六. 动态规划
(Dynamic Programming)

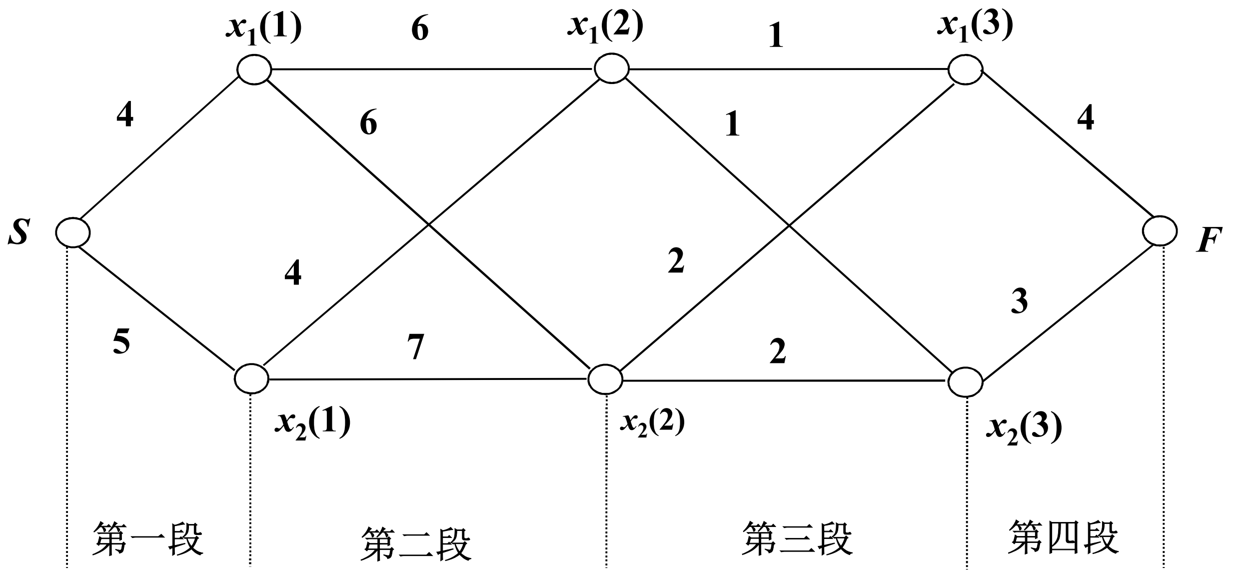
引 言

- 动态规划是最优化理论中的一种重要方法，由美国数学家**Richard Bellman** 于二十世纪**50**年代中期提出。
- 最初动态规划是被用来解决多级决策过程最优化问题，如上一章所述，这一问题等同于离散时间系统的最优控制问题。
- 后来这一方法被推广到连续时间系统最优控制问题，成为解最优控制问题的重要方法之一。
- 动态规划是解最优控制问题的一种“方法”，而不是一种“算法”，其核心是将一个复杂的多级决策问题转化为一系列简单的单级决策问题。
- 动态规划的理论基础是**Bellman**最优性原理，据此可以得到离散时间系统最优控制的一些理论结果，并建立相应的递推优化计算公式。

6.1 多级决策问题

(1) 最短（优）行车路线问题

考虑如图所示的道路网络，其中各段道路的里程如图中数字所示。一车从S站出发，前往F站，要求寻找一条路程最短的行车路线。



最短行车路线多段决策路网

这是一个三段决策问题。在S站、$x_1(1)$或$x_2(1)$站、$x_1(2)$或$x_2(2)$站都要作二选一的决策，即在每一站可供决策至多两个。

对这样一个三段决策问题，要找到最短路径决策，首先考虑采用穷举法求解。

即将所有可能的行车路线全部列出，如$S \to x_1(1) \to x_2(2) \to x_1(3) \to F$或$S \to x_2(1) \to x_1(2) \to x_1(3) \to F$等。所有可能的行车路线共有$2^{(n-1)} = 2^3 = 8$条，其中$n$＝4表示分段总数，最后一段不需作决策。要求出8条路线中的最短路线，需要计算每条路线的长度共8次，比较各条路线的长度7次。以计算机基本运算次数计，即需做加法运算8╳3＝24（即$2^{(n-1)} \times (n-1)$）次，比较运算7（即$2^{(n-1)} - 1$）次。

下面考虑用动态规划法解此问题。

以J表示行车路程（或称代价），并从最后一段开始计算J的值。

从$x_1(3) \to F$，$J = J[x_1(3)] = 4$

从$x_2(3) \to F$，$J = J[x_2(3)] = 3$

此段虽然路径已定，也可以视为作了一选一的决策。

再考察从倒数第二段（即图中第三段）开始。

从$x_1(2) \to F$，有两条路线：

$$x_1(2) \to x_1(3) \to F,\quad J=J[x_1(2)]=1+J[x_1(3)]=1+4=5$$

$$x_1(2) \to x_2(3) \to F,\quad J=J[x_1(2)]=1+J[x_2(3)]=1+3=4$$

有最小代价$J^*[x_1(2)]=4$

从$x_2(2) \to F$，也有两条路线：

$$x_2(2) \to x_1(3) \to F,\quad J=J[x_2(2)]=2+J[x_1(3)]=2+4=6$$

$$x_2(2) \to x_2(3) \to F,\quad J=J[x_2(2)]=2+J[x_2(3)]=2+3=5$$

有最小代价$J^*[x_2(2)]=5$

考虑从倒数第三段（即图中第二段）开始。

从$x_1(1) \to F$，也只考虑两条路线：

$$x_1(1) \to x_1(2) \to F,\quad J=J[x_1(1)]=6+J^*[x_1(2)]=6+4=10$$

$$x_1(1) \to x_2(2) \to F,\quad J=J[x_1(1)]=6+J^*[x_2(2)]=6+5=11$$

有最小代价$J^*[x_1(1)]=10$

从$x_2(1) \to F$，考虑：

$$x_2(1) \to x_1(2) \to F,\quad J=J[x_2(1)]=4+J^*[x_1(2)]=4+4=8$$

$$x_2(1) \to x_2(2) \to F,\quad J=J[x_2(1)]=7+J^*[x_2(2)]=7+5=12$$

有最小代价$J^*[x_2(1)]=8$

最后考虑从倒数第四段开始，即全程决策。

从$S \to F$，考虑：

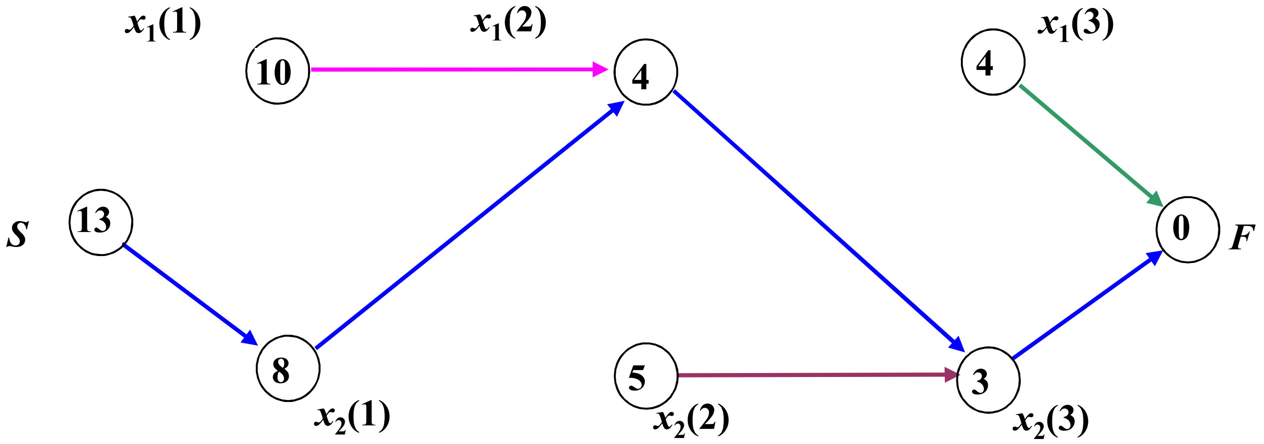
$S \to x_1(1) \to F$， $J = J[S]=4 + J^*[x_1(1)]=4+10=14$

$S \to x_2(1) \to F$， $J = J[S]=5 + J^*[x_2(1)]=5+8=13$

有最小代价$J^*[S]=13$。

则最优决策路线为$S \to x_2(1) \to x_1(2) \to x_2(3) \to F$，最小代价（最短行车路线）为$J^* = J^*[S]=13$。

以上即为采用动态规划方法的求解过程。其中，作决策（比较运算）5次（2*(n-2)＋1＝5），加法10次（4*(n-2)＋2＝10）。



从各站开始的最短行车路线

(2) 动态规划的特点

I. 与穷举法相比，动态规划法的计算工作量大大减少，当段数、决策数都增加时差别尤为明显；

II. 将一个复杂的、难以求解的多级决策问题转化为一系列简单的、易于求解的单级决策问题来处理——数学上称为“不变嵌入原理”（将一条最优行车路线问题嵌入到求一系列最优行车问题中）；

III. 最优路线的后部子路线及其相应决策序列也是自中间站开始的最优路线和最优决策序列——即最优性原理，为动态规划的核心和理论基础；

IV. 动态规划的运算规则是由终点开始采用逆推方法求解最优决策。