动态规划模型的建立与求解

修贤超

https://xianchaoxiu.github.io

例 1

- 某公司有资金 10 万元, 若投资于项目 i $(i=1,2,3)$ 的投资额为 x_i $(i=1,2,3)$ 时，其收益分别为

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

问应如何分配投资数额才能使总收益最大?

- 划分阶段: 引入时段的概念，即将投资项目排序。首先考虑对项目 1 投资，然后对项目 2 和项目 3 投资
- 选择状态变量: 通常选择随递推过程累计的量或者按某种规律变化的量作为状态变量，即把每阶段可供使用的资金定为状态变量 s_k，初始状态 $s_1=10$
- 确定决策变量: 通常可以把决策变量 u_k 定为原静态问题中的变量 x_k，即设 $u_k=x_k$ $(k=1,2,3)$

动态规划建模举例

- 写出状态转移方程: 记u_1 为可分配用于第一种项目的最大资金，则当第一阶段 $(k=1)$ 时，有

$$\begin{cases} s_1 = 10 \\ u_1 = x_1 \end{cases}$$

当第二阶段 $(k=2)$ 时，状态变量 s_2 为可投资余下两个项目的资金，即

$$\begin{cases} s_2 = s_1 - u_1 \\ u_2 = x_2 \end{cases}$$

- 一般地，当第 k 阶段时

$$\begin{cases} s_k = s_{k-1} - u_{k-1} \\ u_k = x_k \end{cases}$$

动态规划建模举例

- **阶段**: $k = 1, 2, 3$
- **状态变量**: s_k 表示第 k 段可以投资于第 k 项到第 3 个项目的资金
- **决策变量**: x_k 表示决定给第 k 个项目投资的资金
- **状态转移方程**: $s_{k+1} = s_k - x_k$
- **指标函数**: $V_{k,3} = \sum_{i=k}^{3} g_i(x_i)$
- **最优指标函数**: $f_k(s_k)$ 表示当可投资金为 s_k 时，投资第 k 个至第 3 个项目所得的最大收益
- **基本方程**:
$$\begin{cases} f_k(s_k) = \max\limits_{0 \le x_k \le s_k} \{g_k(x_k) + f_{k+1}(s_{k+1})\}, \quad k = 3, 2, 1 \\ f_4(s_4) = 0 \end{cases}$$

动态规划建模步骤

- **第一步：划分阶段**
  - 分析题意，识别问题的多阶段特性，按时间或空间的先后顺序适当地划分为满足递推关系的若干阶段，对非时序的静态问题要人为地赋予“时段”概念
- **第二步：正确选择状态变量**
  - 状态变量 s_k 首先应描述研究过程的演变特征，其次应包含到达这个状态前的足够信息并具有无后效性，即到达这个状态前的过程的决策将不影响到该状态以后的决策
  - 状态变量应具有可知性，即规定的状态变量之值可通过直接或间接的方法测知
  - 状态变量可以是离散的，也可以是连续的
  - 建模时，一般从与决策有关的条件中，或者从问题的约束条件中去寻找状态变量，选择随递推过程累计的量或者按某种规律变化的量作为状态变量

动态规划建模步骤

- **第三步：写出状态转移方程**
  - 决策变量 u_k 是对过程进行控制的手段，复杂问题中决策变量可以是多维的向量，它的取值可能离散，也可能连续
  - 每阶段的允许决策集合相当于线性规划问题中的约束条件
  - 写出状态转移方程 $s_{k+1} = T(s_k, u_k)$
- **第四步：确定基本方程**
  - 指标函数 $V_{k,n}$
  - 最优指标函数 $f_k(s_k)$
  - 阶段指标 $d(s_k, u_k)$
  - 列出最优指标函数的递推关系及边界条件（即基本方程）

逆序解法

- 寻优的方向与多阶段决策过程的实际行进方向相反，从最后一段开始计算逐段前推，求得全过程的最优策略，称为逆序解法
- 设已知初始状态为 s_1，并假定最优值函数 $f_k(s_k)$ 表示第 k 阶段的初始状态为 s_k，从 k 阶段到 n 阶段所得到的最大效益
- 从第 n 阶段开始，则有

$$f_n(s_n) = \max_{u_n \in D_n(s_n)} d_n(s_n, u_n)$$

其中 $D_n(s_n)$ 表示状态 s_n 所确定的第 n 阶段的允许决策集合。解此一维极值问题，就得到最优解 $u_n = u_n(s_n)$ 和最优值 $f_n(s_n)$

- 在第 $n-1$ 阶段，有

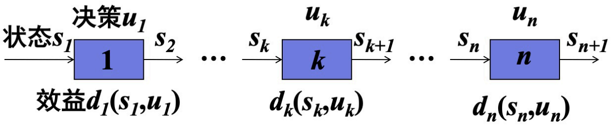
$$f_{n-1}(s_{n-1}) = \max_{u_{n-1}\in D_{n-1}(s_{n-1})} [d_{n-1}(s_{n-1}, u_{n-1}) * f_n(s_n)]$$

其中，$s_n = T_{n-1}(s_{n-1}, u_{n-1})$。解此一维极值问题，就得到最优解 $u_{n-1} = u_{n-1}(s_{n-1})$ 和最优值 $f_{n-1}(s_{n-1})$。这里 $*$ 表示 $+$ 和 $\times$

- 在第 k 阶段，有

$$f_k(s_k) = \max_{u_k\in D_k(s_k)} [d_k(s_k, u_k) * f_{k+1}(s_{k+1})]$$

其中，$s_{k+1} = T_k(s_k, u_k)$。解此一维极值问题，得到最优解 $u_k = u_k(s_k)$ 和最优值 $f_k(s_k)$



逆序解法

- 如此类推，直到第 1 阶段，有

$$f_1(s_1) = \max_{u_1 \in D_1(s_1)} [d_1(s_1, u_1) * f_2(s_2)]$$

其中，$s_2 = T_1(s_1, u_1)$。通过计算得到最优解 $u_1 = u_1(s_1)$ 和最优值 $f_1(s_1)$

- 由于 s_1 已知，那么 $u_1 = u_1(s_1)$ 和 $f_1(s_1)$ 是确定的，从而 $s_2 = T_1(s_1, u_1)$ 也可确定，依此类推就可逐步确定出每阶段的决策及效益

例 2

- 用逆序解法求解问题

$$\max\ z = x_1 \cdot x_2^2 \cdot x_3$$
$$\text{s.t.}\ \begin{cases} x_1 + x_2 + x_3 = c\ (c > 0) \\ x_i \geq 0\ (i = 1, 2, 3) \end{cases}$$

- 按问题的变量个数划分阶段，看作一个三阶段决策问题
- 设状态变量为 s_1, s_2, s_3, s_4，并记 $s_1 = c$
- 取问题中的变量为决策变量 x_1, x_2, x_3，各阶段指标函数按乘积方式结合
- 令最优值函数 $f_k(s_k)$ 表示为第 k 阶段的初始状态为 s_k，从 k 阶段到 3 阶段所得到的最大值

例 2

- **状态变量**: s_1, s_2, s_3, s_4
- **决策变量**: x_1, x_2, x_3
- **阶段指标函数**: $d_k(s_k, u_k)$, x_1, x_2^2, x_3
- **过程指标函数**: $V_{1,3} = x_1 \times x_2^2 \times x_3$
- **分析**: 设 $s_3 = x_3$, $s_3 + x_2 = s_2$, $s_2 + x_1 = s_1 = c$, 于是有 $s_3 = x_3$, $0 \leq x_2 \leq s_2$, $0 \leq x_1 \leq s_1 = c$。采用逆序解法，从后向前依次有

$$f_3(s_3) = \max_{x_3 \in D_3(s_3)} d_3(s_3, u_3) = \max(x_3) = s_3$$

及最优解 $x_3^* = s_3$

例 2

- (续)

$$\begin{aligned} f_2(s_2) &= \max_{x_2 \in D_2(s_2)} \left[d_2(s_2, u_2) \cdot f_3(s_3)\right] \\ &= \max_{0 \le x_2 \le s_2} \left[x_2^2 \cdot f_3(s_3)\right] \\ &= \max_{0 \le x_2 \le s_2} \left[x_2^2 \cdot (s_2 - x_2)\right] \\ &= \max_{0 \le x_2 \le s_2} \; h_2(s_2, x_2) \end{aligned}$$

由 $\frac{dh_2}{dx_2} = 2x_2 s_2 - 3x_2^2 = 0$ **得** $x_2 = \frac{2}{3}s_2$, $x_2 = 0$ **(舍去)**

又 $\frac{d^2 h_2}{dx_2^2} = 2s_2 - 6x_2$**，而** $\frac{d_2^h}{dx_2^2}\Big|_{x_2=\frac{2}{3}s_2} = -2s_2 < 0$**，故** $x_2 = \frac{2}{3}s_2$ **为极大值点，所以**

$$f_2(s_2) = \frac{4}{27}s_2^3, \; x_2^* = \frac{2}{3}s_2$$

例 2

- (续)

$$
\begin{aligned}
f_1(s_1) &= \max_{x_1 \in D_1(s_1)} \left[d_1(s_1, u_1) \cdot f_2(s_2) \right] \\
&= \max_{0 \le x_1 \le s_1} \left[x_1 \cdot f_2(s_2) \right] \\
&= \max_{0 \le x_1 \le s_1} \left[x_1 \cdot \frac{4}{27} (s_1 - x_1)^3 \right] \\
&= \max_{0 \le x_1 \le s_1} \; h_1(s_1, x_1)
\end{aligned}
$$

同理，利用微分法易知

$$
f_1(s_1) = \frac{1}{64} s_1^4, \;\; x_1^* = \frac{1}{4} s_1
$$

例 2

- (续) 由于已知 $s_1 = c$，故按计算的顺序反推可得

$$x_1^* = \frac{1}{4}s_1,\ f_1(c) = \frac{1}{64}c^4$$
$$x_2^* = \frac{2}{3}s_2 = \frac{1}{2}c,\ f_2(s_2) = \frac{1}{16}c^3$$
$$x_3^* = \frac{1}{4}c,\ f_3(s_3) = \frac{1}{4}c$$

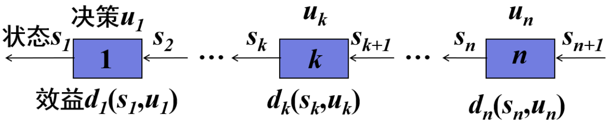
因此，目标函数的最大值为 $f_1(c) = \frac{1}{64}c^4$

顺序解法

- 寻优的方向与多阶段决策过程的实际行进方向相同，从问题的第 1 阶段开始逐段向后递推，计算后一阶段要用到前一阶段的求优结果，最后一段计算的结果就是全过程的最优结果，称为顺序解法
- 已知终止状态用顺序解法与已知初始状态用逆序解法在本质上没有区别。它相当于把实际的起点视为终点，实际的终点视为起点，按逆序解法进行

顺序解法

- 将逆序解法时的 n 阶段决策过程的箭头倒转过来即可，将 s_{k+1} 看作输入，将 s_k 看作输出



- 此时状态变换是逆序解法时的逆变换，记为 $s_k = T_k(s_{k+1}, u_k)$，即由 s_{k+1} 和 u_k 确定 s_k
- 设已知终止状态为 s_{n+1}，并假定最优值函数 $f_k(s)$ 表示第 k 阶段末的结束状态为 s，从 1 阶段到 k 阶段所得到的最大效益

顺序解法

- 从第 1 阶段开始，则有

$$f_1(s_2) = \max_{u_1 \in D_1(s_1)} d_1(s_1, u_1)$$

其中 $s_1 = T_1(s_2, u_1)$，得到最优解 $u_1 = u_1(s_2)$ 和最优值 $f_1(s_2)$

- 在第 2 阶段，有

$$f_2(s_3) = \max_{u_2 \in D_2(s_2)} [d_2(s_2, u_2) * f_1(s_2)]$$

- 如此类推，直到第 n 阶段，有

$$f_n(s_{n+1}) = \max_{u_n \in D_n(s_n)} [d_n(s_n, u_n) * f_{n-1}(s_n)]$$

- 由于终止状态 s_{n+1} 已知，故 $u_n = u_n(s_{n+1})$ 和 $f_n(s_{n+1})$ 是确定的。按过程的相反顺序推算下去，就可逐步求出每段的决策及效益

例 2

- 用顺序解法求解下面问题

$$\max\ z = x_1 \cdot x_2^2 \cdot x_3$$

$$\text{s.t.}\ \begin{cases} x_1 + x_2 + x_3 = c\ (c > 0) \\ x_i \geq 0\ (i = 1, 2, 3) \end{cases}$$

- **状态变量**: s_1, s_2, s_3, s_4
- **决策变量**: x_1, x_2, x_3
- **阶段指标函数**: $d_k(s_k, u_k),\ x_1, x_2^2, x_3$
- **过程指标函数**: $V_{1,3} = x_1 \times x_2^2 \times x_3$

例 2

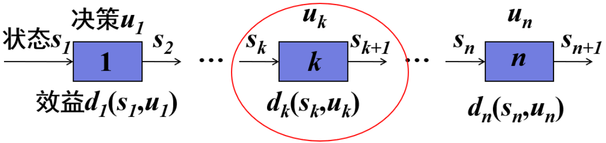
- **分析:** 设$s_4=c$，令最优值函数值 $f_k(s_{k+1})$ 表示第 k 阶段末的结束状态为 s_{k+1}，从 1 阶段到 k 阶段的最大值。设 $s_3+x_3=s_4=c$，$s_2+x_2=s_3$，$s_2=x_1$。状态转移方程为 $s_k=T_k(s_{k+1},u_k)$，有

$$s_2=x_1,\ 0\le x_2\le s_3,\ 0\le x_3\le s_4=c$$

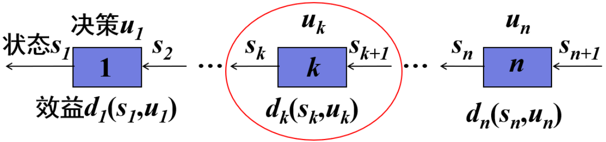
- 用顺序解法，从前向后依次有
  - $f_1(s_2)=\max\limits_{x_1=s_2}(x_1)=s_2$, 最优解 $x_1^*=s_2$
  - $f_2(s_3)=\max\limits_{0\le x_2\le s_3}[x_2^2\cdot f_1(s_2)]=\frac{4}{27}s_3^3$, 最优解 $x_2^*=\frac{2}{3}s_3$
  - $f_3(s_4)=\max\limits_{0\le x_3\le s_4}[x_3\cdot f_2(s_3)]=\frac{1}{64}s_4^4$, 最优解 $x_3^*=\frac{1}{4}s_4$
- 由于已知 $s_4=c$，故可得最优解 $x_1^*=\frac{1}{4}c$，$x_2^*=\frac{1}{2}c$，$x_3^*=\frac{1}{4}c$，相应的最大值为 $\max\ z=\frac{1}{64}c^4$

逆序与顺序解法的区别

- 状态转移方式不同
  - 逆序解法：状态 s_k 到 s_{k+1} 的状态转移方程为 $s_{k+1} = T_k(s_k, u_k)$



  - 顺序解法：状态 s_{k+1} 到 s_k 的状态转移方程为 $s_k = T_k(s_{k+1}, u_k)$



序与顺序解法的区别

- 指标函数的定义不同
  - 逆序解法：定义最优指标函数 $f_k(s_k)$ 表示第 k 段从状态 s_k 出发，到终点后部子过程最优效益值，$f_1(s_1)$ 是整体最优函数值
  - 顺序解法：定义最优指标函数 $f_k(s_{k+1})$ 表示第 k 段从起点到状态 s_{k+1} 的前部子过程最优效益值，$f_n(s_{n+1})$ 是整体最优函数值

逆序与顺序解法的区别

- 基本方程形式不同——当指标函数为阶段指标和形式
  - 在逆序解法中，$v_{k,n} = \sum\limits_{j=k}^{n} v_j(s_j, u_j)$，则基本方程为

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

  - 在顺序解法中，$v_{1,k} = \sum\limits_{j=1}^{k} v_j(s_{j+1}, u_j)$，基本方程为

$$\begin{cases} f_k(s_{k+1}) = \underset{u_k \in D_k}{\text{opt}} \{v_k(s_{k+1}, u_k) + f_{k-1}(s_k)\} \; (k = 1, \ldots, n) \\ f_0(s_1) = 0 \end{cases}$$

逆序与顺序解法的区别

- 基本方程形式不同——当指标函数为阶段指标积形式
  - 在逆序解法中，$v_{k,n} = \prod_{j=k}^{n} v_j(s_j, u_j)$，则基本方程为

$$\begin{cases} f_k(s_k) = \underset{u_k \in D_k}{\text{opt}} \{v_k(s_k, u_k) \cdot f_{k+1}(s_{k+1})\} \ (k = n, \ldots, 1) \\ f_{n+1}(s_{n+1}) = 1 \end{cases}$$

  - 在顺序解法中，$v_{1,k} = \prod_{j=1}^{k} v_j(s_{j+1}, u_j)$，基本方程为

$$\begin{cases} f_k(s_{k+1}) = \underset{u_k \in D_k}{\text{opt}} \{v_k(s_{k+1}, u_k) \cdot f_{k-1}(s_k)\} \ (k = 1, \ldots, n) \\ f_0(s_1) = 1 \end{cases}$$

例 3

- 某公司有资金 10 万元, 若投资于项目 i $(i=1,2,3)$ 的投资额为 x_i $(i=1,2,3)$ 时，其收益分别为

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

- 问应如何分配投资数额才能使总收益最大

例 3——逆序解法

- 当 $k=3$ 时, $f_3(s_3)=\max\limits_{0\le x_3\le s_3} 2x_3^2=2s_3^2$，最优解 $x_3^*=s_3$
- 当 $k=2$ 时,

$$\begin{aligned} f_2(s_2) &= \max_{0\le x_2\le s_2}[9x_2+f_3(s_3)] \\ &= \max_{0\le x_2\le s_2}[9x_2+2(s_2-x_2)^2] \\ &= \max_{0\le x_2\le s_2} h_2(s_2,x_2) \end{aligned}$$

由 $\frac{dh_2}{dx_2}=9-4(s_2-x_2)=0$ 得 $x_2=s_2-\frac{9}{4}$。又 $\frac{d^2h_2}{dx_2^2}=4>0$，故 $x_2=s_2-\frac{9}{4}$ 为极小值点

极大值只可能在 $[0,s_2]$ 端点取得: $f_2(0)=2s_2^2$, $f_2(s_2)=9s_2$

- 当 $f_2(0)=f_2(s_2)$ 时, 解得 $s_2=\frac{9}{2}$
- 当 $s_2>\frac{9}{2}$, $f_2(0)>f_2(s_2)$，有 $x_2^*=0$; $s_2<\frac{9}{2}$, $f_2(0)<f_2(s_2)$，则 $x_2^*=s_2$

例 3——逆序解法

- 当 $k=1$ 时, 有

$$f_1(s_1) = \max_{0\le x_1\le s_1} [4x_1 + f_2(s_2)]$$

- 当 $f_2(s_2) = 9s_2$ 时，有

$$f_1(10) = \max_{0\le x_1\le 10} [4x_1 + 9(s_1 - x_1)]$$
$$= \max_{0\le x_1\le s_1} [9s_1 - 5x_1]$$

易知，$x_1^* = 0,\ f_1(s_1) = 9s_1$
但此时 $s_2 = s_1 - x_1 = 10 - 0 = 10 > \frac{9}{2}$，与 $s_2 < \frac{9}{2}$ 矛盾，所以舍去

例 3——逆序解法

- 当$f_2(s_2) = 2s_2^2$ 时，有

$$f_1(10) = \max_{0\le x_1\le 10} \left[4x_1 + 2(s_1 - x_1)^2\right]$$
$$= \max_{0\le x_1\le 10} h_1(s_1, x_1)$$

由 $\frac{dh_1}{dx_1} = 4x_1 + 4(s_1 - x_1)(-1) = 0$，解得 $x_1 = s_1 - 1$，而 $\frac{d^2h_1}{dx_1^2} = 1 > 0$，所以 $x_1 = s_1 - 1$ 是极小点

比较 $[0, 10]$ 两个端点，$x_1 = 0$, $f_1(0) = 200$; $x_1 = 10$, $f_1(10) = 40$，所以 $x_1^* = 0$

- 再由状态转移方程顺推 $s_2 = s_1 - x_1^* = 10 - 0 = 10$，因为 $s_2 > \frac{9}{2}$，所以 $x_2^* = 0$, $s_3 = s_2 - x_2^* = 10 - 0 = 10$，于是 $x_3^* = s_3 = 10$
- 最优投资方案为全部资金投于第 3 个项目，可得最大收益 200 万元

例 3——顺序解法

- 阶段划分和决策变量的设置同逆序解法，令状态变量 s_{k+1} 表示可用于第 1 到第 k 个项目投资的金额，则有

$$s_4 = 10,\ s_3 = s_4 - x_3,\ s_2 = s_3 - x_2,\ s_1 = s_2 - x_1$$

  即态转移方程为 $s_k = s_{k+1} - x_k$

- 令最优指标函数 $f_k(s_{k+1})$ 表示第 k 段投资额为 s_{k+1} 时第 1 到第 k 个项目所获得的最大收益。此时顺序解法的基本方程为

$$\begin{cases} f_k(s_{k+1}) = \max\limits_{0 \le x_k \le s_{k+1}} [g_k(s_k) + f_{k-1}(s_k)]\ (k = 1, 2, 3) \\ f_0(s_1) = 0 \end{cases}$$

例 3——顺序解法

- 当 $k=1$ 时, 有

$$
\begin{aligned}
f_1(s_2) &= \max_{0\le x_1\le s_2} [g_k(x_1) + f_{k-1}(s_k)] \\
&= \max_{0\le x_1\le s_2} [4x_1] \\
&= 4s_2
\end{aligned}
$$

最优解 $x_1^* = s_2$

- 当 $k=2$ 时, 有

$$
\begin{aligned}
f_2(s_3) &= \max_{0\le x_2\le s_3} [9x_2 + f_1(s_2)] \\
&= \max_{0\le x_2\le s_3} [9x_2 + 4(s_3 - x_2)] \\
&= \max_{0\le x_2\le s_3} [5x_2 + 4s_3] \\
&= 9s_3
\end{aligned}
$$

最优解 $x_2^* = s_3$

例 3——顺序解法

- 当 $k=3$ 时, 有

$$
\begin{aligned}
f_3(s_4) &= \max_{0\le x_3\le s_4}\ [2x_3^2+f_2(s_3)] \\
&= \max_{0\le x_3\le s_4}\ [2x_3^2+9(s_4-s_3)] \\
&= \max_{0\le x_3\le s_4}\ h(s_4,x_3)
\end{aligned}
$$

由 $\frac{dh}{dx_3}=4x_3-9=0$ 得 $x_3=\frac{9}{4}$

又 $\frac{d^2h}{dx_3^2}=4>0$，故 $x_3=\frac{9}{4}$ 为极小值点

极大值只可能在 $[0,s_4]$ 端点取得: $x_3=0, f_3(0)=90$, $x_3=10, f_3(10)=200$,
所以 $x_3^*=10$

例 3——顺序解法

- 再由状态转移方程逆推
  - $s_3 = x_1^* = 10 - x_3^* = 0,\ x_2^* = 0$
  - $s_2 = s_3 - x_2^* = 0,\ x_1^* = 0$
- 优投资方案为全部资金投于第 3 个项目，可得最大收益 200 万元
- 与逆序解法结果相同，但解法简单（为什么）

课堂练习 1

- 用动态规划方法解下面问题

$$\max\ z = 4x_1 - x_2^2 + 2x_3^2 + 12$$

$$\text{s.t.}\ \begin{cases} 3x_1 + 2x_2 + x_3 \leq 9 \\ x_i \geq 0\ (i = 1, 2, 3) \end{cases}$$

小结

- 动态规划建模步骤
- 动态规划基本解法
  - 逆序解法
  - 顺序解法
- 逆序解法和顺序解法的区别
  - 状态转移方式不同
  - 指标函数的定义不同
  - 基本方程形式不同
- 课后作业: P219, 习题 7.8

Q&A

Thank you!

感谢您的聆听和反馈