

基于决策空间变换最近邻方法的 Pareto 支配性预测

李文彬^{1,2} 贺建军¹ 冯彩英² 郭观七²

摘要 为提高在决策空间运用最近邻方法预测多目标优化 Pareto 支配性的精度, 提出一种基于决策空间变换的最近邻预测方法. 在分析目标函数与决策分量相关性的基础上, 提出属性变化趋势模型的构造方法, 建立低计算成本的属性趋势代理模型. 通过属性趋势模型引入决策空间到目标空间的映射知识, 对多目标问题的决策空间进行变换, 使决策空间的最近邻更有效反映目标空间的最近邻. 选取具有不同相关系数特征的典型多目标优化问题, 进行 Pareto 支配性预测的可对比实验, 结果表明在新空间中运用最近邻方法可显著提高分类准确性.

关键词 趋势模型, 空间变换, 最近邻方法, Pareto 支配性

引用格式 李文彬, 贺建军, 冯彩英, 郭观七. 基于决策空间变换最近邻方法的 Pareto 支配性预测. 自动化学报, 2017, 43(2): 294–301

DOI 10.16383/j.aas.2017.c150877

Prediction of Pareto Dominance Using Nearest Neighbor Method Based on Decision Space Transformation

LI Wen-Bin^{1,2} HE Jian-Jun¹ FENG Cai-Ying² GUO Guan-Qi²

Abstract In this paper, nearest neighbor prediction is used to decide Pareto dominance of candidate solutions in expensive multi-objective optimization. For improving the accuracy of predicting Pareto dominance in decision space, a transformation method of decision space is proposed. Based on correlation analysis of objective functions and decision attributes, computationally efficient attribute tendency models of objective functions are set up. These models are used to re-construct the decision space such that the knowledge of objective space is introduced and the neighborhood relation in decision space can more effectively reflect that in the original objective space. Experiments of Pareto dominance prediction are carried out on a group of typical multi-objective optimization problems. The results indicate that the prediction of Pareto dominance using the proposed method is more accurate and efficient than the existing methods.

Key words Tendency model, space transformation, nearest neighbor method, Pareto dominance

Citation Li Wen-bin, He Jian-Jun, Feng Cai-Ying, Guo Guan-Qi. Prediction of Pareto dominance using nearest neighbor method based on decision space transformation. *Acta Automatica Sinica*, 2017, 43(2): 294–301

进化算法是求解多目标优化 (Multi-objective optimization, MO) 问题的主要方法之一^[1–2]. 对于昂贵 MO 问题^[3–4], 目标向量评估所需计算时间或实验成本高昂, 大量昂贵评估必然导致成本灾难. 如何有效解决昂贵 MO 问题, 目前已引起学界的广泛关注. 通过深入研究, 国内外学者提出了基于代理模

型辅助进化算法^[5–9] 和模式分类方法^[10–13] 的解决方案.

在解决昂贵 MO 问题的代理模型辅助进化算法研究方面, Zhang 等将昂贵多目标优化问题分解为多个单目标优化子问题, 用高斯过程模型预测各单目标子问题的解, 对减少某些多目标优化问题的目标评估次数取得了较好效果^[5]. Akhtar 等使用径向基函数迭代更新响应面模型来近似昂贵目标函数, 通过评估少量的个体获得较好的非支配集^[8]. Montemayor-García 等对几种常用代理模型应用于多目标优化问题的准确性、鲁棒性、有效性和可扩展性进行实验研究, 比较各模型的优缺点^[9]. 由于各种代理模型都有各自的优势和缺点, 因此需要在不同优化问题上选择各自合适的模型. 但随着实际问题复杂程度的增强, 建模的难度越来越大, 模型的精度也将难以保证, 导致近似优化往往得不到与真实模型一致的最优解.

在解决昂贵 MO 问题的模式分类方法研究方面, Guo 等首先提出采用模式识别方法进行多目标

收稿日期 2015-12-29 录用日期 2016-08-15
Manuscript received December 29, 2015; accepted August 15, 2016

国家自然科学基金 (60975049, 61174132), 湖南省教育厅科学研究重点项目 (15A079), 湖南省高校科技创新团队支持计划资助

Supported by National Natural Science Foundation of China (60975049, 61174132), Scientific Research Fund of Hunan Provincial Education Department (15A079), and Aid Program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province

本文责任编辑 乔俊飞

Recommended by Associate Editor QIAO Jun-Fei

1. 中南大学信息科学与工程学院 长沙 410083 2. 湖南理工学院信息与通信工程学院 岳阳 414006

1. College of Information Science and Engineering, Central South University, Changsha 410083 2. College of Information and Communication Engineering, Hunan Institute of Science and Technology, Yueyang 414006

优化 Pareto 支配性预测的思路, 通过将 2 个 n 维候选解样本决策向量串接成 $2n$ 维的训练样本, 直接预测 Pareto 支配性, 从而减少对昂贵目标函数的评价次数^[10-11]. 随后, 他们在原有工作基础上提出基于等价分量交叉相似性的 Pareto 支配性预测方法, 通过在决策向量中引入对 Pareto 支配性贡献度相同的等价维, 进行交叉相似性最近邻搜索, 改善具有线性相关等价维特征的决策空间最近邻个体与目标空间最近邻个体的一致性, 提高 Pareto 支配性预测最近邻分类方法的准确率^[12]. 受此启发, 陈志旺等利用最近邻方法实现 Pareto 支配性预测和可行解预测, 并基于决策空间的流形理论, 利用 K-均值聚类 PCA (Principle component analysis) 提出决策空间拥挤距离的计算方法, 改进求解昂贵区间多目标优化问题的 NSGA-II (Non-dominated sorting genetic algorithm II) 算法, 进一步验证了最近邻方法进行 Pareto 支配性预测的有效性^[13]. 然而, 以上研究主要针对原始决策空间^[13] 以及具有线性相关等价维特征决策空间^[12] 中的 Pareto 支配性最近邻预测, 对于具有非线性相关特征的昂贵多目标优化问题, 难以真实度量候选解样本间的距离, 极易出现决策空间最近邻个体与目标空间最近邻个体一致性较差的问题.

为使决策空间最近邻个体与目标空间最近邻个体尽可能一致, 进一步提高 Pareto 支配性预测最近邻分类方法的准确性, 本文研究决策空间不同属性变化对目标空间分类正确率的影响, 利用代理模型方法在离散数据上的拟合能力, 提出属性变化趋势模型的构造方法, 建立一种低计算成本的属性趋势模型. 通过属性趋势模型, 对多目标问题的决策空间进行变换, 让决策空间数据能承载更多的目标空间信息, 提高决策空间与目标空间邻域关系的一致性.

本文余下内容组织如下: 第 1 节简要介绍文章涉及的基本概念. 第 2 节阐述决策属性趋势模型和决策空间变换方法. 第 3 节针对 MO 仿真实例问题进行决策空间变换, 分析决策空间变换的可行性. 第 4 节阐述 Pareto 支配性的预测最近邻算法, 并与现存算法的 Pareto 支配性预测进行对比实验. 最后, 在第 5 节给出全文结论.

1 相关概念

1.1 Pareto 支配性

一般地, 具有 n 维决策变量, m 个目标函数的最小化无约束多目标优化问题, 可以表述为:

$$\begin{aligned} \min \quad & \mathbf{y} = \mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})] \\ \text{s.t.} \quad & \mathbf{x} = [x_1, x_2, \dots, x_n] \in \mathbf{X} \\ & \mathbf{y} = [y_1, y_2, \dots, y_m] \in \mathbf{Y} \end{aligned} \quad (1)$$

其中, $\mathbf{X}$ 为 n 维决策空间, $\mathbf{Y}$ 为 m 维目标空间,

$\mathbf{F}$ 为将决策空间映射到目标空间的一组映射关系, 相关定义如下:

定义 1 (Pareto 支配性). 对任意向量 $\mathbf{u} = [u_1, u_2, \dots, u_m] \in \mathbf{Y}$, $\mathbf{v} = [v_1, v_2, \dots, v_m] \in \mathbf{Y}$, 当且仅当 $\forall i \in \{1, 2, \dots, m\}: u_i \leq v_i \wedge \exists j \in \{1, 2, \dots, m\}: u_j < v_j$ 为真时, 称 $\mathbf{u}$ 支配 $\mathbf{v}$, 记作 $\mathbf{u} \succ \mathbf{v}$; 或称 $\mathbf{v}$ 受支配于 $\mathbf{u}$, 记作 $\mathbf{v} \prec \mathbf{u}$; 否则, 称 $\mathbf{u}$ 与 $\mathbf{v}$ 不可比, 记作 $\mathbf{u} \sim \mathbf{v}$.

定义 2 (Pareto 最优解). $\mathbf{x} \in \mathbf{X}$ 称为 Pareto 最优解 (或 Pareto 非支配解、非劣解), 当且仅当 $\neg \exists \mathbf{x}' \in \mathbf{X}, \mathbf{v} = \mathbf{F}(\mathbf{x}') \succ \mathbf{u} = \mathbf{F}(\mathbf{x})$.

定义 3 (Pareto 最优集). 决策空间中所有最优解的集合称为 Pareto 最优集, 目标空间中与之对应的目标向量集称为 Pareto 前沿或 Pareto 最优面.

1.2 决策属性与目标属性的相关系数

给定 MO 问题 (m 个目标函数, n 维决策变量) 的 N 个训练样本 $(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N)$, 构建样本集的决策矩阵 $\mathbf{X} (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N \in \mathbf{R}^n)$ 及目标矩阵 $\mathbf{Y} (\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^N \in \mathbf{R}^m)$, 设 $\mathbf{M} = [\mathbf{X} \ \mathbf{Y}]$, 则

$$\mathbf{M} = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_n^1 & y_1^1 & \cdots & y_m^1 \\ x_1^2 & x_2^2 & \cdots & x_n^2 & y_1^2 & \cdots & y_m^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_1^N & x_2^N & \cdots & x_n^N & y_1^N & \cdots & y_m^N \end{bmatrix} \quad (2)$$

矩阵 $\mathbf{M}$ 中的前 n 列中, 每一元素表示决策空间向量的一个决策属性 (x_i), 后 m 列的每一元素表示多目标优化问题中的一个目标属性 (y_j). 矩阵中决策属性与目标属性的相关性可通过计算相关系数得到^[14], 如式 (3).

$$r_{x_i, y_j} = \frac{\text{cov}(x_i, y_j)}{\sigma_{x_i} \sigma_{y_j}} = \frac{\mathbf{E}\{(x_i - \bar{x}_i)(y_j - \bar{y}_j)\}}{\sigma_{x_i} \sigma_{y_j}} \quad (3)$$

其中, $\text{cov}(x_i, y_j)$ 为参数 x_i 与 y_j 的协方差; $\sigma_{x_i}, \sigma_{y_j}$ 为属性的标准差; $\bar{x}_i, \bar{y}_j$ 为样本均值; $\mathbf{E}(\cdot)$ 表示数学期望. 计算矩阵 $\mathbf{M}$ 中的前 n 列与后 m 列的相关系数, 得到决策属性与目标属性的相关系数矩阵 $\mathbf{R}$, 表示为

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix} \quad (4)$$

式中 r_{ij} 表示第 i 个决策属性与第 j 个目标属性的相关性. r_{ij} 的性质如下:

- 1) $r_{ij} \in [-1, 1]$;
- 2) $|r_{ij}| = 0$, 表示 x_i 属性关于第 j 个目标函数是非线性相关或者无关;

3) $|r_{ij}| \neq 0$, 表示 x_i 属性关于第 j 个目标函数的线性相关程度。

2 决策空间变换

最近邻分类方法是按照某种度量准则搜索待分类个体在已知空间的最近邻。在有限样本集的决策空间中, 直接搜索待测样本的最近邻个体, 逼近待测样本的目标函数值, 存在决策空间最近邻与目标空间最近邻不一致的问题 (如图 1)。因此, 为使待测样本在决策空间与目标空间最近邻个体尽可能一致, 则需要根据这个目标函数的映射特征对决策空间进行空间变换。

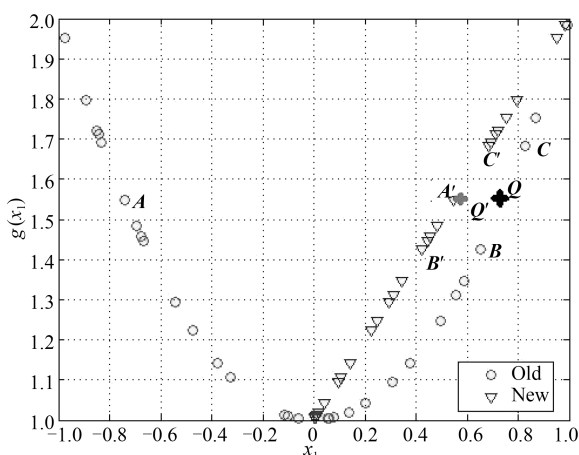


图 1 决策空间的线性化过程 (圆点表示原样本点的分布, 三角表示经过趋势模型 $y = g(x_i)$ 映射后的样本点)

Fig. 1 The linearization process of decision space (Circles are the original sample points and triangles are the mapped sample points by the trend model $y = g(x_i)$.)

2.1 决策属性的趋势模型

在最近邻分类方法中, 待分类个体可以看作是它最近邻个体在系统平衡点的线性增量问题^[15]。若建立非线性属性的趋势模型, 则可有效解决在逼近待测样本目标函数值时, 决策空间与目标空间最近邻不一致的问题。本节给出一种非线性相关属性的趋势模型构造方法, 步骤如下:

步骤 1. 计算样本数据决策属性与目标属性的相关系数;

步骤 2. 对目标函数 $f(x)$, 选择与函数 $f(x)$ 非线性相关的属性 x_i 作为可增量的变量, 并把其余属性看成常量;

步骤 3. 改变 x_i 的值 (采用样本集中已有的不同 x_i 值, 将相近的 x_i 值看作是彼此的线性增量), 观察 $f(x)$ 目标函数值的变化趋势;

步骤 4. 应用最小二乘法进行曲线拟合, 建立 x_i 属性的趋势模型 $y = g(x_i)$ 。

显然, 对决策空间与目标函数的非线性相关属性 x_i , 采用 $\mathbf{x} = (x_1, \dots, x_i, \dots, x_n) \rightarrow \mathbf{x}' =$

$(x_1, \dots, g(x_i), \dots, x_n)$ 的方式进行趋势拟合, 将决策空间进行变换, 这相当于对非线性相关属性 x_i 进行了线性化处理 (如图 1), 将样本决策空间进行了变换, 使决策空间的最近邻更加真实地反映目标空间的最近邻。

在图 1 中, A 、 B 、 C 为原决策空间的三个样本, A' 、 B' 、 C' 分别为采用趋势模型映射后的样本, Q 为待查询点, Q' 为映射后的查询点。若直接在原决策空间中查询 Q 的最近邻, 则其最近邻为样本 B 或 C 。但将 Q 映射为 Q' 后, 查询到的最近邻为 A' , 即原空间中的 A 。

2.2 空间变换

由于决策空间可能是高维的, 考虑建模的复杂性, 不可能对目标函数的每一维属性都建立趋势模型。因此, 根据第 1.2 节相关系数计算方法, 计算决策空间属性与目标函数 $f(\mathbf{x})$ 的相关系数, 得到目标函数 $f(\mathbf{x})$ 的决策属性相关系数向量 $\mathbf{r} = (r(x_1, f(\mathbf{x})), r(x_2, f(\mathbf{x})), \dots, r(x_n, f(\mathbf{x})))$ 。然后选择与函数 $f(\mathbf{x})$ 的相关系数近似于 0 (非线性相关或无关) 以及相等的决策属性进行建模。

因此, 决策空间的变换方法为:

1) 若选择相关系数近似相等的决策属性, 选取任意一个样本 $\mathbf{x}$, 其中属性 $x_i \neq x_j$, 交换 x_i 与 x_j , 比较目标值, 做可交换验证;

2) 若选择的属性变化趋势是近似线性的, 则不建立模型;

3) 若 x_i 增量变化时, $f(\mathbf{x})$ 不变, 则认为属性 x_i 与 $f(\mathbf{x})$ 是无关的, 则去除该属性, 将决策空间进行降维变换;

4) 若存在变化趋势一致的 x_i 和 x_j 属性, 且 x_i 与 x_j 是可交换的, 则建立相同的属性趋势模型 $g(\mathbf{x})$, 对这些属性应用 $g(\mathbf{x})$ 求值并排序, 形成新的样本空间;

5) 若 x_i 是一非线性相关属性, 直接建立属性趋势模型 $g(\mathbf{x})$ 并求值, 形成新的样本空间。

3 多目标决策空间变换的最近邻方法

为验证决策空间变换最近邻方法的有效性, 选取 ZDT1、ZDT3、ZDT6、KUR、UF1、DTLZ2 和 DTLZ4 这 7 个测试函数^[16-19] 进行实验。所用实验平台为: Intel(R) i3 CPU 3.3 GHz, RAM 4 GB, 所有程序代码均在 Win7 系统下的 Matlab 2011a 上运行。

3.1 决策属性与目标函数的相关系数计算

对每一个测试函数, 在定义范围内产生均匀分布的样本集。设二目标优化问题训练样本集与测试样本集的大小 $N = 200$, 三目标优化问题训练样本集与测试集样本的大小 $N = 1000$ 。分别对每个测

试函数进行 100 次实验, 每次实验都重新生成训练样本集. 测试函数的相关系数如表 1 所示.

3.2 基于趋势模型的决策空间变换

由表 1 数据可知, 在这些测试函数的决策属性与目标属性相关系数向量中, 存在线性相关、非线性相关和无关等情况, 按照第 2.2 节方法对这些测试问题的子目标函数进行决策空间变换, 过程如下.

对 ZDT1 和 ZDT3 的子目标函数 $f_1(\mathbf{x})$, 属性 x_1 的相关系数为 1, 其余为 0 且是无关的, 因此可将子目标函数 $f_1(\mathbf{x})$ 的原 n 维决策空间降至 1 维, 实现子目标函数 $f_1(\mathbf{x})$ 的决策空间变化. 对 ZDT1 和 ZDT3 的子目标函数 $f_2(\mathbf{x})$, 除属性 x_1 外, 其余属性的相关系数相等, 观察 $x_i (i \neq 1)$ 变化引起的子目标函数 $f_2(\mathbf{x})$ 值的变化, 发现变化趋势相同且线性, 所以不构造 ZDT1 和 ZDT3 的子目标函数 $f_2(\mathbf{x})$ 关于属性 $x_i (i \neq 1)$ 的趋势模型, 但属性 $x_i (i \neq 1)$ 是可交换的, 因此对属性 $x_i (i = 2, 3, \dots, n)$ 进行排序, 形成新样本空间, 避免在使用最近邻模型距离计算过程中因相同趋势属性之间数据相差太大产生的计算误差.

对 ZDT6 的子目标函数 $f_1(\mathbf{x})$, 属性 $x_i (i = 2, 3, \dots, n)$ 的相关系数近似 0 且是无关的, 因此可将子目标函数 $f_1(\mathbf{x})$ 的原 n 维决策空间降至 1 维, 实现子目标函数 $f_1(\mathbf{x})$ 的决策空间变化. 对 ZDT6 的子目标函数 $f_2(\mathbf{x})$, 除属性 x_1 外, 其余属性的相关系数是相等的, 观察 $x_i (i = 2, 3, \dots, n)$ 变化引起的子目标函数 $f_2(\mathbf{x})$ 值的变化, 发现变化趋势相同且为 $g(\mathbf{x}) = \sqrt{24x}$, 且属性 $x_i (i \neq 1)$ 是可交换的, 因此对属性 $x_i (i = 2, 3, \dots, n)$ 进行变换并排序, 形成新的样本空间.

对 KUR 的子目标函数 $f_1(\mathbf{x})$, 属性 $x_i (i = 1, 3)$ 的趋势模型相同且为 $g(\mathbf{x}) = |\mathbf{x}|$, 属性 $x_i (i = 1, 3)$ 是可交换的, 对属性 $x_i (i = 1, 3)$ 进行变换并排序, 形成新的样本空间. 对 KUR 的子目标函数 $f_2(\mathbf{x})$,

存在相关系数相同的 $x_i (i = 1, 2, 3)$ 属性, 发现变化趋势相同且为 $g(\mathbf{x}) = 5 \sin(\mathbf{x}^3)$, 属性 $x_i (i = 1, 2, 3)$ 是可交换的, 同样对 $x_i (i = 1, 2, 3)$ 属性进行变换并排序, 形成新的样本空间.

对 UF1 的子目标函数 $f_1(\mathbf{x})$, 由于偶数维属性的相关系数近似为 0, 经过分析是无关的, 因此偶数维属性在最近邻模型中不考虑, 直接去除偶数维; 对于除第 1 维外的奇数维属性的相关系数相同, 观察其趋势变化也相同, 但建模复杂, 因此不选择构造趋势函数, 仅对除第 1 维外的奇数维属性进行排序, 形成新的样本空间. 对 UF1 的子目标函数 $f_2(\mathbf{x})$, 由于除第 1 维外的奇数维相关系数近似为 0, 经过分析是无关的, 因此奇数维属性在最近邻模型中不考虑, 直接去除; 而偶数维属性的相关系数基本相同, 观察其趋势变化也相同, 但建模复杂, 因此不选择构造趋势函数, 仅对偶数维属性进行排序, 形成新的样本空间.

对 DTLZ2 和 DTLZ4 的子目标函数 $f_1(\mathbf{x})$, 分为两部分: 属性 x_1 和 x_2 的相关系数相同, 且对子目标函数 $f_1(\mathbf{x})$ 的影响是线性的, 因此不构造趋势模型; 属性 x_3 到 x_n 的相关系数相同, 且变化趋势是关于 0.5 对称的, 因此趋势模型选取为 $g(\mathbf{x}) = 1 - \mathbf{x} (\mathbf{x} > 0.5)$. 对 DTLZ2 和 DTLZ4 的子目标函数 $f_2(\mathbf{x})$, 决策空间的处理方法与子目标函数 $f_1(\mathbf{x})$ 相同. 对 DTLZ2 和 DTLZ4 的子目标函数 $f_3(\mathbf{x})$, 属性 x_2 相关系数为 0, 经过分析是无关的, 则在最近邻模型中不考虑; 属性 x_3 到 x_n 的处理方法与子目标函数 $f_1(\mathbf{x})$ 相同.

由测试函数决策空间变换可知, 对于 MO 问题, 原始决策空间将根据不同子目标函数变为对应的新决策空间.

3.3 决策空间变换的最近邻方法分析

在 MO 问题中, 采用搜索待分类个体在决策空间有限样本集中的最近邻个体, 预测待分类个体的

表 1 测试函数的决策属性与目标属性的相关系数

Table 1 Correlation coefficient of decision attribute and objective attribute of test function

测试函数		$f_1(\mathbf{x})$	$f_2(\mathbf{x})$	$f_3(\mathbf{x})$
ZDT1	$n = 10$	(1, 0, 0, 0, ...)	(-0.59, 0.27, 0.27, 0.27, ...)	
	$n = 30$	(1, 0, 0, 0, ...)	(-0.79, 0.11, 0.11, 0.11, ...)	
ZDT3	$n = 3$	(1, 0, 0)	(-0.28, 0.65, 0.65)	
	$n = 10$	(1, 0, 0, 0, ...)	(-0.49, 0.25, 0.25, ...)	
ZDT6	$n = 3$	(0.47, 0, 0)	(0.02, 0.69, 0.69)	
	$n = 10$	(0.48, 0, 0, 0, ...)	(0.02, 0.69, 0.69)	
KUR	$n = 3$	(0.02, 0, -0.01)	(0.02, 0.02, 0.02)	
UF1	$n = 15$	(0.92, 0, -0.07, 0, -0.07, 0, ...)	(-0.51, -0.11, 0, -0.12, ...)	
DTLZ2	$n = 10$	(-0.64, -0.64, 0.01, -0.01, ...)	(-0.64, 0.64, 0.01, 0.01, ...)	(0.94, 0, 0, 0.01, 0, ...)
DTLZ4	$n = 10$	(-0.64, -0.64, 0.01, -0.01, ...)	(-0.64, 0.64, 0.01, 0, -0.01, ...)	(0.94, -0.01, 0, 0.01, 0, ...)

目标函数值, 存在目标空间中子目标函数不能同时最近邻的问题. 因此, 对 MO 问题的各子目标函数, 根据决策属性的趋势模型进行空间变换. 对 7 个存在不同类型相关系数的 MO 测试问题进行实验, 选择 100 个待测样本, 比较决策空间变换前后在训练样本集中搜索到待测样本最近邻个体在子目标函数值上的平均误差和最大误差.

由表 2 数据可知, 直接去除与子目标函数无关的属性, 相当于对原始决策空间进行降维来实现决策空间的变换, 可减少最近邻方法中的距离计算, 增加最近邻方法预测的准确性. 如 ZDT1、ZDT3 和 ZDT6 的子目标函数 $f_1(\mathbf{x})$ 与 UF1 的子目标函数 $f_1(\mathbf{x})$ 和 $f_2(\mathbf{x})$. 对于与子目标函数非线性相关的属性, 进行趋势拟合, 并对趋势一致的属性进行拟合后排序, 形成新的样本空间, 避免在使用最近邻模型距离计算过程中因相似属性之间数据相差太大产生的计算误差, 增加最近邻方法预测的准确性. 如 ZDT1、ZDT3 和 ZDT6 的子目标函数 $f_2(\mathbf{x})$ 与 KUR、DTLZ2 和 DTLZ4 的所有子目标函数. 因此, 对不同测试函数, 找到其子目标函数的最近邻个体, 子目标函数值的误差均值明显要优于直接在原始决策空间上找到的最近邻个体时的误差均值, 子目标函数值的最大误差也明显减小, 说明候选解个体与最近邻个体的距离更近了. 所以, 该方法能有效改善 MO 问题决策空间的近邻个体不能代表子函数目标空间近邻个体的问题, 可增加最近邻预测方法的准确性.

4 多目标 Pareto 支配性预测实验

4.1 最近邻模型的 Pareto 支配性预测算法

本节主要阐述基于决策空间变换最近邻方法的 Pareto 支配性预测算法, 用两个候选解在决策空间的信息来预测它们在目标空间上的支配关系. 首先, 给定昂贵 MO 问题 (m 个目标函数, n 维决策变量) 的 N 个训练样本 $(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N)$, 计算出子目标函数的相关系数向量 $\mathbf{r}_k$. 然后, 对样本决策空间, 根据第 3.2 节方法, 构造趋势模型进行空间变换. 其次, 对候选解样本根据趋势模型进行相应变换, 并在对应的决策空间中采用最近邻方法搜索最近邻个体. 这样将会从变换后的 m 个决策空间中找到 m 个最近邻个体 $(\mathbf{x}^{ik}|k = 1, \dots, m)$. 对于候选解样本 $\mathbf{x}^i$ 在变换后决策空间中搜索到的每一个最近邻都是对应子目标函数的最近邻 $(\min_{f_j} |f_j(\mathbf{x}^i) - f_j(\mathbf{x})|, j = 1, \dots, m)$. 最后, 比较两个候选解对应子目标函数最近邻个体的目标值进行 Pareto 支配性预测.

因此, 对于两个候选解 $\mathbf{x}^i$ 与 $\mathbf{x}^j$, 其 Pareto 支配性预测算法如算法 1.

算法 1. m 个目标的 MO 问题

- 步骤 1. 生成大小为 N 的训练样本, 计算决策属性与目标属性的相关系数矩阵 R ;
- 步骤 2. 对相关系数矩阵 R 中的 0 分量和相等分量, 在原始训练样本决策空间采用第 3.2 节中的空间变换方法,

表 2 变换前后的最近邻分类误差
Table 2 Nearest neighbor classification error before and after transformation

测试函数		按目标分类的误差均值			按目标分类的最大误差		
		$f_1(\mathbf{x})$	$f_2(\mathbf{x})$	$f_3(\mathbf{x})$	$f_1(\mathbf{x})$	$f_2(\mathbf{x})$	$f_3(\mathbf{x})$
ZDT1	原空间	0.1516	0.4302		0.5115	1.3159	
	变换后	0.0013	0.1725		0.0074	0.9304	
ZDT3	原空间	0.0437	0.4460		0.1837	1.7126	
	变换后	0.0011	0.3282		0.0042	1.5289	
ZDT6	原空间	0.1047	0.1873		0.6864	1.8303	
	变换后	0.0031	0.0765		0.0454	0.4374	
KUR	原空间	0.6601	5.7885		2.2948	23.4230	
	变换后	0.2909	1.6843		1.6411	4.6841	
UF1	原空间	0.1985	0.1753		0.8847	1.0659	
	变换后	0.0753	0.0680		0.2736	0.4554	
DTLZ2	原空间	0.2314	0.2368	0.2344	1.0053	0.8732	1.0964
	变换后	0.0624	0.0569	0.0471	0.3102	0.1942	0.1990
DTLZ4	原空间	0.2072	0.2431	0.2459	0.8283	0.7537	0.8950
	变换后	0.0582	0.0724	0.0488	0.2478	0.2178	0.0488

生成 m 个子目标函数对应的新决策空间;

步骤 3. 对待测候选解样本 $\mathbf{x}^i$ 与 $\mathbf{x}^j$, 首先按子目标函数进行空间变换, 然后采用最近邻方法, 找到对应训练样本新决策空间中的 m 个最近邻个体 ($\mathbf{x}^{ik}|k=1, \dots, m$) 与 ($\mathbf{x}^{jk}|k=1, \dots, m$);

步骤 4. 参照训练样本的目标空间, 分别将 $\mathbf{x}^i$ 与 $\mathbf{x}^j$ 的 m 个最近邻个体对应的目标值作为它们的子函数目标值;

步骤 5. 比较 $\mathbf{x}^i$ 与 $\mathbf{x}^j$ 的各子函数目标值, 判别 Pareto 支配性.

在步骤 1 中, 在计算决策空间属性 (n 维) 与目标属性 (m 维) 的相关系数矩阵时, 计算每一个决策属性与目标属性协方差与标准差的计算复杂度均为 $O(N)$, 所以计算相关系数矩阵的时间复杂度为 $O(nmN)$. 步骤 2 中, 设有 k 个决策属性 ($k \leq n$) 与子目标函数 $f_j(\mathbf{x})$ 是非线性相关且不等价的, 则需建立 k 个趋势函数. 由于属性趋势函数的建立是采用最小二乘方法, 时间复杂度为 $O(N^2)$, 所以完成 m 个分目标函数空间变换的时间复杂度为 $O(kmN^2)$. 步骤 3 中, 对决策空间运用最近邻分类方法对待测候选解样本进行预测的时间复杂度为 $O(mN^2)$. 步骤 4 和步骤 5 的时间复杂度为 $O(m)$. 因此, 两个候选解 $\mathbf{x}^i$ 与 $\mathbf{x}^j$ 的 Pareto 支配性预测的时间复杂度为 $\max\{O(nmN), O(kmN^2)\}$, 这说明最近邻分类 Pareto 支配性预测方法的时间复杂度是可接受的.

4.2 实验研究

选取与第 3 节相同的实验环境和测试函数, 将决策空间变换的最近邻方法 (Nearest neighbor method based on decision space transformation, DSTNN) 与简单欧氏距离相似性测度的最近邻方法 (Nearest neighbor classification based on Euclidean distance measurement, ENNC)^[11]、等价分量最小交叉距离加权求和的最近邻方法 (Nearest neighbor classification based on equivalent similarity, ESNNC)^[12] 和采用高斯过程 (Gaussian process, GP) 模型^[5] 对每一个目标函数分别建模, 用模型的输出值近似真实目标值的方法进行 Pareto 支配性预测比较. 对二目标优化问题设计训练样本集与测试集样本的大小 $N = 200$, 三目标优化问题样本集与测试集样本的大小 $N = 1000$. 分别对每个测试问题进行 100 次实验, 每次实验都重新生成训练样本集, 测试样本集不变, 则 100 次 Pareto 支配性预测实验的 Pareto 支配性的正确性评价如表 3 所示. 参考数值主要有决策空间维数 n 和支配性平均预测精度.

由表 3 可知, 1) DSTNN 算法同其他算法相比, 预测的准确性提高幅度较大, 尤其是对于 MO 优化过程中起重要作用的非支配类, 所以该算法将在 MO 问题的进化优化中具有较大优势; 2) DSTNN 算法在样本空间类别分布不平衡的

DTLZ2、DTLZ4 和 UF1 测试函数上都显示出较好的正确性; 3) Pareto 支配性预测的正确率跟决策空间的维数是负相关的, 这从 ZDT1、ZDT3 和 ZDT6 等函数的维数变化可知, 随着维数的降低, 支

表 3 Pareto 支配性预测的平均精度
Table 3 Average precision of Pareto dominance prediction

测试函数	方法	n	Pareto 支配性 (%)			
			总体正确	非支配类	支配类	不可比类
ZDT1	DSTNN	10	92.32	81.70	88.61	96.64
	ENNC		64.79	40.58	42.30	74.88
	ESNNC		83.23	70.32	75.24	88.31
	GP	30	88.76	65.55	65.55	90.21
	DSTNN		90.85	71.20	81.51	95.20
	ENNC		54.32	17.36	23.26	63.41
	ESNNC		68.90	37.49	40.28	80.54
	GP		73.41	32.97	32.97	84.49
	DSTNN	3	93.13	80.47	93.14	97.19
ZDT3	ENNC		76.84	75.54	69.03	79.20
	ESNNC		68.82	56.95	68.85	72.26
	GP	10	57.73	54.60	54.60	59.48
	DSTNN		83.29	71.53	77.56	88.69
	ENNC		59.33	34.81	38.61	70.03
	ESNNC	3	77.86	72.05	66.85	84.12
	GP		71.10	54.20	54.20	78.56
	DSTNN		95.94	89.83	95.59	99.36
	ENNC	3	53.68	42.23	48.88	55.90
ZDT6	ESNNC		49.78	43.06	49.49	52.97
	GP	10	52.96	36.92	36.92	66.89
	DSTNN		91.43	87.77	93.58	92.33
	ENNC	10	44.30	37.80	47.55	46.10
	ESNNC		64.86	66.12	60.53	66.44
	GP		48.88	47.03	47.03	50.82
	DSTNN	10	86.49	43.89	66.87	89.66
	ENNC		54.89	17.24	17.42	58.87
	ESNNC		90.92	22.68	28.41	95.31
DTLZ2	GP	10	89.50	12.00	12.00	96.75
	DSTNN		85.39	50.14	74.64	88.55
	ENNC	10	53.74	16.93	20.42	57.23
	ESNNC		72.23	48.38	44.56	79.99
	GP		70.87	9.51	9.51	76.65
	DSTNN	10	86.14	71.90	84.17	91.64
	ENNC		60.08	54.86	50.23	65.76
	ESNNC		60.84	65.23	70.69	71.49
	GP	3	57.70	61.53	61.53	51.61
UF1	DSTNN		85.43	79.73	86.79	89.87
	ENNC	3	46.47	53.95	56.71	34.54
	ESNNC		50.73	52.88	55.39	45.02
	GP		44.19	41.79	41.79	47.76

配性识别的准确率会有一定程度的提高; 4) DSTNN 方法没有考虑决策空间降维对 Pareto 支配性预测的影响, 但在实际空间变换过程中, 对无关维的约减即隐含对决策空间降维, ZDT1、ZDT3、ZDT6 和 UF1 测试问题的实验结果表明降维有利于提高正确性. 该方法对于昂贵 MO 问题, 虽然在构造属性趋势模型时增加了昂贵函数额外的计算成本, 但当计算完成后, 则不再需要对原昂贵函数进行评估, 很大程度上减少了昂贵函数的评价次数.

5 结论

昂贵 MO 问题在采用进化算法求解时需要消耗时复杂的目标函数多次评估, 面临适应值评估难的问题, 研究表明采用最近邻分类方法预测候选解间的 Pareto 支配性, 可有效克服计算成本问题. 但现有的最近邻方法在搜索多目标问题的决策空间时, 存在目标空间中子目标函数不能同时最近邻的问题. 本文通过对子目标的决策属性进行分析, 提出基于决策属性趋势模型的空间变换方法, 在决策空间搜索候选解样本的最近邻个体, 预测候选解之间的 Pareto 支配关系. 该代理模型并不要求高精度、易于构造, 使得变换后的决策空间承载了更多的目标空间信息. 结合趋势模型空间变换的最近邻方法进行昂贵 MO 问题的 Pareto 支配性预测, 避免复杂耗时的目标函数评估. 实验结果表明, 该方法显著优于现有的最近邻分类算法和典型代理模型的 Pareto 支配性预测结果.

对于 Pareto 支配性预测方法与优化算法的交互, 我们已开展了相关应用研究, 并取得初步效果. 进一步工作将探索该方法在蛋白质、基因等实际昂贵多目标优化问题中的应用.

致谢

在此, 我们向对本文工作给予支持和建议的同行, 尤其是湖南理工学院信息与通信工程学院的潘理副教授、杨勃副教授等表示感谢.

References

- Gong Mao-Guo, Jiao Li-Cheng, Yang Dong-Dong, Ma Wen-Ping. Research on evolutionary multi-objective optimization algorithms. *Journal of Software*, 2009, **20**(2): 271–289
(公茂果, 焦李成, 杨咚咚, 马文萍. 进化多目标优化算法研究. 软件学报, 2009, **20**(2): 271–289)
- Zhan Wei. Research on Manifold Learning Algorithm for High-Dimensional Multi-objective Optimization Problems [Ph. D. dissertation], China University of Geosciences, China, 2013.
(詹伟. 求解高维多目标优化问题的流形学习算法研究 [博士学位论文], 中国地质大学, 中国, 2013.)
- Tesch M, Schneider J, Choset H. Expensive multiobjective optimization for robotics. In: Proceedings of the 2013 IEEE International Conference on Robotics and Automation. Karlsruhe, Germany: IEEE, 2013. 973–980
- Douguet D. e-LEA3D: a computational-aided drug design web server. *Nucleic Acids Research*, 2010, **38**(Suppl 2): W615–W621
- Zhang Q F, Liu W D, Tsang E, Virginas B. Expensive multi-objective optimization by MOEA/D with Gaussian process model. *IEEE Transactions on Evolutionary Computation*, 2010, **14**(3): 456–474
- Palar P S, Tsuchiya T, Parks G T. A comparative study of local search within a surrogate-assisted multi-objective memetic algorithm framework for expensive problems. *Applied Soft Computing Journal*, 2016, **43**: 1–19
- Rosales-Pérez A, Coello C A C, Gonzalez J A, Reyes-Garcia C A, Escalante H J. A hybrid surrogate-based approach for evolutionary multi-objective optimization. In: Proceedings of the 2013 IEEE Congress on Evolutionary Computation (CEC). Cancun, Mexico: IEEE, 2013. 2548–2555
- Akhtar T, Shoemaker C A. Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection. *Journal of Global Optimization*, 2016, **64**(1): 17–32
- Montemayor-García G, Toscano-Pulido G. A study of surrogate models for their use in multiobjective evolutionary algorithms. In: Proceedings of the 8th International Conference on Electrical Engineering Computing Science and Automatic Control (CCE). Merida, Mexico: IEEE, 2011. 1–6
- Guo G Q, Li W, Yang B, Li W B, Yin C. Predicting Pareto dominance in multi-objective optimization using pattern recognition. In: Proceedings of the 2nd International Conference on Intelligent System Design and Engineering Application (ISDEA). Sanya, China: IEEE, 2012. 456–459
- Guo G Q, Yin C, Yan T S, Li W. Nearest neighbor classification of Pareto dominance in multi-objective optimization. In: Proceedings of the 5th International Conference on Advanced Computational Intelligence (ICACI). Nanjing, China: IEEE, 2012. 328–331
- Guo Guan-Qi, Yin-Cheng, Zeng Wen-Jing, Li Wu, Yan Tai-Shan. Prediction of Pareto dominance by cross similarity of equivalent components. *Acta Automatica Sinica*, 2014, **40**(1): 33–40
(郭观七, 尹呈, 曾文静, 李武, 严太山. 基于等价分量交叉相似性的 Pareto 支配性预测. 自动化学报, 2014, **40**(1): 33–40)

- 13 Chen Zhi-Wang, Bai Xin, Yang Qi, Huang Xing-Wang, Li Guo-Qiang. Strategy of constraint, dominance and screening solutions with same sequence in decision space for interval multi-objective optimization. *Acta Automatica Sinica*, 2015, **41**(12): 2115–2124

(陈志旺, 白锌, 杨七, 黄兴旺, 李国强. 区间多目标优化中决策空间约束、支配及同序解筛选策略. 自动化学报, 2015, **41**(12): 2115–2124)

- 14 Sheng Zhou, Xie Shi-Qian, Pan Cheng-Yi. *Probability and Mathematical Statistics*. Beijing: Higher Education Press, 2005.

(盛骤, 谢式千, 潘承毅. 概率论与数理统计. 北京: 高等教育出版, 2005.)

- 15 Tian Zuo-Hua, Chen Xue-Zhong, Weng Zheng-Xin. *Control Engineering Construction*. Beijing: Tsinghua University Press, 2007.

(田作华, 陈学中, 翁正新. 工程控制基础. 北京: 清华大学出版社, 2007.)

- 16 Zitzler E, Deb K, Thiele L. Comparison of multiobjective evolutionary algorithms: empirical results. *Evolutionary Computation*, 2000, **8**(2): 173–195

- 17 Kursawe F. A variant of evolution strategies for vector optimization. *Parallel Problem Solving from Nature*. Berlin: Springer, 1991.

- 18 Deb K, Thiele L, Laumanns M, Zitzler E. Scalable multi-objective optimization test problems. In: Proceedings of the 2002 IEEE Congress on Evolutionary Computation (CEC). Honolulu, USA: IEEE, 2002. 825–830

- 19 Li H, Zhang Q F. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2009, **13**(2): 284–302



李文彬 中南大学信息科学与工程学院博士研究生. 主要研究方向为进化计算, 多目标优化.

E-mail: wenbin.lii@163.com

(**LI Wen-Bin** Ph.D. candidate at the Institute of Information Science and Engineering, Central South University. His research interest covers evolutionary computation and multi-objective optimization.)



贺建军 中南大学教授. 主要研究方向为复杂工业过程建模与优化控制.

E-mail: jjhe@mail.csu.edu.cn

(**HE Jian-Jun** Professor at Central South University. His research interest covers modeling and optimal control of complex industrial process.)



冯彩英 湖南理工学院信息与通信工程学院硕士研究生. 主要研究方向为计算智能, 多目标优化.

E-mail: fengcy0829@163.com

(**FENG Cai-Ying** Master student at the Institute of Information and Communication Engineering, Hunan Institute of Science and Technology.

Her research interest covers computational intelligence and multi-objective optimization.)



郭观七 湖南理工学院教授. 主要研究方向为计算智能, 多目标优化, 模式识别. 本文通信作者.

E-mail: gq.guo@163.com

(**GUO Guan-Qi** Professor at Hunan Institute of Science and Technology. His research interest covers computational intelligence, multi-objective optimization, and pattern recognition. Corresponding Author of this paper.)