第 6 章　动态规划

当我们视线性规划是一种解决单一阶段单目标规划决策问题的定量分析方法时，则可认为动态规划(Dynamic Programming)是可以解决更复杂的多阶段单目标决策的定量分析方法。做决策时，有一种情况是将决策问题涉及的所有变量一次性处理，寻求最优决策，称为单一阶段决策问题。还有一种情况是将决策问题从时间上或空间上划分为前后几个相互联系的阶段(或部分)，按顺序对每一个阶段做出决策，并考虑每一阶段的决策将对以后各阶段的决策产生的影响，使整个活动过程所取得的效果最优。这样的决策问题称为多阶段决策问题。解决这类问题时显然不能运用线性规划及非线性规划的方法，但用动态规划方法则可以求解。所谓“ 动态 ”就是指某问题需逐个阶段处理(即时间或空间的转移)这一特点。

6.1　动态规划的基本原理

6.1.l 多阶段决策问题的解法

在企业管理的决策活动中，有一类重要问题，对这类问题进行决策将产生广泛、持续性影响。这类问题作决策时，有时需要从时间上划分阶段，按顺序对各个阶段作决策。例如产品生产计划的逐月(季、年)安排这个决策活动，确定了某个月的生产计划，则这个月的生产计划将直接影响到后面各月的计划安排；有时需要从空间上划分为若干部分。对各部分逐个作决策。例如在对一些工程项目进行投资分配的决策中，确定了其中任何一个项目去的投资额，都影响到其他所有项目投资分配额的确定。这两个典型的关于时间和空间的实际问题都具有“ 动态 ”的特性。我们进行决策的最终目标将是在一定的资源条件下，使所安排的逐个时期的生产计划或工程项目的投资分配计划，从整体上能达到最优的效果。上述的两个问题都属于多阶段决策问题。

解决这类具有“ 动态 ”特性的决策问题时，可以按照下面的基本思想进行工作。首先，把较为复杂的决策问题视为多阶段决策问题，按问题的时间或空间关系将问题分解为几个相互联系的阶段，从而使每一阶段上的决策问题都是一个较易求解的“ 子问题 ”，在实际决策时从初始状态开始按顺序逐段进行直到终止状态为止；然后，按“ 动态 ”的特点，有顺序地(逐次地)做出每一阶段上的最优决策，具体求解时通常按逆序进行，即从最后一个阶段开始到最初阶段为止。需要强调的是，做这种最优决策时必须同时考虑并保证这一阶段以前的各阶段上对“ 子问题 ”所做出的决策，从整体来说是最优决策。这样依次地做完每个阶段的最优决策后，它们便构成了整个问题的最优决策。这种方法可用于求解很多运筹学的问题，甚至对复杂的多变量的线性规划问题及非线性规划问

题，也可按动态规划的思想和步骤用动态规划的方法将该问题分解为每次只考虑一个变量的一系列子问题。

动态规划是美国数学家贝尔曼(R·Bellman)于 1957 年首先提出的。贝尔曼在其《动态规划》一书中指出了应用动态规划求优时的最优化原则就是："作为整个过程的最优策略具有这样的性质：即无论初始状态和决策如何，对前面的决策所形成的状态而言，余下的所有决策必须构成一个最优策略。"这个原则便作为我们建立动态规划数学模型的理论基础。下面举例来说明这个最优化原则。

某地拟铺设从地点 A 到地点 F 的地下水管道，有两条路线可供选择，其一是经地点 B_1 及 C_1 到 F；另一是经地点 B_2 及 C_2 到达 F，见图 6-1。

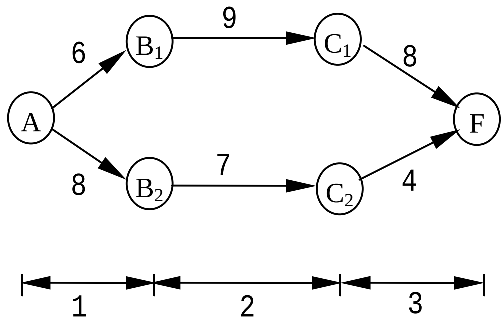

图 6-1

施工中每段路线所需时间如图所示。问题是寻找一条由 A 到 F 的施工时间最短的路线。这是一个多阶段决策问题。此问题可划分为三个阶段(图 6-1)。第一阶段的决策指的就是由 A 到 B_1 或由 A 到 B_2 这两条可能走的路线的选择。一旦选定了其中的某条路线，就直接影响到下一阶段的决策是由 B_1 到 C_1 或由 B_2 到 C_2。这就是说对每一阶段都应做出恰当的决策。由图 6-l 可知由 A 到 B_1 需 6 个时间单位，而由 A 到 B_2 需 8 个时间单位，前者所需时间比后者要少，对本阶段来说似乎应选择由 A 到 B_1 的路线作为第一阶段的最优决策。此阶段决策一经做出，则后面各段决策应接 B_1 到 C_1，C_1 到 F 的路线来施工，所需总时间为 23 个单位。但若选择由 A 到 B_2，再经 C_2 到 F 的路线则所需总时间为 19 个单位。显然，本应选择后一个路线才能使总时间数为最小，而第一阶段所选的由 A 到 B_1 的决策不能保证整个问题为最优决策。这就说明当某阶段决策选定时，它还直接影响到后面各阶段的决策和整个问题的决策。为了使整个问题的决策为优，按照贝尔曼的最优化原则意味着对某阶段作的决策对该阶段本身来说，从表面上看未必是一个最好的选择，甚至需要做出让步。如本例中第一阶段选择由 A 到 B_2 的路线对于本阶段来说表面上看不是最优决策，但是对于从第一阶段的决策由 A 到 B_2 得到的状态 B_2 而言，从 B_2 开始，余下的第二阶段决策由 B_2 到 C_2 及第三阶段的决策由 C_2 到 F 所构成的路线应该在时间上少于由 B_1 经 C_1 到 F 所需的时间。这就是贝尔曼所指的"……余下的所有决策必须构成一个最优策略。"这样再综合第一阶段由 A 到 B_2 的时间，从整体上来看反而为最优路线。相反，对前面阶段虽然选择了一个最优路线如例中的由 A 到 B_1 线路，但并不能保证由此产生的最后结论从整体上看为最优决策。这便是动态规划方法的理论基础。所以按动态规划方法求解时应按逆序逐段从终点向始点方向决策。对于本例，应该是先求第三阶段的最优决策为 $C_2 \to F$，再求第二阶段的最优决策为 $B_2 \to C_2$，最后求得第一阶段的最优决策为 $A \to B_2$，则整个问题的最优决策为 $A \to B_2 \to C_2 \to F$。

6.1.2 最短路线问题

最短路线问题就是说给定了初始点及终点以及由始点到终点的各种可能途径，要求寻找一条由始点到终点的最短路线。对于不同的实际问题也就是求距离（或时间、运输费用）的最小值。

例 6.1 图 6-2 是一个示意图，表示由始点 A 到终点 E 的网络路线。图中 A， B_1， B_2……表示各个地点，线段表示各个地点之间的交通路线，线段上的数字表示各个地点之间的距离（或时间、费用）。要求选择一条从 A 到 E 距离最短的路线（或时间、费用最少）。这是一个多阶段决策问题。

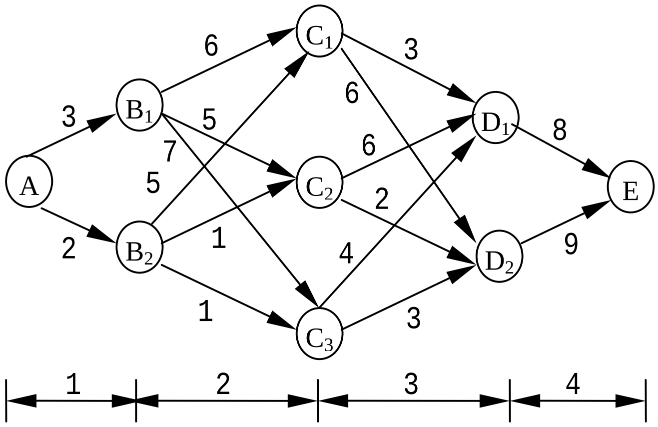

图 6-2

求解这个问题可采用穷举法。就是列出从 A 点到 E 点的所有路线，然后计算每条路线的距离(或时间、费用)，再进行比较，找出距离最短(或时间，费用最少)的路线，则最优解便找到了。此法称为枚举法或穷举法。本例从 A 点到 E 点共有 11 条不同路线，比较它们的长度，最短路线为：

$$A \to B_2 \to C_2 \to D_2 \to E$$

其距离为 14。但这样的计算相当冗繁，特别是网络路线比较复杂时，用穷举法计算甚至无法实现。下面采用动态规划方法求解此例。该法是求解这类最短路线(或最长路线)的最有效的方法之一。

用动态规划方法求解时，根据最优化原则导致最短路线问题应具有这种特性，就是：如果最短路线通过点 P，则这条最短路线从 P 点到终点的部分，对于从 P 点到终点的所有路线(称为 P 的后部路线)而言，必然也是最短的线路(P 的最短后部路线)。例如本例中最短路线通过 B_2， C_2， D_2， 则路线 $C_2 \to D_2 \to E$ 就是 C_2 到终点 E 所有路线中最短的一条。最短路线所具有的这个特性不难理解，用反证法，若从 P 点至终点还有另一条更短的路线存在，把这条更短的路线与原最短路线从起始点到 P 点的那部分路线连接起来，就会形成一条比原来由始点到终点的最短路线还要短的路线，这是不可能的。

现在我们利用这种特性把整个问题按空间关系，即把从 A 到 E 的网络路线划分为几个阶段，然后接逆序从最后一段开始向最初阶段递推，做出各阶段中各点的最优决策即寻找某阶段各点到终点 E 的最短路线上，在下一阶段应经过的点 O 则各点到终点 E 的最短(后部)路线可同时确定，最后便得到起点 A 到终点 E 的最短路线。因此，动态规划方法就是从终点逐段向始点方向寻找最短路线的方法。解题步骤如下：

1.把问题划分为几个阶段。本例可把网络按各点位置划分为四个阶段。第一阶段由 A 到 B_1 及 B_2 的路线构成，……，第四阶段由 D_1，D_2 到 E 的路线构成，如图 6-2 所示。

2.按阶段顺序首先考虑最后阶段即第四阶段的最优决策，也就是走哪条路线最短。先考虑 D_1，D_1 到终点 E 只有一条路线，故 D_1 的最短后都路线就是 $D_1 \to E$，最短距离为 8；同样，从 D_2 到 E 的最短后部路线就是 $D_2 \to E$，最短距离为 9。

3.按阶段顺序依次考虑第三、第二，第一阶段的最优决策，为此只需确定每一阶段上各初始点的最优决策即可。在第三阶段中分别考虑三个点 C_1，C_2 和 C_3 的最优决策。点 C_1 后部路线的第一段路线有两种选择．即 $C_1 \to D_1$ 和 $C_1 \to D_2$。为了确定 C_1 的最优决策即确定出对第四阶段的 D_1，D_2 的某个选择，应比较由 C_1 分别经过 D_1、D_2 到 E 所用的时间，即比较 $C_1 \to D_1$ 的长度与 D_1 的最短后部路线长度之和，以及 $C_1 \to D_2$ 的长度与 D_2 的最短后部路线的长度之和。前者为 3+8=11，后者为 6+9=15，取其中较小者为 C_1 的最短后部路线的长度。所以 C_1 的最优决策是下一阶段到达点 D_1，C_1 的最短后部路线是 $C_1 \to D_1 \to E$，最短距离为 11。用同样的方法求 C_2 的最优决策。由于 C_2 后部路线的第一阶段有两种选择，即 $C_2 \to D_1$，$C_2 \to D_2$。所以比较 $C_2 \to D_1$ 的长度与 D_1 的最短后部路线 $D_1 \to E$ 的长度之和，以及 $C_2 \to D_2$ 的长度与 D_2 的最短后部路线 $D_2 \to E$ 的长度之和，前者为 6+8=14，后者为 2+9=11，取其中较小者为 11，对应的 C_2 的最优决策是到达 D_2，C_2 的最短后部路线为 $C_2 \to D_2 \to E$，其距离为 11。对于 C_3 比较 4+8=12 及 3+9=12，因为二者相等，所以 C_3 的最优决策是到达 D_1 及 D_2 均可，C_3 的最短后部路线可以取其中任一条，$C_3 \to D_1 \to E$ 或 $C_3 \to D_2 \to E$ 距离都是 12。

在第二阶段中讨论点 B_1 及 B_2 的最优决策。点 B_1 的后部路线在第一段上有三条，即 $B_1 \to C_1$，$B_1 \to C_2$，$B_1 \to C_3$。而 C_1，C_2，C_3 各点到终点 E 的最优决策在第三阶段的讨论中已确定了。所以比较 6+11=17，5+11=16 及 7+12=19，知三条路线中长度最短者为 16，则 B_1 的最优决策就是由点 B_1 到点 C_2，点 B_1 的最短后部路线为 $B_1 \to C_2 \to D_2 \to E$，最短距离为 16。点 B_2 的后部路线第一阶段上有三条。比较 $B_2 \to C_1 \to D_1 \to E$，$B_2 \to C_2 \to D_2 \to E$ 及 $B_2 \to C_3 \to D_1 \to E$（或 $B_2 \to C_3 \to D_2 \to E$）这三条后部路线的长短，由 5+11=16，1+11=12，1+12=13 得点 B_2 的最优决策是到点 C_2，点 B_2 的最短后部路线为 $B_2 \to C_2 \to D_2 \to E$，其距离为 12。

最后考虑第一阶段，也就是求始点 A 的最优决策。点可能的决策方案有两个，即到达 B_1 或 B_2，而 B_1 及 B_2 两点到终点 E 的最优决策已在第二阶段确定了，并找到了该两点之间的最短后部路线，因此我们只需在 $A \to B_1 \to C_2 \to D_2 \to E$ 和 $A \to B_2 \to C_2 \to D_2 \to E$ 这两条路线中选择点 B_2，点 A 的最短后部路线为 $A \to B_2 \to C_2 \to E$。至此解题过程结束，整个网络最短路线的决策问题得到解决。从 A 到 E 的最短路线为 $A \to B_2 \to C_2 \to D_2 \to E$，最短距离为 14。

通过此例的求解过程，可以看到用动态规划方法逐段求解时，每个阶段上的求优方

法基本相同，而且比较简单，所以在后面进一步学习这种方法时，虽然会遇到较多的数学符号，但对这些符号还是易于理解的。另外，动态规划方法明显优于枚举法（穷举法），在本例中，每一阶段的计算都要利用上一阶段的计算结果，因而减少了很多计算量。阶段数愈多，这种效果愈明显。

6.2　动态规划的数学模型

6.2.1 动态规划的基本概念

为讨论问题方便起见，先介绍动态规划的基本概念和符号。

1．阶段

用动态规划方法解题，原问题必须能划分为若干阶段。阶段是按问题的时间或空间特征来划分的。通常用 k 表示阶段的变量；阶段的编号在本书中按从左到右的顺序编号。例如前例（最短路线问题）中将整个问题按空间划分为四个阶段，当 $k=1,2,3,4$ 时，分别表示问题处于第一、二、三、四阶段。第一阶段（$k=1$）包括两条支路 $A\to B_1$ 和 $A\to B_2$，第二阶段（$k=2$）包括 6 条支路 $B_1\to C_1$、$B_1\to C_2$、$B_1\to C_3$、$B_2\to C_1$、$B_2\to C_2$、$B_2\to C_3$。依次类推。用动态规划方法求解时，通常按相反方向逐段决策，即从最后一个阶段开始，按逆序到最初阶段为止。

2．状态和状态变量

状态可理解为事物的某种特征。状态的改变意味着事物发生变化。在动态规划问题中，状态是划分阶段的依据，状态的变化就意味着阶段的推移，因此解题时首先应明确每一阶段开始时的一切可能状态。在最短路线问题中，用状态表示某阶段的出发位置，它既是该阶段的初始点也是前一阶段的终止点，通常一个阶段包含若干个状态。如前例的最短路线问题中，第一阶段只有一个状态时点 A；第二阶段有两个状态点 B_1 和 B_2，等等。为了将实际问题转化为数学模型之后便于计算，通常可将状态用数量表示，则状态可视为变量，并称为状态变量 S，记 S_k 为第 k 阶段（有时记 S_i 为第 i 阶段）的状态变量，用 $S_k=\{S_k\}$ 表示 k 阶段上所有状态的集合。例如前例中，当 $k=2$ 时，有两个状态 B_1 和 B_2，则第二阶段状态为 $S_2=B_1$ 和 $S_2=B_2$，状态集合 $\{S_2\}=\{B_1,B_2\}$（或记为 $S_2=B_1,B_2$），若用数量表示，令点 $B_1=1$，点 $B_2=2$（每阶段状态都可记为 $1,2,3,\cdots\cdots n$），则 $\{S_2\}=\{1,2\}$（$S_2=1,2$）。对于有些实际问题状态本身就是数量。例如投资分配问题，第 k 阶段的状态即该段最初的投资数额，仍可将实际数额转换数 i 并记为 $S_k=i\left(i=1,2,\cdots n\right)$。

3．决策和决策变量

对实际问题进行多阶段决策时是从第一阶段的初始状态开始，逐段向后发展，直到最后一个阶段为止。在每一阶段中只达到一个可能的状态，从该状态演变到下一阶段进入哪个状态，取决于这一阶段做出的决策。这就是说，决策就是各阶段对状态演变各种

可能性的选择描述决策的变量，称为决策变量，用 $x_k(S_k)$ 表示第 k 阶段处于 S_k 时的决策变量。用决策集合 $D_k=\{x_k(s_k)\}$ 表示 $x_k(S_k)$ 的取值范围。例如图 6-2 示意的最短路线问题中，当 $k=2$ 时，状态集合 $\{S_2\}=\{B_1,B_2\}$，则状态 B_1 的决策变量 $x_2(B_1)$ 的取值 $x_2(B_1)=C_1$， $x_2(B_1)=C_2$，或 $x_2(B_1)=C_3$，则为 $\{x_2(B_1)\}=\{C_1,C_2,C_3\}$，表示从状态 B_1 出发到终点 E 时，进入到下一阶段所选择的状态，可以是 C_1,C_2 或 C_3，换句话说 $\{x_2(B_1)\}$ 表示从 B_1 点出发到达下一阶段的路线可有三条，所以在第二阶段对于 B_1 点的决策方案有三个，即有 $B_1\to C_1$， $B_1\to C_2$ 或 $B_1\to C_3$。状态 B_1 的最优决策指的是为达到最优目标，对 C_1,C_2,C_3 的某个选择，使得由 A 到 E 的最短路线经过该点。对状态 B_2 的决策变量 $\{x_2(B_2)\}$ 的取值也有三个，就是 $\{x_2(B_2)\}=\{C_1,C_2,C_3\}$。由此可见，决策变量 $x_k(S_k)$ 是状态变量的函数，为了计算的方便，既然状态变量 S_k 的取值可人为地规定为数，我们也可将决策变量 $x_k(S_k)$ 的取值人为地规定为数。例如上面的 $\{x_2(B_2)\}=\{C_1,C_2,C_3\}$，若令 $C_i=i,(i=1,2,3)$，则 $\{x_2(B_2)\}=\{1,2,3\}$。

4．策略

多阶段决策问题中，把每个阶段做出的决策 $x_k(s_k)(k=1,2,\cdots,n)$ 组成的序列称作是策略。把解决问题时产生的效果最优的策略称为最优策略。因此，求解多阶段决策问题便是去寻找该问题的最优策略，而动态规划方法就是用来求得这个最优策略的方法。

设整个问题分化为 n 个阶段 $(n=1,2,\cdots,k,\cdots,n)$，第 k 阶段状态 S_k 开始至最终阶段到 n 的终点为止的过程称为第 k 阶段状态 S_k 的后部过程，相应的第 k 阶段以后的各阶段的决策所构成的决策序列称为状态 S_k 的后部策略(前面提到的点 P 的后部路线便是由状态 P 的后部策略产生的)。从前面的例题可知，动态规划方法的每一个阶段都需找出该阶段每一状态的最优后部策略，即最短后部路线。

5．阶段损益函数(或称阶段效应)

用动态规划方法求解问题时，把用以衡量策略优劣的数量指标称为损益函数，其取值称为损益值。阶段损益函数是某一阶段、某一状态本身在决策选择时所能获得的损失(或效益)值的表达式，通常表示为 $d_k(s_k,x_k)$。 S_k 表示第 k 阶段所处的状态， x_k 表示状态为 S_k 时的决策，也就是该阶段所选择的决策方案。对于不同的决策问题损益函数的损益值有不同的实际含义。例如，效益值可能是效率、产值、利润，而损失值可能是成本、费用等。又如例 2 的最短路线问题中，可视阶段损益值为网络路线中第 k 阶段上某路线的距离，如第三阶段点 C_1 的损益值可记为 $d_3(C_1,x_3(C_1))$，具体就是 $d_3(C_1,D_1)=3$ 以及 $d_3(C_1,D_2)=6$，表示第三阶段由状态 C_1 到 D_1 和到 D_2 的距离分为 3 及 6。

6．最优损益函数

最优损益函数是用以表示某阶段、某状态的最优后部策略的数量指标。它综合考虑了这一阶段本身的阶段损益值以及相应的下一阶段到终点为止的最优损益值两个方面的情况。记最优损益函数为 $f_k(s_k)$ $(k=1,2,\cdots\cdots,n)$ k 表示所处的阶段， S_k 表示第 k 阶段

中所处的状态。例如，在例 2 中，$f_2(B_1)=5+11=16$，表示在第二阶段时，状态 B_1 到终点 E 的最优损益函数值(即点 B_1 到终点 E 的最短后部路线长度)为 16,同理，$f_2(B_1)=12$。

有了上述的基本概念，我们便可将实际问题用下面介绍的动态规划模型来表示。

6.2.2 动态规划的数学模型建立方法

现在我们结合网络最短路线来介绍动态规划方法中所使用的数学模型将如何建立。

例 6.2 建立例 1 的最短路线问题的数学模型，见示意图 6-2。

动态规划的数学模型包括目标函数和约束条件两部分。对于该最短路线问题求的是从始点 A 到终点 E 的距离为最短。又因为这是多阶段决策问题，所以相应的动态规划方法应将问题首先划分为若干个阶段。然后再按上面所述的基本思想，从终点 E 逐段向始点方向寻找每阶段的最优决策。这是一个逆向递推的过程，用动态规划的术语表示出对目标函数的要求就是求得每个阶段的最优损益函数(最短后部路线的长度)。下面的工作就是具体建立这种按阶段表达的数学模型。

动态规划的约束条件可由实际问题决定。例如例 1 中最短路线问题，在网络路线图上用带有箭头的线段表示可行的通路，箭头的方向表示行进的方向。这种方向通常都是单向的，不能反向行进。

具体表达这种具有“递推”特性的数学模型如下：首先将问题划分为四个阶段，分别记为 $k=i(i=1,2,3,4)$，其次，按逆向建模。

$k=4$ 时目标函数的优值为

$$f_4(s_4)=\min\{d_4(s_4,x_4)\} \tag{6-1}$$

其中 $f_4(s_4)$——例 1 中网络路线上第四阶段中状态变量 S_4 取为 D_1、D_2 两点时，由此两点分别到达终点 E 的最短距离。

$d_4(s_4,x_4)$——第四阶段本身的损益函数。实例中即从 D_1、D_2 两点到终点 E 的实际距离。它取决于该段中状态变量 S_4（取 D_1，D_2）和相应的决策变量 x_4（S_4），在本例中 x_4（D_1）及 $x_4(D_2)$ 均为终点 E。

$\min\{d_4(s_4,x_4)\}$——由实际问题决定。因为各状态终点的可行路线可能不止一条，于是就取该阶段中各状态到终点 E 的路线中最短者。所以本例中取 min。如果需要目标函数的取值以大者为优（求最长路线问题），则应改 min 为 max。

具体来看，当 $k=4$ 时，因状态变量可取两个点：$S_4=D_1$，$S_4=D_2$，又因 D_1，D_2 两点到终点 E 的路线分别只有唯一的一条路线 $D_1\to E$ 及 $D_2\to E$，所以 $f_4(D_1)$ 即 $f_4(S_4=D_1)=\min\{d_4(D_1,E)\}$，$f_4(D_2)$ 即 $f_4(S_4=D_2)=\min\{d_4(D_2,E)\}$
分别是 D_1，D_2 到终点 E 的最短距离。

$k=3$ 时，有动态规划的最优决策原则，目标函数优值 $f_3(s_3)$ 应综合考虑第三阶段本身和第四阶段两者结合为整体的最优。所以目标函数优值（反应了第三阶段状态 S_3 的最优后部策略）也就是在前面基本概念中提出的最优损益函数。

所以令 $x_3(s_3)=s_4$ 时

$$f_3(s_3)=\min\{d_3(s_3,x_3)+f_4(s_4)\} \tag{6-2}$$

$f_3(s_3)$——第三阶段目标函数的最优值

$d_3(s_3,x_3)$——第三阶段的阶段损益函数

在本例中即点 C_1，C_2，C_3 分别到 D_1 和 D_2 两点的实际距离。

因此，$k=3$ 时，状态变量 S_3 取值为三点即 C_1，C_2，C_3，相应的目标函数优值有三个：

$f_3(c_1)$ 即 $f_3(s_3=C_1)=\min\{d_3(C_1,D_1)+f_4(D_1),d_3(C_1,D_2)+f_4(D_2)\}$，

$f_3(c_2)$ 即 $f_3(s_3=C_2)=\min\{d_3(C_2,D_1)+f_4(D_1),d_3(C_2,D_2)+f_4(D_2)\}$，

$f_3(c_3)$ 即 $f_3(s_3=C_3)=\min\{d_3(C_3,D_1)+f_4(D_1),d_3(C_3,D_2)+f_4(D_2)\}$，

$k=2$ 时，最优损益函数 $f_2(s_2)$ 应综合考虑该阶段损益 $D_2(s_2,x_2)$ 及其在第三阶段受到影响的状态设为 $s_3=x_2(s_2)$ 到终点 E 的最短距离 $f_3(s_3)$ 整体为优。

所以，第二阶段的递推关系式为

$$f_2(s_2)=\min\{d_2(s_2,x_2)+f_3(s_3)\} \tag{6-3}$$

具体看，$k=2$ 时，状态变量 s_2 取值为两点即 B_1 及 B_2，相应的目标函数优值有两个：

$f_2(B_1)$ 即

$$f_2(s_2=B_1)=\min\{d_2(B_1,C_1)+f_3(C_1),d_2(B_1,C_2)+f_3(C_2),d_2(B_1,C_3)+f_3(C_3)\}$$

$$f_2(B_2)=\min\{d_2(B_2,C_1)+f_3(C_1),d_2(B_2,C_2)+f_3(C_2),d_2(B_2,C_3)+f_3(C_3)\}$$

$k=1$ 时，由始点 A 到终点 E 要经点 B_1 或 B_2，选择经哪个点才能保证 A 经四个阶段后到达 E 所走的距离最短？这就应综合考虑第一阶段的阶段效益 $d_1(s_1,x_1)$ 以及由 $x_1(s_1)$ 影响到的第二阶段的状态（设为 s_2）的最优损益函数 $f_2(s_2)$，才能得到点 A 的最优后部策略的数量表示 $f_1(A)$，也就是始点 A 到终点 E 的最优损益值。

所以，$k=1$ 时，目标函数最优值的一般表达式为：

$$f_1(s_1)=\min\{d_1(s_1,x_1)+f_2(s_2)\} \tag{6-4}$$

由于第一阶段上状态 S_1 只取终点 A，即 $S_1=A$，因而只能列出一个最优损益值

$$f_1(A) \text{ 即 } f_1(s_1=A)=\min\{d_1(A,B_1)+f_2(B_1),d_1(A,B_2)+f(B_2)\}$$

通过上面的讨论，可知求 $f_1(A)$ 的全过程是倒过来（由 $k=4$ 开始向前计算），逐段递推的过程。换句话说，想求得第一阶段的 $f_1(s_1)$，必须先计算第二阶段的 $f_2(s_2)$，而想求 $f_2(s_2)$，应先求第三阶段的 $f_3(s_3)$，而 $f_3(s_3)$ 的求解必须先求得第四阶段的 $f_4(s_4)$。为此我们可以把这种逆推方法推广到阶段数为任意正整数 k 的问题（$k=1,2,\cdots\cdots,n$）中

去。得到第 k 阶段的动态规划数学模型中目标函数优值的表达式为：

$$f_k(s_k)=\min\{d_k(s_k,x_k)+f_{k+1}(s_{k+1})\} \tag{6-5}$$

$$k=1,2,\cdots\cdots,n$$

公式（6-5）中令 $x_k(s_k)=s_{k+1}$，特别是（6-5）中当 $k=n$ 时，(最后一个阶段)，因为 $f_{n+1}(s_{n+1})=0$，所以规定：

$$f_n(s_n)=\min\{d_n(s_n,x_n)\} \tag{6-6}$$

上面这个利用了第 k 阶段与 $k+1$ 阶段之间递推关系的数学模型 (6-5) 称为动态规划的递推公式或函数基本方程。

6.3 动态规划的求解方法

6.3.1 动态规划递推公式的迭代解法

这个解法就是利用递推公式（6-5）直接求解，举例说明如下：

例 6.3 试用公式（6-5）求解§6.1 例 1 种网络最短路线问题（图 6-2)。

解：先按建立模型的要求，确定阶段、状态、决策。将问题划分为四个阶段，状态变量即为各阶段的始点。第一阶段的状态为 A；第二阶段的状态为 B_1，B_2；第三阶段的状态为 C_1，C_2；第四阶段的状态为 D_1，D_2。则所有状态均具有无后效性且能列举出来。每阶段的决策允许集合与§6.2 中建立公式（6-5）时所述一致，在此从略。反映第 k 阶段与第 $k+1$ 阶段的状态转移方程 $s_{k+1}=g_k(s_k,x_k)$ 可由示意图中点的推移过程反映出这种函数关系，不一定非要表达成数学形式。最后，用于计算的递推公式便是 (6-5)。迭代过程如下：

$k=4$ 时,考虑点 D_1，D_2 分别到终点 E 的最短距离 $f_4(s_4)$，此时各有一条路 $D_1\to E$ 及 $D_2\to E$，又因为 $k=4$ 为最后一个阶段，所以由公式（6-6）

$$f_4(D_1)=\min\{d_4(D_1,E)\}=\min\{8\}=8$$

$$f_4(D_2)=\min\{d_4(D_2,E)\}=\min\{9\}=9$$

则 D_1 到 E 的最短路线是 8，路线为 $D_1\to E$，对 D_2，也有 $x_4(D_2)=E$。

$k=3$ 时，考虑从点 C_1，C_2，C_3 到终点 E 的最短距离 $f_3(s_3)$，对点 C_1 到 E 可经过 D_1 或 D_2 两点，由公式（6-5）

$$f_3(C_1)=\min\{d_3(C_1,D_1)+f_4(D_1),d_3(C_1,D_2)+f_4(D_2)\}=\min\{3+8,6+9\}=11$$

即点 C_1 到 E 的最短距离为 11，最短路线为 $C_1\to D_1\to E$，C_1 的决策为 $x_3(C_1)=D_1$。

同理，对点 C_2 有

$$f_3(C_2)=\min\{d_3(C_2,D_1)+f_4(D_1),d_3(C_2,D_2)+f_4(D_2)\}=\min\{6+8,2+9\}=11$$

其最短距离是 11，其路线是$C_2 \to D_2 \to E$，

决策为 $x_3(C_2)=D_1$

对点C_3有

$$f_3(C_3)=\min\{d_3(C_3,D_1)+f_4(D_1),d_3(C_3,D_2)+f_4(D_2)\}=\min\{4+8,3+9\}=12$$

点C_3到E的最短距离是 12，其路线可有两条∶$C_3 \to D_1 \to E$，$C_3 \to D_2 \to E$，决策为$x_3(C_3)=D_1$或$x_3(C_3)=D_2$均可。

$k=2$时，考虑两点B_1，B_2到终点的最短距离$f_2(s_2)$，对点B_1有

$$\begin{aligned}f_2(B_1)&=\min\{d_2(B_1,C_1)+f_3(C_1),d_2(B_1,C_2)+f_3(C_2),d_3(B_1,C_3)+f_3(C_3)\}\\&=\min\{6+11,5+11,7+12\}=16\end{aligned}$$

则最短路线为$B_1 \to C_2 \to D_1 \to E$，且$x_2(B_1)=C_2$

从点B_2出发，有

$$\begin{aligned}f_2(B_2)&=\min\{d_2(B_2,C_1)+f_3(C_1),d_2(B_2,C_2)+f_3(C_2),d_3(B_2,C_3)+f_3(C_3)\}\\&=\min\{5+11,1+11,1+12\}=12\end{aligned}$$

则最短路线为$B_2 \to C_2 \to D_1 \to E$，$x_2(B_2)=C_2$

$k=1$时，考虑始点E的最短距离$f_1(s_1)$，有

$$f_1(A)=\min\{d_1(A,B_1)+f_2(B_1),d_1(A,B_2)+f_2(B_2)\}=\min\{3+16,2+12\}=14$$

所以，始点A决策为$x_1(A)=B_2$，最短距离为 14，其路线为$A \to B_2 \to C_2 \to D_2 \to E$。整个迭代过程至此结束。

最后，可按计算顺序反推，得最优的决策序列$x_1(A)=B_2$，$x_2(B_2)=C_2$，$x_3(C_2)=D_1$，$x_4(C_1)=E$构成了本问题的最优策略。

用数学模型（6-5）直接求解，由于数学公式中符号抽象，所以容易出错。下面再介绍两种较为形象、直观和简捷的计算方法，成为标号法和表格求解法。这两种方法都是根据动态规划的求优原则由递推公式（6-5）演化而来。

6.3.2 动态规划的网络标号法

网络标号法就是直接在网络图上进行计算的方法，现在仍通过求解例 1 种网络最短路线来阐述。

例 6.4　设有一个网络线路(图 6-2)，试用标号法求出由A到E的最短路线。

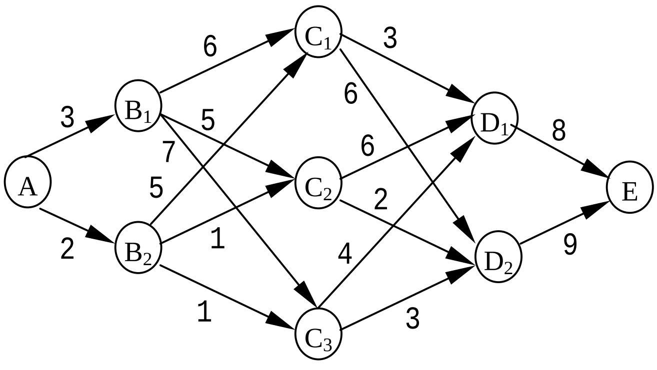

图 6-5

解：首先仍划分为四个阶段，见图 6-5。然后逐段在每阶段的状态点上方标出方格，见图 6-6。求解时仍从第四阶段起逐段在每个状态点上方的方格内填数，以数表示该点到终点 E 的最短距离，直线连接的点表示该点到终 E 点的最短路线，图中的粗线表示由始点 A 到终点 E 的最短路线。

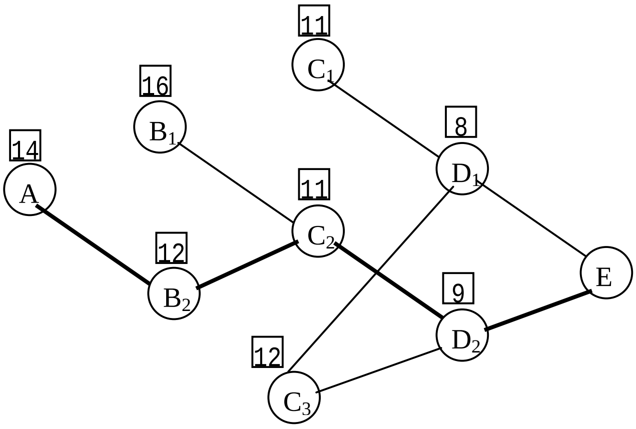

图 6-6

解题与步骤：

先标出第四阶段各状态点 D_1，D_2 分别到终点 E 的最短距离（这是例 2 中计算 $f_4(s_4)$ 的简明表示法）。它们分别是 8 和 9，将此最短距离填入 D_1， D_2 二点上方的方格内，并画出与最短路线相对应的连线。然后在第三阶段的各状态上方标号（在方格内填入最短距离）。需要标号的点到下一阶段已标号的各状态点的各段距离分别加上已标号点方格中的数字而取其中较小者（此过程也就是例 2 中 $f_3(s_3)$ 按地推公式（6-5）进行计算的见解算法）， 就是 $k=3$ 时状态 s_3 到终点 E 的最短距离。将此距离的数字填入点 s_3 上方的方格内（标点），并画出连线表示该点的最短路线在第三阶段上通过的部分。$k=3$ 时，需标号的点有 C_1， C_2， C_3，对于点 C_1 可经过标号点 D_1， D_2 到达 E。其中 C_1 到 D_1 的距离为 3，标号点 D_1 的方格中数字为 8 两者相加为 3+8=11；C_1 到 D_2 的距离为 6，D_2 点的标号数为 9，两者相加是 6+9=15，取 11 及 15 两者中的最小者填入 C_1 上方空格内即得到点 C_1 到终点 E 的最短距离。同时在 C_1，D_1 两点间连线，舍去 C_1 到 D_2 间的支路线。被舍去的路线在第二阶段的计算时就不起作用了。因为由点 C_1 到终点 E 的最短距离为 11，其路线为 $C_1 \to D_1 \to E$ 已确定了。同理，对 C_2 进行标号时，比较 6+8=14 及 2+9=11，取最小者 11 做 C_2 的标号数，并连接 C_2-D_2，舍去 C_2-D_1。对 C_3 进行标号时，比较 4+8=12 及 3+9=12，其最小者认为 12，所以 C_3 的标号数为 12，连接 C_3-D_1 及 C_3-D_2，即点 C_3 到终点 E 的最短距离为 12，最短路线为 $C_3 \to D_1 \to E$ 或 $C_3 \to D_2 \to E$。按照第四阶段和第三阶段的标号方法，逐段逆推，直到算出始点 A 的标号数为止，整个工作就完成了。

第二阶段需要对 B_1， B_2 标号。从点 B_1 到终点 E 有四条路线：$B_1 \to C_1 \to D_1 \to E$，$B_1 \to C_2 \to D_1 \to E$ 和 $B_1 \to C_3 \to D_1 \to E$ 或（$B_1 \to C_3 \to D_2 \to E$）。而 C_1，C_2，C_3 三点在第三阶段的计算中已标过号了，它们到终点 E 的距离分别为 11，11 和 12，所以只要根据 B_1 分别到 C_1，C_2，C_3 三点的距离以及 C_1，C_2，C_3 三个点的标号数，比较 6+11=17，5+11=16，和 7+12=19，取其中最小者 16 作为 B_1 的标号数字填入方格，并连接 B_1-C_2，舍去 B_1-C_1 及 B_1-C_3。同理，B_2 点的标号数取 5+11=16，1+11=12 和 1+12=13 中最小者 12 填入方格，连接 B_1-C_1，舍去 B_2-C_1 及 B_2-C_3。

最后是第一阶段，求始点 A 的标号数。取 3+13=16，2+12=14 种最小者 14 填入 A 的方格内，连接 $A-B_2$，舍去 $A-B_1$，标号工作到此结束，由 A 到终点 E 的最短距离为 14，相应的最短路线为 $A\to B_2\to C_2\to D_2\to E$。

上述方法称标号法，是从终点 E 向始点 A 方向标号的。如果规定从始点 A 到终点 E 为顺行方向，从 E 点出发到 A 点为逆行方向，则称这种标号程序为逆序解法，相应于§6.2 中的递推公式(6-5)。如果以 E 为始点，A 为终点，由第一阶段开始从前面向后面逐段标号，每点的标号数字表示该点到 A 点的最短距离，最后得到 E 的标号数，它表示由 E 到 A 的最短距离，其标号数也应等于 14。称这种由点 A 到点 E 的标号程序为顺行解法，相应于§6.2 中的公式（6-11)。具体标号如图 6-7 所示。

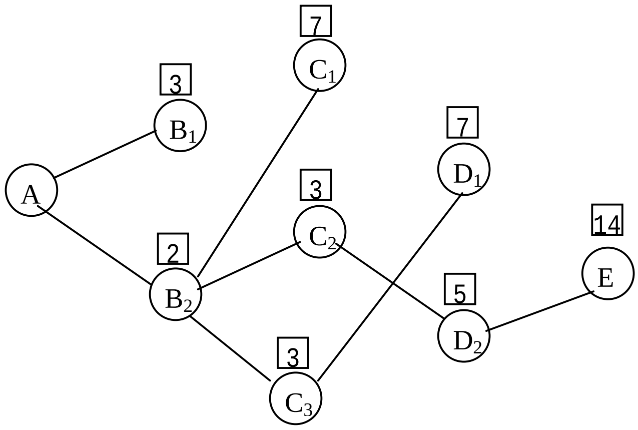

图 6-7

在逆序解法（或顺序解法）中，不仅能得到由始点 A 到终点 E（或由始点 E 到终点 A）的最短路线及相应的距离，而且得到了从所有各中途点到终点的最短路线及相应的距离。这就是说，所求出的不只是一个最优策略，而是一组最优策略。这对许多实际问题是很有用的，有利于帮助分析所得结果。

标号法的优点是形象具体，思路直观，特别是在初学时容易理解，但它只适合求解已有网路的问题。有一些实际管理问题（如设备更新问题）可以转化为网络最短路线来求解。关于这方面的问题本章不作进一步讨论。

6.3.3 动态规划的离散化算法

动态规划是解决多阶段决策的有效方法，但是对于某些具有“静态”特征的单阶段决策问题，例如某些具有特殊结构的线性规划、非线性规划及整数规划问题，也可以按动态规划的思路，将问题先划分为若干个阶段，然后用动态规划的递推公式（6-5）或（6-6）去求最优解。这时，只要将阶段损益 $d_k(s_k,x_k)$ 和最优损益 $f_{k+1}(s_{k+1})$ 用函数的解析式表示即可。对于线性规划、非线性规划及整数规划中的决策变量的取值及状态变量的取值，则有取连续值及离散值之分（前面所述的最短路线问题中 s_k,x_k 总是取可数的有限个离散点（值)），因此应针对具体问题的要求，将 s_k 及 x_k 取值连续化或离散化。此工作可通过建立状态转移方程 $s_{k+1}=g_k(s_k,x_k)$ 或 $s_k=g_k(s_{k+1},x_k)$ 来进行，然后再代入递推公式（6-5)（或（6-11)）求解。这时便产生了动态规划的离散化算法和与之相对应的“连续化”算法，下面分别举例来说明其解法。

例 6.6 用动态规划方法求解变量连续取值问题。

$$\max Z = x_1^2 + x_2^2 - x_3^2$$
$$\begin{cases} x_1 + x_2 + x_3 = 4 \\ x_i \ge 0 \qquad i = (1,2,3) \end{cases}$$

解：假设原线性规划为解决有限资源的利用问题，以使完成的任务 Z 为最多。其中 x_1, x_2, x_3 是三个连续生产阶段的生产数量，各阶段生产一个单位产品所需时间由约束知都是一个时间单位，而边界约束条件 4 表示整个生产过程的时间限制，目标函数 Z 可视为三个阶段所得损益的数量之和。

这是只有一个约束等式的非线性规划问题。非线性规划问题的求解随着决策变量 x_i 的增加将很麻烦，但由于这个问题的特殊结构，可以将它看作是一个多阶段决策问题，用递推公式（6-5）求解。首先把问题化作三个阶段。把原问题的变量当作决策变量，则每一阶段仅对应一个决策变量 $x_k = (k = 1,2,3)$， x_k 可取相应区间上的连续值。状态变量 s_k 可由约束条件来定，仍取连续值，阶段损益 $d_k(s_k, x_k)$ 由目标函数可知。整个过程的流程图如图 6-8 及图 6-9 所示。

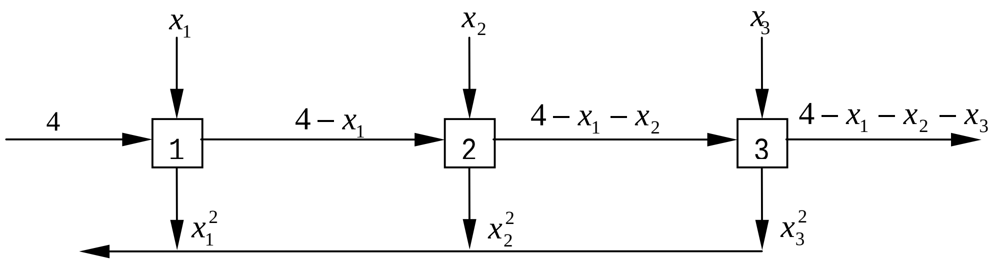

图 6-8

根据流程图 6-8 可说明 s_k， x_k， $d_k(s_k)$ 及 $f_k(s_k)$ 的含义。图 6-8 中先输入了整个生产过程的限制时间为 4，如果生产的数量为 x_1，则留下的时间量为 $4 - x_1$，又进入第二阶段进行生产。在第一阶段的损益为 x_1^2，第二、第三阶段的含义同理。在第三阶段生产完毕剩余的时间量为 $4 - x_1 - x_2 - x_3$，按约束条件的限制，则此剩余时间数量为 0。

可将上面的说明转化为动态规划的表述形式。

状态变量 s_k ：表示可以提供给各阶段的时间，则 s_k 的取值为连续值。

第 1 阶段的状态（即第一阶段的初始状态）为 $s_1 = 4$。

第 2 阶段的状态（即第一阶段的终止状态） $s_1 = 4 - x_1, s_3, s_4$ 均可定义，其中 s_4 表示第 4 阶段（设想的）状态。

决策变量 x_k ：即每阶段生产产品的数量即 $x_i(i = k = 1,2,3)$

状态方程$s_{k+1}=g_k(s_k,x_k)$

因 $\begin{cases} s_1=4 \\ s_2=s_1-x_1 \\ s_3=s_2-x_2 \\ s_4=s_3-x_3 \end{cases}$ 故$s_{k+1}=s_k-x_k, k=1,2,3$ 显然有$x_k\le s_k$

这里的决策变量x_k取值为连续。例如x_1的取值范围由约束条件知为$0\le x_1\le 4$。

边界条件：由于开始时可用时间为 4，而经过三个阶段的生产应将全部时间用完，则$s_1=4$， $s_4=0$。

阶段损益函数$d_k(s_k,x_k): d_1=x_1^2, d_2=x_2^2, d_3=-x_3^2$

最优指标函数： $f_k(s_k)=\max\limits_{sk}\{d_k(s_k,x_k)+f_{k+1}(s_{k+1})\} k=1,2,3$

即为本问题用动态规划求解时使用的递推公式。动态规划的流程图如图 6-9 所示。

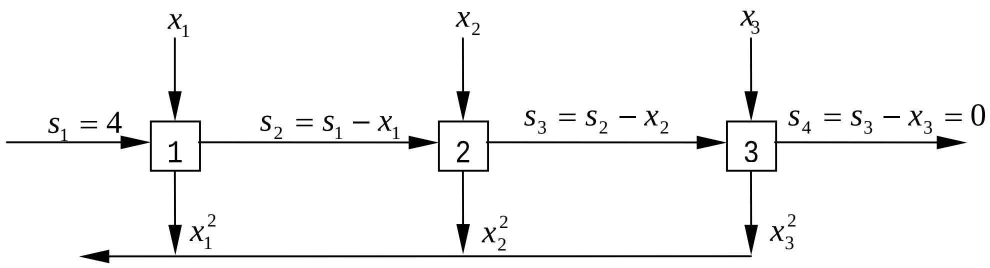

图 6-9

下面用逆序解法求解此例。

$k=3$时，因$s_4=s_3-x_3=0$，故$s_3=x_3$

则 $f_3(s_3)=\max\limits_{x_3=s_3}\{d_3(s_3,x_3)+f_4(s_4)\}=\max\limits_{x_3=s_3}\{-x_3^2+0\}=-s_3^2$

$k=2$时，因$s_3=s_2-x_2\ge 0$，故$0\le x_2\le s_2$

则
$$\begin{aligned} f_2(s_2)&=\max_{0\le x_2\le s_2}\{d_2(s_2,x_2)+f_3(s_3)\}=\max_{0\le x_2\le s_2}\{x_2^2-s_3^2\} \\ &=\max_{0\le x_2\le s_2}\{x_2^2-(s_2-x_2)^2\}=\max_{0\le x_2\le s_2}\{2s_2x_2-s_2^2\} \end{aligned}$$

令$h_2=(s_2,x_2)=2s_2x_2-s_2^2$

由$\dfrac{dh_2}{dx_2}=2s_2\ge 0$

则 h_2 为增函数

故最大值点为 $x_2 = s_2$

而 $f_2(s_2) = s_2^2$

$k = 1$ 时，因 $s_2 = s_1 - x_1 \geq 0$ 故 $0 \leq x_1 \leq s_1$

则 $f_1(s_1) = \max\limits_{0 \leq x_1 \leq s_1} \{d_1(s_1, x_1) + f_2(s_2)\} = \max\limits_{0 \leq x_1 \leq s_1} \{x_1^2 + s_2^2\}$

$= \max\limits_{0 \leq x_1 \leq s_1} \{2x_1^2 - 2s_1x_1 + s_1^2\}$

令 $h_1 = (s_1, x_1) = 2x_1^2 - 2s_1x_1 + s_1^2$

由 $\dfrac{dh_1}{dx_1} = 4x_1 - 2s_1 = 0$ 得 $x_1 = \dfrac{1}{2}s_1$

又 $\dfrac{d^2h_1}{dx_1^2} = 4 > 0$，故 $x_1 = \dfrac{1}{2}s_1$ 为极小值点。

则最大值点可在 $x_1 = 0$ 或 $x_1 = s_1$ 处取

因 $h_1(0) = s_1^2, h_1(s_1) = s_1^2$

则 $x_1 = 0$ 及 $x_1 = s_1$ 均为 $h_1(s_1)$ 的最大值点

且 $f_1(s_1) = s_1^2$ 由 $s_1 = 4$

故 $f_1(4) = 16$ 为最优值

再按计算过程反推得最优决策。

当 $x_1 = 0$ 因 $s_2 = s_1 - x_1$，则 $s_2 = 4$

故由 $x_2 = s_2$ 则 $x_2 = 4$

因 $s_3 = s_2 - x_2$，故 $s_3 = 0$ 则由 $x_3 = s_3$，则 $x_3 = 0$

得一个最优决策为：$x_1 = 0$，$x_2 = 4$，$x_3 = 0$

当 $x_1 = s_1 = 4$ 时，同理可得 $s_2 = 0$

故 $x_2 = 0$，由此得 $s_3 = 0$

则 $x_3 = 0$

得另一个最优决策为：$x_1 = 4$，$x_2 = 0$，$x_3 = 0$

则所得结果就是原非线性规划问题的最优解及优值。

由上例可知，用动态规划解题时，因 s_k 及 x_k 均可取某区间上的连续值，故可以求解如此例中的非线性规划或一些线性规划（除 x_i 的非负要求外，只有一个约束关系式）问题。

如果进一步要求上述规划问题的约束条件中 $x_i \geq 0$，且为整数，则变为求解一个整数规划的问题了。整数规划的求解是较困难的。既然线性、非线性规划（特殊的）问题可以转化求解动态规划问题，那么与这些规划相对应的整数规划问题是否也可以转化为动态规划问题来求解？后面将通过动态规划的应用例题来说明整数规划问题也可用动态规划方法求解，为此出现了动态规划的离散化解法。它的思想方法是取状态变量 s_k 为离散值（不论原问题变量是否为离散值问题），则决策变量 x_k 取值当然也为离散值，离散值通常为整数，然后将离散值分别代入顺序解法的递推公式（6-10），从中选出最优值 $s_k(x_k^*)$。运用计算机求解较复杂的动态规划问题时，可以采用这种离散化解法。

6.4 动态规划在管理决策中的应用举例

动态规划的应用范围很广。它用在管理决策方面可解决生产计划、资源分配、存贮、设备更新、设备可靠性等经常遇到的问题。下面将通过几个例子，说明动态规划方法在这些管理决策问题中的应用。

6.4.1 生产计划问题

如果已知某企业所产产品的生产费用、存贮费用和市场的需求量，在其生产能力和存贮能力许可的前提下，正确确定各个时期的生产量，使既完成交货计划，又使总支出最少，这显然是一个多阶段决策问题。下面通过例子说明如何用动态规划解决这类生产计划问题。

例 6.8 某工厂与用户签订了如表 6-6 所示的交货合同，表中数字为月底的交货量。该厂的生产能力为每月 400 件，该厂仓库的存货能力为 300 件。已知每百件货物的生产费用为 1000 元，在进行生产的月份，工厂还要支出经常费 4000 元，仓库保管费为每百件货物每月 1000 元。假定开始时及 6 月底交货后无存货，试问应在每个月生产多少件物品才能既满足交货任务又使总费用最小？（为了便于计算，以 100 件为单位来考虑这个问题。）

交货任务表　　　　　　　　　　　　　　　　　　　　　　表 6-6

月份	1	2	3	4	5	6
货物量（百件）	1	2	5	3	2	1

解：为了用动态规划的方法解决这一问题，先分析该问题每月的生产活动。工厂每月初有某一存货量，从这一存货量出发，工厂需决定这个月的生产量；原来的存货量加上当月的生产量，应能满足月交货计划的要求；月底按计划交货后，剩下的货物就是下个月初的存货量。这种过程每月重复一次。

以每月为一个阶段，共分六个阶段。对每个阶段i，以月初的存货量为它的状态变量s_i，该月的生产量为其决策变量x_i；若以k_i代表该月底的交货量，则可得状态转移方程为：

$$s_{i+1} = s_i + x_i - k_i \tag{6-15}$$

即，某月初的货物存贮量加上该月的生产量，再减去这个月的交货量，就等于下月初的货物存贮量。

现在来确定状态变量和决策变量的取值范围。根据题意，状态变量s_i必须满足以下条件：

（1）　由于工厂仓库容量的限制：

$$s_i \le 3 \tag{6-16}$$

（2）　不能超过本月以及后面各月交货量的总和：

$$s_i \le \sum_{j=i}^{6} k_j \tag{6-17}$$

（3）　为保证按时交货，月初的存货量加上本月的最大所能生产量不应小于本月的交货量：

$$s_i + 4 \ge k_j \tag{6-18}$$

在确定决策变量的取值范围时，应考虑到以下条件：

（1）　不能超过工厂每月的生产能力：

$$x_i \le 4 \tag{6-19}$$

（2）　存货量不能超过仓库容量：

$$x_i + s_k - k_i \le 3 \tag{6-20}$$

（3）　某月所拥有的货物量，不应超过本月及以后各月交货量的总和：

$$x_i + s_i \le \sum_{j=1}^{6} k_j \tag{6-21}$$

（4） 某月所拥有的货物量，应不少于本月的交货量：

$$x_i + s_i \ge k_i \tag{6-22}$$

此外，根据状态变量和决策变量的定义，显然有 $s_i \ge 0$， $x_i \ge 0$。

由题设，开始时和结束时均无存货，故边界条件 $s_1 = 0$， $s_7 = 0$。

在这个问题中，假定单位货物的生产费和保管费不随阶段而变化。若某个月（阶段）不生产，则仅需支出保管费；如果进行生产，则除了支出生产费之外，还要支出经常费。需要注意的是，生产费放在哪个月都一样，并不影响其总值。总生产费等于

(1+2+5+3+2+1)×10000=140000（元）

这样一来，在确定最优生产计划时，可不考虑生产费。

以每月的费用支出（不包括生产费）为阶段指标（阶段损益）d_i（单位为 1000 元），从而

$$\{d_i = s_i, d_i = s_i + 4, \quad \begin{matrix} \text{当} x_i = 0 \\ \text{当} x_i > 0 \end{matrix} \tag{6-23}$$

指标函数为各阶段指标和形式，得基本方程（递推公式）：

$$\begin{cases} f_i(s_i) = \min\limits_{x_i \in D_i}[d_i + f_{i+1}(s_{i+1})], i = 6,5,4,3,2,1 \\ f_7(s_7) = 0 \end{cases} \tag{6-24}$$

上式中 d_i 为阶段 i 相应状态的决策集合。

下面用逆序递推法进行数量计算。

阶段 6　　　　表 6-7

s_6	X_6	s_7	d_6	f_7	$f_6 = d_6 + f_7$
0	1	0	4	0	4+0=4
1	0	0	1	0	1+0=1

阶段 6：

由式（6-17），$s_6 \le 1$，从而 $s_6 = 0$，1；再由 $s_7 = 0$ 和式（6-15），$x_6 = 1 - s_6$，即 $x_6 = 1$，0。可以验证，上述结果满足（6-15）至（6-22）各式。阶段收益按式（6-23）计算。其计算结果列入表 6-7 中（在以下各表中将 $f_i(s_i)$ 简写为 f_i）。

阶段 5 和阶段 6：

由式（6-17），$s_5 \le \sum_{j=5}^{6} k_5 + k_6 = 2+1=3$，故$s_5 = 0,1,2,3$。然后由（6-21）和（6-22）求出对应的$x_5$。例如，当$s_5 = 1$时，由式(6-21)，$x_5 \le \sum_{j=5}^{6} k_j - s_5 = k_5 + k_6 - s_5 = 2+1-1=2$，故$x_5$所取 0，1，2 三个值；再由式（6-22）求得，$x_5 \ge k_5 - s_5 = 2-1=1$，即$x_5$可取 1，2，⋯，同时考虑上述两个条件，得$x_5 = 1,2$。用同样的方法，可求出和$s_5$的其他值对应的决策变量$x_5$的取值范围，结果见表 6-8。

阶段收益仍按式（6-23）计算。在计算各状态的值之前，应先找出各情况下对应的S_6的值，以便弄清其后续状态。计算中所需的后续状态的最优值由上一个表查出。该问题的计算结果列于表 6-8 中。

对于阶段 4-6、阶段 3-6、阶段 2-6 和阶段 1-6 等问题，可用与上述类似的方法进行分析并实行最优化，其数量计算分别列于表 6-9 至表 6-12 中。和前面两步相同，填入表中的各状态变量和决策变量的数值，应同时满足关系式（6-15）至（6-22）。

阶段 5 至 6　　　　表 6-8

s_5	X_5	s_6	d_5	f_6	$d_5 + f_6$
0	2	0	4	4	4+4=8
0	3	1	4	1	4+1=5*
1	1	0	5	4	5+4=9
1	2	1	5	1	5+1=6*
2	0	0	2	4	2+4=6*
3	1	1	6	1	6+1=7
3	0	1	3	1	3+1=4*

表中*号旁的数字为相应状态的最优值

阶段 4 至 6　　　　表 6-9

s_4	X_4	s_5	d_4	f_5	$d_4 + f_5$
0	3	0	4	5	4+5=9*
	4	1	4	6	4+6=10
1	2	0	5	5	5+5=10*
	3	1	5	6	5+6=11
	4	2	5	6	5+6=11

2	1	0	6	5	6+5=11
	2	1	6	6	6+6=12
	3	2	6	6	6+6=12
	4	3	6	4	6+4=10*
3	0	0	3	5	3+5=8*
	1	1	7	6	7+6=13
	2	2	7	6	7+6=13
	3	3	7	4	7+4=11

阶段 3 至 6 **表 6-10**

s_3	X_3	s_4	d_3	f_4	d_3+f_4
1	4	0	5	9	5+9=14*
2	3	0	6	9	6+9=15*
	4	1	6	10	6+10=16
3	2	0	7	9	7+9=16*
	3	1	7	10	7+10=17
	4	2	7	10	7+10=17

阶段 2 至 6 **表 6-11**

s_2	X_2	s_3	d_2	f_3	d_2+f_3
0	3	1	4	14	4+14=18*
	4	2	4	15	4+15=19
1	2	1	5	14	5+14=19*
	3	2	5	15	5+15=20
	4	3	5	16	5+16=21
2	1	1	6	14	6+14=20*
	2	2	6	15	6+15=21
	3	3	6	16	6+16=22

3	0	1	3	14	3+14=17*
	1	2	7	15	7+15=22
	2	3	7	16	7+16=23

阶段 1 至 6　　表 6-12

s_1	X_1	s_2	d_1	f_2	d_1+f_2
0	1	0	4	18	4+18=22*
	2	1	4	19	4+19=23
	3	2	4	20	4+20=24
	4	3	4	17	4+17=21*

各阶段的最优决策　　表 6-13

月份（阶段）	1	2	3	4	5	6
最优决策 X^*	4	0	4	3	3	0

最后再求出最优决策序列如表 6-13。求优过程计算相反顺序，由阶段 1 对应的表 6-12 开始，按各表逐段求得最优决策 x_k^* 及后一阶段的相应状态 s_{k+1} 及 x_{k+1}^*。例如，由表 6-12 可知，最小费用 $f_1=21$（千元），第一阶段（即一月份）的最优决策是 $x_1^*=4$，即生产货物 400 件。它来自状态 $s_2=3$，由表 6-11 知，第二阶段（2 月份）的最优决策 $x_2^*=0$，即不生产。而相应的 s_2 为 3，在表 6-11 中，对应于 $s_2=3$ 有 $x_2^*=0$，则 $s_3=1$。如此反复进行即可得到表 6-13。

6.4.2 资源分配问题

资源分配问题就是将供应量有限的一种或若干种资源（例如原材料、资金、劳力、时间、空间、运载能力等）分配给若干个使用者，而使某一目标函数最优。

假定有某一资源，其数量为 a，需将它分配给 n 个使用者（或生产活动）而使总收益最大，若分配给第 i 个使用者的数量为 x_i，且由此产生的收益为 $d_i=(x_i)$，则可将这种问题表示为

$$\begin{cases} \max\left[\sum_{i=1}^{n} d_i(x_i)\right] \\ \sum_{i=1}^{n} x_i = a \\ x_i \geq 0, \quad i=1,2,\cdots,n \end{cases}$$

这是一种资源的分配问题，也称一位分配问题。

为了用动态规划的方法解决这种问题，就需要恰当的规定阶段、状态、决策、阶段损益和最优损益函数等在此种情况下所代表的意义，并建立状态转移方程和满足递推关系的动态规划函数方程（即递推公式）。通常以把资源分配给某个或某些使用者（或消费者）这一过程作为一个阶段。若有n个使用者，则为n阶段问题，我们约定，若将物品分配给第i个及以后各个使用者，就说过程处于阶段i，在某一阶段，能投入分配的物品数量由多有少，这就决定了不同的状态，一般要考察各种可能情况（状态），以便找出最优的分配数量。

通常以x_i为阶段i的决策变量。现以s_i为阶段i的状态变量，它表示可用于分配给阶段i（及其以后各阶段）的资源量，显然有

$$0 \le x_i \le s_i$$

在这种问题下，把不同数量的物品分配给某个使用者所产生的损益（阶段损益）往往是已知的。根据最优性原理，可写出计算最优损益函数值的递推公式：

$$\begin{aligned} f_i(s_i) &= \max_{x_i}\left[d_i(s_i, x_i) + f_{i+1}(s_{i+1})\right] \\ &= \max_{x_i}\left[d_i(x_i) + f_{i+1}(s_i - x_i)\right] \end{aligned} \tag{6-26}$$

式中，s_{i+1}为阶段$i+1$的状态变量。

显然 $$s_{i+1} = s_i - x_i \tag{6-27}$$

这就使状态转移方程。

边界条件是：第一阶段的状态变量等于资源总量（将全部资源投入分配）；第$n+1$阶段的状态变量等于零（将全部资源分配完毕），即

$$s_1 = a\ ；\qquad s_{n+1} = 0$$

此外，还有

$$f_n(s_n) = d_n(s_n) \tag{6-28}$$

若状态变量的取值为离散的，计算和以前的例子类似。若分配量从 0 到a连续变化，为了用计算机计算，需将其离散化，例如将其分为$s = 0, 1\Delta, 2\Delta, \cdots, m\Delta = a$。这样，共有$m+1$个状态所取。分别考虑不同的状态，即可从中选出指标值最优时的分配数量。Δ的大小应根据计算机容量、计算时间、对最优解要求的精确度等，视具体问题而定。

例 6.9 某工厂生产A、B和C三种产品，它们都要使用某种原材料，原材料共计 4 吨，将不同数量的这种原材料分配给各种产品产生的收益，如表 6-14 所示（表中空白表示这种分配不增加收益），试确定使总收益最大的分配法。

表 6-14

原料的分配量	产品种类		
(吨)	A	B	C
0	0	0	0
1	10	6	8
2	17	17	11
3	20	18	

解：根据上面的说明，将问题分为三个阶段：阶段 1，2 和 3；它们分别代表将原材料分配给产品 A、B、C。显然 $s_1 = 4$（吨），同时，假定把原材料全部分配，不留作后用，这样就有 $s_{n+1} = s_4 = 0$。其状态变量和决策变量的取值范围应和表 6-14 相一致，计算结果列于表 6-15、表 6-16 和表 6-17 中。

表 6-15

阶段	s_3	X_3	d_3	$f_3 = d_3$
3	0	0	0	0
	1	1	8	8
	2	2	11	11

表 6-16

阶段	s_2	X_2	d_2	$s_3 = s_2 - X_2$	$f_2 = d_2 + f_3$
2	0	0	0	0	0
	1	1	6	0	6+0=6
		0	0	1	0+8=8*
	2	2	17	0	17+0=17*
		1	6	1	6+8=14
		0	0	2	0+11=11
	3	3	18	0	18 0=18
		2	17	1	17+8=25*
		1	6	2	6+11=17

		0	0	3	0+11=11
	4	3	18	1	18+8=2
		2	17	2	17+11=28*
		1	6	3	6+11=17
		0	0	4	0+11=17

表 6-17

s_1	X_1	d_1	$s_2=s_1-X_1$	$f_1=d_1+f_2$	
	4	3	20	1	20+8=28
1		2	17	2	17+7=24
		1	10	3	10+25=35*
		0	0	4	0+28=28

由上述各表可知，最大总收益等于 35，最优决策序列是$x_1^*=1$， $x_2^*=2$， $x_3^*=1$，即分配 1 吨原料给产品A，分配 2 吨原料给产品B，分配 1 吨原料给产品C。根据表 6-14，其总收益等于 10+17+8=35。

例 6.10　考虑下述两种资源的分配问题

$$\begin{cases} \max\left[d_1(x_1,y_1)+d_2(x_2,y_2)\right] \\ x_1+x_3=3 \\ y_1+y_2=3 \\ x_1,x_2,y_1,y_2\geq 0 \ \text{且为整数} \end{cases} \tag{6-38}$$

已知使用不同资料生产的收益如表 6-18 所示。

表 6-18

$d_1(x_1,y_1)$					$d_2(x_2,y_2)$				
x_1 \ y_1	0	1	2	3	x_2 \ y_2	0	1	2	3
0	0	1	5	6	0	0	3	5	6
1	1	3	6	7	1	2	4	7	8
2	3	5	7	8	2	3	5	8	9
3	4	6	8	10	3	4	6	9	10

解：引入拉格朗日乘子λ，以消去第二个约束，这就得到

$$\begin{cases} \max\left[d_1(x_1,y_1)+d_2(x_2,y_2)-\lambda(y_1+y_2)\right] \\ x_1+x_3=3 \\ x_1,x_2,y_1,y_2\geq 0 \quad \text{且为整数} \end{cases} \tag{6-39}$$

其中λ是一个参数。

在该问题中，x_1,x_2,y_1和y_2的所有可能选择由下述条件决定：$x_1+x_2=3$，$x_1\geq 0$，$x_2\geq 0$，$y_1\geq 0$，$y_2\geq 0$，$y_1\leq 3$，$y_2\leq 3$。现令$u=y_1+y_2$ $v=d_1(x_1,y_1)+d_2(x_2,y_2)$。在坐标$uov$中画出对应能够与每一可能选择的点（即图 6-11 中的小黑点），显然，原来二维分配问题的最优解对应于图中的点（3，12）。

现令

$d_1(x_1,y_1)+d_2(x_2,y_2)-\lambda(y_1+y_2)=V-\lambda u=c$

显然，当C和λ取定时，它代表坐标系uov中的一条直路。我们知道，求解问题（6-39），就是对一固定的λ寻求满足（6-39）中所列条件的x_1,x_2,y_1和y_2，而使C取最大值。这就相当于在图 6-11 中将斜率固定为λ的直线尽量平行上移，使它在通过图中黑点的情况下在v轴具有最大的截距。为了得到问题（6-38）的解，最终选取的λ应能和$y_1+y_2=3$相适应，显然$\lambda=2$满足这个要求。

图 6-12 示出了$u=y_1+y_2$和λ的关系。若把问题（6-38）中的第二个约束条件换成$y_1+y_2=1$，则由图 6-11 和图 6-12 均可看出，在这种情况下，用这种方法求不出最优解。

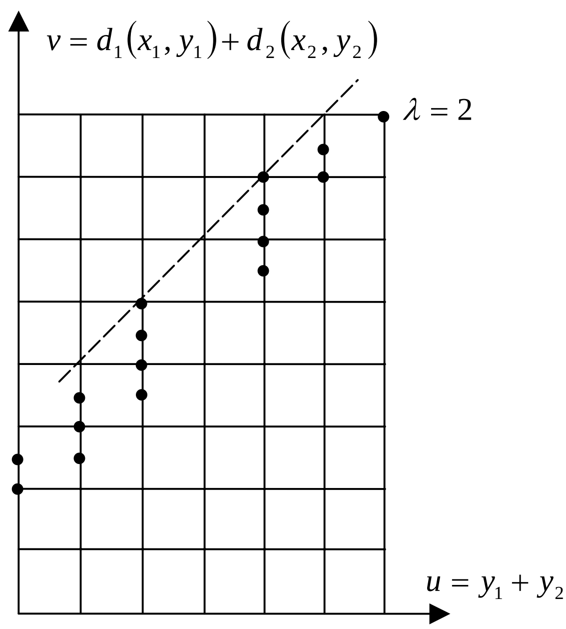

图 6-11

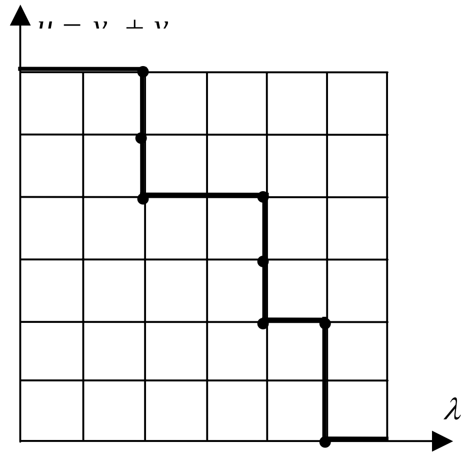

图 6-12