

高维多目标优化中基于稀疏特征选择的目标降维方法

陈小红, 李霞, 王娜

(1. 深圳大学信息工程学院, 广东深圳 518060; 2. 深圳市现代通信与信息处理重点实验室, 广东深圳 518060)

摘要: 目标降维算法通过去除冗余的目标达到简化问题规模的目的, 为求解高维多目标优化问题提供了一种新的思路和方法. 近似解集的几何结构特征和 Pareto 占优关系从不同侧面反映了多目标优化问题的内在结构特性, 而现有算法仅利用其中一种特征分析目标之间的关系, 具有较大局限性. 本文提出基于稀疏特征选择的目标降维方法, 该方法利用近似解集的几何结构特征构建稀疏回归模型, 求解高维目标空间映射为低维目标子空间的稀疏投影矩阵, 依据此矩阵度量目标的重要性, 并利用 Pareto 占优关系改变程度选择满足误差阈值的目标子集, 实现目标降维. 通过与其他已有目标降维算法比较, 实验结果表明本文提出的降维算法具有较高的准确性, 并且受近似解集质量的影响较小.

关键词: 高维多目标优化; 目标降维; 稀疏特征选择

中图分类号: TP181

文献标识码: A

文章编号: 0372-2112 (2015)07-1300-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.07.008

Objective Reduction with Sparse Feature Selection for Many Objective Optimization Problem

CHEN Xiao-hong, LI Xia, WANG Na

(1. College of Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China;

2. Shenzhen Key Lab of Advanced Communications and Information Processing, Shenzhen, Guangdong 518060, China)

Abstract: Objective reduction approach is an effective means for many-objective optimization problems by eliminating redundant objectives with respect to the original objective set. The geometrical structural characteristics and Pareto-dominance relation of approximation set can represent the characteristics of the original problem in different aspects. This paper proposed a new algorithm based on sparse feature selection. It used the geometrical structural characteristics to construct a graph representing the original problem. A sparse projection matrix mapping the high dimensional data into low dimensional space was then learned by a sparse regression model, which was used to measure the importance of each objective. The change of Pareto-dominance relation induced by reduced set was also adopted to identify a minimum set with error not exceeding threshold value. By comparing with other algorithms, the experimental results show that the accuracy of the new algorithm outperforms other dimension reduction techniques, and is scarcely effected by the quality of approximation set.

Key words: many-objective optimization; objective reduction; sparse feature selection

1 引言

近几年, 目标个数多于 4 个的高维多目标优化问题 (Many-Objective Optimization Problems)^[1]引起了学者广泛关注^[2~4]. 研究发现^[5,6], 经典的多目标进化算法 (Multi-Objective Evolutionary Algorithms, MOEAs), 特别是基于 Pareto 占优机制的算法, 虽然能有效地求解 2 或 3 个目标的问题, 但当目标个数增加时, 其收敛能力随之下降. 而且, 目标个数的增多也会带来可视化、决策选择以及

计算复杂度增加等困难^[7]. “维数灾难”是造成高维多目标优化问题求解困难的主要因素.

事实上, 现实世界中存在这样一类多目标优化问题, 其 Pareto 最优前沿仅由较少的几个目标组成, 称为冗余的多目标优化问题. 对于这类问题, 一种可行的解决方法是采用目标降维^[8]技术寻找最能表示原目标集合的子集, 简化问题规模. 目标降维为求解冗余的高维多目标优化问题提供了一种新的思路和方法^[9,10].

目标降维算法依据的样本集来自 MOEAs 求得的近

似解集,它在一定程度上反映了问题的内在结构特性.现有目标降维算法正是利用近似解集的特征分析目标的冲突性和冗余性,包括:(1)几何结构特征.文献[8]提出根据目标的相关系数度量目标之间的冲突程度,进而删除与其他目标最相关的目标.文献[11]对近似解集的谱进行主成分分析,寻找对主成分贡献最大的目标.(2)解集上的 Pareto 占优关系.文献[12]以 Pareto 占优关系改变误差衡量目标子集之间的冲突程度.

现有目标降维算法只利用近似解集的一种特性分析目标之间的关系,具有一定的局限性.因为 MOEAs 求得的近似解集通常与 Pareto 最优前沿有一定的差别,不能准确表征原问题内在结构特征,那么使占优关系不变的目标子集不一定是组成 Pareto 最优前沿的目标子集,组成低维近似解集的目标子集也不一定使其占优关系不变.显然,同时利用这两个特性将有利于提高目标降维的准确性.

因此,本文提出一种新的目标降维算法.该算法基于稀疏特征选择的思想,首先利用近似解集的几何结构特性,构建描述原问题内在结构特征的 Laplacian 谱,提出稀疏回归模型求解高维目标空间映射为低维目标子空间的稀疏投影矩阵,依据此矩阵度量目标的重要性.针对不同降维需求,提出利用 Pareto 占优关系改变误差求使原问题结构改变不超过阈值的最小目标子集,或者选择重要性较大的目标组成使原问题结构改变最小的目标子集.实验结果表明相对于其他目标降维算法,本文的算法具有较高的准确性,并且受近似解集质量的影响较小.

2 目标降维问题的数学模型

不失一般性,最小化多目标优化问题可以描述为:

$$\min \mathbf{F}(\mathbf{a}) = \{f_1(\mathbf{a}), f_2(\mathbf{a}), \dots, f_M(\mathbf{a})\} \quad (1)$$

其中, $\mathbf{a} = [a_1, a_2, \dots, a_n]^T \in Z^n$ 称为决策向量. $F = \{f_1, \dots, f_M\}$ 为目标集合, $\mathbf{F}(\mathbf{a})$ 是 \mathbf{a} 在目标空间的映射.当目标个数 $M > 3$ 时,称为高维多目标优化问题^[1].

决策向量 $\mathbf{a}, \mathbf{b} \in Z$, 如果 $\forall f_i \in F: f_i(\mathbf{a}) \leq f_i(\mathbf{b})$, 并且 $\exists f_j \in F: f_j(\mathbf{a}) < f_j(\mathbf{b})$, 则称 \mathbf{a} 支配 \mathbf{b} , 记作 $\mathbf{a} > \mathbf{b}$. 假设 $\mathbf{a}^* \in A \subseteq Z$ 不被集合 A 中任何个体支配, 则称 \mathbf{a}^* 是集合 A 中的非支配解, 如果 A 中所有解都是非支配的, 则称集合 A 为近似解集, 其在目标空间的映射为 $\mathbf{F}(A) = [f_1(A), \dots, f_M(A)]^T$, 本文中称 $f_i(A)$ 为目标向量. 如果 $\mathbf{z}^* \in Z$ 不被任何解支配, 则称其为 Pareto 最优解, 由所有 Pareto 最优解组成的集合称为 Pareto 最优解集, 该解集在目标空间的映射称为 Pareto 最优前沿.

对于目标个数为 M 的优化问题, 如果其 Pareto 最优前沿的维数小于 M , 称此类问题为冗余的多目标优

化问题. 组成 Pareto 最优前沿的目标称为关键目标, 其他目标称为冗余目标. 以近似解集为样本数据, 寻找关键目标的过程, 称为目标降维^[8]. 根据不同的降维需求, 目标降维问题可以描述为如下两个问题:

$$(P1): F^* = \arg \min_{F' \subset F} |F^*| \quad \text{s.t.} \quad \epsilon(F^*) \leq \epsilon_0 \quad (2)$$

$$(P2): F^* = \arg \min_{F' \subset F} \epsilon(F^*) \quad \text{s.t.} \quad |F^*| = q \quad (3)$$

其中, $\epsilon(F^*)$ 表示目标集合由 F_0 缩减为 F^* 后原问题结构改变的误差. 问题 P1 是寻找一个最小的目标子集 F^* , 使得误差不大于阈值 ϵ_0 . 问题 P2 是寻找目标数为 q 的子集, 并使其误差最小. 这两个问题都已被证明是 NP 难问题(证明过程详见文献[12]).

目标降维依据的样本数据集来自多目标优化算法求得的近似解集, 它从两个方面反映了原问题的内在结构特性. 一方面, 近似解集是对 Pareto 最优前沿的近似, 因此它在一定程度上反映了 Pareto 最优前沿的几何结构特征. 另一方面, 近似解集中个体之间的 Pareto 占优关系也能反映目标的冲突性和冗余性^[15]. 为了实现目标降维, 需要考虑的一个重要问题是利用近似解集的哪种特征定义冗余的目标.

Saxena 等人^[11]从保持 Pareto 最优前沿几何结构不变的角度定义关键目标子集, 提出主成分分析法(MVU-PCA)寻找对主成分贡献最大的目标.

定义 1 原目标集合为 $F_0 = \{f_1, \dots, f_M\}$, 如果由目标子集 $F_e = \{f_{r_1}, \dots, f_{r_d}\} \subset F_0 (d < M)$ 产生的 Pareto 最优前沿与由 F_0 产生的最优前沿完全相同, 则称 F_e 为原问题的关键目标子集. 称 $d = |F_e|$ 为原问题的内在维数, $F_r = F_0 \setminus F_e$ 为冗余目标子集.

Brockhoff 等人^[12]认为即使被去掉也不会对 Pareto 占优关系带来任何改变的目标为冗余的目标, 提出 δ -MOSS 和 KEMOSS 算法分别求解 P1 和 P2 问题. 并定义 Pareto 占优关系改变误差如下:

定义 2 假设决策向量 $\mathbf{a}, \mathbf{b} \in Z$ 在原目标集合 F_0 下是非支配关系, 如果在目标子集 $F' \subset F_0$ 下二者的关系变为 $\mathbf{a} > \mathbf{b}$, 那么由 F' 引起的 Pareto 占优关系的改变误差为:

$$\delta(F') = \max \{\epsilon_i \mid \forall f_i \in F_0: f_i(\mathbf{a}) - \epsilon_i \leq f_i(\mathbf{b})\} \quad (4)$$

Jaimes 等人^[8]则用目标之间的相关系数度量冲突度, 提出基于特征选择的方法(KOSSA), 选择与其他目标最相关的若干个目标予以删除.

通常, MOEAs 求得的近似解集与 Pareto 最优前沿有一定的差别, 往往不能准确表征原问题本身内在的结构特征, 那么使占优关系不变的目标子集不一定是组成 Pareto 最优前沿的目标子集, 组成低维近似解集的目标集合也不一定使其占优关系不变. 而以上算法只利用近似解集的一种特性具有一定的局限性. 如果同时利用近似解集的两种

特性,将有利于提高目标降维的准确性.

3 基于稀疏特征选择的目标降维算法

近几年,稀疏特征选择方法得到了广泛关注和研究^[13~15].这种方法以谱回归(Spectral Regression)降维^[16]为基础,把对高维数据映射为低维数据的线性稀疏投影矩阵的学习,引入到稀疏回归模型中,有效地避免了传统流形学习中出现的稠密矩阵的特征值分解问题.在获得最优投影矩阵后,依据此矩阵评估每个特征的重要性,从而选择最重要的较少个特征表征原特征集合.本文基于稀疏特征选择的思想,提出新的目标降维算法分别求解 P1 和 P2 两个目标降维问题,算法框架如图 1 所示.该算法首先利用近似解集的几何结构特性构建描述原问题结构特征的邻接图,提出稀疏回归模型求解稀疏投影矩阵,依据此矩阵度量每个目标的重要性并排序.对于 P1 问题,利用近似解集上的 Pareto 占优关系改变程度选择误差不大于阈值的最小目标子集;对于 P2 问题,直接根据目标排序选择前 q 个最重要的目标组成误差最小的目标子集,实现目标降维.

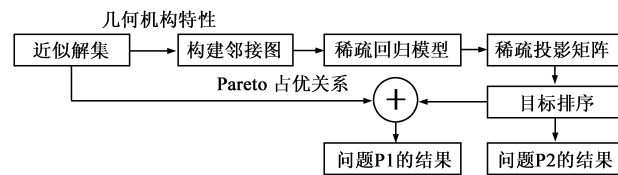


图1 基于稀疏特征选择的目标降维算法框架

3.1 构建邻接图

根据稀疏特征选择的基本思想,首先在原目标集合下构建描述近似解集几何结构特征的邻接图.假设 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathfrak{R}^{M \times N}$ 是一组近似解集, M 是目标个数, N 是解个数, $\mathbf{x}_i \in \mathfrak{R}^M$ 代表第 i 个解. 根据定义 1, 如果 $F_e = \{f_{r_1}, \dots, f_{r_d}\} \subset F_o$ ($d < M$) 是原目标集合 $F_o = \{f_1, \dots, f_M\}$ 的关键目标子集, 那么无论在高维空间还是低维空间, \mathbf{X} 应保持几何结构特征不变, 而且还应保证数据点之间的 Pareto 占优关系不变. 因此, 本文采用近邻保持嵌入(NPE)算法^[17]中邻接图的构造方法刻画高维数据集. 这种方法发现的低维数据能够最大程度的保持原数据的局部几何特征, 同时保持数据点之间的序关系. 具体构造方法如下:

构建邻接图 $G(V, E, S)$, 其中 $V = \{\mathbf{x}_i\}, 1 \leq i \leq N$, E 是连接每一对顶点的边集合, S 是边的权重矩阵. 令 $N(\mathbf{x}_i)$ 为顶点 \mathbf{x}_i 的 k 个近邻点集合, 对于每个数据点 \mathbf{x}_i , 如果 $\mathbf{x}_j \in N(\mathbf{x}_i)$, 则 \mathbf{x}_i 与 \mathbf{x}_j 之间存在一条边, 否则不存在边. 如果 \mathbf{x}_i 与 \mathbf{x}_j 之间不存在边, 则权重矩阵 S 中对应的 $S_{ij} = 0$, 否则 $S_{ij} > 0$, 对于不为零的权重通过求解如式(5)所示的最小化问题得到最优值,

$$\arg \min_{S, \sum S_{ij}=1} \left\| \mathbf{x}_i - \sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} S_{ij} \mathbf{x}_j \right\|_2^2 \quad (5)$$

则图 G 的拉普拉斯矩阵为 $L = (I_{N \times N} - S)^T(I_{N \times N} - S)$, 其中 I 为单位矩阵.

3.2 稀疏投影矩阵

目标降维的目的并非获取近似解集在低维目标子空间的表示, 而是找出最能表示原问题结构特征的目标子集. 假设 $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N] \in \mathfrak{R}^{d \times N}$ 是 \mathbf{X} 的低维嵌入, d 是 \mathbf{Y} 的维数. 为了保持原问题内在结构特征不变, \mathbf{X} 在由 d 个目标组成的低维空间中应仍保持原数据集的局部几何特征, 也就是在低维空间中, 每个样本点 $\mathbf{y}_i \in \mathbf{Y}$ 仍可由其近邻点通过同一个权重矩阵 S 进行重构. 因此低维嵌入 \mathbf{Y} 应满足:

$$\arg \min_{\mathbf{Y}^T = I_{M \times M}} \sum_{i=1}^N \left\| \mathbf{y}_i - \sum_{j=1}^N S_{ij} \mathbf{y}_j \right\|_2^2 \quad (6)$$

最优低维嵌入 \mathbf{Y} 可通过求解 $\text{trace}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T)$ 直接获得.

假设存在投影矩阵 $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_M] \in \mathfrak{R}^{d \times M}$, 使得 $\mathbf{Y} = \mathbf{W}\mathbf{X}$, 其中 $\mathbf{w}_i = [w_{i1}, \dots, w_{id}]$. 用 $\hat{\mathbf{x}}_i = [\hat{x}_{i1}, \dots, \hat{x}_{iN}]^T$ 代表 \mathbf{X} 的第 i 个目标向量, 则 \mathbf{X} 可以改写成 $[\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M]$ 的形式, 那么 $\mathbf{Y} = \mathbf{W}\mathbf{X}$ 可以写成如下的形式:

$$\mathbf{Y} = [\mathbf{w}_1 \quad \dots \quad \mathbf{w}_M] \times \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \vdots \\ \hat{\mathbf{x}}_M \end{bmatrix} = \mathbf{w}_1 \hat{\mathbf{x}}_1 + \dots + \mathbf{w}_M \hat{\mathbf{x}}_M \quad (7)$$

可见, \mathbf{Y} 的形成是 \mathbf{X} 的所有目标向量在 \mathbf{W} 的作用下线性累加而成, \mathbf{w}_i 中各元素的绝对值越大, 对应的第 i 个目标对于 \mathbf{Y} 的形成影响越大, 如果 \mathbf{w}_i 为零向量, 则表明第 i 个目标对于 \mathbf{Y} 的形成无用. 为了实现目标选择, 希望 \mathbf{W} 是稀疏的, 即存在若干个零向量, 那么 $\mathbf{w}_i = \mathbf{0}$ 对应的目标即为冗余的目标. 因此, 在求得稀疏投影矩阵后, 目标的重要性用 \mathbf{w}_i 的 L_2 范数, 即 $\|\mathbf{w}_i\|_2$ 来度量, 以便选择相对重要的目标实现目标降维.

3.3 稀疏回归模型

通过以上分析, 原目标降维问题转化成了求解原高维目标空间映射为低维目标子空间的稀疏投影矩阵. 虽然 \mathbf{Y} 可由 $\text{trace}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T)$ 求得, 但 $\mathbf{Y} = \mathbf{W}\mathbf{X}$ 是超定方程 ($d < M$), 因此无法求得准确的 \mathbf{W} . 一种可行的方法是采用最小二乘法获得近似解, 即:

$$\mathbf{W} = \arg \min \|\mathbf{W}\mathbf{X} - \mathbf{Y}\|_2^2 \quad (8)$$

然而式(8)求得的 \mathbf{W} 不是稀疏的, 无法实现目标降维. 因此希望 \mathbf{W} 同时满足:

$$\mathbf{W} = \arg \min \sum_{i=1}^M \|\mathbf{w}_i\|_2 \quad (9)$$

为了获得稀疏的投影矩阵, 本文采用文献[17]的方法将式(6)、(8)、(9)所示的最优化问题整合为如下所示的稀疏回归模型, 即:

$$L(\mathbf{W}, \mathbf{Y}) = \arg \min_{\mathbf{W}, \mathbf{Y}^T = \mathbf{I}_{M \times M}} \text{trace}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T) + \beta(\|\mathbf{W}\mathbf{X} - \mathbf{Y}\|_2^2 + \alpha \sum_{i=1}^M \|\mathbf{w}_i\|_2) \quad (10)$$

其中 α 和 β 是稀疏回归系数. 文献[17]给出了一种迭代的方法求解式(10), 并证明了代价函数 $L(\mathbf{W}, \mathbf{Y})$ 在迭代过程中是单调递减的, 因此保证了算法的收敛性. 如前所述, 求得稀疏投影矩阵后, 即可对目标的重要性进行评分, 从而选择最重要的目标组成关键目标子集.

3.4 目标降维算法(SORA-1 和 SORA-2)

针对 P1 问题本文提出 SORA-1 算法, 用于寻找原问题改变误差不大于给定阈值的最小目标子集. 由于原问题的内在维数不可知, 本文采用二分查找判定树的方法(decision tree)来确定内在维数, 判定条件是定义 2 中给出的 Pareto 占优关系改变误差 δ . SORA-1 算法流程如图 2(a)所示.

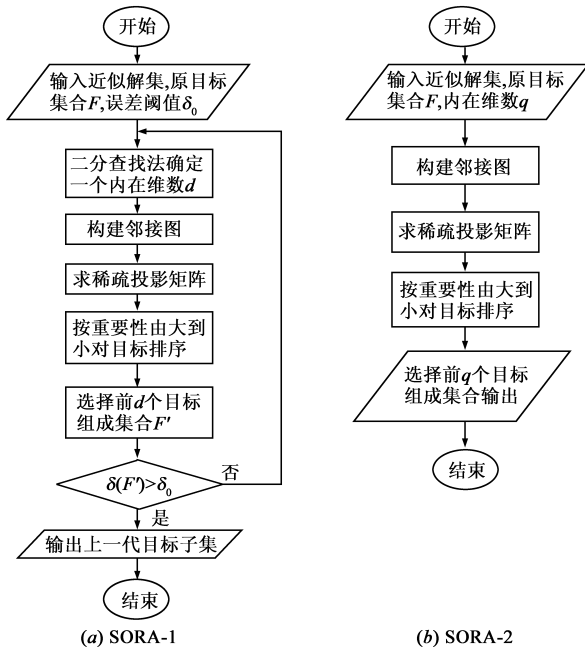


图2 求解P1和P2问题的目标降维算法流程图

P2 问题是寻找原问题改变误差最小的 q 个目标组成的目标子集. 因为给定要寻找的目标个数, 所以省略了 SORA-1 算法中寻找内在维数的过程. 在求得稀疏投影矩阵后, 对所有目标进行排序, 前 q 个目标组成的集合即为关键目标子集. 由于求解问题(10)的过程已被证明是收敛的^[17], 保证了求得的稀疏投影矩阵能够最大程度地描述高维数据与低维数据的映射关系, 因此由其确定的关键目标子集也能够最大程度地表示原问题, 使得原问题的改变误差最小. SORA-2 算法流程如图 2(b)所示.

4 实验结果与分析

本节的实验目的有两个: 一是针对 SORA 算法本身

考察参数对算法性能的影响; 二是通过与其他算法相比较, 考察 SORA 算法的有效性. 实验中采用的测试问题是 DTLZ5 修改版^[18], 即 $\text{DTLZ5}(I, M)$, 其中 I 为原问题的内在维数, M 为原问题的目标个数. 该测试问题不仅目标个数 M 可任意扩展, 而且可通过设置 I 为 2 到 M 之间任意一个整数来控制问题的冗余度, 其关键目标子集是 $\{f_{M-I+1}, f_{M-I+2}, \dots, f_M\}$.

实验中, 我们采用基于指标的多目标进化算法 (IBEA)^[19], 通过进化不同的代数, 获得不同质量的近似解集, 评估近似解集质量对降维准确性的影响. 按照文献[24]的建议, IBEA 的遗传算子采用参数为 5 的模拟二进制交叉算子(概率为 0.9)和参数为 50 的多项式变异算子(概率为 0.1). 用 $\text{DTLZ5}(I, M)-G$ 表示一个测试实例, 其中 M 为目标个数, I 为内在维数, G 为 IBEA 进化代数. 目标个数为 10 时, 种群规模设为 200, 目标个数为 20 时, 种群规模设为 500. 每种测试实例下 IBEA 独立运行 20 次.

4.1 SORA 参数

SORA 算法中有三组参数: 近邻个数 k , 稀疏回归系数 α 和 β , 以及收敛准则. 由于 SORA-1 和 SORA-2 都是依赖目标重要性指标选择目标, 其区别仅在于选择的准则不同, 上述参数只对目标重要性指标有影响, 因此实验中仅利用 SORA-2 对参数进行测试.

4.1.1 近邻个数 k

本组实验考察 k 分别取 $\{1, 2, \dots, 10\}$ 时找到错误关键目标子集的次数. 从图 3 的实验结果可以看出, 当进化代数不大于 100 时, 找到错误关键目标子集的次数随着 k 的增加呈减少的趋势. 表明近似解集质量较差时, 较大的近邻个数能够更准确地描述原高维数据的局部几何结构, 提高降维准确性. 当进化代数大于 100 时, 近邻个数对降维准确性没有影响. 因此对于高质量的近似解集, 近邻个数可设置为较小值, 这样既可以保证较高的降维准确率, 也能在一定程度上降低计算复杂度.

4.1.2 稀疏回归系数 α 和 β

实验中设置 α 的取值范围为 $[1.5, 2.4]$, 变化步长为 0.1, β 的取值范围为 $[0.01, 0.1]$, 变化步长为 0.01. 计算每一种 $\alpha - \beta$ 组合形式下找到错误关键目标子集的次数, 考察 $\alpha - \beta$ 的取值对降维准确性的影响, 实验结果如图 4 所示. 除 $\text{DTLZ5}(8, 10) - 100$ 外, 对于同一个测试实例, 所有 $\alpha - \beta$ 组合下的错误次数都相等. $\text{DTLZ5}(8, 10) - 100$ 的结果显示仅有 6 种 $\alpha - \beta$ 组合形式会引起降维准确性的变化. 当近似解集质量较高时(进化 150 代和 200 代), 无论哪种 $\alpha - \beta$ 组合形式, 降维错误率均为零. 可见, SORA 算法的降维准确性对于 α 和 β 的取值并不十分敏感.

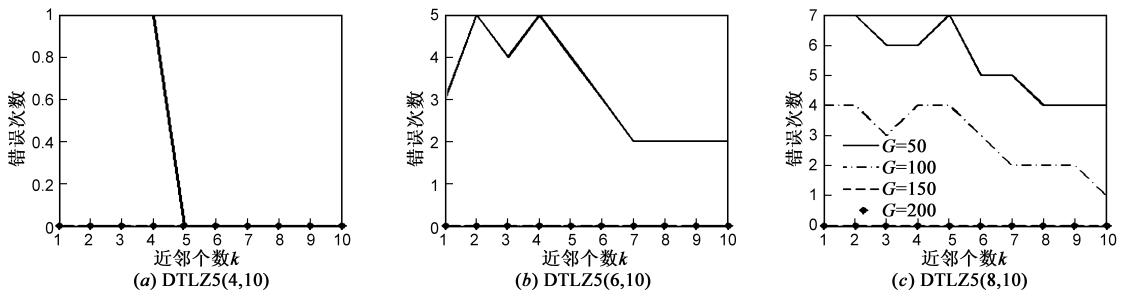


图3 不同近邻个数下求得错误关键目标子集的次数

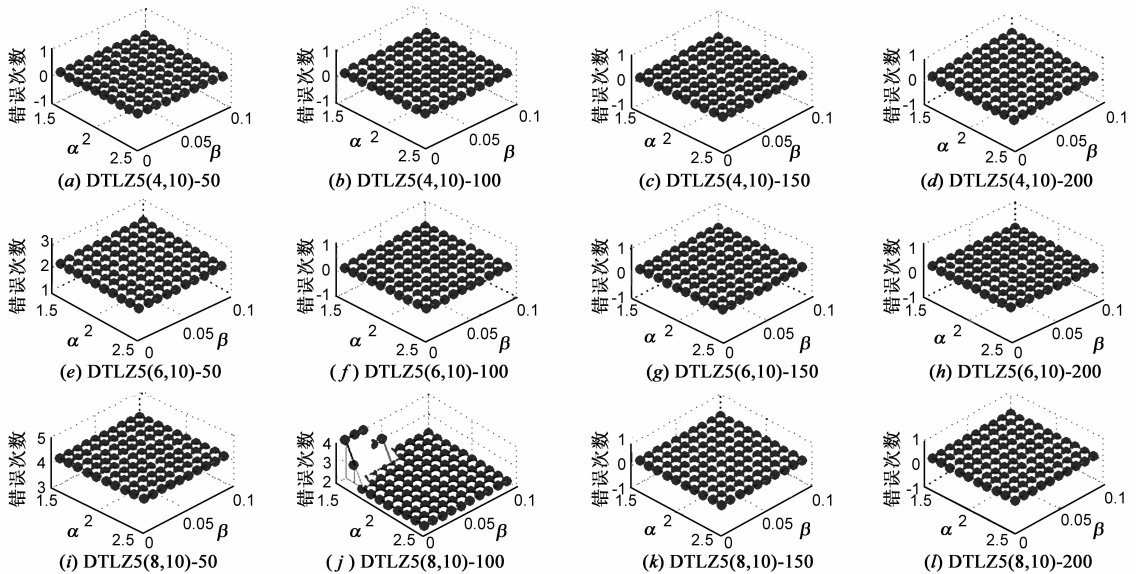


图4 不同稀疏回归系数组合方式下找到的错误最小关键目标子集的次数(纵坐标为错误次数,总共20次独立实验)

4.1.3 收敛准则

SORA算法采用迭代的方式求解满足式(10)的最优稀疏投影矩阵,这个过程是SORA计算量最大的部分,其收敛速度直接影响算法的计算效率。

首先评估算法的收敛速度.设置最大迭代次数为50,记录每次迭代后式(10)的代价函数值,结果如图5

所示.横坐标为迭代次数,纵坐标为20次实验的平均代价函数值,不同的曲线代表IBEA进化的不同代数.可以看出,在所有测试实例下,SORA-2都能在10代之内收敛,表明这种求解方法的计算效率较高,不会占用太多的计算资源和计算时间。

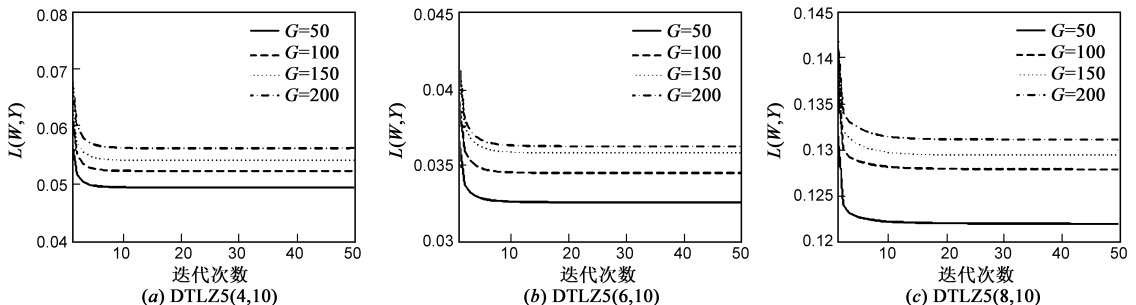


图5 SORA-2算法的收敛特性曲线

但在实际中,设置最大迭代次数作为停止条件通常会浪费计算资源,因此我们考察另一种收敛准则,即当相邻两代的代价函数值的差小于某个阈值 σ 时则停止.为了获得合理的 σ ,我们考察 σ 分别为0.005、0.001

和0.0001时找到错误关键目标子集的次数,并记录对应的迭代次数,实验结果如表1所示.可见,虽然 σ 越小迭代次数越多,但迭代次数的增加并没有使降维的错误次数有明显的下降,特别地,对于DTLZ5(4,10)和

DTLZ5(6,10)测试实例,当 IBEA 进化代数大于 50 时,迭代 4 次即可使得降维错误率为零.可见,SORA-2 算法虽然是一种迭代的计算方法,但其收敛速度较快,计算效率较高.

表 1 不同收敛条件下找到错误关键目标子集的次数以及对应的平均迭代次数

测试实例	$\sigma = 0.005$		$\sigma = 0.001$		$\sigma = 0.0001$	
	错误次数	迭代次数	错误次数	迭代次数	错误次数	迭代次数
DTLZ5(4,10) - 50	3	4.1	3	5.4	3	8.7
DTLZ5(4,10) - 100	0	4.0	0	5.4	0	8.2
DTLZ5(4,10) - 150	0	4.0	0	5.6	0	8.8
DTLZ5(4,10) - 200	0	4.0	0	5.9	0	9.6
DTLZ5(6,10) - 50	3	4.0	2	5.4	4	9.8
DTLZ5(6,10) - 100	0	4.0	0	5.2	0	9.6
DTLZ5(6,10) - 150	0	4.0	0	5.7	0	10.3
DTLZ5(6,10) - 200	0	4.0	0	5.9	0	10.4
DTLZ5(8,10) - 50	3	4.0	3	4.8	4	8.8
DTLZ5(8,10) - 100	3	4.0	3	4.9	2	9.5
DTLZ5(8,10) - 150	1	4.0	1	5.0	0	10.8
DTLZ5(8,10) - 200	0	4.0	0	5.0	0	11.5

4.2 SORA 算法的性能测试

本组实验的目的是通过与其他目标降维算法比较,考察 SORA 算法的有效性.实验中,根据 4.1 节参数的测试结果,SORA-1 和 SORA-2 均采用相同的参数设置,近邻大小 $k = M - 1$,稀疏回归系数 $\alpha = 2.4$, $\beta = 0.01$,收敛条件为迭代相邻代目标函数差值 σ 小于 0.001 时算法停止.每个测试实例下 IBEA 独立运行 20 次.

4.2.1 SORA-1 性能测试

本组实验将 SORA-1 算法与 δ -MOSS 算法^[12]比较,以 Pareto 占优关系改变程度 δ 为误差函数,求解 P1 问题.设置 δ 分别为近似解集中最大差值的 $\{0, 10\%, 20\%, 30\%, 40\%, 50\%\}$.实验中统计找到的目标均为关键目标的次数.由图 6 给出的实验结果可见,在相同允许误差下,SORA-1 算法的准确次数总是大于 δ -MOSS 算法.而且近似解集质量的提高,SORA-1 算法的准确次数增加,甚至达到 100% 的成功率,而 δ -MOSS 算法的成功率基本不超过 50%.

4.2.2 SORA-2 性能测试

本组实验将 SORA-2 与 KOSSA 算法^[8]、KEMOSS 算法^[12]和 MVU-PCA 算法^[11]比较,用于求解 P2 问题,寻找目标个数为 I 的目标子集.

首先,以 Pareto 占优关系改变误差 δ 来评估算法保持原问题结构特性不变的能力, δ 越小,表明降维算法

的有效性越佳.由图 7 的比较结果可以看出,KEMOSS 算法虽然以 Pareto 占优关系改变误差最小为寻找目标子集的目标,但该算法得到的误差并不总是最小的,表明仅利用近似解集的 Pareto 占优关系分析目标的冗余性并非十分有效.另外 MVU-PCA 算法在所有测试实例下求得的目标子集引起的误差最大,表明仅利用近似解集的几何结构特征也不能保证降维的合理性.本文提出的 SORA-2 算法,除 DTLZ5(8,10)和 DTLZ5(10,20)测试问题外,无论 IBEA 算法进化代数多少,总是能获得最小误差的目标子集.可见,SORA-2 算法保持原问题结构特征不变的能力好于其他三种算法.

接下来考察每个算法的降维准确性.表 2 给出了每个测试实例下,算法找到错误目标子集的次数,其中粗体数字表示对应测试实例下的最好结果.可以看出,对于 24 个不同规模、不同冗余度的测试实例,SORA-2 找到正确目标子集的次数仅在 6 个测试实例下稍逊于其他算法,而在其他测试实例下都是最好的,并且错误率基本为零,表明本文算法具有较强的适应性和稳定性.另外,对于 IBEA 进化较少代数(如 50 代)时获得的近似解集,本文的算法仍可以获得较低的错误率.表明,本文算法降维的准确性受近似解集质量的影响较小.

表 2 不同算法找到错误目标子集的次数(总次数为 20)

测试实例	KOSSA	KEMOSS	MVU-PCA	SORA-2
DTLZ5(4,10) - 50	18	17	12	0
DTLZ5(4,10) - 100	17	18	16	0
DTLZ5(4,10) - 150	16	20	16	0
DTLZ5(4,10) - 200	19	18	20	0
DTLZ5(6,10) - 50	20	11	7	2
DTLZ5(6,10) - 100	19	16	3	0
DTLZ5(6,10) - 150	17	12	3	0
DTLZ5(6,10) - 200	17	5	6	0
DTLZ5(8,10) - 50	12	2	14	4
DTLZ5(8,10) - 100	14	5	10	2
DTLZ5(8,10) - 150	14	8	6	0
DTLZ5(8,10) - 200	13	6	9	0
DTLZ5(5,20) - 50	20	18	20	0
DTLZ5(5,20) - 100	20	17	0	0
DTLZ5(5,20) - 150	10	17	0	0
DTLZ5(5,20) - 200	20	11	0	0
DTLZ5(10,20) - 50	20	0	0	3
DTLZ5(10,20) - 100	19	2	0	1
DTLZ5(10,20) - 150	20	2	1	0
DTLZ5(10,20) - 200	19	5	0	0
DTLZ5(15,20) - 50	17	2	0	6
DTLZ5(15,20) - 100	17	0	0	3
DTLZ5(15,20) - 150	18	2	2	3
DTLZ5(15,20) - 200	18	1	3	2

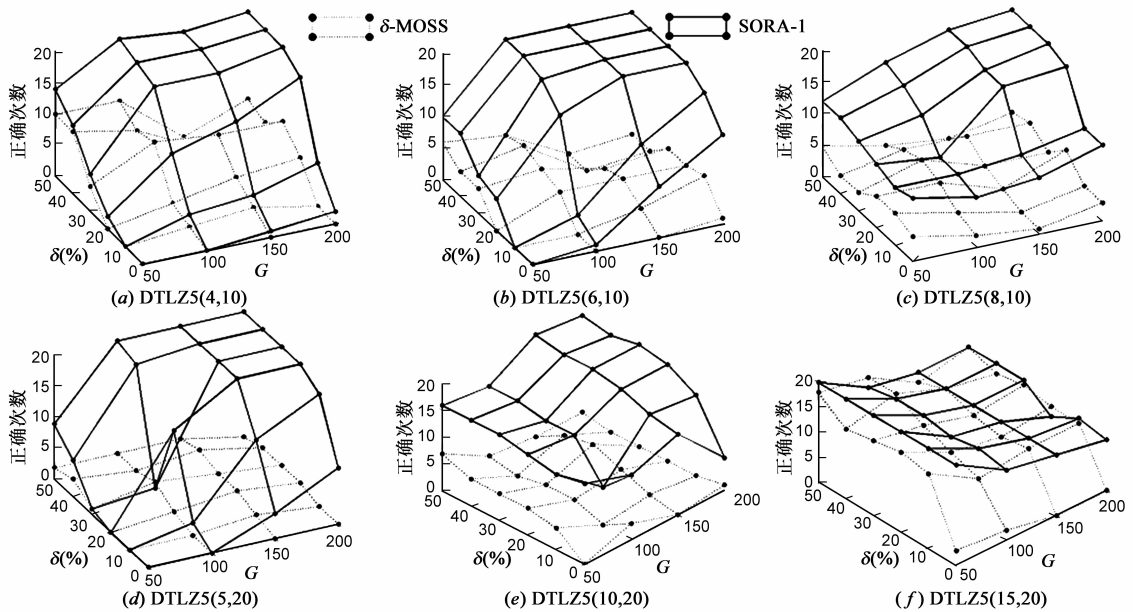
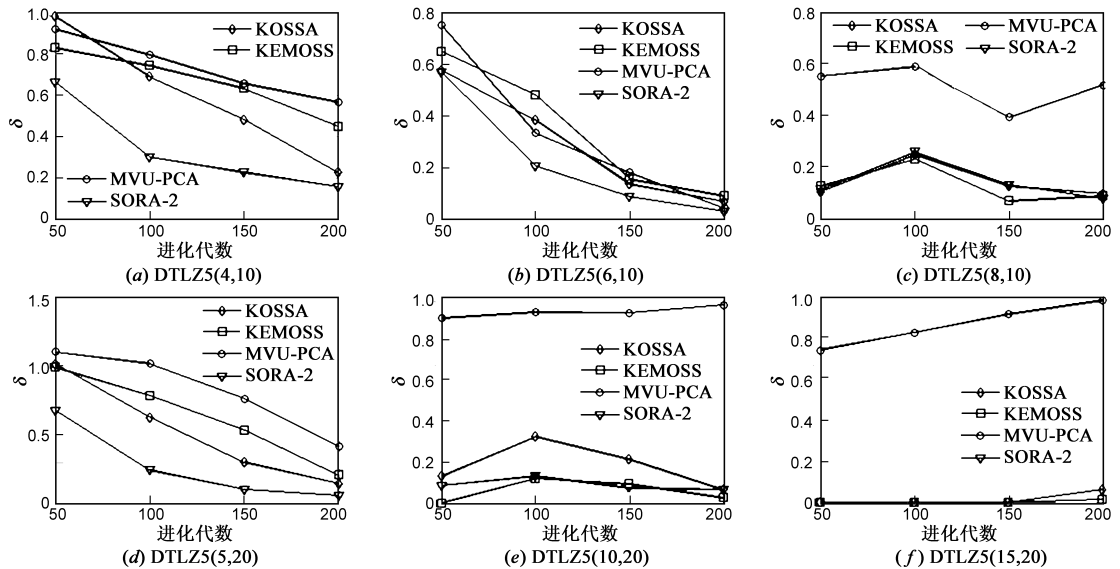


图6 不同允许误差下找到的目标是关键目标的次数

图7 不同算法求得的目标子集引起的Pareto占优关系改变误差 δ

5 结论

本文提出一种新的目标降维算法,该算法基于稀疏特征选择的思想,求解两种目标降维问题.利用近似解集的几何结构特性描述原问题的内在结构特征,提出稀疏回归模型求解稀疏投影矩阵,以此矩阵度量每个目标的重要性,再利用 Pareto 占优关系改变程度选择误差小于阈值的 minimum 目标子集,或直接选择前 q 个目标组成误差最小的目标子集.

实验结果表明,本文算法的性能受参数影响较小.另一方面,通过与其他目标降维算法比较,本文算法对于两种目标降维问题都表现出较高的准确性和鲁棒

性,而且其性能受近似解集的质量影响较小.

参考文献

- [1] Farina M, Amato P. On the optimal solution definition for many-criteria optimization problems[A]. Proceedings of Fuzzy Information Processing Society[C]. IEEE, 2002. 233 - 238.
- [2] 巩敦卫,季新芳,孙晓燕.基于集合的高维多目标优化问题的进化算法[J].电子学报,2014,42(1):77 - 83.
GONG Dun-wei, JI Xin-fang, SUN Xiao-yan. Solving many-objective optimization problems using set-based evolutionary algorithms[J]. Acta Electronica Sinica, 2014, 42(1): 77 - 83. (in Chinese)

- [3] Lopez-Jaimes A, Coello Coello C A. Including preferences into a multiobjective evolutionary algorithm to deal with many-objective engineering optimization problems[J]. Information Sciences, 2014, 277: 1 – 20.
- [4] 孔维健, 丁进良, 柴天佑. 高维多目标进化算法研究综述[J]. 控制与决策, 2010, 25(3): 321 – 326.
KONG Wei-jian, DING Jin-liang, CHAI Tian-you. Survey on large-dimensional multi-objective evolutionary algorithms[J]. Control and Decision, 2010, 25(3): 321 – 326. (in Chinese)
- [5] Wagner T, Beume N, Naujoks B. Pareto-, aggregation-, and indicator-based methods in many-objective optimization [A]. Evolutionary Multi-Criterion Optimization[C]. EMO, Springer, 2007. 742 – 756.
- [6] Ishibuchi H, Tsukamoto N, Nojima Y. Evolutionary many-objective optimization: a short review[A]. IEEE World Congress on Computational Intelligence[C]. IEEE, 2008. 2419 – 2426.
- [7] Bringmann K, Friedrich T. Approximating the least hyper-volume contributor: NP-hard in general, but fast in practice[J]. Theoretical Computer Science, 2012, 425: 104 – 116.
- [8] López Jaimes A, Coello Coello C A, Chakraborty D. Objective reduction using a feature selection technique[A]. Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation[C]. ACM, 2008. 673 – 680.
- [9] Schutze O, Lara A, Coello C A C. On the influence of the number of objectives on the hardness of a multiobjective optimization problem [J]. IEEE Transactions on Evolutionary Computation, 2011, 15(4): 444 – 455.
- [10] Sinha A, Saxena D K, Deb K, et al. Using objective reduction and interactive procedure to handle many-objective optimization problems[J]. Applied Soft Computing, 2013, 13(1): 415 – 427.
- [11] Saxena D K, Duro J A, Tiwari A, et al. Objective reduction in many-objective optimization: linear and nonlinear algorithms [J]. IEEE Transactions on Evolutionary Computation, 2013, 17(1): 77 – 99.
- [12] Brockhoff D, Zitzler E. Objective reduction in evolutionary multiobjective optimization: theory and applications[J]. Evolutionary Computation, 2009, 17(2): 135 – 166.
- [13] Qian M, Zhai C. Robust unsupervised feature selection[A]. Proceedings of 23rd International Joint Conference on Artificial Intelligence[C]. AAAI Press, 2013. 1621 – 1627.
- [14] Hou C, Nie F, Yi D, et al. Feature selection via joint embedding learning and sparse regression[A]. Proceedings of 22nd International Joint Conference on Artificial Intelligence[C]. AAAI Press, 2011. 1324 – 1329.
- [15] 向旭, 李炳南. 主元分析中的稀疏性[J]. 电子学报, 2012, 40(12): 2525 – 2532.
XIANG Nan, LI Bing-nan. Sparsity in principal component analysis: a survey[J]. Acta Electronica Sinica, 2012, 40(12): 2525 – 2532. (in Chinese)
- [16] Cai D, He X, Han J. Spectral regression for dimensionality reduction[R]. Department of Computer Science Technical Report No. 2856, University of Illinois at Urbana-Champaign, 2007.
- [17] He X, Cai D, Yan S, et al. Neighborhood preserving embedding[A]. IEEE International Conference on Computer Vision [C]. IEEE, 2005. 1208 – 1213.
- [18] Saxena D K, Duro J A, Tiwari A, et al. Objective reduction in many-objective optimization: linear and non-linear algorithms [J]. IEEE Transactions on Evolutionary Computation, 2013, 17(1): 77 – 99.
- [19] Zitzler E, Künzli S. Indicator-based selection in multi-objective search[A]. Proceedings on Parallel Problem Solving from Nature[C]. Springer, 2004, 3242: 832 – 842.

作者简介



陈小红 女, 1979 年生于辽宁省锦州市. 现在深圳大学信息工程学院攻读博士学位. 主要研究方向为智能优化算法.

E-mail: chenxh@vip.qq.com



李霞 女, 1968 年出生于四川省乐山市, 毕业于香港中文大学获博士学位, 深圳大学信息工程学院教授, 博士生导师. 主要研究方向为智能计算及应用, 多目标优化.

E-mail: lixia@szu.edu.cn