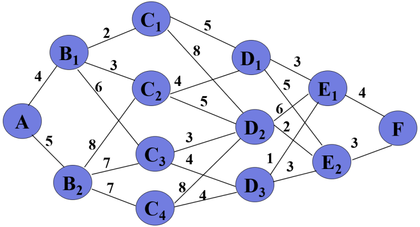
动态规划的基本概念和基本原理

修贤超

https://xianchaoxiu.github.io

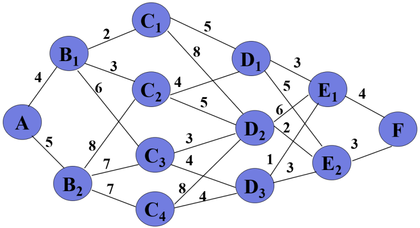
动态规划的基本概念

- 阶段: 将所给问题的过程，按时间或空间特征分解成相互联系的阶段，以便按次序求每阶段的解。记 k 为阶段变量



- $k=1,\ A\to B\ (B_1,B_2)$
- $k=2,\ B\to C\ (C_1,C_2,C_3,C_4)$
- $k=3,\ C\to D\ (D_1,D_2,D_3)$
- $k=4,\ D\to E\ (E_1,E_2)$
- $k=5,\ E\to F$

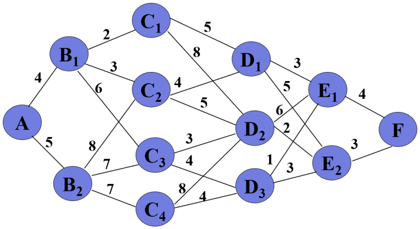
- **状态**: 每个阶段开始时的客观条件，描述了研究问题的状况。记 s_k 为第 k 阶段的状态变量，S_k 为状态变量 s_k 的取值集合
- 当某阶段状态给定以后，在这阶段以后过程的发展不受这段以前各段状态的影响，这称为**无后效性**



- 第一阶段状态为 A，状态变量 s_1 的集合为 $S_1 = \{A\}$
- $S_2 = \{B_1, B_2\}$, $S_3 = \{C_1, C_2, C_3, C_4\}$, $S_4 = \{D_1, D_2, D_3\}$, $S_5 = \{E_1, E_2\}$

动态规划的基本概念

- 决策: 取定各阶段的状态后，就可以做出不同的决定，从而确定下一阶段的状态。记 $u_k(s_k)$ 为第 k 阶段当状态为 s_k 时的**决策变量**，$D_k(s_k)$ 为第 k 阶段从状态 s_k 出发的**允许决策集合**



  - 从第二阶段的状态 B_1 出发，可选择下一阶段的 C_1, C_2, C_3，即其允许决策集合为 $D_2(B_1) = \{C_1, C_2, C_3\}$
  - 如果我们决定选择 C_3，则 $u_2(B_1) = C_3$

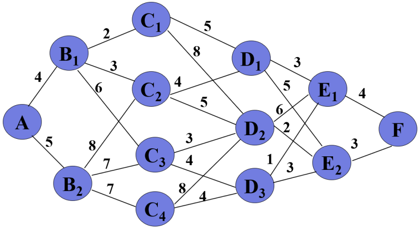
动态规划的基本概念

- 策略: 由所有各阶段组成的决策函数序列

$$p_{1,n}\{u_1(s_1), u_2(s_2), \ldots, u_n(s_n)\} \in P_{1,n}$$

- 最优策略: 使整个问题达到最优效果的策略
- 状态转移方程: 本阶段状态与上一阶段状态和上一阶段决策的关系

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



从 k 阶段到 $k+1$ 阶段的状态转移方程为 $s_{k+1} = u_k(s_k)$

动态规划的基本概念

- **指标函数**: 衡量所选定策略优劣的数量指标
- **阶段指标函数**: 第 k 阶段，从状态 s_k 出发，采取决策 u_k 时的效益，记 $d(s_k, u_k)$
- **过程指标函数**: 一个 n 段决策过程，从 1 到 n 叫做问题的原过程。对于任意一个给定的 k，从第 k 阶段到第 n 阶段的过程称为原过程的一个后部子过程
- 例如，$V_{1,n}(s_1, p_{1,n})$ 表示初始状态为 s_1 采取策略 $p_{1,n}$ 时原过程的指标函数值。$V_{k,n}(s_k, p_{k,n})$ 表示在第 k 阶段状态为 s_k 采取策略 $p_{k,n}$ 时，后部子过程的指标函数值

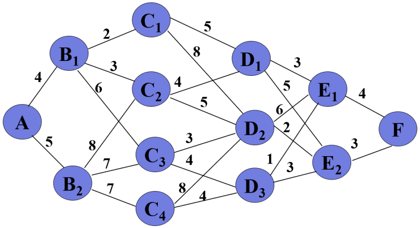
动态规划的基本概念

- **最优指标函数**: 指标函数的最优值
- 例如，$f_k(s_k)$ 表示从第 k 阶段状态 s_k 采用最优策略 $p_{k,n}$ 到过程终止时的最佳效益值。$f_1(s_1)$ 表示从第 1 阶段状态 s_1 采用最优策略 $p_{1,n}$ 到过程终止时的最佳效益值
- 最优指标函数 $f_k(s_k)$ 与 $V_{k,n}(s_k, p_{k,n})$ 的关系

$$
\begin{aligned}
f_k(s_k) &= V_{k,n}(s_k, p^*_{k,n}) \\
&= \mathrm{opt}_{p_{k,n} \in P_{k,n}} V_{k,n}(s_k, p_{k,n})
\end{aligned}
$$

动态规划的基本概念

- 例如指标函数是距离，第 2 阶段，状态为 B_1 时 $d(B_1, C_2)$ 表示由 B_1 出发，采用决策到下一段 C_2 点间的距离。



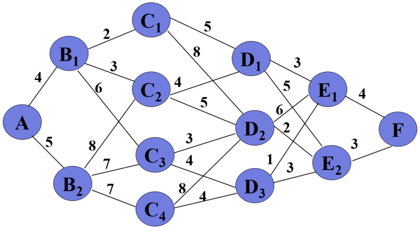
  - $V_{2,5}(B_1)$ 表示从 B_1 到 F 的距离
  - $f_2(B_1)$ 表示从 B_1 到 F 的最短距离
  - $f_1(A)$，即从 A 到终点 F 的最短距离

动态规划的基本思想

- 从过程的最后一段开始，用逆序递推方法求解，逐步求出各段各点到终点 F 的最短路线，最后求得 A 点到 F 点的最短路线。

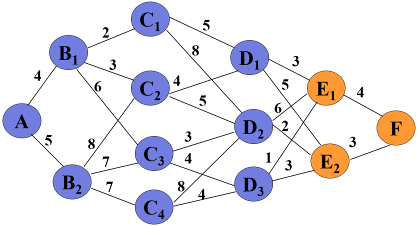
- 当 $k=5$ 时: 状态变量 s_5 可取两种状态 E_1, E_2, 它们到 F 点的路长分别为

$$f_5(E_1)=4,\ f_5(E_2)=3$$



动态规划的基本思想

- 当 $k=4$ 时: 状态变量 s_4 可取三种状态 D_1, D_2, D_3，这是经过一个中途点到达终点 F 的两级决策问题



  - 从 D_1 到 F，其路径为 $D_1 \to E_1 \to F$，相应决策为方程 $u_4^*(D_1) = E_1$

$$\begin{aligned} f_4(D_1) &= \min\{d(D_1, E_1) + f_5(E_1), d(D_1, E_2) + f_5(E_2)\} \\ &= \min\{3+4, 5+3\} = 7 \end{aligned}$$

  - 从 D_2 到 F，其路径为 $D_2 \to E_2 \to F$，相应决策为方程 $u_4^*(D_2) = E_2$
  - 从 D_3 到 F，其路径为 $D_3 \to E_1 \to F$，相应决策为方程 $u_4^*(D_3) = E_1$

动态规划的基本思想

- 当 $k=3$ 时: 有
  - $f_3(C_1)=7$, $u_3^*(C_1)=D_1$
  - $f_3(C_2)=7$, $u_3^*(C_2)=D_2$
  - $f_3(C_3)=7$, $u_3^*(C_3)=D_2$
  - $f_3(C_4)=7$, $u_3^*(C_4)=D_3$
- 当 $k=2$ 时: 有
  - $f_2(B_1)=13$, $u_2^*(B_1)=C_2$
  - $f_2(B_1)=15$, $u_2^*(B_1)=C_3$

动态规划的基本思想

- 当 $k=1$ 时: 只有一个状态点 A，因有

$$
\begin{aligned}
f_1(A) &= \min\{d(A,B_1)+f_2(B_1), d(A,B_2)+f_2(B_2)\} \\
&= \min\{4+13, 5+15\} = 17
\end{aligned}
$$

- 从 A 到 F 的最短距离为 17
- 按计算顺序反推可得最优决策序列 $\{u_k\}$，即

$$
\begin{aligned}
&u_1^*(A)=B_1,\ u_2^*(B_1)=C_2,\ u_3^*(C_2)=D_2 \\
&u_4^*(D_2)=E_2,\ u_5^*(E_2)=F
\end{aligned}
$$

- 最优路线为

$$
A \to B_1 \to C_2 \to D_2 \to E_2 \to F
$$

动态规划的基本思想

- 从本例的计算过程可以看出，在求解的各个阶段，都利用了第 k 段和第 $k+1$ 段的如下关系

$$\begin{cases} f_k(s_k) = \min\{d_k(s_k, u_k) + f_{k+1}(s_{k+1})\}, \ k = 5, 4, 3, 2, 1 \\ f_6(s_6) = 0 \end{cases}$$

- 上式称为动态规划的基本方程
- $f_6(s_6) = 0$ 称为边界条件

动态规划的基本思想

- 将多阶段决策过程划分阶段，恰当的选取状态变量、决策变量及定义最优指标函数，从而将问题化为一族同类型的子问题，然后逐个求解
- 求解时从边界条件开始，逆（或顺）过程行进方向，逐段递推寻优。在每一个子问题求解时，都要使用它前面已求出的子问题的最优结果，最后一个子问题的最优解就是整个问题的最优解
- 既将当前一段与未来各段分开，又将当前效益与未来效益结合起来考虑的一种最优化方法，因此每段的最优决策选取时从全局考虑的，与该段的最优选择一般是不同的
- 动态规划基本方程

$$\begin{cases} f_k(s_k) = \mathrm{opt}_{u_k \in D_k(s_k)} \{v_k(s_k, u_k) + f_{k+1}(s_{k+1})\}, & k = n, n-1, \ldots, 1 \\ f_{n+1}(s_{n+1}) = 0 \end{cases}$$

动态规划的最优化原理

- 作为整个过程的最优策略具有如下性质：不管在此最优策略上的某个状态以前的状态和决策如何，对该状态而言，以后所有的决策必定构成最优子策略
- 对最短路问题而言，从最短路上任一点到终点的部分道路（最短路上的子路）也一定是从该点到终点的最短路

小结

- 基本概念
  - 阶段 k
  - 状态 s_k
  - 决策 u_k
  - 策略 $p_{1,n}$
  - 状态转移方程 $s_{k+1} = T_k(s_k, u_k)$
  - 指标函数 $f_k(s_k)$
- 逆序递推法
- 标号法
- 课后作业: P217, 习题 7.1 (逆序法)

$Q\&A$

$Thank\ you!$

感谢您的聆听和反馈