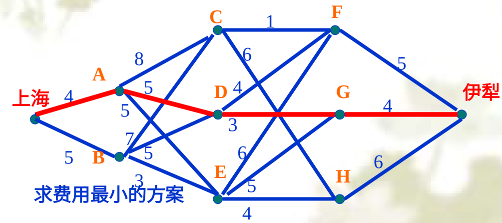
6. 动态规划

引例1： 最短路问题

从上海到伊犁间有一个铁路网,某学生暑假打算到伊犁旅游,出于经济关系只能坐火车,而且要求费用最少。下图标出了各种可能的行车路线,请为这位同学的决策做出指导。



求费用最小的方案

穷举法所有路径的费用?

→多阶段决策问题！

例2 生产计划问题

工厂生产某种产品, 每单位(千件)的成本为1(千元), 每次开工的固定成本为3(千元), 工厂每季度的最大生产能力为6(千件). 经调查, 市场对该产品的需求量第一、二、三、四季度分别为 2, 3, 2, 4(千件). 如果工厂在第一、二季度将全年的需求都生产出来, 自然可以降低成本(少付固定成本费), 但是对于第三、四季度才能上市的产品需付存储费, 每季每千件的存储费为0.5(千元). 还规定年初和年末这种产品均无库存. 试制订一个生产计划, 即安排每个季度的产量, 使一年的总费用(生产成本和存储费)最少.

该规划问题看似一个线性规划或线性整数混合规划问题, 但实际上由于问题的动态特性, 是不容易建立常规的规划问题的(非常复杂)! 但可以分阶段进行决策,即多阶段决策问题. 需要引入新的方法——动态规划方法.

动态规划的基本概念

动态规划(*dynamic programming*)是运筹学的一个分支,是求解决策过程最优化的数学方法。20世纪50年代初美国数学家R.E.Bellman等人在研究多阶段决策过程的优化问题时,提出了著名的最优化原理,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。

多阶段决策过程,是指这样的一类特殊的活动过程,问题可以按时间顺序或空间位置分解成若干相互联系的阶段,在每一个阶段都要做出决策,全部过程的决策是一个决策序列。要使整个活动的总体效果达到最优的问题,称为多阶段决策问题。

动态规划问世以来,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的应用。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它方法求解更为方便。

虽然动态规划主要用于求解以时间划分阶段的动态过程的优化问题,但是一些与时间无关的静态规划(如线性规划、非线性规划),只要人为地引进时间因素,把它视为多阶段决策过程,也可以用动态规划方法方便地求解。

动态规划模型的基本要素

(1) 阶段(Step)——为了表示决策和过程的发展顺序,通常根据决策进行的先后次序、时间顺序或空间特征,将全过程来划分为若干阶段。阶段变量一般用k=1,2,..,n表示。

在例1中由“上海”出发为第一阶段(k=1),由A、B 出发的为第二阶段(k=2),依此下去可得n=4个阶段。在例2中按照第一、二、三、四季度分为k=1,2,3,4,共4个阶段。

(2) 状态(State)——表示每个阶段开始时过程所处的自然状况。它应该能够描述过程的特征并且具有无后向性,即当某阶段的状态给定时,这个阶段以后过程的演变与该阶段以前各阶段的状态无关,即每个状态都是过去历史的一个完整总结。通常还要求状态是直接或间接可以观测的。

描述状态的变量称状态变量。变量允许取值的范围称允许状态集合。用x_k表示第k阶段的状态变量,它可以是一个数或一个向量。用X_k表示第k阶段的允许状态集合。在例1中x_2可取A,B,X_2={A,B}。n个阶段的决策过程有n+1个状态变量,x_{n+1}表示x_n演变的结果,在例1中x_5取目的地“伊犁”。

根据过程演变的具体情况,状态变量可以是离散的或连续的。为了计算的方便有时将连续变量离散化；为了分析的方便有时又将离散变量视为连续的。状态变量简称为状态。

动态规划模型的基本要素(续)

(3)决策(Decision)——当一个阶段的状态确定后,可以作出各种选择从而演变到下一阶段的某个状态,这种选择手段称为决策(decision),在最优控制问题中也称为控制(control)。

描述决策的变量称决策变量. 变量允许取值的范围称允许决策集合.用$u_k(x_k)$表示第k阶段处于状态x_k时的决策变量, 它是x_k的函数, 用$U_k(x_k)$表示了x_k的允许决策集合. 在例1中$u_2(B)$可取C, D, E. 决策变量简称决策.

(4) 策略(policy)——决策组成的序列称为策略. 由初始状态x_1开始的全过程的策略记作$p_{1,n}(x_1)$, 即

$$p_{1,n}(x_1)=\{u_1(x_1),u_2(x_2),\ldots,u_n(x_n)\}.$$

由第k阶段的状态x_k开始到终止状态的后部子过程的策略记作$p_{k,n}(x_k)$, 即

$$p_{k,n}(x_k)=\{u_k(x_k),u_{k+1}(x_{k+1}),\ldots, u_n(x_n)\}.$$

类似地, 由第k到第j阶段的子过程的策略记作

$$p_{k,j}(x_k)=\{u_k(x_k),u_{k+1}(x_{k+1}),\ldots, u_j(x_j)\}.$$

对于每一个阶段k的某一给定的状态x_k. 可供选择的策略$p_{k,j}(x_k)$有一定的范围,称为允许策略集合,用$P_{1,n}(x_1)$,$P_{k,n}(x_k)$,$P_{k,j}(x_k)$表示。

从允许策略集合中找出使问题达到最优效果的策略称为最优策略.

动态规划模型的基本要素(续)

(5) 状态转移方程(equation of state transition)——确定过程由一个状态到下一个状态演变规律的方程. 在确定性过程中, 一旦某阶段的状态和决策为已知, 下阶段的状态便完全确定. 用状态转移方程表示这种演变规律,写作：

$$x_{k+1} = T_k\left(x_k, u_k(x_k)\right), k = 1,2,\cdots,n$$

在例1中状态转移方程为：$x_{k+1}=u_k(x_k)$

(6)指标函数(objective function) ——是衡量过程优劣的数量指标, 它是关于策略的数量函数, 从阶段k到阶段n的指标函数用$V_{k,n}(x_k,p_{k,n}(x_k))$表示, k=1,2,...,n. 能够用动态规划解决的问题的指标函数应具有可分离性, 即$V_{k,n}$可表为x_k, u_k, $V_{k+1,n}$ 的函数,记为：

$$V_{k,n}\left(x_k, p_{k,n}(x_k)\right) = \varphi_k\left[x_k, u_k, V_{k+1,n}\left(x_{k+1}, p_{k+1,n}(x_{k+1})\right)\right]$$

$$即\ V_{k,n} = \varphi_k\left(x_k, u_k, V_{k+1,n}\right)$$

其中 φ_k 是关于变量 $V_{k+1,n}$ 的单调递增函数 .

同时必须注意到 指标函数$V_{k,n}$是状态x_k和策略$p_{k,n}(x_k)$的函数!

动态规划模型的基本要素(续)

(7) 阶段指标(阶段收益)——系统某一阶段的状态一经确定, 执行某一决策所得到的效益或描述该阶段决策的数量指标, 称为阶段指标(阶段收益). 它是整个系统总收益的一部分,是阶段状态和决策变量的函数, 记为

$$v_k(x_k, u_k(x_k))$$

在例1中阶段指标为走完一段路程的花费;

一般地, 指标函数$V_{k,n}(x_k,p_{k,n}(x_k))$是由阶段指标$v_k$(j=k,k+1,..n)组成的, 就常见的形式有:

阶段指标v_k(j=k,k+1,..n)之和:

$$V_{k,n}\left(x_k, p_{k,n}(x_k)\right) = \sum_{j=k}^{n} v_j(x_j, u_j)$$

阶段指标v_k(j=k,k+1,..n)之积:

$$V_{k,n}\left(x_k, p_{k,n}(x_k)\right) = \prod_{j=k}^{n} v_j(x_j, u_j)$$

阶段指标v_k(j=k,k+1,..n) 的极大(或极小):

$$V_{k,n}\left(x_k, p_{k,n}(x_k)\right) = \operatorname*{opt}_{k\le j\le n} v_j(x_j, u_j)$$

动态规划模型的基本要素(续)

(8)最优值函数(optimal value function) ——在x_k给定时指标函数$V_{k,n}$对$p_{k,n}$的最优值称为最优值函数,记作$f_k(x_k)$,即:

$$f_k(x_k) = \operatorname*{opt}_{p_{k,n}(x_k)\in P_{k,n}(x_k)} V_{k,n}\left(x_k, p_{k,n}(x_k)\right)$$

其中opt可根据具体情况取max或min。上式的意义是,对于某个阶段k的某个状态x_k,从该阶段k到最终目标阶段n的最优指标函数值等于从x_k出发取遍所有能策略$p_{k,n}$所得到的指标值中最优的一个。

利用指标函数的性质有:

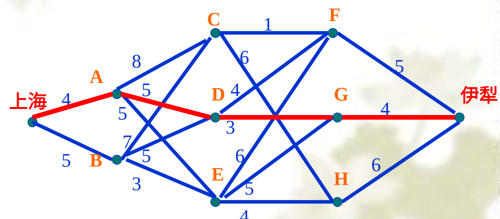
$$V_{k,n}\left(x_k, p_{k,n}(x_k)\right) = \varphi_k\left[x_k, u_k, V_{k+1,n}\left(x_{k+1}, p_{k+1,n}(x_{k+1})\right)\right]$$

$$\Rightarrow f_k(x_k) = \operatorname*{opt}_{p_{k,n}\in P_{k,n}} \varphi_k\left(x_k, u_k, V_{k+1,n}\left(x_{k+1}, p_{k+1,n}(x_{k+1})\right)\right)$$

$$\Rightarrow f_k(x_k) = \operatorname*{opt}_{p_{k,n}\in P_{k,n}} \varphi_k\left(x_k, u_k, f_{k+1}(x_{k+1})\right)$$

动态规划模型的基本要素(续)

(9) 最优策略(optimal policy)和最优轨线(optimal trajectory)——使指标函数$V_{k,n}$达到最优值的策略是从k开始的后部子过程的最优策略,记作$p_{k,n}^*=\{u_k^*,\dots, u_n^*\}$, $p_{1,n}^*$又是全过程的最优策略,简称最优策略。从初始状态$x_1(=x_1^*)$出发,过程按照$p_{1,n}^*$和状态转移方程演变所经历的状态序列$\{x_1^*,x_2^*,\dots, x_{n+1}^*\}$称最优轨线。



动态规划的基本定理

最优性定理——设阶段数为n的多阶段决策过程,其阶段编号为k=1,2,...,n, 对于初始状态$x_1\in X_1$,策略$p_{1,n}^*=\{u_1^*,..u_n^*\}$是最优策略的充要条件是对于任意的k,1<k≤n,有

$$V_{1,n}\left(x_1, p_{1,n}^*\right) = \operatorname*{opt}_{p_{1,k-1}\in P_{1,k-1}(x_1)} \varphi\left[V_{1,k-1}\left(x_1, p_{1,k-1}\right), \operatorname*{opt}_{p_{k,n}\in P_{k,n1}(x_k)} V_{k,n}\left(x_k, p_{k,n}\right)\right]$$

最优化原理(是上定理的推论) ——若$p_{1,n}^*\in P_{1,n}(x_1)$是最优策略, 则对于任意的k,1<k<n, 它的子策略p_{kn}^*对于由x_1和$p_{1,k-1}^*$确定的以x_k^*为起点的第k到n后部子过程而言,也是最优策略。

最优化原理给出了最优策略的必要条件, 通常略述为:不论过去的状态和决策如何, 对于前面的决策形成的当前的状态而言,余下的各个决策必定构成最优策略.

动态规划的基本方程(逆推方程)

$$\begin{cases} f_k(x_k) = \operatorname*{opt}_{u_k\in U_k} \varphi\left(x_k, u_k, f_{k+1}(x_{k+1})\right) \\ x_{k+1} = T_k\left(x_k, u_k(x_k)\right), k = 1,2,\cdots,n \\ f_{n+1}(x_{n+1}) = \delta(x_{n+1}) \ \text{初值} \end{cases}$$

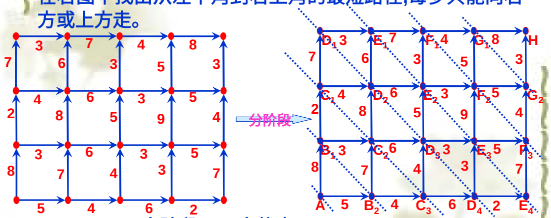
$f_{n+1}(x_{n+1})$是决策过程的终端条件, 往往为一个已知数. 当x_{n+1}只取固定的状态时称固定终端; 当x_{n+1}可在终端集合X_{n+1}中变动时称自由终端.

动态规划的基本方程(顺推方程)

$$\begin{cases} f_k(x_{k+1}) = \operatorname*{opt}_{u_k\in U_k} \varphi\left(x_{k+1}, u_k, f_{k-1}(x_k)\right) \\ x_k = T_k\left(x_{k+1}, u_k(x_k)\right), k = 2,\cdots,n \\ f_1(x_2) = \operatorname*{opt}_{u_1\in U_1} \varphi(x_1, u_1) \end{cases}$$

举例——街道问题

在右图中找出从左下角到右上角的最短路径,每步只能向右方或上方走。



7个阶段、20个状态

决策, 策略, 转移方程, 指标函数, 阶段指标

→最优值函数, 最优策略

求解: $A \to B_2 \to C_3 \to D_3 \to E_3 \to F_3 \to G_2 \to H$

动态规划算法的基本步骤

设计一个标准的动态规划算法,通常可按以下几个步骤进行:

1)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要是有序的或者是可排序的(即无后向性), 否则问题就无法用动态规划求解.

2) 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.

3) 确定决策并写出状态转移方程：之所以把这两步放在一起,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,如果我们确定了决策, 状态转移方程也就写出来了. 但事实上,我们常常是反过来做, 根据相邻两段的各状态之间的关系来确定决策.

4)写出规划方程(包括边界条件):动态规划的基本方程是规划方程的通用形式化表达式.一般说来,只要阶段、状态、决策和状态转移确定了,这一步还是比较简单的。

例2 生产计划问题(求解)

工厂生产某种产品, 每单位(千件)的成本为1(千元), 每次开工的固定成本为3(千元), 工厂每季度的最大生产能力为6(千件). 经调查, 市场对该产品的需求量四个季度分别为 2, 3, 2, 4(千件). 每季每千件的存储费为0.5(千元). 还规定年初和年末这种产品均无库存. 试制订一个生产计划, 使一年的总费用最少.

求解——看成多阶段问题: 我们按照计划时间自然划分n(=4)个阶段, 状态定义为每阶段开始时的存储量x_k, 决策为每个阶段的产量u_k, 记每个阶段的需求量(已知)为d_k，则状态转移方程为：

$$x_{k+1} = x_k + u_k - d_k, (x_k \ge 0), k = 1,2,\cdots,n(=4)$$

设每个阶段开工固定成本费用为a(=3),生产单位数量产品的成本为b(=1), 每阶段单位数量产品的存储费用为c(=0.5), 阶段指标为阶段的生产成本费用和存储费用之和, 即:

$$v_k(x_k,u_k) = cx_k + \begin{cases} a + bu_k, u_k > 0 \\ 0, \quad u_k = 0 \end{cases}$$

指标函数$V_{k,n}$为v_k之和, 最优值函数$f_k(x_k)$为从第k阶段的状态x_k出发到过程终结的最小费用, 则基本逆推方程为:

$$f_k(x_k) = \min_{u_k \in U_k} [v_k(x_k,u_k) + f_{k+1}(x_{k+1})] \quad (f_{n+1}(x_{n+1}) = 0)$$

$$x_{k+1} = x_k + u_k - d_k, \quad u_k \in \{0,1,2,3,4,5,6\}$$

$$v_k(x_k,u_k) = 0.5x_k + \begin{cases} 3 + u_k, u_k > 0 \\ 0, \quad u_k = 0 \end{cases}$$

基本逆推方程为:

$$f_k(x_k) = \min_{u_k \in U_k} [v_k(x_k,u_k) + f_{k+1}(x_{k+1})]$$

$$f_5(x_5) = 0$$

$$x_5 = 0 \Rightarrow x_4 \in \{0,1,2,3,4\}$$

x_4	0	1	2	3	4
u_4^*					
$f_4(x_4)$					

$$x_3 \in \{0,1,2,3,4,5,6\}$$

x_3	0	1	2	3	4	5	6
u_3^*							
$f_3(x_3)$							

$$x_1 = 0, d_1 = 2 \Rightarrow x_2 \in \{0,1,2,3,4\}$$

x_2	0	1	2	3	4
u_2^*					
$f_2(x_2)$					

$$x_1 = 0$$

x_1	0
u_1^*	
$f_1(x_1)$	

$u_1^* = 5$

$u_2^* = 0$

$u_3^* = 6$

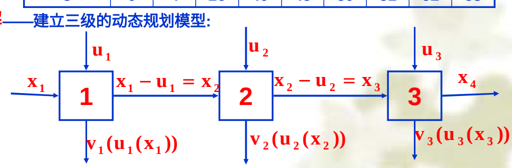
$u_4^* = 0$

例3 资源分配

研究一个人力资源分配方案。某建筑公司有三个工区都需要增加人力资源，对各工区增加人力后，其收益也相应增加，如表所示。请寻求人力资源的最佳分配方案，使全公司增加的总收益最大。

工区 \ 人力	0	1	2	3	4	5	6	7	8
1	0	5	15	40	80	90	95	98	100
2	0	5	15	40	60	70	73	74	75
3	0	4	26	40	45	50	51	52	53

解——建立三级的动态规划模型:



$x_k \in \{0,1,2,3,4,5,6,7,8\} \quad u_k \in \{0,1,2,3,4,5,6,7,8\}$

例4 背包问题

求解如下0-1背包问题

$$\max \quad 7y_1 + 16y_2 + 19y_3 + 15y_4$$

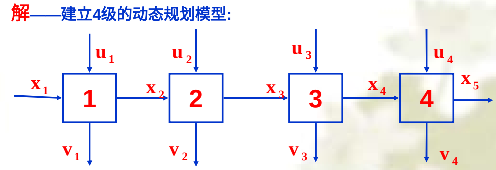
$$3y_1 + 6y_2 + 7y_3 + 5y_4 \le 9$$

$$y_i \in \{0,1\}$$

计算价值/体积比 $\frac{7}{3}, \frac{16}{6}, \frac{19}{7}, \frac{15}{5}$ 　　22　非最优解

近似求解方法　1　0　0　1

解——建立4级的动态规划模型:



$x_k \in \{0,1,2,3,4,5,6,7,8,9\} \quad u_k \in \{0,1\}$