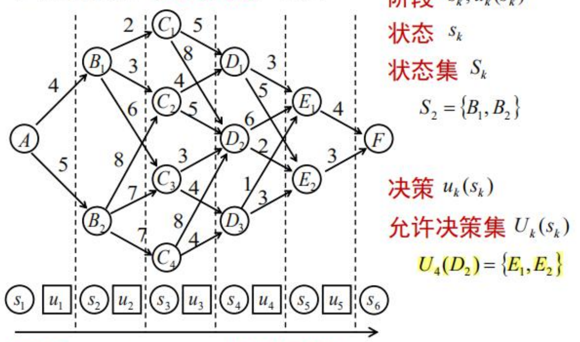
动态规划主要用于求解能够用不同的划分阶段表示的动态过程的优化问题。

一般的动态规划问题可以简单地表示为一个多阶段决策问题，其表示如下：

多阶段决策（序贯决策）问题



策略 $p_{k,5}=\{u_k,\cdots,u_5\}$ 允许策略集 $P_{k,5}=\{U_k(s_k),\cdots,U_5(s_5)\}$

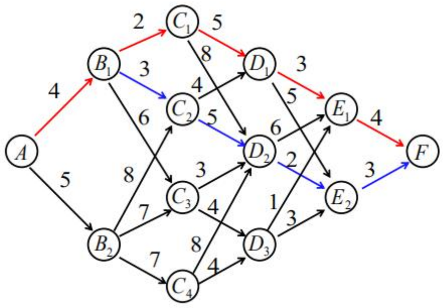
状态转移方程 $s_{k+1}=T_k(s_k,u_k)$ 本问题 $T_k(s_k,u_k)=u_k(s_k)$

阶段指标函数 $d_k(s_k,u_k)$ 如 $d_4(D_2,u_4)=d_4(D_2,E_1)=6$

过程指标函数 $V_{k,5}(s_k,p_{k,5})=\sum_{i=k}^{5}d_i(s_i,u_i)$

比方说，我们要求解如下的最短路问题：

例1: 最短路问题



选择从 A 至 F 的最短路铺设输油管道

用多阶段决策的术语描述最短路问题：

已知 状态集 $S_k, k=1,2,\cdots,6$

允许决策集 $U_k(s_k), \forall s_k\in S_k, k=1,2,\cdots,5$

状态转移方程

$s_{k+1}=T_k(s_k,u_k), \forall s_k\in S_k, u_k(s_k)\in U_k(s_k), k=1,2,\cdots,5$

阶段指标函数

$d_k(s_k,u_k), \forall s_k\in S_k, u_k(s_k)\in U_k(s_k), k=1,2,\cdots,5$

问题 求 $p_{1,5}\in P_{1,5}$ 使下述过程指标函数达到最小

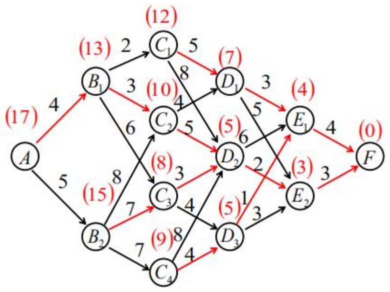
$$V_{1,5}(s_1,p_{1,5})=\sum_{k=1}^{5}d_k(s_k,u_k)$$

用动态规划求解的多阶段模型必须具有无后效性，也被称为马尔可夫（Markov）性。也就是说对于下面的问题：

$$\begin{aligned}\min\quad & V_{1,5}(s_1,p_{1,5})=\sum_{k=1}^{5}d_k(s_k,u_k)\\ \text{s.t.}\quad & s_{k+1}=T_k(s_k,u_k),\quad 1\le k\le 5\\ & s_k\in S_k,\ u_k(s_k)\in U_k(s_k),\ k=1,2,\cdots,5\end{aligned}$$

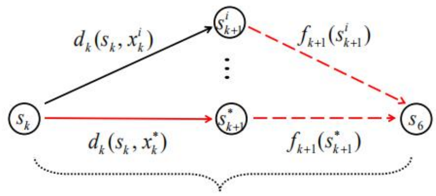
给定 s_k，系统以后的状态就完全由 k 及其以后各阶段的决策所决定，和系统经由什么路径到达 s_k 无关，即和 $s_1, s_2, ..., s_{k-1}$, 的取值无关。

对于无后效的最短路问题，我们可以通过图解法解决：从最后阶段开始逆过程行进方向依次导出到终点的最短距离（最优过程指标函数）及相应路径（最优决策）



其合理性来源于任何满足马尔可夫性的序贯决策问题所具有的最优性原理：对于先前决策所形成的状态而言，其以后的所有决策应能构成最优策略。

求最短路实际上是根据最优性原理进行以下运算：



$$f_k(s_k)=\min_{x_k\in U_k(s_k)}\left\{d_k(s_k,x_k)+f_{k+1}\left(T_k(s_k,x_k)\right)\right\}$$

最优性原理：若 s_{k+1}^* 在 s_k 到 s_6 的最优路径上，那么该路径上自 s_{k+1}^* 以后的部分一定是自 s_{k+1}^* 到 s_6 最优路径

另一方面，通过顺推的方法求出上面最短路问题也是可行的。

定义最优值函数 $\bar{f}(s)$ 为从起点到 s 的最短路程，并根据多阶段结构将其表示为

$$\bar{f}(s)=\bar{f}_k(s),\ \forall s\in S_k,\ k=1,\cdots,6$$

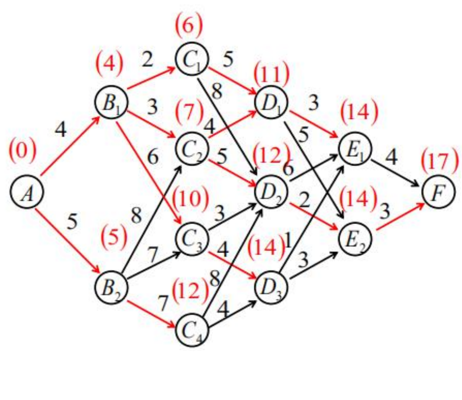
初始条件：$\bar{f}(s)=\bar{f}_1(s)=0,\ \forall s\in S_1$

由于对任意 k 成立 $S_{k+1}=T_k(s_k,U_k(S_k))$ ，所以

$$\bar{f}(s)=\min_{\substack{T(\bar{s},u)=s\\ \bar{s}\in S,u\in U(\bar{s})}}\left\{d(\bar{s},u)+\bar{f}(\bar{s})\right\}$$

$$\Leftrightarrow\ \bar{f}_{k+1}(s)=\min_{\substack{T_k(\bar{s},u)=s\\ \bar{s}\in S_k,u\in U_k(\bar{s})}}\left\{d_k(\bar{s},u)+\bar{f}_k(\bar{s})\right\},\ \forall s\in S_{k+1}$$

例题的顺序求解过程



动态规划不但可以解决一般的分阶段问题，通过对原问题形式的适当变换，还可以用来解非线性规划问题，如下面的投资分配问题：

10 万元资金，投资三个项目，收益分别为

$g_1(x_1)=4x_1,\ g_2(x_2)=9x_2,\ g_3(x_3)=2x_3^2$

如何分配投资额？

静态优化模型

$$\begin{aligned}\max\ & 4x_1+9x_2+2x_3^2\\ \text{s.t.}\ & x_1+x_2+x_3=10\\ & x_i\ge 0,\ i=1,2,3\end{aligned}$$

这是一个非线性优化问题

用动态规划方法解，首先要建立序贯决策模型。

阶段　三个决策阶段，顺序决策三个项目投资额

状态　k 阶段开始可用投资额 s_k，$s_1=10$

决策　k 阶段实际投资额 $x_k=u_k(s_k)$

允许决策集　$0\le x_k\le s_k$

阶段指标　$4x_1$，$9x_2$，$2x_3^2$

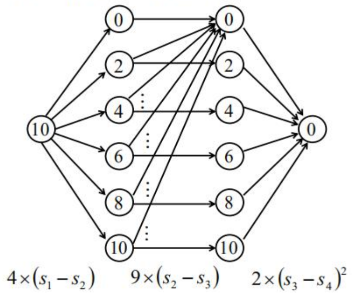
状态转移方程　$s_{k+1}=s_k-x_k$

一般情况可以将连续变量离散化，把 s_k 变成离散点集，用类似求最短路问题的离散变量动态规划。

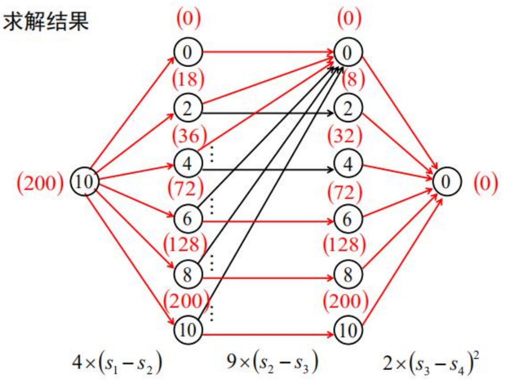
投资分配问题的离散化求解方法

将连续区间 $[0,10]$ 离散化为有限点集 $\{0,2,4,6,8,10\}$

等价于右边的求最长路问题



求解结果



还能求解整数规划问题：

背包问题

10 吨卡车，装三种货物，单位重量和价值如下

货物编号	1	2	3
单位重量（吨）	3	4	5
单位价值	4	5	6

如何装载使总价值最大？

静态优化模型　$\max\ 4x_1+5x_2+6x_3$

$\text{s.t.}\ 3x_1+4x_2+5x_3\le 10$

是一个NP难问题！　x_1,x_2,x_3 非负整数

序贯决策模型（用顺序递推解法）

阶段　三个决策阶段，顺序决策三种货物件数

状态　k 阶段之前已装货物总量 s_k，$s_1=0,\ s_i\le 10,\ \forall i$

决策　k 阶段装货件数 $x_k=u_k(s_k)$

允许决策集　$3x_1\le 10-s_1,\ 4x_2\le 10-s_2,\ 5x_3\le 10-s_3$

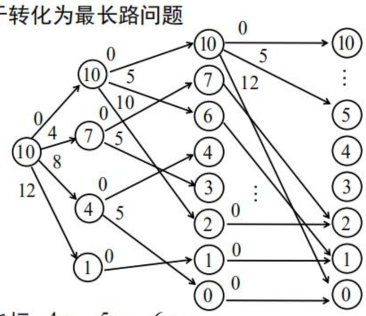
决策变量取非负整数

状态转移方程　$s_2=s_1+3x_1,\ s_3=s_2+4x_2,\ s_4=s_3+5x_3$

阶段指标　$4x_1$，$5x_2$，$6x_3$

$\bar{f}_{k-1}(s_k)$　k 阶段之前装载物品的最大价值

相当于转化为最长路问题

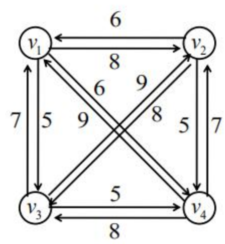


阶段指标　$4x_1$，$5x_2$，$6x_3$

状态转移方程　$s_2=s_1+3x_1,\ s_3=s_2+4x_2,\ s_4=s_3+5x_3$

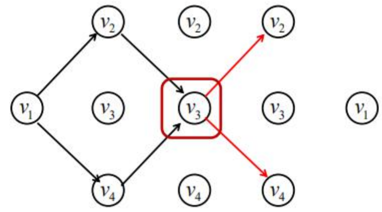
在某些问题中，直接在看似显然的节点直接直接构建状态转移关系可能难以引入无后效性，从而不能直接用动态规划求解，如下面的旅行商问题：

下图四个城市，任何两个城市间均有道路相连，路程由图中数字所示。从 v_1 出发，找出一条经过其他每个城市最终回到 v_1 的最短路线



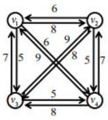
转换成多阶段决策问题

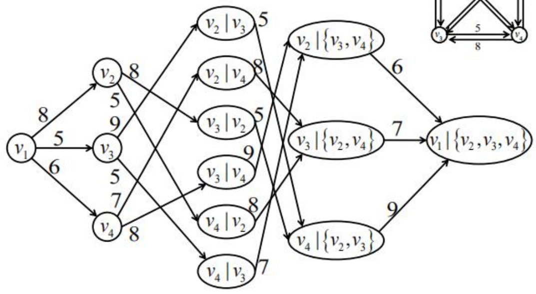
直接以城市为状态会存在后效性问题，如下图，v_3 处的容许决策取决于第二阶段状态值



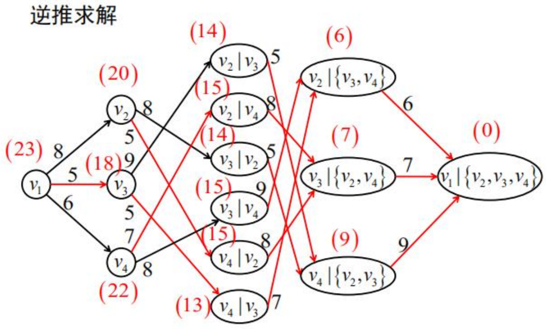
无后效性的状态设置方法

$s_{k,i}$：k 步到达城市 v_i，$1 \le k \le 3, 2 \le i \le 4$



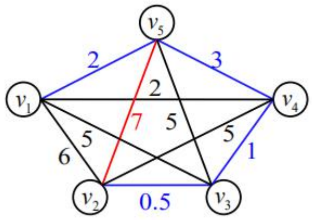


逆推求解



最优路径：$v_1 \to v_3 \to v_4 \to v_2 \to v_1$　最优路程：23

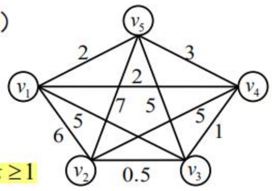
前面问题的网格都比较简单，而实际问题中常见的不定期问题往往没有确定的层状网格结构，如下图五个城市，任何两个城市间均有道路相连，往返路程一样，由图中数字所示。求每个城市到第五个城的最短路线和最短路程。



特点：直接以城市为状态不存在明显的阶段

对于这样的不定期问题我们有函数空间迭代法（值迭代法）和策略空间迭代法（策略迭代法）两种方法。

函数空间迭代法（值迭代法）



首先取 $f_1(v_i) = c_{i5}, 1 \le i \le 5$

然后按下述公式迭代

$$f_{k+1}(v_i) = \min_{1\le j\le 5}\left\{c_{ij} + f_k(v_j)\right\}, \forall k \ge 1$$

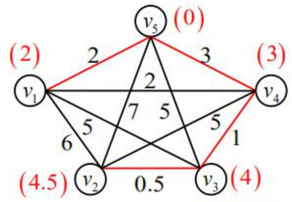
如果到某个 k 满足 $f_{k+1}(v_i) = f_k(v_i),\ \forall 1 \le i \le 5$

那么就成立 $f_k(v_i) = \min\limits_{1\le j\le 5}\left\{c_{ij} + f_k(v_j)\right\},\ \forall 1 \le i \le 5$

于是得到 $f(v_i) = f_k(v_i),\ \forall 1 \le i \le 5$

最优路线可以由最优值方程确定

根据最优路程确定最优路线



最优路径：	最优路程：
$v_1 \to v_5$	2
$v_2 \to v_3 \to v_4 \to v_5$	4.5
$v_3 \to v_4 \to v_5$	4
$v_4 \to v_5$	3

保证不定期最短路问题的值迭代法收敛的充要条件仅为：没有总路程之和小于零的回路。

策略空间迭代法（策略迭代法）

任意选取一个无回路策略 $P_1 = \{p_1(v_i), 1 \le i \le 5\}$

求解线性方程组 $f_k(v_i) = c_{i\,p_k(v_i)} + f_k(v_{p_k(v_i)}),\ \forall 1 \le i \le 4$

$$f_k(v_5) = 0$$

得 $\hat{f}_k(v_i),\ \forall 1 \le i \le 5$（无回路的策略保证方程有唯一解）

利用 $\hat{f}_k(v_i),\ \forall 1 \le i \le 5$ 确定改进的策略

$$P_{k+1} = \{p_{k+1}(v_i), 1 \le i \le 5\}$$

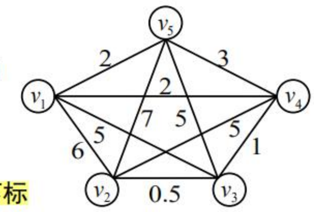
改进途经

$$c_{i\,p_{k+1}(v_i)} + \hat{f}_k(v_{p_{k+1}(v_i)}) = \min_{1\le j\le 5}\left\{c_{ij} + \hat{f}_k(v_j)\right\},\ \forall 1 \le i \le 4$$

重复上述过程直到策略不改变

其中所谓的无回路策略：

无回路策略



一个策略就是在每个点的某种决策构成的集合：

$P = \{p(v_i), 1 \le i \le 5\}$

$p(v_i)$ 表示 v_i 后面城市的下标

所有决策不在一个回路上的策略称为无回路的策略

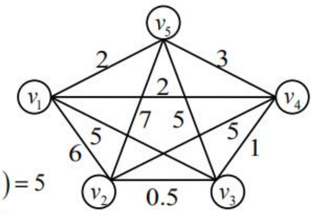
例如：下面不是无回路的策略

$$p(v_1) = 4,\ p(v_2) = 1,\ p(v_3) = 2,\ p(v_4) = 3,\ p(v_5) = 5$$

下面是无回路的策略

$$p(v_1) = 4,\ p(v_2) = 4,\ p(v_3) = 4,\ p(v_4) = 5,\ p(v_5) = 5$$

用策略迭代法解前例



$p_1(v_1) = 4,\ p_1(v_2) = 4$

$p_1(v_3) = 4,\ p_1(v_4) = 5$

$p_1(v_5) = 5$

$$\begin{aligned} f_1(v_1) &= 2 + f_1(v_4) \\ f_1(v_2) &= 5 + f_1(v_4) \\ f_1(v_3) &= 1 + f_1(v_4) \\ f_1(v_4) &= 3 + f_1(v_5) \end{aligned} \Rightarrow \begin{aligned} \hat{f}_1(v_1) &= 5 \\ \hat{f}_1(v_2) &= 8 \\ \hat{f}_1(v_3) &= 4 \\ \hat{f}_1(v_4) &= 3 \end{aligned} \quad \hat{f}_1(v_5) = 0$$

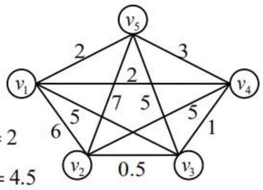
$$c_{i\,p_2(v_i)} + \hat{f}_1(v_{p_2(v_i)}) = \min_{1\le j\le 5}\left\{c_{ij} + \hat{f}_1(v_j)\right\},\ \forall 1 \le i \le 4$$

$$\Rightarrow\ p_2(v_1) = 5,\ p_2(v_2) = 3,\ p_2(v_3) = 4,\ p_2(v_4) = 5,\ p_2(v_5) = 5$$

继续迭代

$p_2(v_1)=5,\ p_2(v_2)=3$

$p_2(v_3)=4,\ p_2(v_4)=5,\ p_2(v_5)=5$



$$\begin{aligned} f_2(v_1)&=2+f_2(v_5)\\ f_2(v_2)&=0.5+f_2(v_3)\\ f_2(v_3)&=1+f_2(v_4)\\ f_2(v_4)&=3+f_2(v_5) \end{aligned} \Rightarrow \begin{aligned} \hat{f}_2(v_1)&=2\\ \hat{f}_2(v_2)&=4.5\\ \hat{f}_2(v_3)&=4\\ \hat{f}_2(v_4)&=3 \end{aligned} \quad \hat{f}_2(v_5)=0$$

$$c_{i\,p_3(v_i)}+\hat{f}_2\left(v_{p_3(v_i)}\right)=\min_{1\le j\le 5}\left\{c_{ij}+\hat{f}_2\left(v_j\right)\right\},\ \forall 1\le i\le 4$$

$\Rightarrow\quad p_3(v_1)=5,\ p_3(v_2)=3,\ p_3(v_3)=4,\ p_3(v_4)=5,\ p_3(v_5)=5$

$P_3=P_2$ 停止

最优策略

$p_3(v_1)=5,\ p_3(v_2)=3,\ p_3(v_3)=4,\ p_3(v_4)=5,\ p_3(v_5)=5$

最优路径： $v_1\to v_5$

$v_2\to v_3\to v_4\to v_5$

$v_3\to v_4\to v_5$

$v_4\to v_5$

同值迭代法结果一样