宁波工程学院资助

动态规划原理及应用

DONGTAI GUIHUA
YUANLI JI YINGYONG

滕　宇　梁方楚◎主编

西南交通大学出版社
Http://press.swjtu.edu.cn

动态规划

原理及应用

DONGTAI GUIHUA
YUANLI JI YINGYONG

滕　宇　梁方楚　主编

图书在版编目（CIP）数据

动态规划原理及应用 / 滕宇，梁方楚主编. —成都：西南交通大学出版社，2011.12
ISBN 978-7-5643-1523-8

Ⅰ. ①动… Ⅱ. ①滕… ②梁… Ⅲ. ①对偶规划 Ⅳ. ①0221

中国版本图书馆 CIP 数据核字（2011）第 260000 号

动态规划原理及应用

滕 宇 梁方楚 主编

责任编辑	张宝华
封面设计	墨创文化
出版发行	西南交通大学出版社 （成都二环路北一段 111 号）
发行部电话	028-87600564 028-87600533
邮政编码	610031
网　　址	http: //press.swjtu.edu.cn
印　　刷	四川森林印务有限责任公司
成品尺寸	148 mm×210 mm
印　　张	5.812 5
字　　数	155 千字
版　　次	2011 年 12 月第 1 版
印　　次	2011 年 12 月第 1 次
书　　号	ISBN 978-7-5643-1523-8
定　　价	18.00 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. 它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，大大简化了求解过程. 而且动态规划方法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他最优化方法难以达到的.

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，并没有统一的算法模型. 为了使读者掌握动态规划方法的基本原理和演算技巧，本书采用了案例引入式的编写方式，并例举了工程技术和经营管理等方面的大量例题，供读者参考.

本书作为动态规划方法的入门教材，考虑到读者对象为一般工程技术人员、经营管理人员和普通大专院校学生等，因此内容力求浅显易懂，在数学上只要具备一般高等数学知识，即可顺利阅读，掌握运用动态规划方法解决问题的一般过程.

在本书的编写过程中，作者得到了荆广珠教授和刘龙章教授的指导和帮助，在此表示衷心感谢.

由于作者水平有限，书中难免存在不足和错误，望读者批评、指正。

作　者

2011 年 9 月

目　录

绪　　论……1

第一章　动态规划的基本概念和基本原理……17

第一节　最短行军路线问题及标号法……17
第二节　动态规划的术语……20
第三节　动态规划的基本方程……27
第四节　动态规划的基本定理和最优化原理……35
第五节　可逆过程及顺序解法……39
习题一……41

第二章　不定期动态规划与无期动态规划……45

第一节　不定期最优路线问题……45
第二节　函数迭代法……48
第三节　策略迭代法……56
第四节　平稳不定期动态规划……66
第五节　无期动态规划……69
习题二……73

第三章　多维动态规划……76

第一节　一维分配问题……76
第二节　二维火力分配问题……81
第三节　多维分配问题……90
习题三……94

第四章　随机动态规划……96

第一节　随机过程……96

第二节 状态概率 …………………………………………………………100
第三节 序贯决策过程 ……………………………………………………105
第四节 何瓦德策略迭代法 ………………………………………………112
习题四 ………………………………………………………………………121

第五章 连续型动态规划 ……………………………………………126

第一节 火箭运行控制问题 ………………………………………………126
第二节 连续型动态规划的最优化原理 …………………………………128
第三节 应用举例 …………………………………………………………132
习题五 ………………………………………………………………………137

第六章 动态规划的应用 ……………………………………………139

第一节 搜索力的最优分配问题 …………………………………………139
第二节 复合系统的可靠性问题 …………………………………………144
第三节 生产计划问题 ……………………………………………………150
第四节 优选组队问题 ……………………………………………………155
第五节 防洪系统联合运行问题 …………………………………………159
第六节 风电投资决策问题 ………………………………………………163
习题六 ………………………………………………………………………167

参考文献 ……………………………………………………………………178

绪　论

一、动态规划的起源与发展

在现实生活中，有一类特殊的活动过程，它可将整个过程分成若干个互相联系的阶段，在它的每一阶段都需要做出决策，从而使整个过程达到最好的活动效果. 因此各个阶段决策的选取不能任意确定，它依赖于当前面临的状态，又影响以后的发展. 当各个阶段决策确定后，就组成一个决策序列，因而也就确定了整个过程的一条活动路线. 这种把一个问题看做是一个前后关联且具有链状结构的多阶段过程就称为多阶段决策过程，这种问题称为多阶段决策优化问题. 在每个阶段中，都要求出本阶段的各个初始状态到过程终点的最短路径和最短距离，当逆序倒推到过程起点时，便得到了全过程的最短路径及最短距离，同时附带得到了一组最优结果，这就是动态规划方法的基本思想. 概括来说就是将待求解问题分解成若干个子问题，先求解子问题，然后从这些子问题的解得到原问题的解. 20 世纪 50 年代初美国数学家 R. E. Bellman 等在研究多阶段决策过程 (multistep decision process) 的优化问题时，提出了著名的最优化原理 (principle of optimality)，把多阶段过程转化为一系列单阶段问题逐个求解，创立了解决这类过程优化问题的新方法 —— 动态规划.

在理查德等人提出动态规划的最优化原理之后，1956 年，C. Pontryagin 提出了最优控制的极大值原理. 1957 年 R. E. Bellman 在美国普林斯顿大学发表了第一本正式的有关动态规划问题的著作. 他首先提出了用目标函数指标来设计控制系统的思想，并解决了非线性和时变系统的设计问题. 1969 年，Merton 在完全市场中股票价

格过程服从扩散过程、股票无红利、投资者也无非资本利得、效用函数为常数、相对风险厌恶、常数绝对风险厌恶等严格条件下，将动态规划方法运用于最优投资与消费选择策略的求解，给出了连续时间下两类资产的最优投资与消费问题的解决办法. 这是最早将动态规划方法运用到最优投资与消费问题的求解，以后，许多学者都运用了此方法. 1973 年，Johnson 等把动态规划方法和模拟技术结合起来使用，确定了联合运用系统的工程规模并取得了成功. 1974 年，HuPpe 采用动态规划方法来规划气田的生产. 1982 年，曾赛星等采用动态规划方法确定内蒙古河套灌区各种作物的灌水定额及灌水次数. 1988 年，黄强把模糊动态规划方法用于求解水电站水库长期优化调度问题，较随机动态规划法简便，计算速度更快. 1989 年，曾赛星等针对内蒙古河套灌区永联试区的具体情况，运用大系统分解协调方法建立了灌区优化灌溉制度及地面水、地下水联合运用的谱系模型，模型中第一层子系统优化采用动态规划方法确定各种作物的灌溉制度. 1991 年，林学钛等在对河南平顶山市地表水与地下水的联合管理研究中，运用动态规划方法对白龟山水库进行了优化调度.

此外，动态规划还是信息学竞赛中选手必须熟练掌握的一种算法，它以其多元性广受出题者的喜爱. 动态规划首次进入信息学奥赛是在 IOI94（数学三角形），在国内首次出现是在 NOI95. 此后，动态规划成为信息学奥赛的必考算法之一.

动态规划自问世以来，在经济管理、生产调度、工程技术和最优控制等方面得到了广泛应用. 例如，最短路线、库存管理、资源分配、设备更新、排序、装载等问题，用动态规划方法比用其他方法求解更为方便. 作为运筹学的一个重要分支，动态规划的发展伴随着它的广泛应用将不断臻善.

二、动态规划与一些算法的比较

动态规划作为诸多算法中的一种，必然和其他一些算法有着诸多联系. 从这些联系中，也可以看出动态规划的一些优缺点.

1. 动态规划与分治

在分治法中，为了解决一个大的问题，总是想方设法把它分成两个或多个更小的问题，然后分别解决每个小问题，再把各小问题的解答组合起来，即可得到原问题的解答. 小问题通常与原问题本质相似，只是规模小些，一般都可以用递归的方法来解决，如汉诺塔问题和快速排序都是应用这种方法的典型例子. 然而常常有些问题难以用分治法来求解. 例如，在把问题分成若干个子问题求解时，如果子问题的数目太多，当把每个子问题再分解时，必将得到更多的子问题，以至于最后所设计的算法需要耗费指数时间. 其实，在这类问题中，子问题的数目只是出现了多个多项式，而且在上述算法中，一定有子问题被重复计算了许多次，如果能够保存已解决的子问题的答案，而在需要时简单查一下，就可以避免大量的重复计算，从而得到多项式时间的算法. 为了实现上述方法，可以用一个数组来记录所有已解决的子问题的答案，无论子问题以后是否被用到，只要它被计算过，就将其结果存入数组中，这就是动态规划. 动态规划算法可以很明显地减少重复计算，以至于在实现的时候完全可以用递推的手段来实现，因此可以说动态规划是减少重复计算的快速算法.

2. 动态规划与递推

动态规划算法也是通过递推来实现其最优解，致使很多人，甚至一些资料书上都往往将递推算法当做动态规划. 实际上，这两种算法还是很容易区分的，下面用一个例子来说明递推法和动态规划的关系.

例 1　模 4 最优路径问题. 在图 1 中找出从第 1 点到第 4 点的一条路径，要求路径长度模 4 的余数最小.

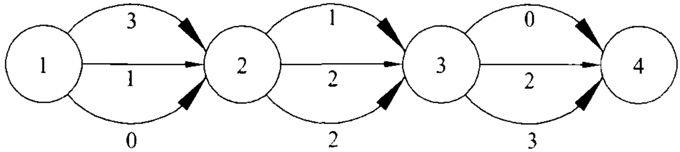


图 1　模 4 最优路径问题

这个图结构是一个多段图，而且是一个特殊的多段图. 虽然这个图的形式比一般的多段图要简单，但是这个最优路径问题却不能用动态规划来求解. 因为一条从第 1 点到第 4 点的最优路径，在它走到第 2 点、第 3 点时，路径长度模 4 的余数不一定是最小，也就是说最优策略的子策略不一定最优 —— 这个问题不满足最优化原理. 但是可以把它转换一下，用递推法来解决，即判断从第 1 点到第 k 点的长度模 4 为 x_k 的路径是否存在，用 $f_k(x_k)$ 来表示，则递推公式如下：

$$f_k(x_k)=\begin{cases} f_{k-1}((x_k-x_{k,1}) \bmod 4) \\ f_{k-1}((x_k-x_{k,2}) \bmod 4) \\ f_{k-1}((x_k-x_{k,3}) \bmod 4) \end{cases}$$

边界条件为

$$f_1(x_1)=\text{true}$$

其中 $x_{k,i}$ 表示从 $k-1$ 点到第 k 点之间的第 i 条边的长度，最后的结果就是可以使 $f_4(x_4)$ 值为真的最小的 x_4 值.

这个递推法的递推公式和动态规划的规划方程非常相似，在这里借用了动态规划的符号也就是为了更清楚地显示这一点. 其实它们的思想也是非常相像的，可以说是递推法借用了动态规划的思想解决了动态规划不能解决的问题.

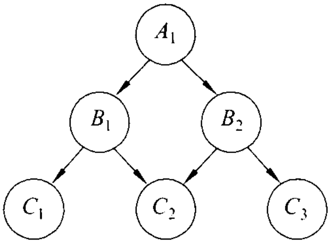
3. 动态规划与搜索

同样是解决最优化问题，有些问题采用动态规划，而有些问题则需要用搜索方法. 撇开时间和空间效率的因素不谈，在解决最优化问题的算法中，搜索可以说是“万能”的，所以如果动态规划可以解决的问题，搜索也一定可以解决.

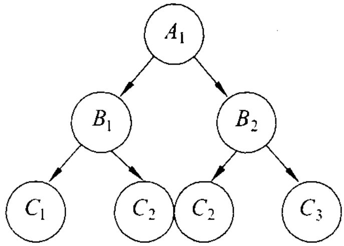
把一个动态规划算法改写成搜索是非常方便的，状态转移方程、规划方程以及边界条件都可以直接“移植”，所不同的只是求解顺序：动态规划是自底向上的递推求解，而搜索则是自顶向下的递归

求解（这里指深度优先搜索，宽度优先搜索类似）. 反过来，也可以把搜索算法改写成动态规划. 状态空间搜索实际上是对图结构中的点进行枚举，这种枚举是自顶向下的. 如果把枚举的顺序反过来，改成自底向上，那么就成了动态规划.

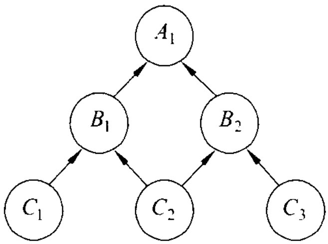
正因为动态规划和搜索有着求解顺序上的不同，也就造成了它们时间效率上的差别，在搜索中，往往会出现图 2 所示的情况.



（a）图结构



（b）搜索算法实现



（c）动态规划实现

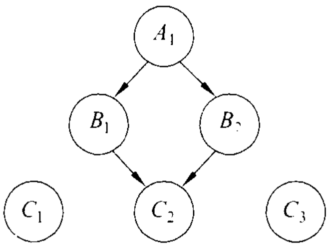
图 2　图结构及实现一

对于图 2（a）这样几个状态构成的一个图结构，用搜索算法就会出现重复，如图 2（b）所示，状态 C_2 被搜索了两次. 在深度优先搜索中，这样的重复会引起以 C_2 为根的整个子搜索树的重复搜索；在宽度优先搜索中，虽然这样的重复可以立即被排除，但是其时间代价也是不小的. 而动态规划就不会出现这个问题，如图 2（c）所示.

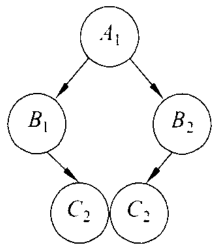
一般来说，动态规划算法在时间效率上的优势是搜索算法无法比拟的，而从理论上讲，任何拓扑有序的图结构中的搜索算法都可

以改写成动态规划. 但事实上，在很多情况下仍然不得不采用搜索算法. 那么，动态规划算法在实现上还有什么障碍吗?

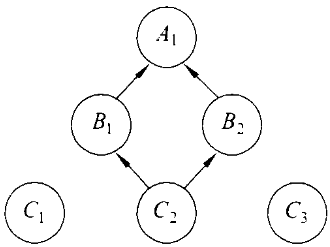
考虑如图 3（a）所示的图结构，其中存在两个从初始状态无法达到的状态. 在搜索算法中，这样的两个状态就不被考虑了，如图 3（b）所示. 但是动态规划由于是自底向上求解，所以就无法估计到这一点，因而遍历了全部的状态，如图 3（c）所示. 动态规划总要遍历所有的状态，而搜索可以排除一些无效状态. 更重要的是搜索还可以剪枝，可能剪去大量不必要的状态，因此在空间开销上往往比动态规划要低很多.



（a）图结构



（b）搜索算法实现



（c）动态规划实现

图 3 图结构及实现二

如何协调好动态规划的高效率与高消费之间的矛盾呢? 有一种折中的办法就是记忆化算法. 记忆化算法在求解的时候还是按着自顶向下的顺序，但是每求解一个状态，就将它的解保存下来，以后再次遇到这个状态的时候，就不必重新求解了. 这种方法综合了搜索和动态规划两方面的优点，因而还是很有实用价值的.

通过上面的分析，我们可以看出动态规划的优点主要表现在以下几个方面:

首先，动态规划方法是一种逐步改善法，它把问题化成一串彼此不相关、结构相似的最优化问题，而这些子问题又化成一串彼此不相关、结构相似的最优化问题，这些子问题的变量个数比原问题少得多，约束集合也简单得多，所以较易于确定其全局最优解. 特别地，当每一个阶段的状态变量个数不多于三个，而每一个允许决策集合不太大时，对每一个问题至少可用全区直接搜索（穷举法）求其全局最优解，然后再用动态规划的递推程序求原问题的解. 基于这一点，目前有相当多的最优化问题，动态规划是求出其全局最优解的唯一方法. 此外，由于子问题可用穷举法求其解，利用动态规划方法时，和一般分析方法相反，对状态变量和决策变量限制越多，子问题的搜索区域就越小，解起来就越容易. 出于同一原因，对于其指标、状态转移和允许决策集合不能用分析形式表出的过程最优化问题，用分析的方法求最优化解几乎成为不可能，而用动态规划方法却很容易.

其次，齐次动态规划方法求出的不仅是对整个过程的某一特定状态的一个解，而且是对所有后部子过程的所有可能出现状态的一族解. 在某些情况下，当这些解族正是实际问题所需要时，用动态规划方法就会比用其他方法节省更大的计算量，即使只求某一个特定解. 这样求出的解族，在分析最优指标值和最优决策序列对于状态变量的稳定性时是非常有用的. 在随机多阶段决策过程中，用策略还是决策序列表示最优解有很大的区别，而策略总是优于决策序列的，所以随机性过程中应用动态规划方法具有更大的优越性.

最后，动态规划方法反映了过程逐段演变的前后联系，较之非线性规划，它与过程的实际特征联系得更紧密，因而在计算中可以更有效地利用实践的经验. 这样做往往能提高求解的效率，特别是在后面的策略迭代法中，初始策略的选取对迭代的收敛速度常常有很大的影响，这时，利用实践经验往往帮助我们选取“好”

的初始决策. 此外，在有的问题中不一定要知道最优解的数值，而更关心最优解的结构，或对于过程的某些参数的依赖关系，或过程历程无限增长时它的渐近性质，这时利用动态规划方程分析就比较方便.

任何思想方法都有一定的局限性，超出了特定条件，它就失去了作用. 同样，动态规划也并不是万能的，它的缺点主要有:

第一，到目前为止，动态规划还没有一个统一的标准模型可供使用. 实际问题不同，其动态规划模型可能各异，虽然理论上说可以把其他数学规划问题化为动态规划模型求解，但是这种转化过程对于复杂的数学规划问题将变得十分困难.

第二，构造动态规划模型时，状态变量必须满足最优化原理和无后效性，这是一个相当强的条件. 无后效性是指将各阶段按照一定的次序排列好之后，对于某个给定的阶段状态，它以前各阶段的状态无法直接影响它未来的决策，而只能通过当前的这个状态. 换句话说，每个状态都是过去历史的一个完整总结. 但是许多实际问题不仅依赖于状态转移规律，还与允许决策集合和指标的结构有关. 当取自然特征作为状态变量时，往往不满足无后效性的条件，减低了动态规划的通用性. 当然，原则上说，采用适当增多状态变量的办法，总能人为地把过程变为无后效的，但动态规划还存在下述局限，所以这种作法的实际意义不大.

第三，用动态规划方法进行数值求解时的主要问题是所谓“维数障碍”. 若状态变量大于 2 或 3，计算时将涉及较大的存储量和计算量. 状态空间维数越高，即描述状态空间的向量分量越多，所遇到的计算困难将变得越大.

尽管动态规划存在一些缺点，但是通过采取一些特殊的方法，可以部分地克服某些应用上的障碍. 作为一种算法，它通过将一个 n 维变量的复杂问题进行分阶段处理，把 n 维变量问题变成求解 n 个单变量问题，大大简化了求解过程，节省了巨大的计算量，这是经典的求解极值方法所做不到的.

三、动态规划的特点

动态规划是求解最优化问题的一种途径、一种方法，而不是一种特殊的算法. 不像前面所述的那些搜索或数值计算那样，具有一个标准的数学表达式和明确清晰的解题方法. 动态规划程序设计往往是针对一种最优化问题，由于各种问题的性质不同，确定最优解的条件也互不相同，因而动态规划的设计方法对不同的问题，有各具特色的解题方法，而不存在一种万能的动态规划算法，可以解决各类最优化问题. 正是由于这个原因，使动态规划算法具有模式性、多样性和技巧性等特点.

1. 动态规划的模式性

动态规划设计有着一定的模式，一般要经历以下几个步骤：

划分阶段：按照问题的时间或空间特征，把问题分为若干个阶段. *注意：这若干个阶段一定是有序的或者是可排序的，否则问题就无法求解.*

选择状态：将问题发展到各个阶段时所处于的各种客观情况用不同的状态表示出来. 当然，状态的选择要满足无后效性.

确定决策并写出状态转移方程：之所以把这两步放在一起，是因为决策和状态转移有着天然的联系，状态转移就是根据上一阶段的状态和决策来导出本阶段的状态. 所以，如果确定了决策，状态转移方程也就写出来了. 但事实上，常常是反过来做，根据相邻两段的各状态之间的关系来确定决策.

写出规划方程（包括边界条件）：根据问题的优化要求，可以直接写出动态规划方程. 一般来说，只要阶段、状态、决策和状态转移确定了，就可以较容易的写出规划方程.

掌握了动态规划的模式性，在用动态规划考虑问题时就可以把主要精力放在理论的设计上面. 一旦设计成熟，问题也就基本解决了，而且在设计算法时也可以按部就班地进行.

2. 动态规划的多样性

对同一个优化问题，由于状态和决策选择的不同，可以建立不同的动态规划模型，下面举例说明.

例 2 花店橱窗布置问题：假设你想以最美观的方式布置花店的橱窗，现在有 F 束不同品种的花束，同时也有至少相同数量的花瓶已按顺序摆成一行. 这些花瓶的位置固定于架子上，并按 1 至 M 顺序编号. M 是花瓶的数目，从左至右排列，则最左边的是花瓶 1，最右边的是花瓶 M. 花束可以移动，并且每束花用 1 至 N 的整数唯一标志. 标志花束的整数决定了花束在花瓶中的顺序，如果 $I<J$，则令花束 I 必须放在花束 J 左边的花瓶中.

假设一束杜鹃花的标志数为 1，一束秋海棠的标志数为 2，一束康乃馨的标志数为 3，所有的花束在放入花瓶时必须保持其标志数的顺序，即杜鹃花必须放在秋海棠左边的花瓶中，秋海棠必须放在康乃馨左边的花瓶中. 如果花瓶的数目大于花束的数目，则多余的花瓶必须空置，且每个花瓶中只能放一束花.

每一个花瓶都有各自的特点，因此，当各个花瓶中放入不同的花束时，会产生不同的美学效果，并以美学值（一个整数）来表示，空置花瓶的美学值为零.

在上述例子中，花瓶与花束的不同搭配具有不同的美学值，如表 1 所示.

表 1 “花店橱窗布置问题”实例数据

花束＼花瓶	1	2	3	4	5
1（杜鹃花）	7	23	−5	−24	16
2（秋海棠）	5	21	−4	10	23
3（康乃馨）	−21	5	−4	−20	20

根据表 1，杜鹃花放在花瓶 2 中，会显得非常好看，但若放在花瓶 4 中则显得十分难看. 为取得最佳美学效果，你必须在保持花

束顺序的前提下，使花束的摆放取得最大的美学值. 如果有不止一种的摆放方式具有最大的美学值，则其中任何一种摆放方式都可以接受，但只要输出任意一种摆放方式即可.

方法 1: 以花束的数目来划分阶段. 在这里，阶段变量 k 表示要布置的花束数目(前 k 束花)，状态变量 x_k 表示第 k 束花所在的花瓶. 而对于每一个状态 x_k，决策就是第 $k-1$ 束花应该放在哪个花瓶，用 u_k 表示. 最优指标函数 $f_k(x_k)$ 表示前 k 束花，其中第 k 束插在第 x_k 个花瓶中所能取得的最大美学值.

状态转移方程为

$$x_{k-1}=u_k$$

规划方程为

$$f_k(x_k)=\max_{k\leqslant u_k<x_k}(f_{k-1}(u_k)+A(k,x_k))$$

其中 $A(i,j)$ 是花束 i 插在花瓶 j 中的美学值.

边界条件为

$$f_0(x_0)=0,\quad 0\leqslant x_0\leqslant M$$

方法 2: 以花瓶的数目来划分阶段. 在这里，阶段变量 k 表示要占用的花瓶数目(前 k 个花瓶)，状态变量 x_k 表示前 k 个花瓶中放了多少花. 而对于任意一个状态 x_k，决策就是第 x_k 束花是否放在第 k 个花瓶中，用变量 $u_k=1$ 或 0 来表示. 最优指标函数 $f_k(x_k)$ 表示前 k 个花瓶中插了 x_k 束花所能取得的最大美学值.

状态转移方程为

$$x_{k-1}=x_k-u_k$$

规划方程为

$$f_k(x_k)=\max_{u_k=0,1}(f_{k-1}(x_k-u_k)+u_k\cdot A(x_k,k))$$

边界条件为