动态规划

第八章

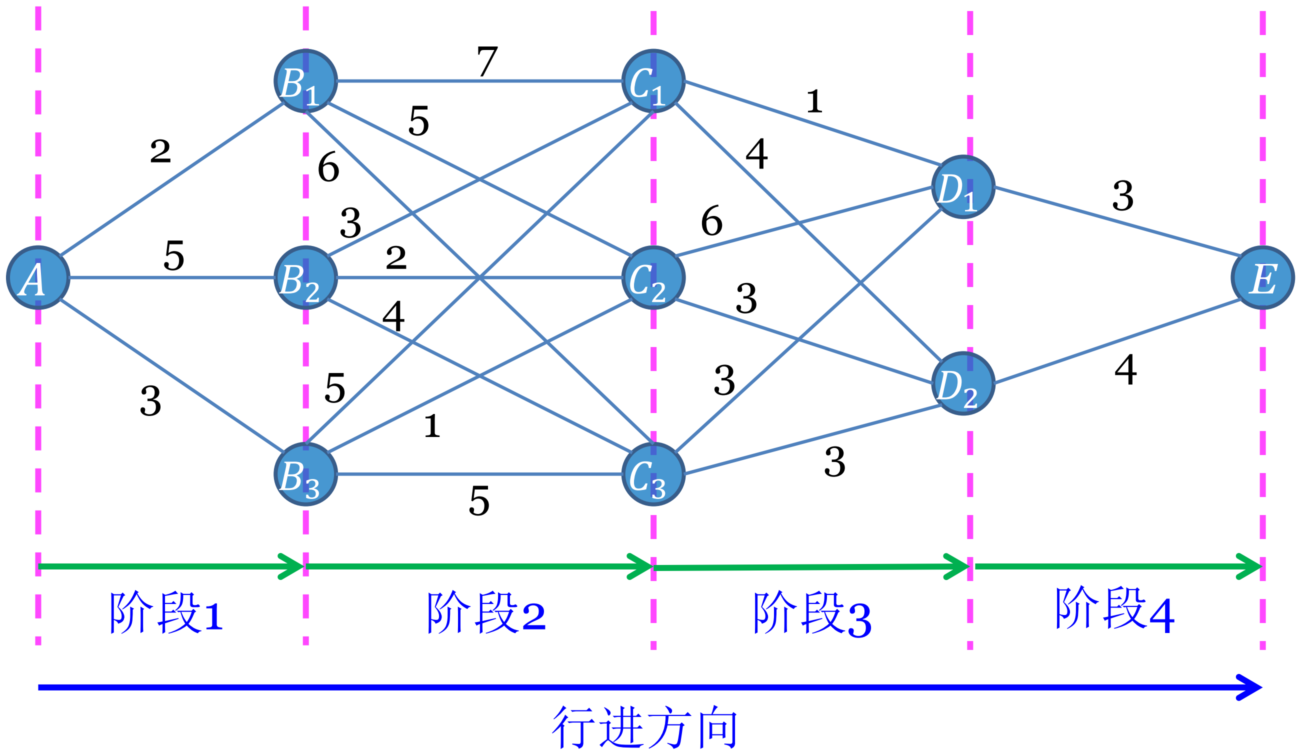
动态规划的基本概念与基本方程

第2节

Haiyan Lu
Jiangnan University

动态规划的基本概念

- 对于**例1**(**最短路线问题**)的分析
  - 该例是动态规划问题的典型的较为直观的实例。



动态规划的基本概念

- 对于**例1**(**最短路线问题**)的分析
  - 从A点到E点可分为4个**阶段**，在每一个阶段都需要根据该阶段所处的状况(**状态**，本例中为该阶段的出发位置)作出**决策**。各个阶段的决策不同，行走路线就不同。
  - 当某个阶段的始点（状态）给定时，它直接影响着后面各阶段的行进路线和整个路线的长短，而后面各阶段路线的发展不受这点以前各阶段路线的影响。
  - 问题要求：在各阶段选取一个恰当的决策，使得由这些决策构成的**决策序列**所决定的一条路线，其总路程最短。
  - 求解方法
    - **穷举法**：需计算3×3×2×1=18条路线的长度并进行比较。但是，当阶段数和各阶段的不同选择很多时，计算将变得极其繁杂，即使在计算机上计算也不现实。
    - **动态规划**：易于求出全局最优解，且可以大大减少计算工作量。

动态规划的基本概念

- 阶段
  - 把所给定问题的过程，恰当地分为若干个相互联系的阶段，以便能按一定的次序去求解。
  - 阶段变量：描述阶段的变量，常用*k*表示。
  - 阶段的划分：一般按**时间**和**空间**的自然特征来划分（要便于把问题的过程转化为多阶段决策过程）。
    - **例1**中问题可分为4个阶段来求解，*k*分别等于1,2,3,4.

动态规划的基本概念

- 状态
  - 表示每个阶段开始时所处的自然状况或客观条件，描述了所研究问题过程的状况，又称**不可控因素**。
  - **状态变量**：描述过程状态的变量，可用一个数、一组数或一个向量(**多维**情形)来描述。常用s_k表示第k阶段的状态变量，S_k表示第k阶段的状态变量的取值范围（可达状态集合）。
    - 如**例1**中，$S_3=\{C_1,C_2,C_3\}$.
  - **马尔科夫性**
    - 动态规划中的**状态**应具有如下性质：若某阶段的状态给定后，则在该阶段以后过程的发展不受该阶段以前各阶段状态的影响。即：过程的过去历史只能通过当前的状态去影响它未来的发展，当前的状态是以往历史的一个总结。该性质称为**马尔科夫性**，又称**无后效性**。**若选定的变量不具备无后效性，则不能作为状态变量。**

动态规划的基本概念

- 决策
  - 当过程处于某一阶段的某个状态时，可以做出的不同决定(或选择)，从而确定下一阶段的状态，这种决定称为**决策**。
    - 在最优控制中，**决策**也称为**控制**。
  - **决策变量：**描述决策的变量，可用一个数、一组数、或一向量来描述。
    - 常用$u_k(s_k)$表示第k阶段当状态处于s_k时的决策变量，$D_k(s_k)$来表示第k阶段状态变量为s_k时决策的取值范围(允许决策集合)。
    - 如**例1**中，$D_2(B_1)=\{C_1,C_2,C_3\}$。

动态规划的基本概念

- 策略
  - 各阶段决策确定后，整个问题的决策序列构成一个**策略**。
  - 问题的**后部子过程(*k*子过程)**：由过程的第***k***阶段开始到终止状态为止的过程。
  - ***k*子过程策略**：由k子过程中每段的决策按顺序排列组成的决策函数序列$\{u_k(s_k),\ldots, u_n(s_n)\}$，简称**子策略**，记为$p_{k,n}(s_k)$，即$p_{k,n}(s_k)= \{u_k(s_k),\ldots, u_n(s_n)\}$。
  - **全过程策略**：k=1时的k子过程策略，简称**策略**，记为$p_{1,n}(s_1)$，即$p_{1,n}(s_1)= \{u_1(s_1),\ldots, u_n(s_n)\}$。
  - **允许策略集合**：可供选择的策略的范围，用P表示。
  - **最优策略**：允许策略集合中达到最优效果的策略。

动态规划的基本概念

- 状态转移方程
  - 是从第k阶段的某一状态值到第$k+1$阶段的某一状态值的转移规律。
  - 给定第k阶段状态变量s_k的值，当该阶段决策变量$u_k(s_k)$的值确定后，第$k+1$阶段的状态变量的值s_{k+1}也就随之确定。显然，s_{k+1}是s_k和$u_k(s_k)$的函数，记为

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

或简记为 $$s_{k+1}=T_k(s_k, u_k)$$

称T_k为**状态转移函数**。

- **例1**中， $s_{k+1}=u_k(s_k)$。

动态规划的基本概念

- 指标函数
  - 是用来衡量所选定策略优劣的一种数量指标。它是定义在**全过程**和**所有后部子过程**上的确定的数量函数，常用$V_{k,n}$表示，即

$$V_{k,n} = V_{k,n}(s_k, u_k, s_{k+1}, u_{k+1}, \ldots, s_{n+1}),\ k=1,2,\ldots,n$$

  - 动态规划的指标函数应具有可分离性，且满足递推关系。即：

$$V_{k,n}(s_k, u_k, s_{k+1},\ldots, s_{n+1})=\psi_k[s_k, u_k, V_{k+1,n}(s_{k+1},\ldots, s_{n+1})]$$

动态规划的基本概念

- 指标函数
  - 最优指标函数：是指标函数的最优值，记为$f_k(s_k)$。它表示从第k阶段的状态s_k开始到第n阶段的终止状态为止的过程，采取最优策略时所得到的指标函数值。即

$$f_k(s_k)=\mathrm{opt}_{\{u_k,\dots,u_n\}}V_{k,n}(s_k, u_k, s_{k+1}, u_{k+1},\dots, s_{n+1})$$

其中“opt”为最优化（optimization），可根据题意而取min或max。

动态规划的基本概念

- 指标函数
  - 指标函数的常见形式
    - $V_{k,n}(s_k, u_k, \dots, s_{n+1}) = \sum_{j=k}^{n} v_j(s_j, u_j)$

      此时有

      $V_{k,n}(s_k, u_k, \dots, s_{n+1}) = v_k(s_k, u_k) + V_{k+1,n}(s_{k+1}, u_{k+1}, \dots, s_{n+1})$
    - $V_{k,n}(s_k, u_k, \dots, s_{n+1}) = \prod_{j=k}^{n} v_j(s_j, u_j)$

      此时有

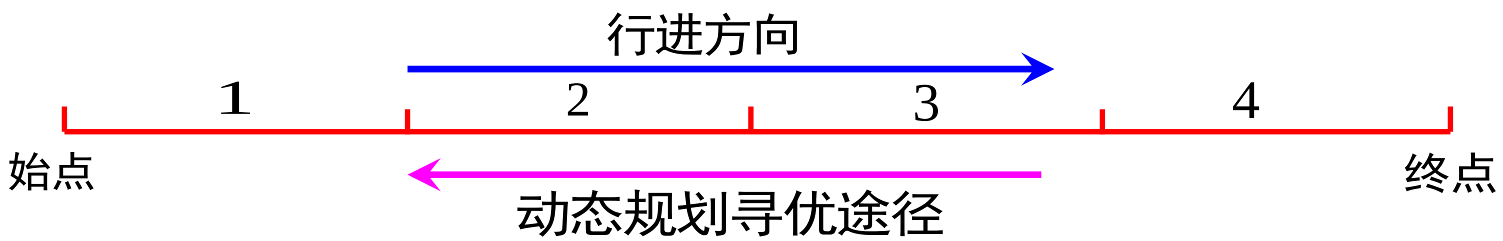
      $V_{k,n}(s_k, u_k, \dots, s_{n+1}) = v_k(s_k, u_k) \cdot V_{k+1,n}(s_{k+1}, u_{k+1}, \dots, s_{n+1})$

      这里$v_j(s_j, u_j)$表示第j阶段的**阶段指标**。

      例1中，用$d_k(s_k, u_k) = v_k(s_k, u_k)$表示由点$s_k$到点$s_{k+1} = u_k(s_k)$的距离，如$d_3(\mathrm{C}_1, \mathrm{D}_2) = 4$.

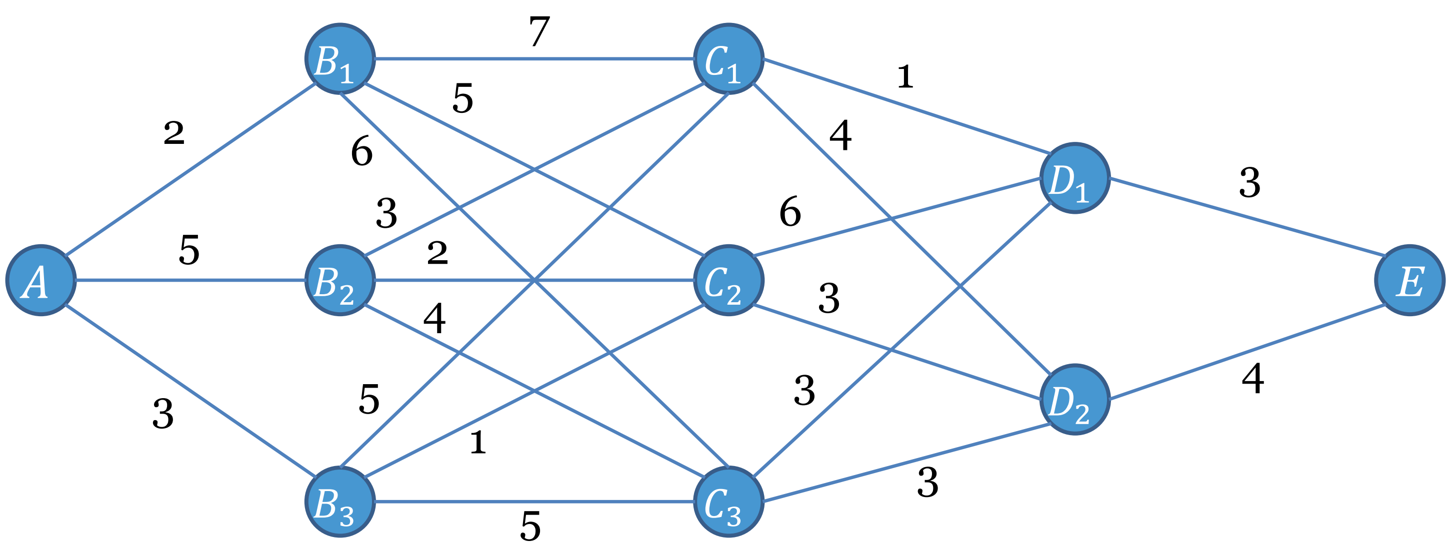
动态规划的基本思想与基本方程

- 举例说明
  - 最短路线的性质
    - 如果由起点A经过P点和H点到达终点E是一条最短路线，则由P点出发经过H点到达终点E的这条子路线，对于从P点出发到达终点E的所有可能选择的不同路线来说，必定也是最短路线。
  - 动态规划方法的求解思路
    - **从最后一段开始，由后向前逐步递推**，求出各点到E点的最短路线，最后求得从A点到E点的最短路线。即从终点逐段向始点方向寻找最短路线。



动态规划的基本思想与基本方程

- (例1)解：用动态规划的方法来求解。
  从A点到E点可分为4个阶段，n=4.
  当$k=4$时，
  $f_4(D_1)=3, u_4(D_1)=E,$
  $f_4(D_2)=4, u_4(D_2)=E.$



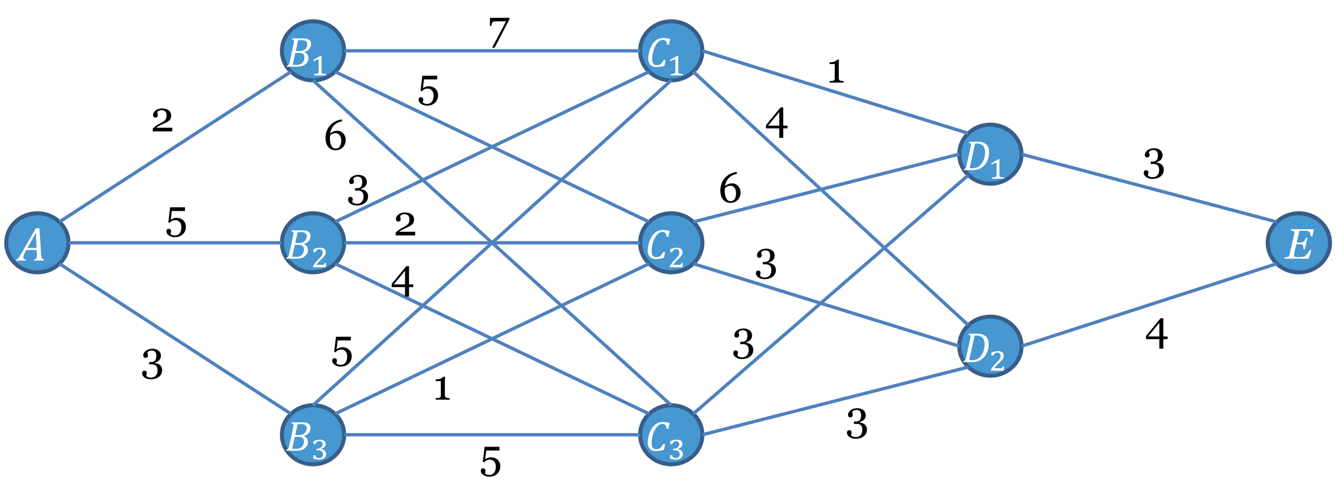
动态规划的基本思想与基本方程

- **解(续)**：当$k=3$时

$$f_3(C_1)=\min\begin{Bmatrix} d_3(C_1,D_1)+f_4(D_1)\\ d_3(C_1,D_2)+f_4(D_2)\end{Bmatrix}=\min\begin{Bmatrix}1+3\\4+4\end{Bmatrix}=4,u_3(C_1)=D_1,$$

$$f_3(C_2)=\min\begin{Bmatrix} d_3(C_2,D_1)+f_4(D_1)\\ d_3(C_2,D_2)+f_4(D_2)\end{Bmatrix}=\min\begin{Bmatrix}6+3\\3+4\end{Bmatrix}=7,u_3(C_2)=D_2,$$

$$f_3(C_3)=\min\begin{Bmatrix} d_3(C_3,D_1)+f_4(D_1)\\ d_3(C_3,D_2)+f_4(D_2)\end{Bmatrix}=\min\begin{Bmatrix}3+3\\3+4\end{Bmatrix}=6,u_3(C_3)=D_1.$$



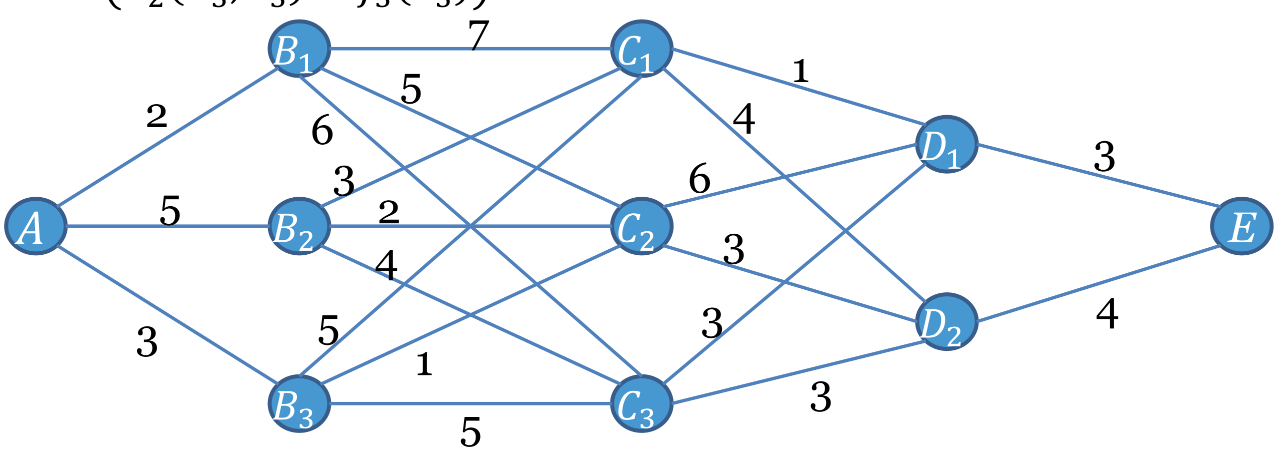
动态规划的基本思想与基本方程

- **解(续)**：当$k=2$时，

$$f_2(B_1)=\min\begin{Bmatrix} d_2(B_1,C_1)+f_3(C_1) \\ d_2(B_1,C_2)+f_3(C_2) \\ d_2(B_1,C_3)+f_3(C_3) \end{Bmatrix}=\min\begin{Bmatrix} 7+4 \\ 5+7 \\ 6+6 \end{Bmatrix}=11\,,u_2(B_1)=C_1,$$

$$f_2(B_2)=\min\begin{Bmatrix} d_2(B_2,C_1)+f_3(C_1) \\ d_2(B_2,C_2)+f_3(C_2) \\ d_2(B_2,C_3)+f_3(C_3) \end{Bmatrix}=\min\begin{Bmatrix} 3+4 \\ 2+7 \\ 4+6 \end{Bmatrix}=7\,,u_2(B_2)=C_1,$$

$$f_2(B_3)=\min\begin{Bmatrix} d_2(B_3,C_1)+f_3(C_1) \\ d_2(B_3,C_2)+f_3(C_2) \\ d_2(B_3,C_3)+f_3(C_3) \end{Bmatrix}=\min\begin{Bmatrix} 5+4 \\ 1+7 \\ 5+6 \end{Bmatrix}=8\,,u_2(B_3)=C_2.$$

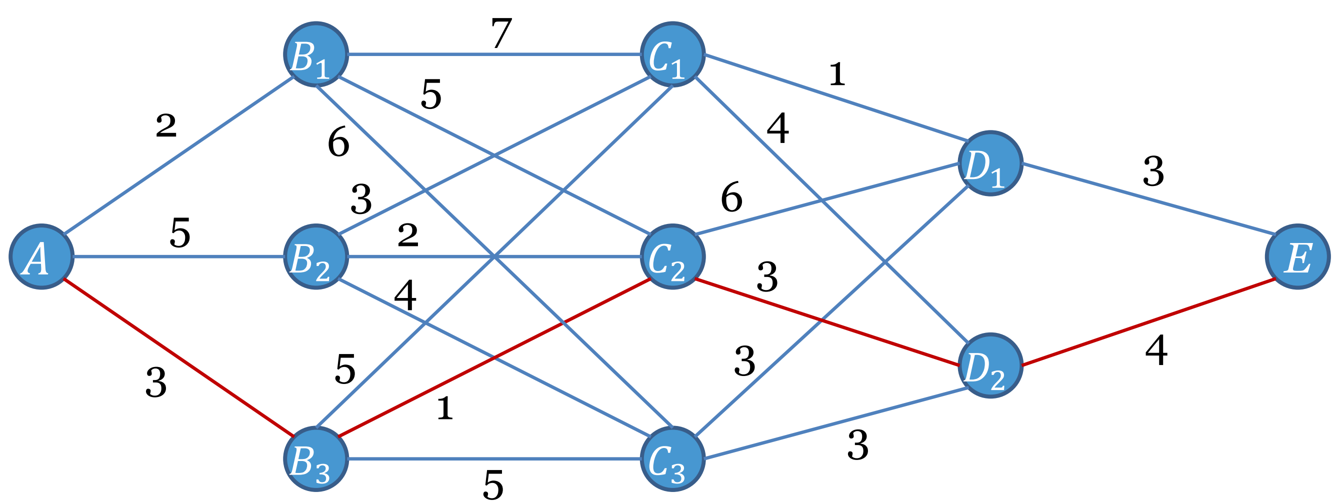


动态规划的基本思想与基本方程

- **解(续)**：当$k=1$时，

$$f_1(A)=\min\begin{Bmatrix} d_1(A,B_1)+f_2(B_1) \\ d_1(A,B_2)+f_2(B_2) \\ d_1(A,B_3)+f_2(B_3) \end{Bmatrix}=\min\begin{Bmatrix} 2+11 \\ 5+7 \\ 3+8 \end{Bmatrix}=11,u_1(A)=B_3.$$

按计算顺序反推，得最优决策序列，从而得到相应的最短路线为
$A\to B_3\to C_2\to D_2\to E$，最短路线的长度为**11**.



动态规划的基本思想与基本方程

- **例1**中，第k阶段与第$k+1$阶段的递推关系

$$\begin{cases} f_k(s_k) = \min\limits_{u_k \in D_k(s_k)} \{d_k(s_k, u_k(s_k)) + f_{k+1}(u_k(s_k))\}, k = 4,3,2,1 \\ f_5(s_5) = 0 \end{cases}$$

- 一般动态规划模型，第k阶段与第$k+1$阶段的递推关系

$$\begin{cases} f_k(s_k) = \operatorname*{opt}\limits_{u_k \in D_k(s_k)} \{v_k(s_k, u_k(s_k)) + f_{k+1}(u_k(s_k))\}, k = n, n-1, \dots, 1 \\ f_{n+1}(s_{n+1}) = 0 \quad \text{（边界条件）} \end{cases}$$

或

$$\begin{cases} f_k(s_k) = \operatorname*{opt}\limits_{u_k \in D_k(s_k)} \{v_k(s_k, u_k(s_k)) \cdot f_{k+1}(u_k(s_k))\}, k = n, n-1, \dots, 1 \\ f_{n+1}(s_{n+1}) = 1 \quad \text{（边界条件）} \end{cases}$$

这种递推关系式称为**动态规划的基本方程**。

(边界条件$f_{n+1}(s_{n+1})$的值要根据具体问题的条件来确定)

动态规划的基本思想与基本方程

- 动态规划的基本思想
  - 正确地写出基本方程(基本递推关系式和边界条件)。
    - 将问题的过程分成几个相互联系的阶段，恰当地选取状态变量和决策变量，定义最优指标函数，从而把一个大问题转化为一族同类型的子问题，然后逐个求解。
    - 从边界条件开始，逐段递推寻优，在每一个子问题的求解中，均利用了其前面子问题的最优化结果，依次进行，最后一个子问题所得的最优解，就是整个问题的最优解。
  - 在多阶段决策过程中，动态规划方法既把当前阶段和未来各段分开，有把当前效益与未来效益结合起来考虑，是**着眼全局**来选取每段的决策，因此与每段的局部最优选择往往不同。

动态规划的基本思想与基本方程

- 动态规划的两种求解方法
  - 逆序解法
    - 从问题的最后一个阶段开始，逆着多阶段决策的实际过程逐段递推寻优。
  - 顺序解法
    - 从问题的最初阶段开始，顺着多阶段决策的实际过程逐段递推寻优。
  - 备注
    - 动态规划的求解较多采用逆序解法。
    - **例1**的上述解法为逆序解法。

动态规划的基本思想与基本方程

- 几点说明
  - **逆序解法**和**顺序解法本质上是相同的**，两者只是对行进方向或始点终点的看法上相反。
    - 即：用动态规划方法求最优解时，无论是采用逆序解法还是顺序解法，都是**在行进方向规定后，均要逆着这个规定的行进方向，逐段递推寻优**。
  - 动态规划方法与穷举法相比有以下**优点**：
    - 大大减少了计算量。
    - 丰富了计算结果。
      - 求出的不仅仅是一个最优策略，而是一族最优策略，对于许多实际问题很有用，有利于帮助分析所得结果。

Haiyan Lu
Jiangnan University

动态规划的基本思想与基本方程

- 动态规划**逆序解法**的基本方程
  （假设指标函数为各阶段指标之和）

$$\begin{cases} f_k(s_k) = \underset{u_k \in D_k(s_k)}{\text{opt}} \{v_k(s_k, u_k) + f_{k+1}(s_{k+1})\}, k = n, n-1, \ldots, 1 \\ f_{n+1}(s_{n+1}) = 0 \quad \text{（边界条件）} \end{cases}$$

其中$s_{k+1} = T_k(s_k, u_k)$.

- 求解过程
  - 根据边界条件，从$k = n$开始，由后向前逆推，从而逐步求得各段的最优决策和相应的最优值，最后求出$f_1(s_1)$时，即得到整个问题的最优解。

动态规划的基本思想与基本方程

- 动态规划**顺序解法**的基本方程
  （假设指标函数为各阶段指标之和）

$$\begin{cases} f_k(s_{k+1}) = \underset{u_k \in D_k^r(s_{k+1})}{\text{opt}} \{v_k(s_{k+1}, u_k) + f_{k-1}(s_k)\}, k = 1,2, \dots, n \\ f_0(s_1) = 0 \quad \text{（边界条件）} \end{cases}$$

其中$s_k = T_k^r(s_k, u_k(s_{k+1}))$, k和s_k的定义与逆序解法中相同。

- 求解过程
  - 根据边界条件，从$k = 1$开始，由前向后顺推，从而逐步求得各段的最优决策和相应的最优值，最后求出$f_n(s_{n+1})$时，即得到整个问题的最优解。

Thank you!

谢谢！