LionSealGrey

Optimal Control 2018

Yury Orlov

Optimal Control 2018

- L1: Functional minimization, Calculus of variations (CV) problem
- L2: Constrained CV problems, From CV to optimal control
- L3: Maximum principle, Existence of optimal control
- L4: Maximum principle (proof)
- L5: Dynamic programming, Hamilton-Jacobi-Bellman equation
- L6: **Linear quadratic regulator**
- L7: Numerical methods for optimal control problems

Exercise sessions (20%):
Solve 50% of problems in advance. Hand-in later.
Mini-project (20%):
Study and present your own optimal control problem.
Written take-home exam (60%).

Summary of L5: HJB equation and viscosity solutions

The value function V of a fixed-time free-end point optimal control is a unique viscosity solution of the HJB equation

$$-V_t(t,x) - \inf_{u \in U}\{L(t,x,u) + < V_x(t,x), f(t,x,u) >\} = 0.$$

with the boundary condition $V(t_1, x) = K(x),\ \forall x \in \mathbb{R}^n$.

Viscosity nonsmooth solutions for the first order PDE

$$F(x, v(x), \nabla v(x)) = 0 \tag{1}$$

were discussed to be approximated by smooth solutions of the viscous fluid equation (what is useful in numerical simulations)

$$F(x, v_\epsilon(x), \nabla v(x)) = \epsilon \Delta v_\epsilon(x) \quad \textbf{as} \quad \epsilon \downarrow 0. \tag{2}$$

Lack of sign symmetry of viscosity solutions is supported by the same of the viscous fluid equation

$$Q_t(t,x) = \epsilon Q_{xx}(t,x) \quad \text{is well-posed}$$

$$\text{whereas} \quad Q_t(t,x) = -\epsilon Q_{xx}(t,x) \quad \text{is ill-posed}$$

Outline

1. FINITE-HORIZON LQR PROBLEM
2. Candidate optimal feedback law
3. Riccati differential equation (RDE)
4. Value function and optimality
5. Global existence of solution for the RDE
6. INFINITE-HORIZON LQR PROBLEM
7. Closed-loop stability
8. Complete result and discussion

Finite-horizon LQR problem

Linear plant dynamics

$$\dot{x} = A(t)x + B(t)u,\ x(t_0) = x_0 \in \mathbb{R}^n$$

Unconstrained control $u \in \mathbb{R}^m$

Target set $S = \{t_1 \times \mathbb{R}^n\}$ (i.e, t_1 is fixed, $x(t_1)$ is free).

Cost functional
$J(u) = \int_{t_0}^{t_1} \Big[x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\Big]dt + x^T(t_1)Mx(t_1)$

Assumptions

$$M = M^T \geq 0,\ Q(t) = Q^T(t) \geq 0,\ R(t) = R^T(t) > 0 \quad \forall t \in [t_0, t_1].$$

Candidate (MP-based) optimal feedback law

Hamiltonian

$$H(t,x,u,p) = p^T A(t)x + p^T B(t)u - x^T Q(t)x - u^T R(t)u$$

where $p_0 = -1$ was chosen due to

Transversality condition
$0 = p^*(t_1) - p_0^* K_x(x^*(t_1)) = p^*(t_1) - 2p_0^* M x^*(t_1)$ to be non-trivial.

Optimality conditions
$0 = H_u|_* = B^T(t)p^* - 2R(t)u^*, \quad 0 \geq H_{uu}|_* = -2R(t)$

Optimal control is thus (if exists) $u^* = \frac{1}{2}R^{-1}B^T(t)p^*(t)$

Adjoint equation $\dot{p}^* = -H_x|* = 2Q(t)x^* - A^T(t)p^*$

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

Next goal: linearity $p^*(t) = -2P(t)x^*(t)$ to be verified for all t rather than just for t_1 where actually $P(t_1) = M$.

Hamiltonian matrix $\mathcal{H}(t)$

Canonical state-costate equations

$$\begin{pmatrix} \dot{x}^* \\ \dot{p}^* \end{pmatrix} = \begin{pmatrix} A(t) & \frac{1}{2}B(t)R_{-1}(t)B^T(t) \\ 2Q(t) & -A^T(t) \end{pmatrix} \begin{pmatrix} x^* \\ p^* \end{pmatrix} =: \mathcal{H}(t) \begin{pmatrix} x^* \\ p^* \end{pmatrix}$$

Hence $\begin{pmatrix} x^*(t) \\ p^*(t) \end{pmatrix} = \Phi(t, t_1) \begin{pmatrix} x^*(t_1) \\ p^*(t_1) \end{pmatrix} =$

$$= \begin{pmatrix} \Phi_{11}(t, t_1) & \Phi_{12}(t, t_1) \\ \Phi_{21}(t, t_1) & \Phi_{22}(t, t_1) \end{pmatrix} \begin{pmatrix} x^*(t_1) \\ p^*(t_1) \end{pmatrix} \qquad (3)$$

where the inverse $\Phi(t, t_1) = \Phi^{-1}(t_1, t)$ of the fundamental matrix $\Phi(t_1, t)$ propagates the solution backward

Substituting the costate boundary condition $p^*(t_1) = -2Mx^*(t_1)$ into (3) yields

$$x^*(t) = \Big(\Phi_{11}(t, t_1) - 2\Phi_{12}(t, t_1)M\Big)x^*(t_1)$$

$$p^*(t) = \Big(\Phi_{21}(t, t_1) - 2\Phi_{22}(t, t_1)M\Big)x^*(t_1)$$

State feedback

Provided that $\exists \left(\Phi_{11}(t,t_1) - 2\Phi_{12}(t,t_1)M\right)^{-1} \forall t$

it follows

$$p^*(t) = \left(\Phi_{21}(t,t_1) - 2\Phi_{22}(t,t_1)M\right)\left(\Phi_{11}(t,t_1) - 2\Phi_{12}(t,t_1)M\right)^{-1} x^*(t)$$

thus concluding that

$$P(t) := -\frac{1}{2}\left(\Phi_{21}(t,t_1) - 2\Phi_{22}(t,t_1)M\right)\left(\Phi_{11}(t,t_1) - 2\Phi_{12}(t,t_1)M\right)^{-1}$$

Summarizing, the closed-loop optimal control is obtained

$$u^*(t) = -R^{-1}(t)B^T(t)P(t)x^*(t)$$

Riccati differential equation

Differentiating

$$p^*(t) = -2P(t)x^*(t) \tag{4}$$

yields

$$\dot{p}^*(t) = -2\dot{P}(t)x^*(t) - 2P(t)\dot{x}^*(t).$$

Let us now use the canonical equations

$$\begin{pmatrix} \dot{x}^* \\ \dot{p}^* \end{pmatrix} = \begin{pmatrix} A(t) & \frac{1}{2}B(t)R_{-1}(t)B^T(t) \\ 2Q(t) & -A^T(t) \end{pmatrix} \begin{pmatrix} x^* \\ p^* \end{pmatrix}$$

to arrive at

$$2Q(t)x^*(t) - A^T(t)p^*(t) =$$
$$-2\dot{P}(t)x^*(t) - 2P(t)A(t)x^*(t) - P(t)B(t)R^{-1}(t)B^T(t)p^*(t)$$

Applying (4) **it follows that** $\Rightarrow$

RDE derivation (continued)

$$Q(t)x^*(t) + A^T(t)P(t)x^*(t) =$$
$$-\dot{P}(t)x^*(t) - 2P(t)A(t)x^*(t) + P(t)R^{-1}(t)B^T(t)B^T(t)P(t)x^*(t)$$

Since x_0 is arbitrary then the state $x^*(t)$ is arbitrary as well as far as the state transition matrix is nonsingular.

$$\Downarrow$$

The RDE must be satisfied for $P(t)$ subject to $P(t_1) = M$:

$$\dot{P}(t) = P(t)B(t)R^{-1}(t)B^T(t)P(t) - P(t)A(t) - A^T(t)P(t) - Q(t).$$

Maximum principle resulted in a unique candidate for an optimal control $u^*(t) = -R^{-1}(t)B^T(t)P(t)x^*(t)$

Other tools should be involved for proving the existence of $P(t)$ for all t as well as for proving the optimality of the control thus derived

Value function and global optimality

LQR-specialized HJB equation

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

Boundary condition $V(t_1, x)) = x^T M x = x^T P(t_1) x$

$R(t) > 0 \Rightarrow$ **the minimizing control** $u = -\frac{1}{2} R^{-1}(t) B^T(t) V_x(t,x)$

LQR-specialized HJB equation is thus simplified to

$$-V_t(t,x) = x^T Q(t) x + \Big(V_x(t,x)\Big)^T A(t) x$$
$$-\frac{1}{4}\Big(V_x(t,x)\Big)^T B(t) R^{-1}(t) B^T(t) V_x(t,x).$$

Just in case if $u^*(t) = -R^{-1}(t) B^T(t) P(t) x^*(t)$ **is the minimizing control, then**

$$\frac{1}{2} V_x(t,x) = P(t) x \quad \Rightarrow \quad V(t,x) = x^T P(t) x.$$

The above quadratic V **does satisfy the HJB equation** provided that $P(t)$ is symmetric (your homework, Exercise 6.2).

Global existence of RDE solutions

Riccati differential equation

$$\dot{P}(t) = P(t)B(t)R^{-1}(t)B^T(t)P(t) - Q(t) - P(t)A(t) - A^T(t)P(t)$$

Subject to $P(t_1) = M,$ **a local solution exists on some** $(\bar{t}, t_1)$**.**

1. To the contrary of the global existence, suppose that $\bar{t} \neq t_0$ and some entries of $P(t)$ escape to infinity as $t \downarrow \bar{t}$;
2. $P(t)$ is known from Exercise 6.2 (homework) to be symmetric and positive semidefinite $\Rightarrow$ all principal minors must be nonnegative;
3. if an off-diagonal entry $P_{ij}(t)$ becomes unbounded near $\bar{t}$, while all diagonal entries stay bounded, then a ceratin 2×2 principal minor must be negative near $\bar{t}$;
4. thus, only diagonal entries, say $P_{ii}(t)$, can be unbounded $\Rightarrow$ the optimal cost-to-go $e_i^T P_{ii}(t) e_i$ from $e_i = (0, \ldots, 1, \ldots, 0)^T$ escapes to infinity as $t \downarrow \bar{t}$;
5. this contradicts to the cost optimality because, e.g., $u \equiv 0$ on $[\bar{t}, t_1]$ would result in a lower finite cost.

Example

$$\dot{x} = u, \quad J(u) = \int_{t_0}^{t_1} [x^2(t) + u^2(t)]dt \to \min$$

$$\Downarrow$$

RDE $\quad \dot{P} = P^2 - 1, \quad P(t_1) = 0$

$$\Downarrow$$

Optimal control $\quad u = -\tanh(t_1 - t)x$

If $R = -1$**, i.e.,** $J(u) = \int_{t_0}^{t_1} [x^2(t) - u^2(t)]dt$**, the RDE** $\dot{P} = -P^2 - 1$ **has no global solutions**

$$\Downarrow$$

Assumption $R > 0$ is thus important.

Infinite-horizon autonomous LQR

Matrices A, B, Q, R are constant and the terminal cost $M = 0$.

RDE $\quad \dot{P} = PBR^{-1}B^TP - Q - PA - A^TP, \quad P(t_1) = 0$

Solution of the above RDE is relabeled as $P(t, t_1)$

Optimal control $\quad u^*_{t_1}(t) = -R^{-1}B^TP(t, t_1)x$

Value function $\quad V^{t_1}(t, x) = x^TP(t, t_1)x$

Finite-horizon optimal cost $\quad V^{t_1}(t_0, x_0) = x_0^TP(t_0, t_1)x_0$

Clearly, the finite-horizon optimal cost is monotonically nondecreasing in t_1.

Moreover, it remains bounded as $t_1 \to \infty$ provided that A and B are controllable. Indeed, it is upperbounded by the cost, matching to $u(t)$, steering the state to the origin by a time instant $\hat{t}$ and which is nullified after $\hat{t}$.

Properties of the limit

Thus, $\exists \lim_{t_1 \to \infty} x^T P(t, t_1) x$. **Moreover,** $\exists \lim_{t_1 \to \infty} P(t, t_1)$.

Indeed, $\exists \quad \lim_{t_1 \to \infty} e_i^T P(t, t_1) e_i = \lim_{t_1 \to \infty} P_{ii}(t, t_1)$ and

$$\exists \quad \lim_{t_1 \to \infty} (e_i + e_j)^T P(t, t_1)(e_i + e_j) = \lim_{t_1 \to \infty} (P_{ii} + 2P_{ij} + P_{jj})$$

Actually, $P(t, t_1) = P(t_1 - t)$ **by virtue of the time-invariance of the RDE and hence there exists a steady state**

$$\lim_{t_1 \to \infty} P(t, t_1) = P \geq 0 \quad \forall t$$

Passing to the limit as $t_1 \to \infty$ **on both sides of the RDE, algebraic Riccati equation (ARE) is obtained for the stedy state** P**:**

$$PA + A^T P + Q - PBR^{-1}B^T P = 0 \tag{5}$$

This is similar to passing from the general HJB equation to its infinite-horizon counterpart.

Our hope that there exists a unique solution $P = P^T \geq 0$ of (5).

Infinite-horizon problem and its solution

$$J(u) = \int_{t_0}^{\infty} \Big[x^T(t) Q x(t) + u^T(t) R u(t) \Big] dt \to \min$$

Is $V(x_0) = x_0^T P x_0$ **optimal cost?**

Is $u^*(t) = -R^{-1} B^T P x$ **optimal control?**

Indeed, $\frac{d}{dt}\Big[(x^*)^T(t) P x^*(t)\Big] = (x^*)^T(t)\Big[P(A - BR^{-1}B^T P) + (A^T$

$-PBR^{-1}B^T)P\Big] x^*(t) = (x^*)^T(t)\Big[PA + A^T P - 2PBR^{-1}B^T P\Big] x^*(t)$

$= -(x^*)^T(t)\Big[Q + PBR^{-1}B^T P\Big] x^*(t)$

It follows $\int_{t_0}^{T} \Big[(x^*)^T(t) Q x^*(t) + (u^*)^T(t) R u^*(t)\Big] dt$

$= \int_{t_0}^{T} (x^*)^T(t)\Big[Q + PBR^{-1}P\Big] x^*(t) dt$

$= -\int_{t_0}^{T} \frac{d}{dt}\Big[(x^*)^T(t) P x^*(t)\Big] dt = x_0^T P x_0 - (x^*)^T(T) P x^*(T) \leq x_0^T P x_0$

Infinite-horizon problem and its solution (cont'd)

Taking the limit as $T \to \infty$, it is thus concluded

$$J(u^*) \leq x_0^T P x_0 \tag{6}$$

On the other hand, $x_0^T P(t_0, t_1) x_0$ is the finite-horizon optimal cost and $\forall x$, subject to the same initial condition, one has

$$x_0^T P(t_0, t_1) x_0 \leq \int_{t_0}^{t_1} \Big[x^T(t) Q x(t) + (u)^T(t) R u(t)\Big] dt$$
$$\leq \int_{t_0}^{\infty} (\Big[x^T(t) Q x(t) + (u)^T(t) R u(t)\Big] dt = J(u)$$

Passing to the limit as $t_1 \to \infty$, it follows

$$x_0^T P x_0 \leq J(u)$$

By virtue of (6), the optimality of u^* is concluded:

$$J(u^*) = x_0^T P x_0 \leq J(u) \quad \forall u$$

Closed-loop stability

Example

$$\dot{x} = x + u,\ J = \int_0^\infty u^2 dt \to \infty$$

Optimal control $u^* \equiv 0 \Rightarrow$ the closed-loop system $\dot{x} = x$ is unstable.

Let system $\dot{x} = Ax + Bu$, be observable with the output $y = Cx$ such that $Q = C^T C$. Its optimal cost functional

$$J(u^*) = \int_{t_0}^\infty \Big[(x^*)^T(t) C^T C x^*(t) + (u^*)^T(t) R u^*(t)\Big] dt < \infty.$$

$$\Downarrow$$

$$y^*(t) = Cx^*(t) \to 0, u^*(t) \to 0 \quad \text{as} \quad t \to \infty$$

$$\Downarrow$$

$x^*(t) \to 0$ as $t \to \infty$ $\Rightarrow$ Closed-loop exponential stability

Infinite-horizon LQR: complete result

System dynamics $\dot{x} = Ax + Bu$

Cost functional

$$J(u) = \int_{t_0}^{\infty} \Big[x^T(t) C^T C x(t) + u^T(t) R u(t) \Big] dt$$

where (A, B) **is controllable,** (A, C) **is observable, and** $R = R^T > 0$**.**

Theorem

1. $\exists P = \lim_{t \to \infty} P(t_0, t_1)$ *of the solution of the RDE with the terminal condition* $P(t_1) = 0$*; this limit is a unique symmetric, positive definite solution of the corresponding ARE;*
2. *The optimal cost* $V(x_0) = x_0^T P x_0$;
3. *The unique optimal control* $u^*(t) = -R^{-1} B^T P x^*(t)$*;*
4. *The closed-loop system* $\dot{x}^* = (A - BR^{-1}B^T P)x^*$ *is exponentially stable.*

Proof of the complete result

All has been proved except the solution uniqueness and positive definiteness of P.

Proving $P > 0$

Suppose that $x_0^T P x_0 = 0$. Then for this initial condition x_0, the optimal cost

$$\int_{t_0}^{\infty} \Big[(x^*)^T(t) C^T C x^*(t) + (u^*)^T(t) R u^*(t) \Big] dt = 0$$

$$\Downarrow$$

$$C x^* \equiv 0, \quad u^* \equiv 0 \quad \text{(since} \quad R > 0)$$

By observability, it follows that $x_0 = 0$, and hence $\Rightarrow P > 0$.

Proof of the complete result (cont'd)

P is a unique positive (semi)definite solution of the ARE.

Suppose that $\exists \bar{P} > 0$ (or even $\bar{P} \geq 0$). Consider the new cost functional

$$\bar{J}^{t_1}(u) := \int_{t_0}^{t_1} \Big[x^T(t)Qx(t) + u^T(t)Ru(t) \Big] dt + x^T(t_1)\bar{P}x(t_1)$$

and its infinite-horizon counterpart:

$$\bar{J}^{\infty}(u) := \lim_{t_1 \to \infty} \Big\{ \int_{t_0}^{t_1} \Big[x^T(t)Qx(t) + u^T(t)Ru(t) \Big] dt + x^T(t_1)\bar{P}x(t_1) \Big\}.$$

Then

$$\bar{J}^{\infty}(u^*) = \int_{t_0}^{\infty} \Big[(x^*)^T(t)Qx^*(t) + (u^*)^T(t)Ru^*(t) \Big] dt \Big\} = x_0^T P x_0$$

is the optimal cost with respect to $\bar{J}^{\infty}$ because

$$\bar{J}^{\infty}(u) \geq \int_{t_0}^{\infty} \Big[x^T(t)Qx(t) + u^T(t)Ru(t) \Big] dt \Big\} \geq x_0^T P x_0.$$

Proof of the uniqueness of P (cont'd)

On the other hand, the optimal cost is given by

$$\bar{J}^{t_1}(u^*) = x_0^T P(t_0; \bar{P}, t_1) x_0$$

where $P(t_0; \bar{P}, t_1)$ denotes the solution of the corresponding RDE subject to $P(t_1) = \bar{P}$.

$$\Downarrow$$

$$P(t_0; \bar{P}, t_1) = \bar{P}$$

because $\bar{P}$ is an equilibrium of the RDE as it satisfies ARE by assumption

$$\Downarrow$$

$P = \bar{P}$ and the P-uniqueness is thus verified

Proof of the complete result (cont'd)

It remains to establish that the optimal control is unique.

By maximum principle, the optimal control satisfies

$$u^*(t) = \arg\max_{u \in U} \Big\{ L(t, x^*, u) + \Big\langle V_x(t, x^*(t)), f(t, x^*(t), u) \Big\rangle \Big\}$$

to presently be specified to

$$u^*(t) = \arg\max_{u \in U} \Big\{ (x^*)^T t) Q x^*(t) + u^T B u$$
$$+ 2(x^*)^T(t) P A x^*(t) + 2(x^*)^T(t) P B u \Big\}$$

The latter uniquely identifies the optimal feedback