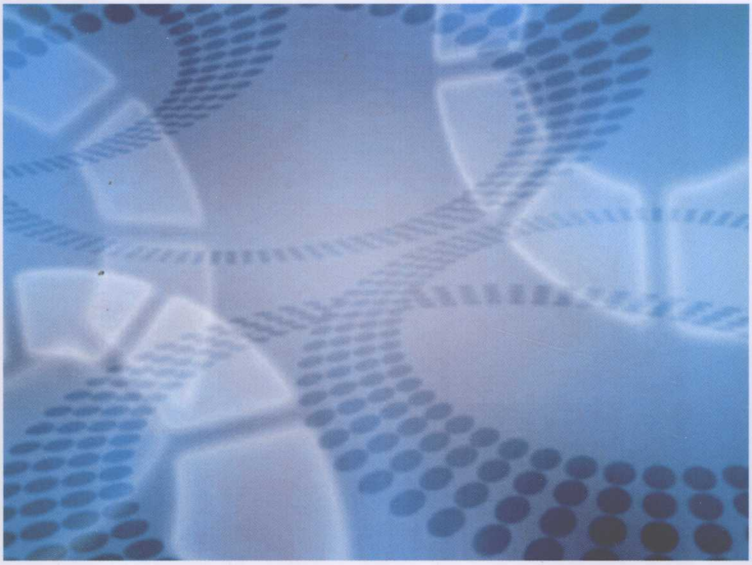
高等学校规划教材
GAODENG XUEXIAO GUIHUA JIAOCAI

运筹学通论

范玉妹　徐　尔　谢铁军　编著



冶金工业出版社
http://www.cnmip.com.cn

高等学校规划教材

运筹学通论

范玉妹　徐尔　谢铁军　编著

北京
冶金工业出版社
2009

内 容 提 要

本书以确定型数学规划模型为基础，介绍了随机型模型中的几个重要分支：动态规划、决策论、对策论、网络规划、网络计划技术、排队论，以及这些随机型数学规划模型的应用案例及计算机实现。书中主要介绍随机型模型中这些主要分支的基本概念、基本思想、基本原理和相应的数学模型；给出求解这些主要分支的主要算法，围绕主要的基本算法讨论其算法的迭代原理、迭代步骤、收敛性和优缺点等。

本书可作为工科院校研究生的教学用书，亦可供从事现代技术和管理工作的科技人员以及相关专业的实验技术人员参考。

图书在版编目(CIP)数据

运筹学通论/范玉妹，徐尔，谢铁军编著. —北京：冶金工业出版社，2009.5

高等学校规划教材

ISBN 978-7-5024-4797-7

Ⅰ.运… Ⅱ.①范… ②徐… ③谢铁军… Ⅲ.运筹学—高等学校—教材 Ⅳ.O22

中国版本图书馆 CIP 数据核字(2009)第 059165 号

出 版 人 曹胜利

地　　址 北京北河沿大街嵩祝院北巷 39 号，邮编 100009

电　　话 (010)64027926 电子信箱 postmaster@cnmip.com.cn

责任编辑 王 优 宋 良 美术编辑 李 心 版式设计 张 青

责任校对 石 静 责任印制 李玉山

ISBN 978-7-5024-4797-7

北京兴华印刷厂印刷；冶金工业出版社发行；各地新华书店经销

2009 年 5 月第 1 版，2009 年 5 月第 1 次印刷

184mm×230mm；16.25 印张；349 千字；249 页；1-2500 册

30.00 元

冶金工业出版社发行部 电话：(010)64044283 传真：(010)64027893

冶金书店 地址：北京东四西大街 46 号(100711) 电话：(010)65289081

（本书如有印装质量问题，本社发行部负责退换）

前　　言

运筹学是用定量的方法,对所研究的各类管理优化问题建立数学模型并进行求解,然后进行定量和定性的分析,为决策者做出合理的决策提供科学的依据。

从方法论的角度来看,运筹学中的数学模型大体可分为两大类:一类是确定型模型,如线性规划、非线性规划、整数规划、几何规划、图论等。这类模型在描述现实世界事物时,或由于事物本身不含随机因素,或事物本身虽含随机因素但并未扮演一个基本重要的角色,因而从数量关系上描述它们的数学模型具有确定性。关于这类确定型模型在范玉妹主编的《数学规划及其应用》(第2版)(冶金工业出版社,2003 年)中给出了详细的介绍。另一类是随机型模型,这类模型由于所描述的现实现象中随机因素扮演了一个基本重要的角色,因而从数量关系上描述它们的数学模型具有随机性。本书主要分7章介绍了随机型模型中的几个主要分支,内容包括:动态规划、决策论、对策论、排队论、网络规划、网络计划技术、应用案例及计算机实现。

本书是编者在多年讲义的基础上编写而成的。2003 年我们启动了《运筹学通论》教材的编写工作,并于2004 年10 月正式用作北京科技大学校内讲义,至今已经在全校研究生中使用了4 届,受到同行与学生的好评,2006 年被评为校优秀讲义。

在编写工作中,我们力求深入浅出,通俗易懂。在选材上,着重介绍了数学模型的基本理论和基本方法,并注意了这些理论和方法的应用;在计算方法上,着重介绍了适用面较广、使用方便、具有实效的方法;第7章给出了实用的应用案例及计算机实现的过程;为便于自学,每章后面都附有习题,书后附有答案。

本书内容简明扼要,取材新颖,内容涉及广泛,注重理论与实践,在适教性上做了有益的探索,收到了一定的实效。

本书的编写工作得到了北京科技大学研究生教育发展基金的资助,在此表示感谢。

由于编者水平所限,书中不妥之处在所难免,敬请读者批评指正。

编　者

2008 年12 月

目　　录

1 动 态 规 划

动态规划是数学规划中的一个分支，主要研究和解决多阶段决策过程的最优化问题。1951年，美国数学家 R. Bellman 等人根据一类多阶段决策问题的特性，提出了解决这类问题的"最优化原理"，并研究和解决了许多实际问题，从而创立了动态规划。

1.1 动态规划的研究对象和特点

动态规划是一种解决复杂系统优化问题的方法，是目前解决多阶段决策过程的基本理论之一。所谓**多阶段决策过程**是指这样一类决策问题：由于它的特性可将过程按时间、空间等标志分为若干个状态相互联系又相互区别的阶段。在它的每一个阶段都需要做出决策，从而使整个过程达到最优；而各个阶段的决策的选取不是任意决定的，它依赖于当前决定的状态，又给以后的发展以影响；当各个阶段决策决定后，就组成了一个决策序列，因而也就决定了整个过程的一条活动路线。这样一个前后关联具有链状结构的多阶段过程（见图 1-1）被称为**多阶段决策过程**，也称**序贯决策过程**。

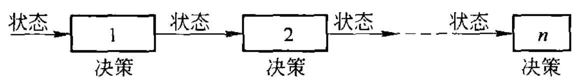


图 1-1　多阶段决策过程

将时间作为变量的决策问题称为**动态决策问题**。多阶段决策问题是一类特殊形式的动态决策问题。由于在动态决策中，决策依赖于当前的状态而又随即引起状态的转移，一个决策序列就是在状态运动变化中产生出来的，故有"动态"的含义。因此处理决策序列的方法称为**动态规划方法**。同时又由于在动态决策中，系统所处的状态和时间都是进行决策的主要因素，即需要在系统发展过程的不同时点，根据系统所处的状态不断地做出决策。因此，多次决策是动态决策的主要特点。但是动态规划也可以解决与时间无关的静态规划中的最优化问题，只要人为地引进"时间"因素把问题划分为若干阶段，也可以把静态规划的问题视为一个多阶段决策问题用动态规划的方法去处理。值得注意的是，多阶段决策过程的发展是通过状态的一系列变换转移来实现的。一般来说，系统在某个阶段的状态转移既与本阶段的状态和决策有关，还可能与系统过去经历的状态和决策有关。因此，问题的求解比较复杂。适用于用动态规划方法求解的是一类特殊的具有无后效性的多阶段决策问题。

无后效性又称为**马尔科夫性**。所谓无后效性，是指系统从某个阶段后的发展，完全是与本阶段所处的状态及其往后的决策无关。这样的问题其求解就方便得多。具有无后效性的多阶段决策过程意味着系统过程的历史只能通过系统现阶段的状态去影响系统的未来，即当前状态就是过程往后发展的初始条件。

实践已证明,动态规划在工程技术、管理、经济、工业生产、军事以及现代控制工程等领域中都有广泛的应用,并获得显著的效果。而且由于动态规划方法有其独特之处,在解决某些实际问题时显得更加方便有效。

例 1-1　最短线路问题。在线路网络图(图 1-2)中,从 A 至 F 有一批货物需要调运,图上所标数字为各节点间的距离,为使总费用最少,试求出一条自 A 到 F 总里程最短的线路。

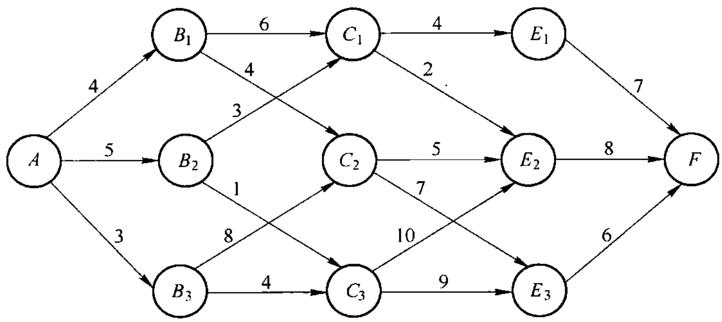


图 1-2　例 1-1 中的线路网络图

例 1-1 是个 4 阶段决策问题。

例 1-2　设飞机处于高度 H_0,并具有速度 v_0,现要使飞机的高度达到 H_n,速度达到 v_n。已知飞机保持速度不变,从任一高度 H_i 升到高度 $H_{i+1}(>H_i)$ 时的燃料费用和保持高度不变而速度从 v_i 提高到 $V_{i+1}(>V_i)$ 时的燃料费用,以及飞机速度、高度同时改变时由 (v_i,H_i) 到 (v_{i+1},H_{i+1}) 的燃料费用,试问:飞机应如何飞行,才能使总的燃料费用最低?

为简便计,将 H_n-H_0 和 v_n-v_0 分别分为 n_1 和 n_2 等分,并假定飞机每一步高度上升的增量 $\Delta H=(H_n-H_0)/n_1$,速度提高的增量为 $\Delta v=(v_n-v_0)/n_2$,而速度与高度同时改变时,每一步的增量为 $(\Delta v,\Delta H)$,相应的燃料费用的数值如图 1-3 所示。

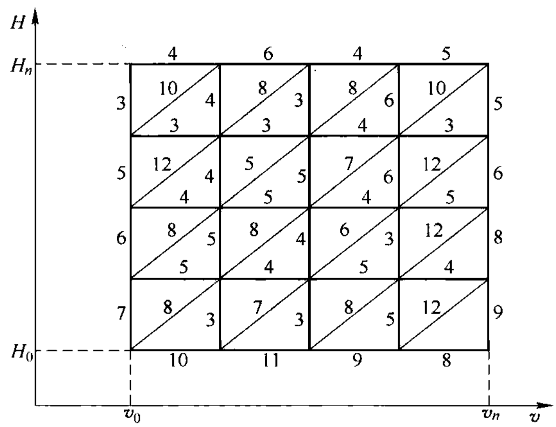


图 1-3　例 1-2 中的燃料费用数值

这里 $n_1=4$,$n_2=4$,显然这也是一个多阶段决策过程。从图 1-3 可以看出,当飞机沿不同的路线飞行时。相应过程的历程数(总的阶段数)是不同的。例如沿网格图的对角线从 (v_0,H_0) 飞到 (v_n,H_n),则过程相应的历程数 $n=4$。而先保持高度 H_n 不变,速度由 v_0 提升至 v_n,则过程相应的历程数 $n=8$,因而它是个不定期多阶段决策过程。

例 1-3　生产-存储问题。某工厂根据市场调查情况,需制订今后 4 个月的生产计划。

据估计,在这 4 个月内,市场对该产品需求量如表 1-1 所示。

表 1-1 例 1-3 中的产品市场需求量

月份(k)	1	2	3	4
需求量 $D(k)$	2	3	2	4

假定生产每批产品的固定成本费为 3000 元,每单位产品的生产成本费为 1000 元,库存费为每月 500 元,并且假定 1 月初和 4 月末均无产品库存。试求该厂如何安排各个月的生产与库存,使总成本费最小?

这显然也是一个多阶段决策问题。

此外,如各种资源(人力、物力)分配问题、背包问题、购货问题、水库调度问题等都具有多阶段决策问题的特性。一般历程为有限的多阶段决策过程,当历程是确定的数字时称为**定期**(固定)**多阶段决策过程**,当历程数不确定时则称为**不定期**(不固定)**多阶段决策过程**。例 1-1 和例 1-3 是定期(固定)多阶段决策过程,例 1-2 是不定期(不固定)多阶段决策过程。本章主要针对定期或固定多阶段决策问题,介绍一些动态规划的基本概念、理论和方法,并讨论它的基本应用。

1.2 动态规划的基本概念

1.2.1 多阶段决策过程

例 1-1 中最短路线问题是动态规划中一个较为直观的多阶段决策过程的典型例子,我们通过讨论它的解法,来说明动态规划方法的基本思想并阐述它的基本概念。

求解例 1-1 这个问题并不困难,从图 1-2 上可看出,问题有 4 次需要做出决策,即在 A 点需要决定下一步是到 B_1、B_2 还是 B_3;同样,若到 B_1 后需要决定走向 C_1 还是 C_2;若在 B_2 时需决定走向 C_1 还是 C_3。以此类推,在 C_i 及 $E_i(1\leqslant i\leqslant 3)$ 时均需要一次决策。这 4 次决策分别划分在 4 个阶段中,令 A 至 B 是第 1 阶段,B 至 C 是第 2 阶段,……。则 A 处在决策的第 1 阶段中,而 B_i、C_i、$E_i(1\leqslant i\leqslant 3)$ 处的决策分别在第 2、3、4 阶段中,在终点 F,虽然不需要作决策了,但为讨论方便,不妨称之为第 5 阶段(即 $n+1$ 阶段)。在图 1-2 中由 A 至 B_i 有 3 条不同的路线,而 $B_i(1\leqslant i\leqslant 3)$ 至 $C_j(1\leqslant j\leqslant 3)$ 及 C_j 至 $E_k(1\leqslant k\leqslant 3)$ 均有两条路线可选择,最后从 E_k 至 F 仅有一种方法可通过。由组合的方法可知从 A 至 F 共有 $3\times 2\times 2\times 1=12$ 条不同的路线。如果把这 12 条路线的距离全算出来,从中取一条最短路线,那么它就是所要找的最短路线。例 1-1 中可算出 $A\to B_2\to C_2\to E_2\to F$ 为最短,其里程为 18。这就是**穷举法**,也叫完全枚举法。它的基本思想是列举出所有可能发生的方案,再针对目的要求对其结果一一比较,求出最优方案。

显然,当网络路线比较复杂和繁多时,其计算量将增加得很快,甚至变得无法计算。

动态规划是解决多阶段决策问题最有效的方法之一。例 1-1 是一个 4 阶段决策过程。

显然，在决策过程中不能仅仅按照第1阶段的距离情况就决定由 A 到 B_1 还是到 B_2、B_3，还必须结合后继过程来考虑，这时如果对所有的后继过程都加以考虑，那就是穷举法了。第2、3阶段的情况也一样，唯最末阶段即第4阶段没有后继过程。因此，动态规划通常总是逆着决策顺序对问题进行求解的。现以 k 表示阶段数，$k=1,2,3,4$，以 U_k 表示 k 阶段的某一个决策点。可以看出，如果最短路线在第 k 阶段经过 U_k，则这一路线中由 U_k 出发到达终点的所有可能子路线来说，也必定是最短的。根据这一特征，对最短路线问题可做如下考虑：即从终点开始，依逆向求出倒数第1阶段、倒数第2阶段……倒数第 $n-1$ 阶段中各点到达终点的最短子路线，最终再正向求出从起点到终点的最短路线。为此，以 $f_k(U_k)$ 表示从第 k 阶段中点 U_k 到达终点的最短子路线的长度。

首先考虑，从 E_1、E_2、E_3 到 F 的最短子路线，由于 E_1、E_2、E_3 到 F 只有一条路可走，所以显然有：

$$f_4(E_1)=d(E_1,F)=7,f_4(E_2)=d(E_2,F)=8,f_4(E_3)=d(E_3,F)=6$$

现在要考虑从 C_1、C_2、C_3 到 F 的最短子路线，由于从 C_i 出发到 F 中间要经过 E_j，而 C_i 到 E_j 的距离为 $d(C_i,E_j)$，所以有：

$$f_3(C_i)=\min_j[d(C_i,E_j)+f_4(E_j)]$$

于是得：

$$f_3(C_1)=\min\begin{Bmatrix}d(C_1,E_1)+f_4(E_1)\\d(C_1,E_2)+f_4(E_2)\end{Bmatrix}=\min\begin{Bmatrix}4+7\\2+8\end{Bmatrix}=10$$

$$f_3(C_2)=\min\begin{Bmatrix}d(C_2,E_2)+f_4(E_2)\\d(C_2,E_3)+f_4(E_3)\end{Bmatrix}=\min\begin{Bmatrix}5+8\\7+6\end{Bmatrix}=13$$

$$f_3(C_3)=\min\begin{Bmatrix}d(C_3,E_2)+f_4(E_2)\\d(C_3,E_3)+f_4(E_3)\end{Bmatrix}=\min\begin{Bmatrix}10+8\\9+6\end{Bmatrix}=15$$

因此，从 C_1、C_2、C_3 到 F 的最短子路线分别为：$C_1\to E_2\to F$，长度为10；$C_2\to E_2\to F$ 或 $C_1\to E_3\to F$，长度为13；$C_3\to E_3\to F$，长度为15。同理，可求得：

$$f_2(B_1)=\min\begin{Bmatrix}d(B_1,C_1)+f_3(C_1)\\d(B_1,C_2)+f_3(C_2)\end{Bmatrix}=\min\begin{Bmatrix}6+10\\4+13\end{Bmatrix}=16$$

$$f_2(B_2)=\min\begin{Bmatrix}d(B_2,C_1)+f_3(C_1)\\d(B_2,E_3)+f_3(C_3)\end{Bmatrix}=\min\begin{Bmatrix}3+10\\1+15\end{Bmatrix}=13$$

$$f_2(B_3)=\min\begin{Bmatrix}d(B_3,C_2)+f_3(C_2)\\d(B_3,C_3)+f_3(C_3)\end{Bmatrix}=\min\begin{Bmatrix}8+13\\4+15\end{Bmatrix}=19$$

所以，从 B_1、B_2、B_3 到 F 的最短子线路分别为：$B_1\to C_1\to E_2\to F$，长度为16；$B_2\to C_1\to E_2\to F$，长度为13；$B_3\to C_3\to E_3\to F$，长度为19。最后求得：

$$f_1(A)=\min\begin{Bmatrix}d(A,B_1)+f_2(B_1)\\d(A,B_2)+f_2(B_2)\\d(A,B_3)+f_2(B_3)\end{Bmatrix}=\min\begin{Bmatrix}4+16\\5+13\\3+19\end{Bmatrix}=18$$

于是从 A 到 F 的最短路线为：$A \to B_2 \to C_1 \to E_2 \to F$，其长度为 18。

以上的计算过程是从后面向前推算的，这种方法称为**逆序法**（反向法）。在动态规划中一般用逆序法求解。例 1-1 也可以从前往后推算，这时称为**顺序法**（前向法）。但对于一些问题，采用逆序法比顺序法更为有效。以上的计算过程说明：动态规划方法的基本思想是把一个比较复杂的问题分解成一系列同一类型的容易求解的子问题，计算过程单一化，便于应用计算机。由于把最优化应用到每个子问题上，于是就系统地删去了若干中间非最优化的方案组合，使计算工作量比穷举法大大减少。图 1-4 是用逆序法求解例 1-1 的图形表示形式，方框中的数字为各点到终点的最短路线。图 1-5 是用顺序法求解例 1-1 的图形表示形式，此时方框中的数字为起点到各点的最短距离，双线条表示起点至各点的最短子路线。

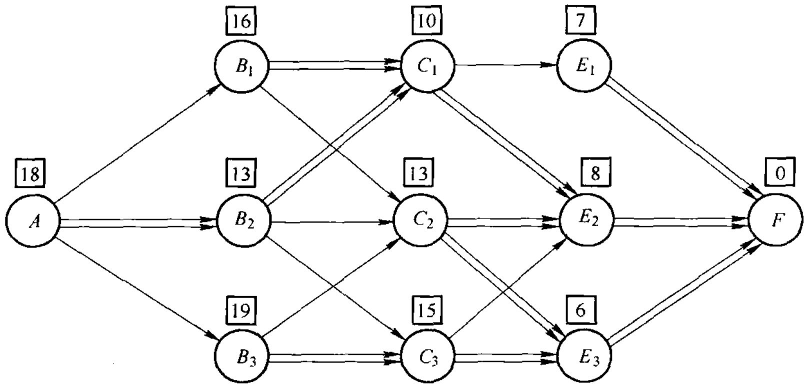


图 1-4　用逆序法求解例 1-1 的图形表示形式

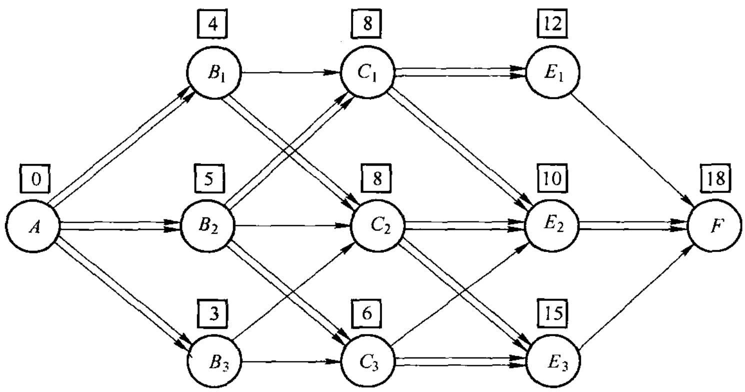


图 1-5　用顺序法求解例 1-1 的图形表示形式

动态规划求多阶段决策过程问题的最优解的特点可归纳如下：

（1）各个阶段初往后的最优决策过程与以前各个阶段所采用的决策无关，但由本阶段开始直到终端必须构成最优决策过程，这就是 Bellman 提出的最优化原理。

(2)在逐段递推过程中,每阶段选择最优决策时,不应只从本阶段的直接效果出发,而应从以后全过程的效果出发,即需要考虑两种效果:一是本阶段初到下阶段初所选决策的**直接效果**,二是由所选决策确定的下阶段初往后直到终点的所有决策过程的总效果,或称为**间接效果**,这两种效果的结合必须是最优的。

(3)经过递推计算获得各阶段有关数据后,反方向即可求出相应的最优决策过程。

1.2.2 多阶段决策过程的基本概念

动态规划中描述多阶段决策过程的基本概念有:阶段和阶段变量,状态和状态变量,决策、决策变量和决策序列,状态转移方程,阶段效应和目标函数等。现结合 1.2.1 节的例子来给出这些常用的基本概念及相应的符号。

1.2.2.1 阶段和阶段变量

在多阶段决策过程中,为了表示决策和过程的发展顺序,引入了阶段概念。一个**阶段**就是需要做出一个决策的子问题部分。通常阶段是按照决策进行的时间或空间上的先后顺序划分的,用阶段变量 k 表示。阶段数等于多段过程中从开始到结束所需要做出决策的数目。n 阶段决策过程就是从过程开始到结束需要依次进行 n 次决策。例 1-1 就是一个 4 阶段决策过程。

1.2.2.2 状态和状态变量

在多阶段决策过程中,决策是根据系统所处情况决定的。**状态**就是描述系统情况所必需的信息。描述多阶段决策过程要求引入与阶段变量相应的状态变量,状态变量必须包含在给定的阶段上确定全部允许决策所需要的信息。即为了做出本阶段的决策,不再需要考虑过去的经历和决策。按照过程进行的先后,每个阶段状态分为**初始状态**和**终止状态**,亦称**输入状态**和**输出状态**。阶段 k 的初始状态表示为 x_k,终止状态表示为 x_{k+1}。例如,阶段 1 的初始状态是 x_1,终止状态是 x_2,阶段 2 的初始状态是 x_2,终止状态是 x_3,等等。由此可见,前一阶段的终止状态又是后一阶段的初始状态。为了清楚起见,通常定义阶段状态为阶段的初始状态。在最短线路问题例 1-1 中,状态就是网络中的各个节点。例如阶段 1 的状态是 A,阶段 2 的状态是 B_1、B_2、B_3 等。

通常,状态变量的取值有一定允许集合或范围,称为**状态可能集**。状态可能集是关于状态的约束条件。状态可能集用相应阶段状态 x_k 的大写字母 X_k 表示。$x_k \in X_k$,状态可能集可以是离散取值的集合,也可以是连续的集合,视所给问题而定。在例 1-1 中 $X_1 = \{A\}$、$X_2 = \{B_1、B_2、B_3\}$、$X_3 = \{C_1、C_2、C_3\}$等。

1.2.2.3 决策、决策变量和决策序列

决策就是确定系统过程发展的方案。多阶段决策过程的发展是用各阶段的状态演变来描述的。因此,决策的实质是关于状态的选择。阶段决策就是决策者从本阶段状态出发对下一阶段状态所做出的选择。因为用状态描述的过程具有无后效性,因此在进行阶段决策时,显然只需根据当前的状态而无需考虑过去的历史。在阶段 k,如果给出了决策变量 u_k

是状态变量 x_k 的函数,称为**决策函数**,记为 $u_k(x_k)$。

和状态变量一样,决策变量的取值也有一定的允许范围,称为**允许决策集合**。允许决策集合是决策的约束条件,u_k 的允许集合记为 U_k,$u_k \in U_k$。U_k 要根据相应的状态可能集 X_k 结合具体问题来确定。

决策序列也叫**策略、政策**。策略有全过程策略和 k 子策略之分。全过程策略是整个 n 段决策过程中依次进行的 n 个阶段决策构成的决策序列,简称**策略**,表示为$\{u_1,u_2,\cdots,u_n\}$,从阶段 k 到阶段 n 依次进行的阶段决策构成的决策序列称为 ***k* 子策略**,表示为:$\{u_k,u_{k+1},\cdots,u_n\}$,显然 $k=1$ 时的 k 子策略就是全过程策略。

在 n 段决策问题中,各阶段的状态可能集和决策允许集确定了策略的允许范围。特别是过程的初始状态不同,决策和策略也就不同,即策略是初始状态的函数。

1.2.2.4 状态转移方程

系统在阶段 k 处于状态 x_k,执行决策 $u_k(x_k)$的结果是系统状态的转移,即系统由阶段 k 的初始状态 x_k 转移到终止状态 x_{k+1},亦即由阶段 k 的状态 x_k 转移到阶段 $k+1$ 的状态 x_{k+1}。多阶段决策过程的发展就是这样用阶段状态的相继演变来描述的。

对于具有无后效性的多阶段决策过程,系统由阶段 k 到阶段 $k+1$ 的状态转移方程为

$$x_{k+1}=T_k(x_k,u_k(x_k)) \tag{1-1}$$

就是说,阶段 $k+1$ 的状态 x_{k+1}完全由阶段 k 的状态 x_k 和决策 $u_k(x_k)$确定,与系统过去的状态 $x_1,x_2,\cdots,x_{k-1}$及其决策 $u_1(x_1),u_2(x_2),\cdots,u_{k-1}(x_{k-1})$无关。例如在例 1-1 中,阶段 $k+1$ 的状态(地点)完全由阶段 k 的决策(地点选择)确定,与阶段 k 以前的情况无关。$T_k(x_k,u_k(x_k))$称为**变换函数**或**变换算子**。变换函数可分为两类,即确定型和随机型,据此形成确定型动态规划和随机型动态规划。

1.2.2.5 阶段效益和目标函数

多阶段决策过程中,在阶段 k 的状态执行决策 u_k,不仅带来系统状态的转移,而且也必然要对决策目的所指的效益函数(目标函数)给予影响。**效益函数**(又称**目标函数**)是用来衡量所实现过程优劣的一种数量指标,它是定义过程上的数量函数,用 R_k 表示:

$$R_k=R(x_k,u_k,x_{k+1},u_{k+1},\cdots,x_n,u_n)$$

当初始状态给定时,过程的策略也随之而确定,因而目标函数是初始状态和策略的函数。

阶段效益(又称**阶段指标**)是衡量该段决策效果的数量指标,它是执行阶段策略时所带来的目标函数的增量,用 $r_k(x_k,u_k)$来表示。

目标函数 R_k 的最优值,称为**最优目标函数值**,记为$f_k(x_k)$,它表示从第 k 阶段状态 x_k 出发到过程结束时所获得的最优目标函数值。在不同问题中,函数值的含义是不同的,可能是距离、利润、资源消耗等。

在具有无后效性的多阶段决策过程中,阶段效益完全由阶段 k 的状态 x_k 和决策 u_k 决定,与阶段 k 以前的状态无关。多阶段决策过程关于目标函数的总效应是由各阶段的阶段

效应累积形成的。适用于动态规划求解的问题的目标函数，必须具有关于阶段效应的可分离形式。k 子过程的目标函数可以表示为：

$$R_k = R(x_k, u_k, x_{k+1}, u_{k+1}, \cdots, x_n, u_n) = r_k(x_k, u_k) \odot r_{k+1}(x_{k+1}, u_{k+1}) \odot \cdots \odot r_n(x_n, u_n)$$

其中 $\odot$ 表示某种运算，可以是加、减、乘、除等。经营管理工程中最常见的目标函数是取各阶段效应之和的形式，即：

$$R_k = R(x_k, u_k, \cdots, x_n, u_n) = \sum_{i=k}^{n} r_i(x_i, u_i) \tag{1-2}$$

例 1-1 中的目标函数是从 A 到 F 的路程，为各阶段路程的和，就具有这种形式。

1.2.2.6　多阶段决策过程的数学模型

综上所述，今后要讨论和求解的问题，即具有无后效性的多阶段决策过程问题的数学模型为如下形式：

$$\begin{aligned} \underset{u_k \sim u_n}{\text{opt}}\ R = r_k(x_k, u_k) \odot r_{k+1}(x_{k+1}, u_{k+1}) \odot \cdots \odot r_n(x_n, u_n) \\ \text{s. t.}\quad & x_{k+1} = T_k(x_k, u_k) \\ & x_k \in X_k \\ & u_k \in U_k \\ & k = 1, \cdots, n \end{aligned} \tag{1-3}$$

式中，opt 是最优化之意，根据具体问题要求而取 max 或 min。所谓求解多阶段决策问题就是求出：

(1) 最优策略或最优决策序列：$\{u_1^*, u_2^*, \cdots, u_n^*\}$；

(2) 最优路线，即执行最优策略时的状态序列：$\{x_1^*, x_2^*, \cdots, x_{n+1}{}^*\}$，其中 $x_{k+1}{}^* = T_k(x_k{}^*, u_k{}^*)$，$k = 1, \cdots, n$；

(3) 最优目标函数值：$R^* = r_k(x_k^*, u_k^*) \odot r_{k+1}(x_{k+1}{}^*, u_{k+1}{}^*) \odot \cdots \odot r_n(x_n^*, u_n^*)$。

1.2.3　建立动态规划模型的基本条件

在明确了动态规划的基本概念和方法后，要把一个实际问题用动态规划的方法求解，首先要构造动态规划的数学模型，而建立动态规划模型时，除了要将问题恰当地划分成若干阶段外，还必须注意以下几点：

(1) 正确选择状态变量 x_k。动态规划中的状态应具有以下三个特性：

1) 要能够用来描述受控过程的演变特征；

2) 要满足无后效性；

3) 可知性，即规定的各阶段状态变量的值由直接或间接都是可以知道的。

(2) 确定决策变量 u_k 及在各阶段允许决策集合 $U_k(x_k)$。

(3) 写出状态转移函数。根据阶段的划分和阶段间演变的规律，写出状态转移函数。

(4)根据题意,列出目标函数关系。一般只有在目标函数满足递推性时才能用动态规划方法来求解。

1.2.4　动态规划的分类

动态规划与线性规划不同,它不存在一种标准的数学形式,因而可以说动态规划方法是一种求解问题的手段。由于多阶段决策过程的特点不同,动态规划的表现形式也就不同,据此可将动态规划做以下分类:

(1)按决策的特性分:可分为时间多段决策过程和空间多段决策过程。

(2)按允许决策集合的连续或不连续分:可分为连续多段决策过程和离散多段决策过程。

(3)按构成决策序列的决策数目分:可分为有限多段决策过程和无限多段决策过程。

(4)按状态变换的确定或随机分:可分为确定性多段决策过程和随机性多段决策过程。

(5)按决策序列与时间起点的关系分:可分为定常(与时间起点无关)多阶段决策过程和非定常多阶段决策过程。

实际的多段决策问题常常归结为各种复合的情况。今后我们只限于讨论定常的、确定性的、有限的多段决策问题。

1.3　动态规划的基本方程

1.3.1　Bellman 函数

为了便于应用最优性原理,建立动态规划基本方程,需要定义如下的辅助函数,即条件最优目标函数,亦称 Bellman 函数$f_k(x_k)$:在阶段 k 从初始状态 x_k 出发,执行最优决策序列或策略,到达过程终点的目标函数取值。

对于目标函数是阶段效益之和的多阶段决策过程而言,有:

$$f_k(x_k) = \operatorname*{opt}_{u_k \sim u_n} \sum_{i=k}^{n} r_i(x_i, u_i) \quad k = 1,2,\cdots,n \tag{1-4}$$

为了区别多段决策过程的任意阶段状态 x_k 的最优策略和最终的最优策略,将前者称为**条件最优策略**,意即处于条件 x_k 时的最优策略。构成条件最优策略的决策称为**条件最优决策**。阶段 k 处于状态 x_k 的条件最优决策表示为 $u'_k(x_k)$,简记为 u'_k,相应的条件最优策略表示为:$\{u'_k, u'_{k+1}, \cdots, u'_n\}$。显而易见,条件最优目标函数值$f_k(x_k)$意即执行条件最优策略时的目标函数值,因此,

$$f_k(x_k) = \sum_{i=k}^{n} r_i(x_i, u'_i) \tag{1-5}$$

执行条件最优策略时的阶段状态序列称为**条件最优路线**,表示为:

$$\{x'_k, x'_{k+1}, \cdots, x'_n\},\text{其中 } x'_{k+1} = T_k(x_k, u'_k), k = 1,2,\cdots,n$$