第八讲：动态规划

最优控制的数学理论之四

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 12 日

Table of Contents

1. 回顾：变分法、极值原理求解最优控制
2. 动态规划方法
3. 动态规划求解离散最优控制
4. 离散时间线性二次性最优控制

下周四，10 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

1 回顾：变分法、极值原理求解最优控制

2 动态规划方法

3 动态规划求解离散最优控制

4 离散时间线性二次性最优控制

最优控制问题

问题 1 (最优控制问题)

1. 被控对象的状态方程为

$$\dot{x}(t) = f(x(t), u(t), t), \quad x(t_0) = x_0.$$

2. 容许控制, $u \in \mathcal{U}, \quad x \in \mathcal{X}$.
3. 目标集, $x(t_f) \in \mathcal{S}$

$$\mathcal{S} = [t_0, \infty) \times \{x(t_f) \in \mathbb{R}^n : m(x(t_f), t_f) = 0\}$$

4. 求分段连续的 u，以最小化性能指标

$$J(u) = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)\,\mathrm{d}t.$$

Pontryagin 极小值原理

定理 1 (庞特里亚金极小值原理)

上述问题得到最优控制 $u(t)$ 的必要条件为（TPBVP）

- 极值条件：对任意容许控制 $u'(t)$

$$\mathcal{H}(x(t),u(t),p(t),t) \le \mathcal{H}(x(t),u'(t),p(t),t).$$

- 规范方程：

$$\text{状态 (state) 方程:}\quad \dot{x}(t) = +\frac{\partial \mathcal{H}}{\partial p}(x(t),u(t),p(t),t),$$

$$\text{协态 (costate) 方程:}\quad \dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial x}(x(t),u(t),p(t),t).$$

- 边界条件 (用于处理目标集)：

$$[\frac{\partial h}{\partial x}(x(t_f),t_f) - p(t_f)] \cdot \delta x_f$$

$$+ [\mathcal{H}(x(t_f),u(t_f),p(t_f),t_f) + \frac{\partial h}{\partial t}(x(t_f),t_f)]\delta t_f = 0.$$

极值原理求解最优控制的过程

- 构造 Hamiltonian，求容许控制的极值条件，以协态和状态表示最优控制
- 最优控制代入规范方程，得到关于最优状态、协态的常微分方程组
- 根据边界条件和初值获得微分方程组的边界条件
- 直接求解或使用数值方法求解两点边值问题

开环控制

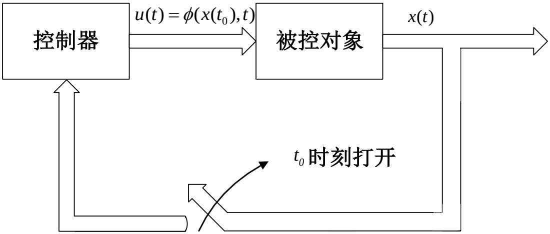
控制的形式：开环控制

定义 1 (开环控制)

若控制律形如

$$u(t) = \phi(x(t_0), t), \tag{1}$$

称之为开环控制



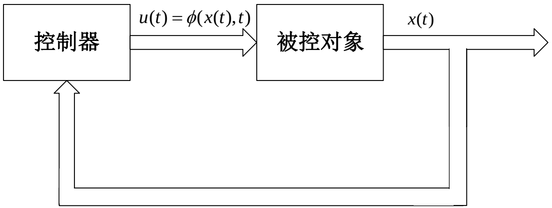
控制的形式：闭环控制

定义 2 (闭环控制)

若控制律形如

$$u(t) = \phi(x(t), t), \tag{2}$$

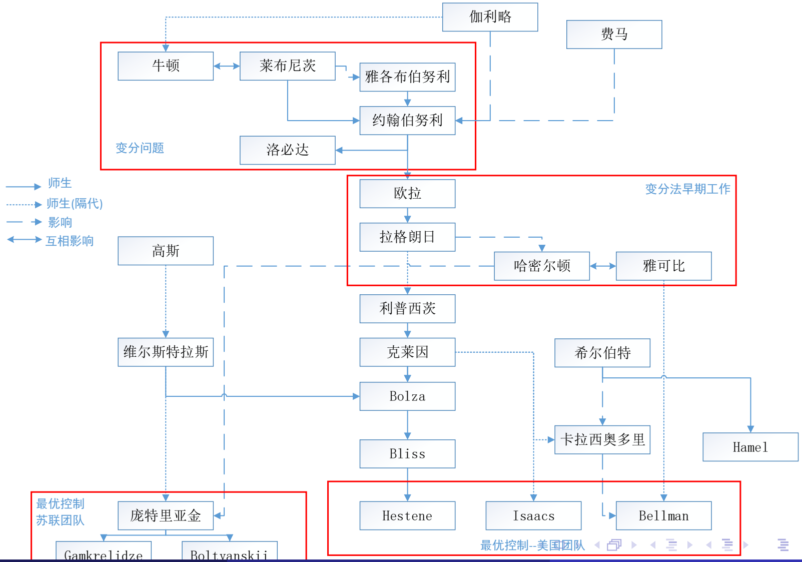
称之为闭环控制



课程内容

- 最优控制的数学理论
  - 经典变分法
  - 庞特里亚金极值原理
  - 动态规划方法
  - 微分博弈
- 最优控制的智能方法
  - 强化学习与自适应动态规划
  - 模型预测控制
  - 模糊控制
  - 平行控制与平行学习

变分法和最优控制的重要人物



庞特里亚金极小值原理

- 第三讲，提出庞特里亚金极小值原理（PMP）
- 第五讲，p47-50，化欧拉方程为没有控制的 PMP
- 第六讲，利用变分法求得开集上的 PMP（一阶条件）
- 第七讲，简要证明了非连续、非开集上的 PMP（极值条件）
- 动态规划方法也可在特殊情况下得到庞特里亚金极小值原理！

Table of Contents

1. 回顾：变分法、极值原理求解最优控制
2. 动态规划方法
3. 动态规划求解离散最优控制
4. 离散时间线性二次性最优控制

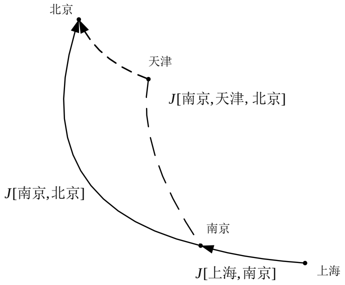
动态规划方法的发展

- 1944 年，冯诺依曼在名著 Theory of Games and Economic Behavior 中使用倒推法 (backward induction) 解决博弈问题，其与 Wald, Arrow 的序贯分析法被认为是动态规划方法的前身
- 1951 年，Rufus Isaacs 首次使用动态规划类似的方法求解微分博弈问题，但当时并未受到重视
- 自 1952 年起，Richard Bellman 等提出可用于离散最优控制的动态规划方法，60 年代得到连续情况下的 Hamilton-Jacobi-Bellman 方程
- 1983 年，Crandall 和 Lions 给出粘性解的概念，完善了 HJB 方程值函数非光滑情况的数学基础

动态规划的最优性原理

定理 2 (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

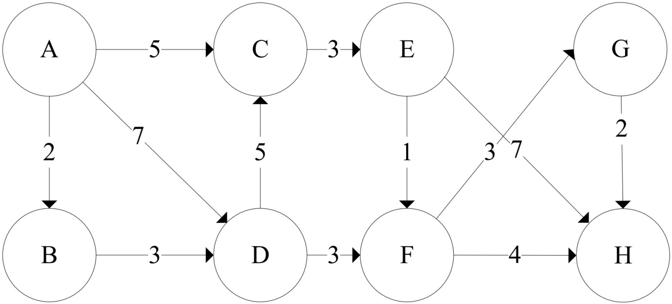


如果南京-天津-北京是南京到北京的最短路，
上海-南京-北京会是最短路吗

一个例子：动态规划求解最短路

例 1 (动态规划求解最短路)

求从图中 A 点到 H 的最短路

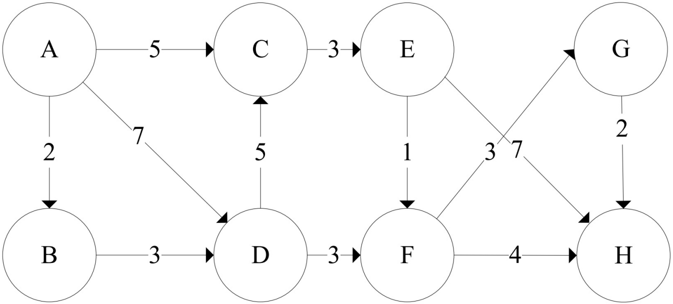


- $V(x)$ 表示从 x 出发到 H 的最短距离——值函数
- $\phi(x)$ 表示位于 x 应走到哪个点——闭环形式最优控制

$1/7\ G \to H$

- 此时已知 $V(H) = 0$。E, F, G 能直达 H
- G 可选控制 H

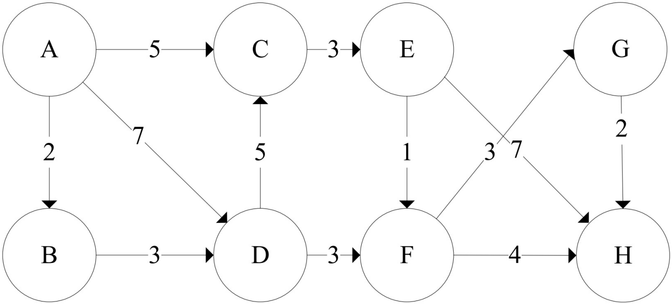
$$V(G) = J[GH] + V[H] = 2 + 0 = 2,$$
$$\phi(G) = H.$$



2/7 $F \to H$

- 此时已知 $V(H)=0, V(G)=2$
- F 可选控制 G, H

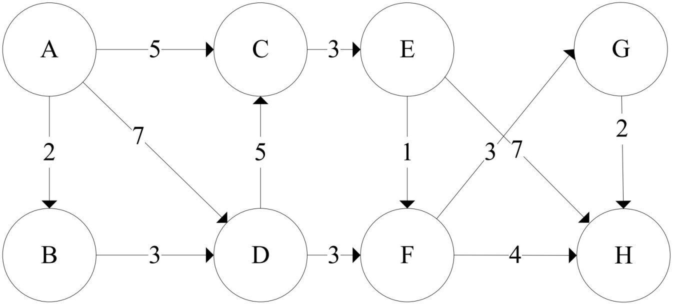
$$V(F)=\min\{J[FG]+V[G], J[FH]+V[H]\}=\min\{3+2, 4+0\}=4.$$
$$\phi(F)=H.$$



$3/7\ E \to H$

- 此时已知 $V(H)=0, V(G)=2, V(F)=4$
- E 可选控制 F, H

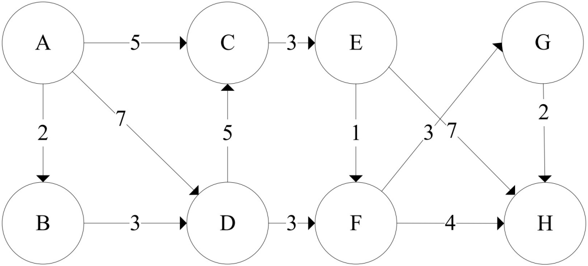
$$V(E)=\min\{J[EF]+V[F], J[EH]+V[H]\}=\min\{1+4, 7+0\}=5.$$
$$\phi(E)=F.$$



$4/7\ D \to H, C \to H$

- $V(H) = 0, V(G) = 2, V(F) = 4, V(E) = 5$。$C, D$ 直达
- D 可选控制 C, F，但 $V(C)$ 未知，暂存
- C 可选控制 E

$$\begin{aligned}&V(C) = J[CE] + V(E) = 3 + 5 = 8.\\&\phi(C) = E.\end{aligned}$$

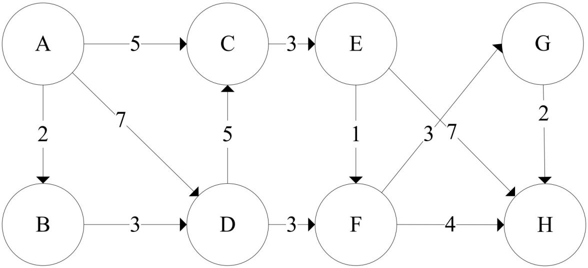


5/7 $D \to H$

- $V(H)=0, V(G)=2, V(F)=4, V(E)=5, V(C)=8$。$A, D$ 直达
- D 可选控制 C, F

$$V(D)=\min\{J[DC]+V(C), J[DF]+V[F]\}=\min\{5+8, 3+4\}=7.$$
$$\phi(D)=F.$$

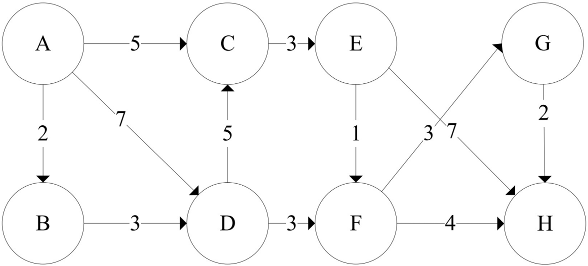
- A 可选控制 B, C, D，但 $V(B)$ 未知，暂存



$6/7\ B \rightarrow H$

- $V(H)=0, V(G)=2, V(F)=4, V(E)=5, V(D)=7, V(C)=8$。$A, B$ 直达
- B 可选控制 D

$$V(B) = J[BD] + V(D) = 3 + 7 = 10.$$
$$\phi(B) = D.$$



$7/7 \; B \to H$

- $V(H) = 0, V(G) = 2, V(F) = 4, V(E) = 5, V(D) = 7, V(C) = 8, V(B) = 10$
- A 可选控制 B, C, D

$$
\begin{aligned}
V(A) &= \min\{J[AB] + V(B), J[AC] + V(C), J[AD] + V(D)\} \\
&= \min\{2 + 10, 5 + 8, 7 + 7\} = 12.
\end{aligned}
$$

$$\phi(A) = B.$$

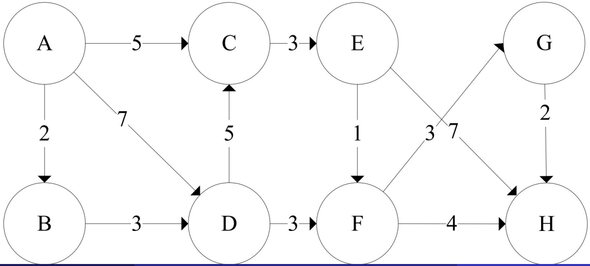


Table of Contents

1 回顾：变分法、极值原理求解最优控制

2 动态规划方法

3 动态规划求解离散最优控制

4 离散时间线性二次性最优控制

离散时间最优控制问题

问题 2 (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*

$$x(k+1) = f_D(x(k), u(k), k), \quad k = 0, \ldots, N-1. \tag{3}$$

(2) *容许控制:*

$$u(k) \in U, \quad x(k) \in X. \tag{4}$$

(3) *目标集:*

$$x(N) \in \mathcal{S}. \tag{5}$$

(4) *性能指标:*

$$J(u; x(k), k) = h_D(x(N), N) + \sum_{i=k}^{N-1} g_D(x(i), u(i), i). \tag{6}$$

（例）最优控制的离散化：时间

$$t \in [t_0, t_f]. \tag{7}$$

$$\Downarrow$$

$$t_0, t_1, t_2, \ldots, t_N = t_f$$

$$\Downarrow$$

记作

$$k = 0, 1, \ldots, N-1. \qquad N \text{ 为终止时刻} \tag{8}$$

例如，$t_0, t_1, \ldots, t_N$ 等分区间 $[t_0, t_f]$

（例）最优控制的离散化：状态方程

$$\dot{x}(t) = f(x(t), u(t), t),\ x(t_0) = x_0. \tag{9}$$

$$\Downarrow$$

$$x(t_{k+1}) \approx x(t_k) + f(x(t_k), u(t_k), t_k)\Delta t$$

$$\Downarrow$$

记作

$$x(k+1) = f_D(x(k), u(k), k) \tag{10}$$

（例）最优控制的离散化：性能指标

$$J = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)dt. \tag{11}$$

$$\Downarrow$$

$$J = h(x(t_f), t_f) + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} g(x(t), u(t), t)dt$$

$$\Downarrow$$

记作

$$J = h_D(x(N), N) + \sum_{k=0}^{N-1} g_D(x(k), u(k), k). \tag{12}$$

Bellman 方程

定理 3 (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”

$$V(x_0, k_0) = \min_{u \in U} J(u; x_0, k_0) \tag{13}$$

根据最优性原理，最优控制满足下列 Bellman 方程：

$$V(x(N), N) = h_D(x(N), N). \tag{14}$$

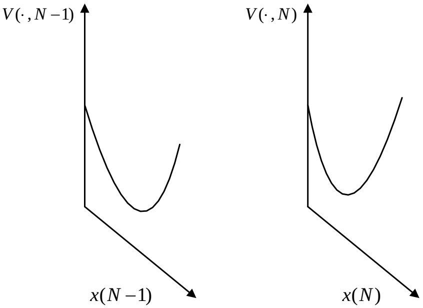
$$V(x(k), k) = \min_{u(k) \in U} \{ g_D(x(k), u(k), k) + V(x(k+1), k+1) \}, \quad k = N-1, \ldots, 0. \tag{15}$$

直接倒推求解 1/2

$$V(x(N),N)=h_D(x(N),N)$$

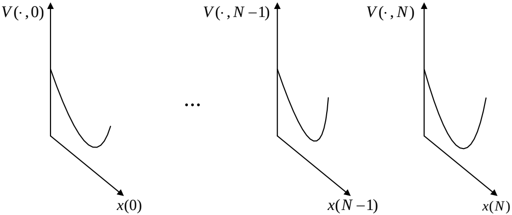
$$V(x(N-1),N-1)=\min_{u(N-1)\in U}\Big\{g_D(x(N-1),u(N-1),N-1)+V(x(N),N)\Big\}$$

$$=\min_{u(N-1)\in U}\Big\{g_D(x(N-1),u(N-1),N-1)+V(f_D(x(N-1),u(N-1)),N)\Big\}.$$



直接倒推求解 2/2

$$
\begin{aligned}
V(x(k),k) &= \min_{u(k)\in U}\left\{g_D(x(k),u(k),k)+V(x(k+1),k+1)\right\}\\
&= \min_{u(k)\in U}\left\{g_D(x(k),u(k),k)+V(f_D(x(k),u(k)),k+1)\right\}.
\end{aligned}
$$



例子：直接倒推求解

例 2 (直接倒推求解 Bellman 方程)

$$\min J = x^2(2) + \sum_{k=0}^{1}(x^2(k) + u^2(k)) \tag{16}$$

$$\text{s.t.} x(k+1) = x(k) + u(k),\ k = 0, 1 \tag{17}$$

1/4 终端时刻 $N=2$ 的值函数

直接由 Bellman 方程的边界条件即可得到对任意的 $x(2)\in\mathbb{R}$,

$$V(x(2),2)=h_D(x(2),2)=x^2(2).$$

2/4 $k=1$ 时刻值函数和最优控制

对任意的 $x(1)\in\mathbb{R}$,

$$\begin{aligned}V(x(1),1) &= \min_{u(1)\in\mathbb{R}}\left\{x^2(1)+u^2(1)+V(x(2),2)\right\}\\ &= \min_{u(1)\in\mathbb{R}}\left\{x^2(1)+u^2(1)+x^2(2)\right\}\\ &= \min_{u(1)\in\mathbb{R}}\left\{x^2(1)+u^2(1)+[x(1)+u(1)]^2\right\}.\end{aligned}$$

求得 $k=1$ 时刻值闭环形式最优控制和值函数

$$u(1) = -\frac{1}{2}x(1),$$

$$V(x(1),1) = x^2(1)+[-\frac{1}{2}x(1)]^2+[x(1)-\frac{1}{2}x(1)]^2 = \frac{3}{2}x^2(1).$$

3/4 求 $k=0$ 时刻值函数和最优控制

对任意的 $x(0)\in\mathbb{R}$,

$$
\begin{aligned}
V(x(0),0) &= \min_{u(0)\in\mathbb{R}}\left\{x^2(0)+u^2(0)+V(x(1),1)\right\}\\
&= \min_{u(0)\in\mathbb{R}}\left\{x^2(0)+u^2(0)+\frac{3}{2}x^2(1)\right\}\\
&= \min_{u(0)\in\mathbb{R}}\left\{x^2(0)+u^2(0)+\frac{3}{2}[x(0)+u(0)]^2\right\}.
\end{aligned}
$$

求得 $k=0$ 时刻闭环形式最优控制和值函数

$$
u(0)=-\frac{3}{5}x(0),
$$

$$
V(x(0),0)=x^2(0)+[-\frac{3}{5}x(0)]^2+\frac{3}{2}[x(0)-\frac{3}{5}x(0)]^2=\frac{8}{5}x^2(0).
$$

4/4 获得闭环形式的最优控制

获得闭环形式的最优控制,

$$u(0) = -\frac{3}{5}x(0), \tag{18}$$

$$u(1) = -\frac{1}{2}x(1), \tag{19}$$

以及值函数

$$V(x(0),0) = \frac{8}{5}x^2(0), \tag{20}$$

$$V(x(1),1) = \frac{3}{2}x^2(1), \tag{21}$$

$$V(x(2),2) = x^2(2). \tag{22}$$

直接倒推求解 Bellman 方程：有约束的情况

例 3 (直接倒推求解 Bellman 方程：有约束的情况)

$$\min J = x^2(2) + \sum_{k=0}^{1}(x^2(k) + u^2(k)) \tag{23}$$

$$\text{s.t.} x(k+1) = x(k) + u(k),\ k = 0, 1 \tag{24}$$

$$-1 \leq x(k) \leq 1,\ k = 0, 1 \tag{25}$$

$$-\frac{1}{2} \leq u(k) \leq \frac{1}{2},\ k = 0, 1. \tag{26}$$

1/4 终端时刻 $N=2$ 的值函数

由 Bellman 方程的边界条件，对 $-1 \le x(2) \le 1$，

$$V(x(2),2) = h_D(x(2),2) = x^2(2).$$

2/4 $k=1$ 时刻值函数和最优控制

对任意的 $-1 \leq x(1) \leq 1$，

$$\begin{aligned} V(x(1),1) &= \min_{-1/2 \leq u(1) \leq 1/2} \left\{ x^2(1) + u^2(1) + x^2(2) \right\} \\ &= \min_{-1/2 \leq u(1) \leq 1/2} \left\{ x^2(1) + u^2(1) + [x(1)+u(1)]^2 \right\}. \end{aligned}$$

计算有约束的条件极值，

$$\begin{aligned} u(1) &= -\frac{1}{2}x(1), \\ V(x(1),1) &= \frac{3}{2}x^2(1). \end{aligned}$$

3/4 $k=0$ 时刻值函数和最优控制

对任意的 $-1 \le x(0) \le 1$,

$$\begin{aligned} V(x(0),0) &= \min_{-1/2 \le u(0) \le 1/2} \left\{ x^2(0) + u^2(0) + \frac{3}{2} x^2(1) \right\} \\ &= \min_{-1/2 \le u(0) \le 1/2} \left\{ x^2(0) + u^2(0) + \frac{3}{2} [x(0) + u(0)]^2 \right\}. \end{aligned}$$

计算有约束的条件极值，对 $-5/6 \le x(0) \le 5/6$,

$$u(0) = -\frac{3}{5} x(0),\ V(x(0),0) = \frac{8}{5} x^2(0).$$

若 $5/6 < x(0) \le 1$， 则

$$u(0) = -\frac{1}{2},\ V(x(0),0) = \frac{5}{2} x^2(0) - \frac{3}{2} x(0) + \frac{5}{8}.$$

若 $-1 \le x(0) < -5/6$， 则

$$u(0) = \frac{1}{2},\ V(x(0),0) = \frac{5}{2} x^2(0) + \frac{3}{2} x(0) + \frac{5}{8}.$$

4/4 获得闭环形式的最优控制

获得闭环形式的最优控制，

$$
\begin{aligned}
u(0) &= \begin{cases} -1/2, & 5/6 < x(0) \leq 1, \\ -\dfrac{3}{5}x(0), & -5/6 \leq x(0) \leq 5/6, \\ +1/2, & -1 \leq x(0) < -5/6. \end{cases} \\
u(1) &= -\frac{1}{2}x(1).
\end{aligned} \tag{27}
$$

以及值函数

$$
\begin{aligned}
V(x(0),0) &= \begin{cases} \frac{5}{2}x^2(0) - \frac{3}{2}x(0) + \frac{5}{8}, & 5/6 < x(0) \leq 1, \\ \frac{8}{5}x(0)^2, & -5/6 \leq x(0) \leq 5/6, \\ \frac{5}{2}x^2(0) + \frac{3}{2}x(0) + \frac{5}{8}, & -1 \leq x(0) < -5/6. \end{cases} \\
V(x(1),1) &= \frac{3}{2}x^2(1). \\
V(x(2),2) &= x^2(2).
\end{aligned} \tag{28}
$$

遍历离散的状态和控制空间，查表法

求解 Bellman 方程的关键在于，在 k 时刻，需要已知 $V(\cdot, k+1)$，

$$V(x(k), k) = \min_{u(k) \in U} \{g_D(x(k), u(k), k) + V(x(k+1), k+1)\}$$

- 若 $x(k) \in \mathbb{R}^n$，值函数是函数，若无解析形式难存储
- 将状态空间离散化为有限取值近似，可建立表格 $V(\cdot, k+1)$

例子：遍历状态和控制 – 直接查表

例 4 (遍历状态和控制)

- 状态方程

$$\dot{x}(t) = u(t), x(0) = 1.5 \tag{29}$$

- 容许控制

$$0 \leq x(t) \leq 1.5, \ -1 \leq u(t) \leq 1. \tag{30}$$

- 最小化性能指标

$$J = x^2(2) + \int_0^2 2u(t)^2 dt. \tag{31}$$

1/5 转化为离散形式

$N = 2, \Delta t = T/N = 1, k = 0, 1$

- 离散状态方程

$$\dot{x} = u \;\Rightarrow\; x(k+1) = x(k) + u(k)\Delta t \tag{32}$$

- 离散化状态空间和容许控制

$$\begin{aligned} &0 \leq x \leq 1.5,\; -1 \leq u \leq 1 \Rightarrow \\ &x(k) \in \{0, 0.5, 1, 1.5\},\; u(k) \in \{-1, -0.5, 0, 0.5, 1\} \end{aligned} \tag{33}$$

- 离散性能指标

$$J(u) = x^2(2) + \int_0^2 2u(t)^2 dt \Rightarrow J(u; 1.5, 0) = x^2(2) + \sum_{k=0}^{1} 2u(k)^2 \tag{34}$$

2/5 求终止时刻 $N=2$ 的值函数

$$J(x(2),2)=x^2(2)$$

$x(2)$	$V(x(2),2)=x^2(2)$
0	0
0.5	0.25
1	1
1.5	2.25

3/5 $k=1$, 状态 $x(1)=1.5$

时刻 $k=1$, 状态 $x(1)=1.5$ 对每个可能控制求性能指标

$x(1)$	$u(1)$	$x(2)$	$J=g_D+V(x(2),2)$	$V(x(1),1)$	$\phi(x(1),1)$
1.5	1	2.5($\times$)		1.5	-0.5
	0.5	2($\times$)			
	0	1.5	$2(0)^2+2.25=2.25$		
	-0.5	1	$2(-0.5)^2+1=1.5$		
	-1	0.5	$2(-1)^2+0.25=2.25$		

4/5 $k=1$, 状态 $x(1)=1,0.5,0$

时刻 $k=1$，遍历其余状态，遍历容许控制，已舍掉不容许的 $x(2)$

$x(1)$	$u(1)$	$x(2)$	$J=g_D+V(x(2),2)$	$V(x(1),1)$	$\phi(x(1),1)$
1	0.5	1.5	$2(0.5)^2+2.25=2.75$	0.75	-0.5
	0	1	$2(0)^2+1=1$		
	-0.5	0.5	$2(-0.5)^2+0.25=0.75$		
	-1	0	$2(-1)^2+0=2$		
0.5	1.	1.5	$2(1)^2+2.25=4.25$	0.25	0
	0.5	1	$2(0.5)^2+1=1.5$		
	0	0.5	$2(0)^2+0.25=0.25$		
	-0.5	0	$2(-0.5)^2+0=0.5$		
0	1	1	$2(1)^2+2=3$	0	0
	0.5	0.5	$2(0.5)^2+0.25=0.75$		
	0	0	$2(0)^2+0=0$		

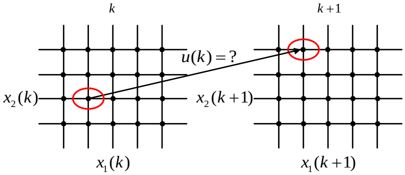
5/5 $k = 0$, 状态 $x(0) = 1.5$

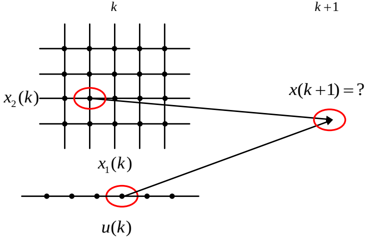
$k = 0$ 时刻，只需求 $x(0) = 1.5$ 对应的每个可能控制求性能指标

$x(0)$	$u(0)$	$x(1)$	$J = g_D + V(x(1), 1)$	$V(x(0), 0)$	$\phi(x(0), 0)$
1.5	0	1.5	$2(0)^2 + 1.5 = 1.5$	1.25	-0.5
	-0.5	1	$2(-0.5)^2 + 0.75 = 1.25$		
	-1	0.5	$2(-1)^2 + 0.25 = 2.25$		

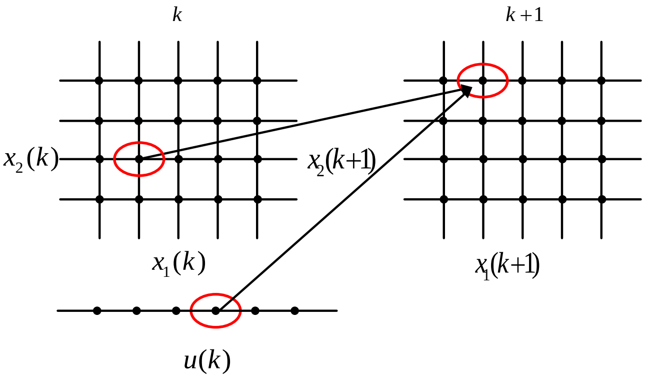
得到，对于初值 $k = 0, x(0) = 1.5$，最优控制 $u(0) = \phi(x(0), 0) = -0.5$，性能指标为 1.25。继而 $x(1) = 1$，再查表得 $u(1) = \phi(x(1), 1) = -0.5$

查表法实现细节





近似查表



遍历状态和控制 – 插值计算 1/2

Remark 1 (插值计算，遍历状态和控制)

并不直接查表或寻找最近的 $x(k+1)$，而是先对 $V(x,k+1)$ 进行插值近似，得到 $\bar{V}(x,k+1)$，使用 $\bar{V}(x(k+1),k+1)$ 近似 $V(x(k+1),k+1)$

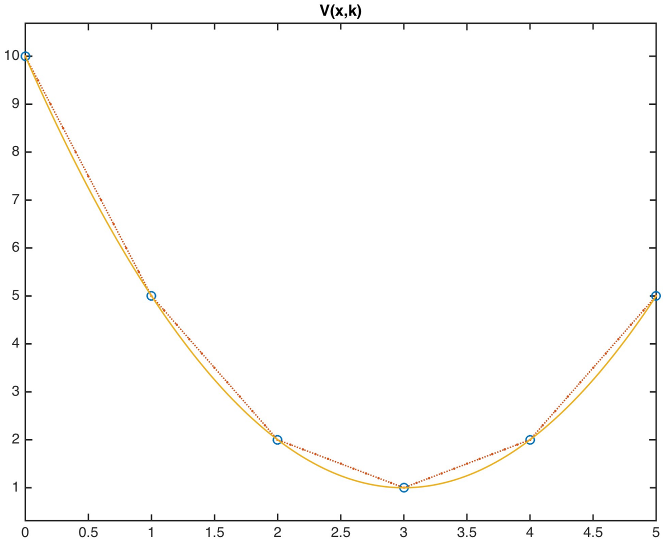
例 5 (线性内插值)

利用经过 (x_1,y_1),(x_2,y_2) 两点的直线

$$\frac{x-x_1}{x_1-x_2}=\frac{y-y_1}{y_1-y_2} \tag{35}$$

由 x 求出 y

遍历状态和控制 – 插值计算 2/2



维数灾难

存储值函数需要 $s^n \times N$ 条数据

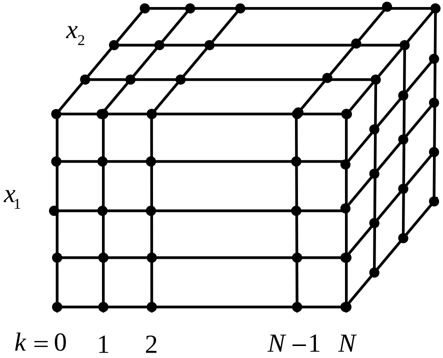


Table of Contents

1 回顾：变分法、极值原理求解最优控制

2 动态规划方法

3 动态规划求解离散最优控制

4 离散时间线性二次性最优控制

离散时间线性二次性最优控制

问题 3 (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

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

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

$$V(x_N, N) = \frac{1}{2} x_N^T H x_N. \tag{36}$$

记

$$P_N = H \tag{37}$$

2/4 $k = N - 1$ 最优控制

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

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

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

$$V(x(N),N)=\frac{1}{2}x^{\mathrm{T}}(N)K(N)x(N). \tag{38}$$

对 $k=N-1,\dots,0,$

$$u(k)=F(k)x(k), \tag{39}$$

$$V(x(k),k)=\frac{1}{2}x^{\mathrm{T}}(k)K(k)x(k). \tag{40}$$

其中

$$K(N)=H, \tag{41}$$

$$F(k)=-[R(k)+B^{\mathrm{T}}(k)K(k+1)B(k)]^{-1}B^{\mathrm{T}}(k)K(k+1)A(k), \tag{42}$$

$$\begin{aligned}K(k)=&\,Q(k)+F^{\mathrm{T}}(k)R(k)F(k)\\&+[A(k)+B(k)F(k)]^{\mathrm{T}}K(k+1)[A(k)+B(k)F(k)].\end{aligned} \tag{43}$$

例子：离散动态规划求解线性二次型

例 6 (离散动态规划求解线性二次型)

状态方程

$$\dot{x} = x + u \Rightarrow x(k+1) = x(k) + [x(k) + u(k)] * \Delta t. \tag{44}$$

最小化性能指标

$$J(u) = \frac{1}{2}x^2(N) + \sum_{k=0}^{N-1} \frac{1}{2}u^2(k)\Delta t. \tag{45}$$

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

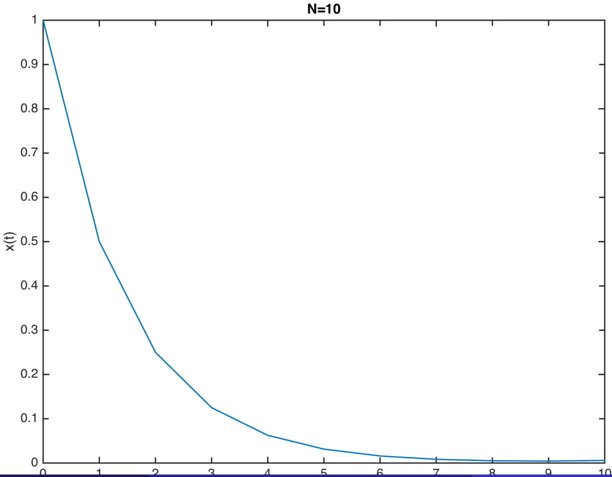
$$J(u) = \frac{1}{2}x^{\mathrm{T}}(N)Hx(N) + \frac{1}{2}\sum_{k=0}^{N-1}[x^{\mathrm{T}}(k)Qx(k) + u^{\mathrm{T}}(k)Ru(k)], \quad (46)$$

其中

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

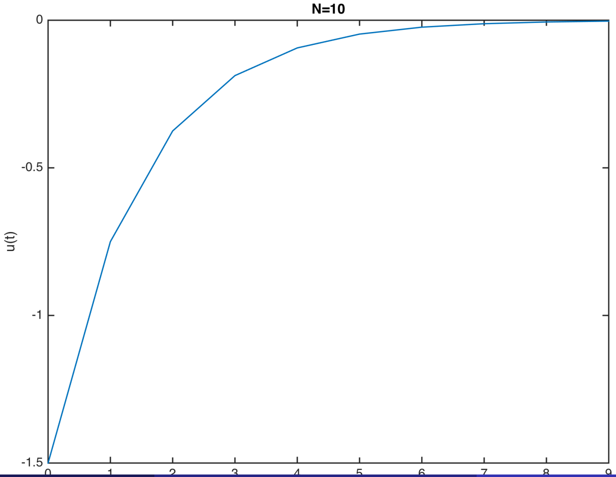
离散动态规划求解线性二次型 2/5

Figure: $\Delta t = 1, N = 10$, 状态变量



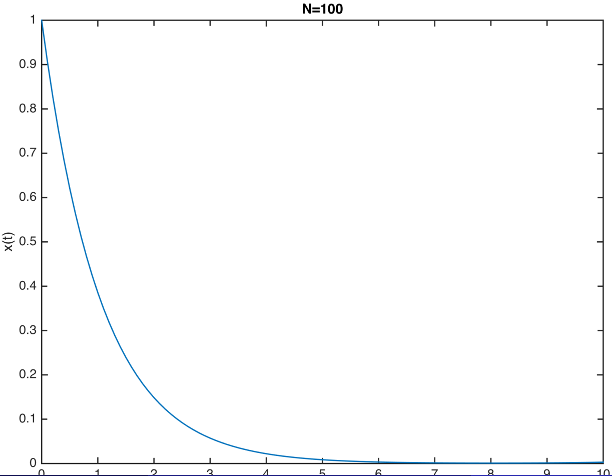
离散动态规划求解线性二次型 3/5

Figure: $\Delta t = 1, N = 10$, 控制变量



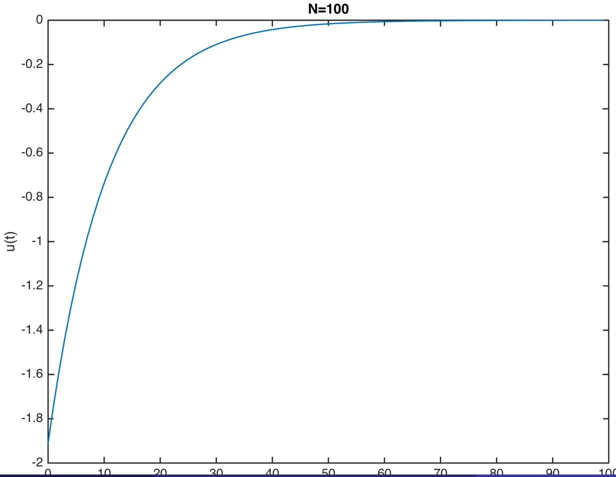
离散动态规划求解线性二次型 4/5

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



动态规划求解最优控制的缺陷

Remark 2 (动态规划求解最优控制的缺陷)

- 离散模型面临维数灾难

- 下次课：连续模型解析求解即可避免“近似”带来的维数灾难
- 自适应动态规划：无法解析求解的模型近似求解贝尔曼方程