

一种基于多代理模型的混合整数规划优化方法

吕志明, 王霖青, 赵 珺[†], 王 伟

(大连理工大学控制科学与工程学院, 辽宁 大连 116024)

摘 要: 提出一种基于多代理模型的优化方法, 求解混合整数规划问题. 首先, 基于群智能优化策略提出一种基于多群体协作模型的采样方法, 保证候选解的正确性和多样性; 其次, 采用基于数据并行的高斯过程建模方法, 在线构造局部代理模型; 再次, 通过多代理模型对候选解进行预筛选, 实现与粒子群算法的协同优化; 最后, 通过 14 个测试问题和一个基于数据驱动模型参数选取问题, 验证所提出方法的有效性.

关键词: 混合整数; 多代理; 粒子群; 高斯过程

中图分类号: TP206 **文献标志码:** A

A multi-surrogates algorithm for mixed-integer programming problems

LYU Zhi-ming, WANG Lin-qing, ZHAO Jun[†], WANG Wei

(School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: A multi-surrogates algorithm is developed to deal with the mixed-integer programming problems. Firstly, a sampling method based on the model of the multi-swarm PSO is developed to ensure the accuracy and diversity of the the samples. Furthermore, the local surrogate models are constructed by an online modeling method based on the data parallel approach. Then, the collaborative optimization is carried out based on the preselecting strategy and PSO. Finally, the effectiveness of the proposed method are verified by the 14 test problems and 1 data driven model parameter selection problems.

Keywords: mixed-integer; multi-surrogates; PSO; Gaussion process

0 引 言

当前, 在最优决策、工程设计、组合优化和能源调度等工程领域, 存在着大量的混合整数规划(Mixed-integer programming, MIP)问题. 这些问题通常是 NP-完全问题, 计算复杂度高. 此外, 在实际应用中, 目标问题通常只有实时的输入和输出, 没有确定性的数学描述, 而且给定输入后, 得到输出响应的的时间很长, 该类问题称为黑箱问题. 近些年, 基于代理模型的全局优化算法在 MIP 黑箱问题中得到了广泛的关注, 并应用于实际的工程优化问题中^[1-3].

基于代理模型的全局优化算法主要是通过自适应采样, 使效益函数最大化. 文献[1]通过优化两个辅助函数进行采样. 但是, 辅助优化过程本身就是一个求解混合整数规划问题的过程, 因此, 优化器设计的合理性影响着所提算法的有效性. 为了避免上述问

题, 文献[2]提出了基于全局和局部网格搜索的协同采样策略; 文献[3]提出了一种随机采样策略. 但是, 上述两种贪婪的采样策略会产生大量的采样点, 进而增加代理模型的训练和预测成本. 为了缓解代理模型构造成本高的问题, 多代理模型技术已成为当前的研究热点^[4]. 文献[5]采用基于数据并行的高斯过程建模方法构造多个局部代理模型; 文献[6]根据分而治之的思想提出了混合专家方法. 但是文献[5]和文献[6]针对连续优化问题, 只考虑了离线建模方法.

针对 MIP 黑箱问题, 本文提出一种基于多代理模型的优化算法. 该方法基于多群体协作, 采用多种进化策略进行联合采样. 此外, 构造一种基于数据并行方法的高斯过程回归模型在线建模方法. 该方法在构造多个局部代理模型的同时自适应地调节局部代理模型的个数. 在序贯优化的过程中, 粒子群算法与

收稿日期: 2017-08-21; 修回日期: 2017-11-26.

基金项目: 国家自然科学基金项目(61603069, 61473056, 61533005, 61522304, U1560102); 国家科技支撑计划项目(2015BAF22B01); 中央高校基本科研业务费专项资金项目(DUT17ZD231).

责任编委: 刘德荣.

作者简介: 吕志明(1988—), 男, 博士生, 从事流程工业过程建模与优化、智能优化算法设计的研究; 王伟(1955—), 男, 教授, 博士生导师, 从事自适应控制、现代集成制造系统和流程工业过程控制等研究.

[†]通讯作者. E-mail: zhaoj@dlut.edu.cn.

代理模型方法协同优化获得的最佳可行解,一方面用于引导采样方法的采样方向,另一方面用于更新代理模型.将本文方法分别用于14个标准测试函数和一个回归模型的参数优化问题,所得结果验证了本文方法的有效性.

1 混合整数规划问题

考虑如下的混合整数规划单目标约束优化问题:

$$\min f(\mathbf{X}). \tag{1}$$

$$\text{s.t. } b_i(\mathbf{X}) \leq 0, i = 1, 2, \dots, m;$$

$$c_j(\mathbf{X}) = 0, j = 1, 2, \dots, n. \tag{2}$$

其中: $f(\mathbf{X})$ 为目标函数的输出响应值;输入变量 $\mathbf{X} = [x_1 \ x_2 \ \dots \ x_p \ x_{p+1} \ \dots \ x_q]$ 的前 p 个变量为离散变量,剩余 $q - p$ 个变量为连续变量; m 和 n 分别为不等式和等式约束的个数.

1.1 离散化处理

本文采用最邻近顶点方法^[7]对属于离散域的前 p 个变量进行离散化处理,有

$$x^{ij} = \begin{cases} x^{iL}, & |x^{ij} - x^{iL}| \leq |x^{ij} - x^{iU}|; \\ x^{iU}, & \text{otherwise.} \end{cases} \tag{3}$$

其中: x^{ij} 为粒子 i 第 j 维的位置变量, x^{iL} 和 x^{iU} 为第 i 个局部超立方体的边界.

1.2 约束处理

为了惩罚约束违背,采用如下的惩罚函数^[8]:

$$F(\mathbf{X}) = f_{\max} + R \sum_{i=1}^m [\min(0, b_i(\mathbf{X}))]^2 + \sum_{j=1}^n [c_j(\mathbf{X})]^2. \tag{4}$$

其中: f_{\max} 为当前可行解中最大目标函数值;式(4)中的第2项为惩罚项, $R > 0$ 为惩罚因子.

2 基于多代理模型的混合整数规划方法

假设目标函数 $f(\mathbf{X})$ 精确评价的时间成本很高,为了在有限的计算成本下获得最优解,本文提出一种基于多代理模型的混合整数规划方法,该方法的结构如图1所示.

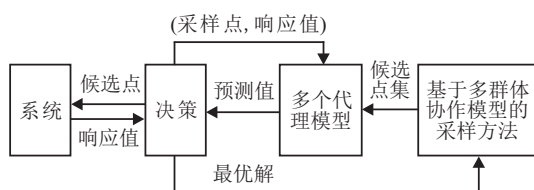


图1 基于多代理模型优化方法的结构

2.1 基于多群体协作模型的采样方法

为了保证采样的准确性和样本分布的多样性,本文提出一种基于多群体协作模型的采样方法.

本文设计一种基于主从式结构的多群体协作模型.考虑到种群的拓扑结构决定了粒子学习样本的选择^[9],在进化过程中,主种群采用具有收缩因子的PSO(CPSO),结合主、从种群的经验,负责局部搜索;从种群分别采用CPSO,基于全局版本拓扑结构的PSO(GPSO)和基于局部版本拓扑结构的PSO(LPSO)独立进化,负责在解空间中全局搜索.基于该多群体协作模型,本文提出一种采样方法,其中,主种群负责引导采样方向,该种群的适应度值采用真实的函数进行计算;而从种群负责采样,该种群的适应度值采用代理模型进行预测.

在迭代时刻 t ,CPSO一方面用于寻优,另一方面为GPSO和LPSO提供初始种群.在迭代时刻 $t + 1$,根据CPSO、GPSO和LPSO的速度和位置更新策略^[9],得到采样策略如下:

$$v_1^{ij} = \chi \{v_1^{ij}(t) + c_1 r_1 [\text{pbest}^{ij}(t) - x^{ij}(t)] + c_2 r_2 [\text{gbest}^j(t) - x^{ij}(t)]\}, \tag{5}$$

$$v_2^{ij} = v_1^{ij}(t) + c_1 r_1 [\text{pbest}^{ij}(t) - x^{ij}(t)] + c_2 r_2 [\text{gbest}^j(t) - x^{ij}(t)], \tag{6}$$

$$v_3^{ij} = \chi \{v_3^{ij}(t) + c_1 r_1 [\text{pbest}^{ij}(t) - x^{ij}(t)] + c_2 r_2 [\text{gbest}_{\text{neighbor}_i}^j(t) - x^{ij}(t)]\}, \tag{7}$$

$$x_1^{ij} = x^{ij}(t) + v_1^{ij}, \tag{8}$$

$$x_2^{ij} = x^{ij}(t) + v_2^{ij}, \tag{9}$$

$$x_3^{ij} = x^{ij}(t) + v_3^{ij}, \tag{10}$$

$$\mathbf{X}_{ps} = \{x_1^{ij}, x_2^{ij}, x_3^{ij} | i = 1 : \text{NP}, j = 1 : q\}. \tag{11}$$

其中: \mathbf{X}_{ps} 为采样得到的候选解集; v_1^{ij} , v_2^{ij} 和 v_3^{ij} 分别表示在迭代时刻 $t + 1$,粒子 i 第 j 维在CPSO、GPSO和LPSO三种速度更新策略下得到的速度矢量; x_1^{ij} , x_2^{ij} 和 x_3^{ij} 分别表示在迭代时刻 $t + 1$,根据三种不同进化策略得到的粒子位置矢量; $\text{pbest}(t)$ 为粒子所经历的最优位置; $\text{gbest}(t)$ 为种群中最优粒子的位置; $\text{gbest}_{\text{neighbor}_i}^j(t)$ 为粒子 i 邻居拓扑中最优粒子的位置;收缩因子 $\chi = \frac{2k}{2 - \phi - \sqrt{\phi^2 - 4\phi}}$, $\phi = c_1 + c_2$; $\phi > 4$; c_1 和 c_2 为学习因子; r_1 和 r_2 为在 $(0, 1]$ 上均匀分布的随机数.

为了保证候选解的采样密度,从种群进化 μ 代,得到候选解集合描述如下:

$$\mathbf{X}_{pss} = \{\mathbf{X}_{ps}\}_{1:\mu}. \tag{12}$$

根据约束条件(2),可得粒子位置 $\mathbf{X}_{pss} = \{\mathbf{X}_{fs}, \mathbf{X}_{ifs}\}$.其中: \mathbf{X}_{fs} 为满足约束条件的解,即可行解; \mathbf{X}_{ifs} 为不满足约束条件的解,即不可行解.对于不可

行解 \mathbf{X}_{ifs} , 通过惩罚函数(4)计算函数值 $F(\mathbf{X}_{\text{ifs}})$; 对于可行解 \mathbf{X}_{fs} , 根据式(1)计算对应的函数值 $f(\mathbf{X}_{\text{fs}})$. 在迭代时刻 t , 根据PSO的寻优历史经验, 构造经验知识库如下:

$$D_t = [D_{t-1}; \mathbf{X}_{\text{ifs}}, F(\mathbf{X}_{\text{ifs}}); \mathbf{X}_{\text{fs}}, f(\mathbf{X}_{\text{fs}})], \quad (13)$$

知识库 D_t 中可行解的个数为 N_t^{fs} .

算法1 基于粒子群拓扑结构的采样方法.

Step 1: 初始化主种群CPSO的速度和位置, 初始化参数 μ .

Step 2: 根据式(5)和(8)更新主种群的速度和位置.

Step 3: 根据式(3)对主种群初始位置进行离散化处理.

Step 4: 根据式(1)、(2)和(4)计算主种群粒子的适应度值并保存种群的个体最优和全局最优粒子信息.

Step 5: 将Step 4得到的种群作为从种群CPSO、GPSO和LPSO的初始种群. 令从种群进化的代数 $\text{sn} = 1$.

Step 6: 根据式(5)~(7)更新粒子的速度.

Step 7: 根据式(8)~(10)得到样本集 \mathbf{X}_{ps} .

Step 8: 令 $\text{sn} = \text{sn} + 1$, 判断 $\text{sn} > \mu$. 如果为真, 则停止采样, 根据式(12)得到粒子位置信息集合 \mathbf{X}_{pss} ; 否则, 返回Step 5.

2.2 构造多代理模型

为了缓解训练样本动态增加导致代理模型构造成本高的问题, 本文提出一种在线的基于数据并行的高斯过程回归模型方法.

为了将输入空间划分为多个子空间, 且子空间中含有相等数量的样本点, GeoCluster(Geometry motivated clustering)聚类算法^[10]需要求解如下的非线性规划方程:

$$\min_{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_m} \sum_{i=1}^m (m_i - m_d)^2. \quad (14)$$

其中: \mathbf{o}_i 和 m_i 为第 i 簇的聚类中心和样本量, m 为聚类的个数, m_d 为每一簇中期望的样本量.

在方程(14)中, 构造局部代理模型的样本量 m_d 是常数. 但是, 在基于代理模型的优化算法中, 总的样本量是动态增加的, 因此, 为了求解方程(14), 聚类个数自适应地更新如下:

$$M_t = \text{fix}\left(\frac{N_t}{m_d}\right). \quad (15)$$

其中: $\text{fix}(A)$ 表示将 A 中的元素向零点方向取整; N_t 表示 D_t 中样本个数; m_d 表示每个子空间中期望的样本个数, 通常 $m_d = 5q$ 且 $m_d \leq 60$.

基于式(15), 改进的GeoCluster聚类算法的聚类中心更新公式描述如下:

$$\mathbf{o}_p(k) = \mathbf{o}_p(k-1) + \alpha \delta \mathbf{v}_p, \quad (16)$$

$$\delta \mathbf{v}_p = \sum_{q=1, q \neq p}^{M_t} \left(\frac{m_q}{m_p}\right) [\mathbf{o}_q(k-1) - \mathbf{o}_p(k-1)], \quad (17)$$

其中 α 为学习因子.

通过对训练样本的聚类, 输入空间被划分为 M_t 个子空间, 其中第 s 个子空间中的样本表示为 $D_s = \{x_i, y_i | i = 1 : N_s\}$, x_i 为输入变量, y_i 为目标函数响应值, N_s 为子空间中样本的个数. 假设存在含噪声的响应函数 $y(x_{1:N_s}) = f(x_{1:N_s}) + \varepsilon$, 其中目标函数服从零均值高斯分布 $f(x_{1:N_s}) \sim \text{GP}(0, k(\cdot, \cdot))$, 噪声表示为 $\varepsilon \sim N(0, \delta_y^2)$, 采用如下的紧支撑协方差函数模型^[10]:

$$\tilde{k}(\mathbf{x}_1, \mathbf{x}_2; c(\cdot), \boldsymbol{\theta}) = \begin{cases} k(\mathbf{x}_1, \mathbf{x}_2; \boldsymbol{\theta}_s), \\ c(\mathbf{x}_1) = c(\mathbf{x}_2) = s; \\ 0, \text{ otherwise.} \end{cases} \quad (18)$$

其中: $\boldsymbol{\theta}_s$ 为第 s 个局部代理模型的超参数矢量, $c(\cdot)$ 为输入变量 \mathbf{x} 与局部代理模型索引之间的映射函数, s 为子空间的索引.

考虑 $M_t = h$ 的情况, 即输入空间被划分为 h 个不相交的子空间. 由式(18)可得如下的协方差矩阵:

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & & \mathbf{K}_s & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \mathbf{K}_h \end{bmatrix}. \quad (19)$$

其中: $\mathbf{K}_s = k(\mathbf{x}_1, \mathbf{x}_2; \boldsymbol{\theta}_s)$, $\mathbf{K}_s \in \mathbf{R}^{N_s \times N_s}$.

为了获得全局代理模型的超参数矢量 $\boldsymbol{\theta}$, 求解如下的最大似然估计问题^[11]:

$$\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = -\frac{1}{2} \log \det \mathbf{K} - \frac{1}{2} \mathbf{y}_{1:N} \mathbf{K}^{-1} \mathbf{y}_{1:N} - \frac{N}{2} \log 2\pi. \quad (20)$$

其中: $L(\boldsymbol{\theta})$ 为似然对数函数, N 为输入空间中的样本个数. 由于 $\det \mathbf{K} = \mathbf{K}_1 \times \dots \times \mathbf{K}_s \times \dots \times \mathbf{K}_h$, 似然估计问题可以分解为 h 个子问题, 即

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}_1) + \dots + L(\boldsymbol{\theta}_s) + \dots + L(\boldsymbol{\theta}_h). \quad (21)$$

显然, 关于全局样本的似然对数函数等于关于局部样本的似然对数函数之和. 其中, 最大化 $L(\boldsymbol{\theta}_s | s = 1, \dots, h)$ 的计算时间复杂度为 $\mathcal{O}(N_s^3)$, 而最大化 $L(\boldsymbol{\theta})$ 的计算时间复杂度为 $\mathcal{O}(N^3)$, 其中 $N_s \ll N$.

已知超参数 $\boldsymbol{\theta}_s$, 则对应的响应函数可以描述为

如下的高斯分布:

$$y(x_{1:N_t}) \sim N(f(x_{1:N_t}), \mathbf{K}_s + \delta_y^2 \mathbf{I}), \quad (22)$$

其中 $\mathbf{K}_s = \mathbf{k}(x_{1:N_t}, x_{1:N_t})$ 为协方差矩阵. 基于高斯噪声假设, 训练输出 $y_{1:N_t}$ 和测试输出 $f(x_*)$ 的联合概率分布^[11]为

$$\begin{bmatrix} y_{1:N_t} \\ f(x_*) \end{bmatrix} \sim N \left(f(x_{1:N_t}), \begin{bmatrix} \mathbf{K}_s & \mathbf{k}(x_{1:N_t}, x_*) \\ \mathbf{k}(x_* x_{1:N_t}) & k(x_*, x_*) \end{bmatrix} \right). \quad (23)$$

由此可以得到第 s 个 GP 模型的后验预测分布为

$$f(x_*) | D_s, x_* \sim N(\mu(x_* | D_s), \delta(x_* | D_s)). \quad (24)$$

其中: 均值和方差描述如下:

$$\mu(x_* | D_s) = \mathbf{k}(x_{1:N_t}, x_*)^T (\mathbf{K}_s + \delta_y^2 \mathbf{I})^{-1} y_{1:N_t}, \quad (25)$$

$$\delta(x_* | D_s) = k(x_*, x_*) - \mathbf{k}(x_{1:N_t}, x_*)^T \times (\mathbf{K}_s + \delta_y^2 \mathbf{I})^{-1} \mathbf{k}(x_{1:N_t}, x_*). \quad (26)$$

考虑到局部代理模型的选择是关于输入变量的函数, 因此, 当已知输入变量 x_p 时, 对应的代理模型选择如下:

$$\eta = \{\eta | \eta = \|x_p - x_j\|, x_j \in D_t\}; \quad (27)$$

$$x' = \operatorname{argmin}(\eta); \quad (28)$$

$$s = c(x'), c(\cdot) \in \{1, \dots, M_t\}; \quad (29)$$

$$\hat{f}(x_p) = \mu(x_p | D_s). \quad (30)$$

其中: η 表示输入变量 x_p 与所有样本点的欧氏距离, x' 表示与输入变量 x_p 最近的样本点, $\hat{f}(x_p)$ 表示输入变量 x_p 的预测值.

2.3 代理模型方法与粒子群算法协同优化

由 2.1 节可知, 主种群不仅为 3 种进化策略提供初始种群, 而且参与全局寻优. 因此, 本文提出一种代理模型方法与粒子群算法的协同优化策略.

已知候选解集 \mathbf{X}_{pss} , 根据式 (27) ~ (30) 对候选解进行预筛选, 则最优的可行解描述如下:

$$x_{\text{pss}} = \operatorname{argmin}[f(\mathbf{X}_{\text{pss}})]. \quad (31)$$

计算可行解 x_{pss} 对应的真实函数值为 $f(x_{\text{pss}})$, 然后更新主种群的学习样本, 则主种群的全局最优个体位置选择如下:

$$\text{gbest}(t) = \begin{cases} \text{gbest}(t), & f(\text{gbest}(t)) \geq f(x_{\text{pss}}); \\ x_{\text{pss}}, & f(\text{gbest}(t)) < f(x_{\text{pss}}). \end{cases} \quad (32)$$

此外, 将 $(x_{\text{pss}}, f(x_{\text{pss}}))$ 用于更新经验知识库, 有

$$D_t = [D_t; x_{\text{pss}}, f(x_{\text{pss}})], \quad (33)$$

知识库 D_t 中可行解的个数为 $N_t^{\text{fs}} = N_t^{\text{fs}} + 1$.

综上所述, 在迭代时刻 t , 基于多代理模型的优化方法得到的最优解为

$$x_{\text{best}}(t) = \text{gbest}(t). \quad (34)$$

算法2 基于多代理模型的优化方法.

Step 1: 随机初始化 CPSO 的速度和位置; 初始化种群规模 NP 以及参数 χ, c_1, c_2, k, η ; 初始化最大可行解评价次数 $N_{\text{max}}^{\text{fs}}$; 每个子空间的样本数 m_d ; 经验知识库 D_0 ; 令 $t = 1$.

Step 2: 执行算法 1, 得到候选解集 \mathbf{X}_{pss} .

Step 3: 根据式 (13) 更新经验知识库 D_t ; 计算可行解的个数 N_t^{fs} .

Step 4: 根据式 (14) ~ (17) 对输入空间分区.

Step 5: 根据式 (22) ~ (26) 构造多代理模型.

Step 6: 根据式 (31) 对候选解集 \mathbf{X}_{pss} 进行预筛选.

Step 7: 根据式 (32) 更新 CPSO 学习样本; 根据式 (33) 更新经验知识库 D_t , 计算可行解的个数.

Step 8: 根据式 (34) 计算基于多代理模型的优化方法得到的最优解 $x_{\text{best}}(t)$.

Step 9: 令 $t = t + 1$. 如果 $N_t^{\text{fs}} \leq N_{\text{max}}^{\text{fs}}$, 则返回 Step 2; 否则算法终止.

3 实验分析

3.1 标准测试问题

为了验证本文方法的有效性, 首先将本文方法 (称为 MSMIO) 与求解混合整数规划问题的代理模型优化算法 (SO-MI)^[3] 进行实验比较. 其中 SO-MI 采用随机采样的方法产生候选解, 并通过构造全局代理模型对候选解进行预筛选. 此外, 考虑到本文将 PSO 与 MSMIO 相结合, 因此, 将本文方法与混合离散粒子群算法 (MDPSO)^[7] 进行实验比较.

本文采用 14 个标准测试问题, 其中 1 ~ 6 为含有盒约束的测试问题, 7 ~ 14 为含有非线性约束的测试问题. 问题 1 ~ 6 和 8 的数学描述参见文献 [12]; 问题 7、9 和 14 的数学描述参见文献 [13]; 问题 10 ~ 13 的数学描述参见文献 [14]. 所有的测试问题如表 1 所示. 其中在问题属性描述中, BC (Box constrain) 表示盒约束, NLC (Nonlinear constrain) 表示非线性约束, LC (Linear constrain) 表示线性约束, UM (Unimodal) 和 MM (Multimodal) 分别表示问题具有单模态和多模态特点.

为了分析 MSMIO、SO-MI 和 MDPSO 三种方法的性能, 本文采用如下的性能评价指标^[7]:

$$\text{Score} = \frac{1}{K} \sum_{\pi \in K} \left| \frac{f_{a_i}^E - f_{\text{best}}^E}{f_{\text{best}}^E} \right| \times 100. \quad (35)$$

其中: K 是不同类型的问题集; π 是问题分类索引; f_a^E

表1 测试问题

问题	维数	搜索空间	离散变量个数	问题属性
1	5	$[-100, 100]^5$	2	BC, MM
2	10	$[3, 9]^{10}$	5	BC, UM
3	10	$[3, 99]^{10}$	5	BC, UM
4	12	$[-1, 3]^{12}$	5	BC, MM
5	12	$[-10, 30]^{12}$	5	BC, MM
6	30	$[-1, 3]^{30}$	10	BC, MM
7	11	$[0, 1]^8 \times [0, 0.997] \times [0, 0.9985] \times [0, 0.9988]$	4	3NLC, 4LC, MM
8	5	$[0, 10]^3 \times [0, 1]^2$	1	2NLC, 3LC, UM
9	3	$[0, 1] \times [0.2, 1] \times [-2.22554, -1]$	2	1NLC, 2LC, MM
10	25	$[0, 10]^{25}$	6	1NLC, 1LC, MM
11	5	$[78, 102] \times [33, 45] \times [27, 45]^3$	2	6NLC, UM
12	2	$[13, 100] \times [0, 100]$	1	2NLC, MM
13	7	$[-10, 10]^7$	3	4NLC, MM
14	11	$[0, 1]^8 \times [0, 10]^3$	4	4NLC, 9LC, MM

表示针对 π 类问题, 在 E (E 为 100, 200 或者 300) 个可行解评价之后, 算法 a_i ($a_{1:3}$ 分别为 SO-MI、MDPSO 和 MSMIO) 获得的最优解时, 对应的目标函数值 (30 次独立实验的平均值); $f_{\text{best}}^E = \min_{a_{1:3}}(f_a^E)$. 最佳算法的评分为 0, 较差算法的评分为正整数. 因此, 评分越小, 算法的性能越好.

本文方法的参数设置如下: CPSO、LPSO 和 GPSO 的种群规模 NP 为 30; c_1 和 c_2 分别为 2.05; 最大速度为 $(x_{\max} - x_{\min})/2$, 其中 x_{\max} 和 x_{\min} 分别为粒子每一维的最大和最小位置; CPSO 的初始种群随机产生; 每一优化周期, 采样频率 η 为 10.

为了实验的公平性, SO-MI、MDPSO 和 MSMIO 三种方法采用相同的初始设计点. 针对 14 个测试问题, 分别进行 3 组实验, 3 组实验的函数评价次数分别为 100、200 和 300, 且每组实验独立进行 30 次. 统计结果如表 2 和表 3 所示, 其中 Mean 和 SEM 分别表示最优解对应函数值的平均值及其对应的标准差.

针对含有盒约束的测试问题, SO-MI、MDPSO 和 MSMIO 三种方法的对比实验的统计结果如表 2 所示. 由表 2 中性能评价指标的结果可知: 在 100 次函数评价后, SO-MI 的收敛性能好于 MSMIO; 而在 200 和 300 次函数评价后, MSMIO 的收敛性能明显好于 SO-MI. 从表 2 中对 SEM 的统计结果可知, SO-MI 方法得到的 SEM 最小, MDPSO 方法得到的 SEM 最大, 表明 SO-MI 的稳定性最好, MDPSO 的稳定性最差, 而 MSMIO 的稳定性差于 SO-MI, 却优于 MDPSO.

针对含有线性和非线性约束的测试问题, SO-MI、MDPSO 和 MSMIO 三种方法的对比实验的统计结果如表 3 所示. 由表 3 中性能评价指标的结果可知, MSMIO 的收敛性能最优. 在 100 和 200 次函数评价后, MDPSO 的收敛性能好于 SO-MI, 而在 300 次函

数评价后, SO-MI 的收敛性能好于 MDPSO. 从表 3 中对 SEM 的统计结果可知, 对于问题 11 和 14, SO-MI 方法得到的 SEM 最小; 而对于问题 7~10、12 和 13, SO-MI、MDPSO 和 MSMIO 三种方法得到的 SEM 较小且非常接近, 表明 3 种算法具有较好的稳定性, 但是 SO-MI 要优于其他两种方法.

3.2 基于 LSSVM 的非线性时间序列回归模型优化

本文以某钢铁企业副产煤气系统为背景, 构造基于最小二乘支持向量机 (LSSVM) 的非线性时间序列回归模型^[15]. 为了得到精确的 LSSVM 模型, 采用所提出方法对样本构造和超参数选取进行联合优化.

在样本构造过程中, 延迟时间 τ 和嵌入维数 d 直接影响相空间重构的质量. 此外, 超参数 (核函数参数 γ 和惩罚系数 δ) 影响 LSSVM 模型的泛化能力. 为了获得精确的 LSSVM 模型, 采用 10 折交叉验证^[16] 方法, 则得到如下的混合整数规划问题:

$$\begin{aligned} \min f(\theta) &= \frac{1}{10} \sum_{i=1}^{10} L(A_{\theta}, D_{\text{train}}^i, D_{\text{valid}}^i); \\ \text{s.t. } \theta_{lb} &\leq \theta \leq \theta_{ub}. \end{aligned} \quad (36)$$

其中: A_{θ} 表示 LSSVM 回归模型; $\theta = [\tau \ d \ \gamma \ \delta]$ 为待优化参数, $\tau, d \in \mathbf{N}$, $\gamma, \delta > 0$; $L(A_{\theta}, D_{\text{train}}^i, D_{\text{valid}}^i)$ 表示根据训练集 D_{train} 和测试集 D_{valid} 得到的 i 折交叉验证误差; θ_{lb} 和 θ_{ub} 分别为 θ 的上下边界.

本文选取某钢铁企业煤气系统转炉煤气消耗流量数据序列, 如图 2 所示. LSSVM 模型的训练样本采用 24 小时煤气消耗流量数据, 样本个数为 2 880. 为了验证本文方法在 LSSVM 模型优化问题中的有效性, 分别采用网格搜索方法、SO-MI、MDPSO 和 MSMIO 四种算法进行实验比较. 其中, 网格搜索的结果作为参考值, 当其他 3 种算法搜索到与网格搜索方法相近的结果时, 算法终止.

表2 针对具有盒约束测试问题的优化结果统计

问题	算法	统计	100次评价	200次评价	300次评价
1	SO-MI	Mean	-386.33	-420.91	-432.58
		SEM	16.74	14.87	15.17
	MDPSO	Mean	-346.50	-415.98	-449.57
		SEM	82.86	75.02	75.41
	MSMIO	Mean	-375.34	-431.67	-461.32
		SEM	89.59	90.56	81.98
2	SO-MI	Mean	-42.92	-42.99	-42.99
		SEM	0.19	0.13	0.13
	MDPSO	Mean	-20.85	-33.50	-39.71
		SEM	3.29	3.37	2.45
	MSMIO	Mean	-43.13	-43.13	-43.13
		SEM	0.00	0.00	0.00
3	SO-MI	Mean	-9581.32	-9584.62	-9584.62
		SEM	5.26	0.79	0.79
	MDPSO	Mean	-5 701.66	-7 547.20	-8 578.57
		SEM	873.59	717.12	515.46
	MSMIO	Mean	-9 576.26	-9 591.72	-9 591.72
		SEM	84.67	0.00	0.00
4	SO-MI	Mean	-4.89	-8.48	-9.63
		SEM	0.33	0.30	0.23
	MDPSO	Mean	-3.46	-6.30	-8.24
		SEM	1.95	1.69	1.38
	MSMIO	Mean	-6.28	-8.47	-10.41
		SEM	1.20	1.00	1.02
5	SO-MI	Mean	27.95	-5.78	-8.27
		SEM	2.85	0.27	0.34
	MDPSO	Mean	335.07	118.77	43.09
		SEM	141.48	59.87	32.75
	MSMIO	Mean	29.22	-11.57	-11.61
		SEM	41.11	0.21	0.21
6	SO-MI	Mean	-6.59	-10.29	-11.80
		SEM	0.37	0.36	0.47
	MDPSO	Mean	9.22	1.23	-4.53
		SEM	5.30	4.46	4.02
	MSMIO	Mean	-2.59	-10.26	-14.23
		SEM	4.60	2.31	1.96
Score in % MDPSO			4	9	10
MSMIO			239	185	63
MSMIO			10	0	0

已知参数边界为 $\tau \in [1, 10], d \in [1, 50], \gamma \in [1, 1000], \delta \in [0.1, 100]$. 仿真实验基于Matlab 2015 b和酷睿双核(i5 6200). 在网格搜索过程中,对连续变量 γ 和 θ 的搜索域进行网格划分,即等分为10份,则产生 $10 \times 50 \times 10 \times 10$ 种参数配置方案. 对于2#高炉的煤气消耗流量数据而言,单次参数配置方案的计算时长为2.4s,则总耗时约33h. 其中,最佳的参数配置方案为 $\tau = 6, d = 50, \gamma = 700, \delta = 10$,其最小误差为1.71. 因此,SO-MI、MDPSO和MSMIO三种算法的终止条件为 $f \leq 1.71$,其参数设置与3.1节相同.

表3 针对具有非线性约束测试问题的优化结果统计

问题	算法	统计	100次评价	200次评价	300次评价
7	SO-MI	Mean	-0.50	-0.66	-0.72
		SEM	0.03	0.02	0.02
	MDPSO	Mean	-0.74	-0.83	-0.87
		SEM	0.06	0.05	0.05
	MSMIO	Mean	-0.77	-0.83	-0.85
		SEM	0.05	0.04	0.04
8	SO-MI	Mean	0.44	0.09	0.04
		SEM	0.06	0.02	0.01
	MDPSO	Mean	0.23	0.09	0.04
		SEM	0.26	0.15	0.10
	MSMIO	Mean	0.12	0.02	0.02
		SEM	0.19	0.09	0.08
9	SO-MI	Mean	2.99	2.95	2.93
		SEM	0.02	0.02	0.01
	MDPSO	Mean	2.89	2.87	2.86
		SEM	0.03	0.02	0.01
	MSMIO	Mean	2.88	2.86	2.85
		SEM	0.02	0.01	0.00
10	SO-MI	Mean	-0.18	-0.25	-0.32
		SEM	0.00	0.01	0.01
	MDPSO	Mean	-0.16	-0.17	-0.18
		SEM	0.02	0.02	0.02
	MSMIO	Mean	-0.17	-0.20	-0.22
		SEM	0.02	0.02	0.02
11	SO-MI	Mean	-29767.98	-29983.69	-30078.49
		SEM	67.99	53.92	51.23
	MDPSO	Mean	-29897.23	-30030.57	-30080.86
		SEM	254.55	256.64	251.89
	MSMIO	Mean	-30 157.97	-30 443.39	-30 506.23
		SEM	183.48	117.89	106.19
12	SO-MI	Mean	-4 156.44	-4 182.99	-4 186.56
		SEM	14.39	9.37	8.12
	MDPSO	Mean	-4 241.05	-4 241.98	-4 242.01
		SEM	1.06	0.05	0.00
	MSMIO	Mean	-4 241.37	-4 241.84	-4 241.95
		SEM	0.96	0.23	0.08
13	SO-MI	Mean	1 057.39	847.65	761.448
		SEM	17.81	15.68	8.63
	MDPSO	Mean	855.65	766.09	749.63
		SEM	78.94	49.43	46.80
	MSMIO	Mean	858.93	723.56	697.04
		SEM	144.55	30.66	15.25
14	SO-MI	Mean	8.05	6.60	6.20
		SEM	0.22	0.10	0.06
	MDPSO	Mean	6.72	6.45	6.32
		SEM	0.48	0.40	0.42
	MSMIO	Mean	6.19	6.04	5.95
		SEM	0.13	0.11	0.11
Score in % MDPSO			45	50	17
MSMIO			14	48	18
MSMIO			1	2	4

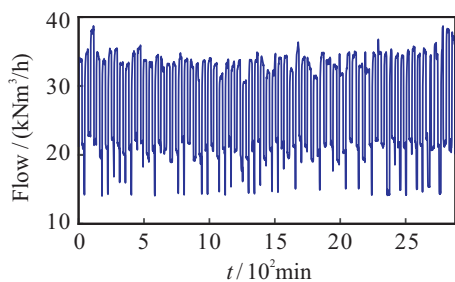


图2 2#高炉煤气消耗使用流量

针对转炉煤气消耗模型的优化问题,SO-MI、MDPSO和MSMIO三种方法的对比实验的统计结果如表4所示.其中:“最大”、“最小”、“平均”、“标准差”和“耗时”分别表示在满足终止条件下,30次独立实验后,参数配置的最大评价次数、参数配置的最小评价次数、参数配置的平均评价次数、参数配置的评价次数的标准差和参数配置的平均耗时(单位min).

表4 LSSVM模型优化结果统计

算法	最大	最小	平均	标准差	耗时/min
SO-MI	738	586	621	103	25
MDPSO	701	498	572	127	23
MSMIO	574	402	495	74	20

由表4可知:与网格搜索方法相比,SO-MI、MDPSO和MSMIO三种算法减少了参数配置的评价次数,并大大减少了模型优化计算成本;与SO-MI和MDPSO相比,在相同的终止条件下,MSMIO所需参数配置的评价次数最少,耗时也最少.由表4中对参数配置的评价次数的标准差的统计结果可知:MSMIO方法得到的标准差最小,表明MSMIO的稳定性最好;MDPSO所需参数配置的评价次数少于SO-MI,但是稳定性比SO-MI差.

4 结论

为了在有限的函数评价次数下求解计算成本高的混合整数规划问题,本文提出了一种基于多代理模型的优化方法.该方法通过基于多代理模型的预选择策略和粒子群优化算法的协同优化,减少了建模成本,提高了优化效率.分别针对14个标准测试函数和一个模型优化问题进行测试,实验结果表明本文方法优于SO-MI和MDPSO.

参考文献(References)

[1] Cetinkaya E. An adaptive multiquadric radial basis function method for expensive black-box mixed-integer nonlinear constrained optimization[J]. *Engineering Optimization*, 2013, 45(2): 185-206.

[2] Holmström K, Quttineh N H, Edvall M M. An adaptive radial basis algorithm(ARBF) for expensive black-box mixed-integer constrained global

optimization[J]. *Optimization & Engineering*, 2008, 9(4): 311-339.

[3] Müller J, Shoemaker C A, Piché R. SO-MI: A surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems[J]. *Computers & Operations Research*, 2013, 40(5): 1383-1400.

[4] Haftka R T, Villanueva D, Chaudhuri A. Parallel surrogate-assisted global optimization with expensive functions — A survey[J]. *Structural & Multidisciplinary Optimization*, 2016, 54(1): 3-13.

[5] Zhou Z, Ong Y S, Nair P B, et al. Combining global and local surrogate models to accelerate evolutionary optimization[J]. *IEEE Trans on Systems, Man, and Cybernetics, Part C*, 2007, 37(1): 66-76.

[6] Liem R P, Mader C A, Martins J R A. Surrogate models and mixtures of experts in aerodynamic performance prediction for aircraft mission analysis[J]. *Aerospace Science and Technology*, 2015, 43(8): 126-151.

[7] Chowdhury S, Tong W, Messac A, et al. A mixed-discrete particle swarm optimization algorithm with explicit diversity-preservation[J]. *Structural & Multidisciplinary Optimization*, 2013, 47(3): 367-388.

[8] Tang Y F, Chen J Q, Wei J H. A surrogate-based particle swarm optimization algorithm for solving optimization problems with expensive black box functions[J]. *Engineering Optimization*, 2013, 45(5): 557-576.

[9] Clerc M, Kennedy J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space[J]. *IEEE Trans on Evolutionary Computation*, 2002, 6(1): 58-73.

[10] Choudhury A, Nair P B, Keane A J. A data parallel approach for large-scale gaussian process modeling[C]. *Siam Int Conf on Data Mining*. Arlington, 2002: 95-111.

[11] Petelin D. Gaussian processes for machine learning[J]. *Int J of Neural Systems*, 2006, 14(6): 3011-3015.

[12] Bussieck M R, Drud A S, Meeraus A. MINLPLib — A collection of test models for mixed-integer nonlinear programming[J]. *Inform J on Computing*, 2003, 15(1): 114-119.

[13] Adjiman C S, Androulakis I P, Floudas C A. Global optimization of mixedinteger nonlinear problems[J]. *AIChE Journal*, 2000, 46(9): 1769-1797.

[14] Mezura-Montes E, Coello C A C. A simple multimembered evolution strategy to solve constrained optimization problems[J]. *IEEE Trans on Evolutionary Computation*, 2005, 9(1): 1-17.

[15] Zhao J, Wang W, Pedrycz W, et al. Online parameter optimization-based prediction for converter gas system by parallel strategies[J]. *IEEE Trans on Control Systems Technology*, 2012, 20(3): 835-845.

[16] Bergstra J, Bengio Y. Algorithms for hyper-parameter optimization[C]. *Int Conf on Neural Information Processing Systems*. Granda, 2011: 2546-2554.