第12卷 第4期 郑 州 工 学 院 学 报 Vol.12 No.4
1991年 12 月 Journal of Zhengzhou Institute of Technology Dec. 1991

策略迭代法在水库优化调度中的应用*

贺北方 石 宾 罗贯英

(郑州工学院水环系)

摘 要: 本文针对综合利用水库的特点，建立了水库优化调度的马尔可夫决策规划模型．文中论述了策略迭代的原理和方法，探讨了策略迭代法的应用．

关键词: 策略迭代法，马尔可夫决策规划，优化调度．

中国图书分类号: TV697

策略迭代法和值迭代法是求解周期性马尔可夫决策规划问题的两类方法．对于鲇鱼山水库的优化调度，我们曾进行过多种方案的计算．考虑面临时段径流预报的马尔可夫决策规划模型，采用了值迭代法求解[3]．不考虑径流预报的马尔可夫决策规划模型，我们采用策略迭代法求解．研制后一方案的目的，在于为鲇鱼山水库运行调度提供不考虑径流预报的优化调度方案，并在求解模型的方法上与值迭代法进行对比研究．

1 马尔可夫决策规划模型的建立

鲇鱼山水库位于淮河支流灌河上，坝址有 1932–1987 年 56 年径流系列．水库优化调度中采用 1953–1987 年计 35 年的实测径流资料．将一年离散为 22 个时段（灌溉期 5–9 月以旬为时段，其余按月划分），经相关分析，可认为鲇鱼山水库的径流过程为独立随机序列．若不考虑面临时段径流预报，可建立如下的马尔可夫决策规划模型：

1．1 阶段变量和状态变量

将年内的离散时段 t 取为阶段变量，t＝1，2，3，……T，全年划分为 22 个阶段（即 T＝22）．以水库蓄水量 $S_t(i)$为状态变量：

$$S_t(i)=\begin{cases}0 & i=1\\ (2i-3)\triangle V/2 & i=2,3,\cdots\cdots m_t-1\\ V & i=m_t\end{cases}$$

上述水库蓄水状态的离散采用了萨瓦林斯基方法．取$\triangle V=15\times10^6m^3$，将兴利库容

---

* 收稿日期: 1990.12.22

$V_{兴}=496\times10^6m^3$ 离散成 $m-2=33$ 等分（$m=35$，是相应于 $Z_{蓄}$时的 m_t 的最大值）. $S_t(1)=0$，代表下边界的库空点，上边界 $S_t(m_t)$对应的是库满点或不同 $Z_{限}$对应的上限库容值V. 其状态离散序数 m_t，视不同时段防洪、兴利水位的限制而采用不同值：

汛 期：6月15日～8月10日，$Z_{限}=104m$， $m_t=26$

8月11日～8月31日，$Z_{限}=105m$， $m_t=29$

9月1日～9月30日， $Z_{限}=106m$， $m_t=32$

非汛期： $Z_{蓄}=107m$， $m_t=m=35$

其它中间状态 i，代表的状态值域为$(i-2)\triangle V\sim(i-1)\triangle V$

1．2 决策变量

取水电站各时段的发电流量 d_t 为决策变量. 为满足灌溉用水要求及水电站过水能力的限制，取$d_t=\{8, 11, 14, 17, \cdots\cdots 62\}$，即每隔$3m^3/s$取一决策流量，共拟定 20 个决策流量. 每个时段、每种状态，可从拟定的可能决策中选取一个最优决策.

1．3 状态转移概率

对任一时段 t，当系统从初始状态 $S_t(i)$出发，取决策 d_t 后，时段末水库蓄水状态将转移到 $S_{t+1}(j)$. 系统状态的转移可用水量平衡方程表示：

$$S_{t+1}(j)=S_t(i)+X_t-d_t\cdot\triangle t$$

令$P_t^d(i, j)$为状态转移概率，它取决于时段径流X_t的概率：

$$P_t^d(i, j)=P\{a\leqslant X\leqslant b\}=\int_a^b f_t(x)dx=F_t(a)-F_t(b)$$

式中 $f_t(x)$为 t 时段入库径流 X_t 的概率密度函数；$F_t(a)$和 $F_t(b)$为相应于 a，b 的入流 X_t 的累积概率；b，a 为蓄水状态 j 所代表的数值域的上下界，可通过水量平衡式确定.

$$\because\ X_t=S_{t+1}(j)-S_t(i)+d_t\cdot\triangle t$$

$$\therefore\ a=X_t-\triangle V/2$$

$$b=X_t+\triangle V/2$$

对于代表库满和库空的上下边界点，其状态转移概率应为：

$$P_t^d(i, 1)=1-P_t^{d'}(i, 1), \qquad \text{(下边界)}$$

$$P_t^d(i, m_t)=P_t^{d'}(i, m_t), \qquad \text{(上边界)}$$

式中 $P_t^{d'}(i, 1)$ 为由 $a=X_t=S_{t+1}(1)=S_t(1)+d_t\cdot\triangle t$ 插值求得的概率；$P_t^{d'}(i, m_t)$ 为由 $b=X_t=S_{t+1}(m_t)-S_t(i)+d_t\cdot\triangle t$ 插值求得的概率.

1．4 报酬函数

伴随水库蓄水状态的转移，水电站面临时段的效益为该时段发电量：

$$r_t^d(i, j)=C\overline{H}d_t\triangle t$$

式中 C 为单位换算系数，$\overline{H}$为时段平均发电水头.

1.5 目标函数

鲇鱼山水库是以防洪灌溉为主结合发电的综合利用水库，其优化调度问题是多目标决策问题．现选水电站多年平均发电量最大为目标函数，而将防洪、灌溉等综合利用要求作为约束条件处理．即

$$\begin{cases} g(d^*) = \max\limits_{d^* \in D} g(d) \\ P(d^*) \geqslant P_{灌} \end{cases}$$

式中 d^* 为优化策略，D 是水库长期运行策略的集合；$g(d^*)$ 为水库按 d^* 长期运行时的最优年期望发电量；$P_{灌}$ 为灌溉用水的设计保证率．

1.6 动态规划的递推方程

马尔可夫决策规划常以随机动态规划作为寻优技术，按照动态规划的最优性原理，随机动态规划的递推方程为:

$$\begin{cases} u_T(i) = \max\limits_{d_T \in D_T} \{q_T^d(i) + \sum\limits_j P_T^d(i, j) V(j)\} \\ \qquad (T = 22) \\ u_t(i) = \max\limits_{d_t \in D_t} \{q_t^d(i) + \sum\limits_j P_t^d(i, j) u_{t+1}(j)\} \\ \qquad (t = 1, 2, \cdots\cdots 21) \end{cases}$$

式中：$q_t^d(i) = \sum\limits_j P_t^d(i, j) \cdot r_t^d(i, j)$，为时段期望发电效益；V(j)为年末余留状态效益．

1.7 约束条件

1.7.1 水位约束

防洪约束　水库水位 $Z < Z_{限}$　（汛期）

兴利要求　$Z_{死} < Z < Z_{蓄}$　（非汛期）

1.7.2 流量约束

$$d_{灌} \leqslant d_t \leqslant d_T$$

$$d_t \leqslant \frac{N''}{7.5(\bar{Z} - 77.0)}$$

式中 $d_{灌}$ 为t时段灌溉流量，d_T 为水电站的最大过水能力，$\bar{Z}$ 为时段平均库水位．

1.7.3 出力约束

$$N' \leqslant N \leqslant N''$$

式中 N' 为水电站的保证出力或满足灌溉要求的最小出力；N'' 为水电站的预想出力，其最大值为水电站装机容量 $N_{装}$．当安排水电站机组定期检修时（1，2 月份 2 台机组检修，11，12 月份 4 台机组停机检修），应扣除检修机组的容量．

1.7.4 可靠性约束

满足灌溉设计保证率 $P_{灌} = 79\%$ 的要求，即：$P(d^*) \geqslant P_{灌}$

2 策略迭代法的原理及应用

对于周期性的马尔可夫决策规划模型，可用值迭代法或策略迭代法求解.

策略迭代法是以系统的策略空间为背景，通过改进策略的迭代演算，最终求得最优策略及稳态运行条件. 该法包括两个子问题:

2．1 策略改进演算

策略改进演算，是由设定的年末状态效益 V(j)及已知的时段径流 X_t 的概率分布，用随机动态规划递推式，进行年内逐时段逆序递推:

$$\begin{cases} u_T(i)=\max\limits_{d^T\in D^T}\{q_T^d(i)+\sum\limits_j P_T^d(i,\ j)V(j)\} \\ \qquad (T=22) \\ u_t(i)=\max\limits_{d_t\in D_t}\{q_t^d(i)+\sum\limits_j P_t^d(i,\ j)u_{t+1}(j)\} \\ \qquad (t=1,\ 2,\ \cdots\cdots 21) \end{cases}$$

式中年末状态效益值 V(j)，在首次迭代时需设定 (如设 V(j) = 0). 经年内逆序递推优选，可求得优化策略 $d^*=\{d_t^*(i),\ i=1,\ 2\cdots m_t;\ t=1,\ 2,\ \cdots\cdots T\}$，以及与 d^* 相应的转移概率 $P_t^{d^*}(i,\ j)$ 和时段期望发电效益 $q_t^{d^*}(i)$.

对于选定策略，按上述随机动态规划递推方程进行逆序逐时段回代，可得年初状态潜在效益 $u_1(i)$ 的表达式:

$$u_1(i)=q(i)+\sum_j P(i,\ j)V(j)$$

若用向量和矩阵表示，上式可写成

$$U_1=Q+PV$$

式中P为年转移概率矩阵，$P=\{P(i,\ j)\}$;

设P_t为时段转移概率矩阵，$P_t=\{P_t(i,\ j)\}$. 则

$$P=P_1\cdot P_2\cdot P_3\cdots\cdots P_{22}=\prod_{t=1}^{22}P_t$$

年综合期望效益$Q=(q(1),\ q(2)\cdots q(m))^T$，表示从年初状态 i 出发，相应于选定策略，经一年的随机运行所伴随的期望效益. 因策略改进演算可求得 U_1 和 P，故 Q 可由下式求得:

$$Q=U_1-PV$$

2．2 确定状态效益值的演算

上述策略改进演算求得的最优策略 d^*，是对应于设定的年末余留效益而言的. 但初始给定的 V(j)未必能反映水库长期运行的状态效益，因此需进行确定状态效益值的演算.

设水库以一选定的策略运行 n 年，则

$$u_1^{(n)}(i) = q(i) + \sum_j P(i, j) u_1^{(n-1)}(i)$$

当n足够大时，

$$u_1^{(n)}(i) = V(i) + ng$$

$$u_1^{(n-1)}(i) = V(i) + (n-1)g$$

将上两式代入前式并整理得：

$$g + V(i) = q(i) + \sum_j P(i, j) V(i)$$

此式是状态效益值演算的基本方程式，它是由 m 个方程组成的方程组．式中 g 为水库长期运行的年期望发电效益，V(i)（亦即 DP 递推式中的年末余留效益 V(j)）为运行期末的状态潜在效益．它们均需通过求解方程组求得．为了在方程式为 m 个而未知数为 m + 1个的方程组中求解，可令V(1) = 0，即以V(i) − V(1)的相对值进行计算．为便于在计算机上运算，作如下变换：

原方程组：

$$\begin{bmatrix} g \\ g \\ \vdots \\ g \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix} + \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \cdots & \cdots & \cdots \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix}$$

变换为：

$$\begin{bmatrix} P_{11}-1 & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22}-1 & \cdots & P_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ P_{m1} & P_{m2} & \cdots & P_{mm}-1 \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix} - \begin{bmatrix} g \\ g \\ \vdots \\ g \end{bmatrix} = -\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix}$$

令V(1) = 0，可化为标准形式：

$$\begin{bmatrix} -1 & P_{12} & \cdots & P_{1m} \\ -1 & P_{22}-1 & \cdots & P_{2m} \\ \vdots & \vdots & & \vdots \\ -1 & P_{m2} & \cdots & P_{mm}-1 \end{bmatrix} \cdot \begin{bmatrix} g \\ V_2 \\ \vdots \\ V_m \end{bmatrix} = -\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix}$$

应用高斯消去法可求得 g 和 V(i)，完成一次迭代计算．

将本次迭代求得的 V(i)作为下一次策略改进演算的起算值，进行下一次迭代计算．循此逐次迭代，直至迭代收敛亦即达稳态运行时为止．其收敛条件是：

① 策略稳定；

② 年末状态余留效益 V(i)稳定；

③ 长期运行的年期望效益满足精度要求

$$\frac{2|G-GG|}{G+GG} \leqslant \varepsilon$$

式中G和GG分别为本次及上次迭代计算得到的g值，ε为精度要求．当满足收敛条件(即系统达到稳态运行时)，该次迭代计算的结果：$d^* = \{d_t^*(i)\}$及$G^* = g(d^*)$即为所求．

3 计算成果及分析

运用策略迭代法对鲇鱼山水库进行了实际计算（在WangVS-300微机上运算，用FORTRAN77语言编写程序），经连续三年的迭代计算，策略已达稳定，g值亦满足精度要求，从而求得了鲇鱼山水库不考虑径流预报的优化运行决策流量表（附表1）．该表是考虑机组检修、最小决策流量为$8m^3/s$的一种优化调度方案，$G^* = g(d^*) = 3026 \times 10^4 KWh$．除此而外，还作了考虑机组检修、最小决策流量为$11m^3/s$的方案，以及不考虑机组检修、最小决策流量为$10m^3/s$的两种方案．(后两方案未附)．制作这三种计算方案的目的，是为了在相互对比中验证成果的合理性，寻求各种因素相互间的关系及变化的规律性．

这三种方案的计算结果差别不大．但从三者的对比分析中可以得到如下几点认识：

3．1 在非灌溉季节及库水位较低时，水电站在满足系统负荷、综合利用、水电站发电水头等要求的前提下，以较小流量发电．因为小流量发电，可以使以后高水头发电，并使发电与灌溉更好地结合．因此，最小决策流量为$8m^3/s$的方案，$G^* = 3026 \times 10^4 KWh$，稍高于最小决策流量为$11m^3/s$的方案（该方案$G^* = 2962 \times 10^4 KWh$）．

3．2 鲇鱼山水库在实际运行中，每年11，12月份，全部机组停机检修，1，2月份，也只安排部分机组发电，空闲机组也可进行检修．这几个月处于非灌溉的枯水季节，机组检修不仅是水电站安全运行所必需，而且因这几个月不发或少发电，增加了水库蓄水量，提高了发电水头，从而使发电效益有所提高．

3．3 灌溉季节应使发电与灌溉用水相结合，只有当灌溉用水大于水电站最大过水能力时，才从灌溉支洞补充放水．汛期，当库水位较高时，水电站出力受装机容量$N_{装}$的限制，即发电流量d_t应满足：$d_t \leqslant N_{装}/7.5(\bar{Z}-77.0)$．因此，最优决策流量在某水位之后随库水位上升而减小．说明维持高水头运行，可以使发电水耗降低，提高发电效益．

3．4 为了检验优化调度成果是否满足灌溉用水的可靠性要求，三个优化调度计算方案均用实测资料进行了模拟计算．结果均为$P(d^*) > P_{灌}$，即满足灌溉用水设计保证率79%的要求．同时，G^*值大的方案，其灌溉用水保证率略低于G^*值小的方案．这种变化趋势，符合发电效益与下游灌溉用水可靠性间的客观情况．

鲇鱼山水库最优决策流量　($d_t(i)$　m^3/s)

附表 1

水位(m)	月旬																			
	1	2	3	4	5			6			7			8			9			10
					上	中	下	上	中	下	上	中	下	上	中	下	上	中	下	
96.82	8	8	8	8	8	8	8	8	11	8	8	8	8	8	8	8	8	8	8	8
97.51	8	8	8	8	8	8	8	8	11	11	8	11	8	8	8	8	8	8	8	8
98.14	8	8	8	8	8	8	8	8	11	11	8	11	8	8	8	8	8	8	8	8
98.71	8	8	8	8	8	8	8	8	11	11	26	11	8	8	8	8	8	8	8	8
99.22	8	8	8	8	8	8	8	8	11	44	44	11	8	8	8	8	8	8	8	8
99.69	8	8	8	8	8	8	8	8	44	44	44	44	8	8	8	8	8	8	8	8
100.13	8	8	8	8	8	8	8	8	44	61	61	44	8	8	8	8	8	8	8	8
100.56	8	8	8	8	8	8	8	26	44	59	59	44	11	8	8	8	8	8	8	8
101.00	8	8	8	8	8	8	26	44	44	58	58	56	11	8	8	8	8	8	8	8
101.46	8	8	8	14	8	26	44	44	44	57	57	56	44	8	11	8	8	8	8	8
101.94	8	8	8	14	26	44	44	44	44	56	56	56	44	8	11	8	8	8	8	8
102.44	8	8	8	20	26	44	44	44	55	55	55	55	44	8	11	8	8	8	8	8
102.95	8	8	8	26	44	44	44	44	54	54	54	54	44	26	11	11	8	8	8	8
103.45	8	8	8	32	44	44	44	44	53	53	53	53	53	26	44	11	8	8	8	8
103.93	8	8	8	41	44	52	44	52	52	52	52	52	52	44	44	11	8	8	8	8
104.36	8	8	8	47	51	51	51	51							44	44	8	8	8	8
104.71	8	8	14	50	50	50	50	50							44	44	8	8	8	20
104.97	8	8	20	50	50	50	50	50							44	44	8	8	8	20
105.14	14	14	26	50	50	50	50	50									26	26	26	20
105.26	20	20	32	50	50	50	50	50									26	26	26	26
105.39	20	20	38	49	49	49	49	49									44	26	26	32
105.67	20	26	49	49	49	49	49	49												38
106.31	20	26	48	48	48	48	48	48												48
106.95	23		47	47	47	47	47	47												47

$G = 3026.334 \times 10^4 kwh$

〔注〕水位 84.00～96.06^m 各月最优决策流量均为 $8m^3/s$；11、12 月份停机检修，$d_t^*(i)=0$，现略去.

4 结　语

策略迭代法的优点是收敛速度比值迭代法快。因为这个方法的每次迭代计算，系由状态效益值解算和策略改进演算两部分所组成，两者互为因果，逐次迭代渐进。正由于迭代过程中增加了状态效益解算的人为干预，从而加快了收敛的进程。但是，该法需要求解m个方程的方程组。当状态变量多、状态离散点数多时，需要解算大型方程组，计算比值迭代法复杂。

目前，策略迭代法在国内研究和应用较少。本文根据鲇鱼山水库的特点，就策略迭代法的原理和应用进行了一些探讨，求得了鲇鱼山水库不考虑径流预报的优化调度成果。此成果可供鲇鱼山水库优化调度中不计径流预报时应用。

参 考 文 献

〔1〕 陈惠源．水库运行的随机优化及可靠性分析．武汉水利电力学院学报．1984年第2期

〔2〕 张勇传等．优化理论在水库调度中的应用．湖南科学技术出版社．1985年

〔3〕 贺北方，王海周，邓保国．综合利用水库优化调度的随机模型．水利经济．1989年第2期

Application of policy Iterative Method to Multi–reservoir Optimal Dispatching

He Beifang　Shi Bin　Luo Guanying

(Zheng zhou Institute of Technology)

Abstract: A mathematical model of Markovian decision programming is proposed in this paper, it is in accordance with the special characterteristic of the multipuropose reservoirs. The principle and method of policy iteration is discussed. At the same time, the application of policy iterative method is discussed.

Keywords: Policy iterative method, markovian decision programming, optimal dispatching.