第二章　动态规划

（Dynamic Programming）

本章重点：动态规划的概念与建模、离散型与连续性问题的求解、主要的应用类型

本章难点：动态规划问题的建模

动态规划是研究多阶段决策问题的一种运筹学方法。本章将介绍动态规划的概念与方法，在此基础上，对动态规划的三种应用类型进行举例分析。

2.1 动态规划的基本概念与方法

2.1.1 多阶段决策问题

1．时间阶段的例子（机器负荷问题）

某厂有 1000 台机器，现需作一个五年计划，以决定每年安排多少台机器投入高负荷生产（产量大但损耗也大）可使五年的总产量最大，如图 2.1 所示。

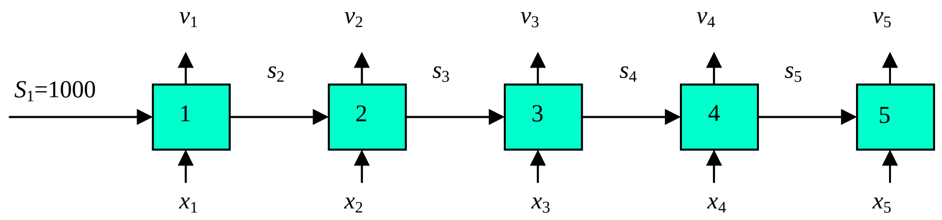

图 2.1 机器负荷问题

2．空间阶段的例子（最短路问题）

如图 2.2 为一线路网络。现要从 A 点铺设一条管道到 E 点，图中两点间连线上数字表示两点间距离。现需选一条由 A 到 E 的铺管线路，使总距离最短。

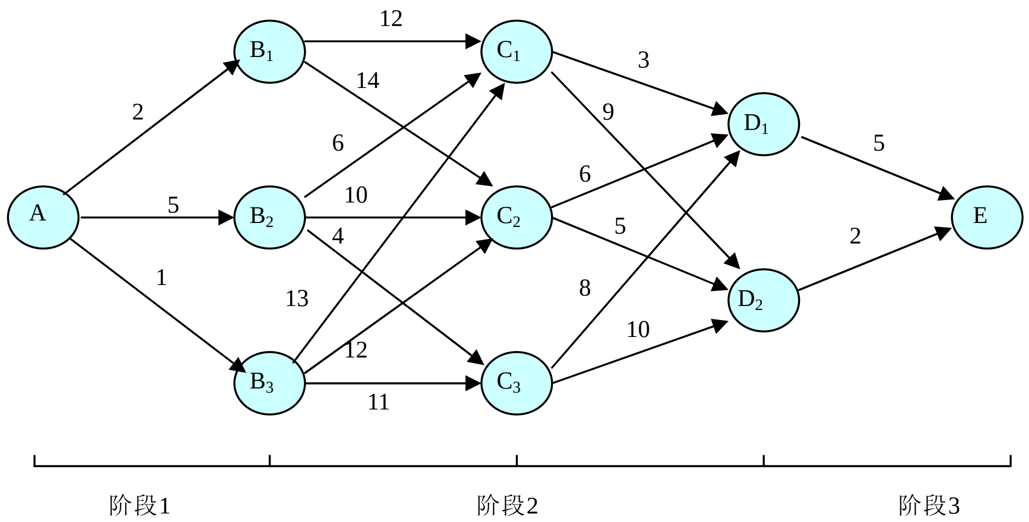

图 2.2 最短路问题

2.1.2 基本概念与方程

1．基本概念

阶段　　——分步求解的过程，用阶段变量 k 表示，k=1，…，n

状态　　——每阶段初可能的情形或位置，用状态变量 S_k 表示。

按状态的取值是离散或连续，将动态规划问题分为离散型和连续型。

决策　　——每阶段状态确定后的抉择，即从该状态演变到下阶段某状态的选择，用决策变量 x_k 表示。

状态转移——由 S_k 转变为 S_k+1 的规律，记 $S_k+1=T(S_k, x_k)$。

策略　　——由各阶段决策组成的序列，记 $P_{1n}=\{x_1, \ldots, x_n\}$，称 $P_{kn}=\{x_k, \ldots, x_n\}$为阶段 k 至 n 的后部子策略。

阶段指标——每阶段选定决策 x_k 后所产生的效益，记 $v_k= v_k(S_k, x_k)$。

指标函数——各阶段的总效益，记相应于 P_{kn} 的指标函数为 $v_{kn}= v_{kn}(S_k, P_{kn})$。

其中最优的称最优指标函数，记 $f_k = f_k(S_k)=\text{opt}\ v_{kn}$。

问题：动态规划的最优解和最优值各是什么？

——最优解：最优策略 P_{1n}；最优值：最优指标 f_1。

2．基本原理与基本方程

（1）基本原理

定理 2.1： $P_{1n}^* = (x_1^*, \cdots, x_n^*)$ 是最优策略 $\Leftrightarrow$ 对任何 k（$1<k<n$) 和允许状态 s_1，有 $f_1 = \underset{p_{1k}}{opt}\{v_{1k} + f_{k+1}\}$。

推论（*Bellman* 最优性原理）：若 P_{1n}^* 是最优策略，则对任何 k（$1<k<n$)，子策略 P_{1n}^* 对于以 s_k^* 为起点的 k 至 n 子过程来说必为最优策略。

以最短路为例说明，如图 2.3 所示。

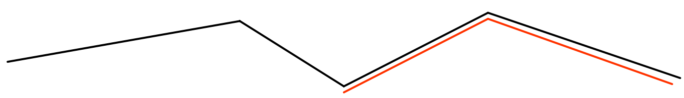

图 2.3 Bellman 最优性原理示意

（2）基本方程

根据最优性原理，可建立从后向前逆推求解的递推公式——基本方程：

$$\begin{cases} f_k = \underset{x_k}{opt}\{v_k + f_{k+1}\} \\ f_{n+1} = 0, \quad k = n, \cdots, 1 \end{cases} \tag{2-1}$$

动态规划求解的一般步骤：

- 确定过程的分段，构造状态变量；
- 设置决策变量，写出状态转移；
- 列出阶段指标和指标函数；
- 写出基本方程，由此逐段递推求解。

2.1.3 求解方法

1. 离散型（用表格方式求解）

例 2.1 用动态规划方法求解如图 2.2 所示的最短路问题。

解：设阶段 k=1，2，3，4 依次表示 4 个阶段选路的过程；

状态 s_k 表示 k 阶段初可能处的位置；

决策 x_k 表示 k 阶段初可能选择的路；

阶段指标 v_k 表示 k 阶段与所选择的路段相应的路长；

指标函数 $v_{k4}=\sum_{i=k}^{4} v_i$ 表示 k 至 4 阶段的总路长；

递推公式：$\begin{cases} f_k = Min\{v_k + f_{k+1}\} \\ f_5 = 0, \quad k = 4,\cdots,1 \end{cases}$

表 2.1 最短路问题的动态规划求解表

k	S_k	x_k	v_k	$v_{kn}=v_k+f_k+1$	f_k	P_{kn}^*
4	D_1	E	5	5+0	5	D_1E
	D_2	E	2	2+0	2	D_2E
3	C_1	D_1	3	3+5	8	C_1D_1E
		D_2	9	9+2		
	C_2	D_1	6	6+5	7	C_2D_2E
		D_2	5	5+2		
	C_3	D_1	8	8+5	12	C_3D_3E
		D_2	10	10+2		
2	B_1	C_1	12	12+8	20	$B_1\ C_1D_1E$
		C_2	14	14+7		
	B_2	C_1	6	6+8	14	$B_2\ C_1D_1E$
		C_2	10	10+7		
		C_3	4	4+12		
	B_3	C_1	13	13+8	19	$B_3\ C_2D_2E$
		C_2	12	12+7		
		C_3	11	11+12		
1	A	B_1	2	2+20	19	$AB_2\ C_1D_1E$
		B_2	5	5+14		
		B_3	1	1+19		

$P_{14}^* = AB_2C_1D_1E$（最短路），f_1=19（最短距离）。最短路如图 2.4 所示。

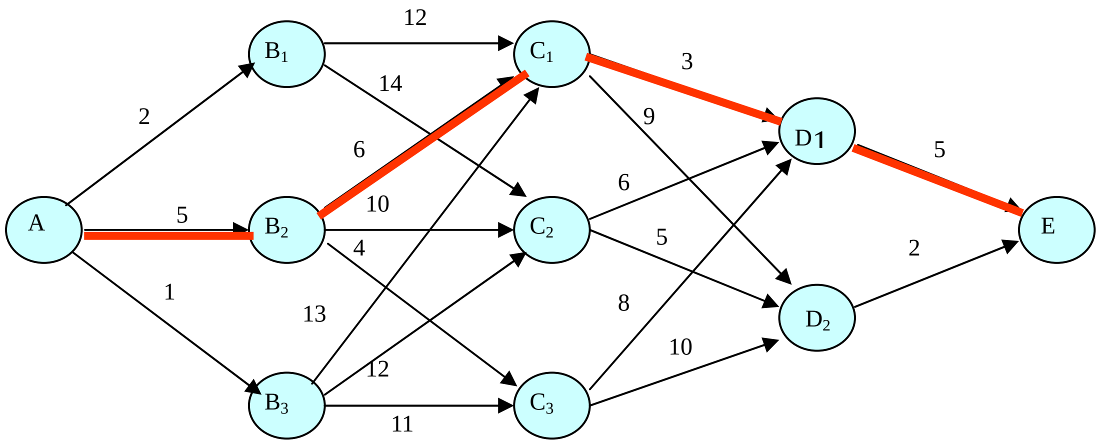

图 2.4　最短路问题的求解结果

2．连续型（用公式递推求解）

例 2.2　用动态规划方法求解前面的机器负荷问题。

某种机器可以在高、低两种负荷下进行生产。高负荷年产量 8，年完好率 0.7；低负荷年产量 5，年完好率 0.9。现有完好机器 1000 台，需制定一个 5 年计划，以决定每年安排多少台机器投入高、低负荷生产，使 5 年的总产量最大。

解：设阶段 k=1，…，5 表示第 k 年安排机器的过程；

状态 s_k 表示第 k 年初的完好机器台数；

决策 x_k 表示第 k 年投入高负荷的机器台数；则投入低负荷的台数为 s_k-x_k ；

状态转移 $s_k+1=0.7x_k+0.9(s_k - x_k)$；

阶段指标 $v_k=8x_k+5(s_k - x_k)$表示第 k 年的产量；

指标函数 v_{kn}=表示第 k 至 5 年的总产量；

递推公式 $\begin{cases} f_k = Max\{v_k + f_{k+1}\} \\ f_6 = 0, \quad k = 5,\cdots,1 \end{cases}$

当 $k=5$，$f_5 = \underset{0 \le x_5 \le s_5}{Max}\{v_5 + f_6\} = \underset{0 \le x_5 \le s_5}{Max}\{8x_5 + 5(s_5 - x_5)\}$

$$= \underset{0 \le x_5 \le s_5}{Max}\{3x_5 + 5s_5\}$$

$\therefore x_5^* = s_5$， $f_5 = 8s_5$，即第 5 年初将全部完好机器都投入高负荷。

当 $k=4$，$f_4 = \underset{0 \le x_4 \le s_4}{Max}\{v_4 + f_5\} = \underset{0 \le x_4 \le s_4}{Max}\{8x_4 + 5(s_4 - x_4) + 8s_5\}$

$$= \underset{0 \le x_4 \le s_4}{Max}\{3x_4 + 5s_4 + 8(0.7x_4 + 0.9(s_4 - x_4))\}$$

$$= \underset{0 \le x_4 \le s_4}{Max}\{1.4x_4 + 12.2s_4\}$$

$\therefore x_4^* = s_4$， $f_4 = 13.6s_4$，即第 4 年初将全部完好机器都投入高负荷。

当 $k=3$，$f_3=\underset{0\le x_3\le s_3}{Max}\{v_3+f_4\}=\underset{0\le x_3\le s_3}{Max}\{3x_3+5s_3+13.6(0.9s_3-0.2x_3)\}$

$$=\underset{0\le x_3\le s_3}{Max}\{0.28x_3+17.24s_3\}$$

$\therefore x_3^*=s_3$， $f_3=17.52s_3$，即第 3 年初将全部完好机器都投入高负荷。

当 $k=2$，$f_2=\underset{0\le x_2\le s_2}{Max}\{v_2+f_3\}=\underset{0\le x_2\le s_2}{Max}\{3x_2+5s_2+17.52(0.9s_2-0.2x_2)\}$

$$=\underset{0\le x_2\le s_2}{Max}\{20.8s_2-0.5x_2\}$$

$\therefore x_2^*=0$， $f_2=20.8s_2$，即第 2 年初将全部完好机器都投入低负荷。

当 $k=1$，$f_1=\underset{0\le x_1\le s_1}{Max}\{v_1+f_2\}=\underset{0\le x_1\le s_1}{Max}\{3x_1+5s_1+20.8(0.9s_1-0.2x_1)\}$

$$=\underset{0\le x_1\le s_1}{Max}\{23.72s_1-1.16x_1\}$$

$\therefore x_1^*=0$， $f_1=23.72s_1$，即第 1 年初将全部完好机器都投入低负荷。

5 年的最大总产量为 23 .72×1000=23720。

2.2 动态规划应用举例

本节将通过动态规划的三种应用类型——资源分配问题、复合系统可靠性问题、设备更新问题，进一步介绍动态规划的特点和处理方法。

2.2.1 资源分配问题

1．问题的一般提法

设有某种资源，总数量为 a，用于生产 n 种产品；若分配数量 x_i 用于生产第 i 种产品，其收益为 $g_i(x_i)$。问应如何分配，可使总收益最大？

2．数学规划模型

决策变量：设分配给第 i 种产品的资源数量为 x_i；

目标函数：$Maxz=\sum_{i=1}^{n}g_i(x_i)$ (2-2)

约束条件 $\begin{cases}\sum_{i=1}^{n}x_i=a\\ x_i\ge 0, i=1,\cdots,n\end{cases}$ (2-3)

模 型 的 特 点
——变量分离

3．用动态规划方法求解

思路如图 2.5。

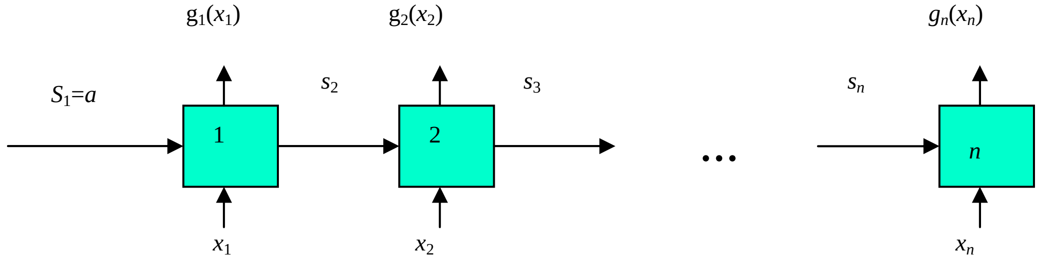

图 2.5 资源分配问题

阶段 k=1,…,n 表示把资源分配给第 k 种产品的过程；

状态 s_k 表示在给第 k 种产品分配之前还剩有的资源量；

决策 x_k 表示分配给第 k 种产品的资源量；

状态转移 $s_k +1 = s_k - x_k$ ；

阶段指标 $v_k = g_k(x_k)$；

指标函数 $v_{kn} = \sum_{i=k}^{n} g_i(x_i)$；

基本方程 $\begin{cases} f_k = \underset{x_k}{Max}\{v_k + f_{k+1}\} \\ f_{n+1} = 0, k = n, \cdots, 1 \end{cases}$

例 2.3 某公司拟将某种高效设备 5 台分配给所属甲、乙、丙 3 厂。各厂获此设备后可产生的效益如表 2.2 所示。问应如何分配，可使所产生的总效益最大？

表 2.2 效益表

效益 厂 / 设备台数	甲	乙	丙
0	0	0	0
1	3	5	4
2	7	10	6
3	9	11	11
4	12	11	12
5	13	11	12

解：阶段 k =1,2,3 依次表示把设备分配给甲、乙、丙厂的过程；

状态 s_k 表示在第 k 阶段初还剩有的可分台数；

决策 x_k 表示第 k 阶段分配的设备台数；

状态转移 $s_k +1 = s_k - x_k$ ；

阶段指标 v_k 表示第 k 阶段分配后产生的效益；

指标函数 $v_{k3} = \sum_{i=\mathrm{k}}^{3} v_i(x_i)$；

基本方程 $\begin{cases} f_k = \underset{x_k}{Max}\{v_k + f_{k+1}\} \\ f_4 = 0, k = 3,2,1 \end{cases}$

问题：本问题是属于离散型还是属于连续型？怎样解？

——离散型，用表格的方式求解。

表 2.3 动态规划的求解表

k	s_k	x_k	v_k	$v_{kn} = v_k + f_k+1$	f_k	P_{kn}^*
3	0	0	0	0+0	0	0
	1	1	4	4+0	4	1
	2	2	6	6+0	6	2
	3	3	11	11+0	11	3
	4	4	12	12+0	12	4
	5	5	12	12+0	12	5
2	0	0	0	0+0	0	0-0
	1	0	0	0+4	5	1-0
		1	5	5+0		
	2	0	0	0+6	10	2-0
		1	5	5+4		
		2	10	10+0		
	3	0	0	0+11	14	2-1
		1	5	5+6		
		2	10	10+4		
		3	11	11+0		
	4	0	0	0+12	16	1-3 2-2
		1	5	5+11		
		2	10	10+6		
		3	11	11+4		
		4	11	11+0		
	5	0	0	0+12	21	2-3
		1	5	5+12		
		2	10	10+11		
		3	11	11+6		
		4	11	11+4		
		5	11	11+0		
1	5	0	0	0+21	21	0-2-3 2-2-1
		1	3	3+16		
		2	7	7+14		
		3	9	9+10		
		4	12	12+5		
		5	13	13+0		

最优策略：P_{13}^* 为 0-2-3 或 2-2-1，即分给甲厂 0 台、分给乙厂 2 台、分给丙厂 3 台；或

分给甲厂 2 台、分给乙厂 2 台、分给丙厂 1 台。

最优值：　　f_1=21。

可见，最优解可以是不唯一的，但最优值是唯一的。

资源分配问题的应用很广泛，例如：

（1）某学生正在备考 4 门功课，还剩 7 天时间，每门功课至少复习 1 天。若他已估计出各门功课的复习天数与能提高的分数之间的关系，问他应怎样安排复习时间可使总的分数提高最多？

（2）背包问题：旅行者携带的背包中能装的物品重量为 a，现他要从 n 种物品中挑选若干数量装入背包，问他应如何挑选可使所带的物品总价值最大？

2.2.2 复合系统工作可靠性问题

1．问题的一般提法

设某工作系统由 n 个部件串接而成，为提高系统的可靠性，在每个部件上装有备用件。已知部件 i 上装有 x_i 个备用件时，其正常工作的概率为 $p_i(x_i)$；每个部件 i 的备用件重量为 w_i，系统要求总重量不超过 W。问应如何安排备用件可使系统可靠性最高？

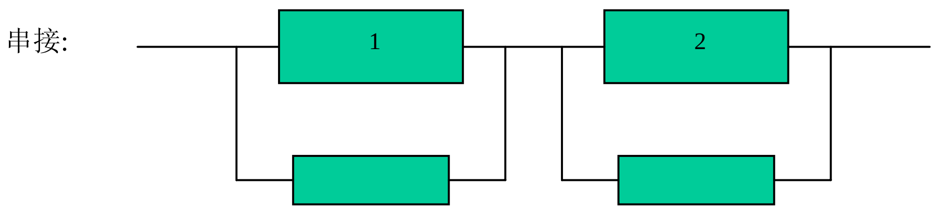

图 2.6 复合系统工作可靠性问题

2．数学规划模型

决策变量：设给第 i 个部件安排 x_i 个备用件；

目标函数：$Max\ z=\prod_{i=1}^{n} p_i(x_i)$　　(2-4)

约束条件 $\begin{cases} \sum_{i=1}^{n} w_i x_i \le W \\ x_i \text{为非负整数} \end{cases}$　　(2-5)

3．用动态规划方法求解

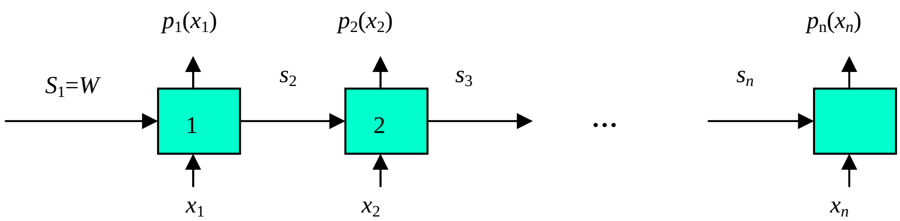

图 2.7 复合系统工作可靠性问题的动态规划求解

阶段 k =1,…,n 表示安排第 k 个部件备用件的过程；

状态 s_k 表示在给第 k 个部件安排之前还剩有的容许重量；

决策 x_k 表示第 s_k 个部件上安排的备用件数量；

状态转移 $s_k+1 = s_k - w_k x_k$ ；

阶段指标 v_k 表示第 k 个部件安排备用件后产生的可靠性；

指标函数 $v_{kn} = \prod_{i=k}^{\mathrm{n}} p_i(x_i)$；

基本方程 $\begin{cases} f_k = \underset{x_k}{Max}\{v_k \times f_{k+1}\} \\ f_{\mathrm{n}+1} = 1, k = n, \cdots, 1 \end{cases}$

模 型 的 特 点
——变量分离

可靠性问题的应用很广泛，例如：

（1）某重要的科研攻关项目正在由 3 个课题组以 3 种不同的方式进行，各组已估计出失败的概率。为减少失败的概率，选派了 2 名高级专家去充实科研力量。若可估计出各组增加专家后的失败概率，问应如何分派专家可使总的失败概率最小？

（2）已知 $x_1 + x_2 + \cdots + x_n = c$，求 $z = x_1 x_2 \cdots x_n$ 的最大值。

2.2.3 设备更新问题

例 2.4 某运输公司购进一批卡车投入运营，公司每年初需对卡车作出更新或继续使用的决定。假设第 k 年中，$r_k(t_k)$表示车龄为 t_k 的车使用一年的收入，$u_k(t_k)$表示车龄为 t_k 的车使用一年的维修费用，$c_k(t_k)$表示车龄为 t_k 的车更新成新车的费用。现公司需制定一个 10 年计划，以决定如何安排使 10 年的总收入最大。

问题：状态和决策怎样设置？

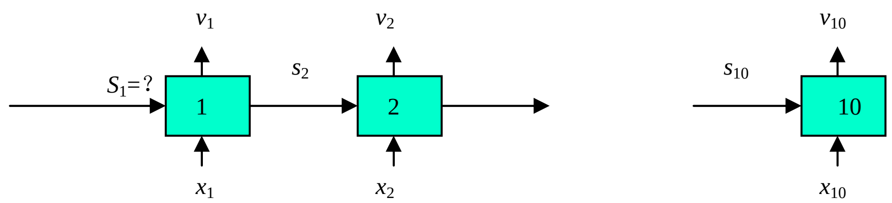

图 2.8 设备更新问题

解： 决策是更新与否，可用 0-1 变量表示；状态可设为车龄。

阶段 $k = 1,\ldots,10$ 表示第 k 年的决策过程；

状态 $s_k = t_k$ 表示第 k 年的车龄；

决策 $x_k = \begin{cases} 1, & 第k年更新 \\ 0, & 第k年不更新 \end{cases}$

状态转移 $t_k+1 = t_k (1-x_k)+1$

阶段指标 $v_k = r_k[t_k(1-x_k)] - u_k[t_k(1-x_k)] - x_k c_k(u_k)$

指标函数 $v_{kn} = \sum_{i=k}^{10} v_i$；

基本方程 $\begin{cases} f_k = \underset{x_k \in \{0,1\}}{Max}\{v_k + f_{k+1}\} \\ f_{11} = 0, k = 10, \cdots, 1 \end{cases}$