第19卷 第3期 自 动 化 学 报 Vol. 19, No. 3
1993年5月 ACTA AUTOMATICA SINICA May, 1993

短 文

基于乘子法的离散动态非凸大系统的递阶优化算法及其应用

仲伟俊 徐南荣 陈森发
(东南大学管理学院，南京 210018)

摘 要

本文针对一类时间上关联的离散动态非凸大规模优化问题，提出了一种将部分约束作为罚项从而将非凸优化问题转化为凸优化问题的方法，研究了它的递阶优化算法，讨论了算法的收敛性及实际应用情况.

关键词: 非凸,动态大系统,递阶优化,收敛性.

一、非凸优化问题及其凸化方法

考虑下面的一类时间上关联的离散动态非凸大规模优化问题

$$\min \sum_{k=0}^{n-1} f_k(\boldsymbol{x}(k), \boldsymbol{u}(k), \boldsymbol{y}(k)), \tag{1a}$$

$$\text{s.t. } \boldsymbol{x}(k+1) = A_k \boldsymbol{x}(k) + B_k \boldsymbol{u}(k), k = 0, \cdots, n-1, \tag{1b}$$

$$h_k(\boldsymbol{x}(k), \boldsymbol{u}(k), \boldsymbol{y}(k)) = 0, \quad k = 0, \cdots, n-1, \tag{1c}$$

$$x(0) = x_0, \ x(n) = x_n. \tag{1d}$$

这里，$\boldsymbol{x}(k) \in R^q$，$\boldsymbol{u}(k) \in R^r$ 及 $\boldsymbol{y}(k) \in R^s$ 分别为关联向量、控制向量及状态向量；f_k，$h_k (k = 0, \cdots, n-1)$ 均为二次连续可微的函数；A_k，$B_k (k = 0, \cdots, n-1)$ 均为常数组成的矩阵. x_0 和 x_n 分别为给定的系统初态和终态. 若 f_k 是非凸的，或 h_k 是非线性的,都将导致问题(1)是一个非凸的优化问题. 目前虽已有几种求解它的方法[1-2],但都存在一些缺陷和不足.

为讨论方便,令 $z(k) = (x^T(k), u^T(k), y^T(k))^T$，$z = (z^T(0), \cdots, z^T(n-1))^T$，$h(z) = (h^T(z(0)), \cdots, h^T(z(n-1)))^T$. 再假定在问题(1)的局部最优解 z^* 处 $\frac{\partial h(z^*)}{\partial z}$ 满秩,且其满足局部最优解应满足的充分条件. 基于上述假定和乘子法[3]的基本原理,由问题(1)可构造出下面的问题:

$$\min \sum_{k=0}^{n-1} f_k(\boldsymbol{x}(k), \boldsymbol{u}(k), \boldsymbol{y}(k)) + \frac{c}{2} \sum_{k=0}^{n-1} \|h(\boldsymbol{x}(k), \boldsymbol{u}(k), \boldsymbol{y}(k))\|^2, \tag{2a}$$

本文于1990年4月24日收到.

$$\text{s.t. } \boldsymbol{x}(k+1)=A_k\boldsymbol{x}(k)+B_k\boldsymbol{u}(k),\quad k=0,\cdots,n-1, \tag{2b}$$

$$h_k(\boldsymbol{x}(k),\boldsymbol{u}(k),\boldsymbol{y}(k))=0,\qquad k=0,\cdots,n-1, \tag{2c}$$

$$x(0)=x_0,\ x(n)=x_n. \tag{2d}$$

其中 $c>0$ 为罚系数；$\|\cdot\|$为二范数．设 $\lambda(k)$，$\mu(k)$ $(k=0,\cdots,n-1)$ 分别为与式(2b),(2c) 有关的拉格朗日乘子向量，问题(2)的拉格朗日函数，即问题(1)的增广拉格朗日函数为

$$G(z,\lambda,\mu,c)=\sum_{k=0}^{n-1}f_k+\frac{c}{2}\sum_{k=0}^{n-1}\|h_k(z(k))\|^2+\sum_{k=0}^{n-1}\mu^T(k)h_k(z(k))$$
$$+\sum_{k=0}^{n-1}\lambda^T(k)(x(k+1)-A_kx(k)-B_ku(k)). \tag{3}$$

易证明问题(1)和问题(2)之间存在下面的关系[4]．

性质 1. 若 $z^*(k)$，$\mu^*(k)$，$\lambda^*(k)$ $(k=0,\cdots,n-1)$ 是问题(1)的平稳点，它也是问题(2)的平稳点．反之也成立．

性质 2. 存在 $c^*>0$ 使得对所有的 $c\geqslant c^*$，问题(2) 是一个具有局部凸结构的优化问题．

定理 1. 存在 $\bar{c}>0$ 使得对所有的 $c\geqslant\bar{c}$，若 (z^*,λ^*,μ^*) 是 $G(z,\lambda,\mu,c)$ 的平稳点，则 z^* 也是问题(1)的局部极小点．

二、分解-协调优化算法

问题(2)是一个具有局部凸结构的、可分解的优化问题． 采用目标协调算法，即将 $\lambda(k)$ $(k=0,\cdots,n-1)$ 作为协调变量，函数 $G(z,\lambda,\mu,c)$ 可以分解为 n 个子函数之和

$$G(z,\lambda,\mu,c)=\sum_{k=0}^{n-1}G_k(z(k),\mu(k),\lambda,c). \tag{4}$$

其中

$$G_0=f_0+\frac{c}{2}\|h_0(u(0),y(0))\|^2+\mu^T(0)h_0(u(0),y(0))-\lambda^T(0)B_0u(0), \tag{5a}$$

$$G_k=f_k+\frac{c}{2}\|h_k(z(k))\|^2+\mu^T(k)h_k(z(k))+[\lambda^T(k-1)$$
$$-\lambda^T(k)A_k]x(k)-\lambda^T(k)B_ku(k), \tag{5b}$$

$$G_{n-1}=f_{n-1}+\frac{c}{2}\|h_{n-1}(z(n-1))\|^2+\mu^T(n-1)h_{n-1}(z(n-1))$$
$$+\lambda^T(n-1)(x(n)-A_{n-1}x(n-1)-B_{n-1}u(n-1))$$
$$+\lambda^T(n-2)x(n-1). \tag{5c}$$

结合目标协调法及乘子法的基本原理，在算法的协调级，一项任务是实现关联平衡．即实现

$$\nabla_{\lambda(k)}G=\boldsymbol{x}(k+1)-A_k\boldsymbol{x}(k)-B_k\boldsymbol{u}(k)=0,\ k=0,\cdots,n-1. \tag{6}$$

上述任务可采用最陡上升算法完成，其第 l 次迭代的计算式为

$$\lambda^{l+1}(k) = \lambda^l(k) + d^{l+1}\nabla^l_{\lambda(k)}G, \quad k = 0,\cdots,n-1. \tag{7}$$

其中 d^{l+1} 为迭代计算的步长．协调级的另一项任务是调整罚系数 c 的值，其计算式为

$$c_{l+1} = \begin{cases} \alpha c_l, & \text{若 } \sum_{k=0}^{n-1} \|h_k(z^{l+1}(k))\|^2 \Big/ \sum_{k=0}^{n-1} \|h(z^l(k))\|^2 > r, \\ c_l, & \text{其它}. \end{cases} \tag{8}$$

这里，α 为大于 1 的常数，$0 < r < 1$．

在算法的第一级，计算任务是对给定的 $\lambda^l(k)$ $(k = 0,\cdots,n-1)$ 及 c_l 的值，求解由式(5)定义的 n 个优化子问题．依据乘子法的基本原理，第 k 个子问题的求解可按下列步骤进行(算法 I)．

1° 设定 $\mu(k)$ 的初值 $\mu^1(k)$ 及 $\varepsilon^* > 0$，令计数器 $j = 1$；

2° 求解下面的无约束优化问题

$$\min_{z(k)} G_k(z(k), \mu^j(k), \lambda^l, c_k); \tag{9}$$

3° 按下式修改拉格朗日乘子 $\mu^j(k)$ 的值

$$\mu^{j+1}(k) = \mu^j(k) + c_l h_k(x^j(k), u^j(k), y^j(k)); \tag{10}$$

4° 判别下列条件是否成立

$$\|\mu^{j+1}(k) - \mu^j(k)\|^2 \leqslant \varepsilon^*. \tag{11}$$

若成立，计算过程结束．否则令 $j = j + 1$，返回到 2°．

由上述讨论，可归纳出求解问题(1)的分解-协调优化算法(算法 II) 步骤如下:

1° 设令 $\lambda^1(k)(k = 0,\cdots,n-1)$ 及 c_1 的值，令计数器 $l = 1$；

2° 利用算法 I 求解 n 个无约束优化问题；

3° 判别下式是否成立

$$\sum_{k=0}^{n-1} \|\nabla^l_{\lambda(k)}G\|^2 + \sum_{k=0}^{n-1} \|h_k(z^l(k))\|^2 \leqslant \varepsilon, \tag{12}$$

若成立，计算过程结束．否则，继续计算；

4° 按式(7)和(8)分别修改 $\lambda^l(k)(k = 0,\cdots,n-1)$ 及 c_l 的值，令 $l = l + 1$，返回到 2° 继续计算．

三、算法的收敛性

可以证明式(8)的迭代计算有与通常的二次罚函数法相同的收敛特性[3]．式(7) 的迭代计算的收敛性由下面的定理给出．

定理 2. 当罚系数的取值充分大时，适当选取 d^l 的值，迭代算法(7)以 $(B-b)^2/(B+b)^2$ 为收敛速率局部线性收敛于 $\lambda(k)$ 的解 $\lambda^*(k)(k = 0,\cdots,n-1)$．其中 B 和 b 是 Hessian 矩阵 $\nabla^2_\lambda G(z^*, \lambda^*, \mu^*, c)$ 的最大和最小的特征值．

算法 I 的收敛性和收敛速率由定理 3 给出．

定理 3. 设 ε 和 δ 是正的标量，且对所有的 $\mu(k) \in B(\mu^*(k),\delta), c_l \geqslant c^*$ 有 $z(k) \in$

$B(z^*(k), \varepsilon)$. 假定 ε 和 δ 充分小，c_l 充分大使得

$$c_l > \max\left\{0, \frac{2}{e_j[D(z(k), \mu^*(k), \lambda)]}\right\}, \tag{13}$$

其中

$$D(z(k), \mu^*(k), \lambda) = -\frac{\partial h_k(z(k))}{\partial z(k)} \nabla^2_{z(k)} \mathscr{L}(z(k), \mu^*(k), \lambda) \left(\frac{\partial h_k(z(k))}{\partial z(k)}\right)^T.$$

$e_j[D(z(k), \mu^*(k), \lambda)]$ 代表矩阵 $D(z(k), \mu^*(k), \lambda)$ 的第 j 个特征值. 若 $\mu^0(k) \in B(\mu^*(k), \delta)$，则由式(10)产生的序列 $\mu^j(k)$ 始终保持在 $B(\mu^*(k), \delta)$ 中，并以线性速率收敛于 $\mu^*(k)$，且当 $c_l \to \infty$ 时，其收敛速率是超线性的.

四、算法的应用实例

本文所讨论的算法已成功地应用于具有下列模型的城市供水系统的优化调度问题

$$\min \sum_{k=1}^{n} \sum_{i \in X} [\nu_i u_i(k) + \gamma_i u_i(k)(p_i(k) - ph_i)], \tag{14a}$$

$$\text{s.t.}\quad x_i(k+1) = x_i(k) + s_i F_i(k),\ k = 0, \cdots, n-1;\ i \in Y, \tag{14b}$$

$$g_i(p(k), u(k), F(k), x(k)) = 0,\ k = 0, \cdots, n-1;\ i = 1, \cdots, m, \tag{14c}$$

$$p_i(k) \geqslant p_i^{\min},\ k = 0, \cdots, n-1;\ i = 1, \cdots, m, \tag{14d}$$

$$u_i^{\min} \leqslant u_i(k) \leqslant u_i^{\max},\ k = 0, \cdots, n-1;\ i \in X, \tag{14e}$$

$$x_i^{\min} \leqslant x_i(k) \leqslant x_i^{\max},\ k = 0, \cdots, n-1;\ i \in Y, \tag{14f}$$

$$x_i(0) = x_i(n) = \text{给定值},\ i \in Y. \tag{14g}$$

这里，n 和 m 分别为时间区间数及节点数；X 和 Y 分别为水源和水库所在节点标号的集合；$u_i(k)$，$F_i(k)$，$x_i(k)$ 及 $p_i(k)$ 分别为在时间区间 k 在节点 i 的水源供水量、水库进出水量、水库水位及节点压力. 式 (14b) 代表水库的进出水量与其水位之间的关系，式 (14c) 代表节点的水流平衡方程；式 (14a) 代表供水系统的运行成本，是一个非凸的函数. 采用本文讨论的算法可求解该问题. 其不等式约束可采用乘子法中的方法对其进行处理. 对图 1 所示的某简单供水系统的仿真结果列成表 1. 仿真计算过程也表明该算法

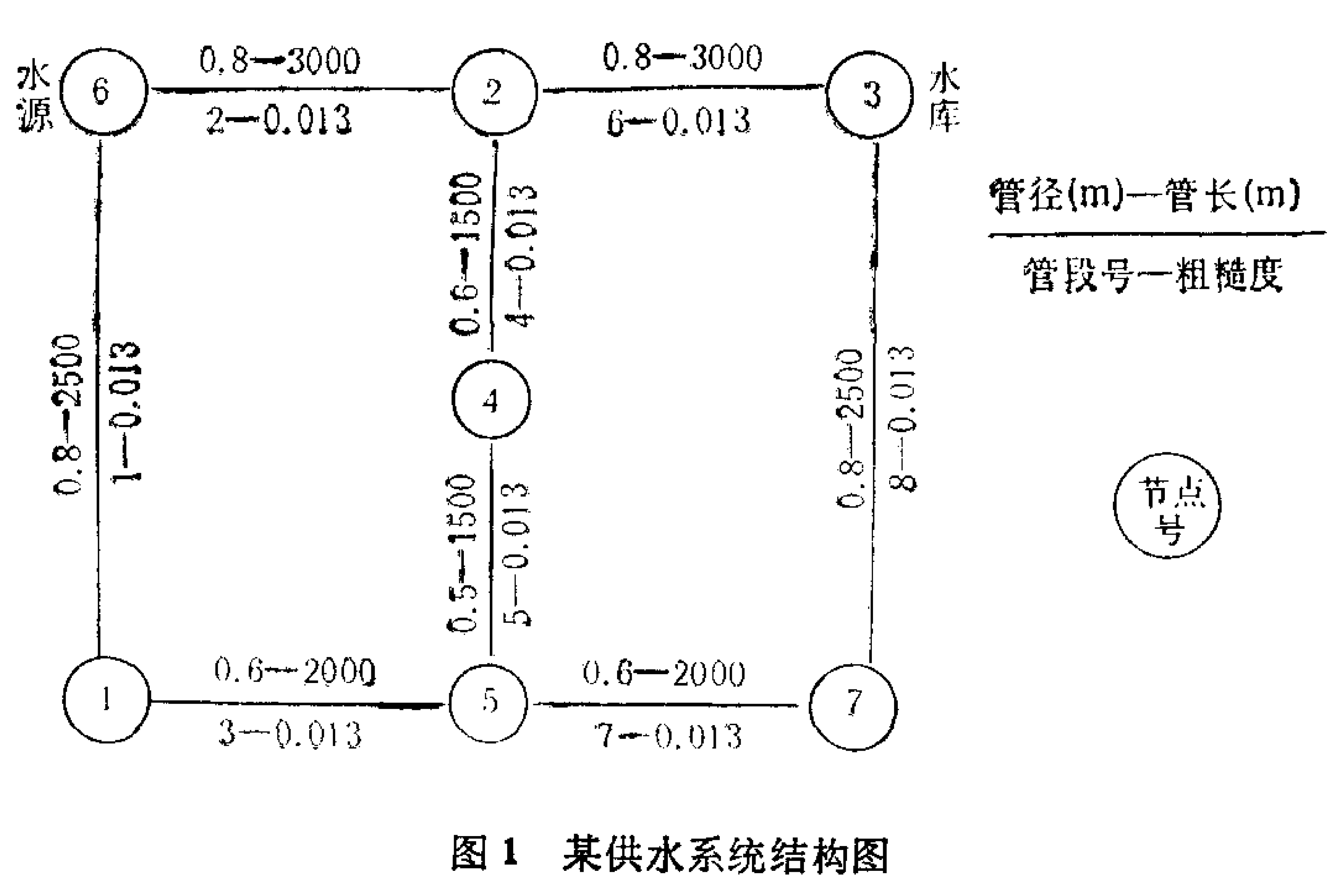

图 1 某供水系统结构图

有良好的收敛性和计算速度. 影响算法计算时间的关键是求解问题(9)所需的计算时间.

表 1　供水系统的负荷及其优化计算结果

时间区间号	节点流量 (m^3/s)					计算结果				
	1	2	4	5	7	供水流量 (m^3/s)	供水压力 (m)	水库进出水量 (m^3/s)	水库水位 (m)	目标函数
1	0.25	0.15	0.20	0.45	1.15	2.359	72.36	−0.159*	26.98	170.65
2	0.65	0.65	0.30	0.85	0.75	3.099	80.32	0.101	26.36	248.90
3	1.00	0.95	0.40	0.95	0.40	3.600	89.88	0.100	25.77	323.52
4	0.60	0.15	0.30	0.30	1.00	2.392	60.90	−0.042*	26.00	145.70

* 一负号表示此时水库进水

参考文献

[1] Bertsekas, D. P., Convexification Procedures and Decomposition Methods for Nonconvex Optimization Problems, *J. Opti. Theory Appl.*, 29(1979), 169—197.

[2] Luca, A. D. and Pielo, G. P., Exact Augmented Lagrangian Approach to Multilevel Optimization of Large-Scale System, *Int. J. Systems Sci.*, 18(1987), 157—176.

[3] Bertsekas, D. P., Multiplier Methods: *A Survey, Automatica*, 12(1976), 133—145.

[4] 仲伟俊、徐南荣、陈森发，城市供水系统调度的分解-协调优化方法，自动化学报，**16**(1990)，217—225.

HIERARCHICAL OPTIMIZATION ALGORITHM OF DISCRETE DYNAMIC AND NONCONVEX LARGE-SCALE SYSTEMS BASED ON MULTIPLIER METHOD AND ITS APPLICATIONS

ZHONG WEIJUN　　XU NANRONG　AND　CHEN SENFA

(School of Management, Southeast University, Nanjing 210018)

ABSTRACT

In this paper, a method to transform a discrete dynamic and nonconvex large-scale optimization problem into convex one is proposed by constructing the penalty terms of the part constraints. The hierarchical optimization algorithm is studied. The convergence and the application in the real world problem of the algorithm are also discussed.

Key words: Nonconvex; dynamic large-scale system; hierarchical optimization; convergence.