Reading group 4
Optimal Control Theory

Antoine Bambade

Willow and Sierra teams

Contents

01

Introduction

Objectives

- Applying previous theory to LQR problem
- Study in two cases :
  - > finite horizon (chapter 1)
  - > infinite horizon (chapter 2)

02

Chapter 1

Ingredient 1 : a linear time-varying system

$$\begin{cases} \dot{x} = f(t, x, u) \\ x(0) = x_0 \end{cases} \implies \begin{cases} \dot{x} = A(t)x + B(t)u \\ x(0) = x_0 \in \mathbb{R}^n \end{cases}$$

Ingredient 2 : quadratic running and terminal costs

$$J : u \in \mathbb{R}^m \to \mathbb{R}$$

$$J(u) \doteq \int_{t_0}^{t_f} L(x(t), u(t), t)dt + K(t_f, x_f)$$

$$= \int_{t_0}^{t_f} (x^T(t)Q(t)x(t) + u^T R(t)u(t))dt + x^T(t_1)Mx(t_1)$$

Ingredient 3 : a fixed-time, free-endpoint target set

$$S = \{t_1\}x\mathbb{R}^n$$

Main ideas

- Finding an optimal control candidate
  - > Necessary condition (Maximum Principle)
  - > Tractable solution (RDE)
- Sufficient solution (HJB)

Recall of main PMP results

If $u^* : [t_0, t_1] \to \mathbb{R}^m$ is an optimal control, $\exists p^* : [t_0, t_1] \to \mathbb{R}^p$ satisfying with $x^* : [t_0, t_1] \to \mathbb{R}^n$:

- Canonical equation :

$$\begin{cases} \dot{x}^* = H_p(x^*, u^*, p^*) \\ \dot{p}^* = -H_x(x^*, u^*, p^*) \end{cases} \text{with} \begin{cases} x^*(t_0) = x_0 \\ H(x, u, p) := p^T f(x, u) - L(x, u) \end{cases}$$

- Transversality equation (fixed time + terminal cost + free variable endpoint) :

$$p^*(t_1) = -K_x(x^*(t_1))$$

- Global maximality :

$$\forall t \in [t_0, t_1], \forall u \in \mathbb{R}^m, H(x^*(t), u^*(t), p^*(t) \geq H(x^*(t), u, p^*(t))$$

Recall of main PMP results

If $u^* : [t_0, t_1] \to \mathbb{R}^m$ is an optimal control, $\exists p^* : [t_0, t_1] \to \mathbb{R}^p$ satisfying with $x^* : [t_0, t_1] \to \mathbb{R}^n$:

- Canonical equation :

$$\begin{cases} \dot{x}^* = H_p = Ax^* + Bu^* \\ \dot{p}^* = -H_x = -(A^T p^* - 2Qx^*) \end{cases} \begin{cases} x^*(t_0) = x_0 \\ H(x,u,p) := p^T(Ax + Bu) \\ -(x^T Qx + u^T Ru) \end{cases}$$

- Transversality equation :

$$p^*(t_1) = -K_x(x^*(t_1)) = -2Mx^*(t_1)$$

- Global maximality :

$$\forall t \in [t_0, t_1], \forall u \in \mathbb{R}^m, H(x^*(t), u^*(t), p^*(t) \geq H(x^*(t), u, p^*(t))$$

Recall of main PMP results

If u^* is an optimal control, $\exists p^*$ satisfying with x^*:

- Canonical equation :

$$\begin{cases} \dot{x}^* = H_p = Ax^* + Bu^* \\ \dot{p}^* = -H_x = -(A^T p^* - 2Qx^*) \end{cases} \quad \begin{cases} x^*(t_0) = x_0 \\ H(x, u, p) := p^T(Ax + Bu) \\ -(x^T Qx + u^T Ru) \end{cases}$$

- Transversality equation :

$$p^*(t_1) = -2Mx^*(t_1)$$

- Global maximality :

$$\begin{cases} H_u(x^*, u^*, p^*) = 0 \\ H_u(x^*, u^*, p^*) = B^T p^* - 2Ru^* \\ H_{uu}(x^*, u^*, p^*) < 0 \end{cases} \implies \begin{cases} u^* = \frac{1}{2}R^{-1}B^T p^* \\ H_{uu}(x^*, u^*, p^*) = -2R \end{cases}$$

Inria

Recall of main PMP results

If $u^* : [t_0, t_1] \to \mathbb{R}^m$ is an optimal control, $\exists p^* : [t_0, t_1] \to \mathbb{R}^p$ satisfying with $x^* : [t_0, t_1] \to \mathbb{R}^n$:

- Canonical equation :

$$\begin{pmatrix} \dot{x}^* \\ \dot{p}^* \end{pmatrix} = \begin{pmatrix} A & \frac{1}{2}BR^{-1}B^T \\ 2Q & -A^T \end{pmatrix} \begin{pmatrix} x^* \\ p^* \end{pmatrix} \text{ with } x^*(t_0) = x_0$$

$$H(x, u, p) := p^T(Ax + pu) - (x^T Qx + u^T Ru)$$

- Transversality equation :

$$p^*(t_1) = -2Mx^*(t_1)$$

- Global maximality :

$$u^* = \frac{1}{2}R^{-1}B^T p^*$$

What would be ideal in practice in terms of **control** ?

a globally asymptotically stable closed loop :

$$\begin{cases} \dot{x}^* = Kx \\ x^*(t_0) = x_0 \end{cases}$$

Here :

$$\begin{cases} \dot{x}^* = Ax^* + Bu^* \\ u^* = \frac{1}{2}R^{-1}B^T p^* \end{cases}$$

Hence, is it true that $p^* = \alpha P x^*, \alpha \in \mathbb{R}^*$? Yes !

Formal construction :

$$\mathcal{H} := \begin{pmatrix} A & \frac{1}{2}BR^{-1}B^T \\ 2Q & -A^T \end{pmatrix}$$

$$\Phi(t) = e^{(t-t_1)\mathcal{H}} = \begin{pmatrix} \Phi_{11}(t) & \Phi_{12}(t) \\ \Phi_{21}(t) & \Phi_{22}(t) \end{pmatrix}$$

$$\begin{pmatrix} x^*(t) \\ p^*(t) \end{pmatrix} = \Phi(t) \begin{pmatrix} x^*(t_1) \\ p^*(t_1) \end{pmatrix} = \begin{pmatrix} (\Phi_{11}(t) - 2\Phi_{12}(t)M)x^*(t_1) \\ (\Phi_{21}(t) - 2\Phi_{22}(t)M)x^*(t_1) \end{pmatrix}$$

$$P(t) = \frac{1}{\alpha}(\Phi_{21}(t) - 2\Phi_{22}(t)M)(\Phi_{11}(t) - 2\Phi_{12}(t)M)^{-1}$$

$p^*(t_1) = -2Mx^*(t_1) \implies \alpha = -2$ and $P(t_1) = M$

inversibility sketch proof when (A,B) controllable

Note : $\mathcal{X}(t) := (\Phi_{11}(t) - 2\Phi_{12}(t)M)$

- By contradiction, let $s \in [t_0, t_1], x^*(t_1) \neq 0 \in \mathbb{R}^n, x^*(s) = \mathcal{X}(s)x^*(t_1) = 0$
- $\frac{d}{dt}(p^{*T}x^*) = (2Qx^* - A^T p^*)^T x^* + p^{*T}(Ax^* + \frac{1}{2}BR^{-1}Bp^*) = 2x^{*T}Qx^* + \frac{1}{2}p^*B^TR^{-1}B^Tp^* \geq 0$
- Integrating from s to t_1 and using $x^*(s) = 0$, $p^*(t_1) = -2Mx^*(t_1)$:

$$0 = 2x^*(t_1)^T Mx^*(t_1) + \int_s^{t_1} 2x^{*T}Qx^* + \frac{1}{2}p^*B^TR^{-1}B^Tp^* \geq 0$$

- as $M, Q \geq 0$ and $R > 0$, $u^* = \frac{1}{2}R^{-1}B^Tp^* = 0$ on $[s, t_1]$,
- hence $\dot{x}^*(t) = Ax^*(t)$ on $[s, t_1]$, with $x^*(s) = 0$ implies $x^* = 0$ on $[s, t_1]$
- hence $x^*(t_1) = 0$

Practical construction

We want : $p^* = -2Px^*$

$$\dot{p}^* = -2\dot{P}x^* - 2P\dot{x}^*$$

$$2Qx^* - A^T p^* = -2\dot{P}x^* - 2PAx^* - PBR^{-1}B^T p^*$$

$$2Qx^* + 2A^T Px^* = -2\dot{P}x^* - 2PAx^* + 2PBR^{-1}B^T Px^*$$

Riccati differential equation (RDE)

$$\dot{P} = -PA - A^T P - Q + PBR^{-1}B^T P$$

Remark

- RDE $=$ nxn matrix quadratic differential equation
- $\frac{d}{dt}\Phi(t) = \mathcal{H}\Phi(t)$ is a 2nx2n linear matrix differential equation

Is our PMP control candidate globally optimal ?

- The HJB equation says a sufficient condition for optimality is that the optimal control $\hat{u}$ makes the value function V satisfy this equation :

$$-V_t(t,x) = \inf_{u \in \mathbb{R}^m} x^T Q(t) x + u^T R(t) u + V_x(t,x)^T (A(t)x + B(t)u)$$

$$\begin{cases} V(t,x) := \inf_{u \in \mathbb{R}^m} \int_t^{t_1} (x^T Q x + u^T R u) ds + x(t_1)^T M x(t_1) \\ x(t) = x \end{cases}$$

- Hence : $V(t_1, x) = x^T M x$
- Here $\hat{u} = -\frac{1}{2} R^{-1}(t) B^T V_x(t,x)$
- Which implies :

$$-V_t = x^T Q x + v_x^T A x - \frac{1}{4} V_x^T B R^{-1} B^T V_x$$

Is our PMP control candidate globally optimal ?

- $\hat{u} = -\frac{1}{2}R^{-1}(t)B^T V_x(t,x)$
- What is the value of V_x ?
- $V = x^T P x$, because it satisfies :

$$-V_t = x^T Q x + v_x^T A x - \frac{1}{4} V_x^T B R^{-1} B^T V_x$$

  as P satisfies the RDE.
- Hence : $u^* = -R^{-1}(t)B^T P x^* = -\frac{1}{2}R^{-1}(t)B^T V_x(t,x) = \hat{u}$ is optimal

Properties of P

- P is unique,
- P is symmetric,
- $P \geq 0$,
- If M is positive definite, $P > 0$.

Reminder

We have seen :

$$\frac{d}{dt}(p^{*T} x^*) \geq 0 \text{ and } p^* = -2Px^*$$

Hence:

$$-2x^{*T} P x^* \leq -2x^*(t_1)^T M x^*(t_1)$$

Does RDE solution P exists for all $t \leq t_1$?

- Locally yes (quadratic matrix differential equation satisfies local Cauchy Lipschitz assumptions)
- Globally ? yes !
  - > By contradiction, if $\exists \hat{t} < t_1$, P(t) becomes unbounded as $t \to \hat{t}+$:
    - – if an off-diagonal entry P_{ij} becomes unbounded while diagonal entries remains bounded, then $\lim_{t\to\hat{t}+} \begin{vmatrix} * & P_{ij}(t) \\ P_{ij}(t) & * \end{vmatrix} = -\infty$
    - – if $P_{ii}(t)$ becomes unbounded. From state $e_i := (0, ..., 1, ..., 0)$ at $t > \hat{t}$, $e_i^T P(t) e_i$ is the optimal cost-to-go from t to t_1. But the zero control on $[t, t_1]$ has a finite cost-to-go associated cost.

$$\int_t^{t_1} e_i^T \Phi_A Q \Phi_A e_i ds + e_i^T \Phi_A^T(t_1) M \Phi_A(t_1) e_i < +\infty$$

Inria

03

Chapter 2

What happens when $t_1 \to +\infty$ for $P(t_0, t_1)$?

- **Monotonicity**: if $t_2 > t_1$, then :

$$x_0^T P(t_0, t_2) x_0 = V^{t_2}(t_0, x_0)$$

$$= \int_{t_0}^{t_2} ((x_{t_2}^*)^T(t) Q (x_{t_2}^*)(t) + (u_{t_2}^*)(t) R (u_{t_2}^*)(t)) dt$$

$$\geq \int_{t_0}^{t_1} ((x_{t_2}^*)^T(t) Q (x_{t_2}^*)(t) + (u_{t_2}^*)(t) R (u_{t_2}^*)(t)) dt$$

$$\geq V^{t_1}(t_0, x_0) = x_0^T P(t_0, t_1) x_0$$

- **Boundedness** (when controllable!).

$$x_0^T P(t_0, t_1) x_0 \leq \int_{t_0}^{\hat{t}} \hat{x}(t)^T Q \hat{x}(t) + \hat{u}(t)^T R \hat{u}(t) dt, \forall t_1 \geq \hat{t}$$

Algebraic Riccati equation (ARE)

- $P = \lim_{t_1 \to +\infty} P(t_0, t_1)$ exists (for each each entry).
- It satisfies the following static equation :

$$PA + A^T P + Q - PBR^{-1}B^T P = 0$$

Are previous results valid ?

- $J(u^*) = \int_{t_0}^{\infty} (x^*)^T Q(x^*) + (u^*)^T R(u^*) \stackrel{?}{=} V(x_0) = x_0^T P x_0$
- is $u^* = -R^{-1}B^T P x^*$ the optimal control ?

Are previous results valid ? Yes !

Consider $\hat{V}(x) := x^T P x$

- As seen in a previous proof sketch, $\frac{d}{dt}\hat{V}(x^*) = -{x^*}^T(Q + PBR^{-1}B^T P)x^*$
- Hence $J(u^*) = \int_{t_0}^{\infty}(x^*)^T(Q + PBR^{-1}B^T P)(x^*) = -\int_{t_0}^{\infty}\frac{d}{dt}\hat{V}(x^*)$

$$= x_0^T P x_0 - (x^*(T))^T P(x^*(T)) \leq x_0^T P x_0$$

- $\forall u, x_0^T P(t_0, t_1)x_0 \leq J(u)$, hence $x_0^T P x_0 \leq J(u)$
- Hence $J(u^*) = x_0^T P x_0$

Under which assumption a closed loop is stable ?

When (A, C) is an observable pair (i.e Kalman criterion :

$rank \begin{pmatrix} C \\ CA \\ ... \\ CA^{n-1} \end{pmatrix} = n$). Main ideas :

- x^* is a linear combinaison of exponential functions of time, hence so do $y^* = Cx^*$, where $Q = C^T C$ and $u^* = \frac{1}{2} RB^T Px^*$.
- Hence, for $J(u^*) = \int_{t_0}^{+\infty} (y^*)^T y^* + (u^*)^T Ru^*$ to be bounded, $\lim_{t\to+\infty} y^*(t) = 0$ and $\lim_{t\to+\infty} u^*(t) = 0$
- By observability, $\exists L, A - LC$ has arbitrary desired eigenvalues with negative real parts.
- Hence : $\dot{x}^* = (A - LC)x^* + Ly^* + Bu^*$ is asymptotically stable.

Main results for infinite-horizon LQR problem

Consider the linear invariant control system $\dot{x} = Ax + Bu$ and cost functional $J(u) = \int_{t_0}^{\infty}((Cx)^T(Cx(t)) + u^T(t)Ru(t))dt$, where (A, B) is controllable, and (A, C) is observable, R symmetric definite positive, then:

- $P := \lim_{t_1 \to +\infty} P(t_0, t_1)$ exists, and is the unique symmetric positive definite solution of the ARE.
- The unique optimal control has the linear time-invariant state feedback $u^*(t) = -R^{-1}B^T Px^*(t)$
- The closed-loop system $\dot{x}^* = (A - BR^{-1}B^T P)x^*$ is exponentially stable.

04

Exercise 6.4

Double integrator

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u \\ J(u) = \int_{t_0}^{\infty} (x_1^2(t) + x_2^2(t) + u^2(t))dt \end{cases}$$

Objective : find P^* the solution of the ARE and illustrate $\lim_{t_1 \to \infty} P(t_0, t_1) = P^*$

Double integrator : other formulation

$$\begin{cases} X = \begin{pmatrix} x_1 \\ \dot{x}_1 \end{pmatrix} \\ \dot{X} = AX + Bu \\ J(u) = \int_{t_0}^{\infty} (X^T(t) Q X(t) + u^2(t))dt \\ A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} R = 1 \end{cases}$$

Double integrator : ARE equation

$$\begin{cases} A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} R = 1 \\ -\dot{P} = PA + A^T P + Q - PBR^{-1}B^T P \\ P = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_4 \end{pmatrix} \end{cases}$$

which implies :

$$\begin{cases} -\dot{p}_1 = 1 - p_2^2 \\ -\dot{p}_2 = p_1 - p_2 p_4 \\ -\dot{p}_4 = 1 + 2p_2 - p_4^2 \end{cases}$$

Double integrator : ARE solution

$$\begin{cases} 0 = 1 - (p_2^*)^2 \\ 0 = p_1^* - (p_2^*)(p_4^*) \\ 0 = 1 + 2p_2^* - (p_4^*)^2 \end{cases} \implies \begin{cases} p_1^* = \sqrt{3} \\ p_2^* = 1 \\ p_4^* = \sqrt{3} \end{cases}$$

Double integrator

ARE solution
$$\begin{cases} p_1^* = \sqrt{3} \\ p_2^* = 1 \\ p_4^* = \sqrt{3} \end{cases}$$
RDE equation
$$\begin{cases} -\dot{p}_1^* = 1 - (p_2^*)^2 \\ -\dot{p}_2^* = p_1^* - p_2^* p_4^* \\ -\dot{p}_4^* = 1 + 2p_2^* - (p_4^*)^2 \end{cases}$$

ARE backward recursion

$$\begin{cases} p_1^*(t-h) = p_1^*(t) + h(1 - p_2^*(t)^2) \\ p_2^*(t-h) = p_2^*(t) + h(p_1^*(t) - p_2^*(t)p_4^*(t)) \\ p_4^*(t-h) = p_4^*(t) + h(1 + 2p_2^*(t) - p_4^*(t)^2) \\ p_1^*(t_1) = p_2^*(t_1) = p_4^*(t_1) = 0 \end{cases}$$

convergence simulations with M = 0

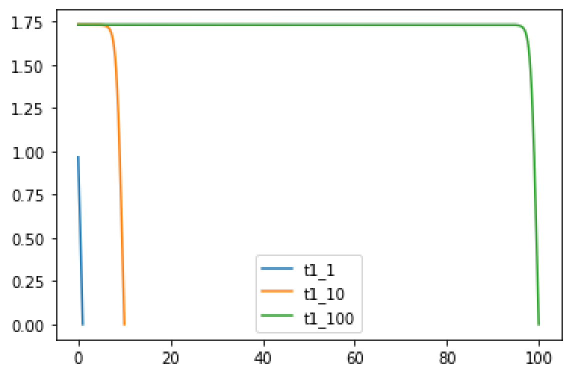

Figure: t$\rightarrow p_1^*(t, t_1), t_1 \in \{1, 10, 100\}$

convergence simulations with M = 0

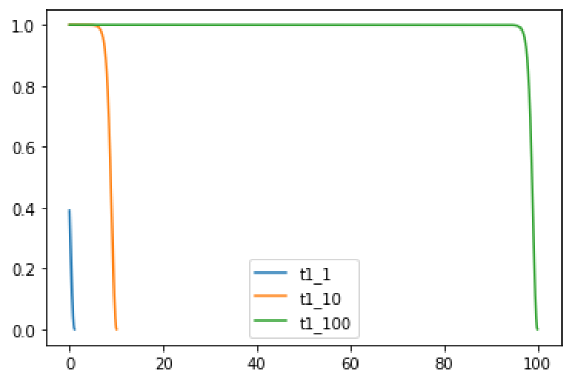

Figure: t→ $p_2^*(t, t_1), t_1 \in \{1, 10, 100\}$

convergence simulations with $m_1 = 5, m_2 = 2,\ m_4 = 10$

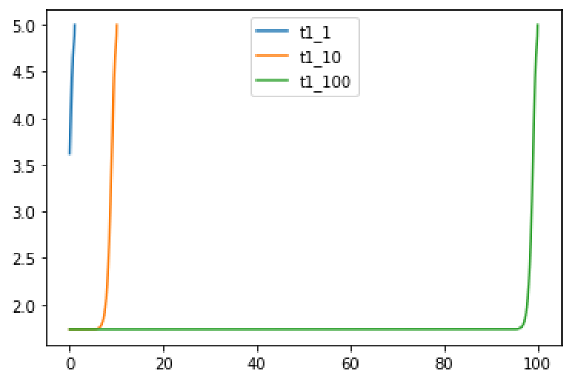

Figure: t$\to p_1^*(t, t_1), t_1 \in \{1, 10, 100\}$

convergence simulations with $m_1 = 5, m_2 = 2,\ m_4 = 10$

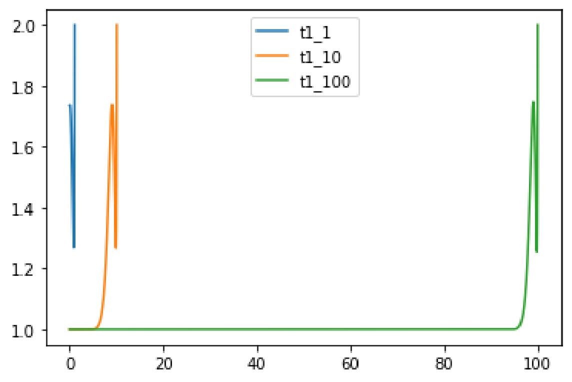

Figure: t$\rightarrow p_2^*(t, t_1), t_1 \in \{1, 10, 100\}$

How to solve analytically the equation when M=0 ?

One idea (J.Nazarzadeh al. 1998 "Solution of the matrix Riccati equation for the linear quadratic control problems") : Consider $Y(t, t_1) = P(t, t_1) - P^*$. Hence, $Y(t_1, t_1) = -P^* < 0$. By monotonicity property $Y(t_0, t_1) < 0$ and WLOG $Y(t, t_1) < 0$ for $t \in [t_0, t_1]$.

$$\begin{cases} -\dot{Y} = YA_c + A_c^T Y - YBR^{-1}B^T Y \\ A_c = A - P^* \end{cases}$$

With change of variable $K = Y^{-1}$ you get:

$$\begin{cases} \dot{K} = A_c K + KA_c^T - BR^{-1}B^T \\ K(t_1) = -P^{*-1} \end{cases}$$

How to solve analytically the equation when M=0 ?

Here, with $K = \begin{pmatrix} k_1 & k_2 \\ k_2 & k_4 \end{pmatrix}$

$$\begin{pmatrix} \dot{k}_1 \\ \dot{k}_2 \\ \dot{k}_3 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ -1 & -\sqrt{3} & 1 \\ 0 & -2 & -2\sqrt{3} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

3 distinct eigenvalues with negative real parts, hence $\forall t \in [t_0, t_1[$:

$$\lim_{t_1 \to +\infty} K^{-1}(t, t_1) = \lim_{t_1 \to +\infty} \frac{1}{k_1 k_4 - k_2^2} \begin{pmatrix} k_4 & -k_2 \\ -k_2 & k_1 \end{pmatrix} = 0_{2x2}$$

Is it true when $M \neq 0$?

At least when (A, C) detectable and (A, B) stabilizable (ref: thoerem 1 p. 237 of book "Linear Optimal Control systems", H. Kwakernaak and R. Sivan).

- Chapter 1: Finite-horizon LQR
  - Definition
    $\dot{x} = A(t)x + B(t)u$
    $x(0) = x_0 \in \mathbb{R}^n$
    $J(u) = \int_{t_0}^{t_1} (x^T Q(t)x + u^T R(t)u)dt + x^T(t_1)Mx(t_1)$
    $M = M^T, Q = Q^T \geq 0, R = R^T > 0$
    $S = \{t_1\}x\mathbb{R}^n$
  - the unique global optimal control has the linear feedback form $u^* = -R^{-1}B^T Px^*$, where P is the unique symmetric positive semi-definite solution of the RDE :

$$\dot{P} = -PA - A^T P - Q + PBR^{-1}B^T P$$

with $P(t_{t_1}) = M$

Chapter 2: Infinite-horizon LQR

- Definition :
  $\dot{x} = Ax + Bu$
  $x(0) = x_0 \in \mathbb{R}^n$
  $J(u) = \int_{t_0}^{\infty} (x^T C^T C x + u^T R u) dt$
  (A, B) is controllable , (A, C) is observable , $R = R^T > 0$
- the unique global optimal control has the linear feedback form $u^* = -R^{-1} B^T P x^*$, where P is the unique, symmetric positive definite solution of the ARE :

$$-PA - A^T P - Q + PBR^{-1} B^T P = 0$$

- The closed-loop system $\dot{x}^* = (A - BR^{-1} B^T P) x^*$ is exponentially stable.