应用动态规划

王日爽 徐 兵 魏权龄 编著

国防工业出版社

应 用 动 态 规 划

王日爽　徐　兵　魏权龄　编著

国防工业出版社

内容简介

本书介绍了动态规划的基本原理与方法。内容包括最优化原则、存在与唯一性定理等，介绍了用动态规划求解多阶段配置、多阶段安排、最短路、背包等知名问题，还包括用动态规划处理数学问题等内容，并介绍了动态规划与马尔可夫过程、动态规划与最优控制的联系。本书着重讲解用动态规划法解决实际问题的思想、原理与方法。书中还列举了动态规划在经济管理中的多种应用范例。

本书可作为理工科、财经等院校有关专业研究生、本科生或专科高年级学生的教材，也可供工业、经济、军事、管理等部门的工程技术人员、管理人员学习参考之用。

应用动态规划

王日爽 徐 兵 魏权龄 编著

*

国防工业出版社出版

新华书店北京发行所发行 各地新华书店经售

国防工业出版社印刷厂印装

*

787×1092 1/32 印张9 1/2 209千字

1987年7月第一版 1987年7月第一次印刷 印数：0,001—2,720册

统一书号：15034·3186 定价：1.95元

前 言

五十年代初，人们根据多阶段决策问题的特性，提出了解决这类问题的“最优化原则”，并研究了许多实际问题和数学模型，从而建立了数学规划的新分支——动态规划。1957年R. Bellman讲述最优化原理的著作是动态规划的第一本专著。

动态规划的方法在工程技术、经济管理、工业生产和军事运筹等方面都有着广泛的应用，其前景不可限量。这种方法日益为人们所观注，它吸引了越来越多的人在各自不同的领域作出新的努力。

在许多问题中利用动态规划要比用线性规划或非线性规划更有成效。但是动态规划并不象线性规划或非线性规划那样具有固定的解法，也就是说动态规划没有统一的处理格式，它必须依据问题本身的特性，利用灵活的数学技巧来处理。

本书是在几届教学讲稿的基础上修订而成。重点是介绍应用动态规划解决问题的思想、原理和方法，而不过多地注重冗长的数学推导。

本书的某些章节自成体系，读者可以根据需要选择其中部分章节阅读。在编写过程中我们参考了国内外一些有影响的著作，在此对这些著作的作者表示感谢。

限于我们的水平，难免有不当或疏漏之处，恳切希望得到批评指正。

目　　录

第一章　基本概念与最优化原则

动态规划是一类多阶段决策过程的最优化方法。所谓多阶段决策过程是指这样一类活动过程：即由于过程的特殊性质，它可以依据时间或空间划分为若干个互相联系的阶段，而在各阶段中，人们都需要作出方案的选择，我们称之为**决策**，并且当一个阶段的决策确定之后，常常影响到下一个阶段的决策，从而影响整个过程的活动。这样，各个阶段所确定的决策就构成一个决策序列，常称之为**策略**。由于各个阶段可供选择的决策往往不止一个，因而就可能有许多策略以供选择。这些可供选择的策略构成一个集合，我们称之为**允许策略集合**（有时简称为**策略集合**）。每一策略都相应地确定一种活动的效果，我们假定这个效果可以用数量指标来衡量。由于不同的策略常常导致不同的效果，因此，如何在允许策略集合中选择一个策略，使其在预定的标准下达到最好的效果，常常是人们所关注的，我们称这样的策略为**最优策略**。上面所描述的一类问题就是所谓的**多阶段决策问题**。

§1.1　几个例子

多阶段决策问题广泛地存在于各类活动中。前已提及，所谓阶段往往是依据时间或空间来划分的，但是有些问题从表面上看本身并不具备时间或空间的概念，似乎不是多阶段决策问题。可是有时人为地引入时间或空间的概念将过程划分为若干个阶段之后，可以将其转化为一个多阶段决策问题。

下面举几个例子：

例 1 多阶段配置问题（投资问题）

这是一个较为直接的例子。

设有数量为 $\bar{x}$ 的某种资源，今要投资到两个项目 A 与 B 中去。若第一次，以数量 y （$0 \leqslant y \leqslant \bar{x}$）投资于 A，设其经济效益为 $g(y)$；以数量 $\bar{x}-y$ 投资于 B，设其经济效益为 $h(\bar{x}-y)$；并称此为活动的第一阶段。那么，在第一阶段可得总经济效益为

$$g(y)+h(\bar{x}-y)$$

我们假定

$$g(0)=h(0)=0$$

（这是合理的，因为不予投资其经济效益即为零。）

显然，对不同的 y （即不同的投资方案），第一阶段的经济效益一般也不同。

又设以数量 y 和 $\bar{x}-y$ 分别投资于 A 与 B 两项生产后，可以回收一定的资源并再投入生产。并设其回收率分别为 a（$0 \leqslant a < 1$）和 b （$0 \leqslant b < 1$）。第一阶段投资后，回收的总资源为 x_1，则有

$$x_1=ay+b(\bar{x}-y)$$

我们将 x_1 再投资于 A 与 B 中去，称之为第二阶段。若第二阶段以资源 y_1（$0 \leqslant y_1 \leqslant x_1$）投资于 A，设其经济效益为 $g(y_1)$；第二阶段以资源 x_1-y_1 投资于 B，设其经济效益为 $h(x_1-y_1)$；则第二阶段的总效益为

$$g(y_1)+h(x_1-y_1)$$

因此，两个阶段的总经济效益之和为

$$g(y)+h(\bar{x}-y)+g(y_1)+h(x_1-y_1)$$

又设第二阶段回收的总资源为 x_2，即有

$$x_2 = ay_1 + b(x_1 - y_1)$$

如此继续下去。假定第 $n-1$ 阶段回收的总资源为 x_{n-1}；并假定第 n 阶段以资源 y_{n-1}（$0 \leqslant y_{n-1} \leqslant x_{n-1}$）投资于 A，经济效益为 $g(y_{n-1})$；而以资源 $x_{n-1} - y_{n-1}$ 投资于 B，经济效益为 $h(x_{n-1} - y_{n-1})$。则前 n 个阶段的总经济效益为

$$g(y) + h(\bar{x} - y) + g(y_1) + h(x_1 - y_1) + \cdots + g(y_{n-1}) + h(x_{n-1} - y_{n-1})$$

其中

$$\left.\begin{aligned} & x_1 = ay + b(\bar{x} - y) \\ & x_2 = ay_1 + b(x_1 - y_1) \\ & \cdots\cdots\cdots\cdots\cdots \\ & x_{n-1} = ay_{n-2} + b(x_{n-2} - y_{n-2}) \\ & 0 \leqslant y \leqslant \bar{x} \\ & 0 \leqslant y_i \leqslant x_i, \quad i = 1, \cdots, n-1 \end{aligned}\right\} \tag{1.1.1}$$

我们希望选择 y，y_1，y_2，$\cdots$，y_{n-1} 以使 n 个阶段投资的总经济效益最大。即希望选择 $[y, y_1, \cdots, y_{n-1}]$，使之在条件（1.1.1）下达到

$$\max\{g(y) + h(\bar{x} - y) + g(y_1) + h(x_1 - y_1) + \cdots + g(y_{n-1}) + h(x_{n-1} - y_{n-1})\} \tag{1.1.2}$$

这显然是一个多阶段决策问题。

满足式（1.1.1）及（1.1.2）的策略 $[y, y_1, \cdots, y_{n-1}]$ 即为最优策略。

上例中数量 $\bar{x}$、a 和 b、以及函数 g 和 h 是已知的。这是一个典型的非线性规划问题。特别当 g 和 h 皆为线性函数时，则上述问题进而是一个线性规划问题。

不难发现此例具有下列特性：当 a、b、g 和 h 均为已

知时，其经济效益依赖于初始资源$\bar{x}$和过程进行的阶段数n。

若设$f_k(x)$为以初始资源x，经过k个阶段，且各阶段投资中均以最优决策投资后所得到的总经济效益。则依定义，应有

$$f_1(x)=\max_{0\leqslant y\leqslant x}[g(y)+h(x-y)] \tag{1.1.3}$$

由于第一阶段以y与$x-y$分别投资于A与B中，以后可以回收$x_1=ay+b(x-y)$，此时还需进行下个阶段的生产，采取最优策略时的总经济效益应为

$$f_1(ay+b(x-y))$$

因此前两个阶段的总经济效益之和为

$$f_2(x)=\max_{0\leqslant y\leqslant x}\{g(y)+h(x-y)+f_1(ay+b(x-y))\}$$

相应地，对任意$k=2$，3，…，n，可得

$$f_k(x)=\max_{0\leqslant y\leqslant x}\{g(y)+h(x-y)+f_{k-1}(ay+b(x-y))\} \tag{1.1.4}$$

事实上，由定义可知

$$f_k(x)\geqslant\max_{0\leqslant y\leqslant x}\{g(y)+h(x-y)+f_{k-1}(ay+b(x-y))\} \tag{1.1.5}$$

其次，设最优策略为$[y^*, y_1^*, \cdots, y_{k-1}^*]$，则

$$f_k(x)=[g(y^*)+h(x-y^*)]+[g(y_1^*)+h(x_1-y_1^*)]+\cdots+[g(y_{k-1}^*)+h(x_{k-1}-y_{k-1}^*)]$$

$$\leqslant [g(y^*)+h(x-y^*)] + f_{k-1}(ay^*+b(x-y^*))$$

$$\leqslant \max_{0\leqslant y\leqslant x}\{g(y)+h(x-y)+f_{k-1}(ay+b(x-y))\} \qquad (1.1.6)$$

由式（1.1.5）与（1.1.6）可知式（1.1.4）成立。

把式（1.1.3）与（1.1.4）写在一起，有

$$\left.\begin{aligned} f_k(x) &= \max_{0\leqslant y\leqslant x}\{g(y)+h(x-y) \\ &\qquad +f_{k-1}(ay+b(x-y))\} \\ &\qquad k=2,\ 3,\ \cdots,\ n \\ f_1(x) &= \max_{0\leqslant y\leqslant x}\{g(y)+h(x-y)\} \end{aligned}\right\} \qquad (1.1.7)$$

显而易见，式（1.1.7）是一个递推关系式，我们的问题是求 $f_n(\bar{x})$。可以设想依据它有可能求得问题的最优策略。

例 2　多阶段安排问题

这也是一个较为直接的例子。

设有两个金矿 A 和 B 可供开采。两矿的含金量分别为 x 和 y，假定开采成功的概率分别为 p_1 与 p_2（开采成功的标准既定，达到标准谓之成功，否则谓之失败，这里简单地把它抽象为成功抑失败的概念），且每阶段从两矿开采出金的数量占当时藏量的百分数分别为 r_1 与 r_2。问题是现有一部采掘机（不能同时用以开采两个矿）应如何安排 n 个阶段的采矿，才能使总的采金量期望值最大？

显然，每一阶段的开采或安排于 A 矿或安排于 B 矿，假定失败前各阶段的采矿均按最优策略安排；又假定 p_1、r_1、p_2、r_2 为已知，则采金的总期望值依赖于 A 矿和 B 矿的藏量

x与y，以及安排采矿的阶段数n。

若设$f_k(x, y)$表示A与B两矿的藏金量分别为x与y，安排k个阶段采矿，且每阶段均按最优决策安排所得的总期望值。那么可知：

如果第一阶段安排开采A矿且成功，则所得金为r_1x，故A矿尚余金量为$(1-r_1)x$，以后$k-1$次若均按最优决策安排，其期望值记为$f_A(x, y)$，则

$$f_1(x, y)=\max\{p_1r_1x, \ p_2r_2y\} \qquad (1.1.8)$$

$$f_A(x, y)=p_1[r_1x+f_{k-1}((1-r)x, \ y)]$$

同理，如果第一阶段安排开采B矿且成功，则得金为r_2y，故B矿尚余金量为$(1-r_2)y$，若以后$k-1$次均按最优决策安排，其期望值记为$f_B(x, y)$，则

$$f_B(x, y)=p_2[r_2y+f_{k-1}(x, \ (1-r_2)y)]$$

因而

$$\begin{aligned} f_k(x, y) &= \max\{f_A(x, y), \ f_B(x, y)\} \\ &= \max\{p_1[r_1x+f_{k-1}(1-r_1)x, \ y)], \\ &\qquad p_2[r_2y+f_{k-1}(x, \ (1-r_2)y)]\} \\ &\qquad k=2, \ 3, \ \cdots, \ n \end{aligned} \qquad (1.1.9)$$

显然式（1.1.9）也是一个递推关系式。可以设想利用式（1.1.9）有可能求解原问题。

例 3　最短路径问题

这是一个间接的例子。

考察如图 1.1 所示的线路网络，各点联线上标明的数字表示它们之间的距离，我们从A点出发，要走到目的地F点，问经过哪些点（走什么路线）才能使总路程最短？又最短路程为多少？

我们先把从A到F的线路人为地分为五个阶段，然后再

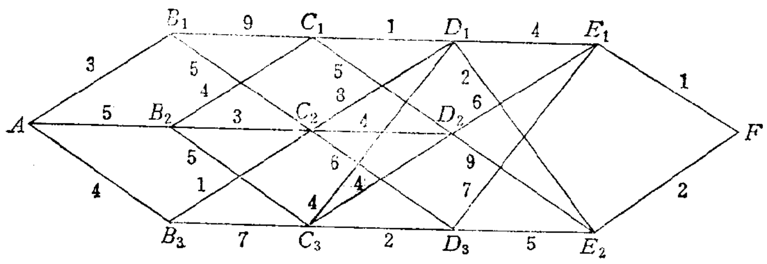

图1.1　最短路径问题

把它化为一个多阶段决策问题来处理。

我们在此例中可以引出阶段变量、状态变量和决策变量等基本概念。

设

n 表示由某点至终点 F 的阶段数，称 n 为**阶段变量**；

s 表示在任一阶段所处的位置，称 s 为**状态变量**；

$x_n(s)$ 表示当状态处于 s，已走了 n 个阶段的最后一个前导点，称 $x_n(s)$ 为**决策变量**；

$x_n(s)=s'$ 称为**状态转移方程**；

$f_n(s)$ 表示从始点 A 算起已走了 n 个阶段到达 s 点的最短路径之长；

$d(s, t)$ 表示从 s 点至下一点 t 的直接距离。

对于此例，在求解的各阶段中，可以建立相继两个阶段之间的如下关系：

$$\begin{cases} f_n(s)=\min\limits_{s}\{f_{n-1}(x_n(s))+d(x_n(s), s)\} \\ \qquad\qquad\qquad n=1, 2, \cdots, 5 \\ f_0(A)=0 \end{cases}$$

我们正是利用上述递推关系来求得最优策略的。§3.1中将给出这个问题的具体解法及计算结果。

§1.2 最优化原则

对于§1.1中多阶段决策问题的诸例，只要分析其数学模型，就可以发现它们都具有一些共同的特性，兹归纳如下：

（1）我们所研究的这些系统，在时间（或空间）过程的每一阶段，都可以由一组状态变量来描述。记之以向量 $\boldsymbol{p}$，并称为**状态向量**，其分量即为状态变量。

（2）在每一阶段，我们都要作出一个决策，使系统由一个状态 $\boldsymbol{p}$ 转移到另一个状态 $\boldsymbol{p}'$。因此状态的变化可以看作是经过一个变换 T_q 而得到的，即一个决策可以看作是一个变换：

$$T_q(\boldsymbol{p})=\boldsymbol{p}', \quad q\in Q \tag{1.2.1}$$

其中 q 表示决策，Q 表示允许决策集合。

（3）对系统在阶段 i 以后所需作出的决策，与系统在此阶段之前的历史无关，具有这种性质的过程又称为**马尔可夫**（Марков）**过程**，简称马氏过程。

（4）决策的目的是最优化状态变量的某个函数，称之为**目标函数**，它依赖于初始状态与阶段变量，我们把从过程开始到结束所做出的一系列允许决策称为一个策略，把使目标函数达到最优值的策略称为**最优策略**。

于是，以上各例可以抽象为统一的函数形式：

$$\left.\begin{aligned} &f_k(\boldsymbol{p})=\sup_{q\in Q}\{g(\boldsymbol{p},q)+h(\boldsymbol{p},q)f_{k-1}(T_q(\boldsymbol{p}))\} \\ &\qquad\qquad k=2,\cdots,n \end{aligned}\right\} \tag{1.2.2}$$

$$f_1(p)=\sup_{q\in Q} g(p,\ q)$$

式（1.2.2）是一个递推关系式，称之为动态规划的函数方程。特别当 $n\to\infty$ 时，可以用其极限形式

$$f(p)=\sup_{q\in Q}\{g(p,\ q)+h(p,\ q)f(T_q(p))\} \tag{1.2.3}$$

来近似之。

函数方程是根据动态规划的如下最优化原则而导出的：

（5）**最优化原则** 一个过程的最优策略，具有以下的性质，即无论系统的初始状态和初始决策如何，对于以初始决策所转移到的状态为以后过程的初始状态，就这后一过程而言，以后诸决策也必须构成最优策略。

例 1 运输问题

设有 m 个发货站，其发货量分别记为 a_1、a_2、…、a_m；还设有 n 个收货站，它们的收货量分别记为 b_1、b_2、…、b_n。

如果 c_{ij} 表示由第 i 个发货站至第 j 个收货站的运输单位费用，那么应如何调度才能使总运费最小？

若设 x_{ij} 表示第 i 个发货站运送到第 j 个收货站的货物数量。则可建立数学模型：

$$\begin{cases} \min \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\ \displaystyle\sum_{i=1}^{n} x_{ij}=b_i, \quad j=1,\ \cdots,\ n \\ \displaystyle\sum_{j=1}^{n} x_{ij}=a_i, \quad i=1,\ \cdots,\ m \\ x_{ij}\geqslant 0, \quad i=1,\ \cdots,\ m;\quad j=1,\ \cdots,\ n \end{cases} \tag{1.2.4}$$

$$\left(\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \right)$$

式（1.2.4）为典型的线性规划问题，通常称为运输问题。但是我们也可以将运输问题化为一个多阶段决策问题来处理。

记 $f_k(a_1, \cdots, a_m)$ 表示由 m 个分别有数量 a_1、$\cdots$、a_m 的发货站到 k 个收货站的最小总运费。

如果只有一个收货站，也就是将所有各发货站的货物都运到同一个收货站，那么易知：

$$f_1(a_1, \cdots, a_m) = \sum_{i=1}^{m} c_{i1} a_i \tag{1.2.5}$$

因而利用最优化原则，可有

$$f_k(a_1, \cdots, a_m) = \min\left\{ \sum_{i=1}^{m} c_{i1} x_{i1} + f_{k-1}(a_1 - x_{11}, a_2 - x_{21}, \cdots, a_m - x_{m1}) \right\} \tag{1.2.6}$$

$$k = 2, 3, \cdots, n$$

其中

$$\sum_{i=1}^{m} x_{i1} = b_1$$

$$0 \leqslant x_{i1} \leqslant a_i, \quad i = 1, 2 \cdots, m$$

依此顺序求出 x_{i1}, x_{i2}, $\cdots$, x_{in}, $i = 1, 2, \cdots, m$

这里 k 是阶段变量；各发货站的现有货物数量 a_i（$i = 1, 2, \cdots, m$）是状态变量；x_{ij}（$i = 1, 2, \cdots, m$；$j = 1, 2, \cdots, n$）是决策变量；$y_i = a_i - \sum_{j=1}^{k-1} x_{ij}$（$i = 1, 2, \cdots, m$）是状态转移方程。