

一个用动态规划方法求解 离散系统最优控制的简单例子

贾晔

物理学院

2005 年 12 月 27 日

问题

- 一个简单的生产-库存模型，状态方程为

$$x(k+1) = x(k) + u(k) - s(k),$$

其中 $x(k)$, $u(k)$, $s(k)$ 分别是库存量和单位时间段内的生产量及销售量, k 是离散化了的时间。

- 一定时间内生产和储存的总费用为

$$J = \sum_{k=0}^{N-1} [au^2(k) + x(k)].$$

- 给定 $s(k)$, 并假定有

$$x(0) = x(N) = b,$$

寻找一个合适的控制函数 $u(k)$, 使总费用 J 最小。

最优控制问题

■ 状态方程

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0.$$

■ 性能指标

$$J[\boldsymbol{u}(\cdot), \boldsymbol{x}_0, t_0] = K(\boldsymbol{x}(t_f), t_f) + \int_{t_0}^{t_f} L(\boldsymbol{x}(t), \boldsymbol{u}(t), t) dt.$$

■ 允许控制

$$\mathcal{U}_{[t_0, t_f]}.$$

- 最优控制：求解满足状态方程的一个控制 $\boldsymbol{u}^*(t) \in \mathcal{U}_{[t_0, t_f]}$ ，使得系统的性能指标 $J[\boldsymbol{u}(\cdot), \boldsymbol{x}_0, t_0]$ 最优（取最大或最小值）。

离散系统的最优控制问题

■ 状态方程

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k), k), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0.$$

■ 性能指标

$$J[\boldsymbol{u}(\cdot), \boldsymbol{x}_0, 0] = K(\boldsymbol{x}(N), N) + \sum_{k=0}^{N-1} L(\boldsymbol{x}(k), \boldsymbol{u}(k), k).$$

■ 允许控制

$$\mathcal{U}.$$

■ 最优控制：求解满足状态方程的一个控制序列

$\boldsymbol{u}^*(k) \in \mathcal{U}$ ，使得系统的性能指标 $J[\boldsymbol{u}(\cdot), \boldsymbol{x}_0, 0]$ 最优（取最大或最小值）。

典型的最优控制问题

- 最短时间控制
- 最少燃料控制
- 最低能量控制
- 线性二次型性能指标最优控制
- 非线性性能指标最优控制
- 最优线性调节器
- 最优线性伺服器
- 镇定问题
- 跟踪问题

最优控制问题的主要解法

- 古典变分法。
- L. C. Pontryagin 的极值原理。
- R. E. Bellman 的动态规划法。

大部分最优控制问题不能求出解析解，或者解析解的形式比较复杂。

我们的问题

- 我们要解决的问题就是一个离散系统的最优控制问题。
- 使用动态规划的解法。

计算机算法中的动态规划

- 用于求解具有最优性质的问题
- 把原问题分解为若干不独立的子问题
- 每个子问题只求解一次
- 保存求得的子问题的解
- 求出一个最优解，而不是所有最优解

动态规划算法的常规步骤

- 刻画最优解的性质和结构特征
- 递归的定义最优值
- 自底向上的计算出最优值
- 利用求得的信息构造最优解

看一个例子——矩阵的连乘

- 利用乘法结合率选择合适的计算顺序，将矩阵连乘的计算量将至最低。
- 计算的顺序用“完全加括号”的矩阵连乘表述。
- “完全加括号”：
 - 单个的矩阵是完全加括号的；
 - 如果矩阵连乘式 A 和 B 都是完全加括号的，那么 (AB) 也是完全加括号的。
- 一次乘法的计算量： $m \times n$ 的矩阵与 $n \times t$ 的矩阵相乘，计算量为 mnt （次数乘）。

完全加括号的方式数

- 给 n 个矩阵连乘式完全加括号的总方式数为

$$P(n) = \begin{cases} 1 & n = 1 \\ \sum_{k=1}^{n-1} P(k)P(n-k) & n > 1 \end{cases}$$

- Catalan 数

$$P(n+1) = C(n) = \frac{1}{n+1} \binom{2n}{n} = \Omega\left(\frac{4^n}{n^{3/2}}\right)$$

- 穷举搜索显然不可行。

Step 1: 分析最优解的结构

- 最优解可以表示为

$$((A_1 \cdots A_k)(A_{k+1} \cdots A_n))$$

- 计算量为 $(A_1 \cdots A_k)$ 的计算量与 $(A_{k+1} \cdots A_n)$ 的计算量之和再加上它们相乘的计算量
- 关键特征：计算 $(A_1 \cdots A_k)$ 和 $(A_{k+1} \cdots A_n)$ 的顺序也应该是最优的
- 原因在于他们分别的计算量与它们相乘的计算量这三者是互相独立的
- 最优子结构性质

Step 2: 建立递归关系

- 如果计算 $(A_i \cdots A_j)$ 所需的最少计算次数为 $m(i, j)$ ，那么原问题的最优值为 $m(1, n)$
- 利用最优子结构性质，有

$$m(i, j) = \begin{cases} 0 & i = j \\ \min_{i \leq k < j} \{m(i, k) + m(k + 1, j) + p_{i-1}p_kp_j\} & i < j \end{cases}$$

Step 3: 计算最优值

- 不同的子问题个数只有 $\theta(n^2)$ 个
- 每个子问题只计算一次
- 自底向上的计算
- 计算结果保存在表中
- 以后用到时只需查表
- 重叠子问题性质

```

1: MatrixChain ( $p$ )
2:  $n \leftarrow \text{length}[p] - 1$ 
3: for  $i \leftarrow 1$  to  $n$  do
4:    $m[i, i] \leftarrow 0$ 
5: for  $l \leftarrow 2$  to  $n$  do
6:   for  $i \leftarrow 1$  to  $n - l + 1$  do
7:      $j \leftarrow i + l - 1$ 
8:      $m[i, j] \leftarrow \text{inf}$ 
9:     for  $k \leftarrow i$  to  $j - 1$  do
10:       $q \leftarrow m[i, k] + m[k + 1, j] + p_{i-1}p_kp_j$ 
11:      if  $q < m[i, j]$  then
12:         $m[i, j] \leftarrow q$ 
13:         $s[i, j] \leftarrow k$ 
14: return  $m$  and  $s$ 

```


离散系统最优控制中的动态规划方法

- 最优化原理：如果 $\mathbf{u}^*(0), \mathbf{u}^*(1), \dots, \mathbf{u}^*(N-1)$ 是一个离散系统的最优控制序列，那么 $\mathbf{u}^*(k), \mathbf{u}^*(k+1), \dots, \mathbf{u}^*(N-1)$ 是这个系统对应于初始状态

$$\mathbf{x}(k) = \mathbf{f}(\mathbf{x}^*(k-1), \mathbf{u}^*(k-1), k-1)$$

的一个最优控制。

递推关系

- 记 $J_i^*(\mathbf{x}(i)) = J^*[\mathbf{u}(\cdot), \mathbf{x}(i), i]$ ，由最优化原理，可以建立递推关系

$$J_i^*(\mathbf{x}(i)) = \begin{cases} K_N(\mathbf{x}(N)) & i = N \\ \min_{\mathbf{u}(i)} \{L_i(\mathbf{x}(i), \mathbf{u}(i)) + J_{i+1}^*(\mathbf{x}(i+1))\} & i < N \end{cases}$$

- 所求最优控制的性能指标即为 $J_0^*(\mathbf{x}_0)$ 。
- 然而按照前面的做法，需要在每个时间 i 时，求出所有可能的 $\mathbf{x}(i)$ 对应的 $J_i^*(\mathbf{x}(i))$ 。
- 通常情况下，每个时间 i 时可取的状态 $\mathbf{x}(i)$ 太多，这样的计算不可行。

- A set of navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other slide controls.

回到我们的问题

- 简单起见，以季度为单位，考虑一年时间内的总花费

$$J = \sum_{k=0}^3 [au^2(k) + x(k)],$$

其中 $a = 0.01$ 。

- $x(k)$ 满足状态方程

$$x(k+1) = x(k) + u(k) - s(k)$$

- 四个季度的订货量分别为

$$s(0) = 550, s(1) = 750, s(2) = 500, s(3) = 1200.$$

- 约束条件

$$x(0) = x(4) = 100, \quad x(1), x(2), x(3) \geq 0; \quad u(k) \geq 0.$$

递推关系

$$J_i^*(x_i) = \begin{cases} 0 & i = 4 \\ J_i(x_i, u_i^*(x_i)) & i < 4 \end{cases},$$

其中

$$J_i(x_i, u_i) = au_i^2 + x_i + J_{i+1}^*(x_i + u_i - s(i)),$$

$u_i^*(x_i)$ 满足

$$\left. \frac{\partial J_i(x_i, u_i)}{\partial u_i} \right|_{u_i=u_i^*(x_i)} = 0.$$

递推: $k = 3$

- $k = 3$ 时的最优控制应由约束条件

$$x_4 = x_3 + u_3 - s(3) = b$$

解出, 得

$$u_3^*(x_3) = b + s(3) - x_3 = 1300 - x_3.$$

- 故

$$J_3^*(x_3) = a(u_3^*)^2 + x_3 = ax_3^2 - 25x_3 + 16900.$$

递推： $k = 2$

■ $k = 2$ 时

$$J_2(x_2, u_2) = au_2^2 + x_2 + J_3^*(x_2 + u_2 - s(2))$$

■ 由 $J_2(x_2, u_2)$ 对 u_2 偏导为零解得

$$u_2^*(x_2) = 875 - \frac{x_2}{2}.$$

■ 故

$$J_2^*(x_2) = au_2^{*2} + x_2 + J_3^*(x_2 + u_2^* - s(2)) = \frac{1}{2} (ax_2^2 - 33x_2 + 33175).$$

递推: $k = 1$

■ $k = 1$ 时

$$J_1(x_1, u_1) = au_1^2 + x_1 + J_2^*(x_1 + u_1 - s(1))$$

■ 由 $J_1(x_1, u_1)$ 对 u_1 偏导为零解得

$$u_1^*(x_1) = 800 - \frac{x_1}{3}.$$

■ 故

$$J_1^*(x_1) = au_1^{*2} + x_1 + J_2^*(x_1 + u_1^* - s(1)) = \frac{a}{3}x_1^2 - 15x_1 + 22175.$$

递推: $k = 0$

■ $k = 0$ 时

$$J_0(x_0, u_0) = au_0^2 + x_0 + J_1^*(x_0 + u_0 - s(0))$$

■ 由 $J_0(x_0, u_0)$ 对 u_0 偏导为零解得

$$u_0^*(x_0) = 700 - \frac{x_0}{4}.$$

■ 故

$$J_0^*(x_0) = au_0^{*2} + x_0 + J_1^*(x_0 + u_0^* - s(0)) = \frac{a}{4}x_0^2 - 13x_0 + 24900.$$

得到的最优解

■ 最少费用

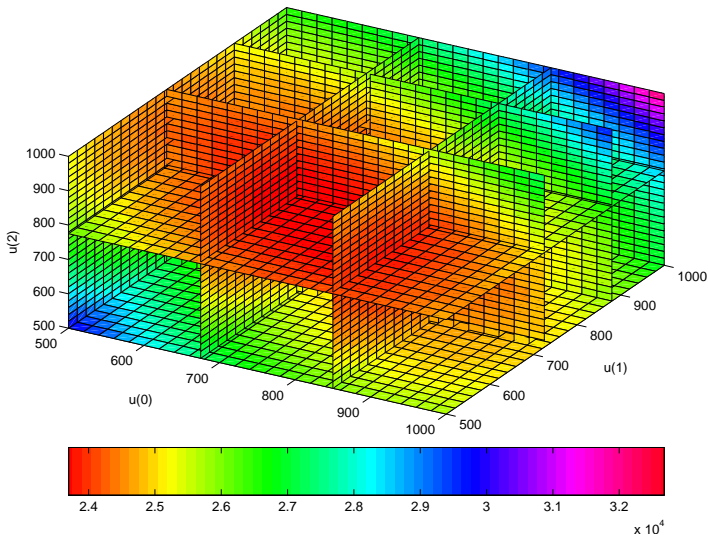
$$J^* = J_0^*(x_0) = 23625.$$

■ 最优控制及相应的系统状态

$$\begin{aligned}x^*(0) &= 100, & u^*(0) &= 675; \\x^*(1) &= 225, & u^*(1) &= 725; \\x^*(2) &= 200, & u^*(2) &= 775; \\x^*(3) &= 475, & u^*(3) &= 825; \\x^*(4) &= 100.\end{aligned}$$

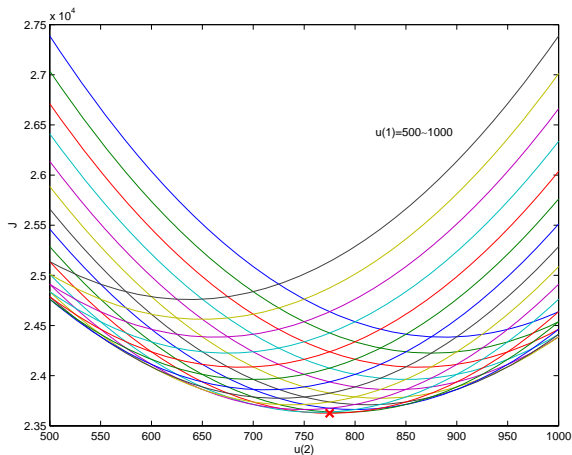
四维切片图

总费用在 $u_0-u_1-u_2$ 空间的四维切片图 ($\sum u_i = \sum s_i$)



u_1 、 u_2 对 J 的影响

固定 $u_0 = u_0^*$, u_1 取不同固定值时，总费用 J 随 u_2 的变化 ($\sum u_i = \sum s_i$)



两种动态规划形式之间的联系

- 本质是相同的。
 - 利用最优化原理或者最优子结构性质把问题分解为若干子问题，每个子问题仍然是一个最优性问题。
 - 子问题之间互相联系，可以建立递推关系。
 - 根据递推关系求出最优值或最优性能指标，然后根据求出的最优值或最优性能指标反向递推求出最优解或最优控制。
- 同一理论的两两种形式。

两者的不同！

其它的动态规划方法

■ 连续系统的动态规划

■ Hamilton-Jaccobi-Bellman 方程

$$-\frac{\partial J^*}{\partial t} = L\left(\mathbf{x}(t), \mathbf{u}^*\left(\mathbf{x}(t), \frac{\partial J^*}{\partial \mathbf{x}}, t\right), t\right) + \frac{\partial J^*}{\partial \mathbf{x}} \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}^*\left(\mathbf{x}(t), \frac{\partial J^*}{\partial \mathbf{x}}, t\right), t\right),$$

$$J^*(\mathbf{x}(t_f), t_f) = K(\mathbf{x}(t_f), t_f).$$

■ $J^*(\mathbf{x}(t), t)$ 的一阶非线性偏微分方程。

■ 微分动态规划

- 将性能指标函数局部展开到二阶，用动态规划方法求解，不断迭代逼近。