动态规划

张润琦　编

北京理工大学出版社

动 态 规 划

张润琦 编

北京理工大学出版社

内 容 简 介

本书主要介绍动态规划的基本方法和技巧，也涉及某些理论问题。第一章和第二章讨论确定性多段决策问题；第三章主要叙述无限段决策过程的两种算法；第四章讨论变量取有限多值的离散随机决策过程；第五章简要介绍时间连续的动态规划。

本书适于作工科研究生和本科高年级学生的教材，列举了较多的典型例题，用以说明基本概念、方法和技巧。本书内容力求全面反映动态规划的基本方面，保持其系统性。叙述力求简明易懂，作到深入浅出。

动 态 规 划

张 润 琦 编

*

北京理工大学出版社出版

新华书店北京发行所发行　各地新华书店经售

国防科工委印刷厂印刷

*

787×1092毫米　32开本　7印张　148千字

1989 年6月 第一版　　1989 年6月 第一次印刷

ISBN 7-81013-221-0/O·36

印数 1——2500册　定价：1.45 元

前　　言

动态规划是运筹学的一个重要分支，它的应用十分广泛。本书是作者近几年来在北京理工大学为研究生讲授动态规划课程时写的教材。全书大约需要讲授60学时，可以根据具体情况予以增删。

本书主要介绍动态规划的基本方法和技巧，也涉及某些理论问题。整个叙述力求简明易懂。全书共五章。第一章和第二章讨论确定性多段决策问题；第三章主要叙述无限段决策过程的两种算法；第四章讨论变量取有限多值的离散随机决策过程；最后一章简要介绍时间连续的动态规划。

在本书的写作过程中，作者得到了吴沧浦教授的指导和帮助，他还认真地审阅了原稿，提出了具体的修改意见。在此表示衷心感谢。

由于作者水平有限，书中一定存在不少缺点和错误，诚恳希望读者给予批评、指正。

目　录

第一章　确定性多段决策问题

第二章　多维决策问题

第三章　函数迭代法和策略迭代法

第一章　确定性多段决策问题

本章讨论决策问题的一种简单模型，它是确定性的、可以分为有限多段的决策问题。我们将介绍动态规划的基本概念、基本方程，并且用它来解实际问题。

§1.1　引　例

为了把问题说得明确具体，我们先列举几个简单例子。

例1.1.1　最短路问题

在图 1-1 中，各结点表示地点，线段表示相应两点间的道路，线段上的数字表示相应距离。求由点 A 到点 G 的最短道路。

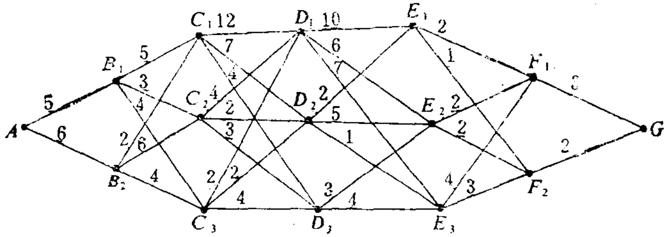

图1-1

自然，解此问题可以先求出一切可能的由点 A 到点 G 的

道路，求出各条道路的总长度，再比较它们的大小，选出其中的最小者即为所求。这种解法(穷举法)想起来十分简单，但真要实现它却很难。特别是当地点多、道路多时，计算量大得惊人，即使用大型计算机，所花费的时间之长也会使求解失去实用价值。

用上述的穷举法解我们的例题，共有96条道路：上例可以分为6段，由点A出发逐段移向G，除D_3外，在同一阶段上从每一点出发的道路数都相同，所以道路总数为

$$(2\times3\times3\times3\times2)-(2\times3\times2)=96$$

上式左端第一个括号中的数为假定D_3到E_1存在一条道路时，由A到G的道路总数；第二个括号中的数为添加道路D_3E_1后，在整个网络中多出的道路数，所以两者之差就是实际的道路总数。如果逐条道路计算长度，则需要的加法次数为

$$96\times5=480$$

如果按段计算，则可省去一些重复计算，所需加法次数为

$$6+18+48+96+96=264$$

例如，到第二段为止，前两段的相应数相加，需要6次加法；到第三段为止又需要18次加法，依次类推就得到上式。比较各条道路的长短需要95次比较运算。

在上述计算中，有许多运算对最短道路的求解并不是必要的。如何把这些不必要的计算省略掉就是简化计算的关键。

要求由A到G的最短路，如果我们知道了由B_1到G以及由B_2到G的最短路，相应长度记为S_{B_1}和S_{B_2}，则我们易于求出由A到G的最短路，这只要比较数值$5+S_{B_1}$和$6+S_{B_2}$，

并取小者

$$\min\begin{pmatrix}5+S_{B_1}\\6+S_{B_2}\end{pmatrix}$$

然而S_{B_1}和S_{B_2}是未知的。要想知道 S_{B_1} 和 S_{B_2}，只要先求得S_{C_1}、S_{C_2}和S_{C_3}，并作比较：

$$S_{B_1}=\min\begin{pmatrix}5+S_{C_1}\\3+S_{C_2}\\4+S_{C_3}\end{pmatrix}$$

$$S_{B_2}=\min\begin{pmatrix}2+S_{C_1}\\6+S_{C_2}\\4+S_{C_3}\end{pmatrix}$$

依次类推。最后，只要求出由 F_1 到G以及由 F_2 到G的最短路。显然有

$$S_{F_1}=3,\quad S_{F_2}=2$$

这样，我们的计算由点F_1、F_2开始，逐步远离点G，最后导向点A。于是得到了由A到G的最短路长度，同时也得到了由任意一点到G的最短路的长度，相应的最短路也可以由此得出。计算过程这里不详细写出，结果标在图 1-2 中。图上

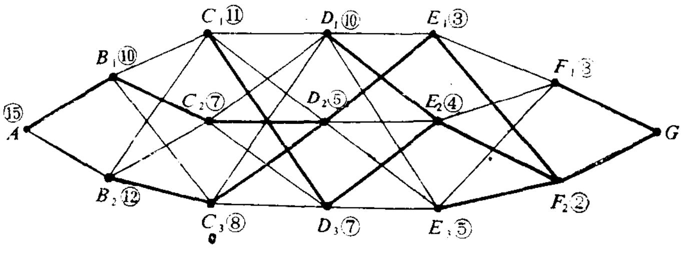

图1-2

某点处圆圈中的数字表示由此点到点G的最短路长度，由此点到点G的最短路用粗线表示。

上述解法所需要的运算次数大为减少。在E_1点求S_{E_1}，要作2次加法和1次比较，在这一段共需要6次加法和3次比较；在下一段共需要8次加法和5次比较、依次类推。共需要的加法次数为

$$6+8+9+6+2=31$$

需要的比较次数为

$$3+5+6+4+1=19$$

如此简化计算的关键在于两个想法。一是所谓的最优性原理：

由A到G的最短路具有这样的性质，不论在A的初始选择如何，由A的下一点开始到G的剩余路必然是由此点到G的最短路。简单地说，由A到G的最优路，其子路必然也是最优的。

利用最优性原理，当我们计算某一点到G的最短路长度时，可以充分利用前一步的计算结果，从而省略大量不必要的计算。

二是所谓的逆序递推算法。由终点开始，逐步导向起点，依次计算相应的最短路长度。

值得注意的是，上述的动态规划算法不仅减少了大量的计算，而且得到了丰富的结果。不仅求得了由A到G的最优路，而且求得了由任意一点到G的最优路。而穷举法仅仅得到了由起点A到G的最优路，如果要求由其它点，如B_2、C_3等，到G的最优路，必须重新计算。动态规划算法的这一优点在解决实际问题时有重要价值。

例1.1.2　背包问题

一旅行者需要带某些物品。为了简单起见，假设可以在 4 种物品中随意挑选。已知每件物品的重量及其效用，效用可以用数量表示出来。又假设旅行者的背包最多只能装10千克物品。表 1-1 给出了相应数据。选取装入背包中的物品及件数，使总效用最大。

表1-1

物　　品	重量（千克）	效　　用
1	5	26
2	3	16
3	2	9
4	1	5

给出的数据十分简单，用试探法易于求解。我们的目的是通过对问题的分析得出用动态规划解此类问题的基本思想。

这是含有 4 个变量的最优问题。为了使问题的求解简化，类似于例 1.1.1 的作法分步求解。首先对某一种物品求最优，假设在背包中装物品 1、2、3 的最优件数已定，须决定应如何装物品 4，使达到的效用最大。对物品 4 求最优时，其限装数量应是 0 到10的所有整数，记它为x_4。相应地，求得的最大效用是x_4的函数，记为$f_4(x_4)$。由此得到

$$f_4(x_4)=\max_{0\leqslant u_4\leqslant x_4}(5u_4)=5x_4 \tag{1-1}$$

记u_4的最优值为u_4^*，则

$$u_4^* = x_4 \tag{1-2}$$

式中的u_4表示第 4 种物品的件数，$x_4 = 0, 1, \cdots, 10$。

第二步考虑背包中装有物品 3 和物品 4。装入物品 3 的件数为u_3，两种物品的总限重为x_3，与x_3相应的最大效用记为$f_3(x_3)$。在x_3、u_3和x_4之间存在下面的关系

$$x_4 = x_3 - 2u_3 \tag{1-3}$$

类似于例 1.1.1 有最优性原理：两种物品 3 和 4 的最优组合，对于相应的单独一种物品 4 来说也是最优的。由此得关系式

$$\begin{aligned} f_3(x_3) &= \max_{0 \leqslant 2u_3 \leqslant x_3} [9u_3 + f_4(x_4)] \\ &= \max_{0 \leqslant u_3 \leqslant [x_3/2]} [9u_3 + f_4(x_3 - 2u_3)] \\ &= \max_{0 \leqslant u_3 \leqslant [x_3/2]} [9u_3 + 5(x_3 - 2u_3)] \\ &= 5x_3, \quad x_3 = 0, 1, \cdots, 10 \end{aligned} \tag{1-4}$$

记u_3的最优值为u_3^*，则

$$u_3^* = 0 \tag{1-5}$$

上面的推导过程用到了式（1-3）和式（1-1）。

第三步考虑背包中装有物品2、3和4。将装入背包中的物品 2 的件数记为u_2，总限重为x_2，相应的最大效用记为$f_2(x_2)$。存在关系式

$$x_3 = x_2 - 3u_2 \tag{1-6}$$

类似地得到

$$f_2(x_2) = \max_{0 \leqslant 3u_2 \leqslant x_2} [16u_2 + f_3(x_3)]$$

$$= \max_{0 \leqslant u_2 \leqslant [x_2/3]} [16u_2 + f_3(x_2 - 3u_2)]$$

$$= \max_{0 \leqslant u_2 \leqslant [x_2/3]} [16u_2 + 5(x_2 - 3u_2)]$$

$$= 5x_2 + \left[\frac{x_2}{3}\right] \qquad x_2 = 0, 1, \cdots, 10 \tag{1-7}$$

记u_2的最优值为u_2^*，则

$$u_2^* = \left[\frac{x_2}{3}\right] \tag{1-8}$$

因为u_2为物品件数，故只可以取整数。

最后考虑背包中装有 4 种物品。类似地，

$$x_2 = x_1 - 5u_1 \tag{1-9}$$

$$f_1(x_1) = \max_{0 \leqslant 5u_1 \leqslant x_1} [26u_1 + f_2(x_1 - 5u_1)] \tag{1-10}$$

x_1为 4 种物品的总限重，按原问题的假设，$x_1 = 10$。只要求$f_1(10)$，得到

$$f_1(10) = \max_{0 \leqslant u_1 \leqslant [10/5]} \left\{26u_1 + 5(10 - 5u_1) + \left[\frac{10 - 5u_1}{3}\right]\right\}$$

$$= \max_{0 \leqslant u_1 \leqslant 2} \left\{50 + u_1 + \left[\frac{10 - 5u_1}{3}\right]\right\}$$

$$= \max(53, 52, 52) = 53 \tag{1-11}$$

$$u_1^* = 0 \tag{1-12}$$

与 4 种物品的最优组合相应，效用为 53。下 面 来 求最优组合。

$u_1^* = 0$表示背包中不装入物品1；由式 (1-9)、(1-8)、(1-

6)、(1-3)和(1-2)分别得出 $x_2=10$，$u_2^*=3$，$u_3^*=0$，$x_3=1$，$x_4=1$和$u_4^*=1$。背包中装入的 4 种物品的最优组合简记为

$$\{u_1^*,u_2^*,u_3^*,u_4^*\}=\{0,3,0,1\}$$

这表明背包中仅装有 3 件物品 2 和 1 件物品4。由数据表1-1知，这显然是正确的。

例1.1.3 机器负荷分配问题

某种机器可以在高低两种负荷下生产，年产量与年初投入生产的机器数有关。在高负荷情况下，设年初投入的完好机器数为u_1，年终的完好机器数为$0.7u_1$，称系数0.7为机器完好率。记年产量为w_1，$w_1=8u_1$。在低负荷下，设年初投入的完好机器数为u_2，完好率为0.9，年产量为$w_2=5u_2$。假设开始时拥有1000台完好机器。要制定五年计划，每年年初决定重新分配在两种负荷下生产的完好机器数，使五年内的总产量最大。

此问题的求解与例 1.1.2 类似。由第 5 年开始计算。设x_5为第 5 年年初的完好机器数，u_5 为年初投入高负荷生产的完好机器数。自然，x_5-u_5 为年初投入低负荷生产的完好机器数。第 5 年的最高产量依赖于 x_5，记为$f_5(x_5)$。由问题给出的条件得

$$f_5(x_5)=\max_{0\leqslant u_5\leqslant x_5}[8u_5+5(x_5-u_5)]$$

$$=8x_5 \tag{1-13}$$

最优点 $$u_5^*=x_5 \tag{1-14}$$

这表明，第 5 年年初要把全部完好机器投入高负荷下生产。这显然是正确的。

对第 4 年，用类似的记号x_4、u_4，得关系式

$$x_5 = 0.7u_4 + 0.9(x_4 - u_4)$$
$$= 0.9x_4 - 0.2u_4 \tag{1-15}$$

记$f_4(x_4)$为第 4 年和第 5 年的最高总产量。此时的最优性原理为：最后两年的高负荷分配的最优组合，对相应的最后一年来说必定也是最优的。由此得出下面的关系式

$$f_4(x_4) = \max_{0 \leqslant u_4 \leqslant x_4} [8u_4 + 5(x_4 - u_4) + f_5(x_5)]$$
$$= \max_{0 \leqslant u_4 \leqslant x_4} [5x_4 + 3u_4 + 8(0.9x_4 - 0.2u_4)]$$
$$= 13.6x_4 \tag{1-16}$$
$$u_4^* = x_4 \tag{1-17}$$

此处用到式 (1-13) 和 (1-15)。

对第 3 年，考虑最后三年的情况，类似地得到关系式

$$x_4 = 0.9x_3 - 0.2u_3 \tag{1-18}$$
$$f_3(x_3) = \max_{0 \leqslant u_3 \leqslant x_3} [8u_3 + 5(x_3 - u_3) + f_4(x_4)]$$
$$= \max_{0 \leqslant u_3 \leqslant x_3} [5x_3 + 3u_3 + 13.6(0.9x_3 - 0.2u_3)]$$
$$= 17.5x_3 \tag{1-19}$$
$$u_3^* = x_3 \tag{1-20}$$

考虑后四年的情况，类似地有

$$x_3 = 0.9x_2 - 0.2u_2 \tag{1-21}$$
$$f_2(x_2) = \max_{0 \leqslant u_2 \leqslant x_2} [8u_2 + 5(x_2 - u_2) + f_3(x_3)]$$
$$= \max_{0 \leqslant u_2 \leqslant x_2} [5x_2 + 3u_2 + 17.5(0.9x_2 - 0.2u_2)]$$
$$= 20.8x_2 \tag{1-22}$$
$$u_2^* = 0 \tag{1-23}$$

考虑五年的情况，此时$x_1=1000$，得到

$$x_2=0.9x_1-0.2u_1 \tag{1-24}$$

$$f_1(1000)=\max_{0\leqslant u_1\leqslant 1000}[8u_1+5(1000-u_1)+f_2(x_2)]$$

$$=\max_{0\leqslant u_1\leqslant 1000}[5000+3u_1+20.8(900-0.2u_1)]$$

$$=23700 \tag{1-25}$$

$$u_1^*=0 \tag{1-26}$$

由此知道，五年内的最高总产量为23700。$u_1^*=0$，表示第1年年初要把所有完好机器投入低负荷下生产；由式(1-24)、(1-21)、(1-20)、(1-18)、(1-17)、(1-15)和(1-14)分别得$x_2=900, u_2^*=0, x_3=810, u_3^*=810, x_4=567, u_4^*=567, x_5=397$和$u_5^*=397$。高负荷生产的完好机器的最优组合简记为

$$\{u_1^*,\cdots,u_5^*\}=\{0,0,810,567,397\}$$

这表明，在前两年年初全部完好机器投入低负荷生产，后三年年初全部完好机器投入高负荷生产。第五年末的完好机器数为

$$0.7\times 397=278(\text{台})$$

在此例中，我们仅考虑了最高产量，而未考虑五年计划后的完好机器数。

§1.2 动态规划基本方程

1. 基本概念

介绍动态规划中常用的术语。

阶段 在处理§1.1中的例题时，我们总是把原来的最

优化问题分解为互相联系的几个小问题。原问题被看作一个过程，小问题就是过程的几个阶段。例1.1.1分为6个阶段，例1.1.2分为4个阶段。阶段往往按时间顺序划分，在实际应用中应视具体情况而定。阶段也常称为段、步、周期、级等。常用k表示它，k取自然数1,2,……。

状态 为了便于对过程作分析，我们要给出描述过程演变的一些参数，这就是要引入的状态变量。在例1.1.1中，各阶段的出发点都是该段的状态，点G是过程的终状态。在例1.1.3中，各年年初的完好机器数就是相应段上的状态变量，其终状态是第5年末的完好机器数，它是不确定的，依赖于人们选择的过程。在例1.1.2中，状态变量是各阶段上的限重。它不象其它两例那样自然，带有更多的人为性质。一般来说，状态变量可以用一个或一组变数表示。如果把例1.1.1的各结点适当编码，使各段上的点与数对应起来，其状态就是一个变数。状态变量常简称为状态。

状态的取值都有一个范围，称为状态集合。状态及状态集合都依赖于阶段k，分别记为x_k和X_k。当各段状态和终状态确定后，整个过程就完全确定了。即过程可以表示为状态序列$\{x_1, x_2, \cdots, x_n, x_{n+1}\}$，其中$x_1$为初状态，也称为起始状态，$x_{n+1}$为终状态。有时也称$x_1$和$x_{n+1}$为边界状态。

状态变量的取法根据具体问题而定，可以有不同的取法，但是都必须满足一个重要性质：由某段状态出发的后续过程（称为后部子过程，简称子过程）不受前面演变过程的影响。这就是说，由第k段的状态x_k出发的k子过程，可以看作是一个以状态x_k为初状态的独立过程$\{x_k, x_{k+1}, \cdots, x_{n+1}\}$。这一性质称为无后效性。它是动态规划中的状态和通常描述