

多阶段决策过程

以 $[\mathbf{a}, T(\mathbf{a}, \mathbf{x})]$ 记一个多阶段决策过程, 其中
 \mathbf{a} 称为状态,
 $T(\mathbf{a}, \mathbf{x})$ 称为状态转移变换,
 \mathbf{x} 称为决策向量.

我们有以下序列,

$$\mathbf{a}_1 = \mathbf{a}_0, \quad (\text{an initial state})$$

$$\mathbf{a}_{i+1} = T(\mathbf{a}_i, \mathbf{x}_i), \quad i = 1, 2, \dots$$

该过程中, 决策向量 \mathbf{x}_i 极大化准则函数
 $R(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N; \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N).$

Bellman 最优性原理

无论系统的过去状态与决策如何，对于前面的决策所形成的当前状态与决策而言，最优决策的余下的各个分量仍构成一个最优策略。

动态规划

由 Richard Bellman (1957) 定义的动态规划可简单的写作

$$\left\{ \begin{array}{l} f_N(\mathbf{a}) = \max_{g_N(\mathbf{a}, \mathbf{x}) \leq 0} r_N(\mathbf{a}, \mathbf{x}) \\ f_n(\mathbf{a}) = \max_{g_n(\mathbf{a}, \mathbf{x}) \leq 0} \{ r_n(\mathbf{a}, \mathbf{x}) + f_{n+1}(T(\mathbf{a}, \mathbf{x})) \} \\ n \leq N - 1. \end{array} \right.$$