第九讲：Hamilton-Jacobi-Bellman 方程

最优控制的数学理论之五

张杰

人工智能学院
中国科学院大学

复杂系统管理与控制国家重点实验室
中国科学院自动化研究所

2017 年 10 月 17 日

Table of Contents

周四，10 月 19 日随堂考试

- 开卷考试，禁止电子设备，自带纸笔
- 计算题、证明题、问答题
- 包括截止考试当日的所有内容
- 折算占平时成绩中的 10 分
- 考试时间，10 月 19 日课上后半节

Table of Contents

离散时间最优控制问题

问题 1 (离散时间最优控制问题)

状态变量为 $x(k) : \mathbb{N} \to \mathbb{R}^n$，控制变量为 $u(k) : \mathbb{N} \to \mathbb{R}^m$

(1) 被控对象的状态方程

$$x(k+1) = f_D(x(k), u(k), k), \quad k = 0, \ldots, N-1. \tag{1}$$

(2) 容许控制：

$$u(k) \in U, \quad x(k) \in X. \tag{2}$$

(3) 目标集：

$$x(N) \in \mathcal{S}. \tag{3}$$

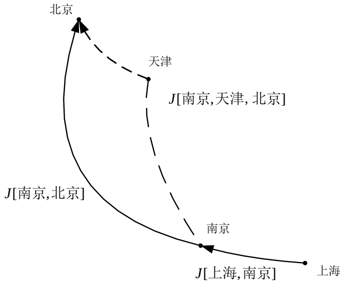
(4) 性能指标：

$$J(u; x(k), k) = h_D(x(N), N) + \sum_{i=k}^{N-1} g_D(x(i), u(i), i). \tag{4}$$

动态规划的最优性原理

定理 1 (最优性原理, Bellman1954)

多级决策过程的最优策略具有如下性质：不论初始状态和初始决策如何，其余的决策对于由初始决策所形成的状态来说，必定也是一个最优策略



如果南京-天津-北京是南京到北京的最短路，
上海-南京-北京会是最短路吗

Bellman 方程

定理 2 (Bellman 方程)

x_0 为初值 k_0 为初始时刻，最优控制下的性能指标记为“值函数”

$$V(x_0, k_0) = \min_{u \in U} J(u; x_0, k_0) \tag{5}$$

根据最优性原理，最优控制满足下列 Bellman 方程：

$$V(x(N), N) = h_D(x(N), N). \tag{6}$$

$$V(x(k), k) = \min_{u(k) \in U} \{g_D(x(k), u(k), k) + V(x(k+1), k+1)\}, \quad k = N-1, \ldots, 0. \tag{7}$$

HJB 方程和 PMP、经典变分的关系

拉格朗日变分法

$\Updownarrow$ 等价

哈密尔顿方程组 $\Leftrightarrow$ 哈密尔顿雅各比方程

$\Downarrow$ + 控制 $\Downarrow$ + 控制

极值原理 $\Leftarrow V$ 二次可微特况* HJB 方程

Table of Contents

Hamilton-Jacobi-Bellman 方程

定理 3 (Hamilton-Jacobi-Bellman 方程)

若最优控制问题有解，值函数是以 t_0 为初始时刻，x_0 为初始状态，在最优控制下的性能指标：

$$V(x_0, t_0) = \min_u J(u; x_0, t_0). \tag{8}$$

若值函数二阶连续可微，则如下 Hamilton-Jacobi-Bellman 方程（简称 HJB 方程）是最优控制的充分必要条件：

$$-\frac{\partial V}{\partial t}(x(t), t) = \min_{u(t)\in\mathbb{R}^m} \mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t), t), \tag{9}$$

$$V(x(t_f), t_f) = h(x(t_f), t_f)\text{(终端代价)} . \tag{10}$$

1/4 HJB 方程必要性-最优性原理

将性能指标分成 $[t, t+\Delta t]$ 和 $[t+\Delta t, t_f]$ 两段

$$
\begin{aligned}
V(x(t),t) &= \min_{u(\tau),\tau\in[t,t_f]} \{h(x(t_f),t_f) + \int_t^{t_f} g(x(\tau),u(\tau),\tau)d\tau\} \\
&= \min_{u(\tau),\tau\in[t,t_f]} \{h(x(t_f),t_f) + \int_t^{t+\Delta t} gd\tau + \int_{t+\Delta t}^{t_f} gd\tau\}
\end{aligned}
$$

由最优性原理，若 u 是以 $x(t), t$ 为初始状态、初始时刻的最优控制，则 u 也是 $x(t+\Delta t), t+\Delta t$ 为初始状态初始时刻的最优控制，于是，后半性能指标等于值函数

$$
V(x(t),t) = \min_{u(\tau),\tau\in[t,t+\Delta t]} \{\int_t^{t+\Delta t} gd\tau + V(x(t+\Delta t), t+\Delta t)\}.
$$

2/4 HJB 方程必要性-泰勒展开

假定 V 二阶连续可微，将值函数在 $x(t), t$ 泰勒展开

$$
\begin{aligned}
V(x(t), t) &= \min_{u(\tau), \tau \in [t, t+\Delta t]} \{ \int_t^{t+\Delta t} g d\tau + V(x(t+\Delta t), t+\Delta t) \} \\
&= \min_{u(\tau), \tau \in [t, t+\Delta t]} \{ \int_t^{t+\Delta t} g d\tau + V(x(t), t) \\
&\quad + \frac{\partial V}{\partial t}(x(t), t)\Delta t \\
&\quad + [\frac{\partial V}{\partial x}(x(t), t)]^T [x(t+\Delta t) - x(t)] \\
&\quad + o(\Delta t) \}
\end{aligned}
$$

3/4 HJB 方程必要性-简单整理

对较小的 Δt, 对 $\tau \in [t, t+\Delta t]$， $u(\tau) \approx u(t)$

$$
\begin{aligned}
V(x(t),t) = \min_{u(t)} \{ & g(x(t),u(t),t)\Delta t + V(x(t),t) \\
& + \frac{\partial V}{\partial t}(x(t),t)\Delta t \\
& + [\frac{\partial V}{\partial x}(x(t),t)]^T [f(x(t),u(t),t)]\Delta t + o(\Delta t)\}
\end{aligned}
$$

于是

$$
\begin{aligned}
-\frac{\partial V}{\partial t}(x(t),t)\Delta t = & \min_{u(t)} \{ g(x(t),u(t),t)\Delta t \\
& + [\frac{\partial V}{\partial x}(x(t),t)]^T [f(x(t),u(t),t)]\Delta t + o(\Delta t)\}
\end{aligned}
$$

4/4 HJB 方程必要性-取极限

两边同除 Δt，取 $\Delta t \to 0$，即可得对于 $t \in [t_0, t_f]$ 都有 HJB 方程

$$-\frac{\partial V}{\partial t}(x(t), t) = \min_{u(t)}\{g(x(t), u(t), t) + [\frac{\partial V}{\partial x}(x(t), t)]^T f(x(t), u(t), t)\}$$

令 $t = t_f$，得到边界条件

$$V(x(t_f), t_f) = h(x(t_f), t_f). \tag{11}$$

1/3 HJB 方程必要性-命题表述

定理 4 (HJB 方程的充分条件)

若存在函数 $V(x,t):\mathbb{R}^n\times[t_0,t_f]\to\mathbb{R}$ 满足 HJB 方程:

$$-\frac{\partial V}{\partial t}(x(t),t)=\min_{\xi}\mathcal{H}\Big(x(t),\xi,\frac{\partial V}{\partial x},t\Big),$$

及边界条件

$$V(x(t_f),t_f)=h(x(t_f),t_f).$$

则,$V(x_0,t_0)$ 是以 t_0 为初始时刻,x_0 为初始状态的值函数

若存在 $x(t),u(t)$ 恰满足哈密尔顿函数最小化,那对其他 $u'(t)$,性能指标均不优于 V

2/3 HJB 方程必要性-处理极值条件

设 $u(t)$ 是满足 HJB 方程的控制变量，$x(t)$ 是对应的状态

$$
\begin{aligned}
-\frac{\partial V}{\partial t}(x(t),t) &= \min_{\xi} \mathcal{H}(x(t),\xi,\frac{\partial V}{\partial x}(x(t),t),t) \\
&= \mathcal{H}(x(t),u(t),\frac{\partial V}{\partial x}(x(t),t),t) \\
&= g(x(t),u(t),t) + \frac{\partial V}{\partial x}(x(t),t)\cdot f(x(t),u(t),t), \\
0 &= g(x(t),u(t),t) + \frac{d}{dt}[V(x(t),t)].
\end{aligned}
$$

$$
0 = \int_{t_0}^{t_f} g(x(t),u(t),t) + [V(x(t_f),t_f) - V(x(t_0),t_0)]
$$

$$
V(x_0,t_0) = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t) = J(u;x_0,t_0) \tag{12}
$$

3/3 HJB 方程必要性-最优性

对于其他 $x'(t), u'(t)$，

$$
\begin{aligned}
-\frac{\partial V}{\partial t}(x'(t), t) &= \min_{\xi} \mathcal{H}(x'(t), \xi, \frac{\partial V}{\partial x}(x'(t), t), t) \\
&\leq g(x'(t), u'(t), t) + \frac{\partial V}{\partial x}(x'(t), t) \cdot f(x'(t), u'(t), t), \\
0 &\leq g(x'(t), u'(t), t) + \frac{d}{dt}[V(x'(t), t)].
\end{aligned}
$$

$$
0 \leq \int_{t_0}^{t_f} g(x'(t), u'(t), t) dt + [V(x'(t_f), t_f) - V(x'(t_0), t_0)]
$$

$$
V(x_0, t_0) \leq h(x'(t_f), t_f) + \int_{t_0}^{t_f} g(x'(t), u'(t), t) = J(u'; x_0, t_0) \quad (13)
$$

于是 u 是最优控制， V 是值函数

1/3 HJB⇒PMP

上述证明过程中出现了极值条件，即对于最优控制 $x(t), u(t)$

$$\mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t), t) \leq \mathcal{H}(x(t), u'(t), \frac{\partial V}{\partial x}(x(t), t), t).$$

若我们能由此推得 $p = \partial V / \partial x$ 满足协态方程和边界条件，则推得极值原理！

$$\dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial x},$$

即，

$$\frac{d}{dt}[\frac{\partial V}{\partial x}] = -\frac{\partial \mathcal{H}}{\partial x}.$$

假定 V 二次连续可微，考察 $n = m = 1$，终端状态时间 free。有

$$\frac{\partial^2 V}{\partial x \partial t} = \frac{\partial^2 V}{\partial t \partial x}, \quad -\frac{\partial V}{\partial t} = g + \frac{\partial V}{\partial x} f$$

2/3 协态方程

$$
\begin{aligned}
\frac{d}{dt}[\frac{\partial V}{\partial x}(x(t),t)] &= \frac{\partial^2 V}{\partial x^2}(x(t),t)\frac{dx}{dt} + \frac{\partial}{\partial t}[\frac{\partial V}{\partial x}(x(t),t)] \\
&= \frac{\partial^2 V}{\partial x^2} f + \frac{\partial}{\partial x}[\frac{\partial V}{\partial t}] \\
&= \frac{\partial^2 V}{\partial x^2} f - \frac{\partial}{\partial x}[g + \frac{\partial V}{\partial x} f] \\
&= \frac{\partial^2 V}{\partial x^2} f - [\frac{\partial g}{\partial x} + \frac{\partial^2 V}{\partial x^2} f + \frac{\partial V}{\partial x}\frac{\partial f}{\partial x}] \\
&= -\frac{\partial g}{\partial x} - \frac{\partial V}{\partial x}\frac{\partial f}{\partial x} \\
&= -\frac{\partial \mathcal{H}}{\partial x}(x(t), u(t), \frac{\partial V}{\partial x}, t)
\end{aligned}
$$

3/3 边界条件

边界条件

$$V(x(t_f), t_f) = h(x(t_f), t_f).$$

于是

$$\frac{\partial V}{\partial x}(x(t_f), t_f) = \frac{\partial h}{\partial x}(x(t_f), t_f),$$

终端状态自由的边界条件。直接令 HJB 方程在 t_f 取值，

$$\begin{aligned} 0 &= \frac{\partial V}{\partial t}(x(t_f), t_f) + \min_{\xi} \mathcal{H}(x(t_f), \xi, \frac{\partial V}{\partial x}(x(t_f), t_f), t_f) \\ &= \frac{\partial V}{\partial t}(x(t_f), t_f) + \mathcal{H}(x(t_f), u(t_f), \frac{\partial V}{\partial x}(x(t_f), t_f), t_f) \end{aligned}$$

Table of Contents

例子：HJB 方程求解最优控制

例 1

$x(t) : [t_0, t_f] \to \mathbb{R}^n$, $u(t) : [t_0, t_f] \to \mathbb{R}^n$

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

其中 $A + A^T = 0$，$\|u(t)\| \leq 1$。求最优控制最小化性能指标

$$J(u) = \int_{t_0}^{t_f} dt.$$

1/3 计算 Hamiltonian，求极小

$$\mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t)) = 1 + \frac{\partial V}{\partial x}(x(t), t) \cdot [Ax(t) + u(t)].$$

在 $\|u(t)\| \leq 1$ 的容许控制下，由 Cauchy-Schwarz 不等式，

$$|\frac{\partial V}{\partial x}(x(t), t) \cdot u(t)|^2 \leq \|\frac{\partial V}{\partial x}(x(t), t)\|^2 \|u(t)\|^2,$$

$$\frac{\partial V}{\partial x}(x(t), t) \cdot u(t) \geq -\|\frac{\partial V}{\partial x}(x(t), t)\| \|u(t)\| = -\|\frac{\partial V}{\partial x}(x(t), t)\|.$$

$$\begin{aligned}
&\mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t)) \\
=&1 + \frac{\partial V}{\partial x}(x(t), t) \cdot [Ax(t)] + \frac{\partial V}{\partial x}(x(t), t) \cdot [u(t)] \\
\geq&1 + \frac{\partial V}{\partial x}(x(t), t) \cdot [Ax(t)] - \|\frac{\partial V}{\partial x}(x(t), t)\|.
\end{aligned}$$

当且仅当

$$\frac{\partial V}{\partial x} = ku(t),$$

$$\|u(t)\| = 1$$

2/3 HJB 方程

$$-\frac{\partial V}{\partial t}(x(t),t) = 1 + \frac{\partial V}{\partial x}(x(t),t)\cdot[Ax(t)] - \|\frac{\partial V}{\partial x}(x(t),t)\|.$$

下面验证，$V(x(t),t) = \|x(t)\|$ 满足 HJB 方程。若 $x(t) \neq 0$，

$$\begin{aligned} V(x(t),t) &= \|x(t)\| = \frac{x^{\mathrm{T}}(t)x(t)}{\|x(t)\|}, \\ \frac{\partial V}{\partial x}(x(t),t) &= \frac{x(t)}{\|x(t)\|}, \\ \frac{\partial V}{\partial t}(x(t),t) &= 0. \end{aligned}$$

3/3 整理得最优控制

则

$$
\begin{aligned}
&\frac{\partial V}{\partial t}(x(t),t)+1+\frac{\partial V}{\partial x}(x(t),t)\cdot[Ax(t)]-\|\frac{\partial V}{\partial x}(x(t),t)\| \\
=&0+1+\frac{x^{\mathrm{T}}(t)Ax(t)}{\|x(t)\|}-\frac{\|x(t)\|}{\|x(t)\|}=\frac{x^{\mathrm{T}}(t)Ax(t)}{\|x(t)\|} \\
=&\frac{1}{\|x(t)\|}\frac{1}{2}x^{\mathrm{T}}(t)[A+A^{\mathrm{T}}]x(t)=0.
\end{aligned}
$$

即，$V(x(t),t)=\|x(t)\|$ 满足 HJB 方程。于是，

$$
u(t)=-\frac{x(t)}{\|x(t)\|}.
$$

例子：HJB 方程求解最优控制

再来看前面使用经典变分和极小值原理求解过的停车问题

例 (小车的能量最优控制)

位置 x_1，速度 x_2，加速度 u。质量为 1 小车的状态方程为：

$$\dot{x}_1(t) = x_2(t), \tag{14}$$

$$\dot{x}_2(t) = u(t). \tag{15}$$

要将状态从初始的 $x(t_0) = x_0$ 在规定的 t_f 到达 $x(t_f) = x_f$。
求最优控制以最小化控制能量：

$$J(u) = \int_{t_0}^{t_f} \frac{1}{2} u^2(t)\,\mathrm{d}t. \tag{16}$$

$t_0 = 0,\ t_f = 2,\ x_0 = [-2, 1]^{\mathrm{T}},\ x_f = [0, 0]^{\mathrm{T}}$

1/8 引入惩罚函数，消除终值约束

HJB 方程没有处理终端状态的约束。我们在原性能指标

$$J(u)=\int_{t_0}^{t_f}\frac{1}{2}u^2(t)\,\mathrm{d}t$$

基础上加上 $x(t_f)=x_f$ 的"惩罚函数"

$$h(x(t_f),t_f)=\frac{b}{2}\|x(t_f)-x_f\|_2^2 \tag{17}$$

b 很大 $x(t_f)\neq x_f$ 时性能指标大幅提升（受到惩罚）

$$J(u)=\frac{b}{2}\|x(t_f)-x_f\|_2^2+\int_{t_0}^{t_f}\frac{1}{2}u^2(t)dt \tag{18}$$

2/8 计算 Hamiltonian，求极小

$$\mathcal{H}(x(t),u(t),p(t),t)=\frac{1}{2}u^2(t)+\frac{\partial V}{\partial x_1}(x(t),t)x_2(t)+\frac{\partial V}{\partial x_2}(x(t),t)u(t)$$

$$0=\frac{\partial \mathcal{H}}{\partial u}\Rightarrow u(t)=-\frac{\partial V}{\partial x_2}(x(t),t) \tag{19}$$

$$\min_u \mathcal{H}=\frac{\partial V}{\partial x_1}(x(t),t)x_2(t)-\frac{1}{2}[\frac{\partial V}{\partial x_2}(x(t),t)]^2 \tag{20}$$

3/8 得到 HJB 方程

将 $\min \mathcal{H}$ 代入，得到 HJB 方程

$$-\frac{\partial V}{\partial t}(x(t),t) = \frac{\partial V}{\partial x_1}(x(t),t)x_2(t) - \frac{1}{2}[\frac{\partial V}{\partial x_2}(x(t),t)]^2 \tag{21}$$

及考虑惩罚函数的边界条件

$$V(x(t_f),t_f) = \frac{b}{2}\|x(t_f) - x_f\|_2^2 = \frac{b}{2}[x_1(t_f)^2 + x_2(t_f)^2]. \tag{22}$$

4/8 求解 HJB 方程

假定 HJB 方程的解为二次形式，即

$$V(x(t),t)=\frac{1}{2}[k_1(t)x_1^2(t)+2k_2(t)x_1(t)x_2(t)+k_3(t)x_2^2(t)]$$

则

$$\frac{\partial V}{\partial t}(x(t),t)=\frac{1}{2}(\dot{k}_1(t)x_1^2(t)+2\dot{k}_2(t)x_1(t)x_2(t)+\dot{k}_3(t)x_2^2(t))$$
$$\frac{\partial V}{\partial x_1}(x(t),t)=k_1(t)x_1(t)+k_2(t)x_2(t)$$
$$\frac{\partial V}{\partial x_2}(x(t),t)=k_2(t)x_1(t)+k_3(t)x_2(t)$$

代入 HJB 方程 (21) 整理得

$$\dot{k}_1x_1^2+2\dot{k}_2x_1x_2+\dot{k}_3x_2^2=-2k_1x_1x_2-2k_2x_2^2+k_2^2x_1^2+2k_2k_3x_1x_2+k_3^2x$$

5/8 求得闭环形式最优控制

由于 HJB 方程对任意 x_1, x_2, t 均成立，得

$$\dot{k}_1 = k_2^2 \tag{23}$$
$$\dot{k}_2 = -k_1 + k_2 k_3 \tag{24}$$
$$\dot{k}_3 = -2k_2 + k_3^2 \tag{25}$$

及终值条件

$$k_1(t_f) = b \tag{26}$$
$$k_2(t_f) = 0 \tag{27}$$
$$k_3(t_f) = b \tag{28}$$

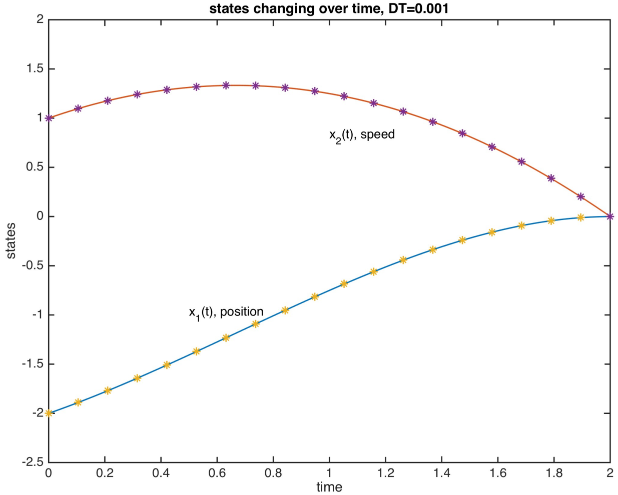
求解该方程，得闭环最优控制 $\boxed{u(t) = -k_2(t)x_1(t) - k_3(t)x_2(t)}$

6/8 常微分方程数值求解

$$\frac{x(t_{k+1}) - x(t_k)}{\Delta t} \approx \phi(x(t_k), t_k) \Rightarrow$$
$$x(t_{k+1}) = x(t_k) + \phi(x(t_k), t_k)\Delta t.$$
$$\frac{x(t_{k+1}) - x(t_k)}{\Delta t} \approx \phi(x(t_{k+1}), t_{k+1}) \Rightarrow$$
$$x(t_k) = x(t_{k+1}) - \phi(x(t_{k+1}), t_{k+1})\Delta t.$$

$$k_1(t_k) \approx k_1(t_{k+1}) - \Big[k_2^2(t_{k+1})\Big]\Delta t$$
$$k_2(t_k) \approx k_2(t_{k+1}) - \Big[-k_1(t_{k+1}) + k_2(t_{k+1})k_3(t_{k+1})\Big]\Delta t$$
$$k_3(t_k) \approx k_2(t_{k+1}) - \Big[-2k_2(t_{k+1}) + k_3^2(t_{k+1})\Big]\Delta t.$$

7/8 动态规划求解最优控制：状态-时间



8/8 动态规划求解最优控制：控制-时间

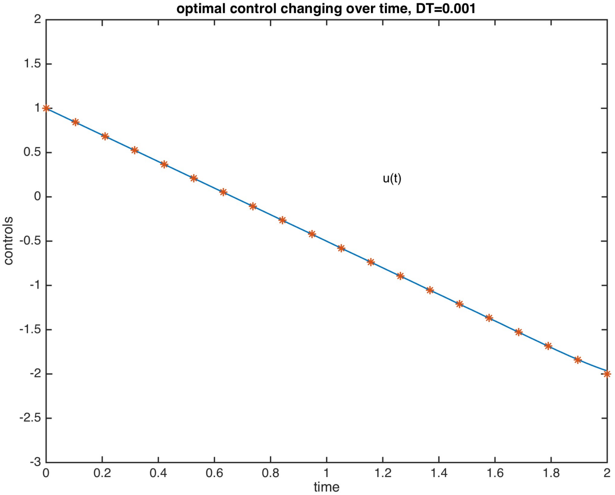


Table of Contents

离散时间线性二次性最优控制

问题 2 (离散时间线性二次型最优控制)

状态变量 $x(k) : \mathbb{N} \to \mathbb{R}^n$，控制变量 $u(k) : \mathbb{N} \to \mathbb{R}^m$。状态方程

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad k = 0, 1, \ldots, N-1.$$

最小化性能指标

$$J(u) = \frac{1}{2}x^{\mathrm{T}}(N)Hx(N) + \frac{1}{2}\sum_{k=0}^{N-1}[x^{\mathrm{T}}(k)Q(k)x(k) + u^{\mathrm{T}}(k)R(k)u(k)].$$

H, Q 半正定，R 正定。固定终端时刻 N，自由终端状态

离散时间线性二次型最优控制

简洁起见，下将 $x(k), u(k), A(k), B(k), Q(k), R(k)$ 简记为 $x_k, u_k, A_k, B_k, Q_k, R_k$。则状态方程可表示为

$$x_{k+1} = A_k x_k + B_k u_k.$$

性能指标可表示为

$$J(u) = \frac{1}{2} x_N^{\mathrm{T}} H x_N + \frac{1}{2} \sum_{k=0}^{N-1} [x_k^{\mathrm{T}} Q_k x_k + u_k^{\mathrm{T}} R_k u_k].$$

性能指标分析

考虑 $m = n = 1$ 情况，H, Q, R 都是实数

$$J = \frac{1}{2} x_N^T H x_N + \frac{1}{2} \sum_{k=0}^{N-1} [x_k^T Q_k x_k + u_k^T R_k u_k].$$

- $\frac{1}{2} x_N^T H x_N$
  终止损失，H 对称半正定，取值越大则终止状态越接近原点
- $\frac{1}{2} x_k^T Q_k x_k$
  过程损失，Q_k 对称半正定，取值越大则状态尽早接近原点
- $\frac{1}{2} u_k^T R_k u_k$
  控制损失，R_k 对称正定，取值越大则能量损耗越小

1/4 N

$$V(x_N, N) = \frac{1}{2} x_N^T H x_N. \tag{29}$$

记

$$K_N = H \tag{30}$$

2/4 $k = N - 1$ 最优控制

$$
\begin{aligned}
& V(x_{N-1}, N-1) \\
&= \min_{u_{N-1}} \left\{ \frac{1}{2} [x_{N-1}^{\mathrm{T}} Q_{N-1} x_{N-1} + u_{N-1}^{\mathrm{T}} R_{N-1} u_{N-1}] + V(x_N, N) \right\} \\
&= \min_{u_{N-1}} \frac{1}{2} \left\{ x_{N-1}^{\mathrm{T}} Q_{N-1} x_{N-1} + u_{N-1}^{\mathrm{T}} R_{N-1} u_{N-1} + x_N^{\mathrm{T}} K_N x_N \right\} \\
&= \min_{u_{N-1}} \frac{1}{2} \Big\{ x_{N-1}^{\mathrm{T}} Q_{N-1} x_{N-1} + u_{N-1}^{\mathrm{T}} R_{N-1} u_{N-1} \\
&\qquad + [A_{N-1} x_{N-1} + B_{N-1} u_{N-1}]^{\mathrm{T}} K_N [A_{N-1} x_{N-1} + B_{N-1} u_{N-1}] \Big\}.
\end{aligned}
$$

$$
\begin{aligned}
0 &= u_{N-1}^{\mathrm{T}} R_{N-1} + [A_{N-1} x_{N-1} + B_{N-1} u_{N-1}]^{\mathrm{T}} K_N B_{N-1}, \\
u_{N-1} &= -[R_{N-1} + B_{N-1}^{\mathrm{T}} K_N B_{N-1}]^{-1} B_{N-1}^{\mathrm{T}} K_N A_{N-1} x_{N-1} \\
&= F_{N-1} x_{N-1}, \\
F_{N-1} &= -[R_{N-1} + B_{N-1}^{\mathrm{T}} K_N B_{N-1}]^{-1} B_{N-1}^{\mathrm{T}} K_N A_{N-1}.
\end{aligned}
$$

3/4 $k = N - 1$ 值函数

将 $k = N - 1$ 时刻最优控制代入，值函数为

$$V(x_{N-1}, N-1) = \frac{1}{2} x_{N-1}^{\mathrm{T}} \Big\{ Q_{N-1} + F_{N-1}^{\mathrm{T}} R_{N-1} F_{N-1} + [A_{N-1} + B_{N-1} F_{N-1}]^{\mathrm{T}} K_N [A_{N-1} + B_{N-1} F_{N-1}] \Big\} x_{N}$$

令

$$K_{N-1} = Q_{N-1} + F_{N-1}^{\mathrm{T}} R_{N-1} F_{N-1} + [A_{N-1} + B_{N-1} F_{N-1}]^{\mathrm{T}} K_N [A_{N-1} + B_{N-1} F_{N-1}].$$

得到

$$V(x_{N-1}, N-1) = \frac{1}{2} x_{N-1}^{\mathrm{T}} K_{N-1} x_{N-1}.$$

4/4 倒推求解最优控制

$$V(x(N), N) = \frac{1}{2} x^{\mathrm{T}}(N) K(N) x(N). \tag{31}$$

对 $k = N-1, \ldots, 0$,

$$u(k) = F(k) x(k), \tag{32}$$

$$V(x(k), k) = \frac{1}{2} x^{\mathrm{T}}(k) K(k) x(k). \tag{33}$$

其中

$$K(N) = H, \tag{34}$$

$$F(k) = -[R(k) + B^{\mathrm{T}}(k) K(k+1) B(k)]^{-1} B^{\mathrm{T}}(k) K(k+1) A(k), \tag{35}$$

$$\begin{aligned} K(k) = {} & Q(k) + F^{\mathrm{T}}(k) R(k) F(k) \\ & + [A(k) + B(k) F(k)]^{\mathrm{T}} K(k+1) [A(k) + B(k) F(k)]. \end{aligned} \tag{36}$$

例子：离散动态规划求解线性二次型

例 2 (离散动态规划求解线性二次型)

状态方程

$$\dot{x} = x + u \Rightarrow x(k+1) = x(k) + [x(k) + u(k)] * \Delta t. \tag{37}$$

最小化性能指标

$$J(u) = \frac{1}{2}x^2(N) + \sum_{k=0}^{N-1} \frac{1}{2}u^2(k)\Delta t. \tag{38}$$

离散动态规划求解线性二次型 1/5

将本例写成矩阵形式。状态方程为

$$x(k+1) = Ax(k) + Bu(k).$$

其中

$$A = 1 + \Delta t, \quad B = \Delta t.$$

性能指标为

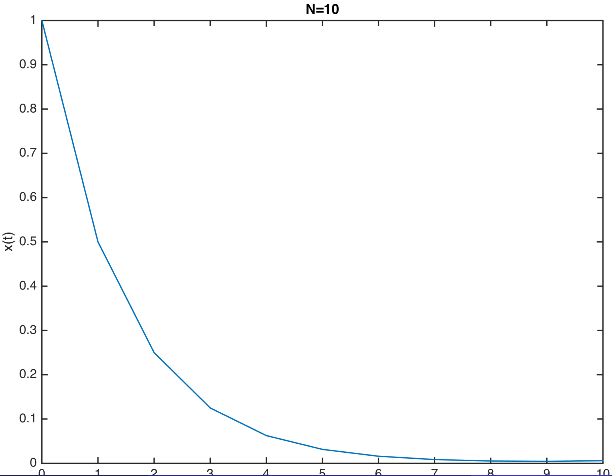
$$J(u) = \frac{1}{2}x^{\mathrm{T}}(N)Hx(N) + \frac{1}{2}\sum_{k=0}^{N-1}[x^{\mathrm{T}}(k)Qx(k) + u^{\mathrm{T}}(k)Ru(k)], \quad (39)$$

其中

$$H = 1, \quad Q = 0, \quad R = 1.$$

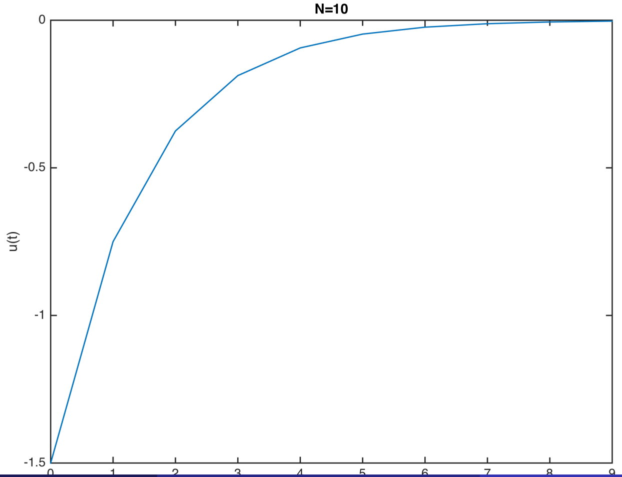
离散动态规划求解线性二次型 2/5

Figure: $\Delta t = 1, N = 10$, 状态变量



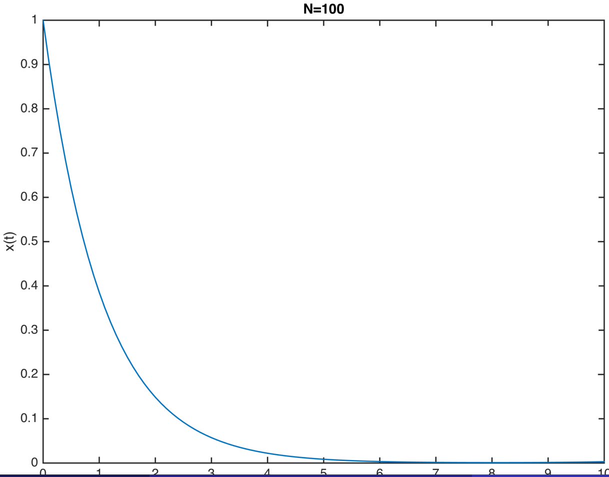
离散动态规划求解线性二次型 3/5

Figure: $\Delta t = 1, N = 10$, 控制变量



离散动态规划求解线性二次型 4/5

Figure: $\Delta t = 0.1, N = 100$, 状态变量



离散动态规划求解线性二次型 5/5

Figure: $\Delta t = 0.1, N = 100$, 控制变量

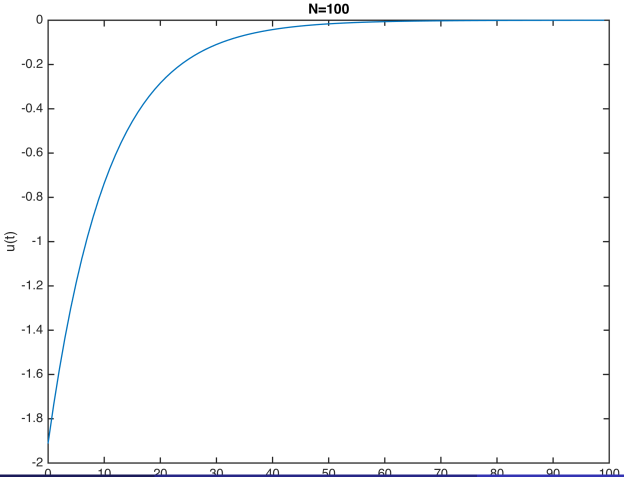


Table of Contents

线性二次型最优控制

问题 3 (线性二次型最优控制)

状态方程

$$\dot{x}(t) = A(t)x(t) + B(t)u(t). \tag{40}$$

最小化性能指标

$$J = \frac{1}{2}x^T(t_f)Hx(t_f) + \frac{1}{2}\int_{t_0}^{t_f} [x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]dt. \tag{41}$$

t_f *fixed*, $x(t_f)$ *free*
其中 H *和* $Q(t)$ *是实对称半正定矩阵,* $R(t)$ *是实对称正定矩阵*

性能指标分析

$m = n = 1$

$$J = \frac{1}{2}x^T(t_f)Hx(t_f) + \frac{1}{2}\int_{t_0}^{t_f}[x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]dt.$$

- $\frac{1}{2}x^T(t_f)Hx(t_f)$
  终止损失，H 对称半正定，取值越大则终止状态越接近原点
- $\frac{1}{2}x^T(t)Q(t)x(t)$
  过程损失，$Q(t)$ 对称半正定，取值越大则状态尽早接近原点
- $\frac{1}{2}u^T(t)R(t)u(t)$
  控制损失，$R(t)$ 对称正定，取值越大则能量损耗越小

连续 v.s. 离散

$$\begin{aligned}\dot{x}(t) &= A(t)x(t) + B(t)u(t).\\ x_{k+1} &\approx x_k + (A(t_k)x_k + B(t_k)u_k)\Delta t\\ &= (I + A(t_k)\Delta t)x_k + B(t_k)\Delta t u_k\end{aligned}$$

$$\begin{aligned}J &= \frac{1}{2}x^T(t_f)Hx(t_f) + \frac{1}{2}\int_{t_0}^{t_f}[x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]dt\\ &\approx \frac{1}{2}x_N^T H x_N + \frac{1}{2}\sum_{k=0}^{N-1}[x_k^T Q(t_k)x_k + u_k^T R(t_k)u_k]\Delta t\\ &= \frac{1}{2}x_N^T H x_N + \frac{1}{2}\sum_{k=0}^{N-1}[x_k^T Q(t_k)\Delta t x_k + u_k^T R(t_k)\Delta t u_k]\end{aligned}$$

1/5 计算 Hamiltonian, 考察极值条件

Hamiltonian

$$\mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t), t) = \frac{1}{2}x^T(t)Q(t)x(t) + \frac{1}{2}u^T(t)R(t)u(t) + [\frac{\partial V}{\partial x}(x(t), t)]^T[A(t)x(t) + B(t)u(t)] \tag{42}$$

一阶条件

$$0 = \frac{\partial \mathcal{H}}{\partial u}, \; 0 = R(t)u(t) + B^T(t)\frac{\partial V}{\partial x}(x(t), t). \tag{43}$$

二阶条件

$$\frac{\partial^2 \mathcal{H}}{\partial u^2} = R(t) \text{ 正定} \tag{44}$$

2/5 求得最优情况的 Hamiltonian

$$u(t) = -R^{-1}(t)B^T(t)\frac{\partial V}{\partial x}(x(t), t). \tag{45}$$

代入得到

$$\begin{aligned}
&\min_{u(t)\in\mathbb{R}^m} \mathcal{H}(x(t), u(t), \frac{\partial V}{\partial x}(x(t), t), t), t) \\
=&\frac{1}{2}x^{\mathrm{T}}(t)Q(t)x(t) + \frac{1}{2}[\frac{\partial V}{\partial x}(x(t), t)]^{\mathrm{T}}B(t)R^{-1}(t)B^{\mathrm{T}}(t)\frac{\partial V}{\partial x}(x(t), t) \\
&\quad + [\frac{\partial V}{\partial x}(x(t), t)]^{\mathrm{T}}[A(t)x(t) - B(t)R^{-1}(t)B^{\mathrm{T}}(t)\frac{\partial V}{\partial x}(x(t), t)] \\
=&\frac{1}{2}x^{\mathrm{T}}(t)Q(t)x(t) + [\frac{\partial V}{\partial x}(x(t), t)]^{\mathrm{T}}A(t)x(t) \\
&\quad - \frac{1}{2}[\frac{\partial V}{\partial x}(x(t), t)]^{\mathrm{T}}B(t)R^{-1}(t)B^{\mathrm{T}}(t)\frac{\partial V}{\partial x}(x(t), t).
\end{aligned}$$

3/5 得到 HJB 方程

$$\frac{\partial V}{\partial t}(x(t),t)+\frac{1}{2}x^{\mathrm{T}}(t)Q(t)x(t)+[\frac{\partial V}{\partial x}(x(t),t)]^{\mathrm{T}}A(t)x(t) \\ -\frac{1}{2}[\frac{\partial V}{\partial x}(x(t),t)]^{\mathrm{T}}B(t)R^{-1}(t)B^{\mathrm{T}}(t)\frac{\partial V}{\partial x}(x(t),t)=0. \tag{46}$$

边界条件为

$$V(x(t_f),t_f)=\frac{1}{2}x^{\mathrm{T}}(t_f)Hx(t_f). \tag{47}$$

假定 V 是二次,

$$V(x(t),t)=\frac{1}{2}x^{\mathrm{T}}(t)K(t)x(t),$$

$$\frac{\partial V}{\partial x}=K(t)x(t),\quad \frac{\partial V}{\partial t}=\frac{1}{2}x^{T}(t)\dot{K}(t)x(t)$$

4/5 求解 HJB 方程

$$\frac{\partial V}{\partial t}(x(t),t)+\frac{1}{2}x^{\mathrm{T}}(t)Q(t)x(t)+[\frac{\partial V}{\partial x}(x(t),t)]^{\mathrm{T}}A(t)x(t) \\ -\frac{1}{2}[\frac{\partial V}{\partial x}(x(t),t)]^{\mathrm{T}}B(t)R^{-1}(t)B^{\mathrm{T}}(t)\frac{\partial V}{\partial x}(x(t),t)=0.$$

$$0=\frac{1}{2}x^T\dot{K}x+\frac{1}{2}x^TQx+x^TKAx-\frac{1}{2}x^TK^TBR^{-1}B^TKx$$

转置得到

$$0=\frac{1}{2}x^T\dot{K}x+\frac{1}{2}x^TQx+x^TA^TKx-\frac{1}{2}x^TK^TBR^{-1}B^TKx$$

两式加和,

连续动态规划求解线性二次型 5/5

闭环形式的最优控制满足 Riccati 微分方程

$$
\begin{aligned}
0 &= \dot{K}(t) + Q(t) - K(t)B(t)R^{-1}(t)B^T(t)K(t) \\
&\quad + K(t)A(t) + A^T(t)K(t) && (48) \\
K(t_f) &= H && (49) \\
u(t) &= -R^{-1}(t)B^T(t)K(t)x(t). && (50)
\end{aligned}
$$

例子：连续动态规划求解线性二次型

例 3 (连续动态规划求解线性二次型)

状态方程

$$\dot{x}_1 = x_2 \tag{51}$$

$$\dot{x}_2 = x_2 + u \tag{52}$$

$t_0 = 0, t_f = 10$，初始状态 $(1, 1)$。最小化性能指标

$$J(u) = \frac{h}{2}x_2^2(t_f) + \int_{t_0}^{t_f} [\frac{q}{2}x_1^2(t) + \frac{r}{2}u^2(t)]dt \tag{53}$$

其中 $q = 1, r = 3, h = 4$

连续动态规划求解线性二次型 1/4

解：（求解常微分方程）

$$
\begin{aligned}
0 &= \dot{K}(t) + Q - K(t)BR^{-1}B^TK(t) + K(t)A + A^TK(t) \quad (54)\\
K(t_f) &= H \quad (55)\\
u(t) &= -R^{-1}B^TK(t)x(t). \quad (56)
\end{aligned}
$$

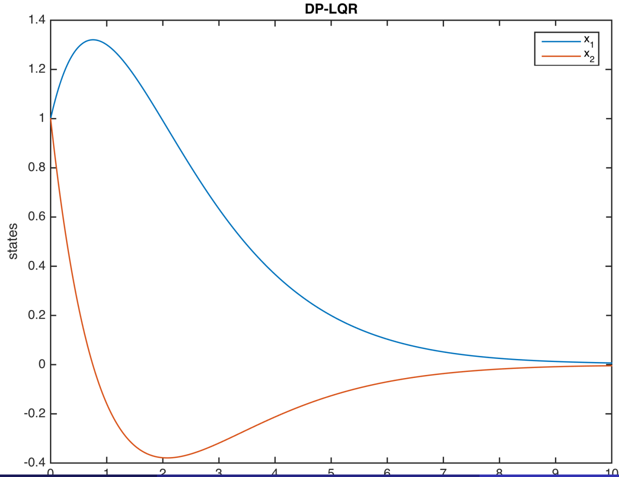
其中

$$
A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (57)
$$

$$
H = \begin{bmatrix} 0 & 0 \\ 0 & h \end{bmatrix},\ Q = \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix},\ R = r \quad (58)
$$

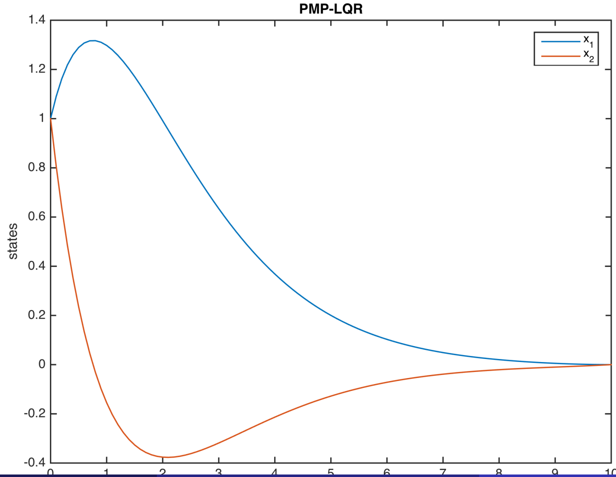
连续动态规划求解线性二次型 2/4

Figure: 连续动态规划求解, $h = 4, q = 1, r = 3$



固定终值的最优状态轨迹 3/4

Figure: 极值原理求解，固定终值 $x_2(t_f) = 0, q = 1, r = 3$



较小终止损失的最优状态轨迹 4/4

Figure: 连续动态规划求解, $h = 0.1, q = 1, r = 3$

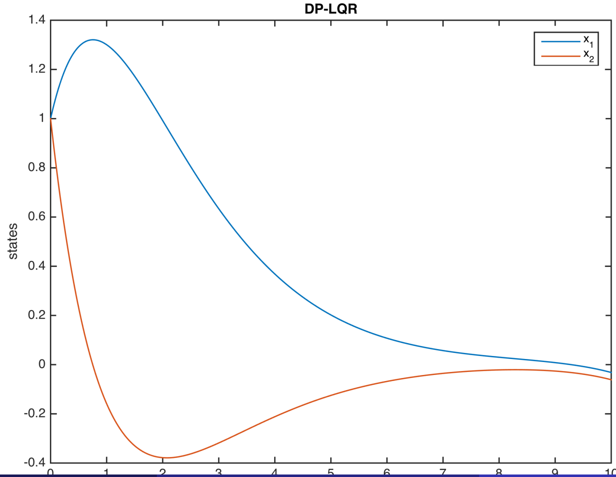


Table of Contents

动态规划求解最优控制的过程

Remark 1 (动态规划求解离散最优控制)

- 若非离散问题，首先将最优控制问题离散化
- 从终止时刻开始倒向计算值函数和闭环最优控制
- 在任意时刻，对给定的状态查表即得最优控制

Remark 2 (动态规划求解连续最优控制)

- 使用惩罚函数法将终值条件转化至目标函数中
- 求 Hamiltonian 极值情况下最优控制 (关于 x, V_x)
- 获得 HJB 方程，并求解
- 得到与 HJB 方程解有关的闭环形式最优控制

动态规划求解最优控制的缺陷

Remark 3 (动态规划求解最优控制的缺陷)

- 离散化模型面临维数灾难
- HJB 方程一般难以求解
- HJB 方程对值函数有可微的要求

动态规划 v.s. 极值原理

	离散 DP	HJB	PMP
必要条件	√	√	√
充分条件	√	√	×
闭环形式	√	√	×
约束条件	极易处理	√	√
值函数不可微	√	×	√
空间复杂度	至少 $O(s^m)$	-	-
计算复杂度	$O(Ns^m)$ 次求极值	解 PDE	解 ODE

HJB 方程和 PMP、经典变分的关系

$\boxed{\text{拉格朗日变分法}}$

$\Updownarrow$ 等价

$\boxed{\text{哈密尔顿方程组}} \quad \Leftrightarrow \quad \boxed{\text{哈密尔顿雅各比方程}}$

$\Downarrow$ + 控制 $\qquad\qquad$ $\Downarrow$ + 控制

$\boxed{\text{极值原理}} \Leftarrow V$ 二次可微特况* $\boxed{\text{HJB 方程}}$