第 11 章　动 态 规 划

动态规划是制定一系列相关决策时运用的一项数学技术，它可以为寻求最优决策组合提供系统化的方法。

与线性规划相比，动态规划问题没有一个标准的数学模型。然而，动态规划是一类通用的问题解决方法，需要建立特定的方程以适应各种情况。因而，只有对动态规划问题一般性结构有深入的理解，才能够更好地判断何时以及如何通过动态规划的方法解决问题。这些能力可以通过大量的动态规划应用以及对其普遍特性的研究而形成。基于这一目的，本章将给出大量实例（其中有些实例可以快速地使用穷举法得到结论，但是运用动态规划的方法可以在面对更复杂的相似问题时，提高效率）。

11.1　动态规划的范例

11.1.1　例 1　驿站马车问题

驿站马车问题是专门构建用来阐述动态规划特征[1]和介绍动态规划术语的特殊问题。

19 世纪中叶，密苏里州的一位淘金者决定去加州西部淘金。淘金者在旅途中，需要乘坐驿站马车经过一些有强盗出没的乡村。虽然他的出发点和目的地已定，但是他有多种路线的选择，这些不同的路线经过不同的州（或后来成为州的地区）。图 11.1 所示为可能路线，每个州用带圆圈的字母表示，旅行的方向从左向右。淘金者无论选择哪条路线从 A 州（密苏里州）出发到达目的地 J 州（加利福尼亚州）都需要经过 3 个驿站，行驶 4 段被驿站隔开的路程，称为 4 个阶段。

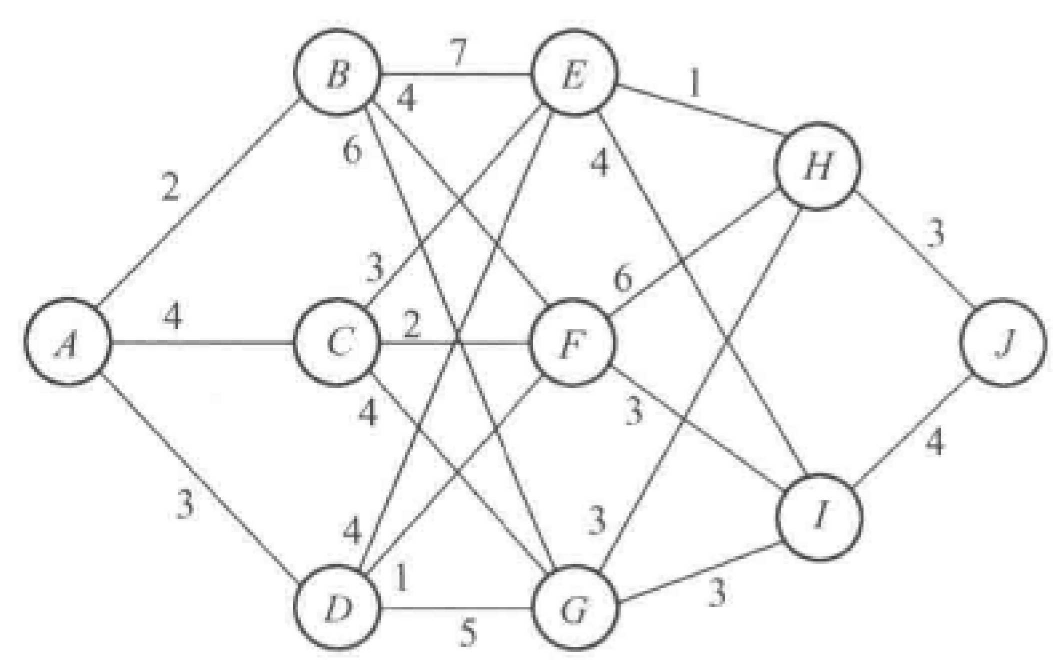


图 11.1　驿站马车问题的路线和成本

淘金者是个谨慎的人，他很担心自己的安全，经过思考，他想到一个巧妙的方法寻找最安全的路线。由于路途凶险，政策为每位驿站马车的乘客提供保险，每条路线保险的成本是充分考虑路线的凶险程度后设定的，因此，最安全的路线应当是保险金额最低的路线。图 11.1 中连接两个驿站之间的横线上的数字为该路线的保险金额。

从 I 州到 J 州驿站马车的保险金额用 c_{ij} 表示：

	B	C	D
A	2	4	3

	E	F	G
B	7	4	6
C	3	2	4
D	4	1	5

	H	I
E	1	4
F	6	3
G	3	3

	J
H	3
I	4

① 该问题由 Harvey M. Wagner 教授在斯坦福大学的时候构建。

现在问题就转化为选哪条路线保险金额之和最小的问题了。

11.1.2　问题的求解

首先需要注意，在每个州出发时选择最低金额的路线组合后并不一定会得到一个整体金额最低的方案，假如按照这种思路，那么，应当选择 $A \to B \to F \to I \to J$ 的路线，这条路线的保险金额为 13，而路线 $A \to D \to F \to I \to J$ 显然要比前者金额更低。

解决这一问题可以使用反复试验的办法①，但是可能的路线数量较多（共 18 条），计算每条路线的总保险金额并不容易。

动态规划提供了一种比反复试验工作量小的解决方案（面对更大规模的问题，减少工作量的优势更加明显）。动态规划首先从原始问题中的一小部分入手，找到小部分的最优解，然后逐渐扩大这一部分，逐步寻找最优解，直到求得原始问题的最优解。

针对驿站马车问题，我们首先从淘金者旅程的最后一个阶段（指抵达目的地前的一个州开始的路线）开始求解，然后每次向前增加一个阶段，逐步扩大问题，从前面已求解的问题找出目前问题的最优解。方法的具体细节如下。

建模：前面提到淘金者从出发到终点需要经过 4 个阶段，设决策变量 x_n（$n=1,2,3,4$）为阶段 n 的直接目的地。这样，选择的路线就是 $A \to x_1 \to x_2 \to x_3 \to x_4$，其中 $x_4 = J$。

假如淘金者在 s 州，即将开始第 n 阶段的路线，且 x_n 为其该阶段目的地，设 $f_n(s, x_n)$ 为该阶段及其之后阶段的整体最优路线的保险金额，并设 x_n^* 为使 $f_n(s, x_n)$ 最小的 x_n（不一定唯一），$f_n^*(s)$ 为相应的 $f_n(s, x_n)$ 最小值，则

$$f_n^*(s) = \min_{x_n} f_n(s, x_n) = f_n(s, x_n^*)$$

式中：$f_n(s, x_n)$ = 阶段 n 金额+最小未来金额（阶段 $n+1$ 至终点）$= c_{sx_n} + f_{n+1}^*(x_n)$。$c_{sx_n}$ 的值可以通过前面表格中 c_{ij} 的值确定。因为第 4 阶段的终点为最终的目的地（J 州），因此 $f_5^*(J) = 0$。

我们的目标是找到 $f_1^*(A)$ 和相应的路线，动态规划通过连续地确定 $f_4^*(s)$、$f_3^*(s)$、$f_2^*(s)$，最终确定 $f_1^*(A)$②。

求解过程：当淘金者的旅程只剩最后一个阶段（$n=4$）时，由于他的最终目的地 $x_4 = J$，路线选择由目前所处的州 s（H 或者 I）决定，最终行驶的路线为 $s_4 \to J$。由于 $f_4^*(s) = f_4(s, J) = c_{sJ}$，则 $n = 4$ 时：

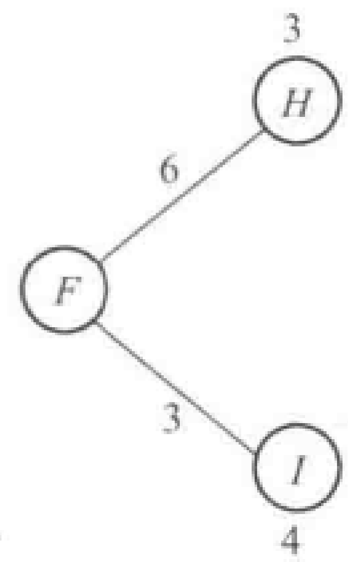
s	$f_4^*(s)$	x_4^*
H	3	J
I	4	J

当淘金者的旅程剩两个阶段时，求解的过程相对复杂一下。例如，当淘金者在 F 州，根据路线图，接下来他必须走到 H 州或者 I 州，直接保险金额分别为 $c_{F,H} = 6$ 或者 $c_{F,I} = 3$，如果他接下来选择到 H 州，则在他到达之后的最小保险金额为 $f_4^*(H) = 3$，这一决策的全部成本为 $6+3=9$；如果他接下来选择到 I 州，则在他到达之后的最小保险金额为 $f_4^*(I) = 4$，这一决策的全部成本为 3+4

① 本问题也可以构建为最短路径问题（见 10.3 节），这里的保险金额相当于最短路径中的距离。10.3 节中的算法实际上运用了动态规划求解的思路。由于这个问题中途径路线的数量一定，因此用动态规划方法更好。

② 因为这一过程为一步步逆向反推的过程，有些作者也将 n 定义为距离目的地剩余的阶段数。本书中简化起见，采用更直观的前向段数定义 n。

$=7$,小于前者。所以,最优的选择是后者,$x_3^*=I$,此时,最小金额$f_3^*(F)=7$。



当然,淘金者旅程剩两个阶段时,还存在 $s=E$ 和 $s=G$ 两种情况,这两种情况也需要进行类似于上述的计算,这里不再介绍具体的计算过程,利用下面的表格将最终的结果进行呈现。当 $n=3$时:

x_3 \ s	$f_3(s,x_3)=c_{sx_3}+f_4^*(x_3)$		$f_3^*(s)$	x_3^*
	H	I		
E	4	8	4	H
F	9	7	7	I
G	6	7	6	H

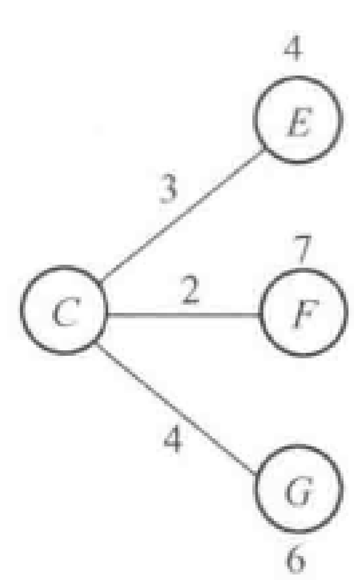
对第 2 阶段($n=2$)问题求解与之前类似,运用公式$f_2(s,x_2)=c_{sx_2}+f_3^*(x_2)$求解。假如淘金者在 C 州($s=C$ 时),如下面路线图所示,他接下来必须走到 E 州、F 州或 G 州,这一阶段的保险金额分别为$c_{C,E}=3$、$c_{C,F}=2$ 或者$c_{C,G}=4$。根据 $n=3$ 时的计算结果:$f_3^*(E)=4$、$f_3^*(F)=7$ 或者$f_3^*(G)=6$,可以得到 $s=C$ 时,3 种不同选择的保险金额计算结果:

$x_2=E$: $f_2(C,E)=c_{C,E}+f_3^*(E)=3+4=7$

$x_2=F$: $f_2(C,F)=c_{C,F}+f_3^*(F)=2+7=9$

$x_2=E$: $f_2(C,G)=c_{C,G}+f_3^*(G)=4+6=10$

最小值为 7,因此,从 C 州到终点的最小保险金额是$f_2^*(C)=7$,该选择下这一阶段的目的地是$x_2^*=E$。

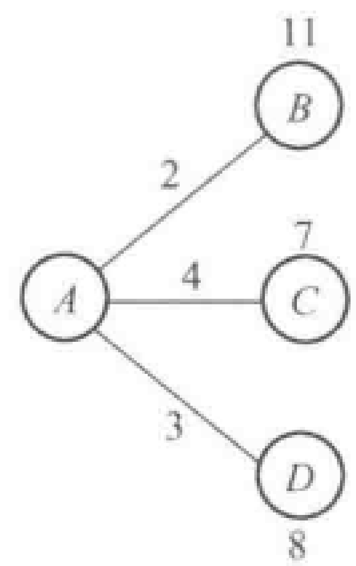


用相同的方法计算从起点 B 或 D 出发的结果,可以得到 $n=2$ 时的计算结果,如下表所列。

x_2 \ s	$f_2(s,x_2)=c_{sx_2}+f_3^*(x_2)$			$f_2^*(s)$	x_2^*
	E	F	G		
B	11	11	12	11	E 或 F
C	7	9	10	7	E
D	8	8	11	8	E 或 F

表格的第 1 行和第 3 行中显示，从 B 州或者 D 州出发，到达 E 州或 F 州都是该出发点下的最佳路线选择。

对于第 1 阶段（$n=1$）如何选择的问题，需要涉及之后的全部 3 个阶段，这一步的计算与第 2 阶段时计算相似，并且只涉及一个出发点。$s=A$ 时的路线图如下所示：



计算过程如下：

$x_1=B$:　$f_1(A,B)=c_{A,B}+f_2^*(B)=2+11=13$

$x_1=C$:　$f_1(A,C)=c_{A,C}+f_2^*(C)=4+7=11$

$x_2=D$:　$f_1(A,D)=c_{A,D}+f_2^*(D)=3+8=11$

最小值为 11，因此 $f_1^*(A)=11$，$x_1^*=$C 或 D，如下表所列：

x_1 / s	$f_1(s,x_1)=c_{sx_1}+f_2^*(x_1)$			$f_1^*(s)$	x_1^*
	B	C	D		
A	13	11	11	11	C 或 D

从 4 个表格中可以得到驿站马车问题的最优解。$n=1$ 的问题的结果表明，淘金者从 A 州出发后应该到 C 州或者 D 州。假如他选择到 C 州，那么，由于 $s=C,n=2$ 时，$x_2^*=E$，接下来应该到 E 州；之后，由于 $s=E,n=3$ 时，$x_3^*=H$，他应当在离开 E 州后来到 H 州；最后从 H 州出发到达最终目的地 J 州。综上所述，选择第一阶段到 C 州之后的最优路线是 $A\to C\to E\to H\to J$。同理，可以得出淘金者第一阶段选择到 D 州之后的最优路线为 $A\to D\to E\to H\to J$ 和 $A\to D\to F\to I\to J$。3 条路线的总保险金额都为 11。

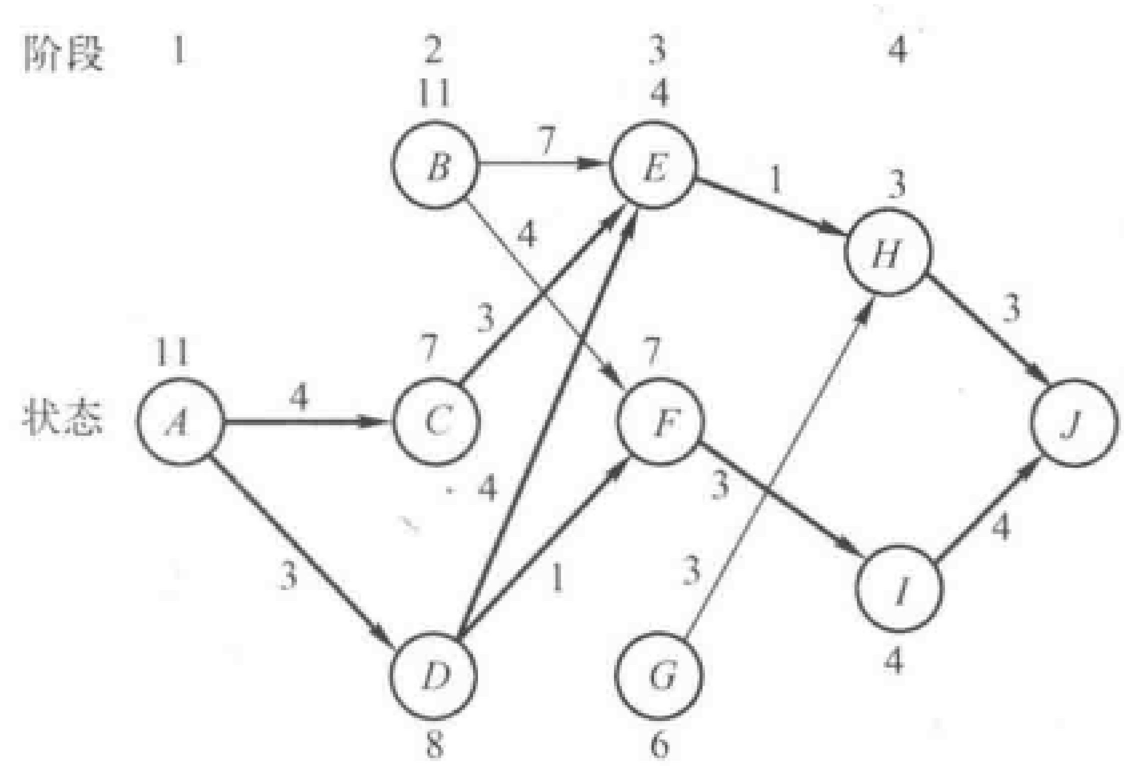


图 11.2　驿站马车问题动态规划求解图示

这一问题的动态规划分析的结果概览如图 11.2所示。图中的箭头和金额数字都可以从之前的 4 个表中找到依据。

图 11.2 中的每个箭头表示处于该状态时的最优决策，数字代表从该状态到终点的保险金额。加粗箭头标出的路线是 A 到 J 的 3 个最优路线，保险金额均为 11。

下一章节中将会对描述这个问题的特殊名词（阶段、状态和策略）进行解释。

11.2　动态规划问题的特性

驿站马车问题是动态规划问题的范例，这一范例的设计是为了对动态规划问题进行具体的

解释。因此,一种识别动态规划问题的方法,就是该问题的结构是否与驿站马车问题的结构类似。

本节对提出动态规划问题的基本特征,并对其进行讨论,具体如下。

(1) 问题可以划分为多个阶段(stages),每个阶段对应一个策略决策(policy decision)。

驿站马车问题中,根据旅程的实际情况,将问题分为了4个阶段。每个阶段对应的策略决策是选择保险金额(根据金额确定马车的下一站)。与此类似,其他的动态规划问题也需要制定一系列的决策,并且这些决策与其问题的阶段划分相一致。

(2) 每个阶段都存在一些与该阶段开始就相关的状态。

驿站马车问题中,每个阶段的状态是该阶段开始时所处的州以及该阶段结束时所在的州。一般来说,状态指的就是动态规划问题中每个阶段各种可能的情况,状态的数量可以是有限的(如驿站马车问题),同时也可以是无限的(之后的例子会涉及)。

(3) 每个阶段的策略决策结果是导致当前状态向下一阶段开始时的状态转变(可能是依据概率分布的)。

淘金者下一个目的地的决策,使他在旅程中从一个州走到下一个州。这一过程表明,动态规划问题可以用第10章中描述的网络来解释,网络中每个节点代表一个状态,网络由节点组成的列构成,每一列代表一个阶段,从网络中的一个点只能向右面的列流动。连接两个点之间的链代表了一种策略决策,而链的价值表示执行这一决策策略后对目标函数的直接贡献。多数情况下,动态规划问题的目标是通过网络寻找最短或最长路径。

(4) 求解问题的过程可以为所有的问题提供一个最优策略,如为每一阶段为每个可能的状态提供最优策略决策。

驿站马车问题求解的过程中,为每个阶段构建了一个表格,指明处于该阶段开始时的最优决策,因此,假如淘金者第一个阶段没有选择最优路线,而是去了 B 州,那么,驿站马车问题的求解结果依旧可以为他提供一个从该阶段开始的最优路线。对任何采用动态规划方法解决的问题,分析结果为该问题提供了每个可能的状态下应该做的策略决定,而并非简单地指出一种从始到终的最优策略。

(5) 已知目前状态的情况下,剩余阶段的最优策略与先前阶段采用的策略无关。因此,最优决策要依据当前的状态,而与如何到达这种状态没有关系,这是动态规划的最优原理。

已知淘金者目前所在的州,之后的旅程中保险金额最低的路线与他如何到这个州没有关系。通常,在动态规划问题中,系统当前的状态所展示出的信息,包含了所有其先前的决定中会对之后决策造成影响的信息(该特性是之后29.2节中讨论的马尔可夫链的性质)。任何不具备这一特征的问题,都不能通过建立动态规划模型来求解。

(6) 求解的过程从为最后一个阶段找出最优策略开始。

最后一个阶段的最优决策,表示该阶段每种可能状态的最优决策策略。这个单一阶段的问题求解方法采用尝试的办法,如驿站马车问题的范例中所示。

(7) 从 $n+1$ 阶段的最优策略,可以通过递推关系确定第 n 阶段的最优策略。驿站马车问题中的递推关系为

$$f_n^*(s)=\min_{x_n}\{c_{sx_n}+f_{n+1}^*(x_n)\}$$

上式是针对特定问题的动态规划递推关系,这里利用类比的方法,对动态规划问题普遍的递推关系进行概括:

N=阶段的数量;

n=当前阶段的标号($n=1,2,\cdots,N$);

s_n = 第 n 阶段的当前状态；

x_n = 第 n 阶段的决策变量；

x_n^* = x_n 的最优值（给定 s_n 的情况下）；

$f_n(s_n, x_n)$ = 从 n 阶段的 s_n 状态出发，直接决策为 x_n 时，阶段 $n, n+1, \cdots, N$ 对目标函数的贡献值，然后据此确定最优决策时的贡献值为

$$f_n^*(s_n) = f_n(s_n, x_n^*)$$

这种递推关系总是表示为

$$f_n^*(s_n) = \max_{x_n}\{f_n(s_n, x_n)\}$$

或者

$$f_n^*(s_n) = \min_{x_n}\{f_n(s_n, x_n)\}$$

式中：$f_n(s_n, x_n)$ 可以用 s_n、x_n、$f_{n+1}^*(s_{n+1})$ 等表示，正是由于 $f_n^*(s_n)$ 和 $f_{n+1}^*(s_{n+1})$ 可以通过这种方式互相表示，所以具备了递推的关系。

处理问题时，一步步的逆序分析，使得这种递推的关系不断重复。直到数字 n 的值降到 1。这种特性在下一个特点中会进一步强调。

(8) 利用递推关系求解动态规划问题时，我们从问题的最后出发，一步一步向前推导，直至找到第 1 阶段的最优策略。第 1 阶段的最优策略就是整个问题的最优解决方案，即初始阶段 s_1 选择 x_1^* 这一决策变量；第 2 阶段 s_2 选择 x_2^* 这一决策变量，依此类推。

驿站马车问题的求解演示了整个逆向递推的过程，[①]对于所有动态规划问题，每个阶段（$n = N, N-1, \cdots, 1$）都可以使用如下所列的表格。

s_n \ x_n	$f_n(s_n, x_n)$	$f_n^*(s_n)$	x_n^*

最初阶段（$n=1$）时，得到这个表格后就解决了动态规划需要解决的问题。

11.3 确定性动态规划

确定性动态规划问题，是指下一阶段的状态完全由当前阶段状态和策略决策决定的动态规划问题。本节深入讨论确定性动态规划问题的方法，下一节会对随机动态规划问题进行讨论。

确定性动态规划可以用图 11.3 描述。在第 n 阶段，在决策变量 x_n 的作用下，状态从 s_n 变成了 s_{n+1}，由于 s_{n+1} 状态下的最优策略的 $f_{n+1}^*(s_{n+1})$ 已经在之前的分析中计算了出来，然后，只需要根据 x_n 对于 $f_n(s_n, x_n)$ 的值的贡献，就可以确定各种情况下最优化的 x_n，寻找到相匹配的 x_n^* 和 $f_{n+1}^*(s_{n+1})$ 之后，求解的过程就可以逆向移动一个阶段。

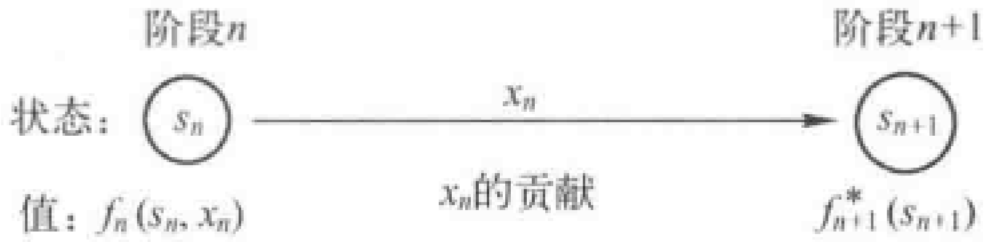


图 11.3　确定性动态规划的基本结构

对确定性动态规划问题进行分类的一种方法可以通过区分不同形式的目标函数实现，如可以划分为最大化各阶段对目标函数贡献的总和或者最小化各阶段对目标函数贡献的总和等类别。另外一种分类方法与各阶段状态的性质有关，典型的类别包含：状态 s_n 可以由离散类型的状态变量描

① 实际上，这个问题也可以选择正向递推的求解过程，但是对于多数动态规划求解的问题必须选择逆向递推的过程。

述(如驿站马车问题中的状态),同时状态也可以由连续的状态变量来描述,甚至有些问题中状态变量是向量性质的。类似地,决策变量($x_1,x_2,\cdots,x_n$)也可以存在离散和连续等不同的形式。

下面会通过几个例子对这些不同的类别的确定性动态规划进行说明,需要注意的是,虽然这些类别表面上存在差异,但其实质上都符合图11.3中描述的确定性动态规划基本结构,只是计算方面存在复杂性的差异。

接下来介绍的第一个例子与驿站马车问题具有相似的数学表达方式,只不过它的决策要求是将目标最大化。

应用案例

1990年8月2日,萨达姆·侯赛因命令伊拉克军队入侵科威特6天之后,美国开始将大批军队和物资运往波斯湾地区。在以美国为首的、来自35个国家的联合国军队集结完毕后,1991年1月17日发动了称为"沙漠风暴"的军事行动。联合国部队攻入伊拉克,科威特被解放,联合国部队获得了决定性的胜利。

将需要的部队和物资快速运往战区是一项艰难的物流任务。在一个典型的空运任务中,从美国本土出发的空运部队和物资到波斯湾往返需要3天,途径至少7个空域,消耗大约100万lb燃料,成本高达280000美元。在整个"沙漠风暴"行动期间,美国空军机动司令部(MAC)指挥着历史上最大规模的空运,平均每天要组织超过100次的类似空运。

为了满足需要,应用运筹学开发的决策支撑系统需要对每个空运任务进行调度和分配航线。驱动该过程的运筹学技术是动态规划方法。动态规划建模中的阶段对应于空运计划的飞行路径网络中的机场。对于一个给定的机场,状态用从机场的起飞时间及当前机组成员的剩余职责描述。要求最小化的目标函数是一个性能测度(包括交货的延误、任务的飞行时间、起落时间及机组成员改变的数量)的权重和。约束包括一次任务的最低载货量、机场可用机组成员及地面物资支持的上限。

动态规划在快速地将所需要的物资和军队运送到波斯湾支持"沙漠风暴"行动中发挥了意想不到的效果。MAC负责运营和运输的总参谋长与该方法的开发者交谈时,说:"我发誓,没有你们的帮助及你们所做的贡献(决策支持系统),我们是无法成功的(在波斯湾的部署),我们绝对无法做到。"

资料来源:M. C. Hilliard, R. S. Solanki, C. Liu, I. K. Busch, G. Harrison, and R. D. Kraemer: "Scheduling the Operation Desert Storm Airlift: An Advanced Autumated Scheduling Support System." Interfaas. 22(1), 131-146. Jan.-Feb. 1992.(以下网址提供本文链接:www. mhhe . com/ hillier。)

11.3.1 例2 医疗队分配问题

世界卫生组织致力于提升发展中国家的医疗水平,为实现这一目的,世界卫生组织现决定将5个医疗队分配到3个国家,用以改善这些国家的医疗服务、健康教育及培训。这种情况下,世界卫生组织需要考虑如何分配现有的队伍资源,以使效益最大化。同时,分配的过程中必须保证每支队伍的完整性,因此分配到每个国家的队伍数量都是整数。

用以评价效益的指标是人们的寿命增加(针对某一特定国家,这一指标的值等于寿命增加的期望值乘以该国人口数量)。表11.1中列出了每个国家的不同医疗队配备数量下,估算的寿命增加值(单位:千人·年)。解决这一问题需要确定如何分配才能使评价指标的数值最大。

表11.1 医疗队分配问题数据

医疗队	增加的人口寿命/(千人·年)		
	国家		
	1	2	3
0	0	0	0
1	45	20	50
2	70	45	70
3	90	75	80
4	105	110	100
5	120	150	130

建模：这个问题的解决需要作出 3 个相关决策，即分别分配多少个医疗队到 3 个国家，虽然这 3 个决策没有先后顺序之分，但是也可以视作动态规划模型的 3 个阶段。决策变量x_n（$n=1,2,3$）是分配给 n 阶段（第 n 个国家）的医疗队数目。

状态的确定稍微困难一些。为了确定状态，我们需要思考是什么推动一个阶段向下一个阶段发展？已知之前状态做出的决定，如何描述目前阶段的状态情况？目前状态的什么信息对决定后面的最优策略是必要的？基于这些考虑，选择状态的一种方法为

$$s_n = \text{用于分配给剩余国家的医疗队数目}$$

基于这种考虑，第 1 阶段（国家 1）开始由于还未进行分配，因此，$s_1=5$，在第 2 阶段和第 3 阶段，s_n是 5 减去前一阶段分配了的医疗队数目，所以状态依次为

$$s_1=5, s_2=5-x_1, s_3=s_2-x_2$$

由于一开始的分配数量不确定性，因此在每个阶段都需要充分考虑所有的分配可能性，即s_n可以为 0,1,2,3,4,5。

图 11.4 反映了每个阶段的所有可能状态，其中的连线表示从一个阶段到下一个阶段可能发生的转变。线段上的数字表示相应转变产生的用以评价的效益，数据来自于表 11.1。解决问题就是找到从状态 5 到状态 0 的路径，使该路径上数量之和的值最大。

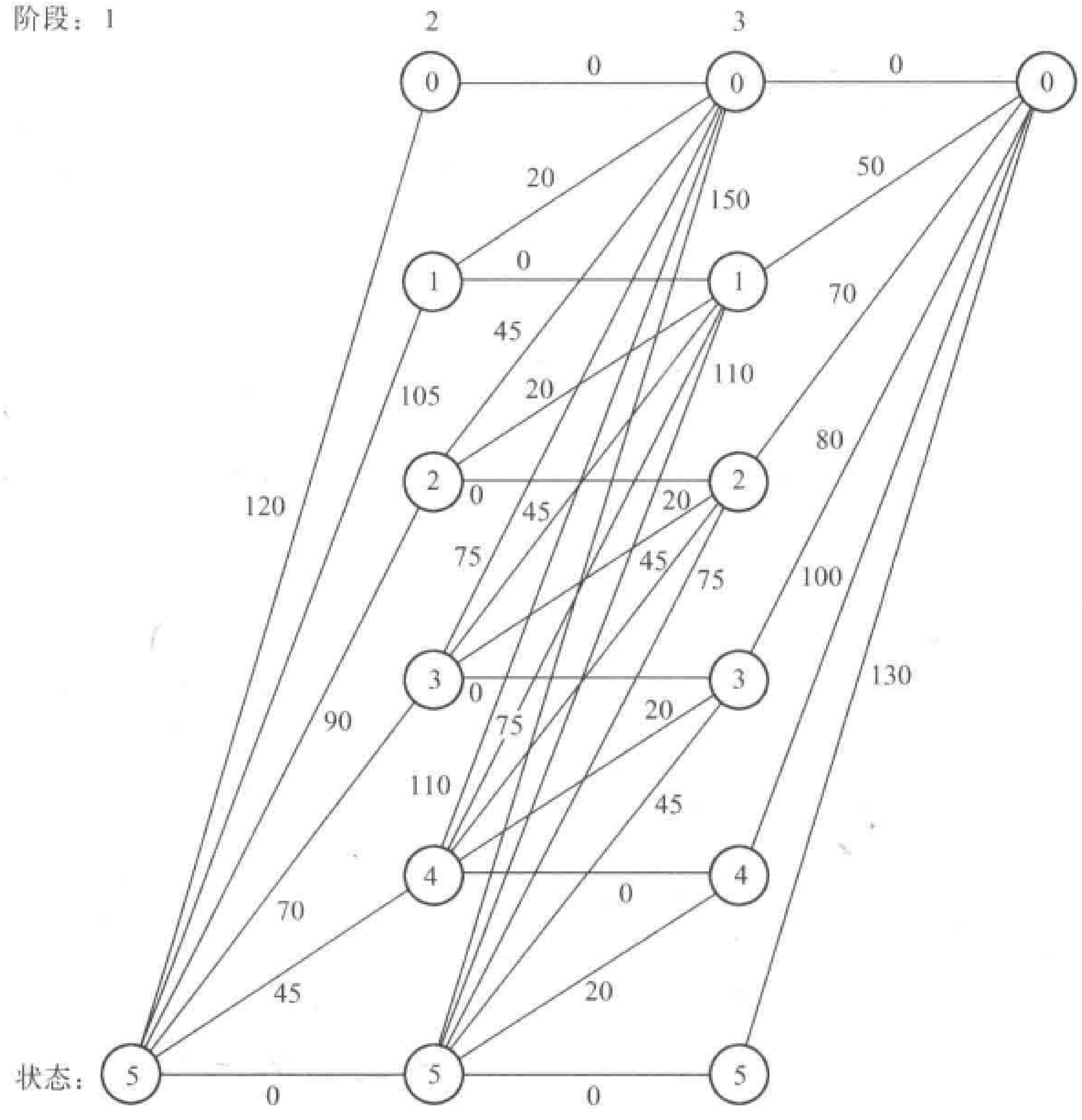


图 11.4　医疗队分配问题的可能状态和可能转化

为了用数学方法描述整个问题，设$p_i(x_i)$为向国家 i 分配x_i个医疗队后的寿命增长值，目标就是选择x_1、x_2、x_3，使

$$\max \sum_{i=1}^{3} p_i(x_i)$$

s. t.

$$\sum_{i=1}^{3} x_i = 5$$

且x_i为非负整数。

运用 11.2 节中提到的内容,$f_n(s_n, x_n)$为

$$f_n(s_n, x_n) = p_n(x_n) + \max \sum_{i=n+1}^{3} p_i(x_i)$$

其中

$$\sum_{i=n}^{3} x_i = s_n$$

且

$$f_n^*(s_n) = \max_{x_n=0,1,\cdots,s_n} f_n(s_n, x_n)$$

因此,有

$$f_n(s_n, x_n) = p_n(x_n) + f_{n+1}^*(s_n - x_n)$$

(将f_4^*定义为 0。)图 11.5 中对这些基本关系进行了总结。

因此,该问题中的递推关系用f_1^*、f_2^*和f_3^*可以表达为

$$f_n^*(s_n) = \max_{x_n=0,1,\cdots,s_n} \{p_n(x_n) + f_{n+1}^*(s_n - x_n)\}, n=1,2$$

对于 $n=3$,有

$$f_3^*(s_3) = \max_{x_3=0,1,\cdots,s_3} p_3(x_3)$$

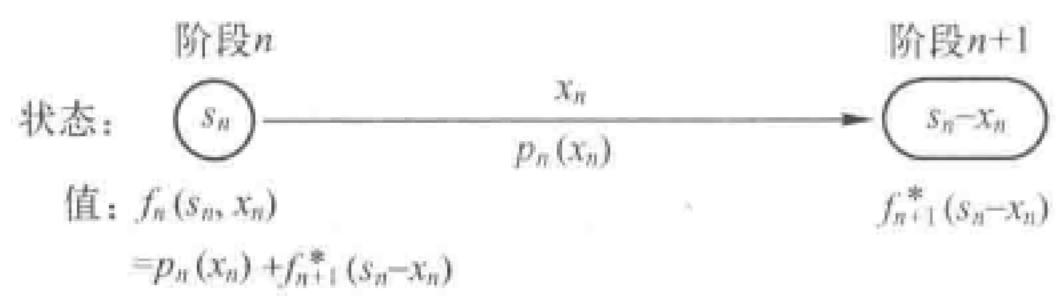


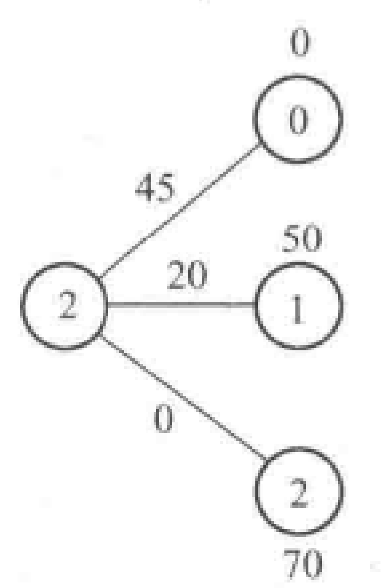
图 11.5 医疗队分配问题的基本结构图

下面对动态规划问题的计算过程进行介绍。

求解过程:从最后一个阶段($n=3$)开始,$p_3(x_3)$的值在表 11.1 的最后一列给出,在向国家 3 分配s_3个医疗队时,由于分配的医疗队越多$p_3(x_3)$的值越大,因此,$x_3^* = s_3$,且$f_3^*(s_3) = p_3(s_3)$,如下表所列:

s_3	$f_3^*(s_3)$	x_3^*
0	0	0
1	50	1
2	70	2
3	80	3
4	100	4
5	130	5

接下来是倒数第 2 个阶段($n=2$),找到x_2^*需要计算所有可能的x_2的$f_2(s_2, x_2)$值,并对其进行比较。这里选择$s_2=2$时进行说明。



上图截选自图 11.5,如果$x_2=0$,则第 3 阶段的状态将是$s_2 - x_2 = 2-0=2$;$x_2=1$,则第 3 阶段状态将是 1;$x_2=2$,则第 3 阶段状态将是 0。图中线段上的数字表示$p_2(x_2)$的值,第 3 阶段$f_3^*(s_2-x_2)$

的值在第 3 阶段状态点的上方标出。计算概括如下。

公式：$f_2(2,x_2)=p_2(x_2)+f_3^*(2-x_2)$

$p_2(x_2)$的值可以从表 11.1 的国家 2 那一列得到。

$f_3^*(2-x_2)$的值可以从第 3 阶段的 $n=3$ 的表中给出。

$x_2=0$: $f_2(2,0)=p_2(0)+f_3^*(2)=0+70=70$

$x_2=1$: $f_2(2,1)=p_2(1)+f_3^*(1)=20+50=70$

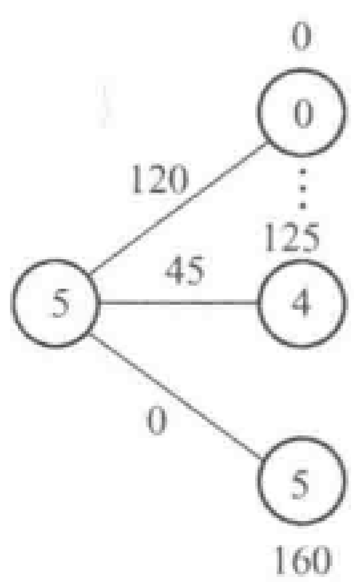
$x_2=2$: $f_2(2,2)=p_2(2)+f_3^*(0)=45+0=70$

为使目标函数最大化，所以$x_2^*=0$ 或者 1，$f_2^*(2)=70$。

用相同的方法可以得到s_2的其他可能值，如下表所列：

s_2 \ x_2	$f_2(s_2,x_2)=p_2(x_2)+f_3^*(s_2-x_2)$						$f_2^*(s_2)$	x_2^*
	0	1	2	3	4	5		
0	0						0	0
1	50	20					50	0
2	70	70	45				70	0 或 1
3	80	90	95	75			95	2
4	100	100	115	125	110		125	3
5	130	120	125	145	160	150	160	4

接下来从第 1 阶段开始求解整个问题。第 1 阶段，只需要考虑$s_1=5$ 的最初状态，如下图所示。



向第 1 个国家分配x_1个医疗队，使得第 2 阶段状态为$5-x_1$，因此，当$x_1=0$ 时，对应图中右侧最下方一个点，依此类推，$x_1=5$ 时，对应右侧最上方一个点。$p_1(x_1)$的值可以从表 11.1 中获取，图中标示在直线上，$f_2^*(s_2)$的值可以从第 2 阶段分析结束时的表格中获取，图中标示在右侧节点上方。计算过程与 $n=2$ 时相似，需要对决策变量的每个可能值进行计算，计算过程如下。

公式：$f_1(5,x_1)=p_1(x_1)+f_2^*(5-x_1)$

$p_1(x_1)$的值可以从表 11.1 的国家 1 那一列得到。

$f_2^*(5-x_1)$的值可以从第 2 阶段 $n=2$ 的表中给出。

$x_1=0$: $f_1(5,0)=p_1(0)+f_2^*(5)=0+160=160$

$x_1=1$: $f_1(5,1)=p_1(1)+f_2^*(4)=45+125=170$

⋮

$x_1=5$: $f_1(5,5)=p_1(5)+f_2^*(0)=120+0=120$

为使目标函数最大化，所以$x_1^*=1$，$f_1^*(5)=170$。具体结果见下表。

s_1 \ x_1	$f_1(s_1,x_1)=p_1(x_1)+f_2^*(s_1-x_1)$						$f_1^*(s_1)$	x_1^*
	0	1	2	3	4	5		
5	160	170	165	160	155	120	170	1

因此,实现最优解的方法是$x_1^*=1$,使$s_2=5-1=4$,从而$x_2^*=3$,使$s_3=4-3=1$,从而$x_3^*=1$,这种情况下,$f_1^*(5)=170$。最后的结论是向3个国家分别分配1个、3个、1个医疗队,会为总寿命带来大概170000人·年的增加量,比次优的方案多出5000人·年。图11.6对动态分析结果进行了概括。

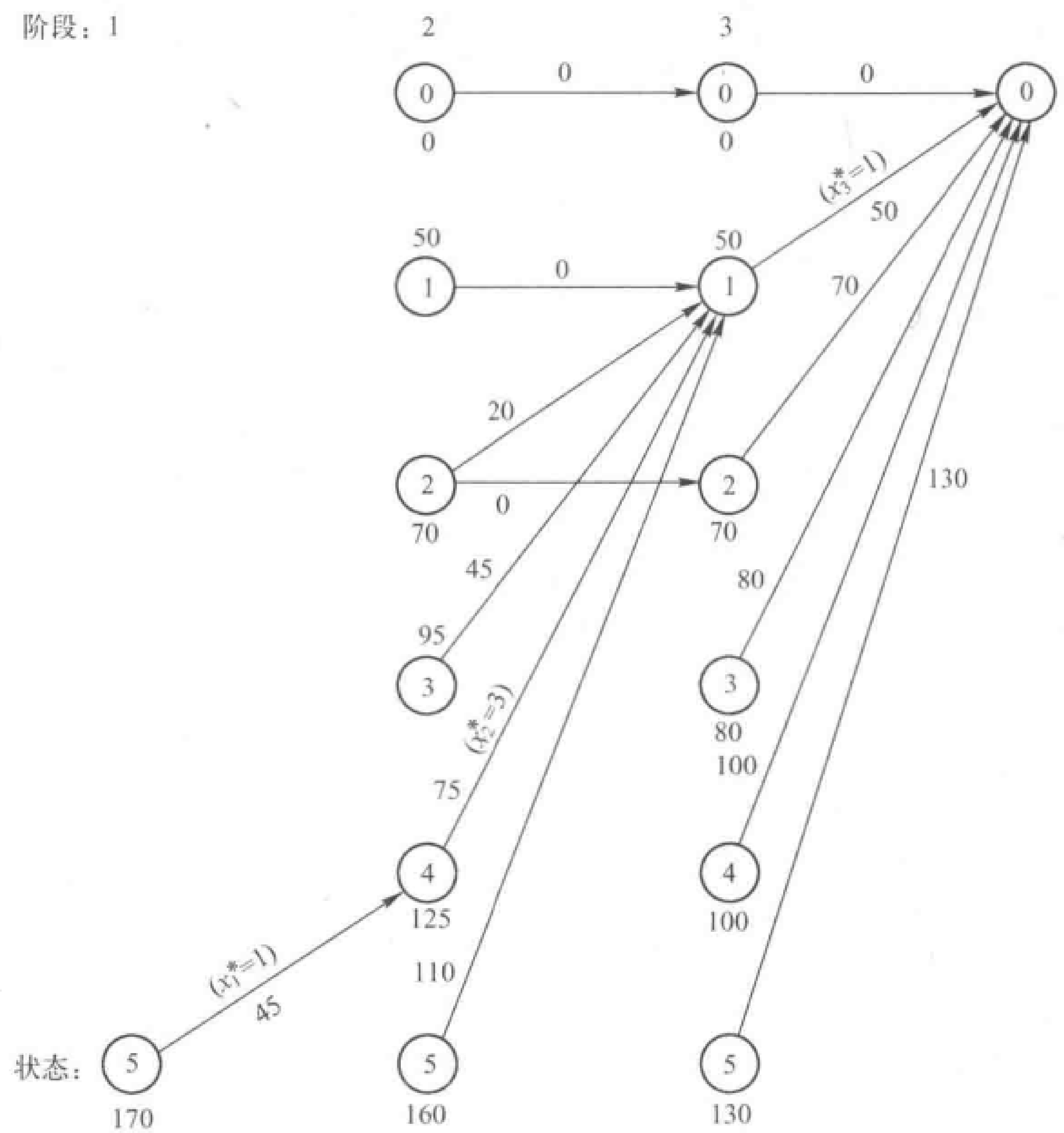


图11.6 医疗队分配问题动态规划求解方法的图形演示

11.3.2 一种常见的问题范例——工作分配问题

前面的例子展示了一种常见类型的动态规划问题,称为工作分配问题,这种类型的问题中,一种资源被分配到许多活动中,这一类型问题的目标是决定如何在活动中最有效地分配工作(资源)。医疗队分配问题的范例中,资源指的是医疗队,涉及的3个活动是在3个国家的医疗保健工作。

假设:分配资源给活动的这种解释是第3章开头的线性规划问题中的典型解释。但是工作分配问题和线性规划问题之间存在一些区别。在这里借助工作分配问题和线性规划问题间的区别,阐述动态规划和其他数学规划间的区别。

区别之一是工作分配问题仅涉及一种资源的分配(一个约束),线性规划可以处理上千种资源(原则上,动态规划问题也可以处理多种资源问题,如例5中的问题,但是随着资源数量的增长,采用动态规划的方法解决问题效率会变的很低)。

此外,工作分配问题远比线性规划问题普遍,3.3 节中提到线性规划问题需要具备比例性、可加性、可分性和确定性 4 个假设,但实际上,仅比例性这一个假设几乎就与所有的动态规划问题相违背,如表 11.1 中的数值就不符合比例性。可分性也经常不能够满足,如例 2 中,决策变量必须为整数,与可分性相违背,但当可分性成立时(如例 4),动态规划的计算将变的更为复杂。虽然我们应该在确定性的假设条件下考虑工作分配问题,但是这并非必需,并且许多动态规划问题都不符合这一假设(如 11.4 节中所述)。

线性规划的 4 个假设中,只有可加性是工作分配问题(或其他动态规划的问题)必须满足的,因为只有满足可加性的条件下,动态规划的最优性原则才能够成立(11.2 节中特点 5)。

模型:工作分配问题总是为多个活动分配一种资源,工作分配问题具有下述符合动态规划规则的公式(其中活动的排序是任意的):

阶段 n=活动 $n(n=1,2,\cdots,N)$

x_n=分配到活动 n 的资源数

状态s_n=分配到剩余活动中的可用资源数$(n,n+1,\cdots,N)$

这种方式定义的s_n正是时间当前状态的信息,我们需要用这种信息对剩余活动进行分配决策。

当系统从阶段 n 的状态s_n开始时,选择x_n将使位于 $n+1$ 阶段的下一个状态$s_{n+1}=s_n-x_n$,描述如下①:

上图与医疗队分配问题案例的结构具有一定相似性,不同的地方在于$f_n(s_n,x_n)$和$f_{n+1}^*(s_n-x_n)$之间的关系,以及f_{n+1}^*和f_n^*之间的递推关系。这些关系的确定取决于特定问题的目标函数,因此这里图中没有给出。

下面的例子与医疗队分配问题相似,也是工作分配问题,区别在于其目标是最小化相关阶段的产品数量。这个例子中涉及到概率的问题,会造成该问题不属于动态规划研究的范畴,但它的确符合我们关于动态规划问题的定义,例子中下一阶段的状态完全由当前阶段的状态和决策策略决定。

11.3.3　例 3　向科研小组分配科学家

一项政府空间计划正在进行某项工程问题的攻关,这个问题是人类登陆火星的关键。3 个科研小组目前正在尝试 3 种不同的方法解决这个问题,各组估计的(分别称为 1 组、2 组和 3 组)失败概率分别为 0.40、0.60 和 0.80。因此,整个问题攻关失败的概率为(0.40)(0.60)(0.80)=0.192,为最小化问题攻关失败的概率,现有另外 2 名科学家被分配到该工程项目中。

由于新加入的科学家需要全身心投入该组的工作,因此分配科学家的数目必须为整数,表 11.2中列出了各组在有 0 名、1 名或 2 名科学家加入时的失败概率。需要解决的问题是如何分配两名科学家才能够最小化整个工程项目失败的概率。

① 这种描述中假定s_n、x_n用相同单位表达,另一种更为方便的定义为:x_n是分配于第 n 项活动的资源总量,数量为$a_n x_n$,于是,有$s_{n+1}=s_n-a_n x_n$。

表 11.2 政府空间计划问题的数据

新科学家数目	失败概率		
	小组		
	1	2	3
0	0.40	0.60	0.80
1	0.20	0.40	0.50
2	0.15	0.20	0.30

建模:例 2 和例 3 都是工作分配问题,其基本结构具有较大相似性。例 3 中,科学家代替医疗队成为了被分配的资源,科研小组代替了国家成为了资源分配的去向,这两个问题之间唯一的区别是它们的目标函数。

例 3 中涉及到的科学家和小组数量较少,采用枚举法可以很容易地解决问题,但采用动态规划的方法解决此问题,有利于掌握此类问题的通用解法。

本例中,阶段 $n(n=1,2,3,)$ 与科研小组编号 n 相一致,状态s_n是能够被分配到剩余小组中的科学家数目,决策变量$x_n(n=1,2,3)$是额外分配到组 n 中的科学家数量。

设$p_i(x_i)$为向组 i 分配x_i个科学家后的失败概率值,数值由表 11.2 给出。政府的目标是选择合适的x_1、x_2、x_3,使

$$\min \prod_{i=1}^{3} p_i(x_i) = p_1(x_1)\ p_2(x_2)\ p_3(x_3)$$

s. t.

$$\sum_{i=1}^{3} x_i = 2$$

且x_i为非负整数。

因此,该问题中,$f_n(s_n,x_n)$为

$$f_n(s_n,x_n) = p_n(x_n) \cdot \min \prod_{i=n+1}^{3} p_i(x_i)$$

其中

$$\sum_{i=n}^{3} x_i = s_n$$

对 $n=1,2,3$,有

$$f_n^*(s_n) = \min_{x_n=0,1,\cdots,s_n} f_n(s_n,x_n)$$

其中

$$f_n(s_n,x_n) = p_n(x_n) \cdot f_{n+1}^*(s_n-x_n)$$

(将f_4^*定义为 1。)图 11.7 中对这些基本关系进行了总结。

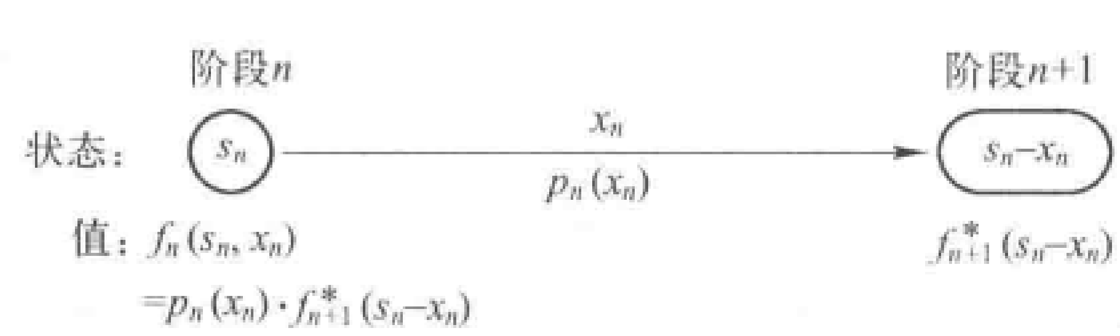


图 11.7 政府空间计划问题的基本结构

因此,该问题中的递推关系用f_1^* f_2^* 和f_3^* 可以表达为

$$f_n^*(s_n) = \min_{x_n=0,1,\cdots,s_n} \{p_n(x_n) \cdot f_{n+1}^*(s_n-x_n)\},\ n=1,2$$

对于 $n=3$，有

$$f_3^*(s_3)=\min_{x_3=0,1,\cdots,s_3} p_3(x_3)$$

求解过程：动态规划计算如下。

$n=3$ 时：

s_3	$f_3^*(s_3)$	x_3^*
0	0.80	0
1	0.50	1
2	0.30	2

$n=2$ 时：

s_2 \ x_2	$f_2(s_2,x_2)=p_2(x_2)\cdot f_3^*(s_2-x_2)$			$f_2^*(s_2)$	x_2^*
	0	1	2		
0	0.48			0.48	0
1	0.30	0.32		0.30	0
2	0.18	0.20	0.16	0.16	2

$n=1$ 时：

s_1 \ x_1	$f_1(s_1,x_1)=p_1(x_1)\cdot f_2^*(s_1-x_1)$			$f_1^*(s_1)$	x_1^*
	0	1	2		
2	0.064	0.060	0.060	0.060	1

因此，最优解的情况是$x_1^*=1, s_2=1; x_2^*=0, s_3=1; x_3^*=1$，最优的分配策略是小组 1 和小组 3 分别接收一名科学家，这种情况下，3 个组都失败的概率为 0.060。

之前的所有案例每个阶段都有一个离散的状态变量 s，并且这些案例都具有可逆性，体现在求解过程可以向后或者向前一步一步完成，并且活动的阶段可以按照需要任意排序，而不用考虑资源分配的先后顺序。这个可逆性是工作分配问题的一般特性。

下面的例子与之前的例子不同，阶段 n 的状态变量s_n没有必须为整数的约束条件，是可以任意取值的连续变量，这就导致分别考虑每一个可行值的办法变得不可行。这个例子不再是可逆的了，它的阶段与时间周期一致，求解过程必须从前向后进行。

在介绍下一个例子之前，可以先去本书（英文版）网站的案例解答部分学习另外两个关于确定性动态规划的例子。其中一个例子是关于多个时间阶段产品和存货的规划问题，这个例子中状态变量和决策变量都是离散的，但是由于案例中阶段与时间直接相关，因此是不可逆的过程，不属于工作分配问题；第 2 个案例属于非线性规划问题，具有两个变量和一个约束条件，过程虽然可逆，但是状态和决策变量都是连续的，与接下来要介绍的案例相比，这个案例只有两个阶段，求解思路和计算都比较简单。

11.3.4　例 4　车间雇佣问题

地方加工车间的工作量受季节影响波动较大，但是机床工人招聘难度较大，且培训费用高。因此，加工车间的经理不愿意在淡季时裁员，同时也不愿意为不需要的机床工人支付薪水。由于所有工作按照订单完成，无法在淡季时库存产品留给旺季用，所以在确定雇佣策略时，经理进退

两难。

下表是未来一年不同季节中,加工车间的最小雇佣需求的估算值(单位:个)。

季节	春	夏	秋	冬	春
需求	255	220	240	200	255

雇佣的数量不能够小于表中的数值,当雇佣数量高于表中的数值时,每人每个季节需要耗费掉约2000美元。从一个季节到下一个季节,雇佣数量发生变化时,成本为200乘以雇佣数量变化的平方。雇佣数量可以出现非整数的情况,因为雇佣中会存在一些兼职的工人,这使得成本的数据也可以为非整数。

建模:雇佣的人数超过旺季的需求显然是不划算的,因此,作为每年旺季的春季雇佣人数应当为255人(不能少于最低需求),问题就简化成了确定其他3个季节雇佣人数的问题。

从动态规划模型的角度来说,季节应该是阶段。问题随着时间无限延伸,阶段的数量应该是无限个,但是由于阶段以一年为周期循环,所以只需要考虑4个阶段。春季的雇佣人数是确定的,因此,将春季作为一个循环周期中的最后一个阶段,因为最后一个阶段的决策变量的最优值必须是已知或者不考虑其他阶段影响就可以得到的,其他季节的最优雇佣数量都需要考虑后面季节的雇用成本决定,具体如下:

阶段1=夏季

阶段2=秋季

阶段3=冬季

阶段4=春季

x_n=阶段 n 的雇佣数量,其中$x_4=255$。

设r_n为第 n 阶段最小雇佣数量,前面表中已经给出了每个阶段的r_n值。这样,x_n的取值范围为

$$r_n \leqslant x_n \leqslant 255$$

根据之前叙述部分提供的信息,可以确定

阶段 n 的成本$=200\,(x_n-x_{n-1})^2+2000(x_n-r_n)$

从上式可以看出,阶段 n 的成本仅取决于当前需要作出决策的x_n和上一阶段的雇佣数量x_{n-1},因此,当前阶段决策所需的全部信息都可以从之前阶段的状态中获得。第 n 阶段的状态为

$$s_n=x_{n-1}$$

当 $n=1$ 时,$s_1=x_0=x_4=255$。

为了便于理解,表11.3中列出了4个阶段的数据并进行了归纳。

表11.3 车间雇佣问题中的数据

n	r_n	x_n	$s_n=x_{n-1}$	成本
1	220	$220\leqslant x_1\leqslant 255$	$s_1=255$	$200\,(x_1-255)^2+2000(x_1-220)$
2	240	$240\leqslant x_2\leqslant 255$	$220\leqslant s_1\leqslant 255$	$200\,(x_2-x_1)^2+2000(x_2-240)$
3	200	$200\leqslant x_3\leqslant 255$	$240\leqslant s_2\leqslant 255$	$200\,(x_3-x_2)^2+2000(x_3-200)$
4	255	$x_4=255$	$200\leqslant s_3\leqslant 255$	$200\,(255-x_3)^2$

这个问题的目标是选择合适的x_1、x_2和x_3,使

$$\min \sum_{i=1}^{4} \left[200\,(x_i - x_{i-1})^2 + 2000(x_i - r_i) \right]$$

s. t.
$$r_i \leqslant x_i \leqslant 255, i=1,2,3,4$$

对第 n 阶段，由于$s_n = x_{n-1}$，有

$$f_n(s_n, x_n) = 200\,(x_n - s_n)^2 + 2000(x_n - r_n) + \min_{r_i \leqslant x_i \leqslant 255} \sum_{i=n+1}^{4} \left[200\,(x_i - x_{i-1})^2 + 2000(x_i - r_i) \right]$$

当 $n=4$ 时，上式没有意义。由于

$$f_n^*(s_n) = \min_{r_n \leqslant x_n \leqslant 255} f_n(s_n, x_n)$$

因此，有

$$f_n(s_n, x_n) = 200\,(x_n - s_n)^2 + 2000(x_n - r_n) + f_{n+1}^*(x_n)$$

由于阶段 4 之后的成本与分析过程无关，因此，将f_5^* 定义为 0。为便于理解，在图 11.8 中对这些关系进行了概括。

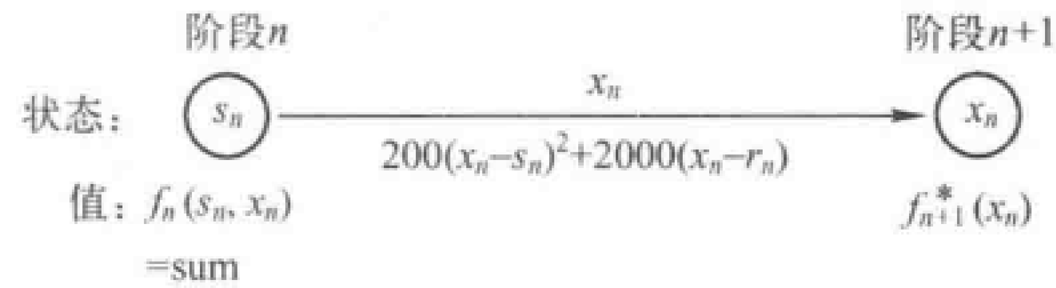


图 11.8　车间雇佣问题的基本结构

因此，函数f_n^* 的递推关系为

$$f_n^*(s_n) = \min_{r_n \leqslant x_n \leqslant 255} \{ 200\,(x_n - s_n)^2 + 2000(x_n - r_n) + f_{n+1}^*(x_n) \}$$

动态规划中运用这一关系确定$f_4^*(s_4)$、$f_3^*(s_3)$、$f_2^*(s_2)$、$f_1^*(s_1)$和它们对应的x_n值。

求解过程：阶段 4：由于已知$x_4^* = 255$，因此其结果为

s_4	$f_4^*(s_4)$	x_4^*
$200 \leqslant s_4 \leqslant 255$	$200\,(255-s_4)^2$	255

阶段 3：问题仅包含后两个阶段时，递推关系可以简化为

$$f_3^*(s_3) = \min_{200 \leqslant x_3 \leqslant 255} \{ 200\,(x_3 - s_3)^2 + 2000(x_3 - 200) + f_4^*(x_3) \}$$
$$= \min_{200 \leqslant x_3 \leqslant 255} \{ 200\,(x_3 - s_3)^2 + 2000(x_3 - 200) + 200\,(255 - x_3)^2 \}$$

其中

$$240 \leqslant s_3 \leqslant 255$$

对于特定的s_3，可以使用图解法确定使$f_3(s_3, x_3)$最小的x_3的值，如图 11.9 所示。

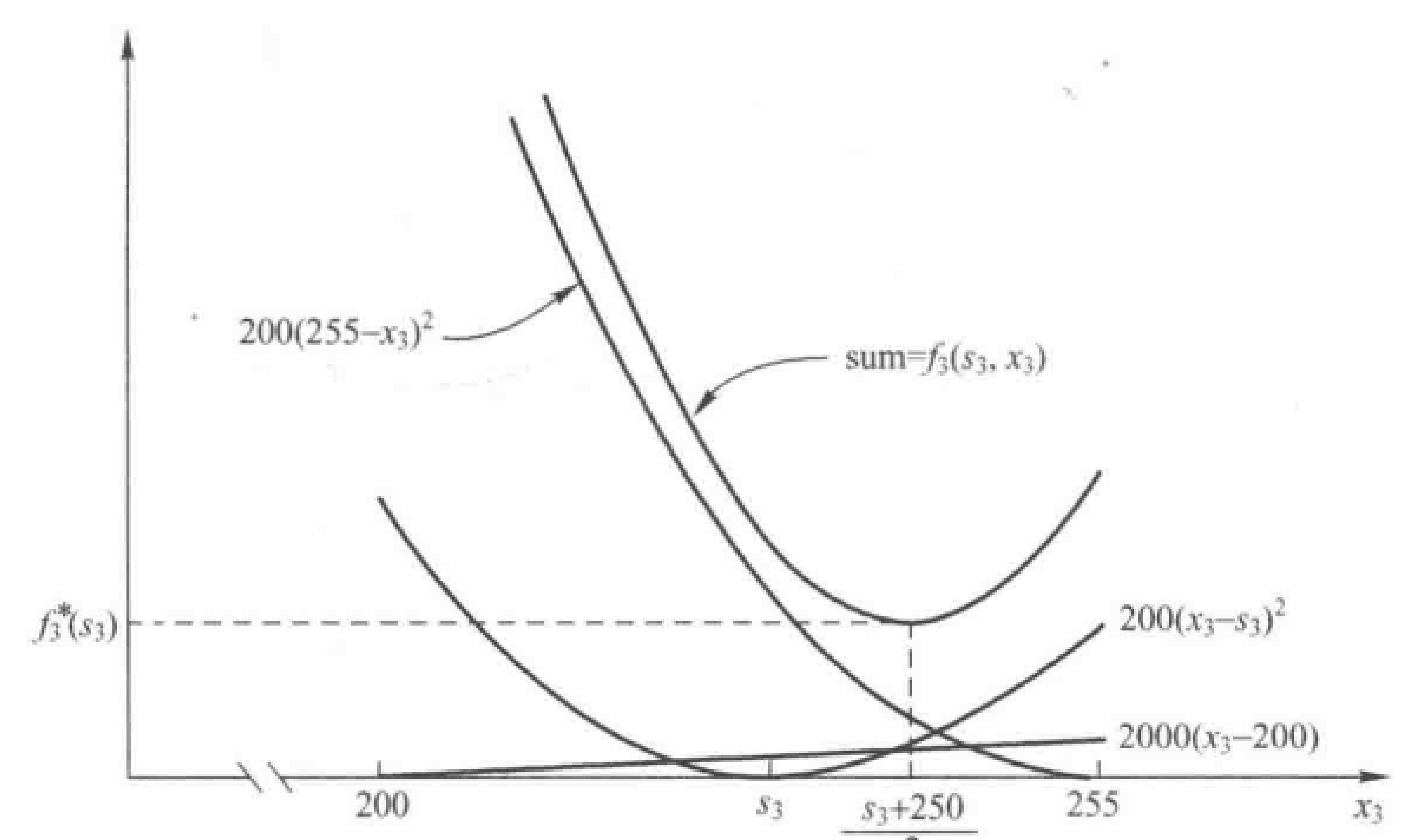


图 11.9　车间雇佣问题阶段 3 图解法

然而,使用微积分的方法可以更快地得到结果。由于阶段3开始时s_3是定值(目前不明确),因此最终得到x_3的最小值是用s_3表示的,当$f_3(s_3,x_3)$对x_3的偏导数为0时,解出的x_3为使$f_3(s_3,x_3)$值最小时的x_3,即

$$\begin{aligned}\frac{\partial f_3(s_3,x_3)}{\partial x_3}&=400(x_3-s_3)+2000-400(255-x_3)\\&=400(2x_3-s_3-250)\\&=0\end{aligned}$$

由于2阶导数大于零,得出$x_3^*=\dfrac{s_3+250}{2}$,由于对于任何s_3的可能值,x_3^*都在x_3的可行域内,因此求得的x_3^*符合条件。

从阶段3的求解过程就可以看出,这个例子与之前例子的不同,这个例子中由于s_3可以取无限多连续的值,因此x_3^*也存在无限多的值,而不能和之前例子一样,通过有限个s_3,一一尝试之后就可以确定每个x_3^*。这里我们将s_3作为未知变量,用它对x_3^*进行了表示,同时,$f_3^*(s_3)$也可以用s_3进行表示,即

$$f_3^*(s_3)=f_3(s_3,x_3^*)=200\left(\frac{s_3+250}{2}-s_3\right)^2+200\left(255-\frac{s_3+250}{2}\right)^2+2000\left(\frac{s_3+250}{2}-200\right)$$

对上式进行简化,可以得到第3阶段的分析结果,如下表所列:

s_3	$f_3^*(s_3)$	x_3^*
$240\leqslant s_3\leqslant 255$	$50(250-s_3)^2+50(260-s_3)^2+1000(s_3-150)$	$\dfrac{s_3+250}{2}$

阶段2:第2阶段和第1阶段求解的方法类似。$n=2$时,有

$$\begin{aligned}f_2(s_2,x_2)&=200(x_2-s_2)^2+2000(x_2-r_2)+f_3^*(x_2)\\&=200(x_2-s_2)^2+2000(x_2-240)+50(250-x_2)^2+50(260-x_2)^2+1000(x_2-150)\end{aligned}$$

s_2的可能值是$220\leqslant s_2\leqslant 255$,$x_2$的可行域是$240\leqslant x_2\leqslant 255$,现在需要找出满足如下关系的$x_2$,即

$$f_2^*(s_2)=\min_{240\leqslant x_2\leqslant 255}f_2(s_2,x_2)$$

与第3阶段相同,采用微积分的方法,令$f_2(s_2,x_2)$关于x_2的偏导数为0,则

$$\begin{aligned}\frac{\partial f_2(s_2,x_2)}{\partial x_2}&=400(x_2-s_2)+2000-100(250-x_2)-100(260-x_2)+1000\\&=200(3x_2-2s_2-240)=0\end{aligned}$$

2阶导数大于零,得出$x_2=\dfrac{2s_2+240}{3}$,由于仅$240\leqslant s_2\leqslant 255$时,该$x_2$可以落在可行域内,针对剩余的$s_2$还需要找出使$f_2(s_2,x_2)$最小的$x_2$可行值。

当$s_2<240$时,有

$$\frac{\partial f_2(s_2,x_2)}{\partial x_2}>0$$

其中

$$240\leqslant x_2\leqslant 255$$

于是,使$f_2(s_2,x_2)$最小的x_2可行值是240。将上述两种情况下的x_2值带入$f_2(s_2,x_2)$,可以求得$f_2^*(s_2)$,具体结果如下表所列:

s_2	$f_2^*(s_2)$	x_2^*
$220\leqslant s_2\leqslant 240$	$200\ (240-s_2)^2+115000$	240
$240\leqslant s_2\leqslant 255$	$\frac{200}{9}[\ (240-s_2)^2+(225-s_2)^2+(270-s_2)^2\ ]+2000(s_2-195)$	$\frac{2s_2+240}{3}$

阶段 1：对于第 1 阶段，有

$$f_1(s_1,x_1)=200\ (x_1-s_1)^2+2000(x_1-r_1)+f_2^*(x_1)$$

其中$r_1=220$，x_1需满足 $220\leqslant x_1\leqslant 255$，但是$x_1$在 $220\leqslant x_1\leqslant 240$ 和 $240\leqslant x_1\leqslant 255$ 两个区间里，$f_2^*(x_1)$具有不同的表达形式。

因此，当 $220\leqslant x_1\leqslant 240$ 时，有

$$f_1(s_1,x_1)=200\ (x_1-s_1)^2+2000(x_1-220)+200(240-x_1)^2+115000$$

当 $240\leqslant x_1\leqslant 255$ 时，有

$$f_1(s_1,x_1)=200\ (x_1-s_1)^2+2000(x_1-220)+\frac{200}{9}[\ (240-x_1)^2+(225-x_1)^2+(270-x_1)^2\ ]+2000(x_1-195)$$

考虑第 1 种情况，当 $220\leqslant x_1\leqslant 240$ 时，有

$$\frac{\partial f_1(s_1,x_1)}{\partial x_1}=400(x_1-s_1)+2000-400(240-x_1)=400(2x_1-s_1-235)$$

其中已知$s_1=255$，$\frac{\partial f_1(s_1,x_1)}{\partial x_1}=800(x_1-245)$，考虑$x_1$的取值范围，$\frac{\partial f_1(s_1,x_1)}{\partial x_1}$恒小于 0，因此，$x_1=240$ 时，$f_1(s_1,x_1)$最小。

当 $240\leqslant x_1\leqslant 255$ 时，有

$$\begin{aligned}\frac{\partial f_1(s_1,x_1)}{\partial x_1}&=400(x_1-s_1)+2000-\frac{400}{9}[\ (240-x_1)+(225-x_1)+(270-x_1)\ ]+2000\\&=\frac{400}{3}(4x_1-3s_1-225)\end{aligned}$$

由于 2 阶导数大于 0，令

$$\frac{\partial f_1(s_1,x_1)}{\partial x_1}=0$$

得出$x_1=\frac{3s_1+225}{4}$，由于$s_1=255$，$x_1=247.5$，满足可行域的要求，因此，当 $240\leqslant x_1\leqslant 255$ 时，$x_1=247.5$ 使$f_1(s_1,x_1)$最小。

比较$x_1=240$ 和$x_1=247.5$ 时的$f_1(s_1,x_1)$的大小，可以发现，$x_1=247.5$ 时，$f_1(s_1,x_1)$更小。它的最终值为

$$\begin{aligned}f_1^*(225)=&\ 200\ (247.5-255)^2+2000(247.5-220)\\&+\frac{200}{9}[\ 2\ (250-247.5)^2+(265-247.5)^2+30(742.5-575)\ ]=185000\end{aligned}$$

第 1 阶段的结论如下表所列：

s_1	$f_1^*(s_1)$	x_1^*
255	185000	247.5

根据第 1 阶段的结论，并设$s_n=x_{n-1}^*$，可以一次求得 4 个阶段的x_n^*分别为 247.5、245、247.5、255，每个周期的成本为 185000 美元。

虽然之前已经讲解了多个关于动态规划的案例,并且在本章接下来的部分,还将继续进行案例的分析,但是这些案例仅仅涉及动态规划表层的东西。第2章的参考文献[2]中描述了可运用动态规划求解的47种不同问题(文章中还提供了求解这些问题的软件工具),这些动态规划应用的目的都是为一系列相关的活动高效地寻找最优的决策组合。

11.4 随机性动态规划

随机性动态规划不同于确定性动态规划,它在下一阶段的状态不完全由当前阶段状态和决策策略决定,下一阶段会按照一定的概率存在多种状态,但是概率分布由当前阶段的状态和决策策略决定。随机性动态规划的基本结构如图11.10所示。

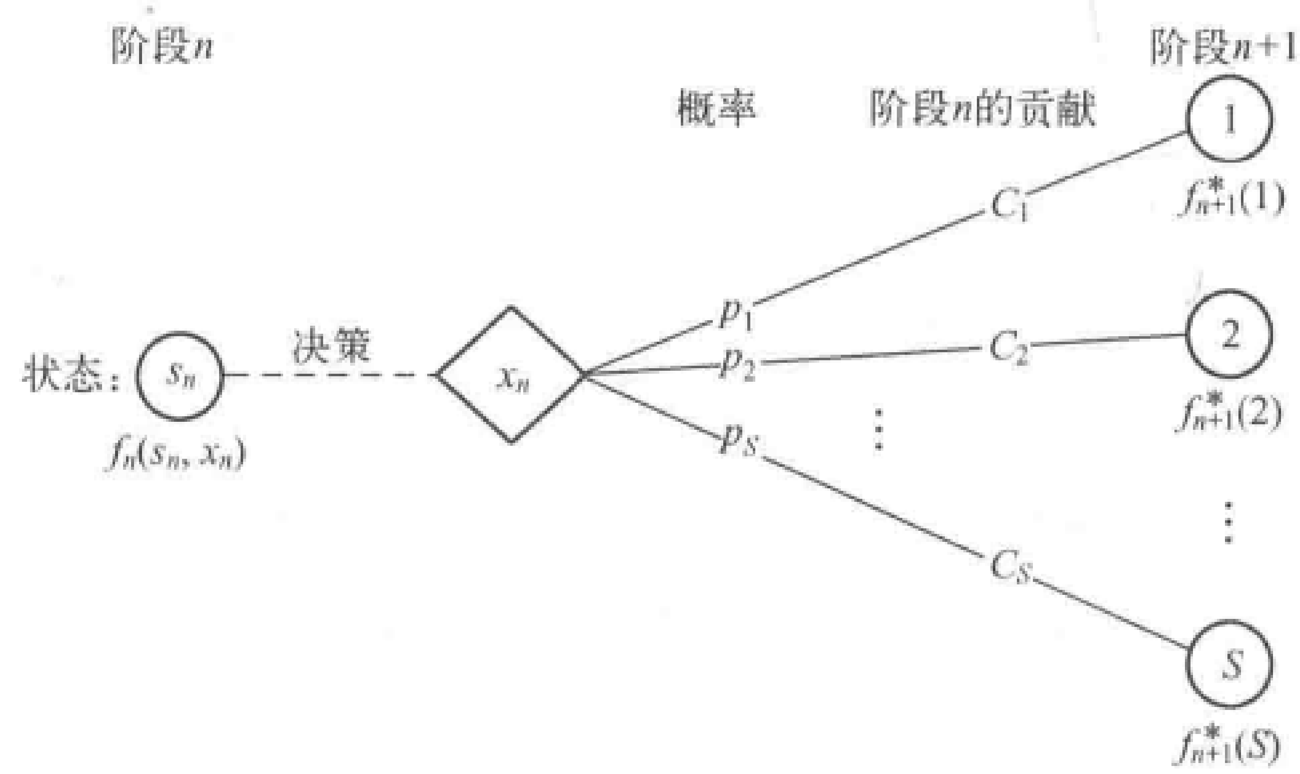


图11.10 随机性动态规划基本结构图

基于图11.10,设S为阶段$n+1$的可能状态,将这些状态标记为$1,2,\cdots,S$,在阶段n状态s_n下,做出x_n的决策后,系统有$p_i(i=1,2,\cdots,S)$的概率进入状态i,这种情况下,设C_i为阶段n对目标函数的贡献。

当图11.10扩展到包含所有阶段的可能状态和决策时,可以称为决策树。由于随机性的原因,$f_n(s_n,x_n)$和$f^*_{n+1}(s_{n+1})$之间的关系要比确定性动态规划复杂,其精确形式取决于目标函数。为了说明问题,这里假设目标是每个阶段某要素求和的最小化,用$f_n(s_n,x_n)$表示阶段n之前的要素和,已知第n阶段状态和决策策略,则

$$f_n(s_n,x_n)=\sum_{i=1}^{S} p_i[C_i+f^*_{n+1}(i)]$$

其中

$$f^*_{n+1}(i)=\min_{x_n} f_{n+1}(i,x_{n+1})$$

接下来介绍的例5就是这种形式,例6属于另外一种形式。

11.4.1 例5 确定次品限额

HIT-AND-MISS公司接受了一份特殊类型的产品供应订单,订单客户对于质量的要求十分严格,公司需要生产多个产品以保证生产出合格的产品。生产周期内生产的额外产品的数量称为次品限额。

公司估计每个该类型的产品生产时,成品率为$\frac{1}{2}$,次品率也为$\frac{1}{2}$。批量生产L件产品成品的

数量符合二项式分布，全部次品概率为$\left(\frac{1}{2}\right)^{L}$。

每件产品的边际成本为 100 美元（即使产品不合格），额外的产品没有任何价值。每个产品设立生产过程需要花费 300 美元的固定成本，如果所有产品被检出不合格，要以同样的成本设立新的生产过程。由于时间所限，公司只有 3 个生产周期的时间，如果经过 3 个生产周期（只有 3 次设立生产过程的时间），仍未能产出一件合格的产品，则不仅没有收入，还会被罚款 1600 美元。

问题是确定每个生产周期应制造几件产品（次品数量+1），使总的生产成本最小。

建模：这个问题的动态规划模型是：

阶段 n = 第 n 个生产周期（$n=1,2,3$）

x_n = 阶段 n 的产品数量

s_n = 从阶段 n 开始仍需要生产的成品数量（1 或者 0）

在阶段 1，$s_1=1$，如果后续至少得到 1 个成品，则$s_n=0$，此后没有额外成本增加。

该问题的目标是：

$f_n(s_n,x_n)$ = 阶段 $n,\cdots,3$ 的全部预期成本，如果系统第 n 阶段开始于状态s_n，直接决策为x_n，之后的最优决策为

$$f_n^*(s_n)=\min_{x_n=0,1,\cdots} f_n(s_n,x_n)$$

式中：$f_n^*(0)=0$。简单起见，接下来的分析中用 100 美元为单位，不考虑下一状态的情况下，第 n 阶段的对成本的贡献是$[K(x_n)+x_n]$，其中 $K(x_n)$ 是关于x_n的函数，具体形式为

$$K(x_n)=\begin{cases}0, x_n=0\\3, x_n=1\end{cases}$$

因此，对于$s_n=1$ 的情况，有

$$\begin{aligned}f_n(1,x_n)&=K(x_n)+x_n+\left(\frac{1}{2}\right)^{x_n}f_{n+1}^*(1)+\left[1-\left(\frac{1}{2}\right)^{x_n}\right]f_{n+1}^*(0)\\&=K(x_n)+x_n+\left(\frac{1}{2}\right)^{x_n}f_{n+1}^*(1)\end{aligned}$$

如果最终没有得到合格的产品$f_4^*(1)=16$，图 11.11 对基本关系进行了概括。

对于 $n=1,2,3$，动态规划计算的递推关系为

$$f_n^*(1)=\min_{x_n=0,1,\cdots}\left\{K(x_n)+x_n+\left(\frac{1}{2}\right)^{x_n}f_{n+1}^*(1)\right\}$$

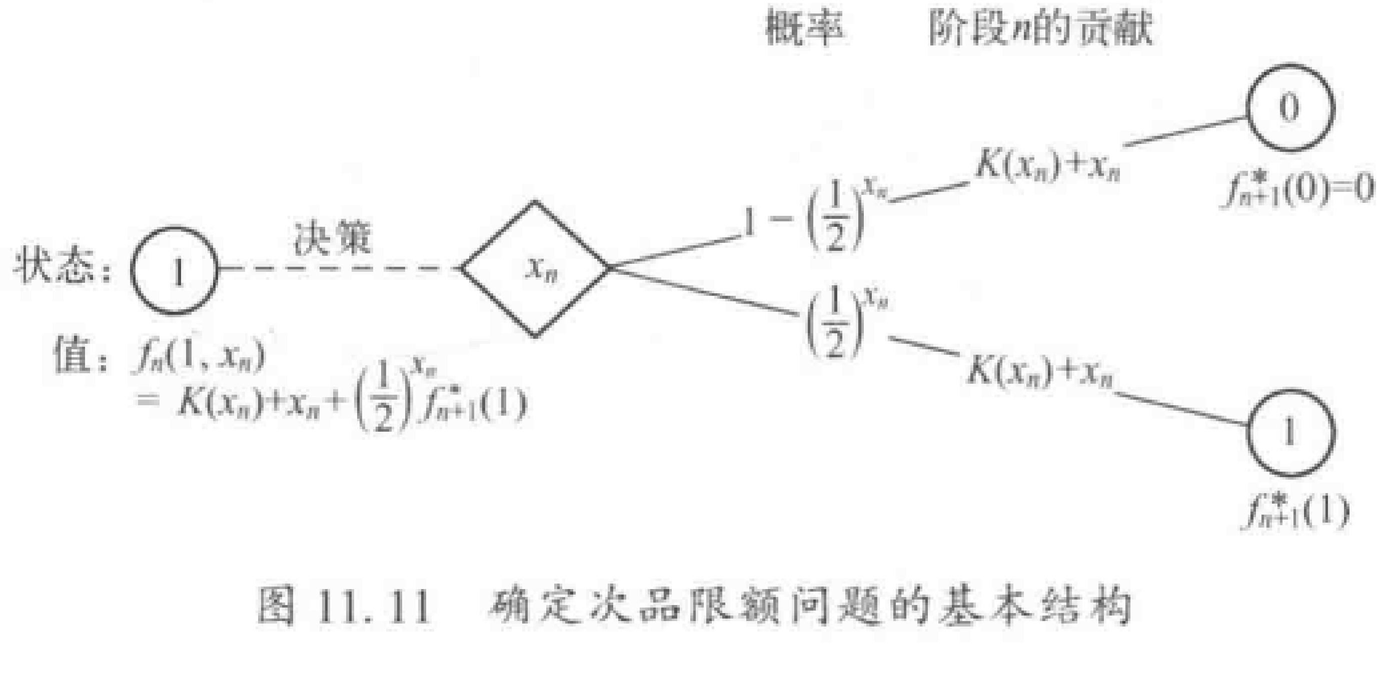


图 11.11　确定次品限额问题的基本结构

求解过程：用上述递推关系进行计算，结果如下。

$n=3$ 时：

s_3 \ x_3	$f_3(1,x_3)=K(x_3)+x_3+16\left(\frac{1}{2}\right)^{x_3}$						$f_3^*(s_3)$	x_3^*
	0	1	2	3	4	5		
0	0						0	0
1	16	12	9	8	8	$8\frac{1}{2}$	8	3 或 4

$n=2$ 时:

s_2 \ x_2	$f_2(1,x_2)=K(x_2)+x_2+\left(\frac{1}{2}\right)^{x_2}f_3^*(1)$					$f_2^*(s_2)$	x_2^*
	0	1	2	3	4		
0	0					0	0
1	8	8	7	7	$7\frac{1}{2}$	7	2 或 3

$n=1$ 时:

s_1 \ x_1	$f_1(1,x_1)=K(x_1)+x_1+\left(\frac{1}{2}\right)^{x_1}f_2^*(1)$					$f_1^*(s_1)$	x_1^*
	0	1	2	3	4		
1	7	$7\frac{1}{2}$	$6\frac{3}{4}$	$6\frac{7}{8}$	$7\frac{7}{16}$	$6\frac{3}{4}$	2

得出结论,最优策略时,第 1 个生产周期中生产 2 个产品;如果都不合格,第 2 个生产周期中生产 2 个或者 3 个产品;如果还是都不合格,第 3 个生产周期生产 3 个或 4 个产品。该策略预计总期望成本为 675 美元。

11.4.2 例 6 在拉斯维加斯赢钱

一位年轻的统计学家相信自己开发的一套系统可以在拉斯维加斯的一款常见赌博游戏中赢钱。她的同事们不相信,于是和她打了一个赌。打赌的条件是:如果她开始时有 3 个筹码,在 3 局之后,她剩余的筹码数不会超过 5 个。每局可以将手上有的筹码以任意数量下注,游戏的结果要么输掉下注的筹码,要么赢得相同数量的筹码。统计学家相信她的系统每一局都有$\frac{2}{3}$概率的胜算。

假设她估计的概率是正确的,现在使用动态规划决定 3 局游戏每次下注多少个筹码。每局的决策应该考虑上一局的结果,目标是使她赢得同事的可能性最大。

建模。这个问题的动态规划模型是:

阶段 n=第 n 局游戏($n=1,2,3$)

x_n=第 n 局游戏需要下的筹码数

s_n=第 n 局下注前手中的筹码数

用第 n 局下注前手中的筹码数定义状态,可以为该阶段提供决策前所需的所有信息。

$f_n(s_n,x_n)$= 3 局结束后有至少 5 个筹码的概率。目标是使该概率最大,即

$$f_n^*(s_n)=\max_{x_n=0,1,\cdots,s_n} f_n(s_n,x_n)$$

假如某一局输了,下一阶段的状态就变为了 s_n-x_n,完成时至少有 5 个筹码的概率应该是 $f_{n+1}^*(s_n-x_n)$,如果赢下这一局,则下一阶段状态变为了 s_n+x_n,完成时至少有 5 个筹码的概率应该

是$f_{n+1}^*(s_n+x_n)$，那么，有

$$f_n(s_n,x_n)=\frac{1}{3}f_{n+1}^*(s_n-x_n)+\frac{2}{3}f_{n+1}^*(s_n+x_n)$$

式中：当$s_4<5$时，$f_4^*(s_4)=0$；当$s_4\geqslant 5$时，$f_4^*(s_4)=1$。从表达式中可以看出，阶段 n 的目标函数在下一状态才产生影响。图 11.12 概括了这种关系。

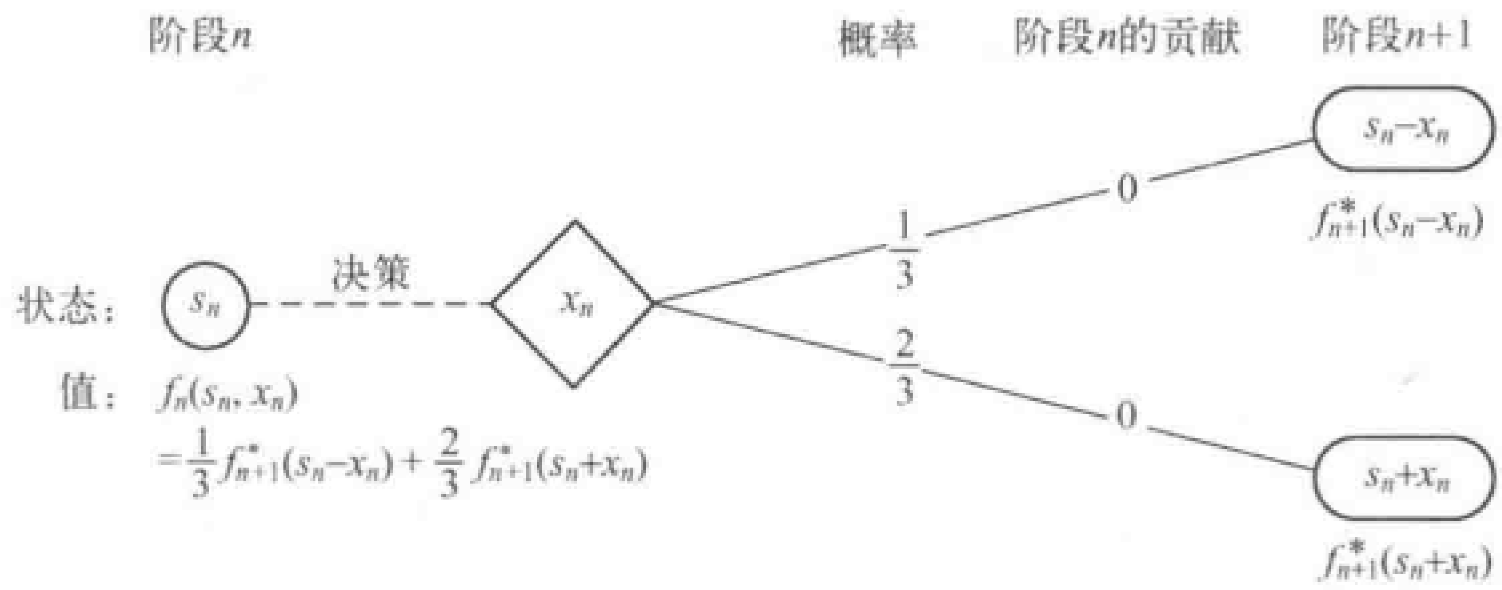


图 11.12　拉斯维加斯赌博问题结构

这个问题的递推关系为

$$f_n^*(s_n)=\max_{x_n=0,1,\cdots,s_n}\left\{\frac{1}{3}f_{n+1}^*(s_n-x_n)+\frac{2}{3}f_{n+1}^*(s_n+x_n)\right\}$$

求解过程：利用递推关系，可以得到如下计算结果。

$n=3$ 时：

s_3	$f_3^*(s_3)$	x_3^*
0	0	—
1	0	—
2	0	—
3	$\frac{2}{3}$	2 或更多
4	$\frac{2}{3}$	1 或更多
≥5	1	0

$n=2$ 时：

s_2 \ x_2	$f_2(s_2,x_2)=\frac{1}{3}f_3^*(s_2-x_2)+\frac{2}{3}f_3^*(s_2+x_2)$					$f_2^*(s_2)$	x_2^*
	0	1	2	3	4		
0	0					0	—
1	0	0				0	—
2	0	$\frac{4}{9}$	$\frac{4}{9}$			$\frac{4}{9}$	1 或 2
3	$\frac{2}{3}$	$\frac{4}{9}$	$\frac{2}{3}$	$\frac{2}{3}$		$\frac{2}{3}$	0、2 或 3
4	$\frac{2}{3}$	$\frac{8}{9}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{8}{9}$	1
5	1					1	0 或s_2-5

$n=1$ 时：

s_1 \ x_1	$f_1(s_1,x_1)=\frac{1}{3}f_2^*(s_1-x_1)+\frac{2}{3}f_2^*(s_1+x_1)$				$f_1^*(s_1)$	x_1^*
	0	1	2	3		
3	$\frac{2}{3}$	$\frac{20}{27}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{20}{27}$	1

因此，最佳策略为

$$x_1^*=1\begin{cases}\text{赢}, x_2^*=1\begin{cases}\text{赢}, x_3^*=0\\ \text{输}, x_3^*=2\text{ 或 }3\end{cases}\\ \text{输}, x_2^*=1\text{ 或 }2\begin{cases}\text{赢}, x_3^*=\begin{cases}2\text{ 或 }3(x_2^*=1\text{ 的情况})\\ 1,2,3,4(x_2^*=2\text{ 的情况})\end{cases}\\ \text{输，最后无法赢得赌局}\end{cases}\end{cases}$$

这种策略下，统计学家赢得和同事赌局的概率为$\frac{20}{27}$。

11.5 结　论

动态规划对制定一系列相关的决策非常有用，相比于穷举法，它能够大大节省计算量，例如，某问题包含 10 个阶段、10 个状态，这样每个阶段就会具有 10 个可能的策略，假如使用穷举法解决问题就需要考虑 100 亿个组合，而运用动态规划的方法，仅需要不超过 1000 次的计算。

本章仅考虑了阶段数量有限的动态规划。

部分参考文献

[1] Denardo, E. V.: *Dynamic Programming: Models and Applications*, Dover Publications, Mineola, NY, 2003.

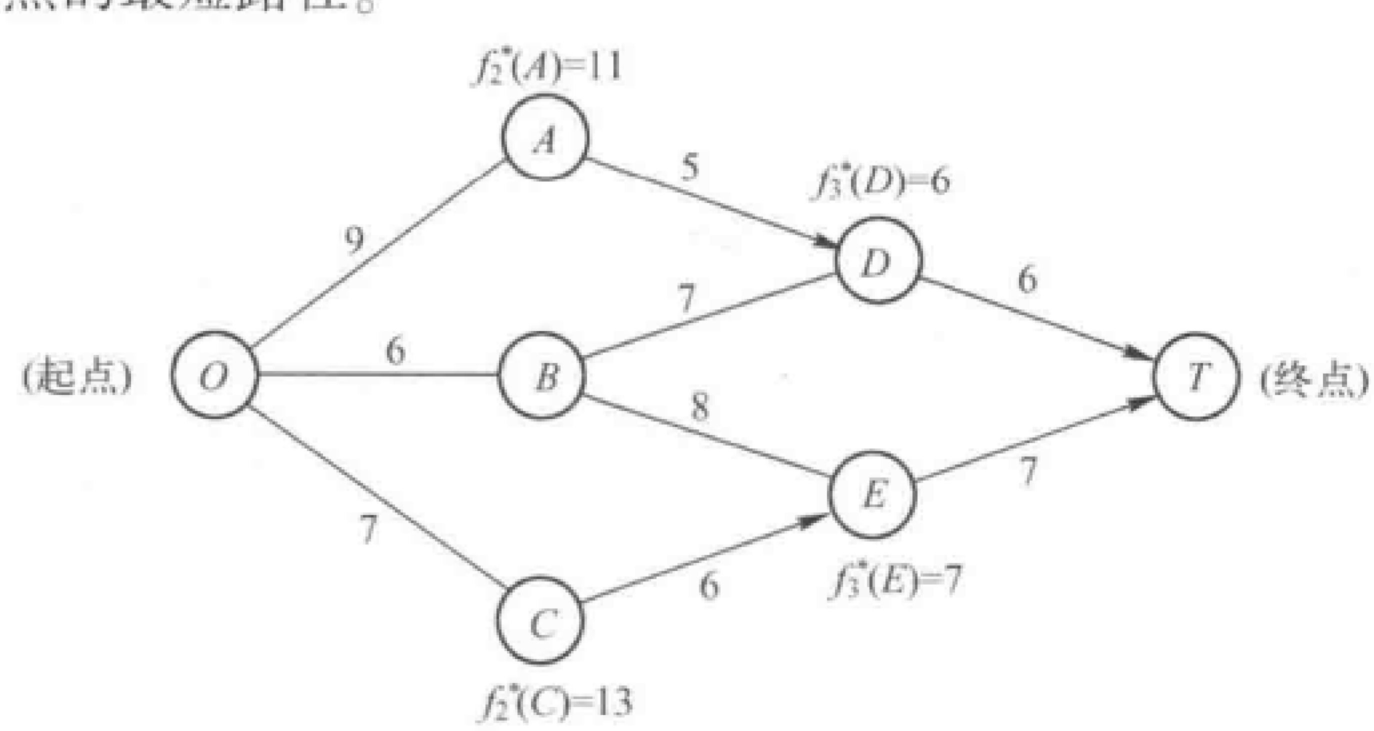
[2] Lew, A., and H. Mauch: *Dynamic Programming: A Computational Tool*, Springer, New York, 2007.

[3] Sniedovich, M.: *Dynamic Programming: Foundations and Principles*, Taylor & Francis, New York, 2010.

习　题

习题上有星号表示书后至少给出该题一部分答案。

11.2-1　考虑下面的网络图，其中沿线的每个数字表示了通过线段连接的两点之间的距离，找出从始点到终点的最短路径。



(a) 该问题的动态规划模型的阶段和状态是什么？

运用动态规划解决这个问题。用图解法表示求解过程（图 11.2）。图中有 4 个节点，已知 $f_n^*(s_n)$，求解 $f_2^*(B)$ 和 $f_1^*(O)$，然后用箭头在图中标明最优路线。

(b) 运用动态规划方法，建立 $n=3$、$n=2$、$n=1$ 时的表格求解这个问题。

(c) 使用 9.3 节中最短路径法解决这个问题，并与(b)和(c)中的方法比较。

11.2-2　某大学教材出版社的销售经理手下有 6 名推销员，他将这些推销员分配到 3 个不同的区域。每个区域应至少派一名推销员，并且每名推销员只能严格负责该区域的推销工作。现在他需要考虑，为使销售量最大，每个区域应该分配多少名推销员？

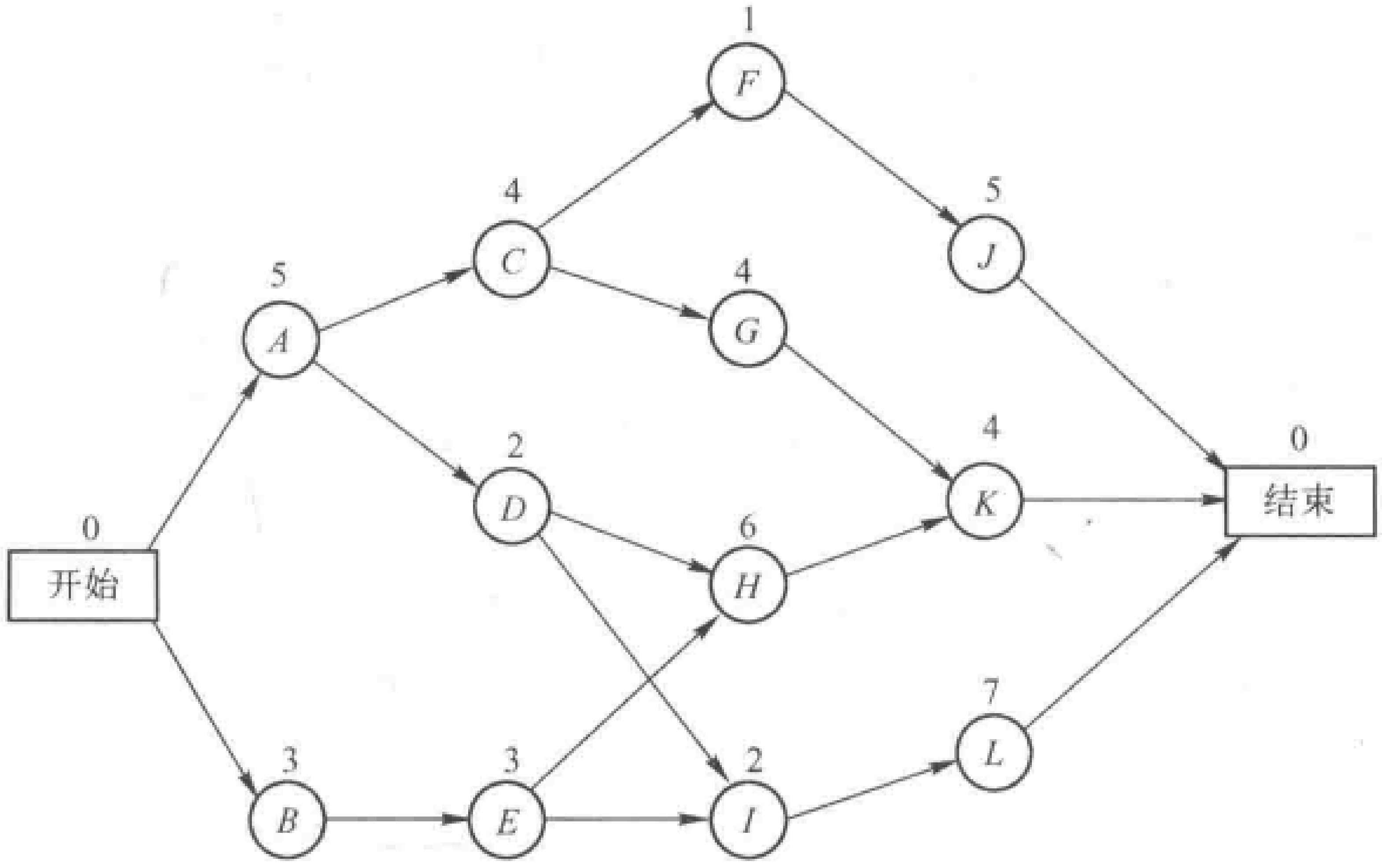
下表给出了每个区域分配到一定数量推销员时的销量增长值。

销售人数	地　区		
	1	2	3
1	40	24	32
2	54	47	46
3	78	63	70
4	99	78	84

利用动态规划解决该问题。通过绘制类似于习题 11.2-2 中的网络图进行求解，要求按照节点求解 $f_n^*(s_n)$，并在图中每个节点处进行标示，用箭头标示出最优路线，并指出最优路径及最优解。

运用动态规划方法，建立 $n=3$、$n=2$、$n=1$ 时的表格求解这个问题。

11.2-3　下面的工程网络图（如 10.8 节中所述），每点上的数字是相应活动所需的时间，找到这个网络从开始到结束的最长路径（全部时间最长）。



(a) 问题中动态规划模型的阶段和状态是什么？

(b) 利用动态规划的图解法解决该问题。要求按照节点求解 $f_n^*(s_n)$，并在图中每个节点处进行标示；标示出最优路线，并指出最优路径及最优解。如果存在不止一条最优路线，全部进行标出。

(c) 运用动态规划方法，建立 $n=4$、$n=3$、$n=2$、$n=1$ 时的表格求解这个问题。

11.2-4　考虑下列关于动态规划问题求解的陈述。判断每条陈述的对错，并从本章的内容中寻找证明（标出对错及证明的引用页码）。

(a) 求解的过程使用递推关系,已知阶段 n 的最优策略时,能够解出 $n+1$ 阶段的最优策略。

(b) 完成求解后发现某一阶段错误地做出非最优决策,重新求解时需要对该阶段之后的阶段重新确定最优决策。

(c) 整个问题的最优策略确定后,需要说明的是,特定阶段的最优策略是该阶段的状态和前一阶段的决策。

11.3-1 通过阅读参考文献,了解11.3节中插图部分故事的完整内容,简述故事中动态规划如何运用,然后列出动态规划的运用为其带来的财务和非财务方面的好处。

11.3-2* 一位水果店主有3个连锁水果店,他有5箱新鲜草莓需要在腐烂前售出。3家水果店潜在的销售能力不同,但店主可以接受有水果店分不到草莓的情况,同时,不希望在售出之前将整箱的草莓拆箱。因此,店主想知道如何分配这5箱草莓,才能够使最终的收益最大。

下表给出了不同的店分配到不同箱数草莓时的利润,运用动态规划的方法,将5箱草莓分配给3家店,使总利润最大。

箱 数	水 果 店		
	1	2	3
0	0	0	0
1	5	6	4
2	9	11	9
3	14	15	13
4	17	19	18
5	21	22	20

11.3-3 距离期末考试还有7天时间,某学生需要参加4门课的考试,她想要尽可能有效地分配学习时间。每门课至少需要花一天时间复习,且该学生喜欢每天仅复习一门课程,因此,每门课程可分配的时间为1天、2天、3天和4天。由于学习了运筹学课程,该学生决定运用动态规划的方法制定时间分配的策略,使4门课程总的分数提高最大。下表是该学生估计的每门课复习一定天数后的成绩提升数据,试用动态规划帮她解决这一问题。

复习天数	估计可提升的分数			
	课程			
	1	2	3	4
1	1	5	4	4
2	3	6	6	4
3	6	8	7	5
4	8	8	9	8

11.3-4 一项竞选活动进入了最后阶段,这是一场势均力敌的选举。其中一名候选人还有足够的资金购买5个黄金时段的电视时间,备选的电视时间包含4个不同地区电视台。基于选举信息,在电视播放地区赢得额外选票的数量与黄金时段的数量相关,下表给出了估算数据,其中数据单位为千张选票。运用动态规划的方法确定如何在4个地区电视台分配5个黄金时段,使最终增加的选票数最多。

时间段数量	地区			
	1	2	3	4
0	0	0	0	0
1	4	6	5	3
2	7	8	9	7
3	9	10	11	12
4	12	11	10	14
5	15	12	9	16

11.3-5　一个政党的地区主席正在为即将到来的总统竞选制定计划。为了选区工作，她准备将 6 名志愿者分派到 4 个选区，以获得最大的选票数的提升。为提高工作效率，每名志愿者只能去一个选区工作，并且允许出现不向某些选区分配志愿者的情况。下表给出了每个选区被分配一定数量志愿者后的选票增加数量。运用动态规划的方法，找出 6 名志愿者的分配策略，使选票数提升总数最大。

志　愿　者	选　　区			
	1	2	3	4
0	0	0	0	0
1	4	7	5	6
2	9	11	10	11
3	15	16	15	14
4	18	18	18	16
5	22	20	21	17
6	24	21	22	18

11.3-6　采用动态规划的方法求解 9.1 节中提出的北方航空公司生产计划问题（表 9.7），假设生产数量必须为 5 的倍数。

11.3-7*　一家公司即将向某竞争激烈的市场推出新产品，目前正在制定营销策略。目前的策略准备分为 3 个阶段推出产品，第 1 阶段为吸引首次购买的顾客，将会以非常低的价格推出某款介绍性产品；第 2 阶段将运用广告攻势，以使首次购买产品的客户能够按照正常价格购买产品；据了解，第 2 阶段结束时，另一家公司将推出一款竞争产品，所以第 3 阶段将在广告攻势的基础上增加促销活动，以确保客户不转向购买竞争产品。

本次市场营销的经费预算大概为 400 万美元，现在需要决定如何以最有效的方式，将这笔经费合理分配给 3 个阶段。用 m 表示第一阶段结束后的市场份额（用百分比表示），f_2 和 f_3 分别表示在 2、3 阶段结束后，剩余的市场份额比例。用动态规划的办法决定如何分配 400 万美元，使 3 个阶段的营销结束后，市场份额最大，即 $\max\ m f_2 f_3$。

(a) 假设这些经费必须以 100 万美元的整数倍花在每个阶段，其中第 1 阶段花费的最小允许倍数为 1，其他阶段为 0。下表给出了每个阶段花费对市场份额的影响，求解题干中所述的动态规划问题。

花费/百万美元	对市场份额的影响		
	m	f_2	f_3
0	—	0.2	0.3
1	20	0.4	0.5
2	30	0.5	0.6
3	40	0.6	0.7
4	50	—	—

(b) 假如预算内的经费可以以任何数量花在每一阶段,其中x_i表示第 i 阶段($i=1,2,3$)花费的金额(单位:百万美元),根据下面花费和影响的关系,求解题干中的动态规划问题。

$$m=10x_1-x_1^2$$
$$f_2=0.40+0.10x_2$$
$$f_3=0.60+0.07x_3$$

11.3-8 某电子系统由4个部件组成,系统的运行需要每个部分都运行。该系统的可靠性可以通过在部件中安装并行单元提升,下表给出了不同部件安装不同数量的并行单元后的运行概率。

单元数量	运行概率			
	部件1	部件2	部件3	部件4
1	0.5	0.6	0.7	0.5
2	0.6	0.7	0.8	0.7
3	0.8	0.8	0.9	0.9

系统运行的概率为各部件运行概率的乘积。下表给出了每一个部件分别安装一定数量的并行单元的成本(以100美元为单位)。

单元数量	成本			
	部件1	部件2	部件3	部件4
1	1	2	1	2
2	2	4	3	3
3	3	5	4	4

由于预算的限制,最多可以花费1000美元。使用动态规划的方法确定每个部件应该安装多少个并行单元,使系统的可靠运行概率最大。

11.3-9 考虑下列非线性整数规划问题。

max $Z=3x_1^2-x_1^3+5x_2^2-x_2^3$

s.t. $x_1+2x_2\leqslant 4$

且

$x_1\geqslant 0, x_2\geqslant 0$

x_1、x_2为整数。

使用动态规划求解这个问题。

11.3-10 考虑下列非线性整数规划问题。

max $Z=18x_1-x_1^2+20x_2+10x_3$

s. t. $2x_1+4x_2+3x_3\leqslant 11$

x_1、x_2、x_3为非负整数。

使用动态规划求解这个问题。

11.3-11* 考虑下列非线性规划问题。

$$\max \quad Z=36x_1+9x_1^2-6x_1^3+36x_2-3x_2^3$$

s. t. $x_1+2x_2\leqslant 3$

且

$x_1\geqslant 0, x_2\geqslant 0$

使用动态规划求解这个问题。

11.3-12 例 4 的问题中,假如从一个季节到下一个季节,改变雇员数量的全部成本变为 100 美元乘以雇佣数量差值的平方时,利用动态规划的方法重新求解该问题。

11.3-13 考虑下列非线性规划问题。

$$\max \quad Z=2x_1^2+2x_2+4x_3-x_3^2$$

s. t. $2x_1+x_2+x_3\leqslant 4$

且

$x_1\geqslant 0, x_2\geqslant 0, x_3\geqslant 0$

使用动态规划求解这个问题。

11.3-14 考虑下列非线性规划问题。

$$\min \quad Z=x_1^4+2x_2^2$$

s. t. $x_1^2+x_2^2\geqslant 2$

(没有非负的限制)使用动态规划求解这个问题。

11.3-15 考虑下列非线性规划问题。

$$\max \quad Z=x_1^3+4x_2^2+16x_3$$

s. t. $x_1x_2x_3=4$

且

$x_1\geqslant 1, x_2\geqslant 1, x_3\geqslant 1$

当 3 个变量都为整数时,使用动态规划求解这个问题。

无整数约束条件,使用动态规划求解这个问题。

11.3-16 考虑下列非线性整数规划问题。

$$\max \quad Z=x_1(1-x_2)x_3$$

s. t. $x_1-x_2+x_3\leqslant 1$

x_1、x_2、x_3非负。

使用动态规划求解这个问题。

11.4-1 一个西洋双陆棋棋手今晚有连续 3 场比赛。每次比赛,他都有机会进行他会赢得比赛的赌局。投注金额可以为他手中的钱的任意值,每场比赛,他赢得比赛的概率都是 1/2,赢得比赛后他可以赢得投注金额,输掉比赛后会输掉投注金额。他以 75 美元开始比赛,目标为最终拥有 100 美元(由于是友谊赛,他不想在结束比赛时超过 100 美元)。使用动态规划找出最优打赌策略,使他 3 场比赛之后正好赢得 100 美元的概率最大。

11.4-2 假设你有 5000 美元用于投资,接下来的 3 年中,每年选择 2 种投资方式中的一种(A 或 B)进行投资。两种投资的回报都不确定:对于 A 投资,一年结束后可能失去全部资金或者收回 10000 美元(利润为 5000 美元);对于 B 投资,一年结束后收回 5000 美元或 10000 美元。回

报概率如下表所列:

投　资	回收金额/美元	概　率
A	0	0.3
	10000	0.7
B	5000	0.9
	10000	0.1

每年只能够进行一次投资,每次只能投资 5000 美元(额外的资金不能够进行投资)。

(a) 使用动态规划寻找投资策略,使 3 年后拥有的资金最多。

(b) 使用动态规划寻找投资策略,使 3 年后拥有至少 10000 美元的概率最大。

11.4-3* 假设经过仔细分析,例 5 中生产每件成品的概率为$\frac{2}{3}$,而不是$\frac{1}{2}$,因此批量生产 L 件商品都为次品的概率为$\left(\frac{1}{3}\right)^L$,并且时间仅够两个生产周期。在这种情况下,运用动态规划的方法确定最优策略。

11.4-4 例 6 中,假如打赌改为:开始有两个筹码,5 局之后,她的筹码数小于 5 个。结合例 6 的计算结果,为统计学家确定新的最优策略。

11.4-5 Profit & Gambit 公司有一款主打产品,近期由于销量下降一直在亏本。本季度,销量将低于保本点 400 万个单位,每个单位将导致 5 美元的亏损(由于每个单位产品售出后边际收入比边际成本高 5 美元),总损失将达到 2000 万美元。管理层有两种方法扭转亏损局面:一种方法是放弃生产,停产将导致 2000 万美元的费用;另外一种方法是通过广告营销,增加产品的销售,当广告营销不成功时放弃产品(当成本达到 2000 万美元)。广告营销的初步计划已经形成,广告的时间将持续之后的 3 个季度(也可能在广告营销不成功的情况下提前取消),每个季度的费用为 3000 万美元,预计第一季度销量会增加 300 万个;第二季度再增加 200 万个;第三季度再增加 100 万个。进一步分析表明,选择任一发展方向,估计每个季度都会以 200 万个单位结束生产(假设每个季度增量为独立的随机变量,3 个季度额外增长分别处于 100 万~500 万、0~400 万、-100 万~300 万),假如实际增量太少,可停止广告活动,并在下两个季度中的任一季度都可以放弃该产品。

如果开始广告营销持续到其结束,估计后来一段时间里的销量将于第三季度保持同水平,因此,如果第三个季度的销量仍低于保本点,产品将被放弃。其后,假如销量高于第三季度保本点,则每个产品的利润为 40 美元。

运用动态规划确定最优策略,使期望的利润最高。