to \mathbb{R}$ is called the *terminal cost*, $u \in L^2([0,T], \mathbb{R}^m)$, f is a complete vector field in $\mathbb{R}^n$.

When $T \to +\infty$, (1)(2) is called *infinite-horizon* optimal control problem.

When $x(t)$ in (2) is a linear system and ℓ in (1) is a quadratic function, it is called a **linear quadratic regulator** *(LQR) problem.*

Necessary and Sufficient Conditions

- Pontryagin's Maximum Principle (1950s-1960s): Formulated by Lev Pontryagin and collaborators, it provided **necessary conditions** for optimality by introducing **adjoint variable** and **Hamiltonian dynamics**.

- Dynamical Programming (1960s): Richard Bellman introduced a recursive optimization framework based on the **principle of optimality**, leading to the **Hamilton–Jacobi–Bellman (HJB)** equation, a nonlinear PDE describing the value function.

Outline

Mathematical formulation and solution of the LQR:

1. Review of linear control systems: **Controllability**
2. The **Hamilton-Jacobi-Bellman (HJB)** equation: A continuous-time form of the dynamic programming equation
3. The **Algebraic Riccati Equation (ARE)**: A solution to the HJB for the LQR

Review of Linear Control Systems

Review: Finite-dimensional linear control

Let $n, m \in \mathbb{N}^*$ and $T > 0$. Consider the following finite dimensional system:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \quad t \in (0,T), \\ x(0) = x^0. \end{cases} \tag{3}$$

In (3), A is a real $n \times n$ matrix, B is a real $n \times m$ matrix and x^0 a vector in $\mathbb{R}^n$. $x : [0,T] \to \mathbb{R}^n$ represents the *state* and $u : [0,T] \to \mathbb{R}^m$ the *control*. Both are vector functions of n and m components respectively depending exclusively on time t.

In practice usually $m \leqslant n$. The most desirable goal is, of course, controlling the system by means of a minimum number m of controls.

Review: Controllability

Given an initial datum $x^0 \in \mathbb{R}^n$ and a vector function $u \in L^2(0, T; \mathbb{R}^m)$, system (3) has a unique solution $x \in H^1(0, T; \mathbb{R}^n)$ characterized by the variation of constants formula:

$$x(t) = e^{At}x^0 + \int_0^t e^{A(t-s)}Bu(s)ds, \quad \forall t \in [0, T]. \tag{4}$$

Definition

System (3) is **exactly controllable** in time $T > 0$ if given any initial and final one $x^0, x^1 \in \mathbb{R}^n$ there exists $u \in L^2(0, T, \mathbb{R}^m)$ such that the solution of (3) satisfies $x(T) = x^1$.

According to this definition the aim of the control process consists in driving the solution x of (3) from the initial state x^0 to the final one x^1 in time T by acting on the system through the control u.

Let A^* be the adjoint matrix of A, i.e. the matrix with the property that $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x, y \in \mathbb{R}^n$. Consider the following homogeneous adjoint system of (3):

$$\begin{cases} -\varphi' = A^*\varphi, & t \in (0,T) \\ \varphi(T) = \varphi_T. \end{cases} \tag{5}$$

The following is an equivalent condition for exact controllability .

Lemma

An initial datum $x^0 \in \mathbb{R}^n$ of (3) is driven to zero in time T by using a control $u \in L^2(0,T)$ if and only if

$$\int_0^T \langle u, B^*\varphi\rangle \mathrm{d}t + \langle x^0, \varphi(0)\rangle = 0, \quad \forall\varphi. \tag{6}$$

Proof:

Let φ_T be arbitrary in $\mathbb{R}^n$ and φ the corresponding solution of (5). By multiplying (3) by φ and (5) by x we deduce that

$$\langle x', \varphi\rangle = \langle Ax, \varphi\rangle + \langle Bu, \varphi\rangle; \quad -\langle x, \varphi'\rangle = \langle A^*\varphi, x\rangle.$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x, \varphi\rangle = \langle Bu, \varphi\rangle$$

which, after integration in time, gives that

$$\langle x(T), \varphi_T\rangle - \langle x^0,\, \varphi(0)\rangle = \int_0^T \langle Bu, \varphi\rangle dt = \int_0^T \langle u, B^*\varphi\rangle \mathrm{d}t. \qquad (7)$$

We obtain that $x(T) = 0$ if and only if (6) is verified for any $\varphi_T \in \mathbb{R}^n$.

Identity (6) is in fact an optimality condition for the critical points of the quadratic functional $J : \mathbb{R}^n \to \mathbb{R}$,

$$J(\varphi_T) = \frac{1}{2}\int_0^T \mid B^*\varphi \mid^2 \mathrm{d}t + \langle x^0, \varphi(0)\rangle$$

where φ is the solution of the adjoint system (5) with initial data φ_T at time $t = T$.

More precisely:

Lemma

Suppose that J has a minimizer $\widehat{\varphi}_T \in \mathbb{R}^n$ and let $\widehat{\varphi}$ be the solution of the adjoint system (5) with initial data $\widehat{\varphi}_T$. Then

$$u = B^*\widehat{\varphi} \tag{8}$$

is a control of system (3) with initial data x^0.

If $\widehat{\varphi}_T$ is a point where J achieves its minimum value, then

$$\lim_{h\to 0} \frac{J\left(\widehat{\varphi}_T + h\varphi_T\right) - J\left(\widehat{\varphi}_T\right)}{h} = 0, \quad \forall \varphi_T \in \mathbb{R}^n.$$

This is equivalent to

$$\int_0^T \langle B^*\widehat{\varphi}, B^*\varphi\rangle dt + \langle x^0, \varphi(0)\rangle = 0, \quad \forall \varphi_T \in \mathbb{R}^n,$$

which, in view of Lemma 2, implies that $u = B^*\widehat{\varphi}$ is a control for (3).

Observability

But minimizing the functional J requires its coercivity.

Definition

System (5) is said to be **observable** in time $T > 0$ if there exists $c > 0$ such that

$$\int_0^T \mid B^* \varphi \mid^2 dt \geq c \mid \varphi(0) \mid^2, \tag{9}$$

for all $\varphi_T \in \mathbb{R}^n$, φ being the corresponding solution of (5).

In the sequel (9) will be called the **observation** or **observability inequality**.

It guarantees that the solution of the adjoint problem at $t = 0$ is uniquely determined by the observed quantity $B^* \varphi(t)$ for $0 < t < T$.

The following remark is important in the context of finite dimensional control.

Proposition

Inequality (9) is equivalent to the following unique continuation principle:

$$B^*\varphi(t) = 0, \ \forall t \in [0,T] \Rightarrow \varphi_T = 0. \tag{10}$$

This is an uniqueness or unique continuation property. Unfortunately the equivalence is not true for infinite-dimensional systems (PDE, distributed parameter systems), as we use that $\partial B(0,1)$ is compact to prove that (10) implies (9).

Kalman's Rank Condition

Question: *Are there algebraic conditions on the state matrix A and the control one B for it to be true?*

The following classical result is due to R. E. Kalman and gives a complete answer to the problem of exact controllability of finite dimensional linear systems.

Theorem

System (3) is exactly controllable in some time T if and only if

$$\textit{rank}\,[B,\, AB, \cdots, A^{n-1}B] = n. \tag{11}$$

Consequently, if system (3) is controllable in some time $T > 0$ it is controllable in any time.

Proof of Theorem 5: " $\Leftarrow$".

Suppose now that rank$([B,\, AB, \cdots, A^{n-1}B]) = n$. It is sufficient to show that system (5) is observable.

Assume $B^*\varphi = 0$ and $\varphi(t) = e^{A^*(T-t)}\varphi_T$, it follows that $B^*e^{A^*(T-t)}\varphi_T \equiv 0$ for all $0 \leq t \leq T$. By computing the derivatives of this function in $t = T$ we obtain that

$$B^*[A^*]^k\varphi_T = 0 \quad \forall k \geq 0.$$

But since rank$(\big[B,\, AB, \cdots, A^{n-1}B\big]) = n$ we deduce that

$$\mathrm{Ker}(\big[B^*,\, B^*A^*, \cdots, B^*(A^*)^{n-1}\big]) = \{0\}$$

and therefore $\varphi_T = 0$. Hence, (10) is verified.

Proof of Theorem 5: " $\Rightarrow$ "

Suppose that rank$([B,\ AB, \cdots, A^{n-1}B]) < n$. Then the kernel of the adjoint is non trivial. Accordingly, there is a non-trivial φ_T such that

$$B^*[A^*]^k \varphi_T = 0 \quad \forall k \geq 0.$$

Let φ be the corresponding solution of the adjoint system.

From Cayley-Hamilton Theorem we deduce that there exist constants $c_1, \cdots, c_n$ such that, $A^n = c_1 A^{n-1} + \cdots + c_n I$. Expanding the exponential in power series it is the easy to see that for all $t \in (0, T)$:

$$B^*\varphi(t) = B^* \exp(A^*(T-t))\varphi_T \equiv 0.$$

Multiplying the state equation by φ it is easy to see that

$$\begin{aligned}\langle x(t), \varphi(t)\rangle =& x(t)^*\varphi(t) = (x^0)^* e^{A^* t}\varphi(t) + \int_0^T u(s)^* B^* e^{A^*(t-s)}\varphi(t) ds \\ =& (x^0)^* e^{A^* t}\varphi(t) + \int_0^T u(s)^* B^* e^{A^*(T-s)}\varphi_T = (x^0)^* e^{A^* t}\varphi(t)\end{aligned}$$

Consequently, the product $\langle x(t), \varphi(t)\rangle$ is invariant of the control, which excludes the controllability property to hold. The proof of Theorem 5 is now complete. $\square$

Remark

The set of controllable pairs (A, B) is open and dense.

- Most systems are controllable;
- The controllability property is robust, i. e. it is invariant under small perturbations of A and/or B.

When controllability holds,

$$\| u \|_{L^2(0,T)} \leq C|e^{AT}x^0 - x^1| \tag{12}$$

for any initial data x^0 and final objective x^1.

Remark

Linear scalar equations of any order provide examples of systems that are controllable with only one control:

$$x^{(k)} + a_1 x^{(k-1)} + \ldots + a_{k-1}x = u.$$

Exercise: Check that the Kalman condition is fulfilled in this case.

Linear Quadratic Regulator

The "Regulator" Problem: Intuition

What is the Goal?

The core goal of a "regulator" is to **stabilize** a system by driving its state $x(t)$ to a desired setpoint (defined as zero) and keeping it there in the presence of disturbances.

Example 1: Inverted Pendulum

- **State (x):** Angle and angular velocity.
- **Goal:** Keep the pendulum upright (angle $= 0$).
- **Control (u):** Force applied to the cart.

Example 2: Satellite Attitude

- **State (x):** Roll, pitch, yaw angles.
- **Goal:** Stop tumbling and point at a target.
- **Control (u):** Torque from reaction wheels.

Goal: to return to zero optimally.

The "L": Linear System Dynamics

The LQR problem assumes the system's dynamics to be linear and time-invariant (LTI).

State-Space Equations

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \\ t &\geqslant 0, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{aligned} \tag{13}$$

Where:

- $\boldsymbol{x}(t) \in \mathbb{R}^n$: the state vector (e.g., [position, velocity]).
- $\boldsymbol{u}(t) \in \mathbb{R}^m$: the control input vector (e.g., [force, torque]).
- $\boldsymbol{A} \in \mathbb{R}^{n \times n}$: the *system matrix* (how the state evolves on its own).
- $\boldsymbol{B} \in \mathbb{R}^{n \times m}$: the *input matrix* (how control affects the state).

To guarantee the existence of the optimal control, we need to assume that (13) *is controllable, namely* $(\boldsymbol{A}, \boldsymbol{B})$ *forms a controllable pair.*

The "Q": Quadratic Cost Function

To define "optimal", we create a *quadratic* cost function J that we want to minimize over an infinite time horizon.

The Cost Functional

$$J = \int_0^\infty \left(\boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R} \boldsymbol{u} \right) dt \tag{14}$$

- $\boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x}$: **Cost of being away from zero.**
  - $\boldsymbol{Q} \succeq 0$ (positive semi-definite) is the state-weighting matrix.
- $\boldsymbol{u}^\top \boldsymbol{R} \boldsymbol{u}$: **Cost of control effort.**
  - $\boldsymbol{R} \succ 0$ (positive definite) is the control-weighting matrix.

The LQR Problem: Find the control law $\boldsymbol{u}(t)$ that minimizes J.

Intuition: the Q vs. R Trade-off

The "art" of LQR is choosing the $\boldsymbol{Q}$ and $\boldsymbol{R}$ matrices. This is a trade-off:

High Q (Penalize State Error)

- $\boldsymbol{Q} \gg \boldsymbol{R}$
- Forces $\boldsymbol{x}(t) \to 0$ very quickly.
- Leads to a very **aggressive** controller.
- Uses a lot of control effort (high $\boldsymbol{u}$), which may be expensive or infeasible.

High R (Penalize Control Effort)

- $\boldsymbol{R} \gg \boldsymbol{Q}$
- Uses minimal control effort $\boldsymbol{u}(t)$.
- Leads to a very **conservative** (slow) controller.
- $\boldsymbol{x}(t)$ will drift to 0 slowly.

Bellman's Principle of Optimality

The Goal

Consider the optimal control problem

$$\min_{\boldsymbol{u}} \int_0^{\infty} \left(\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u}\right) dt$$
$$\text{s.t. } \dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \quad t \geqslant 0, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

How to find a function $\boldsymbol{u}(t)$ for all $t \in [0, \infty)$ that minimizes the *entire* future cost?

The Principle

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

This allows us to break a massive, infinite problem into a series of smaller, sequential decisions.

The Value Function V(x)

Define a **Value Function**, $V(\boldsymbol{x}(t))$, as the optimal "cost-to-go" from state $\boldsymbol{x}$ at time t.

The Value Function

$$V(\boldsymbol{x}(t)) = \min_{\boldsymbol{u}} \int_t^\infty \left(\boldsymbol{x}(\tau)^\top \boldsymbol{Q}\boldsymbol{x}(\tau) + \boldsymbol{u}(\tau)^\top \boldsymbol{R}\boldsymbol{u}(\tau)\right) d\tau$$

What does $V(\boldsymbol{x})$ represent?

- $V(\boldsymbol{x})$ is the cost to accumulate *from now until forever*, assuming we act optimally from now on.
- If $V(\boldsymbol{x})$ is high, we are in a "bad" state (far from zero).
- $V(\mathbf{0}) = 0$: we are at the goal, and the future cost is zero.
- The goal is to find the control $\boldsymbol{u}$ that minimizes $V(\boldsymbol{x})$.

Dynamic Programming Intuition

By the Bellman's principle, the total optimal cost $V(\boldsymbol{x}(t))$ can be split into two parts:

1. The **immediate cost** incurred over a tiny time step Δt.
2. The **optimal future cost** from the *new* state $\boldsymbol{x}(t+\Delta t)$.

The DP Equation (Discrete-Time View)

$$V(\boldsymbol{x}(t)) = \min_{\boldsymbol{u}(t)} \Bigg[\underbrace{\left(\boldsymbol{x}(t)^\top \boldsymbol{Q}\boldsymbol{x}(t) + \boldsymbol{u}(t)^\top \boldsymbol{R}\boldsymbol{u}(t)\right)\Delta t}_{\text{Cost from } t \text{ to } t+\Delta t} + \underbrace{V(\boldsymbol{x}(t+\Delta t))}_{\text{Optimal cost from } t+\Delta t \text{ onwards}} \Bigg]$$

By extending it to continuous-time, one can derive the Hamilton-Jacobi-Bellman (HJB) equation.

From DP to the HJB Equation

Start with our DP equation:
$V(\boldsymbol{x}(t)) = \min_{\boldsymbol{u}}[(\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u})\Delta t + V(\boldsymbol{x}(t+\Delta t))]$

Use a Taylor expansion for $V(\boldsymbol{x}(t+\Delta t))$:
$V(\boldsymbol{x}(t+\Delta t)) \approx V(\boldsymbol{x}(t)) + \nabla V(\boldsymbol{x})^\top \dot{\boldsymbol{x}}\Delta t$

Substitute this back into the DP equation:
$V(\boldsymbol{x}) \approx \min_{\boldsymbol{u}}[(\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u})\Delta t + V(\boldsymbol{x}) + \nabla V^\top \dot{\boldsymbol{x}}\Delta t]$

Subtract $V(\boldsymbol{x})$, divide by Δt, and let $\Delta t \to 0$:
$0 = \min_{\boldsymbol{u}}[\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u} + \nabla V^\top \dot{\boldsymbol{x}}]$

Substitute $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}$ to get the HJB equation:

The Hamilton-Jacobi-Bellman (HJB) Equation

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

The *Hamiltonian function*:
$\mathcal{H}(\boldsymbol{x}, \boldsymbol{u}, \nabla V) = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u} + \nabla V(\boldsymbol{x})^\top (\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u})$

Solving the HJB (1): Minimizing u

The HJB gives a minimization problem at each t. $\min_{\boldsymbol{u}} [\mathcal{H}(\boldsymbol{x}, \boldsymbol{u}, \nabla V)] = 0$

Let $\mathcal{H}$ be the Hamiltonian. To find the minimum, take the partial derivative w.r.t. $\boldsymbol{u}$ and set it to zero:

$$\frac{\partial \mathcal{H}}{\partial \boldsymbol{u}} = \frac{\partial}{\partial \boldsymbol{u}} \left(\boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R} \boldsymbol{u} + \nabla V^\top \boldsymbol{A} \boldsymbol{x} + \nabla V^\top \boldsymbol{B} \boldsymbol{u} \right) = 0$$

The only terms that depend on $\boldsymbol{u}$ are $\boldsymbol{u}^\top \boldsymbol{R} \boldsymbol{u}$ and $\nabla V^\top \boldsymbol{B} \boldsymbol{u}$:

$$\frac{\partial \mathcal{H}}{\partial \boldsymbol{u}} = 2\boldsymbol{R}\boldsymbol{u} + \boldsymbol{B}^\top \nabla V = 0$$

This gives us the **optimal control policy** $\boldsymbol{u}^*$ in terms of ∇V:

Optimal Control Policy

$$\boldsymbol{u}^*(\boldsymbol{x}) = -\frac{1}{2} \boldsymbol{R}^{-1} \boldsymbol{B}^\top \nabla V(\boldsymbol{x})$$

Solving the HJB (2): Guessing the Value Function

Problem: there is no information about $V(\boldsymbol{x})$.

Let's make an educated guess. Since the *cost* J is quadratic in $\boldsymbol{x}$ and $\boldsymbol{u}$, it is reasonable to assume the *value* $V(\boldsymbol{x})$ is also quadratic in $\boldsymbol{x}$.

The quadratic guess

Assume $V(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{P} \boldsymbol{x}$, where $\boldsymbol{P}$ is some symmetric, positive-definite matrix.

Why this form?

- If $\boldsymbol{x} = \boldsymbol{0}$, then $V(\boldsymbol{0}) = 0$, which is correct (zero state, zero cost-to-go).
- The gradient is simple and linear in $\boldsymbol{x}$:

Gradient of the quadratic value function

$$\nabla V(\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^\top \boldsymbol{P} \boldsymbol{x}) = 2\boldsymbol{P}\boldsymbol{x}$$

Solving the HJB (3): Finding the Optimal Control

Now we combine our findings.

(1) $\boldsymbol{u}^* = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \nabla V(\boldsymbol{x})$; (2) $\nabla V(\boldsymbol{x}) = 2\boldsymbol{P}\boldsymbol{x}$

Substituting (2) into (1) we have

The Optimal LQR Control Law

$$\boldsymbol{u}^*(\boldsymbol{x}) = -(\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P})\boldsymbol{x} = -\boldsymbol{K}\boldsymbol{x}$$

Meaning: the optimal control is a **linear state feedback**.

- The control $\boldsymbol{u}$ is just a constant matrix $\boldsymbol{K}$ times the current state $\boldsymbol{x}$.
- $\boldsymbol{K} = \boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}$ is the **optimal feedback gain**.
- All we need is to find the matrix $\boldsymbol{P}$.

Deriving the Riccati Equation

We find $\boldsymbol{P}$ by plugging everything back into the HJB equation:
$0 = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + (\boldsymbol{u}^*)^\top \boldsymbol{R}\boldsymbol{u}^* + (2\boldsymbol{P}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}^*)$

Substitute $\boldsymbol{u}^* = -\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\boldsymbol{x}$:

$$\begin{aligned} 0 = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + &\ (\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1})\boldsymbol{R}(\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\boldsymbol{x}) \\ + &\ 2\boldsymbol{x}^\top \boldsymbol{P}(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}(-\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\boldsymbol{x})) \end{aligned}$$

Simplify the terms:

$$\begin{aligned} 0 = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + &\ \boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\boldsymbol{x} \\ + &\ 2\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\boldsymbol{x} \end{aligned}$$

Combine and use $2\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top(\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A})\boldsymbol{x}$:

$$0 = \boldsymbol{x}^\top \left(\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\right)\boldsymbol{x}$$

The Algebraic Riccati Equation (ARE)

The previous equation must hold for **any** state $\boldsymbol{x}$.

$$\boldsymbol{x}^\top \underbrace{\left(\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}\right)}_{\text{This entire term must be zero}} \boldsymbol{x} = 0$$

This gives the final equation we need to solve for $\boldsymbol{P}$.

The Continuous Algebraic Riccati Equation (ARE)

$$\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P} = 0$$

- This is a (non-linear) matrix equation.
- We solve it (numerically) for the unique symmetric, positive-definite matrix $\boldsymbol{P}$.
- Once $\boldsymbol{P}$ is derived, we have the solution to the entire LQR problem.

The LQR Design Recipe

Step 1: Define the Problem

- Identify system matrices: $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}$
- Choose cost matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ based on performance goals.

Step 2: Solve the ARE

- Use a numerical solver (e.g., `lqr` in MATLAB, Python) to solve the ARE for $\boldsymbol{P}$:
- $\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P} = 0$

Step 3: Calculate the Optimal Gain

- Compute the optimal feedback gain matrix $\boldsymbol{K}$:
- $\boldsymbol{K} = \boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}$

Step 4: Implement the Control Law

- Your optimal controller is: $\boldsymbol{u}(t) = -\boldsymbol{K}\boldsymbol{x}(t)$

Summary & Key Properties

What We Did

1. Defined the LQR problem (Linear System, Quadratic Cost).
2. Used Bellman's Principle (Dynamic Programming) to define a Value Function $V(\boldsymbol{x})$.
3. Derived the HJB equation, the continuous form of Bellman's principle.
4. Assumed $V(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{P} \boldsymbol{x}$ and solved the HJB.
5. This revealed the optimal control is a linear feedback $\boldsymbol{u} = -\boldsymbol{K}\boldsymbol{x}$.
6. The solution $\boldsymbol{P}$ is found from the Algebraic Riccati Equation (ARE).

Summary & Key Properties

Why LQR is Fundamental

- **Optimality:** By definition, it is the "best" controller for the given quadratic cost.
- **Guaranteed Stability:** If the system $(\boldsymbol{A}, \boldsymbol{B})$ is *controllable* and $(\boldsymbol{A}, \boldsymbol{C})$ (where $\boldsymbol{Q} = \boldsymbol{C}^\top \boldsymbol{C}$) is *observable*, the resulting closed-loop system $(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{K})$ is guaranteed to be stable.
- **Robustness:** LQR controllers have good, well-defined robustness margins (e.g., $\geqslant 60^\circ$ phase margin).
- **Further Application:** It is the basis for most modern optimal and robust control (e.g., LQG, H_∞).

Finite-horizon LQR

Consider a **finite-horizon** case of the LQR problem:

$$\min_{\boldsymbol{u}} \int_0^{T} \boldsymbol{x}^\top(t)\boldsymbol{Q}\boldsymbol{x}(t) + \boldsymbol{u}^\top(t)\boldsymbol{R}\boldsymbol{u}(t)\,\mathrm{d}t + \boldsymbol{x}(T)^\top \boldsymbol{W}\boldsymbol{x}(T) \tag{15}$$

$$\text{s. t.} \quad \dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0,\ t \in [0, T] \tag{16}$$

where $\boldsymbol{Q}$, $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ are positive definite matrices.

Question: *How to adapt the Riccati theory to solve the above problem?*

Solving Finite-horizon LQR

The optimal control law $\boldsymbol{u}^*(t)$ is again a linear state feedback:

Optimal Control Law

$$\boldsymbol{u}^*(t) = -\boldsymbol{K}(t)\boldsymbol{x}(t) \tag{17}$$

$\boldsymbol{K} : [0, T] \to \mathbb{R}^{m \times n}$ is the *optimal feedback gain matrix*. This is still a closed-loop controller.

The problem is reduced to finding the time-varying matrix $\boldsymbol{K}(t)$.

Finding Feedback K: The DRE

To find $\boldsymbol{K}$, we must first solve for a unique, symmetric, positive-definite matrix $\boldsymbol{P}$, as in the infinite-horizon case.

This $\boldsymbol{P}$ is the solution to the **Differential Riccati Equation (DRE)**.

Differential Riccati Equation (DRE)

$$\dot{\boldsymbol{P}}(t) = \boldsymbol{A}^\top \boldsymbol{P}(t) + \boldsymbol{P}(t)\boldsymbol{A} - \boldsymbol{P}(t)\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}(t) + \boldsymbol{Q} \tag{18}$$

with boundary condition $\boldsymbol{P}(T) = \boldsymbol{W}$.

- This is a matrix quadratic equation (note the $\boldsymbol{P} \ldots \boldsymbol{P}$ term)
- It is solvable when $(\boldsymbol{A}, \boldsymbol{B})$ forms a controllable pair

Finding Feedback K: The Gain

Once the numerical solver finds the unique, positive-definite solution $\boldsymbol{P}$ to the DRE, the optimal gain $\boldsymbol{K}$ is calculated directly.

Optimal Gain Calculation

$$\boldsymbol{K}(t) = \boldsymbol{R}^{-1}\boldsymbol{B}^{\top}\boldsymbol{P}(t) \tag{19}$$

The Final Closed-Loop System

By substituting $\boldsymbol{u}(t) = -\boldsymbol{K}(t)\boldsymbol{x}(t)$ back into the original dynamics:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}(-\boldsymbol{K}(t)\boldsymbol{x}(t)) = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{K}(t))\boldsymbol{x}(t)$$

The stability of the system is now determined by the eigenvalues of $(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{K}(t))$.

Result: *when $t \to \infty$, $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{K}(t)$ is a stable matrix.*

Recap of Key Equations

System: $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}$, with $(\boldsymbol{A}, \boldsymbol{B})$ controllable

Cost: $J = \int_0^\infty (\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u})\mathrm{d}t$ or
$J_T = \int_0^T (\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{R}\boldsymbol{u})\mathrm{d}t + \boldsymbol{x}(T)^\top \boldsymbol{W}\boldsymbol{x}(T)$

ARE: $0 = \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}$
DRE: $\dot{\boldsymbol{P}} = \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P}$, $\boldsymbol{P}(T) = \boldsymbol{W}$

Solution: $\boldsymbol{u} = -(\boldsymbol{R}^{-1}\boldsymbol{B}^\top \boldsymbol{P})\boldsymbol{x} = -\boldsymbol{K}\boldsymbol{x}$