第 13 卷第 1 期　　控制理论与应用　　Vol. 13, No. 1
1996 年 2 月　　CONTROL THEORY AND APPLICATIONS　　Feb., 1996

非凸稳态优化控制新方法——双迭代法

黄正良　吴　坚　　　万百五　韩崇昭
(西南工学院科研处·四川绵阳,621002)　(西安交通大学系统工程研究所·西安,710049)

摘要：本文针对带有输出关联不等式约束的非凸稳态优化控制问题,提出了一种新的双迭代算法,该算法具有全局单调收敛性,应用条件较弱.最后的仿真结果也说明了该算法的有效性.

关键词：稳态优化；双迭代法；目标系数

1　引　言

自从 Roberts 提出修正两步法以来,工业过程稳态优化控制的各种改进算法相继问世[1].但针对带有输出关联不等式约束且目标函数为非凸情形的问题至今还没有能保证迭代收敛的算法.其主要原因是算法的收敛性往往需要目标函数的严格凸性来保证[2].解决这一问题的自然想法是凸化目标函数[3].本文利用一严格凸的函数序列去逐步逼近原来的目标函数,所提的算法不仅具有全局单调收敛性,还能保证收敛到局部最优点.

2　算法的构成

带有输出关联不等式约束的工业过程稳态优化控制问题可描述如下[1]

$$\text{(ROP)}\quad \begin{cases} \min Q(c,y), \\ \text{s.t.}\quad y=F_*(c), \\ (c,y)\in CY=\{(c,y)\in \mathrm{R}^n\times \mathrm{R}^m \mid G(c,y)\leqslant 0\}, \\ c\in C=\{v\in \mathrm{R}^n \mid f(v)\leqslant 0\} \quad (C\text{为紧集}). \end{cases} \tag{1}$$

系统的近似稳态模型为 $y=F(c,\alpha),\alpha\in\mathscr{A}$($\mathscr{A}$ 为 R^m 中的紧集).假设模型是点参数的,即 $\forall\, v\in C$,存在 $\alpha\in\mathscr{A}$ 使 $F_*(v)=F(v,\alpha)$.进一步假设映射 $Q(\cdot,\cdot),F_*(\cdot),F(\cdot,\cdot),G(\cdot,\cdot)$ 和 $f(\cdot)$ 都是二阶连续可微的.记 $q(c,\alpha)=Q(c,F(c,\alpha)),g(c,\alpha)=G(c,F(c,\alpha)),q_*(c)=Q(c,F_*(c))$ 和 $g_*(c)=G(c,F_*(c))$.在上述假设下,显然(ROP)问题等价于下列(EOP)问题:

$$\text{(EOP)}\quad \begin{cases} \min q(c,\alpha), \\ \text{s.t.}\quad F(c,\alpha)=F_*(c), \\ g(c,\alpha)\leqslant 0,\quad c\in C. \end{cases} \tag{2}$$

由于输入输出特性 $F_*(\cdot)$ 是未知的,求解(EOP)问题只能采用迭代搜索方式.为此首先任取设定点 $v\in C$,和一非负向量 $\xi\in\mathrm{R}^p_+=\{x\mid x\in\mathrm{R}^p,x\geqslant 0\}$,让模型输出和实际系统输出匹配确定参数集合 $\alpha(v)$,一般项记为 $\hat\alpha$.进一步记

$$P(c,v)=g_*(v)+\nabla^{\mathrm T} g_*(v)(c-v); \tag{3}$$

$$J(c,v,\hat\alpha)=q(c,\hat\alpha)+(\nabla g_*(v)-\nabla_c q(v,\hat\alpha))^{\mathrm T}c+\rho\|c-v\|^2/2; \tag{4}$$

本文于 1994 年 8 月 25 日收到.1995 年 5 月 17 日收到修改稿.

$$J_1(c,v,\hat{\alpha},\xi)=J(c,v,\hat{\alpha})+\xi^{\mathrm{T}}P(c,v). \tag{5}$$

这里 ∇ 表示梯度算子，$\|\cdot\|$ 是一般的欧氏范数，$\rho>0$. 将 v 加到实际过程中，测量到 $F_*(v)$，利用摄动技术可获取过程导数 $\nabla F_*(v)$，从而可得到(3)～(5)式. 基于已知模型的优化控制问题为

$$\text{(MOP)}\qquad \begin{cases}\min\ J(c,v,\hat{\alpha}),\\ \text{s.t.}\quad P(c,v)\leqslant 0,\quad c\in C.\end{cases} \tag{6}$$

记解为 $\hat{C}(v,\hat{\alpha})$，并记 $\hat{C}(v)=\bigcup\limits_{\hat{\alpha}\in\alpha(v)}\{\hat{c}(v,\hat{\alpha})\}$. 由于 $J(\cdot,v,\hat{\alpha})$ 的严格凸性(适当选择 ρ)，其解 $\hat{c}(v,\hat{\alpha})$ 是唯一存在的. 求解(MOP)问题有时十分困难，特别当变量维数过大时，采用对偶法将问题进行分解有明显的优点. (MOP)问题的 Lagrange 对偶问题为

$$\text{(DOP)}\qquad \begin{cases}\max\ L(v,\hat{\alpha},\xi),\\ \text{s.t.}\quad \xi\in \mathrm{R}_+^{r}.\end{cases} \tag{7}$$

对偶函数 $L(v,\hat{\alpha},\xi)$ 由下列优化问题确定

$$L(v,\hat{\alpha},\xi)=\min_{C}J_1(c,v,\hat{\alpha},\xi). \tag{8}$$

上述解集记为 $c(v,\hat{\alpha},\xi)$. 显然 $L(v,\hat{\alpha},\cdot)$ 是 R_+^r 上的凹函数，且 $c(v,\hat{\alpha},\xi)$ 是唯一存在的，$P(c(v,\hat{\alpha},\xi),v)$ 为其梯度. 进一步若约束 $P(c,v)\leqslant 0$ 为正则约束，即存在 $\tilde{v}\in C$ 使 $P(\tilde{v},v)<0$，那么依强对偶定理[5]知(MOP)问题与(DOP)问题是等价的.

求解(DOP)问题可采用如下梯度法

$$\xi=\xi+\varepsilon_\xi\hat{P}(c(v,\hat{\alpha},\xi),v). \tag{9}$$

这里 ε_ξ 是事先给定的步长因子，$\hat{P}(c(v,\hat{\alpha},\xi),v)$ 为 ξ 处的可行上升方向，其含义如下：

$$\hat{P}_i(c(v,\hat{\alpha},\xi),v)=\begin{cases}P_i(c(v,\hat{\alpha},\xi),v), & \text{若 } \xi_i>0,\\ \max[0,P_i(c(v,\hat{\alpha},\xi),v)], & \text{若 } \xi_i=0.\end{cases} \tag{10}$$

这里 P_i 表示 P 的第 i 个分量.

若存在 $\bar{\xi}\in R_+^r$ 使 $\hat{P}(c(v,\hat{\alpha},\bar{\xi}),v)=0$，则 $c(v,\hat{\alpha},\bar{\xi})$ 必为(DOP)的解. 事实上，由 $\hat{P}(c(v,\hat{\alpha},\bar{\xi}),v)=0$，可知 $P(c(v,\hat{\alpha},\bar{\xi}),v)\leqslant 0$ 和 $\bar{\xi}_iP_i(c(v,\hat{\alpha},\bar{\xi}),v)=0$，从而(DOP)问题的一阶 kuhn-Tucker 条件在 $\bar{\xi}$ 处成立. 又因为 $L(v,\hat{\alpha},\cdot)$ 为凹函数，所以 $\bar{\xi}$ 是(DOP)问题的最优解，即 $c(v,\hat{\alpha},\bar{\xi})$ 为(MOP)问题的解优解，于是有 $c(v,\hat{\alpha},\bar{\xi})=\hat{c}(v,\hat{\alpha})$. 这种梯度法的收敛性是显然的[5]，求解过程是离线迭代，即所谓内环迭代.

外环迭代可描述如下：

首先选择初始值 $v^0\in C$ 和 $\xi^0\in \mathrm{R}_+^r$，外环第 k 步迭代为

1) 将 v^k 加到实际系统中，测量输出稳态值 $F_*(v^k)$；利用模型输出与实际系统输出匹配确定参数集合 $\alpha(v^k)$；利用摄动技术估计 $\nabla F_*(v^k)$.

2) 求解(MOP)问题，更新设定值

$$v^{k+1}=v^k+\varepsilon_v(\hat{c}(v^k,\hat{\alpha})-v^k),\quad \hat{\alpha}\in\alpha(v^k). \tag{11}$$

终止准则：　$\hat{c}(v^k,\hat{\alpha})=v^k$.

内环迭代可描述如下：

对于由外环迭代产生的 v^k 和对应于 $P(c,v^k)\leqslant 0$ 的拉氏乘子 ξ^k，取初始值 $\xi_0^k=\xi^k$，内环迭代第 j 步为

1）确定 $L(v^k,\hat{\alpha},\xi_j^k)$ 和 $c(v^k,\hat{\alpha},\xi_j^k)$；

2）更新迭代；

$$\xi_{j+1}^k = \xi_j^k + \varepsilon_\xi \hat{P}(c(v^k,\hat{\alpha},\xi_j^k),v^k). \tag{12}$$

终止准则： $\hat{P}(c(v^k,\hat{\alpha},\xi_j^k),v^k) = 0.$

如在第 j 步终止迭代，记 $\xi^{k+1} = \xi_j^k$，否则记 $\xi^{k+1} = \lim\limits_{j\to\infty}\xi_j^k$，自然有 $c(v^k,\hat{\alpha},\xi^{k+1}) = \hat{c}(v^k,\hat{\alpha})$.

3 算法的最优性与收敛性分析

记算法解集为 $\Omega = \{v|\hat{c}(v,\hat{\alpha}) = v, v\in C, \hat{\alpha}\in\alpha(v)\}$，(ROP) 问题的最优解集记为 Ω^*.

定理 1 如果 $f(c)\leqslant 0$ 为正则约束，则任意算法解必为 (ROP) 问题的一阶 Kuhn-Tucker 点.

证 $\forall\ v\in\Omega$，则存在 $\hat{\alpha}\in\alpha(v)$ 使 $\hat{c}(v,\hat{\alpha}) = v$，又因 $\hat{c}(v,\hat{\alpha})$ 为(MOP) 问题的解，则必存在乘子 $\eta\in R_+^L$，$\xi\in R_+^P$ 使得

$$\nabla_c J(v,v,\hat{\alpha}) + \nabla_c P(v,v)\xi + \nabla f(v)\eta = 0; \tag{13}$$

$$P(v,v)\leqslant 0,\quad \xi_i P_i(v,v) = 0,\quad i = 1,\cdots,P; \tag{14}$$

$$f(v)\leqslant 0,\quad \eta_j f_j(v) = 0,\quad j = 1,\cdots,L. \tag{15}$$

由(3)式和(4)式知

$$P(v,v) = g_*(v),\quad \nabla_c P(v,v) = \nabla g_*(v),\quad \nabla_c J(v,v) = \nabla q_*(v). \tag{16}$$

从而有

$$\nabla q_*(v) + \nabla g_*(v)\xi + \nabla f(v)\eta = 0; \tag{17}$$

$$g_*(v)\leqslant 0,\quad \xi_i g_{*i}(v) = 0,\quad i = 1,\cdots,P. \tag{18}$$

由(15)式，(17)式和(18)式知 v 为(ROP) 问题的一阶 Kuhn-Tucker 点.

定理 2 如果 $g_*(c)\leqslant 0$ 和 $f(c)\leqslant 0$ 为正则约束，且 $f(c)$ 为凸函数，则任意最优解可由算法导出.

证 设 v 为任意最优解，则必存在 $\xi\in R_+^P$ 和 $\eta\in R_+^L$ 使得(15)式，(17)式和(18)式成立，从而(13)～(15)式成立. 又因为 $p(\cdot,v)$ 和 $f(\cdot)$ 是凸的，并且 $J(\cdot,v,\hat{\alpha})$ 为严格凸函数，则 $v\in\Omega$.

定理1意味着任意算法解必是原问题的局部最优解. 在实际工业过程中，由于原材料的变更和触媒的老化等原因引起的最优工况漂移，一般也是局部的. 所以在工作点附近寻找到的局部最优点就是全局最优工作点，定理1有着重要的实际意义.

下面给出全局单调收敛性定理.

定理 3 如果下列假设成立：

1）近似模型是点参数的；

2）参数 ε_v 和 ρ 满足下列关系

$$\overline{M} - \rho + M\varepsilon_v/2 < 0. \tag{19}$$

这里 $M = \max\limits_v \|\nabla^2 q_*(v)\|$，$\overline{M} = \max\limits_{v,\alpha}\|\nabla^2 q(v,\alpha)\|$，则上述迭代算法或者有限步终止于 Ω 上，或产生一个无限序列 $\{v^k\}$ 以致于

1）$\{v^k\}$ 的所有聚点属于 Ω；

2）$\lim\limits_{k\to\infty} q_*(v^k) = q_*(\bar{v})$，$\bar{v}$ 为 $\{v^k\}$ 的任意聚点.

证　如果算法有限步不能终止于 Ω 上，那么 $\forall\ k \geqslant 0$，必有 $\hat{c}(v^k, \hat{\alpha}) \neq v^k, \hat{\alpha} \in \alpha(v^k)$. 下面的证明分两步，先证明 $q_*(v^k)$ 严格单调下降；后证明由(11)式确定的映射是闭映射，再由文[5]知定理 3 成立. 首先有

$$q_*(v^{k+1}) - q_*(v^k) \leqslant \nabla^{\mathrm{T}} q_*(v^k)(v^{k+1} - v^k) + \frac{1}{2} M \| v^{k+1} - v^k \|^2. \tag{20}$$

又因为 $\hat{c}(v^k, \hat{\alpha})$ 是(MOP)问题的解，则有[5]

$$(\nabla_c q(\hat{c}(v^k, \hat{\alpha}), \hat{\alpha}) + \nabla q_*(v^k) - \nabla_c q(v^k, \hat{\alpha}) + \rho(\hat{c}(v^k, \hat{\alpha}) - v^k))^{\mathrm{T}}(v^k - \hat{c}(v^k, \hat{\alpha})) \geqslant 0. \tag{21}$$

于是有

$$q_*(v^{k+1}) - q_*(v^k) \leqslant \varepsilon_v(\overline{M} - \rho + M\varepsilon_v/2) \| \hat{c}(v^k, \hat{\alpha}) - v^k \|^2 < 0, \tag{22}$$

即 $q_*(v^k)$ 严格单调下降.

由(11)式确定的点到集的映射为

$$\tau(v): v \to \hat{V}(v) = \{v + \varepsilon_v(\hat{c}(v, \hat{\alpha}) - v) | \hat{\alpha} \in \alpha(v), a \leqslant \varepsilon_v \leqslant 2(\rho - \overline{M})/M\}. \tag{23}$$

这里 $a > 0$ 为某一事先选定的常数. 下面证明 $\tau(\cdot)$ 是闭的，首先由 $F_*(\cdot)$ 和 $F(\cdot, \cdot)$ 的连续性及点参数性可知 $v \to \alpha(v)$ 是 C 上的闭映射；其实 $\hat{c}(\cdot, \cdot)$ 是 $C \times \mathscr{A}$ 上的连续映射[5]，于是点到集的映射 $v \to \hat{C}(v)$ 是 C 上的闭映射. 事实上，任取 $\bar{v} \in C, v^k \in C, \hat{c}^k \in \hat{C}(v^k)$ 且 $v^k \to \bar{v}, \hat{c}^k \to \bar{c}$. 因 $\hat{c}^k \in \hat{C}(v^k)$，所以存在 $\hat{\alpha}^k \in \alpha(v^k)$ 使 $\hat{c}^k = \hat{c}(v^k, \hat{\alpha}^k)$. 因 $\mathscr{A}$ 为紧集，所以存在 $\hat{\alpha}^k$ 的聚点 $\bar{\alpha}$. 又因 $\hat{c}(\cdot, \cdot)$ 是连续的，从 $\hat{c}^k = \hat{c}(v^k, \hat{\alpha}^k)$ 两边按子序列取极限，必有 $\bar{c} = \hat{c}(\bar{v}, \bar{\alpha})$. 又因 $\alpha(\cdot)$ 是闭映射，必有 $\bar{\alpha} \in \alpha(\bar{v})$，所以 $\bar{c} \in \hat{C}(\bar{v})$，即 $\hat{C}(\cdot)$ 是 C 上的闭映射. 从(23) 式可知 $\tau(\cdot)$ 是 C 上的闭映射. 于是由文[5]知上述定理成立.

从定理 3 可看出，本文所提算法是全局单调收敛的，且适用于非凸问题. 定理中的条件(1)是自然的；条件(2)只需合理选择 ρ 和 ε_v 就能得到满足. 从(21)式可看出，如果 $\nabla q(\cdot, \alpha)$ 是严格单调的，只需选择 $\rho = 0, \varepsilon_v < 2b/M$($b$ 为单调常数). 这里 ρ 相当于一个罚系数，它能抑制算法的迭代步长，保证迭代收敛[3].

考虑线性二次型稳态优化控制问题

$$\begin{aligned}
&F(c, \alpha) = Ec + H\alpha, \\
&F_*(c) = E_* c + d_*, \\
&Q(c, y) = (c - c_d)^{\mathrm{T}} S_c (c - c_d) + (y - y_d)^{\mathrm{T}} S_y (y - y_d), \\
&G(c, y) = Tc + Wy + z \leqslant 0.
\end{aligned}$$

这里 S_c 和 S_y 为半正定矩阵，从而有

$$\begin{aligned}
&M = 2\lambda_{\max}(S_c + E_*^{\mathrm{T}} S_y E_*); \\
&b = z\lambda_{\min}(S_c + E^{\mathrm{T}} S_y E).
\end{aligned}$$

如果 $b \neq 0$，可取 $\rho = 0, \varepsilon_v < 2b/M$；如果 $b = 0$，取 $\varepsilon_v < 2\rho/M$.

4　仿真研究[4,7]

考虑如下系统的数字仿真

例 1　实际系统为

$$\begin{aligned}
&y_1 = c_1 - c_2 + 2y_2 - 0.5c_1^2 + 0.5(c_1 + c_2 - 2)y_2, \\
&y_2 = c_3 - c_4 + y_1 - 3y_4,
\end{aligned}$$

$$y_3 = 2c_4 - c_5 - y_1 + y_4,$$
$$y_4 = c_6 - 4y_3 + 0.5c_6y_3.$$

近似模型(Ⅰ)为

$$y_1 = 1.4375c_1 - 0.1875c_2 + 1.5y_2 + \alpha_1,$$
$$y_2 = 0.5c_3 - 1.5c_4 + y_1 - 2y_4 + \alpha_2,$$
$$y_3 = 2.5c_4 - 0.5c_5 - y_1 + 1.5y_4 + \alpha_3,$$
$$y_4 = 1.25c_6 - 3y_3 + \alpha_4.$$

近似模型(Ⅱ)为

$$y_1 = 1.7c_1 + 0.2c_2 + 1.5y_2 + \alpha_1,$$
$$y_2 = 0.5c_3 - 1.5c_4 + y_1 - 2y_4 + \alpha_2,$$
$$y_3 = 2.5c_4 - 0.5c_5 - y_1 + 1.5y_4 + \alpha_3,$$
$$y_4 = 0.8c_6 - 3y_3 + 0.4c_6y_3 + \alpha_4.$$

约束为

$$CY = \{(c,y) \in \mathrm{R}^{10} \mid 1.006c_1 - y_2 \geqslant 0, 0.375 + 2.25c_6 - 2.75y_3 - y_4 \geqslant 0\},$$
$$C = \{c \in \mathrm{R}^6 \mid |c_i| \leqslant 0.5 \quad i = 1,4,5,6, 0 \leqslant c_2 \leqslant 2.5, 0 \leqslant c_3 \leqslant 2\}.$$

目标函数为

$$Q(c,y) = c_1^2 + (c_2 - 2)^2 + 2(c_3 - 2)^2 + c_4^2 + 3c_5^2 + (c_6 + 1)^2 + 4y_1^2 + (y_2 - 1)^2 + (y_3 - 1)^2 + y_4^2.$$

因 $Q(c,y)$ 是二次函数,所以 $q_1(c,\alpha) = Q(c,F_1(c,\alpha))$ 是 C 上的二次函数,且梯度 $\nabla_c q_1(c,\alpha)$ 是严格单调的,所以取 $\rho = 0$;$q_2(c,\alpha) = Q(c,F_2(c,\alpha))$,其梯度不再是严格单调的,所以 $\rho = 10$. 初始条件取零值,迭代收敛准则取误差上界为 10^{-4}. 其仿真结果见表1.

表1 例1仿真结果

	在线迭代次数	ε_r	ρ	算法值	实际最优值	模型
Lin 方法	43			2.1405	2.1405	Ⅰ
Brdy's 方法	32			2.1405		
双迭代方法	29	0.7	0	2.1405		
	24	0.5	10	2.1405		Ⅱ
Lin 方法,Brdy's 方法,发散						

5 结论

本文研究了带有输出关联不等式约束的非凸稳态优化控制问题,给出了一种具有全局单调收敛的双迭代法,应用条件较弱,能解决以往算法不能解决的非凸优化问题.

参考文献

[1] 万百五,林杰. 大规模工业过程的稳态递阶控制综述. 自动化学报,1990,16(2):1—7

[2] Brdys, M.. et al.. Convergence and Optimality of Modified Two-Step Algorithm for Integrated System Optimization and Parameter Estimation. Int. J. Syst. Sci., 1987, 18(7): 1305—1322

[3] 黄正良,万百五,韩崇昭.工业过程稳态优化新方法——函数逼近法.自动化学报,1994,20(2):227-230

[4] Brdys, M.. et al.. An Extension to the Modified Two-Step Algorithm for Steady-State System Optimization and Parameter Estimation. Int, J. Syst. Sci., 1986, 17(8): 1229—1243

[5] Bbaazaraa, M.. et al.. Nonlinear Programming Theory and Algorithm. John and Wiley & Sons, 1979

[6] Clande, B.. Topological Space. Oliver and Boyd Ltd, 1963

[7] Lin, J.. et al.. Modified Algorithm for Steady-State Integrated System Optimisation and Parameter Estimation. IEE Proc., 1988, 135(2): 119—126

A New Double Iterative Algorithm for Steady-State Optimizing Control of Industrial Processes with Nonconvex Index

HUANG Zhengliang　and　WU Jian

(Administration of Achievements in Scientific Research of Southwest Institute of Technology, Mianyang Sichuan, 621002, PRC)

WAN Baiwu　and　HAN Chongzhao

(Institute of Systems Engineering Xi'an Jiaotong University · Xi'an 710049, PRC)

Abstract: In this paper, a new double iterative algorithm for steady-state optimizing control of industrial processes with output dependent constraints and nonconvex index is presented. Under very mild conditions, the algorithm possesses the global convergence. Finally, the simulation results show that the new approach is very efficient.

Key word: steady-state optimizing control; double iterative algorithm; index

本文作者简介

黄正良　1962 年生.副教授.1983 年毕业于武汉建材学院,1988 年毕业于东北工学院,1992 年西安交通大学博士研究生毕业.现为西南工学院院长助理.主要研究方向为动态对策,非线性系统辨识、工业过程稳态优化.发表论文 30 余篇.

吴　坚　1943 年生.副教授.1966 年毕业于重庆大学,现任西南工学院副院长.主要从事微机在工业过程控制中的应用研究,并有多项成果获奖,发表论文 10 余篇.

万百五　1928 年生.教授.1951 年毕业于交通大学电信研究所.现任西安交通大学系统工程研究所大系统室主任.目前主要研究方向为大系统智能控制,国内外发表论文 120 余篇.

韩崇昭　1943 年生.教授.1968 年毕业于西安交通大学,1981 年中科院研究生院硕士研究生毕业.现为西安交通大学系统工程研究所副所长.主要研究方向为自适应控制和非线性频谱分析等.