西安交通大学出版社

DONGTAI
GUIHUA
YUANLI
JI
YINGYONG

# 动态规划原理及应用

姜衍智 编

# 动态规划原理及应用

姜衍智　编

西安交通大学出版社

# 内 容 简 介

本书系统地介绍了数学规划论的重要分支——动态规划的原理及其在工程技术和经济管理方面的应用。全书分三章，第一章介绍作为动态规划基础的分级决策方法，第二章介绍了目前已有的各种动态规划方法，第三章中列举了大量的应用实例。

本书取材适当，重点突出，由浅入深，概念清晰，便于自学。

本书可作为高等院校工科和管理专业教材，也适用于工程技术和经济管理人员阅读。

**动态规划原理及应用**

姜衍智 编

责任编辑 陆 薇

西安交通大学出版社出版
(西安市咸宁路28号)
国营五二三厂印装
陕西省新华书店发行 各地新华书店经售

开本787×1092 1/32 印张9.375 字数198千字
1988年1月第1版 1988年4月第1次印刷
印数：1—2,000册

ISBN1-5605-0029-3/TK-5
书号：15340·146

定价：1.55元

# 前　言

在工程技术和经营管理领域中，常常会遇到时序上连续而各时刻的状态又随时间变化的动态系统。其中大多数可以采用分级决策的方法寻求该系统的最佳决策。分级决策方法是把整个研究过程(时期)分成若干阶段，根据各阶段的状态与其他阶段的关联确定各阶段的决策值，并由各阶段的决策组成整个研究过程的“决策序列”。当求得的决策序列能使系统的总效益为最佳时，则称该决策序列为最佳决策序列，也就是系统的最佳解。

分级决策方法是建立在“整体优化”的基础上的。在寻求某一阶段的决策值时，不能仅从当前局部效益出发，只就本阶段的条件选择最佳决策。而是在选择任一阶段的决策时，都必须考虑到全系统的总体最佳的要求。

动态规划（Dynamic Programming, DP）是综合了分级决策方法和美国数学家里卡德·贝尔曼（Bellman·R.）在1957年提出的最佳化原理而形成的一种数学方法，它不仅可用于分析连续的时序问题，而且已广泛应用于任何可以分级处理的决策问题。

动态规划通过对一个多变量的复杂问题进行分级处理，把多变量的复杂问题变成了求解多个单变量的问题，使求解过程大大简化了。

动态规划的另一个优点是能够顺利地确定出全局最大（最小）值，而不是局部最大（最小）值。这一优点是现有的其它最佳化方法难以达到的。

动态规划的第三个重要优点是对于处理变量为整数的数学规划问题有着特别的功效。相反，变量的某些约束（如整数、非负等）对其它优化方法的应用都将产生困难。

尽管动态规划方法在很多领域中已经获得广泛的应用，但是迄今为止，并没有通用算法。为了使读者掌握动态规划方法的基本知识和演算技巧，本书采集了涉及工程技术和经营管理等方面的大量例题（多取材于参考文献中所列参考书目），供读者参考。

本书作为动态规划方法的入门书，考虑到读者对象为一般工程技术人员、经营管理人员和大专学生等，因此内容力求浅显易懂，便于读者自学，在数学上只要具备一般高等数学知识，即可顺利阅读。

作者对认真审阅书稿并提出宝贵修改意见的蒲义书副教授吴传绪副教授以及参与校订、整理书稿的曹明涛和赵庆美同志表示衷心的谢意。

限于编者水平，错误在所难免，望读者指正。

编　者 1984 年 2 月

# 目　　录

# 第一章　多级系统的最佳解

## §1.1　分级决策方法

为了说明分级决策的概念，让我们先分析一个“求最短路径”的例子。图 1-1 是单向的从 $P$ 市通往 $Q$ 市的路径图。图中 $A$、$B$、$C$、$D$、$E$、$F$ 都是市镇的名称。

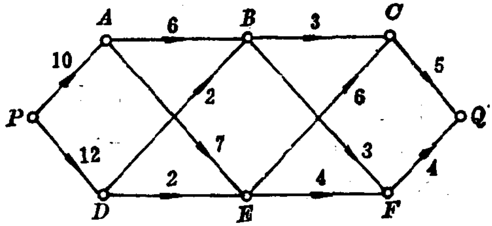


图 1-1

路径上的数字则是该段路径的长度。现在的问题是要求出从 $P$ 市到 $Q$ 市的最短路径。

从图 1-1 可以直观地看到，从 $P$ 市到 $Q$ 市的可行路径共有八条。为了清晰地说明这些可行路径，可以把图 1-1 改画为图 1-2，这样就很容易得到问题的最佳解（即最佳路径）。在图 1-2 中，对八条可行路径进行比较可知，从 $P$市到 $Q$ 市的最短路径应为 $P \to D \to B \to F \to Q$，其总长度为 21。

通常把这样一种方法，即列举出所有的可行方案，逐个

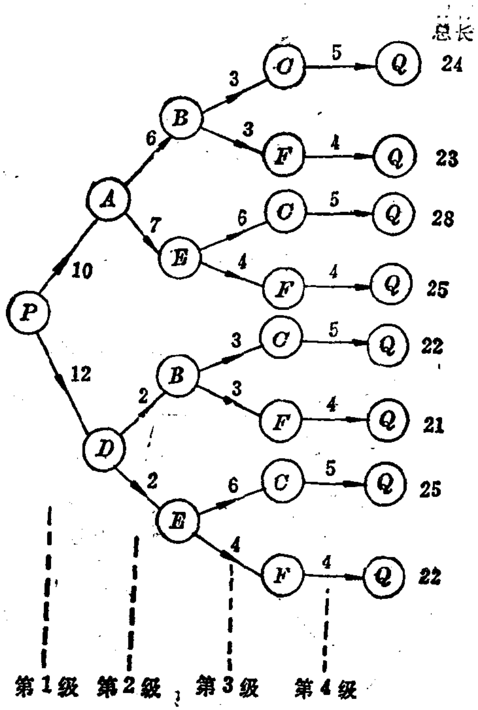


图 1-2

方案进行比较，并从中选出最佳方案的方法称为“穷举法”。这种方法非常直观，非常简单，而且所求得的结果总是全局最佳解。但是当可行解很多时，要把可行方案一一列出，而又要对它们一一进行比较以求出最佳方案来，则需要很大的计算工作量。让我们分析一下上述问题的计算工作量。在图 1-2 中，我们把从 $P$ 市到 $Q$ 市的路径中每通过一个市镇称为走过一个阶段，这样，从 $P$ 市到 $Q$ 市共需要四个阶段，我们通称这些阶段为“级”。因此，本题可以分为

四级，如图 1-2 所示。从图上也可以看到，在前三级从任一市镇出发时，总有两条路可供选择，而第四级的市镇（$C$ 和 $F$）各自通向 $Q$ 的道路只有一条。因此总的可行方案是 $2^3=8$ 个。当然从八个可行方案中求取最佳方案，其所需的计算工作量并不很大（仅需 24 次加法运算和 7 次比较运算），但当级数（$N$）增大，或者从一个市镇出发的可供选择的道路数（$M$）增加时，则可行方案数（$M^{(N-1)}$）也将大大增加，使计算工作量大为增加。

下面以一个较为复杂的问题为例来加以说明。

图 1-3 是一个较复杂的道路网络，求从 $A$ 到 $B$ 的最短路径。当采用"穷举法"时，可列出的可行方案共有 20 个，读者可以自行作出图 1-3 的分析图，并求出其最短路径为 $A\to C\to F\to J\to M\to O\to B$，全长为 $1+4+2+2+2+2=13$。当采用"穷举法"求解上述问题时，由于该问题可以分为 6 级（$N=6$），因此加法运算次数为 $20\times5=100$ 次，比较运算次数为 19 次，其计算工作量已远较前例为大。当 $N$ 或 $M$ 的值进一步增大时，由于加法和比较运算的次数都大大增多而使"穷举法"难以实际应用。

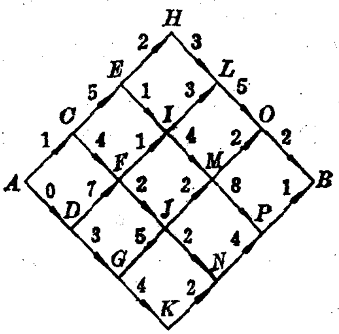

图 1-3

但是"穷举法"为分级决策方法提供了一个基础。我们仍以图 1-1 的最短路径问题为例，把决定最短路径问题分成四个阶段（级）来进行考虑，按一般通用的符号画成图 1-4 所示图形。图中 $x_i$ 称为状态变量，$d_i$ 称为决策变量，$J_i$ 称

为级收益。

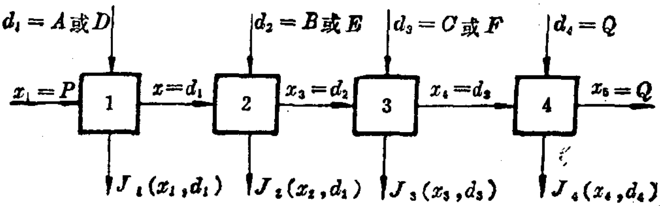


图 1-4

方框中的编号表示级的编号。对照图 1-1 来分析，第 1 级表示从 $P$ 走向 $A$ 或 $D$ 的过程。取第 1 级的输入状态变量 $x_1=P$，表示这一级的过程是从 $P$ 市开始走的。在这一级中，人们可以走向 $A$ 市，也可以走向 $D$ 市。为了使从 $P$ 市到 $Q$ 市的总距离为最短，这就需要在走向 $A$ 或走向 $D$ 这两个方案中进行决策。因此在图上写为 $d_1=A$ 或 $D$。而第 1 级的级收益则是这一过程中走过的路径长度，它是输入状态变量 $x_1$ 和决策变量 $d_1$ 的函数。显然，当 $d_1=A$ 时，即从 $P$ 出发后决定走向 $A$ 时，则级收益为 $J_1(P, A)=10$。当 $d_1=D$ 时，即从 $P$ 出发后决定走向 $D$ 时，则级收益为 $J_1(P,D)=12$。

由于我们待求的最短路径是指从 $P$ 到 $Q$ 的总路径长度为最短，因此单考虑第 1 级并不能确定最佳决策，必须考虑到后面所有各级的影响，才能正确地得出最佳决策值。

因此我们在第 1 级分析中，尚不能确定其决策 $d_1$ 是 $A$ 或是 $D$，也就是说在第二个阶段（第 2 级）的分析中，无法确定其出发点是 $A$ 还是 $D$，因此只能取第 2 级的输入状态变量为 $x_2=d_1$，即只有在 $d_1$ 确定后 $x_2$ 才能确定，这就是为什么称其为状态“变量”的缘故。从图 1-1 可知，不论第 2

级的 $x_2=A$ 或 $D$，它们可能通往的市镇总是 $B$ 或 $E$，因此取 $d_2=B$或 $E$。而第 2 级的级收益函数 $J_2(x_2, d_2)$则有四个可能值：

当 $x_2=A$，$d_2=B$（即从 $A$ 出发到达 $B$）则 $J_2(A, B)=6$。

当 $x_2=A$，$d_2=E$（即从 $A$ 出发到达 $E$）则 $J_2(A, E)=7$。

当 $x_2=D$，$d_2=B$（即从 $D$ 出发到达 $B$）则 $J_2(D, B)=2$。

当 $x_2=D$，$d_2=E$（即从 $D$ 出发到达 $E$）则 $J_2(D, E)=2$。

对于第 3 级的输入状态变量，应取 $x_3=d_2$，也就是第三阶段的出发点取决于第二阶段的决策值，但不论其出发点为 $B$ 或 $E$，可能通往的市镇则是 $C$ 或 $F$，相应的级收益函数有以下四种可能值

$J_3(B, C)=3$，　　　　$J_3(B, F)=3$，

$J_3(E, C)=6$，　　　　$J_3(E, F)=4$。

同理，第 4 级的输入状态变量，应取 $x_4=d_3$。也就是第四阶段的出发点取决于第三阶段的决策值。不论其出发点为 $C$ 或 $F$，可能通往的市镇只有 $Q$市，因此对第 4 级来说，只能确定决策 $d_4=Q$，或者说最后一级没有决策问题。该级的级收益函数只有以下两个可能的值

$J_4(C,Q)=5$，　　$J_4(F,Q)=4$。

为了求解该问题的最佳方案——从 $P$ 到 $Q$ 的最短路径，我们可以分级求出其最佳决策。因此称为“分级决策方法”。该问题的求解过程如下：

对于求解第1级输入状态变量（$x_1$）为已知（$x_1=\tilde{P}$）的这类问题（数学上称为初值问题），计算程序总是从最后一级开始进行。

第一步：从第 4 级开始计算。通往 $Q$ 市的路径共有两条，当

$x_4=C$ 时 $J_4(C,Q)=5$。

$x_4=F$ 时 $J_4(F,Q)=4$。

第二步：分析第 3 级。其输入状态变量 $x_3=B$ 或 $E$，而决策变量 $d_3=C$ 或 $F$。因此通往 $Q$ 市的路径就有四种方案：

（1）$x_3=B$，$d_3=C$，则从 $B$ 到 $Q$ 的总长度为

$$J_3(B,C)+J_4(C,Q)=3+5=8,$$

（2）$x_3=B$，$d_3=F$，则从 $B$ 到 $Q$ 的总长度为

$$J_3(B,F)+J_4(F,Q)=3+4=7。$$

以上两方案都是从 $B$ 通往 $Q$ 的长度，显然路径 $B\to F\to Q$ 较路径 $B\to C\to Q$ 短。因此，在以后的计算中，只要路径通往 $B$ 点，必将选择 $B\to F\to Q$ 路径到达 $Q$ 市，而不宜再选择 $B\to C\to Q$ 路径了。

（3）$x_3=E$，$d_3=C$，则从 $E$ 到 $Q$ 的总长度为

$$J_3(E,C)+J_4(C,Q)=6+5=11$$

（4）$x_3=E$，$d_3=F$，则从 $E$ 到 $Q$ 的总长度为

$$J_3(E,F)+J_4(F,Q)=4+4=8$$

以上两方案都是从 $E$ 通往 $Q$ 的长度，显然路径 $E\to F\to Q$ 较路径 $E\to C\to Q$ 短。因此在以后的讨论中，只要路径通过 $E$ 点，必将选择 $E\to F\to Q$ 路径到达 $B$ 市，而不宜再选择 $E\to$ $\to C\to Q$ 路径了。

采用分级决策方法后，可以在逐级计算中，不断地淘汰

一些非最佳方案，因而可使总的计算工作量减少。

第三步：分析第 2 级，其输入状态变量 $x_2=A$ 或 $D$，而决策变量 $d_2=B$ 或 $E$。因此通往 $Q$ 市的路径也有四种方案。

（1）$x_2=A$，$d_2=B$，则从 $A$ 经过 $B$ 到 $Q$ 的总长度为

$$J_2(A,B)+[J_3(B,F)+J_4(F,Q)]=6+7=13$$

这里必须指出，由于在第二步计算中已经选出从 $B$ 到 $Q$ 的最短路径为 $B\to F\to Q$。因此第三步中计算从 $A$ 经 $B$ 到 $Q$ 的路径时即应沿着 $A\to B\to F\to Q$ 进行。因为只有这样才能保证从 $A$ 经 $B$ 到 $Q$ 的路径为最短。式中方括号中的值是从 $B$ 到 $Q$ 的最短路径长度。

（2）$x_2=A$，$d_2=E$ 则从 $A$ 经过 $E$ 到 $Q$ 的总长度为

$$J_2(A,E)+[J_3(E,F)+J_4(F,Q)]=7+8=15$$

式中方括号中的值是从 $E$ 到 $Q$ 的最短路径长度。

比较（1）、（2）结果可知，从 $A$ 市到 $Q$ 市的最短路径应为 $A\to B\to F\to Q$（其总长为13），而不是 $A\to E\to F\to Q$（其总长为 15）。因此在以后的计算中，凡是经过 $A$ 市到达 $Q$ 市的最短路径都应选择 $A\to B\to F\to Q$。

（3）$x_2=D$，$d_2=B$，则从 $D$ 经过 $B$ 到 $Q$ 的总长度为

$$J_2(D,B)+[J_3(B,F)+J_4(F,Q)]=2+7=9$$

式中方括号中的值是从 $B$ 到 $Q$ 的最短路径长度。

（4）$x_2=D$，$d_2=E$ 则从 $D$ 经过 $E$ 到 $Q$ 的总长度为

$$J_2(D,E)+[J_3(E,F)+J_4(F,Q)]=2+8=10$$

式中方括号中的值是从 $E$ 到 $Q$ 的最短路径长度。

比较（3）、（4）结果可知，从 $D$ 市到 $Q$ 市的最短路径应为 $D\to B\to F\to Q$（其总长为 9），而不是 $D\to E\to F\to Q$

（其总长为 10）。因此在以后的计算中，凡是经过 $D$ 市到达 $Q$ 市的最短路径都应选择 $D\to B\to F\to Q$。

第四步：分析第 1 级，其输入状态变量仅有一个单一值，即 $x_1=P$，而决策变量 $d_1=A$ 或 $D$。因此通往 $Q$ 市的路径只有两种方案：

（1）$x_1=P$，$d_1=A$，则从 $P$ 经 $A$ 到 $Q$ 的总长度为

$$J_1(P,A)+[J_2(A,B)+J_3(B,F)+J_4(F,Q)]$$
$$=10+13=23$$

式中方括号中的值是从 $A$ 到 $Q$ 的最短路径长度。

（2）$x_1=P$，$d_1=D$，则从 $P$ 经 $D$ 到 $Q$ 的总长度为

$$J_1(P,D)+[J_2(D,B)+J_3(B,F)+J_4(F,Q)]$$
$$=12+9=21$$

式中方括号中的值是从 $D$ 到 $Q$ 的最短路径长度。

比较（1）、（2）结果可知，从 $P$ 到 $Q$ 的最短路径应为 $P\to D\to B\to F\to Q$，其全长为 21。这就是我们求得的最佳解，而各级的决策值 $d_1=D$，$d_2=B$，$d_3=F$，$d_4=Q$ 可以写成序列形式 $[d_1, d_2, d_3, d_4]=[D, B, F, Q]$，称为该问题的“最佳决策序列”。

分析这种分级决策方法，加法运算仅需 10 次，比较运算仅需 6 次，可见计算工作量较穷举法为少。

下一章介绍的动态规划方法就是以分级决策方法为基础建立的。

## §1.2 分级决策问题的数学模型

从采用分级决策方法求解最短路径问题的启示中，可以

作出图 1-5 的模型

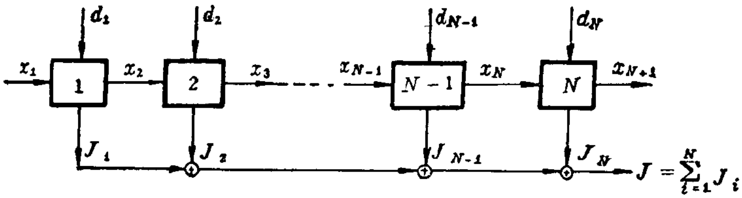


图 1-5

图中 $d_1$，$d_2$，…，$d_N$ 为 $N$ 个决策变量。一个问题中，只有那些可以直接控制、调整的变量才能作为决策变量。分级决策问题就是在一定的条件下（如 $x_1$ 为给定值），寻求一组可行决策变量 $\{d_1, d_2, \cdots, d_N\}$ 使问题的总收益 $J$ 为最佳。

图中 $x_1$，$x_2$，…，$x_N$，$x_{N+1}$ 为状态变量，可以看到它常常是一个分级的输入而又是另一个分级的输出。或者说在各级之间起着信息传递的作用。

图中 $J_i(i=1, 2, \cdots, N)$ 为第 $i$ 级的级收益。各个级收益的和为该问题的总收益，即

$$J=J_1+J_2+\cdots+J_N=\sum_{i=1}^{N} J_i$$

为了应用分级决策方法，要求各级的级收益是该级输入状态变量和决策变量的函数，即级收益函数为

$$J_i=R_i(x_i, d_i)$$

图 1-5 的模型清楚地显示出采用分级决策方法可以把一个含有 $n$ 个变量的问题转化为求解 $n$ 个单变量的问题，只要原问题满足分级条件，并按照一定的要求将问题分成 $N$ 级，这样求解 $n$ 个单变量的问题当然比求解一个 $n$ 变量的

问题要容易得多。

应用分级决策方法的条件是什么呢？它们是：状态变量的传递性和收益函数的可分性。

（一）状态变量的传递性。从图 1-5 可知，每级的输入状态变量也就是前级的输出状态变量。对于能够应用分级决策方法的问题，必须要求各级状态变量有一定的函数关系，即

$$x_{i+1}=g_i(x_i, d_i) \qquad i=1, 2, \cdots, N \tag{1-1}$$

我们称式（1-1）为状态传递方程。

根据级间状态变量的传递方程，在给定了初始状态 $x_1$ 后（称为初值问题），当已知决策 $d_1$，即可求出 $x_2$，当已知 $d_2$ 后，又可求出 $x_3$，…。因此有

$$\begin{aligned} x_{i+1}&=g_i(x_i, d_i)\\ &=g_i(g_{i-1}(x_{i-1}, d_{i-1}), d_i)\\ &=g_i(g_{i-1}(g_{i-2}(x_{i-2},d_{i-2}),d_{i-1}),d_i)\\ &=g_i(g_{i-1}(g_{i-2}(\cdots g_1(x_1,d_1),d_2),\cdots),d_i) \end{aligned} \tag{1-2}$$

$$i=1, 2, \cdots, N$$

式(1-2)从另一方面说明了级间状态变量之间的关系，只要知道了初始状态 $x_1$ 及决策序列 $\{d_1, d_2, \cdots, d_i\}$，就可以求得 $x_i$ 的值。

在分级决策问题中选择状态变量是一个较为复杂但又很关键的问题，目前并没有普遍可遵循的规则，主要依靠解题的技巧和经验。

（二）总收益函数的可分性。从图 1-5 可知，为了采用分级决策方法，必须能把总收益（J）划分成 $n$ 个级收益，在总收益函数为下列形式时，总是可以满足这一要求的，即