

求解昂贵区间多目标优化问题的高斯代理模型

陈志旺^{1,2}, 白 铨^{1†}, 杨 七¹, 黄兴旺¹, 李国强¹

(1. 燕山大学 工业计算机控制工程河北省重点实验室, 河北 秦皇岛 066004;

2. 燕山大学 国家冷轧板带装备及工艺工程技术研究中心, 河北 秦皇岛 066004)

摘要: 本文将数据挖掘(高斯过程回归建模)和智能进化算法(GA, NSGA-II)进行结合, 用于解决优化函数未知的昂贵区间多目标优化问题. 首先利用高斯过程对采用中点和不确定度表示的未知目标函数和约束函数进行建模, 由于相关性和准确性是区间函数模型的两个必备条件, 故提出一种融合多属性决策的双层种群筛选策略, 并将其嵌入到遗传算法求解高斯模型参数的过程中, 第1层根据相关性属性排除候选解集中部分劣解, 第2层根据准确性属性排除候选解集中其余超出种群规模的劣解, 两属性的权重系数决定两层排除劣解的比例. 然后将所建模型作为优化对象的代理模型引导区间NSGA-II算法优化求解, 从而获得所需的Pareto前沿.

关键词: 多目标优化; 区间规划; 第2代非支配排序进化算法(NSGA-II); 高斯过程; 多属性决策; 代理模型

中图分类号: TP301 文献标识码: A

Gaussian surrogate models for expensive interval multi-objective optimization problem

CHEN Zhi-wang^{1,2}, BAI Xin^{1†}, YANG Qi¹, HUANG Xing-wang¹, LI Guo-qiang¹

(1. Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao Hebei 066004, China;

2. National Engineering Research Center for Equipment and Technology of Cold Strip Rolling, Yanshan University, Qinhuangdao Hebei 066004, China)

Abstract: In this paper data mining (Gaussian process regression modeling) and intelligent evolutionary algorithm (GA, NSGA-II) are combined to solve the expensive interval multi-objective optimization problem with unknown optimization functions. Firstly, Gaussian process (GP) is used to model the objective functions and constraint functions represented by the midpoint and uncertainty. Because relevance and accuracy are two essential factors of interval function models, A kind of double steps screening strategy based on multiple attribute decision making (MADM) is proposed and it is embedded into the genetic algorithm to identify the parameters of the GP model. In the first step, inferior solutions in candidate solutions are excluded according to relevance. In the second step, the rest of inferior solutions beyond population quantity are excluded according to accuracy. And the proportion of inferior solutions excluded in the two steps is decided by the weight coefficient of two factors. Then, the built GP models for optimization objects are used as surrogate models in the NSGA-II optimization algorithm, so that Pareto front can be found.

Key words: multi-objective optimization; interval programming; non-dominated sorting genetical algorithm II (NSGA-II); Gaussian process; multiple attribute decision making; surrogate model

1 引言(Introduction)

随着学者们对多目标优化问题的研究, 昂贵区间多目标优化问题成为当今多目标优化的研究热点. 该问题不仅包含了区间多目标优化和昂贵多目标优化的难点, 而且由于两问题嵌合, 使得求解时需互相兼顾, 从而加深了问题的优化难度. 其中区间多目标优化指的是优化问题中目标函数和约束函数含有不确

定区间变量^[1]; 而昂贵多目标优化指的是优化问题精确数学模型难以获得且与优化对象进行评估实验的成本高昂, 大量的评估实验必然会造成成本灾难^[2-3].

对于昂贵多目标优化问题, 近年来学者们将数据挖掘的思想引入到智能进化领域^[4]. 其主要思路是利用回归算法对优化对象的经验数据(经验数据由解及其所对应的真实目标向量组成)进行学习训练而辨识

收稿日期: 2015-05-13; 录用日期: 2016-09-05.

†通信作者. E-mail: 15233013272@163.com; Tel: +86 15233013272.

本文责任编辑: 刘丁.

国家自然科学基金(61403331, 61573305), 河北省自然科学基金青年基金(F2014203099), 燕山大学青年教师自主研究计划课题(13LGA006)资助. Supported by National Natural Science Foundation of China (61403331, 61573305), Natural Science Foundation for Young Scientist of Hebei Province, China (F2014203099) and Independent Research Program for Young Teachers of Yanshan University, China (13LGA006).

出原始目标函数的近似模型,并将此模型作为代理模型运用于算法优化.所谓代理模型是指通过学习解空间上的已知区域,建立相关逼近模型来预测未知区域的解分布情况,并进一步指导算法的搜索方向,引导算法绕过局部极值而向全局最优值方向搜索.目前基于代理模型的多目标优化进化算法备受学术界关注^[5],其中文献[6]作者为解决模型训练的维数灾难问题,对局部高斯模型进行了改进,并将其应用于求解昂贵多目标优化问题.文献[7]作者提出了基于双层代理模型的粒子群优化算法,第1层全局模型用于全局搜索,避免算法局部收敛,第2层局部模型用于指导算法快速收敛.文献[8]作者先将多目标优化问题转换成多个单目标优化子问题,利用模糊聚类和高斯过程对每个单目标优化子问题建立预测模型,基于预测模型利用MOEA/D进行求解.

以上基于代理模型的多目标优化研究中,优化问题多为确定性的,而对考虑了不确定性优化函数的建模研究少.其中文献[9]作者采用半监督协同学习机制训练两个径向基函数神经网络,构造精度高泛化能力强的区间上、下界模型.但区间函数的代理模型构建需考虑两个因素:1)模型预测的准确性(误差越小越优);2)预测下界大于预测上界出现的概率(概率越小越优)即相关性.该方法只是以提高准确性为目标,上下界的建模缺少相关性.为此文献[10]中引入多输出高斯模型,虽考虑上下界相关性,但由文献[11]可知该模型相比单输出高斯模型待优化超参数过多,造成建模过程具有高复杂度.文献[12]提出一种区间ARX(IARX)建模法,并利用其建立运动模式的区间模型,该建模方法将区间函数表示为区间中点和不确定度,然后以预测准确性为目标函数,模型相关性为约束条件进行建模,但回归建模过程中没有考虑准确性和相关性的权重问题.

对于代理模型建模算法的选取,由文献[13]可知高斯过程相比神经网络和支持向量机等具有容易实现、超参数自适应获取以及输出具有概率意义等优点.而高斯模型的构建常通过共轭梯度法对模型超参数进行求解,并通过最小二乘法获得其余未知参数.然而共轭梯度法从单个初始值出发,依梯度下降迭代求最优解,容易陷入局部最优,并且其优化效果及迭代的收敛性过于依赖初值.比较而言,遗传算法是一种公认的具有较好的全局搜索能力的算法,从问题的初始解集开始并行搜索,覆盖面大,具有较大概率找到全局最优解;并且其优化结果对初值依赖性小^[14].文献[14–15]作者为提高模型预测精度,均利用遗传算法替代共轭梯度法实现高斯模型的构建,主要原因是遗传算法求解过程与优化问题形式无关,而且还可以根据具体问题将其它求解策略纳入遗传操作框架内,建立更为有效的求解方法^[16].

综上所述,为了获得具有低建模复杂度且兼顾准确性和相关性的区间模型,本文采用单输出高斯模型,该模型参数少、建模简单^[17],然后提出了一种区间函数高斯建模法,利用嵌入双层筛选策略的遗传算法求解.通过该建模方法可以构建兼顾准确性和相关性的高斯模型且两者权重可调.最后将所建模型作为优化对象的代理模型引导区间NSGA-II算法优化.从而利用数据挖掘和智能算法的双向结合实现对昂贵区间多目标优化问题的求解.

2 问题描述(Problem statement)

区间多目标优化问题的一般表达式为

$$Q: \min_{\mathbf{x}} \mathbf{F}(\mathbf{x}, \mathbf{u}) = (f_m(\mathbf{x}, \mathbf{u})), m = 1, 2, \dots, z,$$

$$\text{s.t.} \begin{cases} g_j(\mathbf{x}, \mathbf{u}) \geq a_j = [\underline{a}_j, \bar{a}_j], j = 1, 2, \dots, n, \\ h_k(\mathbf{x}, \mathbf{u}) = b_k = [\underline{b}_k, \bar{b}_k], k = 1, 2, \dots, n^*, \\ \mathbf{x} = (x_1, \dots, x_q) \in \mathbb{R}^q, x_t \in [\underline{x}'_t, \bar{x}'_t], \\ t = 1, 2, \dots, q, \\ \mathbf{u} = (u_1, \dots, u_p) \in \mathbb{R}^p, u_l \in [\underline{u}'_l, \bar{u}'_l], \\ l = 1, 2, \dots, p, \end{cases} \quad (1)$$

其中: $\mathbf{F}(\mathbf{x}, \mathbf{u})$ 为目标函数向量, $f_m(\mathbf{x}, \mathbf{u})$ 为第 m 个区间目标函数. $\mathbf{x} = (x_1, \dots, x_q)$ 为 q 维解, x_t 是 \mathbf{x} 的第 t 个变量, \bar{x}'_t 和 \underline{x}'_t 为变量 x_t 的上下边界. $\mathbf{u} = (u_1, \dots, u_p)$ 为函数 $\mathbf{F}(\mathbf{x}, \mathbf{u})$ 的 p 维内部扰动向量, u_l 是扰动向量 \mathbf{u} 的第 l 个变量, \bar{u}'_l 和 \underline{u}'_l 为变量 u_l 的上下边界. $g_j(\mathbf{x}, \mathbf{u})$ 是第 j 个区间不等式约束, $h_k(\mathbf{x}, \mathbf{u})$ 是第 k 个区间等式约束.

若式(1)中的区间函数 $\mathbf{F}(\mathbf{x}, \mathbf{u})$, $g_j(\mathbf{x}, \mathbf{u})$ 及 $h_k(\mathbf{x}, \mathbf{u})$ 的内部结构是未知的,且评估实验成本高昂,则该优化问题可归为昂贵多目标优化问题.文献[18]提出的区间NSGA-II算法只能求解优化函数已知的优化问题.对于昂贵区间多目标优化,该算法并不适用.因为在优化函数未知的情况下,该区间NSGA-II算法为了获得Pareto前沿需要对优化对象进行大量的评估实验,若评估成本高昂必然会造成成本灾难.而代理模型可以将输入、输出变量与优化目标、约束之间的未知关系通过显式函数近似表达.故可通过构建优化对象的代理模型,引导NSGA-II算法优化.

3 区间NSGA-II相关定义(Related definition of interval NSGA-II)

令 $X' = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\text{pop}}\}$ 为NSGA-II算法中的种群解集, pop 为种群规模. \mathbf{x}_i 和 \mathbf{x}_j ($i, j = 1, 2, \dots, \text{pop}$)为种群 X' 中的两个解.为比较解 \mathbf{x}_i 和 \mathbf{x}_j 的优劣性, \mathbf{x}_i 和 \mathbf{x}_j 第 m 维上的区间目标函数可转换为

$$f_m(\mathbf{x}_i, \mathbf{u}) = [f_m(\mathbf{x}_i, \mathbf{u}), \overline{f_m(\mathbf{x}_i, \mathbf{u})}],$$

$$f_m(\mathbf{x}_j, \mathbf{u}) = [\underline{f_m(\mathbf{x}_j, \mathbf{u})}, \overline{f_m(\mathbf{x}_j, \mathbf{u})}].$$

定义 1^[18] 对于一个区间数 $a = [\underline{a}, \bar{a}]$, 设 $a^c = \frac{\bar{a} + \underline{a}}{2}$, $a^r = \frac{\bar{a} - \underline{a}}{2}$, $a^w = \bar{a} - \underline{a}$, 则称 a^c, a^r, a^w 分别为区间数 a 的中点、半径和不确定度. 当 $\bar{a} > \underline{a}$ 时, $a^w > 0$. 称对 \underline{a} 和 \bar{a} 的要求 “ $\bar{a} > \underline{a}$ ” 为区间上下界的相关性, 简称相关性.

定义 2^[18] 令 $P(f_m(\mathbf{x}_i, \mathbf{u}) \leq f_m(\mathbf{x}_j, \mathbf{u}))$ 为区间数 $f_m(\mathbf{x}_i, \mathbf{u})$ 小于等于区间数 $f_m(\mathbf{x}_j, \mathbf{u})$ 区间的区间可能度. 根据文献[19]中区间6种位置关系可得区间可能度的计算公式如式(2).

$$P(f_m(\mathbf{x}_i, \mathbf{u}) \leq f_m(\mathbf{x}_j, \mathbf{u})) = \begin{cases} 0, & \overline{f_m(\mathbf{x}_j, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_i, \mathbf{u})}, \\ \frac{(\overline{f_m(\mathbf{x}_j, \mathbf{u})} - \underline{f_m(\mathbf{x}_i, \mathbf{u})})^2}{8 \cdot f_{rm}(\mathbf{x}_i, \mathbf{u}) \cdot f_{rm}(\mathbf{x}_j, \mathbf{u})}, & \underline{f_m(\mathbf{x}_j, \mathbf{u})} < \underline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})} < \overline{f_m(\mathbf{x}_i, \mathbf{u})}, \\ \frac{f_{cm}(\mathbf{x}_j, \mathbf{u}) - \underline{f_m(\mathbf{x}_i, \mathbf{u})}}{2 \cdot f_{rm}(\mathbf{x}_i, \mathbf{u})}, & \underline{f_m(\mathbf{x}_j, \mathbf{u})} < \underline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})} < \overline{f_m(\mathbf{x}_i, \mathbf{u})}, \\ \frac{f_{cm}(\mathbf{x}_i, \mathbf{u}) - \underline{f_m(\mathbf{x}_j, \mathbf{u})}}{2 \cdot f_{rm}(\mathbf{x}_j, \mathbf{u})}, & \underline{f_m(\mathbf{x}_j, \mathbf{u})} < \underline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})} < \overline{f_m(\mathbf{x}_i, \mathbf{u})}, \\ 1 - \frac{(\overline{f_m(\mathbf{x}_i, \mathbf{u})} - \underline{f_m(\mathbf{x}_j, \mathbf{u})})^2}{8 \cdot f_{rm}(\mathbf{x}_i, \mathbf{u}) \cdot f_{rm}(\mathbf{x}_j, \mathbf{u})}, & \underline{f_m(\mathbf{x}_i, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_j, \mathbf{u})} < \overline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})}, \\ \frac{\overline{f_m(\mathbf{x}_j, \mathbf{u})} - f_{cm}(\mathbf{x}_i, \mathbf{u})}{2 \cdot f_{rm}(\mathbf{x}_j, \mathbf{u})}, & \underline{f_m(\mathbf{x}_j, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})} < \overline{f_m(\mathbf{x}_i, \mathbf{u})}, \\ \frac{\overline{f_m(\mathbf{x}_j, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_i, \mathbf{u})} < \overline{f_m(\mathbf{x}_j, \mathbf{u})}, & \underline{f_m(\mathbf{x}_i, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_j, \mathbf{u})}, \\ 1, & \overline{f_m(\mathbf{x}_i, \mathbf{u})} \leq \underline{f_m(\mathbf{x}_j, \mathbf{u})}. \end{cases} \quad (2)$$

定义 3^[18] 两区间数的区间距离公式如下:

$$d(f_m(\mathbf{x}_i, \mathbf{u}), f_m(\mathbf{x}_j, \mathbf{u})) = \sqrt{D^2 + \frac{1}{3}F - \frac{2}{3}(T^r)^2}, \quad (3)$$

其中: $D = f_{cm}(\mathbf{x}_i, \mathbf{u}) - f_{cm}(\mathbf{x}_j, \mathbf{u})$, $F = f_{rm}^2(\mathbf{x}_i, \mathbf{u}) + f_{rm}^2(\mathbf{x}_j, \mathbf{u})$, $T = (f_m(\mathbf{x}_i, \mathbf{u}) \cap f_m(\mathbf{x}_j, \mathbf{u}))$ 为区间 $f_m(\mathbf{x}_i, \mathbf{u})$ 和 $f_m(\mathbf{x}_j, \mathbf{u})$ 的交集, T^r 表示 T 的半径.

定义 4^[18] 设 \mathbf{x}_i 和 \mathbf{x}_j 为式(1)的两个解, 如果所有 $f_m(\mathbf{x}_i, \mathbf{u})$ 小于等于 $f_m(\mathbf{x}_j, \mathbf{u})$ 的区间可能度 $P(f_m(\mathbf{x}_i, \mathbf{u}) \leq f_m(\mathbf{x}_j, \mathbf{u}))$ (式(2), 以下简称为 P) 均不小于 0.5, 且存在 $f_m(\mathbf{x}_i, \mathbf{u})$ 小于等于 $f_m(\mathbf{x}_j, \mathbf{u})$ 区间可能度 P 大于 0.5, 则称 \mathbf{x}_i 以 P 支配 \mathbf{x}_j , 记为 $\mathbf{x}_i \succ_P \mathbf{x}_j$, 也可称 \mathbf{x}_j 以 P 被 \mathbf{x}_i 支配, 记为 $\mathbf{x}_j \prec_P \mathbf{x}_i$, 即

$$\mathbf{x}_i \succ_P \mathbf{x}_j \Leftrightarrow$$

$$\begin{aligned} \forall i \in (1, 2, \dots, z), \exists: P(f_m(\mathbf{x}_i, \mathbf{u}) \leq f_m(\mathbf{x}_j, \mathbf{u})) \geq 0.5, \\ \exists i \in (1, 2, \dots, z), \exists: P(f_m(\mathbf{x}_i, \mathbf{u}) \leq f_m(\mathbf{x}_j, \mathbf{u})) > 0.5. \end{aligned} \quad (4)$$

定义 5^[18] 目标函数为区间数的拥挤距离如下: 具有相同序值的种群个体 i 周围的个体 $i - 1$ 与 $i + 1$ 区间距离的 z 维欧几里得距离, 记做 SD_i :

$$SD_i = \sqrt{\sum_{m=1}^z d^2[f_m(\mathbf{x}_{i-1}, \mathbf{u}), f_m(\mathbf{x}_{i+1}, \mathbf{u})]}. \quad (5)$$

4 区间优化函数建模 (Interval optimization functions modeling)

4.1 建模数据集 (Modeling dataset)

对于区间函数的建模, 文献[20]通过建模算法分别辨识出区间函数的上、下界模型, 但是其预测出的上下界缺少相关性, 当区间值上下界接近时, 可能出现某解预测区间上界小于下界的情况, 称该类解为异常解. 此时若不及时纠正最终会影响建模的准确性. 对于该类问题文献[21]令预测区间上下界相加除 2 取平均值, 将区间数变为确定数, 这无疑会增加模型的预测误差.

由异常解的概念可知, 当分别对区间上下界建模时, 很容易出现异常解, 即使在建模的策略设计中考虑异常解, 由于上下界分别建模, 策略的设计也会涉及到许多困难. 由定义 1 可知区间数除可用区间上下界表示外还可表示为中点和不确定度. 对不确定度进行建模, 通过降低其预测值小于 0 的概率, 从而可体现上下界的相关性, 这是区间上下界独立建模所不具备的. 故本文对未知函数区间中点和不确定度进行高斯建模. 由于未知约束函数与未知目标函数建模方法相同, 本文以未知目标函数为例进行论述. 在建模前首先需要获得用于建模的数据集. 对于数据集一般取经验数据. 若评估实验允许可通过评估实验获取. 由定义 1 可知式(1)中的第 m 个目标函数的表达式可转换为以下形式:

$$f_m(\mathbf{x}, \mathbf{u}) = [f_{cm}(\mathbf{x}, \mathbf{u}), f_{wm}(\mathbf{x}, \mathbf{u})], \\ m = 1, 2, \dots, z,$$

其中: $f_{cm}(\mathbf{x}, \mathbf{u})$ 为目标函数 $f_m(\mathbf{x}, \mathbf{u})$ 输出区间的中点, $f_{wm}(\mathbf{x}, \mathbf{u})$ 为目标函数 $f_m(\mathbf{x}, \mathbf{u})$ 输出区间的确定度. 为了辨识区间函数, 可分别辨识区间中点 $f_{cm}(\mathbf{x}, \mathbf{u})$ 的模型和区间不确定度 $f_{wm}(\mathbf{x}, \mathbf{u})$ 的模型, 利用得到的两个模型描述区间函数. 故将对应的建模数据集分成两部分即中点数据集和不确定度数据集, 如式(6):

$$\begin{cases} D_{cm} = \{(\mathbf{x}_i, f_{cm}(\mathbf{x}_i, \mathbf{u}))\}, \\ D_{wm} = \{(\mathbf{x}_i, f_{wm}(\mathbf{x}_i, \mathbf{u}))\}, \end{cases} \quad i = 1, 2, \dots, \zeta,$$

$$\begin{aligned} \mathbf{X} &= [\mathbf{x}_1 \cdots \mathbf{x}_\zeta], \\ \begin{cases} \mathbf{F}_{cm} &= [f_{cm}(\mathbf{x}_1, \mathbf{u}) \cdots f_{cm}(\mathbf{x}_\zeta, \mathbf{u})]^T, \\ \mathbf{F}_{wm} &= [f_{wm}(\mathbf{x}_1, \mathbf{u}) \cdots f_{wm}(\mathbf{x}_\zeta, \mathbf{u})]^T, \end{cases} \end{aligned} \quad (6)$$

其中: D_{cm} 为中点数据集, D_{wm} 为不确定度数据集, \mathbf{x}_i 表示数据中的第*i*个*q*维数据输入值, 所有 \mathbf{x}_i 构成输入矩阵 \mathbf{X} , \mathbf{F}_{cm} 为中点观测输出矩阵, \mathbf{F}_{wm} 为不确定度观测输出矩阵.

4.2 高斯建模(Gaussian modeling)

本文采用高斯过程构建优化对象的代理模型. 由于区间不确定度的建模相对区间中点建模要考虑异常解情况, 故本节主要对区间不确定度的建模进行的说明. 设第*m*个目标函数不确定度模型输出矩阵为

$$\begin{cases} \mathbf{y}_{wm} &= [y_{wm}(\mathbf{x}_1, \mathbf{u}) \cdots y_{wm}(\mathbf{x}_\zeta, \mathbf{u})]^T, \\ y_{wm}(\mathbf{x}_i, \mathbf{u}) &= f_{wm}(\mathbf{x}_i, \mathbf{u}) + \psi_{wm}, \quad i=1, 2, \cdots, \zeta, \end{cases} \quad (7)$$

$f_{wm}(\mathbf{x}_i, \mathbf{u})$ 为输入值是 \mathbf{x}_i 时, 第*m*个目标函数 $f_m(\mathbf{x}_i, \mathbf{u})$ 的输出区间不确定度, ψ_{wm} 为与 $f_{wm}(\mathbf{x}_i, \mathbf{u})$ 不相关的高斯白噪声, 可表示为 $\psi_{wm} \sim \text{GP}(0, \sigma_{wm}^2)$, σ_{wm}^2 为噪声方差, $y_{wm}(\mathbf{x}_i, \mathbf{u})$ 为 \mathbf{x}_i 所对应模型的输出值. 将区间不确定度 $f_{wm}(\mathbf{x}_i, \mathbf{u})$ 的先验概率分布 $p(f_{wm}(\mathbf{x}_i, \mathbf{u}))$ 看作高斯分布, 该函数所对应的已知有限数据集为 D_{wm} . 此数据集中可将 $f_{wm}(\mathbf{x}_1, \mathbf{u})$, $f_{wm}(\mathbf{x}_2, \mathbf{u})$, \cdots , $f_{wm}(\mathbf{x}_\zeta, \mathbf{u})$ 构成一个具有联合高斯分布的变量集合, 它们形成的随机过程称之为高斯过程. 该过程满足均值为

$$e_{wm}(\mathbf{x}_i) = \text{E}(f_{wm}(\mathbf{x}_i, \mathbf{u})). \quad (8)$$

由于区间上界必然大于或等于区间下界即区间的不确定度大于或等于0, 故要求 $e_{wm}(\mathbf{x}_i) \geq 0$. 协方差函数为

$$\begin{aligned} k_{wm}(\mathbf{x}_i, \mathbf{x}_j) &= \text{E}((f_{wm}(\mathbf{x}_i, \mathbf{u}) - e_{wm}(\mathbf{x}_i)) \cdot \\ &\quad (f_{wm}(\mathbf{x}_j, \mathbf{u}) - e_{wm}(\mathbf{x}_j))), \\ &\quad i=1, 2, \cdots, \zeta, \quad j=1, 2, \cdots, \zeta, \end{aligned}$$

则可记为

$$f_{wm}(\mathbf{x}_i, \mathbf{u}) \sim \text{GP}(e_{wm}(\mathbf{x}_i), k_{wm}(\mathbf{x}_i, \mathbf{x}_j)). \quad (9)$$

$f_{wm}(\mathbf{x}_i, \mathbf{u})$ 的协方差函数可用矩阵表示:

$$\begin{aligned} \mathbf{K}_{wm}(\mathbf{X}, \mathbf{X}) &= \\ &\begin{bmatrix} k_{wm}(\mathbf{x}_1, \mathbf{x}_1) & k_{wm}(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k_{wm}(\mathbf{x}_1, \mathbf{x}_\zeta) \\ k_{wm}(\mathbf{x}_2, \mathbf{x}_1) & k_{wm}(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k_{wm}(\mathbf{x}_2, \mathbf{x}_\zeta) \\ \vdots & \vdots & \ddots & \vdots \\ k_{wm}(\mathbf{x}_\zeta, \mathbf{x}_1) & k_{wm}(\mathbf{x}_\zeta, \mathbf{x}_2) & \cdots & k_{wm}(\mathbf{x}_\zeta, \mathbf{x}_\zeta) \end{bmatrix}. \end{aligned} \quad (10)$$

由于 ψ_{wm} 与 $f_{wm}(\mathbf{x}_i, \mathbf{u})$ 同为高斯分布, 且相互独

立, 因此式(7)中, y_{wm} 同样为高斯分布, 可记为

$$\begin{cases} y_{wm} \sim \text{GP}(e_{wm}(\mathbf{x}_i), c_{wm}(\mathbf{x}_i, \mathbf{x}_j)), \\ c_{wm}(\mathbf{x}_i, \mathbf{x}_j) = k_{wm}(\mathbf{x}_i, \mathbf{x}_j) + \sigma_{wm}^2 \delta_{ij}, \end{cases} \quad (11)$$

其中: δ_{ij} 为Kronecker因子, 当 $i=j$ 时, $\delta_{ij}=1$, 反之 $\delta_{ij}=0$. y_{wm} 的协方差函数用矩阵表示如下:

$$\begin{aligned} \mathbf{C}_{wm}(\mathbf{X}, \mathbf{X}) &= \mathbf{K}_{wm}(\mathbf{X}, \mathbf{X}) + \sigma_{wm}^2 \mathbf{I} = \\ &\begin{bmatrix} k_{wm}(\mathbf{x}_1, \mathbf{x}_1) + \sigma_{wm}^2 & \cdots & k_{wm}(\mathbf{x}_1, \mathbf{x}_\zeta) \\ k_{wm}(\mathbf{x}_2, \mathbf{x}_1) & \cdots & k_{wm}(\mathbf{x}_2, \mathbf{x}_\zeta) \\ \vdots & \vdots & \vdots \\ k_{wm}(\mathbf{x}_\zeta, \mathbf{x}_1) & \cdots & k_{wm}(\mathbf{x}_\zeta, \mathbf{x}_\zeta) + \sigma_{wm}^2 \end{bmatrix}. \end{aligned} \quad (12)$$

将建模数据集(式(6))分成学习数据集(式(13))和测试数据集(式(15)). 式(13)中 D_{cm}^{learn} 为中点学习数据集, D_{wm}^{learn} 为不确定度学习数据集, \mathbf{X}_1 为学习数据集输入矩阵, $\mathbf{F}_{cm}^{\text{learn}}$ 为中点学习数据集观测输出矩阵, $\mathbf{F}_{wm}^{\text{learn}}$ 为不确定度学习数据集观测输出矩阵, $\mathbf{F}_{cm}^{\text{learn}}$ 和 $\mathbf{F}_{wm}^{\text{learn}}$ 中的元素仍然分别满足联合高斯分布.

$$\begin{cases} D_{cm}^{\text{learn}} = \{(\mathbf{x}_i, f_{cm}(\mathbf{x}_i, \mathbf{u}))\}, \\ D_{wm}^{\text{learn}} = \{(\mathbf{x}_i, f_{wm}(\mathbf{x}_i, \mathbf{u}))\}, \\ i = 1, 2, \cdots, \phi; \quad 0 < \phi < \zeta, \\ \mathbf{X}_1 = [\mathbf{x}_1 \cdots \mathbf{x}_\phi], \\ \begin{cases} \mathbf{F}_{cm}^{\text{learn}} &= [f_{cm}(\mathbf{x}_1, \mathbf{u}) \cdots f_{cm}(\mathbf{x}_\phi, \mathbf{u})]^T, \\ \mathbf{F}_{wm}^{\text{learn}} &= [f_{wm}(\mathbf{x}_1, \mathbf{u}) \cdots f_{wm}(\mathbf{x}_\phi, \mathbf{u})]^T. \end{cases} \end{cases} \quad (13)$$

由式(7)可知学习数据集所对应的模型输出矩阵为

$$\begin{cases} \mathbf{y}_{cm}^{\text{learn}} = [y_{cm}(\mathbf{x}_1, \mathbf{u}) \cdots y_{cm}(\mathbf{x}_\phi, \mathbf{u})]^T, \\ \mathbf{y}_{wm}^{\text{learn}} = [y_{wm}(\mathbf{x}_1, \mathbf{u}) \cdots y_{wm}(\mathbf{x}_\phi, \mathbf{u})]^T. \end{cases} \quad (14)$$

测试数据集如式(15):

$$\begin{cases} D_{cm}^{\text{test}} = \{(\mathbf{x}_i, f_{cm}(\mathbf{x}_i, \mathbf{u}))\}, \\ D_{wm}^{\text{test}} = \{(\mathbf{x}_i, f_{wm}(\mathbf{x}_i, \mathbf{u}))\}, \\ i = \phi + 1, \phi + 2, \cdots, \zeta, \\ \mathbf{X}_2 = [\mathbf{x}_{\phi+1}, \cdots, \mathbf{x}_\zeta], \\ \begin{cases} \mathbf{F}_{cm}^{\text{test}} &= [f_{cm}(\mathbf{x}_{\phi+1}, \mathbf{u}) \cdots f_{cm}(\mathbf{x}_\zeta, \mathbf{u})]^T, \\ \mathbf{F}_{wm}^{\text{test}} &= [f_{wm}(\mathbf{x}_{\phi+1}, \mathbf{u}) \cdots f_{wm}(\mathbf{x}_\zeta, \mathbf{u})]^T, \end{cases} \end{cases} \quad (15)$$

其中: D_{cm}^{test} 为中点测试数据集, D_{wm}^{test} 为不确定度测试数据集, \mathbf{X}_2 为测试数据集输入矩阵, $\mathbf{F}_{cm}^{\text{test}}$ 为中点测试数据集观测输出矩阵, $\mathbf{F}_{wm}^{\text{test}}$ 为不确定度测试数据集观测输出矩阵.

由式(9)和式(11)可知 $\mathbf{y}_{wm}^{\text{learn}}$, $\mathbf{F}_{wm}^{\text{test}}$ 服从联合高斯分布. 设 $\mathbf{y}_{wm}^{\text{learn}}$ 和 $\mathbf{F}_{wm}^{\text{test}}$ 的均值向量分别为 $\boldsymbol{\mu}_{wm}^{\text{learn}} = \boldsymbol{\mu}_{wm}^{\text{learn}} \times \mathbf{I}$, $\boldsymbol{\mu}_{wm}^{\text{test}} = \boldsymbol{\mu}_{wm}^{\text{test}} \times \mathbf{I}$, \mathbf{I} 为单位列向量, 其中

$$\mu_{wm}^{\text{learn}} = \text{sum}(\mathbf{y}_{wm}^{\text{learn}})/\phi. \quad (16)$$

由式(8)可知 $\mu_{wm}^{\text{learn}}, \mu_{wm}^{\text{test}}$ 均为非负矩阵, 得

$$\begin{bmatrix} \mathbf{y}_{wm}^{\text{learn}} \\ \mathbf{F}_{wm}^{\text{test}} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{wm}^{\text{learn}} \\ \mu_{wm}^{\text{test}} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{wm}(\mathbf{X}_1, \mathbf{X}_1) & \mathbf{K}_{wm}(\mathbf{X}_1, \mathbf{X}_2) \\ \mathbf{K}_{wm}(\mathbf{X}_2, \mathbf{X}_1) & \mathbf{K}_{wm}(\mathbf{X}_2, \mathbf{X}_2) \end{bmatrix}\right). \quad (17)$$

根据联合分布的性质^[17]可得 $\mathbf{F}_{wm}^{\text{test}}$ 在 $\mathbf{y}_{wm}^{\text{learn}}$ 条件下的条件分布为

$$\begin{aligned} p(\mathbf{F}_{wm}^{\text{test}} | \mathbf{y}_{wm}^{\text{learn}}) &\sim \\ N(\mu_{wm}^{\text{test}} + \mathbf{K}_{wm}(\mathbf{X}_2, \mathbf{X}_1) \cdot \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \\ (\mathbf{y}_{wm}^{\text{learn}} - \mu_{wm}^{\text{learn}}), \mathbf{K}_{wm}(\mathbf{X}_2, \mathbf{X}_2) - \\ \mathbf{K}_{wm}(\mathbf{X}_2, \mathbf{X}_1) \cdot \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \mathbf{K}_{wm}(\mathbf{X}_1, \mathbf{X}_2)). \end{aligned}$$

对于某单个输入数据 \mathbf{x}_i , 其对应的函数输出 $f_{wm}(\mathbf{x}_i, \mathbf{u})$ (简写为 f_{wm}^i)同样满足以上所述性质, 即

$$\begin{bmatrix} \mathbf{y}_{wm}^{\text{learn}} \\ f_{wm}^i \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{wm}^{\text{learn}} \\ \mu_{wm} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{wm}(\mathbf{X}_1, \mathbf{X}_1) & \mathbf{K}_{wm}(\mathbf{X}_1, \mathbf{x}_i) \\ \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) & \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{x}_i) \end{bmatrix}\right), \quad (18)$$

式中: $\mathbf{C}_{wm}(\mathbf{X}_1, \mathbf{X}_1)$ 由 $c_{wm}(\mathbf{x}_i, \mathbf{x}_j)$ 构成. f_{wm}^i 在 $\mathbf{y}_{wm}^{\text{learn}}$ 条件下的条件分布为

$$\begin{aligned} p(f_{wm}^i | \mathbf{y}_{wm}^{\text{learn}}) &\sim \\ N(\mu_{wm} + \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \\ (\mathbf{y}_{wm}^{\text{learn}} - \mu_{wm}^{\text{learn}}), \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{x}_i) - \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \\ \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \mathbf{K}_{wm}(\mathbf{X}_1, \mathbf{x}_i)), \end{aligned}$$

从而得到单个测试数据的预测均值为

$$e_{wm}^*(\mathbf{x}_i) = \mu_{wm} + \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot (\mathbf{y}_{wm}^{\text{learn}} - \mu_{wm}^{\text{learn}}). \quad (19)$$

预测方差为

$$\begin{aligned} \sigma_{wm}^{*2}(\mathbf{x}_i) &= \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{x}_i) - \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \\ &\mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \mathbf{K}_{wm}(\mathbf{X}_1, \mathbf{x}_i). \end{aligned}$$

预测均值 $e_{wm}^*(\mathbf{x}_i)$ 即为区间不确定度模型在输入为 \mathbf{x}_i 时的预测输出, 而 $\sigma_{wm}^{*2}(\mathbf{x}_i)$ 为模型预测置信水平. 同理得到区间中点模型的预测均值为

$$e_{cm}^*(\mathbf{x}_i) = \mu_{cm} + \mathbf{K}_{cm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \mathbf{C}_{cm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot (\mathbf{y}_{cm}^{\text{learn}} - \mu_{cm}^{\text{learn}}). \quad (20)$$

预测方差为

$$\begin{aligned} \sigma_{cm}^{*2}(\mathbf{x}_i) &= \mathbf{K}_{cm}(\mathbf{x}_i, \mathbf{x}_i) - \mathbf{K}_{cm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \\ &\mathbf{C}_{cm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot \mathbf{K}_{cm}(\mathbf{X}_1, \mathbf{x}_i). \end{aligned}$$

由式(9)(11)(20)知, $c_{wm}(\mathbf{x}_i, \mathbf{x}_j), c_{cm}(\mathbf{x}_i, \mathbf{x}_j), \mu_{cm}, \mu_{wm}$ 是未知项. 而函数建模的过程为未知项的求取过程. 未知项的具体求取过程见第4.3, 4.4节.

4.3 协方差函数的选择(The choice of Covariance functions)

由式(11)可得GP模型的全协方差函数为

$$c_{wm}(\mathbf{x}_i, \mathbf{x}_j) = k_{wm}(\mathbf{x}_i, \mathbf{x}_j) + \sigma_{wm}^2 \delta_{ij},$$

其中 $k_{wm}(\mathbf{x}_i, \mathbf{x}_j)$ 为协方差函数也称核函数. 核函数中的未知量称为超参数. 由于均方指数协方差函数是无限可微的, 它的随机过程非常平稳, 因此它是最常用的协方差函数. 故本节选其作为高斯过程核函数:

$$\begin{aligned} k_{wm}(\mathbf{x}_i, \mathbf{x}_j) &= \\ v_{wm}^2 \cdot \exp\left(-\frac{1}{2} \cdot (\mathbf{x}_i - \mathbf{x}_j) \cdot \mathbf{M} \cdot (\mathbf{x}_i - \mathbf{x}_j)^T\right). \end{aligned} \quad (21)$$

令

$$\mathbf{M} = \begin{bmatrix} \lambda_{wm}^{-2} & & \\ & \ddots & \\ & & \lambda_{wm}^{-2} \end{bmatrix},$$

则式(21)变为

$$\begin{aligned} k_{wm}(\mathbf{x}_i, \mathbf{x}_j) &= \\ v_{wm}^2 \cdot \exp\left(-\frac{1}{2 \cdot \lambda_{wm}^2} \cdot (\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{x}_i - \mathbf{x}_j)^T\right). \end{aligned} \quad (22)$$

故全协方差函数为

$$\begin{aligned} c_{wm}(\mathbf{x}_i, \mathbf{x}_j) &= k_{wm}(\mathbf{x}_i, \mathbf{x}_j) + \sigma_{wm}^2 \delta_{ij} = \\ v_{wm}^2 \cdot \exp\left(-\frac{1}{2 \cdot \lambda_{wm}^2} \cdot (\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{x}_i - \mathbf{x}_j)^T\right) &+ \sigma_{wm}^2 \delta_{ij}, \end{aligned} \quad (23)$$

其中未知的参数有 $v_{wm}, \lambda_{wm}, \sigma_{wm}$. 对于可能度模型式(19)还存在1个未知的变量 μ_{wm} . 将 μ_{wm} 与 $v_{wm}, \lambda_{wm}, \sigma_{wm}$ 合并到一起, 令得到集合为未知参数集即 $\omega_{wm} = \{\mu_{wm}, v_{wm}, \lambda_{wm}, \sigma_{wm}\}$. 同理对于区间中点的未知参数集为 $\omega_{cm} = \{\mu_{cm}, v_{cm}, \lambda_{cm}, \sigma_{cm}\}$. 通过建模数据集对GP模型的训练就是获取未知参数集的过程.

4.4 模型参数求解策略(Solving strategy of model parameters)

超参数通常用共轭梯度法对添加了白噪声的GP模型似然函数最大化得到^[17], 然后用最小二乘法获得其余未知参数. 该方法速度较快, 但区间函数的代理模型构建需考虑两个因素: 1) 模型预测的准确性; 2) 异常解出现的次数即相关性. 两种因素前者预测准确性越高越好(误差越小), 后者异常解出现次数越低越好. 两种因素相关但并不完全一致, 即某些算法中模型预测准确性虽然很高, 但异常解出现的次数也很

大. 而利用共轭梯度法进行建模只是以提高模型的预测准确性为准, 辨识后的区间模型缺少相关性, 如文献[6]. 而遗传算法作为一种仿生最优化算法可有效克服经典方法的局限性, 为此本节提出一种基于区间中点、不确定度的遗传算法建模策略. 该策略在考虑模型预测准确性的同时兼顾区间函数的相关性.

由于区间中点和不确定度建模方法类似, 本节以不确定度建模为例进行论述. 对于不确定度的建模, 需先按照式(13)和式(15)将 D_{wm} 分为学习样本集 D_{wm}^{learn} 和测试样本集 D_{wm}^{test} , 利用 D_{wm}^{learn} 根据式(19)对 D_{wm}^{test} 的函数输出值进行预测, 对于未知参数集需以保证模型对 D_{wm}^{test} 预测值和实际值的误差最小为主, 还需尽量降低不确定度预测值出现负值(异常解)的次数. 其中未知参数集为待优化值, 适应度函数可取模型的预测误差函数:

$$\begin{aligned} Q_{wm} : \chi_{wm} = \arg \min_{\omega_{wm}} \{ \tau_{wm} \}, \\ \tau_{wm} = \sum_{i=1}^{\zeta-\phi} |f_{wm}(\mathbf{x}_i, \mathbf{u}) - e_{wm}^*(\mathbf{x}_i)| = \\ \sum_{i=1}^{\zeta-\phi} |f_{wm}(\mathbf{x}_i, \mathbf{u}) - (\mu_{wm}^i + \mathbf{K}_{wm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \\ \mathbf{C}_{wm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot (\mathbf{y}_{wm}^{\text{learn}} - \boldsymbol{\mu}_{wm}^{\text{learn}}))|, \end{aligned} \quad (24)$$

其中: $f_{wm}(\mathbf{x}_i, \mathbf{u})$ 见式(15), $e_{wm}^*(\mathbf{x}_i)$ 为通过式(19)预测得到的区间不确定度, $|\cdot|$ 为求绝对值操作.

由于排除劣解时需考虑准确性和相关性两种属性. 因此种群解的筛选问题可以看成多属性决策问题, 对于该问题可通过几何加权算子进行求解, 但是考虑到两种属性的量级不同, 若处理不当, 最终决策可能受到部分属性主导, 故为获取优势解本节采用双层筛选策略. 设待筛选种群规模为 pop^R , 需保留的优势解的个数为 pop^* , 双层筛选策略如下:

第1层相关性筛选策略: 为体现相关性, 遗传算法在种群筛选时, 需先根据异常解出现次数剔除 $\omega_{\text{weight}} * (\text{pop}^R - \text{pop}^*)$ ($0 \leq \omega_{\text{weight}} \leq 1$) 个劣解, 异常解出现次数越少越优.

第2层准确性筛选策略: 为保证模型的预测准确性, 第1层策略筛选后, 需根据模型预测误差再次排序, 剔除另外 $(1 - \omega_{\text{weight}}) * (\text{pop}^R - \text{pop}^*)$ 个劣解, 预测误差越小越优, 从而获得 pop^* 个优势解. 其中系数 ω_{weight} 将决定两种属性的权重, ω_{weight} 越大相关性的权重越大, 准确性权重越小, 反之 ω_{weight} 越小, 相关性的权重越小, 准确性权重越大, 这里称其为权重系数.

对于区间中点建模, 区别于区间不确定度的地方有3点:

- 1) 建模数据集取 D_{cm} ;
- 2) 适应值函数取式(25);

$$\begin{aligned} Q_{cm} : \chi_{cm} = \arg \min_{\omega_{cm}} \{ \tau_{cm} \}, \\ \tau_{cm} = \sum_{i=1}^{\zeta-\phi} |f_{cm}(\mathbf{x}_i, \mathbf{u}) - e_{cm}^*(\mathbf{x}_i)| = \\ \sum_{i=1}^{\zeta-\phi} |f_{cm}(\mathbf{x}_i, \mathbf{u}) - (\mu_{cm}^i + \mathbf{K}_{cm}(\mathbf{x}_i, \mathbf{X}_1) \cdot \\ \mathbf{C}_{cm}^{-1}(\mathbf{X}_1, \mathbf{X}_1) \cdot (\mathbf{y}_{cm}^{\text{learn}} - \boldsymbol{\mu}_{cm}^{\text{learn}}))|, \end{aligned} \quad (25)$$

其中: $f_{cm}(\mathbf{x}_i, \mathbf{u})$ 见式(15), $e_{cm}^*(\mathbf{x}_i)$ 为通过式(20)预测得到的区间中心.

3) 种群解的筛选只需考虑预测准确性, 即双层筛选策略中只考虑第2层即可.

5 算法步骤(Steps of algorithm)

5.1 建模步骤(Steps of modeling)

利用遗传算法对区间函数的不确定度建模过程为:

步骤1 取用于建模的数据集为 D_{wm} (见第4.1节), 进而得到 D_{wm}^{learn} 和 D_{wm}^{test} . 适应度函数为式(24). 设置未知参数集中各元素的取值范围, 将未知参数集作为待优化的变量, 初始化种群 $N^*(k)$, 解的个数为 pop^* , 设置遗传初始代数 $k = 0$, 遗传总代数 G^* .

步骤2 对种群 $N^*(k)$ 进行交叉变异得到种群 $N_m^*(k)$. 合并 $N^*(k)$ 和 $N_m^*(k)$ 得到种群 $R^*(k)$.

步骤3 按照双层筛选策略第1层策略, 计算得到种群 $R^*(k)$ 的规模 pop^R , 将每个解代入模型对测试样本集 D_{wm}^{test} 的函数输出值进行预测, 根据预测出现负值(异常解)的个数, 按照从小到大的顺序对种群 $R^*(k)$ 进行排序, 排除后 $\omega_{\text{weight}} * (\text{pop}^R - \text{pop}^*)$ 个解, 从而更新种群 $R^*(k)$.

步骤4 按照双层筛选策略第2层策略, 将更新后的种群 $R^*(k)$ 按照适应度从小到大排序并选出适应度较小的 pop^* 个解, 构成下一代种群 $N^*(k+1)$. 判断算法是否进化了 G^* 代, 如果满足, 转步骤5; 否则, 令 $k = k + 1$, 转到步骤2.

步骤5 输出适应值最小的未知参数集.

对于区间函数中点的建模过程, 需将建模数据集由 D_{wm} 换为 D_{cm} , 适应度函数由式(24)换为式(25), 且建模步骤中无步骤3. 其他与区间不确定度的建模过程相同. 目标函数和约束函数的建模过程相同.

5.2 改进的NSGA-II算法步骤(Steps of improved NSGA-II algorithm)

本文将所设计的基于区间中点和不确定度的高斯代理模型引入区间多目标NSGA-II算法^[18], 算法具体过程如下:

步骤1 初始化种群 $N(k)$, 包含的解数为 N , 通过代理模型预测优化函数的输出区间值, 设置初始遗传代数 $k = 0$, 遗传总代数 G , 每个约束函数的约束允许违背度 θ_i^{max} .

步骤2 算出 $N(k)$ 中满足约束条件^[18]的解数 N_{o1} , 并记录满足约束条件解.

步骤3 对种群 $N(k)$ 进行交叉变异得到种群 $N_m(k)$. 利用代理模型算出 $N_m(k)$ 中满足约束条件的解数 N_{o2} , 并记录满足约束条件解.

步骤4 合并所有满足约束的解得到种群 $N_w(k)$, 判断是否 $N_{o1} + N_{o2} \geq N$, 若不满足, 返回步骤3继续求满足约束条件的解数 N_{o2i} , 设 i 为步骤4的循环次数, 继续判断 $N_{o1} + \sum_{i=1}^n N_{o2i} \geq N$, 直到满足条件进入第5步.

步骤5 针对步骤4中满足约束条件的解, 先按照解的序值排序, 序值相同的解再根据区间拥挤距离排序, 并选取前 N 个优势解, 构成下一代种群 $N(k) (k=1)$, 转到步骤6.

步骤6 对种群 $N(k)$ 进行交叉变异后, 利用代理模型去除不满足约束的解得到种群 $N_w(k)$. 合并 $N(k)$ 和 $N_w(k)$. 得到种群 $W(k)$, 计算种群 $W(k)$ 中解的序值和拥挤距离.

步骤7 针对种群 $W(k)$, 按照精英策略选取前 N 个优势解, 构成下一代种群 $N(k+1)$.

步骤8 判断算法是否进化了 G 代, 如果满足, 转步骤9, 否则, 令 $k = k + 1$, 转到步骤6.

步骤9 对优化对象评估实验输出Pareto解集和前沿.

上述算法需要说明的内容如下:

1) 步骤4中在保留了可行解的同时进行了可行解数目的判断, 若数目过少则重新通过遗传操作补充可行解, 这样做不仅保证了种群有足够数量的可行解, 而且补充解是通过遗传操作获得, 其要优于随机获得.

2) 步骤9中, 最终获得的优势解要与优化对象(即目标函数和约束函数)而不是代理模型进行实验评估, 因为实际应用中算出的优势解要用于实际优化问题中.

6 仿真分析(Simulation analysis)

本文以优化问题 $Q1$ ^[18], $Q2$ ^[22]为例, 验证所提算法的有效性.

$$\begin{aligned}
 Q1: \min f_1(\mathbf{x}, \mathbf{u}) &= u_1(x_1 + x_2 - 7.5)^2 + \\
 &\quad u_2^2(x_2 - x_1 + 3)^2/4, \\
 f_2(\mathbf{x}, \mathbf{u}) &= u_1^2(x_1 - 1)^2/4 + u_2^3(x_2 - 4)^2/2, \\
 \text{s.t.} \quad &\begin{cases} g_1(\mathbf{x}, \mathbf{u}) = \\ u_1^2(x_1 - 2)/2 + u_2x_2 - 2.5 \leq [0, 0.3], \\ g_2(\mathbf{x}, \mathbf{u}) = \\ u_1^3x_2 + u_2^2x_1 - 3.85 - \\ 8u_2^2(x_2 - x_1 + 0.65)^2 \leq [0, 0.3], \\ x_1 \in [0, 5], x_2 \in [0, 3], u_1, u_2 \in [0.9, 1.1], \end{cases}
 \end{aligned} \tag{26}$$

$$\begin{aligned}
 Q2: \min f_1(\mathbf{x}, \mathbf{u}) &= u_1[\sin(5u_1x) - \cos(5u_1x)], \\
 f_2(\mathbf{x}, \mathbf{u}) &= e^{2+u_2 \cos(5u_2x)}, \\
 x &\in [-5, 5], u_1, u_2 \in [0.99, 1.01].
 \end{aligned} \tag{27}$$

模型辨识参数为: 由于 μ_{wm} 和 μ_{cm} 为联合高斯分布函数输出值的真实均值, 可根据已有的经验数据给出范围(经验输出值的最大值作为上界, 最小值作为下界), 为提高最优 μ_{wm} 能在定义域内搜索到的概率, 这里将上下界做了扩展:

$$\begin{aligned}
 \mu_{wm} &\in \\
 &[\min(\mathbf{F}_{wm}^{\text{learn}}) - 0.2(\max(\mathbf{F}_{wm}^{\text{learn}}) - \min(\mathbf{F}_{wm}^{\text{learn}})), \\
 &\max(\mathbf{F}_{wm}^{\text{learn}}) + 0.2(\max(\mathbf{F}_{wm}^{\text{learn}}) - \min(\mathbf{F}_{wm}^{\text{learn}}))], \\
 \mu_{cm} &\in \\
 &[\min(\mathbf{F}_{cm}^{\text{learn}}) - 0.2(\max(\mathbf{F}_{cm}^{\text{learn}}) - \min(\mathbf{F}_{cm}^{\text{learn}})), \\
 &\max(\mathbf{F}_{cm}^{\text{learn}}) + 0.2(\max(\mathbf{F}_{cm}^{\text{learn}}) - \min(\mathbf{F}_{cm}^{\text{learn}}))].
 \end{aligned}$$

如不特殊说明, 其余参数的范围给定如下:

$$\begin{aligned}
 v_{cm} &\in [-50, 50], \lambda_{cm} \in [-50, 50], \sigma_{cm} \in [-10, 10], \\
 v_{wm} &\in [-50, 50], \lambda_{wm} \in [-50, 50], \sigma_{wm} \in [-10, 10], \\
 \zeta &= 160, \phi = 80, G^* = 20, \text{pop}^* = 40, \\
 \omega_{\text{weight}} &= 0.5, \\
 \text{变异概率} &= 0.2.
 \end{aligned}$$

NSGA-II算法参数为

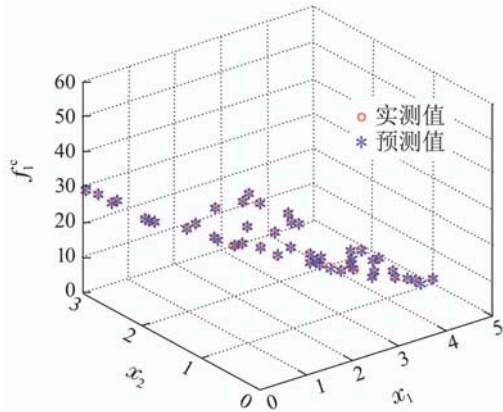
$$G = 150, \text{pop} = 40, \theta_1^{\max} = 0.5, \theta_2^{\max} = 0.5,$$

其它参数同文献[18].

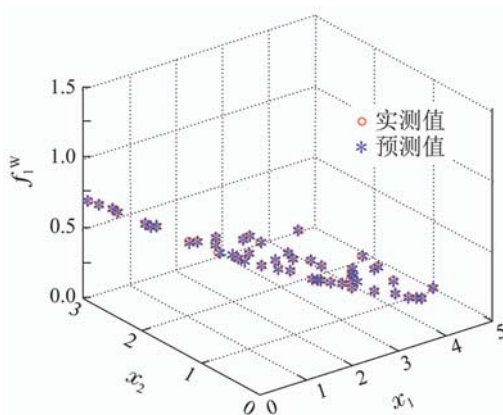
针对目标函数为区间数这一特点, 本文引入了6种测度: E 测度^[18]、 C 测度^[18]、 D 测度^[18]、最大误差率^[10]、平均误差率^[10]、均方根误差率^[10], 以检验所提出改进NSGA-II的性能. 以下实验数据若无特殊说明, 均为算法独立运行20次获得的测度均值.

6.1 代理模型辨识分析(Analysis of surrogate model identification)

图1为 $Q1$ 中 f_1 的代理模型辨识准确性, 其中*为预测值, o 为实测值. $u_1, u_2 \in [0.99, 1.01]$. 由图1可知无论是区间中点还是不确定度, 本文代理模型均可准确预测. 为了具体分析, 表1-2列出了 $Q1, Q2$ 所建模型的预测性能测度, 其中本文所介绍的高斯代理模型为模型1, 以文献[6]所介绍的高斯代理模型为模型2. 由表可知优化问题 $Q1$ 中 f_1 模型1的3种误差率略差, 优化问题 $Q1$ 中 g_2 模型1的最大误差略差, 而对于其余测度, 模型1均要好于模型2, 尤其是模型1所有函数异常解次数明显比模型2少. 故本文所述区间函数建模法能有效解决代理模型中区间函数辨识缺少相关性的问题.



(a) Q1: f_1 区间中点的辨识准确性



(b) Q1: f_1 区间不确定的辨识准确性

图1 代理模型的辨识准确性

Fig. 1 The identification accuracy of surrogate models

表1 优化问题Q1的模型测度

Table 1 Measures of model for Q1

目标函数	f_1		f_2	
代理模型	模型1	模型2	模型1	模型2
最大误差	0.0396	0.0299	0.0022	0.0038
平均误差	0.0007	0.0006	0.0003	0.0004
均方误差	0.0061	0.0052	0.0024	0.0025
异常次数	0	3.5000	0	0

目标函数	g_1		g_2	
代理模型	模型1	模型2	模型1	模型2
最大误差	1.9026	12.9722	10.0732	9.0748
平均误差	0.0046	0.0218	0.0205	0.0306
均方误差	0.0055	0.0068	0.0299	0.0610
异常次数	1.0000	34.4500	1.1500	30.3000

表2 优化问题Q2的模型测度

Table 2 Measures of model for Q2

目标函数	f_1		f_2	
代理模型	模型1	模型2	模型1	模型2
最大误差	4.5752	25.7602	0.0543	0.1436
平均误差	0.0152	0.0391	0.0049	0.0246
均方误差	0.0114	0.0103	0.0790	0.2562
异常次数	0.1000	1.1000	0.9500	54.2000

6.2 权重系数对代理模型的影响(The influence of the weight coefficient for surrogate model)

权重系数 ω_{weight} 决定着相关性和准确性的权重大小,权重赋值的合理与否,对结果的科学性和合理性起着至关重要的作用,若某一因素的权重发生变化,将会影响整个算法的结果.因此,权重的赋值必须要做到科学和客观,这就要求寻求合适的权重确定方法.表3和表4给出了3种不同权重系数下的模型预测误差率以及异常解次数.由表可知对于大多数函数的误差率影响, ω_{weight} 并非越大越好也非越小越好, ω_{weight} 需根据具体问题而定.对于异常解次数由表可知Q1, Q2均满足规律:随着权重系数的增加即相关性权重增加,异常解次数降低.这说明了本文所提策略可有效解决区间函数建模缺少相关性的问题.考虑到当取 $\omega_{weight} = 0.5$ 时,多数模型的预测准确性较优,故以下仿真中 ω_{weight} 均取0.5.

表3 权重系数对Q1模型的影响

Table 3 Influence of weight coefficient for Q1

目标函数	f_1			f_2		
权重系数	0.2	0.5	0.8	0.2	0.5	0.8
最大误差	0.0227	0.0396	0.0507	0.0029	0.0022	0.0043
平均误差	0.0004	0.0007	0.0008	0.0004	0.0003	0.0006
均方误差	0.0046	0.0061	0.0062	0.0023	0.0024	0.0034
异常次数	0	0	0	0	0	0

目标函数	g_1			g_2		
权重系数	0.2	0.5	0.8	0.2	0.5	0.8
最大误差	4.7445	1.9026	59.8678	5.8941	10.0732	28.7662
平均误差	0.0069	0.0046	0.0715	0.0162	0.0205	0.0451
均方误差	0.0031	0.0055	0.0110	0.0287	0.0299	0.0448
异常次数	2.2500	1.0000	0.6000	3.7000	1.1000	0.8000

表4 Q2权重系数对代理模型的影响

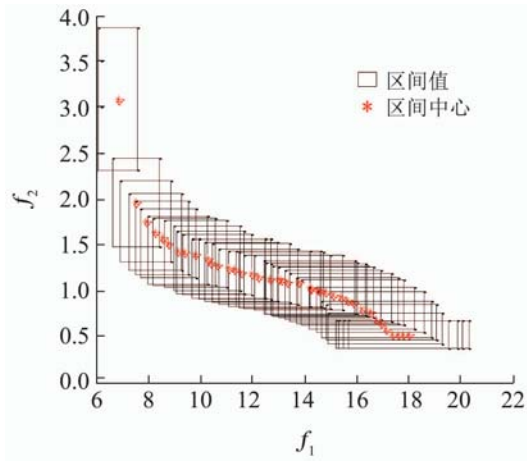
Table 4 Influence of weight coefficient for Q2

目标函数	f_1			f_2		
权重系数	0.2	0.5	0.8	0.2	0.5	0.8
最大误差	19.7916	4.5752	74.3361	0.0724	0.0543	0.2263
平均误差	0.0291	0.0152	0.1451	0.0044	0.0049	0.0347
均方误差	0.0084	0.0114	0.0313	0.1126	0.0790	0.2804
异常次数	0.4000	0.1000	0	1.6000	0.9000	0.5500

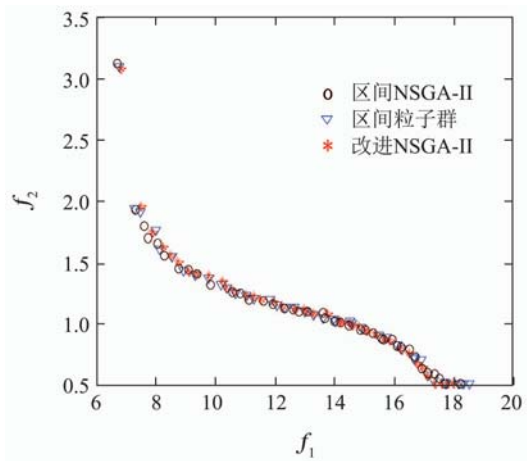
6.3 代理模型的有效性分析 (Effectiveness analysis of surrogate model)

本文算法是在优化函数未知的条件下优化求解的,为了说明本文所建代理模型的有效性,取文献[18]中区间NSGA-II算法以及文献[22]中区间粒子群算法进行对比说明.这两种文献算法均是在优化函数已知的条件下进行优化求解,本节将区间NSGA-II所得前沿作为参考前沿求C测度.图2为两种算法的Pareto前沿分布图,其中图(a)(c)为优化问题Q1, Q2经本文算法获得的Pareto前沿,图(b)(d)为本文算法、区间NSGA-II

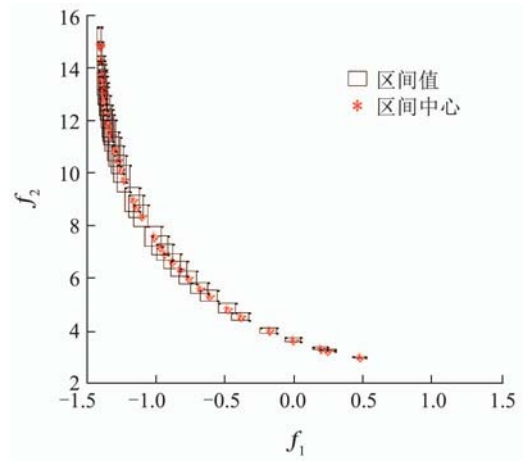
II算法以及区间粒子群算法的前沿中心分布图. 表5为前沿所对应的性能测度, 由图2知本文算法的Pareto前沿非常接近参考前沿, 3种算法的前沿中心几乎重合.



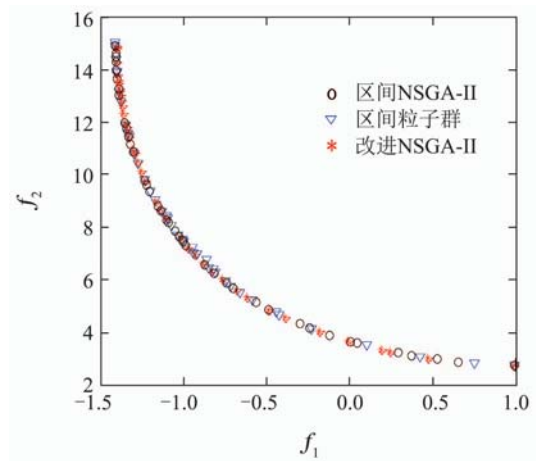
(a) Q1本文算法Pareto前沿



(b) Q1 Pareto前沿中心



(c) Q2本文算法Pareto前沿



(d) Q2 Pareto前沿中心

图 2 算法的Pareto前沿

Fig. 2 The Pareto front of algorithms

表 5 算法的测度比较

Table 5 The comparison of algorithms on performance measures

测度	本文算法			区间NSGA-II算法			区间粒子群算法		
	E	D	C	E	D	C	E	D	C
Q1	0.2102	11.6321	0.3075	0.2012	11.9637	0.2892	0.2236	11.9292	0.2982
Q2	0.2367	12.1129	0.2599	0.2062	12.3690	0.1939	0.3145	12.4047	0.2330

就表5中的前沿测度而言, 本文算法Q1, Q2的 E 测度均要好于区间粒子群算法, 其余测度较差, 但是差距均很小. 这说明对于目标函数未知且评估成本高昂的区间多目标优化问题, 本文算法利用待优化问题的经验数据便可获得与优化函数已知区间多目标算法相接近的Pareto前沿.

7 结论(Conclusions)

针对优化函数未知的昂贵区间多目标优化问题, 本文将代理模型的思想引入区间多目标NSGA-II算

法. 为了构建兼顾预测准确性和相关性的区间函数中点和不确定度的代理模型, 文中提出双层筛选策略, 并将其嵌入遗传算法用于求解高斯模型模型参数. 经仿真实验可知该建模方法相比传统的高斯建模方法不仅有效降低区间函数模型异常解出现的概率, 而且模型预测准确性也有所提升. 将该代理模型应用到NSGA-II算法中, 所得前沿趋近于优化函数已知区间多目标算法, 因此在实际工程中, 比优化函数已知的优化算法应用范围更广泛.

参考文献(References):

- [1] ZHONG Y F. *Multi-objective optimized applications in the safety design of vehicle collision* [D]. Hunan: Hunan University Press, 2013.
- [2] KONG B. *Research on car-body styling design based on aerodynamics* [D]. Wuhan: Wuhan University of Technology, 2008.
- [3] DOUGUET D. e-LEA3D: a computational-aided drug design web server [J]. *Nucleic Acids Research*, 2010, 38(5): 615 – 621.
- [4] LIM D, JIN Y, ONG Y S, et al. Generalizing surrogate-assisted evolutionary computation [J]. *IEEE Transactions on Evolutionary Computation*, 2010, 14(3): 329 – 355.
- [5] KUNAKOTE T, BUREERAT S. Surrogate-assisted multi-objective evolutionary algorithms for structural shape and sizing optimization [J]. *Mathematical Problems in Engineering*, 2013, 20(1): 27 – 53.
- [6] LU J, LI B, JIN Y. An evolution strategy assisted by an ensemble of local Gaussian process models [C] // *Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation Conference*. New York: ACM, 2013: 447 – 454.
- [7] SUN C, JIN Y, ZENG J, et al. A two-layer surrogate-assisted particle swarm optimization algorithm [J]. *Soft Computing*, 2014, 19(6): 1461 – 1475.
- [8] ZHANG Q, LIU W, TSANG E, et al. Expensive multi-objective optimization by MOEA/D with Gaussian process model [J]. *IEEE Transactions on Evolutionary Computation*, 2010, 14(3): 456 – 474.
- [9] SUN X, GONG D, JIN Y, et al. A new surrogate-assisted interactive genetic algorithm with weighted semisupervised learning [J]. *IEEE Transactions on Systems Man & Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems Man & Cybernetics Society*, 2012, 43(2): 685 – 698.
- [10] SUN Xiaoyan, CHEN Shanshan, GONG Dunwei, et al. Weighted multi-output Gaussian process-based surrogate of interactive genetic algorithm with individuals interval fitness [J]. *Acta Automatica Sinica*, 2014, 40(2): 172 – 184.
(孙晓燕, 陈姗姗, 巩敦卫, 等. 基于区间适应值交互式遗传算法的加权多输出高斯过程代理模型 [J]. *自动化学报*, 2014, 40(2): 172 – 184.)
- [11] NGE0 J, TAMEI T, SHIBATA T. Estimation of continuous multi-DOF finger joint kinematics from surface EMG using a multi-output Gaussian process [C] // *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Chicago: IEEE, 2014: 3537 – 3540.
- [12] XU Zhengguang, SUN Changping, WU Jinxia. Moving pattern measured by interval number for modeling and control [J]. *Control Theory & Applications*, 2012, 29(9): 1115 – 1124.
(徐正光, 孙昌平, 吴金霞. 基于区间数度量的运动模式建模与控制 [J]. *控制理论与应用*, 2012, 29(9): 1115 – 1124.)
- [13] HE Zhikun, LIU Guangbin, ZHAO Xijing, et al. Overview of Gaussian process regression [J]. *Control and Decision*, 2013, 28(8): 1121 – 1129, 1137.
(何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述 [J]. *控制与决策*, 2013, 28(8): 1121 – 1129, 1137.)
- [14] GAN Di, KE Deping, SUN Yuanzhang, et al. Short-term wind speed probabilistic forecasting based on EEMD and coupling GA-GPR [J]. *Transactions of China Electrotechnical Society*, 2015, 30(11): 138 – 147.
(甘迪, 柯德平, 孙元章, 等. 基于集合经验模式分解和遗传-高斯过程回归的短期风速概率预测 [J]. *电工技术学报*, 2015, 30(11): 138 – 147.)
- [15] XU Chong, LIU Baoguo, LIU Kaiyun, et al. Intelligent analysis model of landslide displacement time series based on coupling PSO-GPR [J]. *Rock and Soil Mechanics*, 2011, 32(6): 1669 – 1675.
(徐冲, 刘保国, 刘开云, 等. 基于粒子群-高斯过程回归耦合算法的滑坡位移时序分析预测智能模型 [J]. *岩土力学*, 2011, 32(6): 1669 – 1675.)
- [16] YAN G W. *Knowledge-based multi-agent mind evolutionary algorithm and its applications in engineering* [D]. Taiyuan: Taiyuan University of Technology, 2007
- [17] LI Jun, ZHANG Youpeng. Single-step and multi-step prediction of chaotic time series using Gauss process model [J]. *Chinese Journal of Physics*, 2011, 66(7): 150 – 159.
(李军, 张友鹏. 基于高斯过程的混沌时间序列单步与多步预测 [J]. *物理学报*, 2011, 66(7): 150 – 159.)
- [18] CHEN Zhiwan, CHEN Lin, BAI Xin. Interactive multi-attribute decision-making NSGA-II for constrained multi-objective optimization with interval numbers [J]. *Control and Decision*, 2015, 30(5): 865 – 870.
(陈志旺, 陈林, 白铨. 求解约束多目标区间优化的交互多属性决策NSGA-II算法 [J]. *控制与决策*, 2015, 30(5): 865 – 870.)
- [19] JIANG C. *Theories and algorithms of uncertain optimization based on interval* [D]. Hunan: Hunan University, 2008.
- [20] SUN Xiaoyan, REN Jie, GONG Dunwei. Interval-fitness interactive genetic algorithms with varying population size based on semisupervised learning [J]. *Control Theory & Applications*, 2011, 28(5): 610 – 618.
(孙晓燕, 任洁, 巩敦卫. 基于半监督学习的变种群规模区间适应值交互式遗传算法 [J]. *控制理论与应用*, 2011, 28(5): 610 – 618.)
- [21] GONG Dunwei, REN Jie, SUN Xiaoyan. Neural network surrogate models of interactive genetic algorithms with individual. sinterval fitness [J]. *Control and Decision*, 2009, 24(10): 1522 – 1525.
(巩敦卫, 任洁, 孙晓燕. 区间适应值交互式遗传算法神经网络代理模型 [J]. *控制与决策*, 2009, 24(10): 1522 – 1525.)
- [22] ZHANG Yong, GONG Dunwei, HAO Guosheng, et al. Particle swarm optimization for multi-objective systems with interval parameters [J]. *Acta Automatica Sinica*, 2008, 34(8): 921 – 928.
(张勇, 巩敦卫, 郝国生, 等. 含区间参数多目标系统的微粒群优化算法 [J]. *自动化学报*, 2008, 34(8): 921 – 928.)

作者简介:

陈志旺 (1978–), 男, 博士, 副教授, 目前研究方向为多属性决策、区间数优化, E-mail: czwaaaron@ysu.edu.cn;

白铨 (1988–), 男, 硕士研究生, 目前研究方向为多目标优化, E-mail: 15233013272@163.com;

杨七 (1988–), 男, 硕士研究生, 目前研究方向为多属性决策, E-mail: yang0000qi@126.com;

黄兴旺 (1991–), 男, 硕士研究生, 目前研究方向为多目标优化, E-mail: xingwang22@foxmail.com;

李国强 (1984–), 男, 博士, 讲师, 目前研究方向为复杂工业系统的建模与控制, E-mail: zhihuiyuan@163.com.