最优控制理论与系统

第四章 动态规划

张永韡

江苏科技大学
电子信息学院
yongwzhang@gmail.com

November 23, 2016

本章主要内容

- 多级决策的例子——最短时间问题
- 最优性原理
- 用动态规划解资源分配问题
- 用动态规划求离散最优控制
- 连续系统的动态规划
- 动态规划与极小值原理
- 小结

动态规划

动态规划是贝尔曼 (Bellman) 在五十年代为解决多级决策过程而提出来的。它可以解决很多领域中的问题，如生产过程的决策，收益和投资问题，有多级反应器的化工装置的设计，多级轧钢机的最速轧制问题，资源分配、机器负荷分配、生产计划编制，特别是控制工程问题。

它和极小值原理一样，可解决控制变量受约束的最优控制问题，而且在这两种方法之间存在某种内在的联系。动态规划的中心思想是利用所谓“最优性原理”，把一个 N 级决策过程化为 N 个单级决策过程，从而使问题简单。

多级决策的例子——最短时间问题

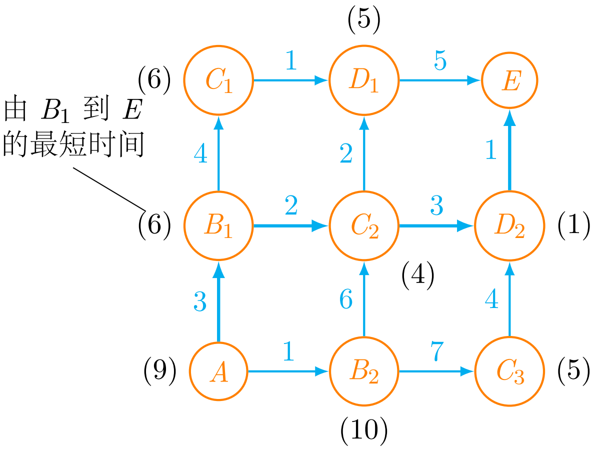

图 **4-1:** 按最短时间的路径选择

（一）穷举法

从 A 走到 E 一共有六条路线，每条路线由四段组成。这六条路线和对应的行车时间如下

路线	行车时间（小时）
$AB_2C_3D_2E$	13
$AB_2C_2D_2E$	11
$AB_2C_2D_1E$	14
$AB_1C_1D_1E$	13
$AB_1C_2D_1E$	12
$AB_1C_2D_2E$	9

（一）穷举法

- 显然最优路线是 $AB_1C_2D_2E$，它所花时间为 9 小时。
- 这里每条路线由四段组成，也可以说是四级决策。
- 为了计算每条路线所花时间，要做三次加法运算，为了计算六条路线所花的时间要作 3 6=18 次运算。这种方法称为“穷举法”。
- 显然当段数很多时，计算量是很大的。这种方法的特点是从起点站往前进行，而且把这四级决策一起考虑。应注意从 A 到下一站 B_2 所花的时间为 1，而到 B_1 所花时间为 3，但最优路线却不经过 B_2。
- 这说明只看下一步的“眼前利益”来作决策是没有意义的。

（二）动态规划法

为将问题表达得清楚，引进下面的术语。

- 令 n 表示由某点 E 到终点的段数（如 C_2 到 E 为 2 段）。
- 令 x 表示当前所处点的位置（如 A_1，B_1，C_1），称为状态变量。
- 令 $u_n(x)$ 为决策（控制）变量，它表示当处在 x 位置而还有 n 段要走时，所要选取的下一点。例如，从 C_2 出发，下一点为 D_2 时，则表示为 $u_2(C_2) = D_2$。
- 令 $T_n(x)$ 表示在位置 x，向终点还有 n 段要走时，由 x 到终点 E 的最短时间。例如，从 C_2 到 E 的最短时间为 4，可表示为 $T_2(C_2) = 4$。
- 令 $t_n(x, u_n)$ 表示从点 x 到点 $u_n(x)$ 的时间。例如，从 C_2 到 D_2 的时间为 $t_n(C_2, D_2) = 3$

有了这些术语后，就可用动态规划来解这个例子。从最后一段出发进行计算，并将计算出的最短时间 $T_n(x)$ 用括号表示在相应的点 x 处（见图4-1）。

1 $n = 1$（倒数第一段）

考虑从 D_1 和 D_2 到 E 的路线，由定义可知，最短时间分别为

$$T_1(D_1) = 5 \quad T_1(D_2) = 1$$

2 （倒数第二段）

考虑从 C_1、C_2 或 C_3 到 E 的路线。由 C_2 到 E 有两种路线：C_2D_1E，C_2D_2E。两种路线中的最短时间由下式确定：

$$T_2(C_2) = \min_{D_1,D_2} \begin{Bmatrix} t(C_2, D_1) + T_1(D_1) \\ t(C_2, D_2) + T_1(D_2) \end{Bmatrix} = \min \begin{Bmatrix} 2+5 \\ 3+1 \end{Bmatrix} = 4$$

最优决策为 $u_2(C_2) = D_2$。
由 C_1 到 E 只有一种路线 C_1D_1E，其时间为 $T_2(C_1) = 1 + 5 = 6$
由 C_3 到 E 也只有一种路线 C_3D_2E，其时间为 $T_2(C_3) = 4 + 1 = 5$

3（倒数第三段）

考虑从 B_1 或 B_2 到 E 的路线。B_1 到 E 有两种路线：B_1C_2E 和 B_1C_1E。最短时间为

$$T_3(B_1) = \min_{C_1,C_2} \begin{Bmatrix} t(B_1, C_1) + T_2(C_1) \\ t(B_1, C_2) + T_2(C_2) \end{Bmatrix} = \min \begin{Bmatrix} 4+6 \\ 2+4 \end{Bmatrix} = 6$$

最优决策 $u_3(D_1) = C_2$

从 B_2 到 E 有两种路线：B_2C_2E 和 B_2C_3E。最短时间为

$$T_3(B_2) = \min_{C_2,C_3} \begin{Bmatrix} t(B_2, C_3) + T_2(C_3) \\ t(B_2, C_2) + T_2(C_2) \end{Bmatrix} = \min \begin{Bmatrix} 7+5 \\ 6+4 \end{Bmatrix} = 10$$

最优决策为 $u_3(B_2) = C_2$。

4（倒数第四段）

从 A 到 E 的路线有两种：AB_1E 和 AB_2E。最短时间为：

$$T_4(A) = \min_{B_1,B_2} \begin{Bmatrix} t(A,B_1) + T_3(B_1) \\ t(A,B_2) + T_3(B_2) \end{Bmatrix} = \min \begin{Bmatrix} 3+6 \\ 1+10 \end{Bmatrix} = 9$$

最优决策为 $u_4(A) = B_1$。

至此求出了 A 到 E 的最短时间为 9，最优路线为 $AB_1C_2D_2E$。在图4-1中用粗线表示。这里，为决定最优路线进行了 10 次加法，比穷举法的 18 次少了 8 次。当段数 n 更多时，节省计算将会更多。

从上面解题过程可见，动态规划解题的两个特点：它是从最后一级往后倒着计算的；它把一个 N 级决策问题（这里是决定一整条路线）化为 N 个单级决策问题，即把一个复杂问题化为多个简单问题来求解。我们可看出 n 阶段与 $n-1$ 阶段有下面的关系（$n=1,2,\cdots,N$）

$$T_n(x) = \min_{u_n(x)}\{t[x, u_n(x)] + T_{n-1}[u_n(x)]\} \tag{4-1}$$

$T_1(x) = t_1(x, E)$ 表示最后一级

(4-1)式称为函数方程，从(4-1)式可见, 在选择了决策 $u_n(x)$ 后有两个影响，其一是直接影响下一段的时间（眼前利益），其二是影响以后 $n-1$ 段的最短时间 T_{n-1}（未来利益）。因此动态规划方法可以说是把眼前利益和未来利益区分开来又结合起来考虑的一种优化方法。这些特点都是由动态规划法的基本原理——最优性原理所决定的。

最优性原理

贝尔曼的最优性原理可叙述如下：

“一个多级决策问题的最优决策具有这样的性质：当把其中任何一级及其状态作为初始级和初始状态时，则不管初始状态是什么，达到这个初始状态的决策是什么，余下的决策对此初始状态必定构成最优策略。”

以上面的最短时间问题为例，如把 C_2 当作初始状态，则余下的决策 C_2D_2E 对 C_2 来讲是最优策略；如把 B_1 当初始状态，则余下的决策 $B_1C_2D_2E$ 对 B_1 来讲也构成最优策略。一般来说，如果一个最优过程用状态 $x_0, x_1, \cdots, x_N$ 来表示，最优决策为 $u_0, u_1, \cdots, u_{N-1}$，则对状态 x_k 来讲，$u_k, u_{k+1}, \cdots, u_{N-1}$ 必定是最优的，这可用图4-2表示。

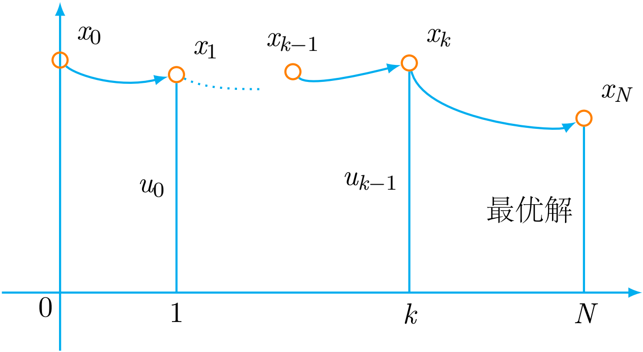

图 **4-2:** 最优性原理示意图

在多数实际问题中，N 级决策的性能指标 J 取如下形式

$$J_N = \sum_{k=0}^{N-1} F(x_k, u_k)$$

$F(\cdot,\cdot)$ 是由某级状态和决策决定的性能函数，要求寻找决策 $u_0, u_1, \cdots, u_{N-1}$ 使 J 取极小值 J_N^*。
最优性原理可表示为

$$\begin{aligned} J_N^* &= \min_{u_0,\cdots,u_{N-1}} \{F(x_0,u_0) + F(x_1,u_1) + \cdots + F(x_{N-1},u_{N-1})\} \\ &= \min_{u_0}\{F(x_0,u_0) + \min_{u_1,\cdots,u_{N-1}} [F(x_1,u_1) + \cdots + F(x_{N-1},u_{N-1})]\} \quad (4\text{-}2) \\ &= \min_{u_0}\{F(x_0,u_0) + J_{N-1}^*\} \end{aligned}$$

根据上式就可证明最优性原理的正确性。

若以 x_0 为初态时，余下的决策 $u_1, \cdots, u_{N-1}$ 不是最优的，那么就存在另一决策序列 $\bar{u}_1, \cdots, \bar{u}_{N-1}$ 所决定的指标值 $\bar{J}_{N-1} < J^*_{N-1}$，于是

$$J^*_N = \min_{u_0}\{F(x_0, u_0) + J^*_{N-1}\} > \bar{J}_N = \min_{u_0}\{F(x_0, u_0) + \bar{J}_{N-1}\}$$

这与 J^*_N 是极小值发生矛盾，所以余下的决策必须是最优的。
(4-1)式的函数方程与(4-2)式所表示的最优性原理是一致的，只是表示方法不同。(4-1)式中 $u_n(x)$ 的下标 n 表示离终点的级数，(4-2)式中 u_k 的下标 k 表示离起点的级数。两式的对照留给读者去做。
将(4-2)式进一步分解为

$$\begin{aligned}\min_{u_0,\cdots,u_{N-1}} J_N = & \min_{u_0}[F(x_0, u_0) + \min_{u_1}[F(x_1, u_1) \\ & + \cdots + \min_{u_{N-1}}[F(x_{N-1}, u_{N-1})]]]\end{aligned} \tag{4-3}$$

由上式可见，最优化的过程是从最里面的方括号开始向外扩展的，即寻找最优控制的次序是 $u_{N-1}, u_{N-2}, \cdots, u_1, u_0$。因此根据最优性原理，动态规划是从最后一级倒退计算的。

用动态规划解资源分配问题

我们提到过，动态规划的应用范围是非常广的。这里介绍用动态规划解决资源分配问题。假定有 m 种资源用来生产 n 种产品（资源可以指工人、机床、资金等，每种资源的总数为 $x_1, x_2, \cdots, x_m$）。如果生产第 i 种产品时投入的各种资源量为 $x_{i1}, x_{i2}, \cdots, x_{im}$ 则可以得到的收益为 g_i，它是所分配的资源量的函数，可写成 $g_i(x_{i1}, x_{i2}, \cdots, x_{im})$。现在问应如何分配这些资源到各个产品上，使得所有产品的总收益为最大？
写成数学形式，即要使

$$J = \sum_{i=1}^{n} g_i(x_{i1}, x_{i2}, \cdots, x_{im}) \quad g_i > 0 \tag{4-4}$$

取最大，其中满足约束

$$\sum_{i=1}^{n} x_{ij} \le x_j \quad j = 1, 2, \cdots, m \tag{4-5}$$

$$x_{ij} \ge 0 \quad i = 1, \cdots, n; \quad j = 1, \cdots, m \tag{4-6}$$

上面的问题可以用动态规划求解。为了说明问题简单起见，这里只考虑单资源分配问题，即如何将一种资源分配给 n 种产品，使总收益最大。设这种资源的总数为 x，分配给第 i 种产品的数量为 x_i，则性能指标为

$$J = \sum_{i=1}^{n} g_i(x_i) \quad g_i \geq 0 \tag{4-7}$$

取最大，约束条件是

$$\sum_{i=1}^{n} x_i = x \quad x_i \geq 0 \tag{4-8}$$

为了用动态规划求解，引进一个函数 $f_K(x_T)$，它表示将资源量 $x_T \leq x$ 分配给第 1 至第 K 种产品时所能得到的最大收益。显然 $f_n(x)$ 表示将总资源 x 分配到所有 n 种产品上所得到的最大收益，即

$$f_n(x) = J_{\max} \tag{4-9}$$

容易看出，函数 $f_K(x_T)$ 有下列性质

1. $f_K(0) = 0$ 即没有资源投入时收益为零。
2. $f_0(x_T) = 0$ 即不生产产品时收益为零。
3. $f_1(x_T) = g_i(x_T)$ 这表明将资源量只用于生产一种产品时的总收益，就是这种产品本身收益。

这些性质构成了以后解题的边界的条件。

现在来推导 $f_K(x_T)$ 所应满足的关系式。已知投入前 K 种产品的资源量为 x_T。如果投入第 K 种产品的资源量为 x_K，则投入前 $K-1$ 种产品的资源量为 $x_T - x_K$。
如果把 K 种产品的资源分配看成是 K 步决策，则 $f_K(x_T)$ 表示 K 步决策的指标最优值，$g_K(x_K)$ 表示用决策量 x_K 时第 K 步的指标值，$f_{K-1}(x_T - x_K)$ 表示余下 $K-1$ 步决策的指标最优值，根据最优性原理 (对照(4-2)式)，则有

$$f_K(x_T) = \max_{x_K}[g_K(x_K) + f_{K-1}(x_T - x_K)] \tag{4-10}$$

这表明若 x_T 在第 1 至 K 种产品上的最优分配为 $x_1, \cdots, x_K$，则 $x_1, \cdots, x_{K-1}$ 一定是资源量 $x_T - x_K$ 在前 $K-1$ 种产品上的最优分配。

例 4 - 1

假定某一种资源的量有四个单元（如重量单元千克，体积单元公升等），把它分配到三种产品的生产中，每种产品的收益函数 $g_i(x)$ 如下表所表示，x 表示所分配的资源的单元数。问怎样分配资源才能使总收益最大？

x(投入资源单元数)	1	2	3	4
$g_1(x)$(第一种产品增益)	8	18	22	24
$g_2(x)$(第二种产品增益)	3	6	9	12
$g_3(x)$(第三种产品增益)	6	7	8	10

解：由边界条件知 $f_1(x)=g_1(x)$。现在考虑 $f_K(x_T)=f_2(1)$，它表示用 1 个单元资源分配到 2 个产品上的最大收益。$x_K=x_2$ 表示投入第 2 个产品的资源，则 x_2 可取值 1 或 0，对应地将有下表。

x_2	x_T-x_2	第一产品收益	第二产品收益
1	0	$g_1(0)=f_1(0)$	$g_2(1)$
0	1	$g_1(1)=f_1(1)$	$g_2(0)$

根据(4-10)式可得

$$\begin{aligned} f_2(1) &= \max\{g_2(1)+f_1(0), g_2(0)+f_1(1)\} \\ &= \max\{3+0, 0+8\} = 8 \end{aligned}$$

$f_2(2)$ 表示用 2 个单元的资源分配到 2 个产品上，显然 x_2 可取值 2、1、0。类似地可得

$$\begin{aligned} f_2(2) &= \max\{g_2(2)+f_1(0), g_2(1)+f_1(1), g_2(0)+f_1(2)\} \\ &= \max\{6+0, 3+8, 0+18\} = 18 \end{aligned}$$

同理，当可取值 3，4 时可求得

$$\begin{aligned} f_2(3) &= \max\{g_2(3)+f_1(0), g_2(2)+f_1(1), g_2(1)+f_1(2), g_2(0)+f_1(3)\} \\ &= \max\{9, 14, 21, 22\} = 22 \\ f_2(4) &= \max\{12, 17, 24, 25, 24\} = 25 \end{aligned}$$

再考虑 3 个产品的资源分配，可得这三个产品投入资源的单元数为 1，2，3，4 时的最优值如下

$$
\begin{aligned}
f_3(1) &= \max\{g_3(1)+f_2(0), g_3(0)+f_2(1)\} \\
&= \max\{6+0, 0+8\} = 8 \\
f_3(2) &= \max\{g_3(2)+f_2(0), g_3(1)+f_2(1), g_3(0)+f_2(2)\} \\
&= \max\{7+0, 6+8, 0+18\} = 18 \\
f_3(3) &= \max\{8, 15, 24, 22\} = 24 \\
f_3(4) &= \max\{10, 16, 25, 28, 25\} = 28
\end{aligned}
$$

可见

$$
f_3(4) = g_3(1)+f_2(3) = g_3(1)+g_2(0)+f_1(3) = g_3(1)+g_2(0)+g_1(3)
$$

即把一个单元的资源分配给第三种产品，把三个单元的资源分配给第一种产品，第二种产品不分配资源，这时总收益达最大值 28。

用动态规划求离散最优控制

离散系统的状态方程为

$$x(k+1) = f[x(k), u(k)] \quad k = 0, 1, \cdots, N-1 \tag{4-11}$$

性能指标为

$$J = \sum_{k=0}^{N-1} J_k[x(k), u(k)] \tag{4-12}$$

寻找最优控制序列 $u(0), u(1), \cdots, u(N-1)$ 使 J 最小，这正是多级决策问题，利用最优性原理，根据递推方程(4-2)，从最后一步开始，逐步求出 $u(N-1), u(N-2), \cdots, u(0)$。下面用例子来说明求解过程。

例 4 - 2

系统方程为

$$x(k+1) = x(k) + u(k) \quad x(0)\text{给定} \tag{4-13}$$

$$J = \frac{1}{2}cx^2(2) + \frac{1}{2}\sum_{k=0}^{1} u^2(k) \tag{4-14}$$

要求用动态规划寻找最优控制序列 $u(0), u(1)$ 使 J 最小。

解：先考虑最后一步，即从 $x(1)$ 到 $x(2)$。这时由(4-13)、(4-14)式得

$$x(2) = x(1) + u(1)$$

$$J_1 = \frac{1}{2}cx^2(2) + \frac{1}{2}u^2(1) = \frac{1}{2}c[x(1) + u(1)]^2 + \frac{1}{2}u^2(1)$$

求 $u(1)$ 使 J 最小，得

$$\frac{\partial J_1}{\partial u(1)} = c[x(1)+u(1)] + u(1) = 0$$

$$u(1) = -\frac{cx(1)}{1+c}$$

将 $u(1)$ 代入 J_1 和 $x(2)$ 可得，相应于上面最优控制的性能指标与最优状态转移为

$$J_1^* = \frac{c}{2} \cdot \frac{x^2(1)}{1+c} \tag{4-15}$$

$$x(2) = \frac{x(1)}{1+c} \tag{4-16}$$

再考虑倒数第二步，即从

$$x(1) = x(0) + u(0)$$

$$J = J_0 + J_1^* = \frac{1}{2}u^2(0) + \frac{c}{2}\frac{x^2(1)}{1+c} = \frac{1}{2}u^2(0) + \frac{c}{2(1+c)}[x(0)+u(0)]^2$$

求 $u(0)$ 使 J 最小，得

$$\frac{\partial J}{\partial u(0)} = u(0) + \frac{c}{1+c}[x(0)+u(0)] = 0$$

$$u(0) = -\frac{cx(0)}{1+2c}$$

于是最优性能指标与最优状态转移为

$$J^* = \frac{cx^2}{2(1+2c)} \tag{4-17}$$

$$x(1) = x(0) + u(0) = \frac{1+c}{1+2c}x(0) \tag{4-18}$$

这个结果与用离散极小值原理求解结果完全一样。

连续系统的动态规划

设系统的状态方程和性能指标为

$$\dot{X} = f(t, X, U) \quad X(t_0) = X_0 \tag{4-19}$$

$$J = \phi[X(t_f), t_f] + \int_{t_0}^{t_f} F(X, U, t)\, dt \tag{4-20}$$

U 受约束，可写成 $U \in \Omega$，Ω 为某一闭集。要寻找满足此约束且使 J 最小的最优控制 U。

设时间 t 在区间 $[t_0, t_f]$ 内，则根据最优性原理，从 t 到 t_f 这一段过程应当是最优过程。把这段最优指标写成 $V(X, t)$，则

$$V(X, t) \triangleq J^*(x, t) = \min_{u \in \Omega} \{\phi[X(t_f), t_f] + \int_{t}^{t_f} F(X, U, \tau)\, d\tau\} \tag{4-21}$$

显然 $V(X, t)$ 满足终端条件

$$V[X(t_f), t_f] = \phi[X(t_f), t_f] \tag{4-22}$$

通常假定 $V(X, t)$ 对 X 及 t 的二阶偏导数存在且有界。

现在，考虑系统从 t 出发，到 t_f 分两步走：先从 t 到 $t+\Delta t$，再 $t+\Delta t$ 从到 t_f，Δt 是小量，则

$$\begin{aligned} V(X,t) &= \min_{u\in\Omega}\{\phi[X(t_f),t_f] + \int_t^{t+\Delta t} F(X,U,\tau)\,d\tau \\ &\quad + \int_{t+\Delta t}^{t_f} F(X,U,\tau)\,d\tau\} \end{aligned} \tag{4-23}$$

根据最优性原理，从 $t+\Delta t$ 到 t_f 也应是最优过程。
因 $X(t+\Delta t)\approx X+\dot{X}\Delta t$ 故

$$V(X+\dot{X}\Delta t, t+\Delta t) = \min_{u\in\Omega}\{\phi[X(t_f),t_f] + \int_{t+\Delta t}^{t_f} F(X,U,\tau)\,d\tau\}$$

这样，式(4-23)可写成

$$V(X,t)=\min_{u\in\Omega}\{F(X,U,t)\Delta t+V(X+\dot{X}\Delta t,t+\Delta t)\} \tag{4-24}$$

$$=\min_{u\in\Omega}\{F(X,U,t)\Delta t+V(X,t)+\left(\frac{\partial V}{\partial X}\right)^T\Delta X+\frac{\partial V}{\partial t}\Delta t+0(\Delta t^2)\}$$

注意到 $\Delta X=\dot{X}\Delta t=f(X,U,t)\Delta t$，上式右边括号中 $V(X,t)$ 表示最优指标，其中 U 为最优控制，不需再选择，$\frac{\partial V}{\partial t}\Delta t$ 也与 U 选择无关。故

$$\begin{aligned} V(X,t) &= \min_{u\in\Omega}\{F(X,U,t)\Delta t+\left(\frac{\partial V}{\partial X}\right)^T f(X,U,t)\Delta t\} \\ &\quad +V(X,t)+\frac{\partial V}{\partial t}\Delta t+0(\Delta t^2) \end{aligned}$$

从上式两端消去 $V(X,t)$，除以 Δt，再令 $\Delta t\to 0$，可得

$$-\frac{\partial V}{\partial t}=\min_{u\in\Omega}\{F(X,U,t)+\left(\frac{\partial V}{\partial X}\right)^T f(X,U,t)\} \tag{4-25}$$

引用以前使用过的哈密顿函数

$$H(X, U, \lambda, t) = F(X, U, t) + \lambda^T f(X, U, t) \tag{4-26}$$

$$\lambda = \frac{\partial V}{\partial X} \tag{4-27}$$

则(4-25)可写成

$$-\frac{\partial V}{\partial t} = \min_{u \in \Omega} H(X, U, \frac{\partial V}{\partial X}, t) = H^*(X, U, \frac{\partial V}{\partial X}, t) \tag{4-28}$$

(4-25)或(4-28)称为哈密顿—雅可比—贝尔曼方程，边界条件是(4-22)式。哈密顿—雅可比—贝尔曼方程在理论上很有价值，但它是 $V(X, t)$ 的一阶偏微分方程并带有取极小的运算，因此求解是非常困难的，一般情况得不到解析解，只能用计算机求数值解。对于线性二次问题，可以得到解析解，而且求解结果与用极小值原理或变分法所得结果相同。这时，哈密顿——雅可比——贝尔曼方程可归结为黎卡提方程。在实际计算线性二次问题时，一般用直接求解黎卡提方程来求最优控制。

动态规划与极小值原理

动态规划和极小值原理是最优控制理论的两大基石，它们都可以解决有约束的最优控制问题，虽然在形式上和解题方法上不同，但却存在着内在的联系。下面我们从动态规划来推演极小值原理，不过要说明这种推演是基于最优指标 V 针对 X 二次连续可微这个条件的。
于是最优性能指标与最优状态转移为

$$\dot{X} = f(X, U, t), \quad X(t_0) = X_0 \tag{4-29}$$

要求确定 $U(t)$ 使性能指标

$$J = \phi[X(t_f), t_f] + \int_{t_0}^{t_f} F(X, U, t)\, dt \tag{4-30}$$

极小。其中，t_0，t_f 固定，$x(t_f)$ 自由，U 可以有约束，也可以没有。

用极小值原理求解的结果可以由(4-31)-(4-35)来表示，因这里 t_f 固定，故不需最优终端时刻条件；$X(t_f)$ 自由，故无终端约束方程 $G[X(t_f)=0]$。将最优解的条件再写在下面以对照。

1. $$\dot{X} = \frac{\partial H}{\partial \lambda} \quad \text{(状态方程)} \tag{4-31}$$
2. $$\dot{\lambda} = -\frac{\partial H}{\partial X} \quad \text{(协态方程)} \tag{4-32}$$
3. $$X(t_0) = X_0 \quad \text{边界方程} \tag{4-33}$$
4. $$\lambda(t_f) = \frac{\partial \phi}{\partial X(t_f)} \quad \text{(横截条件)} \tag{4-34}$$
5. $$\min_{u \in \Omega} H(X^*, \lambda^*, U, t) = H(X^*, \lambda^*, U^*, t) \quad \text{极值条件} \tag{4-35}$$

用动态规划求解的结果已在上节中得到，现在归纳一下：在动态规划中协态变量 λ 满足

$$\lambda = \frac{\partial V}{\partial X}$$

哈密顿—雅可比—贝尔曼方程(4-28)本身说明了哈密顿函数在最优控制上取极值的条件，故等同于上面极小值原理所得的条件 5，不过(4-28)还多给出了一点信息，即 $H^* = -\frac{\partial V}{\partial t}$。

下面由动态规划法来推出协态方程。

由(4-27)

$$\dot{\lambda} = \frac{d\lambda}{dt} = \frac{d}{dt}\left[\frac{\partial V(X,t)}{\partial X}\right] = \frac{\partial}{\partial t}\left[\frac{\partial V(X,t)}{\partial X}\right] + \frac{\partial^2 V(X,t)}{\partial X \partial X^T}\frac{dX}{dt}$$

因假设对 X 两次连续可微，因此上式成立，且可交换求导次序，得

$$
\begin{aligned}
\dot{\lambda} &= \frac{\partial}{\partial X}\left[\frac{\partial V(X,t)}{\partial t}\right] + \frac{\partial^2 V(X,t)}{\partial X \partial X^T} f(X,U,t) \\
&= -\frac{\partial}{\partial X}\left[F(X,U,t) + \left(\frac{\partial V}{\partial X}\right)^T \cdot f(X,U,t)\right] + \frac{\partial^2 V(X,t)}{\partial X \partial X^T} f(X,U,t) \\
&= -\left[\frac{\partial F}{\partial X} + \left(\frac{\partial V}{\partial X}\right)^T \frac{\partial f}{\partial X}\right] \\
&= -\frac{\partial}{\partial X}\left[F + \left(\frac{\partial V}{\partial X}\right)^T f\right] \\
&= -\frac{\partial H}{\partial X}
\end{aligned}
$$

即协态方程(4-32)（因都是最优解条件。故省去 * 号)。

由(4-22)和(4-27)再来推横截条件

$$\lambda(t_f) = \frac{\partial V[X(t_f), t_f]}{\partial X(t_f)} = \frac{\partial \phi[X(t_f), t_f]}{\partial X(t_f)}$$

即横截条件(4-34)。其它条件如状态方程和初始条件都是给定的。故由动态规划推出了极小值原理的全部条件。应该强调，这不是说用动态规划可证明极小值原理。因为上面的推演要求 $V(X, t)$ 对 X，t 二次连续可微，而极小值原理的证明本身不需要这一条件。

小结

1. 动态规划是把多级决策问题化为多个单级决策问题来求解的，而单级问题比多级问题容易处理得多。这种把一个复杂的特定问题化为（又可称为嵌入）一系列性质相似的易于求解的问题的做法称为“不变嵌入”法。
2. 动态规划的基础是最优性原理。这个原理告诉我们：在多级最优决策中，不管初始状态是什么，余下的决策对此状态必定构成最优决策。根据这个原理，动态规划解决多级决策问题（包括离散系统最优控制）是从最后一级开始倒向计算的。
3. 连续系统的动态规划可导出哈密顿——雅可比——贝尔曼方程，这个方程一般只能有数值解。从它可推演出极小值原理，不过要假定 $V(X, t)$ 对 X，t 二次连续可微。
4. 动态规划比穷举法的计算量是少了不少，但对复杂问题（状态变量和控制变量的数目多，级数多），它的计算量和存储量仍旧非常大，有时用一般计算机也解决不了。这种情况称为“维数灾难”。