第 30 卷 第 2 期
2004 年 3 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 30, No. 2
Mar., 2004

基于模糊模型的大系统关联平衡法的收敛性

顾佳晨[1] 万百五[2]

[1](冶金自动化研究设计院 北京 100071)

[2](西安交通大学系统工程研究所 西安 710049)

(E-mail: gujiachen@263.net)

摘 要 结合模糊规划与非线性规划的收敛性分析方法,给出了基于模糊模型的关联平衡法的收敛性分析.首先证明了经去模糊处理后形成的约束集合与子过程原有的约束集合必有交集,并且此交集是凸集.在此基础上,分析和证明了基于模糊模型的关联平衡法可用于求解基于模糊模型的稳态大工业过程递阶优化问题.继而通过定义迭代序列的 A-内积,证明了基于模糊模型的关联平衡法是收敛的.同时给出了保证迭代收敛的迭代系数取值范围.

关键词 模糊数,模糊规划,关联平衡法

中图分类号 TP273

Convergence of Interaction Balance Method Based on Fuzzy Model for Large Scale Systems

GU Jia-Chen[1] WAN Bai-Wu[2]

[1](*Automation Research and Design Institute of Metallurgical Industry, Beijing* 100071)

[2](*Systems Engineering Institute, Xi'an Jiaotong University, Xi'an* 710049)

(E-mail: gujiachen@263.net)

Abstract This paper studies the convergence of Interaction Balance Method (IBM) based on fuzzy model by combining the analysis of convergence of fuzzy programming and nonlinear programming. Firstly, it is proven that the intersection of the defuzzified and the original constraint sets is nonempty and convex. Secondly, the applicable condition of IBM based on fuzzy model is established. Finally, the convergence of IBM based on fuzzy model is proven by using the definition of the A-norm of iterative sequence, and the range of iteration gain is found for which convergence is guaranteed.

Key words Fuzzy numbers, fuzzy programming, interaction balance method

1 引言

为解决稳态大工业过程递阶优化中的模型-实际差异问题,文献[1]提出了基于模糊模

收稿日期 2002-06-25 收修改稿日期 2002-11-12
Received June 25, 2002; in revised form November 12, 2002

型的关联平衡法,但并未证明其收敛性,本文是对它的完善. 由于基于模糊模型的关联平衡法是非线性模糊规划与大系统稳态递阶关联平衡法结合的产物,因此为证明其收敛性必然要考虑如何将这两种方法的收敛性证明结合起来. 目前,模糊规划问题主要是先将其转化为确定性规划问题,然后用普通的优化算法进行求解[2]. 但分析模糊规划问题收敛性的文章较少,文献[3,4]分别对约束和目标函数中含有模糊参数的多目标非线性规划问题进行了稳定性分析,在定义了问题的可行参数集后,得出了模糊参数的稳定集.

基于模糊模型的关联平衡法,经过去模糊后得到的实际上是一个非线性规划问题. 在求解非线性规划问题时,有关迭代算法的全局收敛性分析,Zangwill 给出了统一的分析方法[5].

大系统的关联平衡法实际上是将原问题转化为它的对偶问题来求解. 由于 Zangwill 全局收敛性定理是一种证明迭代求极值算法收敛性的一般性方法,对于本文这种特定的规划问题,用其证明将比较繁琐. 本文从另一个角度出发,首先通过满足度函数的定义,证明了去模糊后得到的约束集合是凸集;其后在算法收敛性的证明方法上分两步来证明. 首先,证明了去模糊后得到的 α-OP 问题经对偶分解后可用基于模糊模型的关联平衡法求解;继而,通过定义迭代序列的 A 内积和 A 范数,完成了基于模糊模型的关联平衡法的收敛性证明,并给出了迭代系数的取值范围.

2 问题描述

稳态大工业过程的优化问题(称为 OP 问题)可描述如下:找到控制量 $\hat{\boldsymbol{c}}$,关联输入量 $\hat{\boldsymbol{u}}$,关联输出量 $\hat{\boldsymbol{y}}$,使过程的整体性能指标达到最优

$$Q(\hat{\boldsymbol{c}},\hat{\boldsymbol{u}},\hat{\boldsymbol{y}}) = \min_{\boldsymbol{c},\boldsymbol{u},\boldsymbol{y}} Q(\boldsymbol{c},\boldsymbol{u},\boldsymbol{y}) = \min_{\boldsymbol{c},\boldsymbol{u},\boldsymbol{y}} \sum_{i=1}^{N} Q_i(\boldsymbol{c}_i,\boldsymbol{u}_i,\boldsymbol{y}_i) \tag{1}$$

同时需要满足约束

1) 子过程方程

$$y_{ik} = F_{ik}(\boldsymbol{c}_i,\boldsymbol{u}_i),\ k \in \overline{1,n_i},\ i \in \overline{1,N} \tag{2}$$

2) 关联耦合方程

$$P(\boldsymbol{u},\boldsymbol{y}) = \sum_{i=1}^{N} P_i(\boldsymbol{u}_i,\boldsymbol{y}_i) = 0 \tag{3}$$

其中比较常见的形式为 $\boldsymbol{u}=H\boldsymbol{y}$;

3) 局部子过程约束

$$(\boldsymbol{c}_i,\boldsymbol{u}_i) \in CU_i \triangleq \{(\boldsymbol{c}_i,\boldsymbol{u}_i) \in C_i \times U_i : G_i(\boldsymbol{c}_i,\boldsymbol{u}_i) \in S_i\},\quad i = \overline{1,N} \tag{4}$$

$$(\boldsymbol{c},\boldsymbol{u}) \in CU = CU_1 \times \cdots \times CU_N$$

上式中 $\boldsymbol{y}_i \in Y_i$ 是子过程 i 输出;n_i 是子过程 i 的输出维数;$\boldsymbol{c}_i \in C_i$ 是控制输入;$\boldsymbol{u}_i \in U_i$ 是关联输入;$F_i: C_i \times U_i \to Y_i$ 是子过程 i 的输入-输出映射;$\boldsymbol{y} \triangleq (\boldsymbol{y}_1,\cdots,\boldsymbol{y}_N) \in Y_1 \times \cdots \times Y_N \triangleq Y$;$P_i: U_i \times Y_i \to P$;$H=[H_1 \cdots H_N]^{\mathrm{T}}$;$\boldsymbol{u} \triangleq (\boldsymbol{u}_1,\cdots,\boldsymbol{u}_N) \in U_1 \times \cdots \times U_N \triangleq U$;$\boldsymbol{u}_i = H_i\boldsymbol{y}$,$H_i$ 是子过程 i 的关联子矩阵;$Q_i: C_i \times U_i \times Y_i \to R$ 为子过程 i 的性能指标;C_i, U_i, Y_i, P 都是实 Hilbert 空间.

对于上述的精确 OP 问题,由于存在模型-实际差异,即式(2)所描述的模型输入输出映射与实际过程的输入输出映射

$$y_i^* = F_i^*(c_i, u_i^*)\ i \in \overline{1,N} \tag{5}$$

有差异，造成最终的解达不到实际过程最优解. 为有效地解决这个问题，提出将子过程模型模糊化的方法即将式(2)中每一方程的各项系数处理为模糊数. 子过程 i 的第 k 方程模糊后可表示为

$$\begin{aligned} y_{ik} - \widetilde{F}_{ik}(\boldsymbol{c}_i, \boldsymbol{u}_i, \widetilde{\boldsymbol{a}}_{ik}) = y_{ik} - (\widetilde{a}_{ik1} \otimes c_{i1} \oplus \cdots \oplus \\ \widetilde{a}_{ikp} \otimes c_{ip} \oplus \widetilde{a}_{ikp+1} \otimes u_{i1} \oplus \cdots \oplus \widetilde{a}_{ikp+q} \otimes u_{iq}) \,\widetilde{\subset}\, \widetilde{0},\ k \in \overline{1,n_i} \end{aligned} \tag{6}$$

式中 $\widetilde{\boldsymbol{a}}_{ik} \triangleq (\widetilde{a}_{ik1}, \cdots, \widetilde{a}_{ikp+q})$和 $\widetilde{0}$ 是 L-R 型模糊数，n_i子过程 i 的输出维数，$\otimes$和$\oplus$分别表示基于扩张原理的模糊数的数量积与数量和运算.

称具有式(6)这样模糊等式约束的 OP 问题为 FOP(fuzzy optimization problem). 要解 FOP 问题需将式(6)去模糊. 对于给定的满足度函数的 α-水平截集，模糊方程 $y_{ik} - \widetilde{F}_{ik}(\boldsymbol{c}_i, \boldsymbol{u}_i, \widetilde{\boldsymbol{a}}_{ik}) \,\widetilde{\subset}\, \widetilde{0}$ 可化为不等式

$$M_{ik}(\boldsymbol{c}_i, \boldsymbol{u}_i, \boldsymbol{y}_{ik}) \geqslant \alpha \tag{7}$$

去模糊后的 FOP 问题在本文中称为 α-OP 问题，表述如下：

找到控制量 $\hat{\boldsymbol{c}}$，关联输入量 $\hat{\boldsymbol{u}}$，关联输出量 $\hat{\boldsymbol{y}}$，使过程的整体性能指标达到最优

$$Q(\hat{\boldsymbol{c}}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}) = \min_{c,u,y} Q(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{y}) = \min_{c,u,y} \sum_{i=1}^{N} Q_i(\boldsymbol{c}_i, \boldsymbol{u}_i, \boldsymbol{y}_i)$$

同时需要满足约束

1)
$$(\boldsymbol{c}_i, \boldsymbol{u}_i) \in CU_i \cap \Omega_{i\alpha} \triangleq \Phi_{i\alpha}\ i \in \overline{1,N} \tag{8}$$

2) 关联耦合方程

$$P(\boldsymbol{u}, \boldsymbol{y}) = \sum_{i=1}^{N} P_i(\boldsymbol{u}_i, \boldsymbol{y}_i) = 0$$

其中比较常见的形式为 $\boldsymbol{u} = H\boldsymbol{y}$，式中 $\Omega_{i\alpha} = \Omega_{i1\alpha} \cap, \cdots, \cap \Omega_{in_i\alpha}$，$\Omega_{ik\alpha}$ 表示子过程 i 的第 k 个方程的满足度函数的 α-水平截集，n_i是子过程 i 的输出维数.

在本文中称用于解 α-OP 问题的关联平衡法为基于模糊模型的关联平衡法(FIBM)，描述如下

定义全局问题的 Lagrangian 函数为

$$L(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{y}, \boldsymbol{\lambda}) = Q(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{y}) + \langle \boldsymbol{\lambda}, P(\boldsymbol{u}, \boldsymbol{y}) \rangle \tag{9}$$

式中 $\boldsymbol{\lambda} \in P$ 为 Lagrangian 乘子向量，$\langle \cdot, \cdot \rangle$为内积算子；假设问题是可分解的，则 Lagrangian 函数可分解为

$$L(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{y}, \boldsymbol{\lambda}) = L(\boldsymbol{w}, \boldsymbol{\lambda}) = \sum_{i=1}^{N} L_i(\boldsymbol{c}_i, \boldsymbol{u}_i, \boldsymbol{y}_i, \boldsymbol{\lambda}) = \sum_{i=1}^{N} L_i(\boldsymbol{w}_i, \boldsymbol{\lambda}) \tag{10}$$

式中 $\boldsymbol{w} \triangleq (\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{y})$，$\boldsymbol{w}_i \triangleq (\boldsymbol{c}_i, \boldsymbol{u}_i, \boldsymbol{y}_i)$；FIBM 的下层局部决策单元问题($\mathrm{LP}_i$)为

$$(\mathrm{LP}_i)\begin{cases} \text{对于给定的 } \boldsymbol{\lambda} \in P,\ \alpha \in R \text{ 找到 } \hat{\boldsymbol{w}}_i(\boldsymbol{\lambda}) = (\hat{\boldsymbol{c}}_i(\lambda), \hat{\boldsymbol{u}}_i(\lambda), \hat{\boldsymbol{y}}_i(\boldsymbol{\lambda})) \text{ 使得} \\ L_i(\hat{\boldsymbol{w}}_i(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = \min\limits_{w_i \in \Phi_{i\alpha}} L_i(\boldsymbol{w}_i, \boldsymbol{\lambda}) \end{cases} \tag{11}$$

将式(11)的解定义为 $\overline{\Phi}_{i\alpha}(\boldsymbol{\lambda})$，则上层的协调单元问题(CP)为

$$(\mathrm{CP})\begin{cases} \text{找到 } \hat{\boldsymbol{\lambda}} \in P \text{ 使解集 } \overline{\Phi}_{\alpha}(\hat{\boldsymbol{\lambda}}) = \overline{\Phi}_{1\alpha}(\hat{\boldsymbol{\lambda}}) \times \cdots \times \overline{\Phi}_{N\alpha}(\hat{\boldsymbol{\lambda}}) \text{ 非空，} \\ \text{并且 } \forall\, \boldsymbol{w}(\hat{\boldsymbol{\lambda}}) \in \overline{\Phi}_{\alpha}(\hat{\boldsymbol{\lambda}}) \text{，成立 } P(\boldsymbol{u}, \boldsymbol{y}) = 0 \end{cases} \tag{12}$$

3 基于模糊模型的关联平衡法的可应用性证明

定义 1. 称 FIBM 可用于解 α-OP 问题,如果由式(12)描述的 CP 问题至少有一个解 $\hat{\boldsymbol{\lambda}}$.

引理 1. 若 LP 问题可解,则 Φ_α是非空集合.

证明. 由满足度函数和模糊数的定义可知,当 $\alpha=1$,并且各模糊参数都取为模糊单点时,由 $M_{ik}(\boldsymbol{c}_i,\boldsymbol{u}_i,y_{ik})\geqslant\alpha$ 描述的约束就转变为原来的精确的等式约束(即子过程数学模型).因此,若原精确问题 LP 问题可解,则 Φ_α必是非空集合. 证毕.

引理 2. 若集合 CU 是凸集,则 Φ_α也是凸集.

证明. 首先证明 $\Omega_{i\alpha}$是凸集. 为简化符号,定义 $\boldsymbol{x}=(\boldsymbol{c}_i,\boldsymbol{u}_i,y_{ik})$,$\tilde{\boldsymbol{a}}=(\tilde{a}_{ik1},\cdots,\tilde{a}_{ikp+q})$. 设 $(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)\in\Omega_{ik\alpha}$,$(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)\in\Omega_{ik\alpha}$,则由约束集合定义

$$\Rightarrow M_{ik}(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)\geqslant\alpha,\ M_{ik}(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)\geqslant\alpha$$

设在$(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)$和$(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)$两组参数下的模糊等式约束 $y_{ik}-\tilde{F}_{ik}(\boldsymbol{c}_i,\boldsymbol{u}_i,\tilde{\boldsymbol{a}}_{ik})\,\widetilde{\subseteq}\,\tilde{0}$ 的左边经过模糊数的运算后得到的隶属函数分别为$\mu_1(F)$和$\mu_2(F)$. 给定常数 $0<\lambda<1$,因为$(1-\lambda)>0$,且 $\tilde{a}_1$,$\tilde{a}_2$为 L-R 型模糊数,则由 L-R 型模糊数运算的线性性质得到,由$(1-\lambda)(x_1,\tilde{a}_1)$这组参数运算得出的隶属函数为$(1-\lambda)\mu_1(F)$;同理可得,由 $\lambda(x_2,\tilde{a}_2)$这组参数得出的隶属函数为 $\lambda\mu_2(F)$. 所以,由$(1-\lambda)(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)+\lambda(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)$得出的隶属函数为$(1-\lambda)\mu_1(F)+\lambda\mu_2(F)$. 根据满足度函数的定义有

$$M_{ik}((1-\lambda)(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)+\lambda(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2))=\frac{\int_{-\infty}^{+\infty}((1-\lambda)\mu_1(F)+\lambda\mu_2(F))\psi_i(F)\mathrm{d}F}{\int_{-\infty}^{+\infty}((1-\lambda)\mu_1(F)+\lambda\mu_2(F))\mathrm{d}F}=$$

$$\frac{(1-\lambda)\int_{-\infty}^{+\infty}\mu_1(F)\psi_i(F)\mathrm{d}F+\lambda\int_{-\infty}^{+\infty}\mu_2(F)\psi_i(F)\mathrm{d}F}{(1-\lambda)\int_{-\infty}^{+\infty}\mu_1(F)\mathrm{d}F+\lambda\int_{-\infty}^{+\infty}\mu_2(F)\mathrm{d}F}$$

由 $M_{ik}(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)\geqslant\alpha$ 和 $M_{ik}(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)\geqslant\alpha$ 得 $\dfrac{\int_{-\infty}^{+\infty}\mu_1(F)\psi_i(F)\mathrm{d}F}{\int_{-\infty}^{+\infty}\mu_1(F)\mathrm{d}F}\geqslant\alpha$, $\dfrac{\int_{-\infty}^{+\infty}\mu_2(F)\psi_i(F)\mathrm{d}F}{\int_{-\infty}^{+\infty}\mu_2(F)\mathrm{d}F}\geqslant\alpha$

为书写方便,令

$$\int_{-\infty}^{+\infty}\mu_1(F)\mathrm{d}F=a,\ \int_{-\infty}^{+\infty}\mu_2(F)\mathrm{d}F=b,\ \int_{-\infty}^{+\infty}\mu_1(F)\psi_i(F)\mathrm{d}F=x,\ \int_{-\infty}^{+\infty}\mu_2(F)\psi_i(F)\mathrm{d}F=y$$

由此可得

$$a>0,\ b>0,\ x>0,\ y>0,\ \frac{x}{a}\geqslant\alpha,\ \frac{y}{b}\geqslant\alpha\Rightarrow x\geqslant a\alpha,\ y\geqslant b\alpha$$

$$\Rightarrow(1-\lambda)x\geqslant(1-\lambda)a\alpha,\lambda y\geqslant\lambda b\alpha\Rightarrow(1-\lambda)x+\lambda y\geqslant(1-\lambda)a\alpha+\lambda b\alpha$$

$$\Rightarrow\frac{(1-\lambda)x+\lambda y}{(1-\lambda)a+\lambda b}\geqslant\alpha$$

于是可以得出

$$M_{ik}((1-\lambda)(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)+\lambda(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2))\geqslant\alpha\Rightarrow(1-\lambda)(\boldsymbol{x}_1,\tilde{\boldsymbol{a}}_1)+\lambda(\boldsymbol{x}_2,\tilde{\boldsymbol{a}}_2)\in\Omega_{ik\alpha}$$

因此，$\Omega_{ik\alpha}$是凸集，$\Omega_{i\alpha}$是凸集. 又由 CU 是凸集及引理 1，可以推出 Φ_α是凸集. 证毕.

假设 1. 集合 Φ_α是紧集，映射 P 在 $U\times Y$ 上是连续的，Q 在 Φ_α上是连续的.

引理 3. 如果 1) 假设 1 成立，并且 Q 在 Φ_α上是有界的；2) $\forall \boldsymbol{h}\in P$，$\|\boldsymbol{h}\|=k_1>0$，存在 $\boldsymbol{h}_1\in P$，$\|\boldsymbol{h}_1\|\leqslant k_2<k_1$，并且存在$(\boldsymbol{u},\boldsymbol{y})\in\Phi_\alpha$使得 $P(\boldsymbol{u},\boldsymbol{y})=\boldsymbol{h}+\boldsymbol{h}_1$；那么存在 $\hat{\boldsymbol{\lambda}}\in P$，使得

$$\varphi(\hat{\boldsymbol{\lambda}})=\max_{\lambda\in P}\varphi(\boldsymbol{\lambda})$$

式中 $\varphi:P\to R$ 是 $L(\boldsymbol{w},\boldsymbol{\lambda})$的对偶函数，定义为

$$\varphi(\boldsymbol{\lambda})=\inf_{w\in\Phi_\alpha}L(\boldsymbol{w},\boldsymbol{\lambda})$$

证明. 首先证明 $\varphi(\boldsymbol{\lambda})$在 P 上是凹的.

$\forall\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in P$，$\forall w\in\Phi_\alpha$，$\forall\rho\in[0,1]$，有

$$L(\boldsymbol{w},\rho\boldsymbol{\lambda}_1+(1-\rho)\boldsymbol{\lambda}_2)=Q(\boldsymbol{w})+\rho\langle\boldsymbol{\lambda}_1,P\rangle+(1-\rho)\langle\boldsymbol{\lambda}_2,P\rangle=$$
$$\rho Q(\boldsymbol{w})+(1-\rho)Q(\boldsymbol{w})+\rho\langle\boldsymbol{\lambda}_1,P\rangle+(1-\rho)\langle\boldsymbol{\lambda}_2,P\rangle=\rho L(\boldsymbol{w},\boldsymbol{\lambda}_1)+(1-\rho)L(\boldsymbol{w},\boldsymbol{\lambda}_2)$$

因此

$$\varphi(\rho\boldsymbol{\lambda}_1+(1-\rho)\boldsymbol{\lambda}_2)=\inf_{w\in\Phi_\alpha}L(\boldsymbol{w},\rho\boldsymbol{\lambda}_1+(1-\rho)\boldsymbol{\lambda}_2)\geqslant$$
$$\rho\inf_{w\in\Phi_\alpha}L(\boldsymbol{w},\boldsymbol{\lambda}_1)+(1-\rho)\inf_{w\in\Phi_\alpha}L(\boldsymbol{w},\boldsymbol{\lambda}_2)=\rho\varphi(\boldsymbol{\lambda}_1)+(1-\rho)\varphi(\boldsymbol{\lambda}_2)$$

所以 $\varphi(\boldsymbol{\lambda})$在 P 上是凹的，同时 $\forall\boldsymbol{\lambda}\in P$ 集合 Φ_α是非空的且 $\varphi(\boldsymbol{\lambda})>-\infty$. 由于 $\forall\boldsymbol{\lambda}\in P$ 可以找到 $\boldsymbol{h}\in P(\|\boldsymbol{h}\|=k_1)$，使得$\langle\boldsymbol{\lambda},\boldsymbol{h}\rangle=-k_1\|\boldsymbol{\lambda}\|$. 由条件 2) $P(\boldsymbol{u},\boldsymbol{y})=\boldsymbol{h}+\boldsymbol{h}_1$，$\|\boldsymbol{h}_1\|\leqslant k_2$，得到

$$\varphi(\boldsymbol{\lambda})\leqslant L(\boldsymbol{w},\boldsymbol{\lambda})=Q(\boldsymbol{w})+\langle\lambda,P(\boldsymbol{u},\boldsymbol{y})\rangle=Q(\boldsymbol{w})+\langle\boldsymbol{\lambda},\boldsymbol{h}+\boldsymbol{h}_1\rangle=$$
$$Q(\boldsymbol{w})+\langle\boldsymbol{\lambda},\boldsymbol{h}\rangle+\langle\boldsymbol{\lambda},\boldsymbol{h}_1\rangle\leqslant Q(\boldsymbol{w})-k_1\|\boldsymbol{\lambda}\|+\|\boldsymbol{\lambda}\|\ \|\boldsymbol{h}_1\|\leqslant k'-k_1\|\boldsymbol{\lambda}\|+k_2\|\boldsymbol{\lambda}\|$$

式中 $k'=\sup_{\Phi_\alpha}Q(\boldsymbol{w})$. 另外定义 $k''=\min_{\Phi_\alpha}Q(\boldsymbol{w})$，$r=\dfrac{k'-k''}{k_1-k_2}$，所以当$\|\lambda\|>r$ 时，有

$$\varphi(\boldsymbol{\lambda})\leqslant k'-(k_1-k_2)\|\boldsymbol{\lambda}\|<k'-(k'-k'')=k''=\varphi(0)$$

因此 $\varphi(\boldsymbol{\lambda})$在 P 上也是弱上半连续的，这意味着下式成立

$$\sup_{\lambda\in P}\varphi(\boldsymbol{\lambda})=\sup_{\lambda\in\bar{B}(0;r)}\varphi(\boldsymbol{\lambda})$$

$B(0;r)=\{\boldsymbol{\lambda}\in P:\|\boldsymbol{\lambda}\|<r\}$，$\bar{B}(0;r)$是 $B(0;r)$的闭包，而球 $\bar{B}(0;r)$是紧的，同时 $\varphi(\boldsymbol{\lambda})$又是上半连续的，因此存在 $\hat{\boldsymbol{\lambda}}$ 使得

$$\varphi(\hat{\boldsymbol{\lambda}})=\max_{\lambda\in\bar{B}(0;r)}\varphi(\boldsymbol{\lambda})=\max_{\lambda\in P}\varphi(\boldsymbol{\lambda})$$

并且$\|\hat{\boldsymbol{\lambda}}\|\leqslant r$. 证毕.

定理 1. 如果 1) 假设 1 成立；2) Q 在 Φ_α上是严格凸的，并且 P 是仿射映射；那么由式(11)，式(12)定义的基于模糊模型的关联平衡法可用于解 α-OP 问题.

证明. 由条件 2)推得 $\forall\boldsymbol{\lambda}\in P$，$L(\boldsymbol{w},\boldsymbol{\lambda})$在 Φ_α上是严格二次凸的. 又由引理 2 可知由式(11)和(12)定义的基于模糊模型的关联平衡法是一个凸规划问题，并且由引理 3 可知它有解，因此由式(11)和(12)定义的基于模糊模型的关联平衡法的解集 $V(\bar{\Phi}_\alpha(\hat{\boldsymbol{\lambda}}))$非空并且是单点集. 同时，由定义 1 可知结论成立. 证毕.

4 算法收敛性证明

根据上一节的结论，基于模糊模型的关联平衡法(FIBM)得到的解就是 α-OP 问题的

解.其下层的局部决策单元任务是一个普通的非线性优化问题,用适当的非线性优化算法都可以得到其解,将其解表示为 $\hat{\boldsymbol{w}}(\boldsymbol{\lambda})=(\hat{\boldsymbol{c}}(\boldsymbol{\lambda}),\hat{\boldsymbol{u}}(\boldsymbol{\lambda}),\hat{\boldsymbol{y}}(\boldsymbol{\lambda}))$.上层的协调问题可用梯度法求解,由引理 2 知 $\hat{\boldsymbol{\lambda}}=\arg\max\limits_{\lambda\in P}\varphi(\boldsymbol{\lambda})$,又由定理 1 知局部决策单元任务是一个凸规划问题,$\forall\boldsymbol{\lambda}\in P$ 有唯一解 $\hat{\boldsymbol{w}}(\boldsymbol{\lambda})$.因此,$\forall\boldsymbol{\lambda}\in P$,$\nabla\varphi(\boldsymbol{\lambda})=P(\hat{\boldsymbol{u}},\hat{\boldsymbol{y}})$.所以协调器可按下式产生迭代序列 $\{\boldsymbol{\lambda}^n\}$

$$\boldsymbol{\lambda}^{n+1}=\boldsymbol{\lambda}^n+\varepsilon AP(\hat{\boldsymbol{u}}(\boldsymbol{\lambda}^n),\hat{\boldsymbol{y}}(\boldsymbol{\lambda}^n)) \tag{13}$$

式中 $A:P\to P$ 是一正定对称矩阵,$\varepsilon>0$ 是步长系数.

令 $D(\boldsymbol{\lambda}^n)=-P(\hat{\boldsymbol{u}}(\boldsymbol{\lambda}^n),\hat{\boldsymbol{y}}(\boldsymbol{\lambda}^n))$,由式(13)产生的迭代增量为 $\boldsymbol{h}(\boldsymbol{\lambda}^n)=-\varepsilon AD(\boldsymbol{\lambda}^n)$.为书写方便,在以下的定理及证明当中,省去上标 n.

定理 2. $\forall\boldsymbol{\lambda},\boldsymbol{\lambda}+\boldsymbol{h}\in P_0\subset P$($P_0$是 P 的开凸子集),若 A 是一正定对称阵,则由式(13)描述的迭代序列 $\{\boldsymbol{\lambda}^n\}$ 在 $0<\varepsilon<\dfrac{\rho_1^2\delta_2}{\rho_2^2\delta_1^2}<1$ 的条件下收敛.

证明. 因为 A 是正定对称阵,可以在 P 上定义 A 内积

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle_A=\langle\boldsymbol{x},A\boldsymbol{y}\rangle\ \forall\boldsymbol{x},\boldsymbol{y}\in P$$

由此内积定义可诱导出 P 上的 A 范数 $\|\cdot\|_A$

$$\|\boldsymbol{x}\|_A\triangleq\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle_A}=\sqrt{\langle\boldsymbol{x},A\boldsymbol{x}\rangle},\quad\forall\boldsymbol{x}\in P$$

因为 A 是正定对称阵,所以 $\forall\boldsymbol{\lambda}\in P$ 满足下面的不等式

$$\rho_1\|\boldsymbol{\lambda}\|^2\leqslant\langle\boldsymbol{\lambda},A\boldsymbol{\lambda}\rangle\leqslant\rho_2\|\boldsymbol{\lambda}\|^2,\quad\rho_1>0$$

显然下面的不等式亦成立

$$\rho_1\|\boldsymbol{\lambda}\|_A^2\leqslant\langle\boldsymbol{\lambda},A\boldsymbol{\lambda}\rangle=\|\boldsymbol{\lambda}\|_A^2\leqslant\|\boldsymbol{\lambda}\|_A\|A\boldsymbol{\lambda}\|_A\leqslant\|A\|_A\|\boldsymbol{\lambda}\|_A^2\leqslant\rho_2\|\boldsymbol{\lambda}\|_A^2$$

$$\|A\boldsymbol{\lambda}\|_A^2\leqslant(\|A\|_A\|\boldsymbol{\lambda}\|_A)^2\leqslant\rho_2^2\|\boldsymbol{\lambda}\|_A^2,\quad\|A\boldsymbol{\lambda}\|_A\geqslant\rho_1\|\boldsymbol{\lambda}\|_A$$

由以上 A 内积及 A 范数的定义可得不等式

$$\forall\boldsymbol{x},\boldsymbol{y}\in P\ \langle\boldsymbol{x},A\boldsymbol{y}\rangle=\langle\boldsymbol{x},\boldsymbol{y}\rangle_A\leqslant\|\boldsymbol{x}\|_A\|\boldsymbol{y}\|_A \tag{14}$$

任取两常数 δ_1,δ_2满足 $\delta_1>0$,$\delta_1>\delta_2$,$\delta_1^2\rho_2^2>\delta_2^2\rho_1^4$,使下列不等式成立:

a) $\|D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda})\|\leqslant\delta_1\|\boldsymbol{h}\|$

b) $\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle\geqslant\delta_2\|\boldsymbol{h}\|^2$

由此可以确定内积 $\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle$ 的上界

$$\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle=\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),-\varepsilon AD(\boldsymbol{\lambda})\rangle=$$
$$\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\varepsilon AD(\boldsymbol{\lambda}+\boldsymbol{h})-\varepsilon AD(\boldsymbol{\lambda})-\varepsilon AD(\boldsymbol{\lambda}+\boldsymbol{h})\rangle=$$
$$\varepsilon\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),A[D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda})]\rangle-\varepsilon\langle D(\boldsymbol{\lambda}+\boldsymbol{h}),AD(\boldsymbol{\lambda}+\boldsymbol{h})\rangle+\varepsilon\langle D(\boldsymbol{\lambda}),AD(\boldsymbol{\lambda}+\boldsymbol{h})\rangle$$

由 A 范数的定义 $\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),A[D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda})]\rangle=\|D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda})\|_A^2$,$\langle D(\boldsymbol{\lambda}+\boldsymbol{h}),AD(\boldsymbol{\lambda}+\boldsymbol{h})\rangle=\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2$,由不等式(14)可知 $\langle D(\boldsymbol{\lambda}),AD(\boldsymbol{\lambda}+\boldsymbol{h})\rangle\leqslant\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A$,于是可得

$$\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle\leqslant$$
$$\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda})\|_A^2-\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2+\varepsilon\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A\leqslant$$
$$\varepsilon\delta_1^2\|\boldsymbol{h}\|_A^2-\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2+\varepsilon\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A\leqslant$$
$$\varepsilon^3\delta_1^2\|AD(\boldsymbol{\lambda})\|_A^2-\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2+\varepsilon\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A\leqslant$$
$$\varepsilon^3\delta_1^2\rho_2^2\|D(\boldsymbol{\lambda})\|_A^2-\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2+\varepsilon\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A$$

然后确定内积 $\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle$ 的下界

$$\langle D(\boldsymbol{\lambda}+\boldsymbol{h})-D(\boldsymbol{\lambda}),\boldsymbol{h}\rangle\geqslant\delta_2\|\boldsymbol{h}\|_A^2=\delta_2\varepsilon^2\|AD(\boldsymbol{\lambda})\|_A^2\geqslant\delta_2\varepsilon^2\rho_1^2\|D(\boldsymbol{\lambda})\|_A^2$$

结合以上两个不等式，得

$$-\varepsilon\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2+\varepsilon\|D(\boldsymbol{\lambda})\|_A\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A+\varepsilon^3\delta_1^2\rho_2^2\|D(\boldsymbol{\lambda})\|_A^2\geqslant\delta_2\varepsilon^2\rho_1^2\|D(\boldsymbol{\lambda})\|_A^2$$

$$\Rightarrow\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A^2-\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A\|D(\boldsymbol{\lambda})\|_A+\frac{1}{4}\|D(\boldsymbol{\lambda})\|_A\leqslant(\frac{1}{4}+\varepsilon^2\delta_1^2\rho_2^2-\varepsilon\delta_2\rho_1^2)\|D(\boldsymbol{\lambda})\|_A^2$$

$$\Rightarrow(\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A-\frac{1}{2}\|D(\boldsymbol{\lambda})\|_A)^2\leqslant\Delta\|D(\boldsymbol{\lambda})\|_A^2$$

式中 $\Delta=\varepsilon^2\rho_2^2\delta_1^2-\varepsilon\delta_2\rho_1^2+\frac{1}{4}$，因为 $\rho_2^2\delta_1^2>\rho_1^4\delta_2^2$，则由 $\varepsilon^2-\varepsilon\frac{\rho_1^2\delta_2}{\rho_2^2\delta_1^2}+\frac{1}{4\rho_2^2\delta_1^2}=\frac{\Delta}{\rho_2^2\delta_1^2}\Rightarrow$

$$\left(\varepsilon-\frac{\rho_1^2\delta_2}{2\rho_2^2\delta_1^2}\right)^2+\frac{\rho_2^2\delta_1^2-\rho_1^4\delta_2^2}{4\rho_2^4\delta_1^4}=\frac{\Delta}{\rho_2^2\delta_1^2}\Rightarrow\Delta>0$$

因此，ε 取任何值，$\Delta>0$. 于是可以得到 $\|D(\boldsymbol{\lambda}+\boldsymbol{h})\|_A\leqslant\left(\frac{1}{2}+\sqrt{\Delta}\right)\|D(\boldsymbol{\lambda})\|_A$. 要使迭代收敛，必须 $\sqrt{\Delta}<\frac{1}{2}$. 由此得 $\Delta<\frac{1}{4}\Rightarrow\varepsilon^2\rho_2^2\delta_1^2-\varepsilon\rho_1^2\delta_2<0\Rightarrow\varepsilon<\frac{\rho_1^2\delta_2}{\rho_2^2\delta_1^2}$

于是，当选择 $0<\varepsilon<\frac{\rho_1^2\delta_2}{\rho_2^2\delta_1^2}<1$，可使由式(13)描述的迭代序列 $\{\boldsymbol{\lambda}^n\}$ 收敛.　　证毕.

5　结束语

本文在证明基于模糊模型的大系统关联平衡法的收敛性时，并未采用 Zangwill 的全局收敛性定理，这是因为关联平衡法是由协调器算法和决策单元算法组合而成，证明它是一个闭的点集-映射则较困难. 采用本文的证明方法，可以避开这一繁琐的证明. 另外，采用本文的证明方法可以同时给出保证迭代收敛的系数取值范围.

References

1　Gu Jia-Chen, Wan Bai-Wu, Guan Xiao-Hong. Interaction balance method for large-scale industrial processes with fuzzy parameters. *Acta Automatica Sinica*, 2002, **28**(4): 569～574(in Chinese)

2　Tanaka H, Asai K. Fuzzy linear programming problems with fuzzy numbers. *Fuzzy Sets and Systems*, 1984, **13**(1): 1～10

3　Kassem M A, Ammar E E. Stability of multiobjective nonlinear programming problems with fuzzy parameters in the constraints. *Fuzzy Sets and Systems*, 1995, **74**(3): 343～351

4　Kassem M A, Ammar E E. A parametric study of multiobjective NLP problems with fuzzy parameters in the objective functions. *Fuzzy Sets and Systems*, 1996, **80**(2): 187～196

5　Zangwill W I. Nonlinear Programming: A Unified Approach. Englewood Cliffs, NJ: Prentice-Hall, 1969

顾佳晨　西安交通大学系统工程研究所获得博士学位. 主要研究领域为智能控制、大系统的智能优化控制.

(**GU Jia-Chen**　Received his Ph. D. degree from Institute of Systems Engineering, Xi'an Jiaotong University in 2001. His research interests include intelligent control and data mining applications.)

万百五　简介见本刊 1999 年第 25 卷第 1 期.

(**WAN Bai-Wu**　Refer to the biography after the paper on Vol. 29, No. 1 of this journal.)