动态规划

- 1951年美国数学家（R.Bellman）提出
- 同时提出解决的方法“最优化原理”
- 应用范围十分广泛。在企业管理方面：最优路径问题，资源分配问题，生产调度问题，库存问题，装载问题，排序问题，设备更新问题
- 优点：通常比线性规划和非线性规划更有效。特别是离散性问题。
- 缺点：没有标准的数学表达式，没有统一的处理方法。若变数的维数太大，则无法求解；最致命缺点是最优化原理没有经过理论证明，其适用范围还不清楚。

动态规划

本章内容重点

多阶段决策过程的最优化

动态规划的基本概念和基本原理

动态规划方法的基本步骤

动态规划方法应用举例

1. 多阶段决策过程的最优化

一、多阶段决策问题

动态规划是把多阶段决策问题作为研究对象。

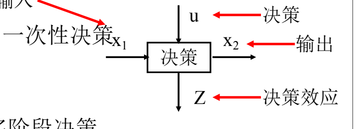
所谓多阶段决策问题，是指这样一类活动过程，即根据问题本身的特点，可以将其求解的全过程划分为若干个相互联系的阶段(即将问题划分为许多个相互联系的子问题)，

在它的每一阶段都需要作出决策；

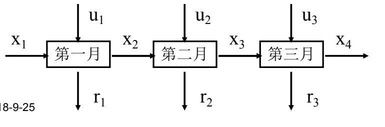
并且在一个阶段的决策确定以后再转移到下一个阶段。

动态规划的概念与模型

- 静态决策　一次性决策

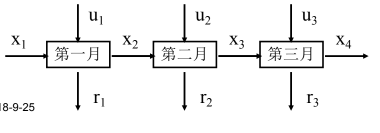


- 动态决策　多阶段决策



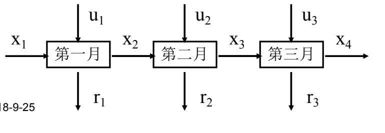
多阶段决策过程的最优化

往往前一个阶段的决策要影响到后一个阶段的决策，从而影响整个过程。人们把这样的决策过程称做多阶段决策过程(Multi-Stage decision process)。



多阶段决策过程的最优化

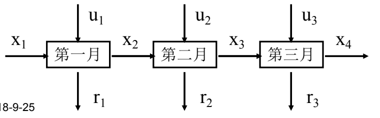
各个阶段所确定的决策就构成了一个决策序列，称为一个策略。一般来说，由于每一阶段可供选择的决策往往不止一个，因此，对于整个过程，就会有许多可供选择的策略。



多阶段决策过程的最优化

若对应于一个策略，可以由一个量化的指标来确定这个策略所对应的活动过程的效果，那么，不同的策略就有各自的效果。

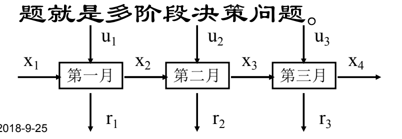
$U=(u_1, u_2 \ldots\ldots u_n)$
$R=r_1+r_2\ldots\ldots+r_n$



多阶段决策过程的最优化

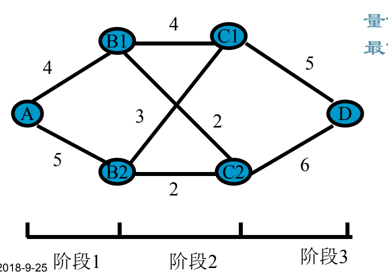
在所有可供选择的策略中，对应效果最好的策略称为最优策略。

把一个问题划分成若干个相互联系的阶段选取其最优策略，这类问题就是多阶段决策问题。



例如.从A运送物资到D

划分阶段
单阶段决策
策略
量化指标
最优策略



1. 多阶段决策过程的最优化

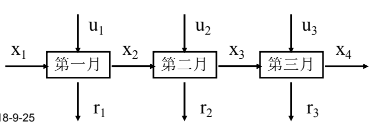
多阶段决策过程最优化的目标是要达到整个活动过程的总体效果最优。

不应仅考虑本阶段最优，还应考虑对最终目标的影响，从而作出对全局来讲是最优的决策。

动态规划就是符合这种要求的一种决策方法。

多阶段决策问题举例

1) 工厂生产过程：由于市场需求是一随着时间而变化的因素，因此，为了取得全年最佳经济效益，就要在全年的生产过程中，逐月或者逐季度地根据库存和需求情况决定生产计划安排。



1. 多阶段决策过程的最优化

2) 设备更新问题：一般企业用于生产活动的设备，刚买来时故障少，经济效益高，即使进行转让，处理价值也高，随着使用年限的增加，就会逐渐变为故障多，维修费用增加，可正常使用的工时减少，加工质量下降，经济效益差，并且，使用的年限越长、处理价值也越低，自然，如果卖去旧的买新的，还需要付出更新费．因此就需要综合权衡决定设备的使用年限，使总的经济效益最好。

1. 多阶段决策过程的最优化

3) 连续生产过程的控制问题： 一般化工生产过程中，常包含一系列完成生产过程的设备，前一工序设备的输出则是后一工序设备的输入，因此，应该如何根据各工序的运行工况，控制生产过程中各设备的输入和输出，以使总产量最大。

1. 多阶段决策过程的最优化

以上所举问题的发展过程都与**时间因素**有关，因此在这类多阶段决策问题中，阶段的划分常取时间区段来表示，并且各个阶段上的决策往往也与时间因素有关，这就使它具有了“动态”的含义，所以把处理这类动态问题的方法称为**动态规划方法**。

实际中尚有许多不包含时间因素的一类“**静态**”决策问题，就其本质而言是一次决策问题，是非动态决策问题，但是也可以人为地引入阶段的概念当作多阶段决策问题，应用动态规划方法加以解决。

1. 多阶段决策过程的最优化

4）**资源分配问题：** 便属于这类静态问题。如：某工业部门或公司，拟对其所属企业进行稀缺资源分配，为此需要制定出收益最大的资源分配方案。这种问题原本要求一次确定出对各企业的资源分配量，它与时间因素无关，不属动态决策，但是，我们可以人为地规定一个资源分配的阶段和顺序，从而使其变成一个多阶段决策问题。

1. 多阶段决策过程的最优化

5）**运输网络问题：** 如图所示的运输网络，点间连线上的数字表示两地距离（也可是运费、时间等），要求从 v_1 至 v_{10} 的最短路线。

这种运输网络问题也是静态决策问题。但是，按照网络中点的分布，可以把它分为4个阶段，而作为多阶段决策问题来研究。

1. 多阶段决策过程的最优化

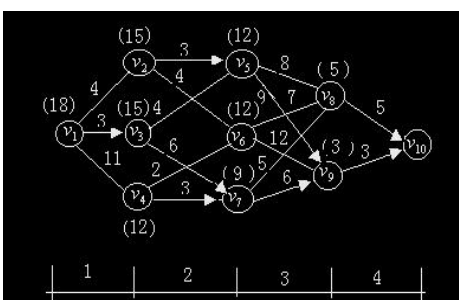


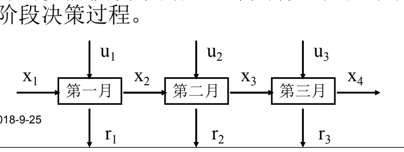
图1 运输网络图示

动态规划求解的多阶段决策问题的特点

通常多阶段决策过程的发展是通过状态的一系列变换来实现的。

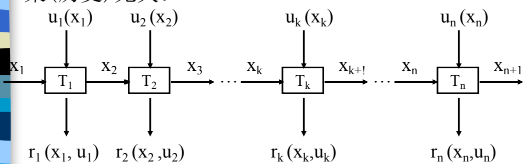
一般情况下，系统在某个阶段的状态转移除与本阶段的状态和决策有关外，还可能与系统过去经历的状态和决策有关。因此，问题的求解就比较困难复杂。

而适合于用动态规划方法求解的只是一类特殊的多阶段决策问题，即具有“无后效性”的多阶段决策过程。



具有无后效性的多段决策过程

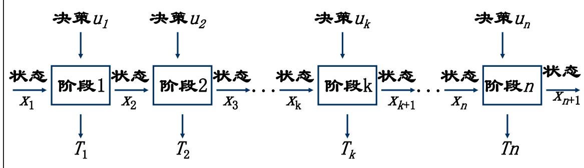
所谓无后效性，又称马尔柯夫性，是指系统从某个阶段往后的发展，仅由本阶段所处的状态及其往后的决策所决定，与系统以前经历的状态和决策(历史)无关。



$X_{k+1}=T_k(x_k, u_k)$

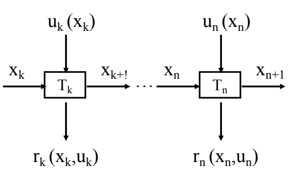
1. 多阶段决策过程的最优化

多阶段决策过程特点：



要点：阶段，状态，决策，状态转移方程，k-后部子过程

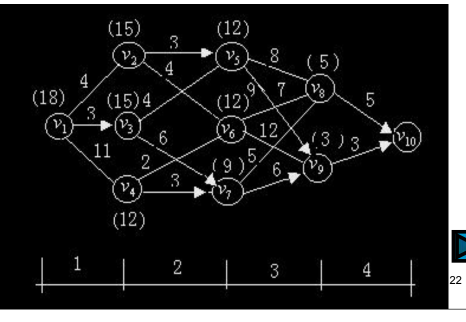
K后部子过程



多段决策过程中从第k阶段到最终阶段的过程称为k-后部子过程，简称k-子过程。

动态规划方法导引

例1：为了说明动态规划的基本思想方法和特点，下面以图1所示为例讨论的求最短路问题的方法。

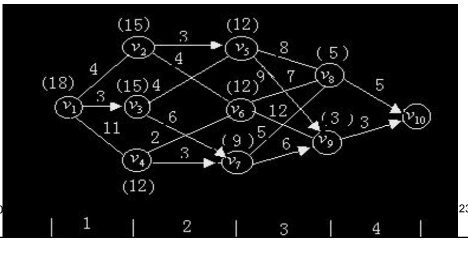


多阶段决策过程的最优化

第一种方法称做全枚举法或穷举法。

它的基本思想是列举出所有可能发生的方案和结果，再对它们一一进行比较，求出最优方案。

最优路线是 $v_1 \to v_3 \to v_7 \to v_9 \to v_{10}$，
最短距离是18.

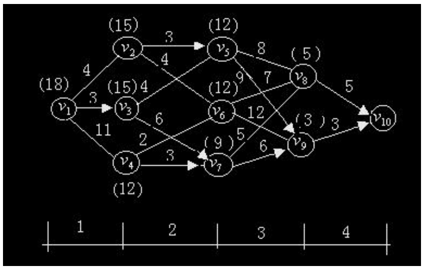


1. 多阶段决策过程的最优化

显然，当组成交通网络的节点很多时，用穷举法求最优路线的计算工作量将会十分庞大，而且其中包含着许多重复计算.

第二种方法即所谓“局部最优路径”法，是说某人从k出发，他并不顾及全线是否最短，只是选择当前最短途径，“逢近便走”，错误地以为局部最优会致整体最优，在这种想法指导下，所取决策必是 $v_1 \to v_3 \to v_5 \to v_8 \to v_{10}$，全程长度是20；显然，这种方法的结果常是错误的.

???



1. 多阶段决策过程的最优化

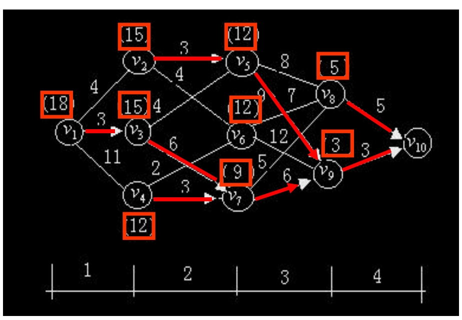
第三种方法是动态规划方法。

基本思想：

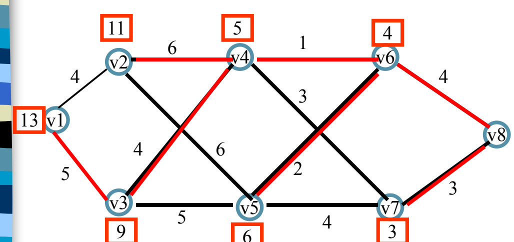
首先将问题划分为4个阶段，每次的选择总是综合后继过程的一并最优进行考虑，

在各段所有可能状态的最优后继过程都已求得的情况下，全程的最优路线便也随之得到。

为了找出所有可能状态的最优后继过程，动态规划方法总是从过程的最后阶段开始考虑，然后逆着实际过程发展的顺序，逐段向前递推计算直至始点。



练习



多阶段决策过程的最优化

整个求解过程分为两个阶段：

先按整体最优的思想逆序地求出各个子问题中所有可能状态的最优决策与最优路线值，

然后再顺序地求出整个问题的最优策略和最优路线。

2. 动态规划的基本概念

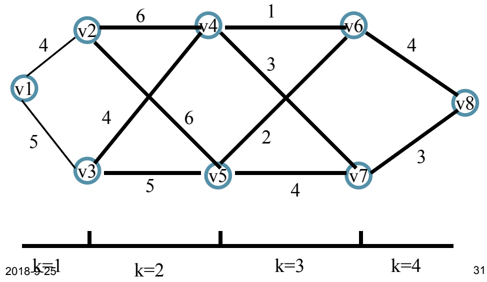
（一）阶段和阶段变量

为了便于求解和表示决策及过程的发展顺序，而把所给问题恰当地划分为若干个相互联系又有区别的子问题，称之为多段决策问题的阶段。

一个阶段，就是需要作出一个决策的子问题，通常，阶段是按决策进行的时间或空间上先后顺序划分的。

用以描述阶段的变量叫作阶段变量，一般以k表示阶段变量.

阶段数等于多段决策过程从开始到结束所需作出决策的数目。



2. 动态规划的基本概念

（二）状态、状态变量和可能状态集

1. 状态与状态变量。

用以描述事物(或系统)在某特定的时间与空间域中所处位置及运动特征的量，称为状态。

反映状态变化的量叫做状态变量。

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。

按照过程进行的先后，每个阶段的状态可分为初始状态和终止状态，或称输入状态和输出状态，阶段k的初始状态记作s_k（或x_k），终止状态记为s_{k+1}（或x_{k+1}）。

2. 动态规划的基本概念

2．可能状态集

一般状态变量的取值有一定的范围或允许集合，称为可能状态集，或可达状态集。

可能状态集实际上是关于状态的约束条件。通常可能状态集用相应阶段状态 s_k的大写字母S_k表示，$s_k \in S_k$，可能状态集可以是一离散取值的集合，也可以为一连续的取值区间，视具体问题而定.

2. 动态规划的基本概念

在上述最短路问题中，

第一阶段状态为v_1，状态变量s_1的状态集合$S_1=\{v_1\}$；

第二阶段则有三个状态：v_2, v_3, v_4，状态变量s_2的状态集合$S_2=\{v_2, v_3, v_4\}$；

第三阶段也有三个状态：v_5, v_6, v_7，状态变量s_3的状态集合$S_3=\{v_5, v_6, v_7\}$；

第四阶段则有二个状态：v_8, v_9，状态变量s_4的状态集合$S_4=\{v_8, v_9\}$；

2. 动态规划的基本概念

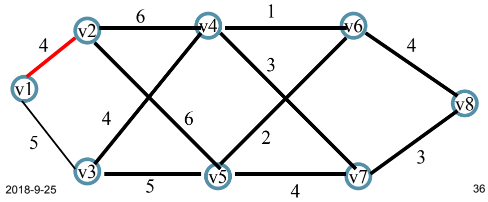
（三）决策、决策变量和允许决策集合

所谓决策，就是确定系统过程发展的方案。决策的实质是关于状态的选择，是决策者从给定阶段状态出发对下一阶段状态作出的选择。

用以描述决策变化的量称之决策变量. 和状态变量一样，决策变量可以用一个数，一组数或一向量来描述，也可以是状态变量的函数，记以$u_k= u_k(s_k)$，表示于阶段k处于状态s_k时的决策变量。

决策变量的取值往往也有一定的允许范围，称之允许决策集合。决策变量$u_k(s_k)$的允许决策集用$U_k(s_k)$表示，$u_k(s_k) \in U_k(s_k)$允许决策集合实际是决策的约束条件。

- $u_1(v1)$=v2
- $U_1(v1)$={v2, v3}
- $u_2(v3)$=v4
- $U_2(v3)$={v4, v5}



2. 动态规划的基本概念

（四）、策略和允许策略集合

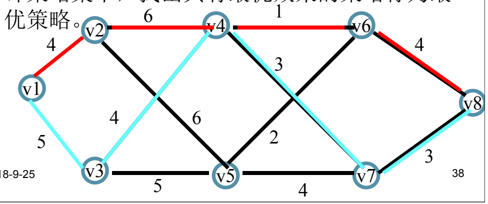
策略(Policy)也叫决策序列.

策略有全过程策略和k部子策略之分.

全过程策略是指具有n个阶段的全部过程，由依次进行的n个阶段决策构成的决策序列，简称策略，表示为$p_{1,n}\{u_1, u_2, ..., u_n\}$。

从k阶段到第n阶段，依次进行的阶段决策构成的决策序列称为k部子策略，表示为$p_{k,n}\{u_k, u_{k+1}, ..., u_n\}$，显然当$k$=1时的$k$部子策略就是全过程策略。

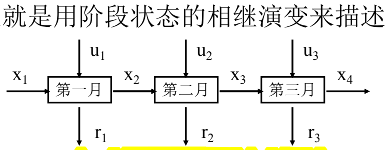
- 在实际问题中，由于在各个阶段可供选择的决策有许多个，因此，它们的不同组合就构成了许多可供选择的决策序列(策略)，由它们组成的集合，称之允许策略集合，记作$P_{1,n}$，从允许策略集中，找出具有最优效果的策略称为最优策略。



2. 动态规划的基本概念

（五）状态转移方程

系统在阶段k处于状态s_k，执行决策$u_k(s_k)$的结果是系统状态的转移，即系统由阶段k的初始状态s_k转移到终止状态s_{k+1}，或者说，系统由k阶段的状态s_k转移到了阶段k+1的状态s_{k+1}，多阶段决策过程的发展就是用阶段状态的相继演变来描述的。



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

2. 动态规划的基本概念

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的状态转移完全由阶段k的状态s_k和决策$u_k(s_k)$所确定，与系统过去的状态$s_1, s_2, ...$，s_{k-1}及其决策$u_1(s_1)$, $u_2(s_2)...u_{k-1}(s_{k-1})$无关。系统状态的这种转移，用数学公式描述即有：

$$s_{k+1} = T_k(s_k, u_k(s_k))$$

有些问题的状态转移方程不一定存在数学表达式，但是它们的状态转移，还是有一定规律可循的。

2. 动态规划的基本概念

（六）指标函数

用来衡量策略或子策略或决策的效果的某种数量指标，就称为指标函数。

它是定义在全过程或各子过程或各阶段上的确定数量函数。对不同问题，指标函数可以是诸如费用、成本、产值、利润、产量、耗量、距离、时间、效用，等等。

2. 动态规划的基本概念

(1)阶段指标函数（也称阶段效应）。

用$g_k(s_k, u_k)$表示第k段处于s_k状态且所作决策为$u_k(s_k)$时的指标，则它就是第k段指标函数，简记为g_k。

例如$g_2(v_2, v_5)$=3, 即v_2到v_5的距离为3。

(2)过程指标函数（也称目标函数）。用$R_k(s_k, u_k)$表示第k子过程的指标函数。

如例1的$R_k(s_k, u_k)$表示处于第k段s_k状态且所作决策为u_k时，从s_k点到终点v_{10}的距离。由此可见，$R_k(s_k, u_k)$不仅跟当前状态s_k有关，还跟该子过程策略$p_k(s_k)$有关，因此它是s_k和$p_k(s_k)$的函数，严格说来，应表示为：

$$R_k(s_k, p_k(s_k))$$

2. 动态规划的基本概念

实际应用中指标函数往往表示为$R_k(s_k, u_k)$或$R_k(s_k)$。

它与第k子过程上各段指标函数有关，过程指标函数$R_k(s_k)$通常是描述k后部子过程效果优劣的数量指标，它是由各阶段的阶段指标函数$g_k(s_k, u_k)$累积形成的，

适于用动态规划求解的问题的过程指标函数（即目标函数），必须具有关于阶段指标的可分离形式．对于k部子过程的指标函数可以表示为:

$$R_{k,n} = R_{k,n}(s_k, u_k, s_{k+1}, u_{k+1}, \cdots, s_n, u_n)$$
$$= g_k(s_k, u_k) \oplus g_{k+1}(s_{k+1}, u_{k+1}) \oplus \cdots \oplus g_n(s_n, u_n)$$

式中，⊕表示某种运算，可以是加、减、乘、除、开方等。

2. 动态规划的基本概念

多阶段决策问题中，常见的目标函数形式之一是取各阶段效应之和的形式，即：

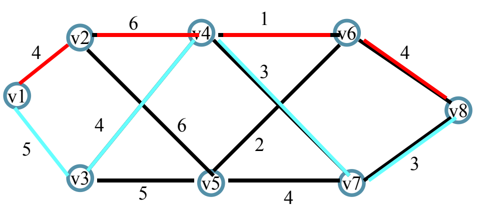
$$R_k = \sum_{i=k}^{n} g_i(s_i, u_i)$$

有些问题，如系统可靠性问题，其目标函数是取各阶段效应的连乘积形式，如:

$$R_k = \prod_{i=k}^{n} g_i(s_i, u_i)$$

总之，具体问题的目标函数表达形式需要视具体问题而定。

阶段指标



2. 动态规划的基本概念

（七）最优解

用$f_k(s_k)$表示第k子过程指标函数 $R_k(s_k, p_k(s_k))$ 在状态s_k下的最优值，即

$$f_k(s_k) = \underset{p_k \in P_K(s_k)}{opt} \{R_k(s_k, p_k(s_k))\}, k = 1,2,\cdots,n$$

称$f_k(s_k)$为第k子过程上的最优指标函数；与它相应的子策略称为s_k状态下的最优子策略，记为$p_k^*(s_k)$ ；而构成该子策略的各段决策称为该过程上的最优决策记为 ;

$$p_k^*(s_k) = \{u_k^*(s_k), u_{k+1}^*(s_{k+1}), \cdots; u_n^*(s_n)\} k = 1,2,\cdots;n$$

简记为

$$p_k^* = \{u_k^*, u_{k+1}^*, \cdots, u_n^*\}, k = 1,2,\cdots,n$$

2. 动态规划的基本概念

特别当k=1且s_1取值唯一时，$f_1(s_1)$就是问题的最优值，而p_1*就是最优策略。如例1只有唯一始点v_1即s_1取值唯一，故$f_1(s_1)$=18就是例的最优值，而

$$p_1^* = \{v_3, v_7, v_9, v_{10}\}$$

就是例1的最优策略。

我们把最优策略和最优值统称为问题的最优解。

2. 动态规划的基本概念

（八）多阶段决策问题的数学模型

综上所述，适于应用动态规划方法求解的一类多阶段决策问题，亦即具有无后效性的多阶段决策问题的数学模型呈以下形式：

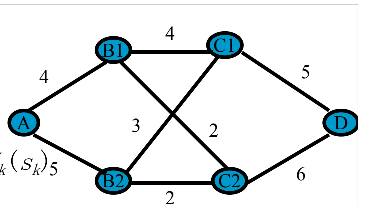
$$f = \underset{u_1 \sim u_n}{opt}\ R = R(s_1, u_1, s_2, u_2, \cdots, s_n, u_n)$$

$$s.t.\begin{cases} s_{k+1} = T_k(s_k, u_k) \\ s_k \in S_k \\ u_k \in U_k \\ k = 1,2,\cdots,n \end{cases} \qquad (5\text{-}5)$$

2. 动态规划的基本概念

式中“OPT”表示最优化，视具体问题取max或min。

上述数学模型说明了对于给定的多阶段决策过程，求取一个(或多个)最优策略或最优决策序列 $\{u_1^*, u_2^*, \cdots, u_n^*\}$



- $p_k^*(s_k)$
- $u_k(s_k) \in U_k(s_k)$
- s_k（或x_k）
- $s_k \in S_k$
- $g_k(s_k, u_k)$
- $R_k(s_k, u_k)$或$R_k(s_k)$
- $p_{k,n}\{u_k, u_{k+1}, \ldots, u_n\}$

$$s_{k+1} = T_k(s_k, u_k(s_k))$$

$$f_k(s_k) = \underset{p_k \in P_k(s_k)}{opt}\ \{R_k(s_k, p_k(s_k))\}, k = 1,2,\cdots,n$$

2. 动态规划的基本原理

二、动态规划的最优化原理与基本方程

1．标号法

只适用于例1这类最优路线问题的特殊解法—标号法。

标号法是借助网络图通过分段标号来求出最优路线的一种简便、直观的方法。通常标号法采取“逆序求解”的方法来寻找问题的最优解，即从最后阶段开始，逐次向阶段数小的方向推算，最终求得全局最优解。

2. 动态规划的基本原理

下面给出标号法的一般步骤：

1. 从最后一段标起，该段各状态(即各始点)到终点的距离用数字分别标在各点上方的方格内，并用粗箭线连接各点和终点。

2. 向前递推，给前一阶段的各个状态标号。每个状态上方方格内的数字表示该状态到终点的最短距离。即为该状态到该阶段已标号的各终点的段长，再分别加上对应终点上方的数字而取其最小者。将刚标号的点沿着最短距离所对应的已标号的点用粗箭线连接起来，表示出各刚标号的点到终点的最短路线。

2. 动态规划的基本原理

3. 逐次向前递推，直到将第一阶段的状态(即起点)也标号，起点方格内的数字就是起点到终点的最短距离，从起点开始连接终点的粗箭线就是最短路线。

2. 动态规划的基本原理

用标号法来求解下例

例5.2： 网络图表示某城市的局部道路分布图。一货运汽车从S出发，最终到达目的地E。其中A_i(i=1，2，3)，B_j(j=1，2)和C_k(k=1，2)是可供汽车选择的途经站点，各点连线上的数字表示两个站点间的距离。问此汽车应走哪条路线，使所经过的路程距离最短？

2. 动态规划的基本原理

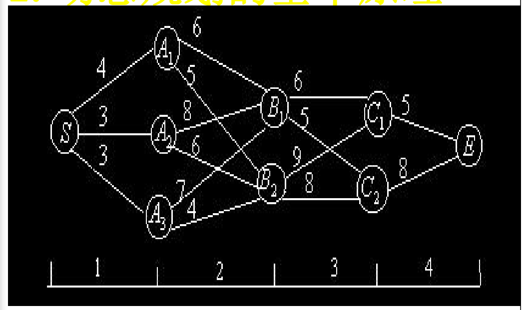
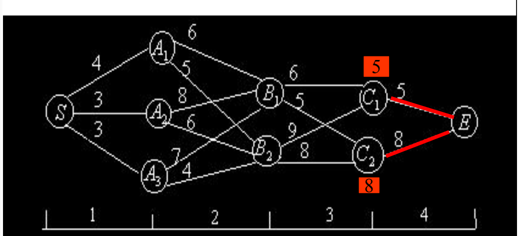


图5-2　某城市的局部道路分布图

2. 动态规划的基本原理

第一步：先考虑第四阶段，即k=4，

该阶段共有两个状态：C_1、C_2，设$f_4(C_1)$和$f_4(C_2)$分别表示C_1、C_2到E的最短距离，显然有$f_4(C_1)$=5和$f_4(C_2)$=8，边界条件$f_5(E)$=0 。



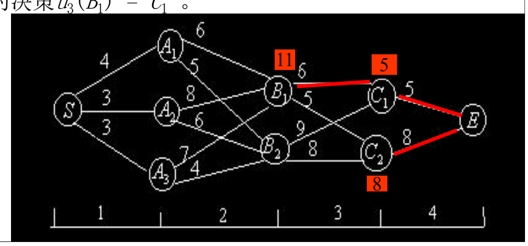
第二步：即k=3，该阶段共有两个状态：B_1，B_2

从B_1出发有两种决策：$B_1 \to C_1$，$B_1 \to C_2$ 。

$d_3(B_1, C_1)$表示B_1到C_1的距离，$B_1 \to C_1$的阶段指标函数为$d_3(B_1, C_1)=6$，$B_1 \to C_2$的阶段指标函数为$d_3(B_1, C_2)=5$。

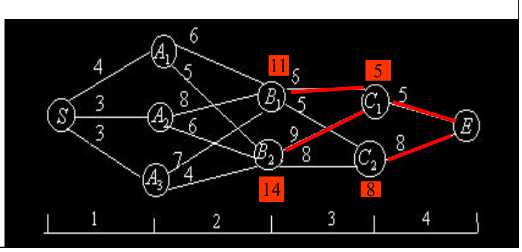
$f_3(B_1)=\min\{d_3(B_1, C_1)+f_4(C_1), d_3(B_1, C_2)+f_4(C_2)\}=\min(6+5，5+8)=11$。

那么，从B1出发到E的最短路线是$B_1 \to C_1 \to E$，对应的决策$u_3(B_1) = C_1$ 。



2. 动态规划的基本原理

(2) 从B_2出发也有两种决策：$B_2 \to C_1$，$B_2 \to C_2$同理，有$f_3(B_2)=\min\{d_3(B_2, C_1)+f_4(C_1), d_3(B_2, C_2)+f_4(C_2)\}=\min(9+5，8+87)=14$，那么，从$B_2$出发到$E$的最短路线是$B_2 \to C_1 \to E$，且$u_3(B_2)=C_1$ 。



第三步：即k=2，该阶段共有三个状态：A_1, A_2, A_3

（1）$A_1 \to B_1, A_1 \to B_2$。则

$f_2(A_1)=\min\{d_2(A_1, B_1)+f_3(B_1), d_2(A_1, B_2)+f_3(B_2)\}=\min\{6+11，5+14\}=17$，即$A_1$到$E$的最短路线为$A_1 \to B_1 \to C_1 \to E$且$u_3(A_1)=B_1$ 。

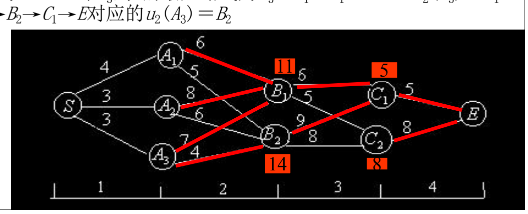
（2）$A_2 \to B_1, A_2 \to B_2$。

$f_2(A_2)=\min\{d_2(A_2, B_1)+f_3(B_1), d_2(A_2, B_2)+f_3(B_2)\}=\min\{8+11，6+14\}=19$，即$A_2$到$E$的最短路线为$A_2 \to B_1 \to C_1 \to E$，且$u_3(A_2)=B_1$。

(3) $A_3 \to B_1, A_3 \to B_2$ 此时

$f_2(A_3)=\min\{d_2(A_3, B_1)+f_3(B_1), d_2(A_3, B_2)+f_3(B_2)\}=\min\{7+11，4+14\}=18$，即$A_3$到$E$的最短路线为$A_3 \to B_1 \to C_1 \to E$ ，　$u_2(A_3)=B_1$

$A_3 \to B_2 \to C_1 \to E$对应的$u_2(A_3)=B_2$

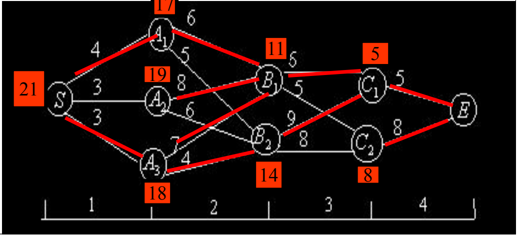


第四步：即$k=1$，该阶段只有一个状态S，从S出发有三种决策：$S\to A_1$，$S\to A_2$，$S\to A_3$，那么，

$f_1(S)=\min\{d_1(S,A_1)+f_2(A_1),d1(S,A_2)+f_2(A_2),d1(S,A_3)+f_2(A_3)\}=\min\{4+17, 3+19, 3+18\}=21$，

那么，从S到E共有三条最短路线：

此时，$u_1(S)=A_1$　$u_1(S)=A_3$，最短距离为21。

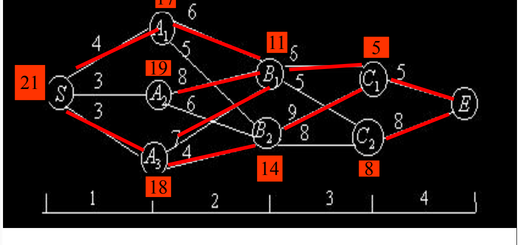


2. 动态规划的基本原理

每个状态上方的方格内的数字表示该状态到E的最短距离，首尾相连的粗箭线构成每一状态到E的最短路线。

因此，标号法不但给出起点到终点的最短路线和最短距离，同时也给出每一状态到终点的最短路线及其最短距离。

2. 动态规划的基本原理



S---A1---B1---C1---E

2. 动态规划的基本原理

2．最优化原理　（贝尔曼最优化原理）

作为一个全过程的最优策略具有这样的性质：对于最优策略过程中的任意状态而言，无论其过去的状态和决策如何，余下的诸决策必构成一个最优子策略。

即若某一全过程最优策略为：

$$p_1^*(s_1)=\{u_1^*(s_1),u_2^*(s_2),\cdots,\boxed{u_k^*(s_k),\cdots u_n^*(s_n)}\}$$

则对上述策略中所隐含的任一状态而言，第k子过程上对应于该状态的最优策略必然包含在上述全过程最优策略p_1^*中，即为

$$p_k^*(s_k)=\{u_k^*(s_k),u_{k+1}^*(s_{k+1}),\cdots,u_n^*(s_n)\}$$

2. 动态规划的基本原理

逆序递推法：这里可以指出，该法的关键在于给出一种递推关系。一般把这种递推关系称为动态规划的**函数基本方程**。

2. 动态规划的基本原理

第k子过程的最优指标

函数基本方程

上例中，用标号法求解最短路线的

第k+1子过程的最优指标

可概括写成：

边界条件

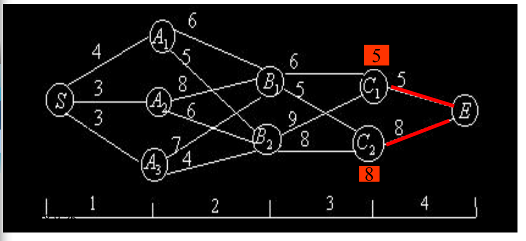
$$\begin{cases} f_5(s_5)=0 \\ f_k(s_k)=\min\limits_{u_k\in U_k(s_k)}\{g_k(s_k,u_k(s_k))+f_{k+1}(\cancel{u}_{k+1}(s_{k+1}))\} \quad k=4,3,2,1 \end{cases} \quad (5\text{-}6)$$

K阶段指标

其中，$g_k(s_k,u_k(s_k))$在这里表示从状态s_k到由决策$u_k(s_k)$所决定的状态s_{k+1}之间的距离，

边界条件表示全过程到第四阶段终点结束。

$$\begin{cases} f_5(s_5)=0 \\ f_3(B_1)=\min\limits_{u_3\in U_3(B_1)}\{g_3(B_1,u_3(B_1))+f_4(u_4(s_4))\} \\ \quad =\min\limits_{u_3\in\{C_1,C_2\}}\{6+5,5+8\} \\ \quad =11 \end{cases}$$



2. 动态规划的基本原理

一般地，对于n个阶段的决策过程，假设只考虑指标函数是“和”与“积”的形式，第k阶段和第k+1阶段间的递推公式可表示如下：

(1) 当过程指标函数为下列“和”的形式时，

$$f_k(s_k)=\underset{p_k\in P_K(s_k)}{opt}\{R_k(s_k,p_k(s_k))\}=\sum_{i=k}^{n} g_i(s_i,u_i)$$

相应的函数基本方程为

$$\begin{cases} f_{n+1}(s_{n+1})=0 \\ f_k(s_k)=\underset{u_k\in U_k}{opt}\{\boxed{g_k(s_k,u_k(s_k))}+\boxed{f_{k+1}(\cancel{u}_{k+1}(s_{k+1}))}\},k=n,n-1,\cdots,2,1 \end{cases}$$

2. 动态规划的基本原理

(2) 当过程指标函数为下列“积”的形式时，

$$f_k(s_k)=\underset{p_k\in P_k(s_k)}{opt}\{R_k(s_k,p_k(s_k))\}=\prod_{i=k}^{n} g_i(s_i,u_i)$$

相应的函数基本方程为

$$\begin{cases} f_{n+1}(s_{n+1})=1 \\ f_k(s_k)=\underset{u_k\in U_k}{opt}\{g_k(s_k,u_k(s_k))\bullet f_{k+1}(u_{k+1}(s_{k+1}))\} k=n,n-1,\cdots,2,1 \end{cases}$$

可以看出，和、积函数的基本方程中边界条件(即 $f_{n+1}(s_{n+1})$ 的取值)是不同的。

3. 动态规划方法的基本步骤

一．动态规划的建模

标号法仅适用于求解象最短路线问题那样可以用网络图表示的多阶段决策问题。但不少多阶段决策问题不能用网络图表示。此时，应该用函数基本方程来递推求解.

首先，要有效地建立动态规划模型，然后再递推求解，最后得出结论.

构造成动态规划模型，这是非常重要的一步。正确地建立一个动态规划模型，往往问题也就解决了一大半，而一个正确的动态规划模型，应该满足哪些条件呢？

3. 动态规划方法的基本步骤

1．**应将实际问题恰当地分割成*n*个子问题(*n*个阶段)**。通常是根据时间或空间而划分的，或者在经由静态的数学规划模型转换为动态规划模型时，常取静态规划中变量的个数n，即k=n。

2．**正确地定义状态变量s_k，使它既能正确地描述过程的状态，又能满足无后效性**．动态规划中的状态与一般控制系统中和通常所说的状态的概念是有所不同的，动态规划中的状态变量必须具备以下三个特征：

3. 动态规划方法的基本步骤

(1) 要能够正确地描述受控过程的变化特征。

(2) 要满足无后效性。即如果在某个阶段状态已经给定，那么在该阶段以后，过程的发展不受前面各段状态的影响，如果所选的变量不具备无后效性，就不能作为状态变量来构造动态规划的模型。

(3) 要满足可知性。即所规定的各段状态变量的值，可以直接或间接地测算得到。一般在动态规划模型中，状态变量大都选取那种可以进行累计的量。此外，在与静态规划模型的对应关系上，通常根据经验，线性与非线性规划中约束条件的个数，相当于动态规划中状态变量 s_k的维数．而前者约束条件所表示的内容，常就是状态变量 s_k 所代表的内容。

3. 动态规划方法的基本步骤

3. 正确地定义决策变量及各阶段的允许决策集合$U_k(s_k)$，根据经验，一般将问题中待求的量，选作动态规划模型中的决策变量。或者在把静态规划模型（如线性与非线性规划）转换为动态规划模型时，常取前者的变量x_j为后者的决策变量u_k。

4. 能够正确地写出状态转移方程，至少要能正确反映状态转移规律。如果给定第k阶段状态变量s_k的值，则该段的决策变量u_k一经确定，第k+1段的状态变量s_{k+1}的值也就完全确定，即有$s_{k+1}=T_k(s_k, u_k)$

3. 动态规划方法的基本步骤

5. 根据题意，正确地构造出目标与变量的函数关系——目标函数，目标函数应满足下列性质：

(1) 可分性，即对于所有k后部子过程，其目标函数仅取决于状态s_k及其以后的决策u_k，u_{k+1}，┄，u_n，就是说它是定义在全过程和所有后部子过程上的数量函数。

(2) 要满足递推关系，即

$$R_{k,n}(s_k,u_k,s_{k+1},u_{k+1},\cdots,s_{n+1})=\varphi_k[s_k,u_k,R_{k+1}(s_{k+1},\cdots,s_{n+1})]$$

(3) 函数 $\varphi_k[s_k,u_k,R_{k+1}(s_{k+1},\cdots,s_{n+1})]$对其变元$R_{k+1}$来说要严格单调。

3. 动态规划方法的基本步骤

6. 写出动态规划函数基本方程

例如常见的指标函数是取各段指标和的形式

$$R_k(s_k)=\sum_{i=k}^{n} g_i(s_i,u_i)$$

其中 $g_i(s_i,u_i)$ 表示第i阶段的指标，它显然是满足上述三个性质的。所以上式可以写成：

$$R_k=g_k(s_k,u_k)+R_{k+1}(s_{k+1},\cdots\cdots,s_{n+1})$$

3. 动态规划方法的基本步骤

二．动态规划方法的基本步骤

例5.3：有某种机床，可以在高低两种不同的负荷下进行生产，

在高负荷下生产时，产品的年产量为g，与年初投入生产的机床数量u_1的关系为$g=g(u_1)=8u_1$，

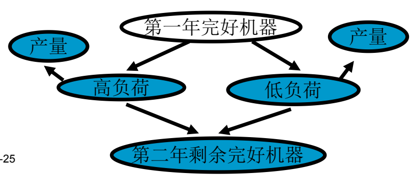
这时，年终机床完好台数将为au_1，（a为机床完好率，$0<a<1$，设a=0.7）.

在低负荷下生产时，产品的年产量为h，和投入生产的机床数量u_2的关系为$h=h(u_2)=5u_2$，相应的机床完好率为b（$0<b<1$，设b=0，9），一般情况下$a<b$。

3. 动态规划方法的基本步骤

假设某厂开始有x=1000台完好的机床，现要制定一个五年生产计划，

问每年开始时如何重新分配完好的机床在两种不同的负荷下生产的数量，以使在5年内产品的总产量为最高。



解：首先构造这个问题的动态规划模型。

1．变量设置

(1) 设阶段变量k表示年度，因此，阶段总数n=5。

(2) 状态变量s_k表示第k年度初拥有的完好机床台数，同时也是第k-1年度末时的完好机床数量。

3. 动态规划方法的基本步骤

(3) 决策变量u_k，表示第k年度中分配于高负荷下生产的机床台数。

$s_k- u_k$便为该年度中分配于低负荷下生产的机床台数.

这里s_k与u_k均取连续变量，当它们有非整数数值时．可以这样理解：如s_k=0.6，就表示一台机器在k年度中正常工作时间只占6/10；u_k=0.4时，就表示一台机床在 k年度只有4/10的时间于高负荷下工作.

2．状态转移方程为

$$s_{k+1} = au_k + b(s_k - u_k) = 0.7u_k + 0.9(s_k - u_k)$$

k=1，2，…，6

3. 动态规划方法的基本步骤

3．允许决策集合，在第k段为

$$U_k(s_k) = \{u_k \mid 0 \le u_k \le s_k\}$$

4．目标函数。设$g_k(s_k, u_k)$为第k年度的产量，则$g_k(s_k, u_k)$=8u_k+5(s_k-u_k)，因此，目标函数为

$$R_k = \sum_{i=k}^{5} g_k(s_k, u_k)$$

k=1, 2, . . . , 5

5．条件最优目标函数递推方程。

令$f_k(s_k)$表示由第k年的状态s_k出发，采取最优分配方案到第5年度结束这段时间的产品产量，根据最优化原理有以下递推关系：

$$f_k(s_k) = \max_{0 \le u_k \le s_k} \{[8u_k + 5(s_k - u_k)] + f_{k+1}[0.7u_k + 0.9(s_k - u_k)]\}$$

k=1, 2, 3, 4, 5

3. 动态规划方法的基本步骤

6. 边界条件为 $f_{5+1}(s_{5+1}) = 0$

下面采用逆序递推计算法，从第5年度开始递推计算。

k=5时有 $f_5(s_5) = \max\limits_{0 \le u_k \le s_5} \{[8u_5 + 5(s_5 - u_5)] + f_6(s_k)\}$

显然，当u_5*=s_5时，$f_5(s_5)$有最大值，相应的有

$f_5(s_5)$=8s_5 $= \max\limits_{0 \le u_k \le s_5} \{[8u_5 + 5(s_5 - u_5)]$

k=4时有

$$f_4(s_4) = \max_{0 \le u_4 \le s_4} \{[8u_4 + 5(s_4 - u_4)] + f_5[0.7u_4 + 0.9(s_4 - u_4)]\}$$
$$= \max_{0 \le u_4 \le s_4} \{[8u_4 + 5(s_4 - u_4)] + 8[0.7u_4 + 0.9(s_4 - u_4)]\}$$
$$= \max_{0 \le u_4 \le s_4} \{13.6u_4 + 12.2(s_4 - u_4)\}$$

因此，当u_4*=s_4时，有最大值$f_4(s_4)$=13.6s_4

3. 动态规划方法的基本步骤

k=3 时有

$$f_3(s_3) = \max_{0 \le u_3 \le s_3} \{[8u_3 + 5(s_3 - u_3)] + f_4(s_4)\}$$
$$= \max_{0 \le u_3 \le s_3} \{17.55u_3 + 17.22(s_3 - u_3)\}$$

可见，当u_3*=s_3时，$f_3(s_3)$有最大值$f_3(s_3)$ =17.55s_3.

k=2 时有

$$f_2(s_2) = \max_{0 \le u_2 \le s_2} \{[8u_2 + 5(s_2 - u_2)] + f_3(s_3)\}$$
$$= \max_{0 \le u_2 \le s_2} \{[8u_2 + 5(s_2 - u_2)] + 17.55[0.7u_2 + 0.9(s_2 - u_2)]\}$$
$$= \max_{0 \le u_2 \le s_2} \{[20.25u_2 + 20.8(s_2 - u_2)]$$

当u_2*=0时，$f_2(s_2)$有最大值$f_2(s_2)$ =20.8 s_2

3. 动态规划方法的基本步骤

k=1时有

$$f_1(s_1) = \max_{0 \le u_1 \le s_1} \{8u_1 + 5(s_1 - u_1) + 20.8[0.7u_1 + 0.9(s_1 - u_1)]\}$$
$$= \max_{0 \le u_1 \le s_1} \{22.55u_1 + 23.7(s_1 - u_1)\}$$

当取u_1*=0时，$f_1(s_1)$有最大值，即$f_1(s_1)$=23.7s_1，因为s_1=1000，故$f_1(s_1)$=23700个产品.

按照上述计算顺序寻踪得到下述计算结果：

$u_1^* = 0$ $s_1 = 1000$ $g_1(s_1, u_1) = 5000$ $f_1(s_1) = 23700$

$u_2^* = 0$ $s_2 = 900$ $g_2(s_2, u_2) = 4500$ $f_2(s_2) = 20.8s_2 = 18720$

3. 动态规划方法的基本步骤

$$u_3^* = s_3 \; s_3 = 810 \quad g_3(s_3,u_3) = 6480 \quad f_3(s_3) = 17.55s_3 = 14216$$

$$u_4^* = s_4 \; s_4 = 567 \quad g_4(s_4,u_4) = 4536 \quad f_4(s_4) = 13.6s_4 = 7711$$

$$u_5^* = s_5 \;\; s_5 = 397 \quad g_5(s_5,u_5) = 3176 \quad f_5(s_5) = 13.6s_5 = 3176$$

$$s_6 = 0.7u_5 + 0.9(s_5 - u_5) = 0.7x_5 = 278$$

上面所讨论的最优决策过程是所谓始端状态**s_1**固定，终端状态**s_6**自由．如果终端也附加上一定的约束条件，那么计算结果将会与之有所差别．例如，若规定在第五个年度结束时，完好的机床数量为**500**台(上面只有**278**台)，问应该如何安排五年的生产，使之在满足这一终端要求的情况下产量最高？

3. 动态规划方法的基本步骤

解：由状态转移方程

$$s_{k+1} = au_k + b(s_k - u_k) = 0.7u_k + 0.9(s_k - u_k)$$

有得

$$s_6 = 0.7u_5 + 0.9(s_5 - u_5) = 500$$

$$u_5 = 4.5s_5 - 2500$$

显而易见，由于固定了终端的状态s_6，第五年的决策变量U_5的允许决策集合$U_5(s_5)$也有了约束，上式说明$U_5(s_5)$已退化为一个点，即第五年投入高负荷下生产的机床数只能由式U_5=4.5s_5−2500作出一种决策，故

3. 动态规划方法的基本步骤

当k=5时有

$$f_5(s_5) = \max_{0 \le u_k \le s_5}\{8u_5 + 5(s_5 - u_5)\} = 8(4.5u_5 - 2500) + 5(s_5 - 4.5u_5 + 2500)$$

$$= 18.5s_5 - 7500$$

当k=4时有

$$f_4(s_4) = \max_{0 \le u_4 \le s_4}\{[8u_4 + 5(s_4 - u_4)] + f_5(s_5)\}$$

$$= \max_{0 \le u_4 \le s_4}\{8u_4 + 5(s_4 - u_4) + 18.5s_5 - 7500\}$$

$$= \max_{0 \le u_4 \le s_4}\{8u_4 + 5(s_4 - u_4) + 18.5[0.7u_4 + 0.9(s_4 - u_4)] - 7500\}$$

$$= \max_{0 \le u_4 \le s_4}\{21.7s_4 - 0.75u_4 - 7500\}$$

显然，只有取u_4^*=0 ，$f_4(s_4)$有最大值，即$f_4(s_4)$=21.7s_4−7500。同理类推

3. 动态规划方法的基本步骤

k=3时有

$$f_3(s_3) = \max_{0 \le u_3 \le s_3}\{[8u_3 + 5(s_3 - u_3)] + f_4(s_4)\}$$

$$= \max_{0 \le u_3 \le s_3}\{[8u_3 + 5(s_3 - u_3)] + 21.7[0.7u_3 + 0.9(s_3 - u_3)] - 7500\}$$

$$= \max_{0 \le u_3 \le s_3}\{-1.3u_3 + 24.5s_3 - 7500\}$$

可知，当u_3^*=0时，$f_3(s_3)$有最大值$f_4(s_4)$=24.5s_3−7500.

k=2时有

$$f_2(s_2) = \max_{0 \le u_2 \le s_2}\{[8u_2 + 5(s_2 - u_2)] + f_3(s_3)\}$$

$$= \max_{0 \le u_2 \le s_2}\{[8u_2 + 5(s_2 - u_2)] + 24.5[0.7u_2 + 0.9(s_2 - u_2)] - 7500\}$$

$$= \max_{0 \le u_2 \le s_2}\{-1.9u_2 + 27.1s_2 - 7500\}$$

此时，当u_2^*=0时有最大值，即

$f_2(s_2)$=27.1s_2−7500

3. 动态规划方法的基本步骤

k=1时有

$$f_1(s_1) = \max_{0 \le u_1 \le s_1}\{8u_1 + 5(s_1 - u_1) + f_2(s_2)\}$$

$$= \max_{0 \le u_1 \le s_1}\{8u_1 + 5(s_1 - u_1) + 27.1[0.7u_1 + 0.9(s_1 - u_1)] - 7500\}$$

$$= \max_{0 \le u_1 \le s_1}\{-2.4u_1 + 29.4s_1 - 7500\}$$

只有取u_1^*=0时，$f_1(s_1)$有最大值，即

$f_1(s_1)$=29.4s_1−7500 。

由此可见，为了使下一个五年计划开始的一年有完好的机床500台，其最优策略应该为：在前4年中，都应该把全部机床投入低负荷下生产，在第5年，只能把部分完好机投入高负荷下生产。根据最优策略，从始端向终端递推计算出各年的状态，即算出每年年初的完好机床台数，因为s_1=1000台，于是有

3. 动态规划方法的基本步骤

$$u_1^* = 0 \qquad s_1 = 1000\text{（台）}$$

$$u_2^* = 0 \qquad s_2 = 0.7u_1^* + 0.9(s_1 - u_1^*) = 0.9s_1 = 900\text{（台）}$$

$$u_3^* = 0 \qquad s_3 = 0.7u_2^* + 0.9(s_2 - u_2^*) = 0.9s_2 = 810\text{（台）}$$

$$u_4^* = 0 \qquad s_4 = 0.7u_3^* + 0.9(s_3 - u_3^*) = 0.9s_3 = 729\text{（台）}$$

$$s_5 = 0.7u_4^* + 0.9(s_4 - u_4^*) = 0.9s_4 = 656\text{（台）}$$

$$s_6 = 0.7u_5 + 0.9(s_5 - u_5) = 0.7x_5 = 278\text{（台）}$$

因此，u_5^*=4.5s_5-2500=425(台)，这就是说第5年里还有204台投入低负荷下生产，否则不能保证s_6=0.7u_5+0.9(u_5−s_5)=500(台)。

在上述最优决策下，5年里所得最高产量为：

$f_1(s_1)$＝29.4s_1−7500＝29400-7500＝21900(个)。

可见，附加了终端约束条件以后，其最高产量$f_1(s_1)$比终端自由时要低一些。

1. 动态规划的四大要素

① 状态变量及其可能集合 $x_k \in X_k$

② 决策变量及其允许集合 $u_k \in U_k$

③ 状态转移方程 $x_{k+1} = T_k(x_k, u_k)$

④ 阶段效应 $r_k(x_k, u_k)$

2. 动态规划基本方程

$f_{n+1}(x_{n+1}) = 0$ (边界条件)

$f_k(x_k) = \mathrm{opt}_u\{r_k(x_k, u_k) + f_{k+1}(x_{k+1})\}$

$k = n,\ldots,1$