

面向多极值质量特性的过程参数全局优化研究^①

崔庆安

(郑州大学管理工程研究所, 郑州 450001)

摘要: 对于作用关系复杂, 而且质量特性拥有多个极值的制造过程, 现有质量改进方法只能实现参数的局部优化, 产品质量仍有较大改进空间. 本文采用支持向量机(SVM)作为复杂作用关系过程的近似模型, 提出基于支持向量聚类(SV)与序列二次规划(SQP)的参数全局性优化方法. 首先建立了复杂过程的SVM近似模型; 而后根据 ε 管道理论, 通过对聚类过程谱系图的分析, 确定了聚类的最小相似度水平及合适的聚类数目, 将过程各极值点邻域内的支持向量分别聚为一类; 最后由各聚类中心出发, 并行进行SQP寻优以发现过程的多极值. 仿真研究表明, 所提方法能够全面反映过程的极值分布, 实现参数的全局性优化; 寻优结果与实际极值的绝对偏差及相对偏差的平均值分别为0.15%和1.28%, 并且偏差的大小与过程极值的数目无关, 说明方法具有较高的精确度和稳定性; 此外, 通过支持向量聚类, 不仅保证了SQP寻优结果对于过程全部极值的遍历性, 而且将寻优的次数降低了50%以上, 提高了寻优效率.

关键词: 多极值质量特性; 全局优化; 支持向量机; 聚类分析; 质量改进

中图分类号: F406.2 **文献标识码:** A **文章编号:** 1007-9807(2012)09-0046-12

0 引言

制造业中, 控制和调整合适的过程参数可以得到期望的输出质量特性, 因此过程参数优化是提高产品质量的重要途径之一. 随着技术不断进步, 在机械电子、医药化工等行业出现了大量复杂的制造过程^[1], 其输出的质量特性呈多极值分布, 而且过程参数与质量特性高度非线性相关, 但是对其物理或工程机理不甚明晰. 上述特性对于过程参数优化提出了新要求. 例如, 对于某类化工生产过程, 最终产品的收率有多个极值点, 在不同极值处, 过程参数值(包括溶液的添加量、反应温度、反应时间)差异较大, 导致生产成本也差异较大. 如果能够发现收率的多极值分布情况, 综合考察收率与成本, 进而确定合适的过程参数值, 即可实现效益的最大化; 再如, 对于产品质量波动较大的过程, 如果能够发现质量特性均值与方差的多

极值分布情况, 在均值达到目标与方差最小之间实现折中, 提高过程的稳健性, 比单纯性地优化质量特性均值更具实际意义. 总的来说, 如何能够尽可能真实地描述过程作用关系, 探究过程质量特性的多极值分布情况, 进而实现过程参数的全局性优化, 成为制造业迫切需要解决的问题.

对于没有具体作用关系模型的过程, 参数优化的主要方法是基于实验设计(design of experiments, DOE)的过程建模和寻优^[2]. 常用的方法包括因子实验^[3]、响应曲面法^[4-5]、均匀设计^[6]、正交实验^[7]等, 但是由于多项式拟合模型及寻优策略的限制, 此类方法较好适用于单极值过程, 很难在可行域全局内描述质量特性的多极值分布. 近年来, 研究者开始利用非参数模型来近似过程的复杂作用关系^[8], 包括核函数回归^[9-10]、人工神经网络^[11-12]、局部多项式回归^[13]、Kriging模

① 收稿日期: 2010-12-02; 修订日期: 2011-06-17.

基金项目: 国家自然科学基金重点资助项目(70931004); 国家自然科学基金资助项目(71171180); 河南省高等学校青年骨干教师计划资助项目(2012GGJS-020).

作者简介: 崔庆安(1974—), 男, 山西襄垣人, 博士, 副教授. Email: cuiqa@zhu.edu.cn

型^[14-15]等等. 此类方法需要大的样本量才能保证模型性能, 但是由于时间及成本的约束, 实验设计所得样本量通常较少. 此外, 非参数模型的解析形式高度复杂, 给最终的过程寻优带来较大困难. 为改进上述不足, 研究者^[16-17]提出基于支持向量机 (support vector machine, SVM), 面向复杂作用关系制造过程的全局性建模方法. 研究表明, 该方法不仅可以在全局范围内描述过程的复杂作用关系, 而且建模所需样本量最少; 另一方面, SVM 模型的解析形式相对简单, 且处处可导, 有利于应用各类优化算法进行寻优. 但是, 前期研究侧重于过程建模, 尚未涉及如何对模型进行寻优以实现参数全局性优化的问题. 而现有关于最优化方法的研究, 例如基于梯度信息^[18]的最速下降法、牛顿法, 基于计算智能^[19]的模拟退火算法、遗传算法等, 研究者在改善寻优结果^[20]、拓展应用领域^[21]等方面做了大量工作. 但是仍然是以如何发现模型的单一极值(即所谓的全局最优)为目标, 无法反映过程的多极值分布, 也无法实现参数的全局性优化. 本文在前期研究^[16]的基础上, 研究如何基于已建立的 SVM 模型, 利用约束最优化算法, 实现复杂作用关系过程的全局性优化参数. 首先简要回顾了基于 SVM 的全局性建模方法, 而后分析复杂作用关系过程的参数全局性优化的特点, 研究 SVM 模型中支持向量与质量特性极值的对应分布关系, 提出基于支持向量聚类与序列二次规划的过程寻优与参数全局性优化方法; 最后通过仿真研究验证了方法的正确性与有效性.

1 理论简介

1.1 统计学习理论与支持向量机

统计学习理论及 SVM 技术^[22]目前被广泛应用于模式识别^[23]、建模与预测等领域^[24]. 按照应用类型不同, 可以分为支持向量分类机和支持向量回归机. 设过程输入 $x = [x_1, x_2, \dots, x_m]^T$ 与输出 y 之间存在某种未知的依赖关系 $f(x^*)$ 或联合概率分布 $F(x, y)$, n 个独立同分布的样本为

$$S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (1)$$

回归估计就是根据式 (1) 寻找某个实值函数 $f(x)$, 而后用 $\hat{y} = f(x)$ 来估计输入 x 所对应的输出值 y . 统计学习理论认为, 在寻找最优的 $f(x)$

时, 应保证它对于未知样本的预测风险最小, 该预测风险为

$$R = \int L(y, f(x)) dF(x, y) \quad (2)$$

其中 $L(y, f(x))$ 为损失函数, 一般将其定义为不敏感参数 ε ($\varepsilon > 0$), 即

$$L(y, f(x)) = |y - f(x)|_\varepsilon = \begin{cases} 0, & |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon, & |y - f(x)| > \varepsilon \end{cases} \quad (3)$$

SVM 是统计学习理论的实现工具, 在 SVM 中, 实值函数 $f(x)$ 的一般形式为

$$f(x) = wT(x) + b \quad (4)$$

其中: $T(x)$ 为某一非线性变换函数; w, b 可以通过求解如下的最优化问题来得到

$$\begin{aligned} \min_{a_i^*} & \frac{1}{2} \sum_{i,j=1}^n (a_i^* - a_i)(a_j^* - a_j) k(x_i, x_j) + \\ & \varepsilon \sum_{i=1}^n (a_i^* + a_i) - \sum_{i=1}^n y_i (a_i^* - a_i) \quad (5) \\ \text{s. t.} & \sum_{i=1}^n (a_i - a_i^*) = 0 \\ & 0 \leq a_i, a_i^* \leq C, i = 1, 2, \dots, n \end{aligned}$$

其中 $k(x_i, x_j) = T(x_i)T(x_j)$, 称 $k(x_i, x_j)$ 为核函数, 用以简化非线性变换后的高维计算, 常用的有 高斯径向基核、多项式核、傅立叶核等. C 称为惩罚参数, 且有 $C > 0$.

求解式 (5), 可以在模型的拟合性能与推广能力之间取得折中, 从而得到最小的预测风险. 式 (5) 的最优解为

$$(\bar{a}, \bar{a}^*) = (\bar{a}_1, \bar{a}_1^*, \bar{a}_2, \bar{a}_2^*, \dots, \bar{a}_n, \bar{a}_n^*) \quad (6)$$

计算 $\bar{w} = \sum_{i=1}^n (\bar{a}_i^* - \bar{a}_i) k(x_i, x)$, 选择位于区间 $(0, C)$ 中的 \bar{a}_i 或 \bar{a}_i^* , 计算

$$\bar{b} = y_j - \sum_{i=1}^n (\bar{a}_i^* - \bar{a}_i) k(x_i, x) - \varepsilon \quad (7)$$

则 $f(x)$ 变为

$$f(x) = \sum_{i=1}^n (\bar{a}_i^* - \bar{a}_i) k(x_i, x) + \bar{b} \quad (8)$$

式 (8) 即为关于样本集 S 的 SVM 拟合方程. 其中 $\bar{a}_i \neq 0$ 或 $\bar{a}_i^* \neq 0$ 所对应的样本 x_i 称为支持向量, 其余样本称为非支持向量.

在统计学习理论中, ε 管道模型^[22]描述了式

(3) 中 ε 与支持向量的关系. 根据式 (3) 在精度 ε 下, 用函数 $f(x)$ 去拟合未知依赖关系 $f(x^*)$, 有 $f(x^*) \leq f(x) \pm \varepsilon$, 从几何上来看, 就是 $f(x^*)$ 处于以 $f(x)$ 为轴线的, 直径为 ε 的弹性管道内, 而该管道的弹性规律就是核函数 $K(x_i, x)$. 要构造 $f(x)$, 就是把 $f(x^*)$ 套在管道中. 而管道在趋于平坦的过程中, 会接触到 $f(x^*)$ 的一些点, 这些点阻止了管道的进一步变化, 支撑了管道的形状, 而此时管道的轴线就定义了 $f(x^*)$ 的 ε 逼近 $f(x)$, 而支撑点处的坐标则对应着支持向量样本.

1.2 基于 SVM 的全局性建模方法

设 $x = [x_1, x_2, \dots, x_m]^T$ 代表制造过程参数的组合, y 代表过程输出质量特性的值. 则可以采用如下的方法来得到 SVM 拟合方程.

步骤 1 采用均匀空间网格的实验设计方法获取式 (1) 所示的样本集 S ;

步骤 2 估计过程误差 σ_p ;

步骤 3 选择 Gauss 函数

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$$

作为式 (5) 的核函数. 根据 σ_p 和样本集 S , 并由文献 [16] 的方法确定参数 σ , 不敏感参数 ε 、惩罚参数 C 的值;

步骤 4 利用样本集 S 和 ε, C 进行 SVM 拟合, 得到拟合方程 (8), 计算拟合误差、预测误差.

2 过程参数全局性优化方法研究

2.1 基于 SVM 模型的参数全局性优化的问题表述

根据样本集建立了拟合模型以后, 参数优化即转化为对拟合模型的寻优. 对多极值质量特性, 就是通过寻找拟合模型输出的全部极值 $f(x^*) = [x_1^*, x_2^*, \dots, x_k^*]$, 得到其对应的输入向量 $x^* = [x_1^*, x_2^*, \dots, x_k^*]$, 也即是对应的参数值. 再结合 x 的可行域约束范围, 参数优化问题即转换为一类约束最优化问题. 即(以望大型问题为例)

$$\begin{aligned} \max_{x^*} f(x) &= \sum_{i=1}^n (\bar{\alpha}_i^* - \bar{\alpha}_i) k(x_i, x) + \bar{b} \\ \text{s. t. } &\begin{cases} x_{\min} \leq x \leq x_{\max} \\ c_i(x) \geq 0; i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (9)$$

式中的 $\forall x^*$ 表示需要求出目标函数中的所

有的局部极大值点, x_{\min}, x_{\max} 表示可行域的边界, $c_i(x) \geq 0; i = 1, 2, \dots, m$ 表示各参数之间的线性及非线性约束. 显然, 式(9) 不同于现有一般意义上的约束最优化问题. 现有最优化问题所提供的优化结果均是单个的优化点, 而且不一定是全局最优点. 具体来说, 式(9) 通过寻找目标函数的全部极值来保证得到全局性极值, 同时多个极值也提供了质量特性的全局性分布规律, 有助于在产品质量与制造成本之间取得折中, 也有助于综合考察各极值点处质量特性方差的分布情况, 进而实现稳健性设计.

进一步地, 根据 1.2 小节 SVM 核函数的选择, 结合统计学习中的 ε 管道模型, 式(9) 所示的约束最优化问题有如下两点特性.

1) 由于采用 Gauss 径向基核作为核函数, 因此式(9) 的目标函数又可以写为

$$f(x) = \sum_{i=1}^l (\bar{\alpha}_i^* - \bar{\alpha}_i) \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) + \bar{b} \quad (10)$$

其中 $\bar{\alpha}_i^*, \bar{\alpha}_i, \sigma, \bar{b}$ 为参数. 针对特定的样本集是固定的, 因此, 该拟合方程是一系列不含未知参数, 关于 x 的指数函数的线性组合. 此外

$$\|x_i - x\|^2 = (x_{i1} - x_1)^2 + (x_{i2} - x_2)^2 + \dots + (x_{im} - x_m)^2 \quad (11)$$

在 x 的可行域内连续, 拥有高阶导数, 因此拟合方程(9) 的目标函数在 x 的可行域内任意点处连续, 且存在梯度信息.

2) ε 管道模型指出, 支持向量的样本点支撑了 ε 管道的形状, 进而确定了 ε 管道的轴线, 而该轴线就是拟合方程的曲面形状. 因此, 支持向量分布反映了曲面形状以及过程极值分布, 或者说, 在过程极值附近存在着支持向量. 因此, 对于式(10) 中的支持向量样本 x_i ($\bar{\alpha}_i \neq 0$ 或 $\bar{\alpha}_i^* \neq 0$ 所对应的样本) 在其附近极有可能存在着过程的局部极值.

上述两特性有助于寻优算法的选择及寻优策略的确定.

2.2 全局性寻优算法的选择

求解约束最优化问题的方法, 包括基于梯度信息的确定性算法^[18] 和基于计算智能的随机性算法^[19]. 前者包括最速下降法、牛顿法、罚函数法、序列二次规划法^[25] (sequential quadratic programming, SQP) 等, 其特点是要求目标函数

在其可行域内可导,优化时从某个初始点出发,根据目标函数的梯度信息进行迭代运算,最终收敛至某个极值点;后者包括模拟退火算法、禁忌搜索算法、遗传算法等,其特点是不需要目标函数的解析表达式和梯度信息,根据某些准则,通过随机性的搜索实现优化。两类算法在应用中均存在一定的局限性:对于确定性的优化算法,初始迭代点对于最终结果影响较大,如果初始点在某个局部极值附近,那么算法可以很快收敛到该极值点,但无法达到全局最优;对于随机性的优化算法,虽然理论上可以收敛至全局最优,但是算法参数(例如模拟退火中的温度参数、禁忌搜索算法中的禁忌长度、遗传算法中的初始种群)对最终寻优结果影响较大,实际应用中很难得到最优的参数设置,因此也不一定都收敛至全局最优。

对于式(9)所示的约束最优化问题,由于要求寻找到目标函数的全部极值,因此要保证优化算法寻优结果的遍历性,但是上述两类算法均是以发现过程的单一极值为目标,并不能够满足发现过程所有极值的研究,因此不能直接应用,需要做进一步改进。由于式(9)所示的目标函数的解析形式明显,且在可行域内处处可导,存在梯度信息,因此既可以采用确定性寻优算法,也可以采用随机性寻优算法。然而,随机性寻优算法虽然可以在整个可行域内进行寻优,但是由于搜索准则的随机性,不一定能够遍历可行域内的全部极值点;而对于确定性搜索算法,虽然初始点附近的局部极值会使寻优陷入局部最优,但这一局限性恰恰有助于发现多个极值。或者说,通过选择合适的初始迭代点的集合,可使算法依次收敛于迭代点附近的极值点,进而发现可行域内的多个极值。综合上述分析,可选择确定性的优化算法作为式(9)的寻优算法。进一步地,SQP^[26]是目前求解约束优化的最为有效的确定性优化算法之一,在每次迭代时,均利用原有约束构造一个较简单的近似优化问题,由此修正当前迭代点,直至收敛至某个点。它具有迭代次数少,收敛速度快等优点,因此可以作为基本优化算法并加以改进,以求解式(9)所示的约束最优化问题。

2.3 全局性寻优的实现

2.3.1 初始迭代点集合的确定

初始迭代点集合的确定对于寻找到目标函数

的全部极值具有关键影响。根据2.1小节中的分析,在SVM模型的支持向量样本附近存在过程极值的概率较大,这提示可以分别以各个支持向量作为初始迭代点,再利用SQP算法发现其附近的极值。由于在过程极值附近均存在着一定数目的支持向量,因此采用上述方法有助于发现可行域内的全部极值。此外,由于各支持向量不存在先后关系,因此SQP算法可以并行进行,以缩短寻优时间。

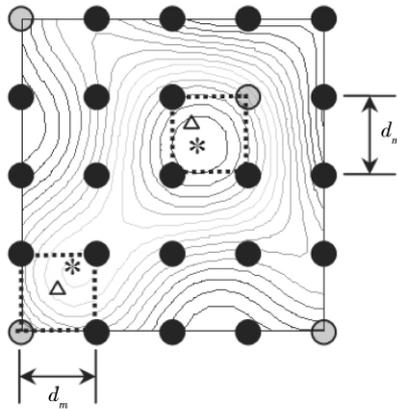
若支持向量的数量较多,则需要并行运行多次SQP算法,耗时较多,因此需要提高算法效率。进一步分析,如果在某个极值点附近存在多个支持向量,那么以这些支持向量为起点进行寻优,都将收敛至该极值点。因此可以首先对支持向量集合进行聚类预处理,将极值点附近的若干支持向量聚为一类,形成若干个聚类,然后再以各聚类中心为起点进行SQP寻优,这样既可以减少寻优次数,又可以发现多个极值。

聚集分层聚类分析^[27-28]可以作为实现这一设想的工具,其基本原理是根据某些准则,将样本集分为若干聚类,每个聚类中的样本具有较高相似程度,而不同聚类之间的样本相似程度则较低。聚类分析算法涉及到连接方式、相似度水平以及聚类数目的确定。根据质量优化及SVM拟合方程的特点,这里采用欧氏距离来度量样本之间的距离;采用完全连接法,即由两个聚类中相距最远的两个样本的距离来确定聚类与聚类之间的距离。而对于最终聚集数目,则可以通过选择聚类的最小相似度水平来确定。

具体来说,设过程有 m 个参数,采用均匀空间网格实验设计获取样本,空间网格的最小间距设为 d_m 。对于过程的任一极值,将其投影在参数的可行域上,在以其投影为中心的邻域内存在一系列空间网格单元,而其中最小的单元长度就是 d_m ,此单元由 2^m 个样本点包围而成,其中至多存在 2^m 个支持向量。以这些支持向量的中心为起点进行SQP寻优,将以较大的概率收敛至该极值点。因此,通过对聚类过程谱系图的分析,选择合适的最小相似度水平,使最终形成的各聚类所含支持向量的数目不大于 2^m ,这样既可以降低寻优的迭代次数,又不至于错过各个极值。

例如图1所示的含有2个参数(即 $m=2$)的

过程等高线图(相当于过程质量特性在参数的可行域上的投影图),两个极大值点在参数的可行域上的投影分别以“*”表示.可以看出,包含每个极值点的投影的最小单元网格,均由 $2^2 = 4$ 个样本点组成.或者说,在过程极值附近的较小的邻域内,若这4个样本均为支持向量样本,则这些样本点的中心,即为SQP算法的初始迭代点.



● : 非支持向量 ● : 支持向量 * : 极值点
 ... : 极值点邻域内最小单元网格(长度为 d_m)的边界线
 Δ : 最小单元网格包含的支持向量中心点(SQP初始迭代点)

图 1 SQP 算法的初始迭代点的确定示意图

Fig. 1 Determination of initial iterating points in SQP

2.3.2 实现步骤

根据上述分析,提出基于SVM拟合方程的全局性参数优化方法如下.

步骤 1 确定过程参数向量 x 的约束范围.

在实际制造过程的参数优化中 x 必须首先满足可行域的约束,设 x_{\min} x_{\max} 分别代表可行域的上下限,则 x 须满足 $x_{\min} \leq x \leq x_{\max}$.若各参数间存在一定的约束条件,则 x 还应满足线性或非线性约束条件 $c_i(x) \geq 0, i = 1, 2, \dots, m;$

步骤 2 根据质量特性优化的要求,拟合SVM方程,识别出样本集中的支持向量集合.

1) 对于望大型问题,由样本集(1)及2.2的步骤得到过程的SVM拟合方程,并识别出其中的支持向量样本集合,其中支持向量的数目为 l 个

$$S_{sv} = \{ (x_{sv_1}, y_{sv_1}), (x_{sv_2}, y_{sv_2}), \dots, (x_{sv_l}, y_{sv_l}) \} \quad (12)$$

2) 对于望小型问题,在样本集(1)中,首先令

$$y'_i = -y_i, i = 1, 2, \dots, n \quad (13)$$

可将其转化为望大型问题

3) 对于望目型问题,设过程输出的质量特性需达到的目标值为 T ,令

$$y'_i = -(y_i - T)^2, i = 1, 2, \dots, n \quad (14)$$

从而形成新的样本集 S' ,即可将原有问题转化为望大型问题.

步骤 3 结合 x 的约束条件,形成如下的约束优化问题

$$\begin{aligned} \max_{\forall x^*} f(x) &= \sum_{i=1}^l (\bar{\alpha}_i^* - \bar{\alpha}_i) \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) + b \\ \text{s. t. } &\begin{cases} x_{\min} \leq x \leq x_{\max} \\ c_i(x) \geq 0; i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (15)$$

式中, $\forall x^*$ 表示需要寻找到过程所有的局部极大值点.

步骤 4 对式(14)中 $f(x)$ 的支持向量样本集进行聚类分析,形成多个聚类,计算各聚类的几何中心点.

1) 生成参加聚类的样本集.对于 l 个支持向量,则参加聚类的样本集 C_{sv} 定义为

$$C_{sv} = \{ (x_{sv_i}, y_{sv_i}), i = 1, 2, \dots, l \} \quad (16)$$

2) 以欧氏距离 d_{ij} 来度量样本点与样本点之间的距离 d_{ij} 的表达式为

$$d_{ij} = [\|x_{sv_i} - x_{sv_j}\|^2 + (y_{sv_i} - y_{sv_j})^2]^{1/2} \quad (17)$$

此外,采用完全连接法来度量两个聚类之间的距离,即以两个聚类中样本之间的最大距离作为两个聚类之间的距离;

3) 利用聚集分层聚类法对 C_{sv} 进行聚类分析,根据聚类分析结果和谱系图,选择合适的最小相似性水平,使得最终聚类中各聚类所包含的最大样本数不超过 2^m ,由此得到 p 个聚类

$$C_{sv_1}, C_{sv_2}, \dots, C_{sv_p} \quad (18)$$

4) 计算每个聚类 $C_{sv_i}, i = 1, 2, \dots, p$ 在 x 维度上的几何中心,然后以各几何中心为初始迭代点,并行运行SQP算法,得到若干极值点;

5) 对SQP寻优得到的极值点进行分析,删除其中重复的极值点,得到过程不同的极值点

$$Opt = \{ (x_{opt_1}, y_{opt_1}), (x_{opt_2}, y_{opt_2}), \dots, (x_{opt_q}, y_{opt_q}) \} \quad (19)$$

3 案例研究

3.1 算例介绍

选择一个较为复杂的仿真函数,通过对其寻

优来考察方法的有效性. 假设过程输出的质量特性 y (望大型) 与参数 $\mathbf{x} = [x_1, x_2]^T$ 存在如下关系

$$y = 2 + 4x_1 + 4x_2 - x_1^2 - x_2^2 + \theta_1 \sin(\theta_2 x_1) \sin(\theta_2 x_2) + \varepsilon \quad (20)$$

式中: θ_1, θ_2 为可变系数, 用以改变过程的极值数目; $\varepsilon \sim N(0, 0.01^2)$, 为过程噪声; $\mathbf{x} = [x_1, x_2]^T$ 代表参数的组合, 且 $\mathbf{x}_{\min} = [-1, -1]^T, \mathbf{x}_{\max} = [1, 1]^T$. 图2分别给出了 $[\theta_1, \theta_2] = [2, 2], [\theta_1, \theta_2] = [3, 2.5], [\theta_1, \theta_2] = [4, 3]$ 的过程曲面形状以及等高线图, 其过程极大值的数目分别为 2、5、8. 可以看出, 这是 3 个典型的多极值过程.

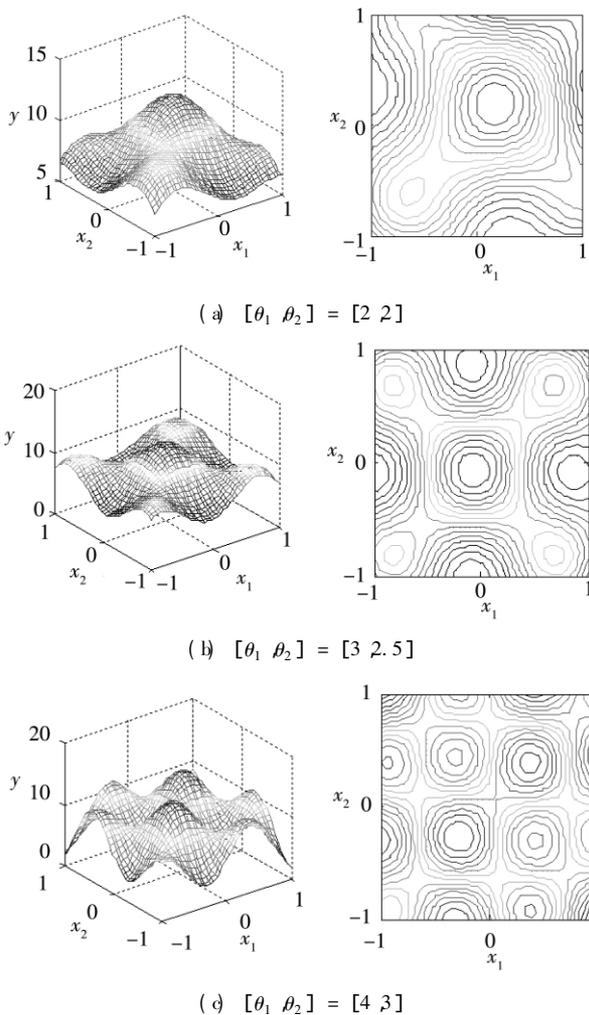


图2 不同 θ_1, θ_2 时的过程曲面形状图及等高线图

Fig.2 Surface and contour plots of processes with different θ_1, θ_2

3.2 评价指标

1) 绝对偏差 d_{xy} . 将寻优得到的极值点与过程实际极值点之间的欧氏距离作为度量寻优效果的绝对指标, 即

$$d_{xy} = \sqrt{\|\mathbf{x}_{\text{opt}} - \mathbf{x}^*\|^2 + (y_{\text{opt}} - y^*)^2} \quad (21)$$

2) 相对偏差 D_y (%). 为进一步考察方法的效果, 引入关于质量特性 y 的相对偏差指标

$$D_y = \left| \frac{y_{\text{opt}} - y^*}{y^*} \right| \times 100\% \quad (22)$$

3) 偏差的比例指标 P_x, P_y . 由于

$$d_{xy}^2 = d_x^2 + d_y^2 \quad (23)$$

于是可以令

$$\begin{cases} P_x = \frac{d_x^2}{d^2} \times 100\% \\ P_y = \frac{d_y^2}{d^2} \times 100\% \end{cases} \quad (24)$$

以考察在参数 (x 维度) 与质量特性 (y 维度) 的偏差占总偏差的比例, 且有

$$P_x + P_y = 1 \quad (25)$$

4) 标准泛化误差 SEP (%). 选择与质量特性 y 的度量无关的标准泛化误差 (standard error of prediction, SEP) 来评价模型的预测性能

$$SEP = \frac{1}{y} \sqrt{\frac{\sum_{i=1}^l (\hat{y}_i - y_i)^2}{l}} \times 100\% \quad (26)$$

其中: l 代表样本量; \hat{y}_i 代表质量特性的估计值; y_i 代表质量特性的实际值; \bar{y} 代表所有质量特性实际值的均值.

3.3 仿真结果

对于图2所示的3种过程, 为使建模的预测误差 SEP 尽可能接近, 各采用了3种间距的空间网格进行建模, 其样本量分别为 (36, 49, 64), (49, 64, 81), (64, 81, 100). 表1给出了过程实际极值 y^* 及其对应的参数值 \mathbf{x}^* , 不同样本量时的 d_{xy}, D_y, P_x, P_y ; 表2给出了不同样本量时的 SEP 值, 支持向量数, 聚类数, 最小相似度水平, 聚类后所减少的 SQP 寻优次数的比率 d_{xy}, D_y, P_x, P_y 的均值及总均值.

为详细说明算法的聚类及寻优的实现过程, 对于 $[\theta_1, \theta_2] = [3, 2.5]$ 的过程 (如图2(b)所示), 以样本量 $n = 49$ 为例, 表3给出了建模形成的40个支持向量及对应的聚类别别, 聚类中心, 相似度水平, 以及聚类前后的寻优结果; 图3给出了聚类过程的谱系图; 图4在过程等高线图上给出了最终15个聚类的分布及其中心, 寻优得到的极值点的分布.

表1 不同 θ_1, θ_2 及不同样本量时过程的实际极值与寻优结果的偏差
Table 1 Actual extremes and corresponding optimization deviations with different θ_1, θ_2 and sample size

$[\theta_1, \theta_2]$	x^*		y^*	d_{xy}	D_y (%)	P_x (%)	P_y (%)	D_{xy}	D_y (%)	P_x (%)	P_y (%)	D_{xy}	D_y (%)	P_x (%)	P_y (%)
				$n = 36$				$n = 49$				$n = 64$			
[2, 2]	-0.614 4	-0.614 4	9.689 6	0.25	1.24	77.41	22.59	0.12	1.01	28.75	71.25	0.10	0.97	8.01	91.99
	0.189 4	0.189 4	11.797 0	0.09	0.66	30.10	69.90	0.04	0.32	9.85	90.15	0.04	0.37	1.63	98.37
[3, 2.5]				$n = 49$				$n = 64$				$n = 81$			
	-0.819 2	-0.819 2	9.613 3	0.58	5.85	5.45	94.55	0.36	3.54	8.49	91.51	0.20	1.97	13.44	86.56
	-0.821 0	0.682 4	10.131 6	0.28	2.65	10.75	89.25	0.15	1.33	20.70	79.30	0.05	0.17	88.01	11.99
	0.682 4	-0.821 0	10.131 6	0.30	2.82	9.03	90.97	0.13	1.14	25.39	74.61	0.06	0.31	67.98	32.02
	0.683 9	0.683 9	10.651 0	0.08	0.18	93.68	6.32	0.13	1.09	13.95	86.05	0.15	1.42	2.74	97.26
	-0.069 3	-0.069 3	12.976 1	0.11	0.87	0.80	99.20	0.06	0.46	0.01	99.99	0.05	0.40	0.50	99.50
[4, 3]				$n = 64$				$n = 81$				$n = 100$			
	1.000 0	1.000 0	8.595 5	0.09	1.10	0.00	100.00	0.09	1.10	0.00	100.00	0.06	0.67	0.00	100.00
	-0.930 5	-0.930 5	9.874 1	0.32	3.11	9.29	90.71	0.33	3.23	8.68	91.32	0.33	3.18	8.93	91.07
	-0.2691	1.0000	11.0955	0.20	1.76	0.00	100.00	0.07	0.67	0.06	99.94	0.09	0.83	0.00	100.00
	1.000 0	-0.269 1	11.095 5	0.18	1.62	0.01	99.99	0.10	0.86	0.02	99.98	0.08	0.73	0.00	100.00
	-0.931 7	0.389 7	11.574 2	0.16	1.27	18.48	81.52	0.24	1.96	8.54	91.46	0.21	1.80	2.70	97.30
	0.389 7	-0.931 8	11.574 2	0.17	1.37	16.39	83.61	0.25	2.10	7.58	92.42	0.20	1.71	5.02	94.98
	0.390 2	0.390 2	13.276 5	0.15	1.10	2.15	97.85	0.09	0.65	1.64	98.36	0.06	0.48	0.56	99.44
-0.271 0	-0.271 0	13.651 0	0.31	2.27	0.02	99.98	0.09	0.65	0.07	99.93	0.10	0.75	0.06	99.94	

表2 不同 θ_1, θ_2 及不同样本量时的聚类及寻优结果的均值
Table 2 Averages of cluster and optimization with different θ_1, θ_2 and sample size

$[\theta_1, \theta_2]$	n	过程极值数	SEP (%)	支持向量数	最小相似度水平 (%)	聚类数(SQP寻优次数)	聚类后SQP次数的降低率 (%)	d_{xy} 均值	D_y 均值 (%)	P_x 均值 (%)	P_y 均值 (%)
[2, 2]	36	2	2.46	32	62.00	11	65.63	0.17	0.95	53.75	46.25
	49	2	1.35	33	68.10	13	60.61	0.08	0.66	19.30	80.70
	64	2	0.94	41	73.00	18	56.10	0.07	0.67	4.82	95.18
[3, 2.5]	49	5	2.51	40	73.57	15	62.50	0.27	2.47	23.94	76.06
	64	5	1.79	46	70.00	13	71.74	0.17	1.51	13.71	86.29
	81	5	1.20	55	76.00	23	58.18	0.10	0.85	34.53	65.47
[4, 3]	64	8	2.49	54	76.00	22	59.26	0.20	1.70	5.79	94.21
	81	8	1.60	57	73.00	22	61.40	0.16	1.40	3.32	96.68
	100	8	1.25	61	83.00	30	50.82	0.14	1.27	2.16	97.84
总均值							60.69	0.15	1.28	17.92	82.08

为了对比算法对于过程多极值的遍历性,表4给出了基于支持向量聚类的SQP算法、随机选取初始点的SQP算法以及遗传算法最终寻优结果的分布.其中,随机选取初始点的SQP算法利用均匀分布在可行域内随机产生初始迭代点.遗传算法以SVM拟合方程作为适应度函数,采用均匀分布的方式在可行域内生成初始种群,种群数为20;采用随机遍历法进行选择操作、离散重组方式进行基因交叉,以及Gaussian函数作为变异

函数.对比时,以基于支持向量聚类的SQP算法所需的运行次数(也就是支持向量的聚类数目)为基准,对比其他两种算法在相同运行次数下收敛于不同极值数目的概率(重复100轮以计算概率),以此考察算法的遍历性.例如,当 $[\theta_1, \theta_2] = [3, 2.5]$ 时,过程共有5个极值,在样本量 $n = 49$ 时,SVM拟合方程的支持向量形成了15个聚类,因此,各种算法均运行15次 \times 100轮;具体来说,基于支持向量的SQP算法在100轮重复时,每轮

的 15 次迭代均发现了 5 个极值点; 而随机选取初始点的 SQP 算法在 100 轮的重复下, 以 1% 的概率发现了 3 个极值点, 30% 的概率发现了 4 个极值

点, 69% 的概率发现了 5 个极值点; 遗传算法在 100 轮的重复下, 以 97% 的概率发现了 1 个极值点, 3% 的概率发现了 2 个极值点.

表 3 $[\theta_1 \ \theta_2] = [3 \ 2.5]$ $n = 49$ 时的支持向量聚类以及聚类前后的寻优结果

Table 3 Clusters of support vectors and optimization results fore-and-aft clustering with $[\theta_1 \ \theta_2] = [3 \ 2.5]$ and $n = 49$

支持向量 编号	支持向量			未经聚类的寻优结果			聚类 类别	相似度 (%)	聚类中心			聚类后寻优结果		
1	-1.000 0	-1.000 0	8.189 8	-0.061 6	-0.062 8	12.863 8	C_{sv_1}	87.07	-0.833 3	-0.833 3	8.544 5	-0.724 5	-0.722 3	9.051 0
2	-1.000 0	-0.666 7	8.465 8	-0.724 6	-0.722 3	9.051 0								
8	-0.666 7	-1.000 0	8.439 7	-0.725 2	-0.722 5	9.051 0								
9	-0.666 7	-0.666 7	9.082 6	-0.724 4	-0.722 3	9.051 0								
3	-1.000 0	-0.333 3	5.873 0	-0.061 6	-0.062 8	12.863 8	C_{sv_2}	86.02	-1.000 0	-0.166 7	5.447 3	-0.724 5	-0.722 3	9.051 0
4	-1.000 0	0.000 0	5.021 6	-0.753 9	0.617 3	9.862 8								
5	-1.000 0	0.333 3	7.421 6	-0.754 1	0.617 5	9.862 8	C_{sv_3}	73.59	-0.888 9	0.777 8	7.527 6	-0.754 0	0.617 4	9.862 8
7	-1.000 0	1.000 0	7.289 1	-0.754 0	0.617 3	9.862 8								
13	-0.666 7	1.000 0	7.872 2	-0.754 1	0.617 4	9.862 8								
6	-1.000 0	0.666 7	9.417 9	-0.754 1	0.617 4	9.862 8	C_{sv_4}	73.88	-0.666 7	0.500 0	9.343 6	-0.754 0	0.617 4	9.862 8
11	-0.666 7	0.333 3	8.699 6	-0.754 1	0.617 4	9.862 8								
12	-0.666 7	0.666 7	9.705 8	-0.754 1	0.617 4	9.862 8								
17	-0.333 3	0.333 3	9.551 0	-0.061 6	-0.062 8	12.863 8								
10	-0.666 7	-0.333 3	7.718 5	0.619 2	-0.756 9	9.846 4	C_{sv_5}	83.50	-0.500 0	-0.500 0	7.718 1	-0.061 7	-0.062 7	12.863 8
15	-0.333 3	-0.666 7	7.717 7	-0.754 0	0.617 3	9.862 8								
14	-0.333 3	-1.000 0	5.882 9	-0.061 6	-0.062 7	12.863 8	C_{sv_6}	86.01	-0.166 7	-1.000 0	5.456 5	-0.724 7	-0.722 6	9.051 0
20	0.000 0	-1.000 0	5.030 1	0.619 3	-0.756 9	9.846 4								
16	-0.333 3	0.000 0	11.377 3	-0.061 6	-0.062 8	12.863 8	C_{sv_7}	83.50	-0.111 1	-0.111 1	11.843 6	-0.061 6	-0.062 8	12.863 8
21	0.000 0	-0.333 3	11.394 7	-0.061 6	-0.062 8	12.863 8								
22	0.000 0	0.000 0	12.758 6	-0.061 7	-0.062 8	12.863 8								
18	-0.333 3	0.666 7	7.148 1	-0.754 1	0.617 4	9.862 8	C_{sv_8}	86.62	-0.222 2	0.777 8	6.621 2	-0.754 0	0.617 4	9.862 8
19	-0.333 3	1.000 0	6.425 0	0.631 4	0.631 4	10.631 8								
23	0.000 0	0.666 7	6.290 6	-0.061 6	-0.062 8	12.863 8								
24	0.333 3	-1.000 0	7.401 3	0.619 4	-0.757 0	9.846 4	C_{sv_9}	100.00	0.333 3	-1.000 0	7.401 3	0.619 4	-0.757 0	9.846 4
25	0.333 3	-0.666 7	8.681 5	0.619 3	-0.756 8	9.846 4	$C_{sv_{10}}$	73.88	0.500 0	-0.666 7	9.335 5	0.619 4	-0.756 9	9.846 4
26	0.333 3	-0.333 3	9.549 6	-0.061 6	-0.062 7	12.863 8								
30	0.666 7	-1.000 0	9.423 0	0.619 3	-0.757 0	9.846 4								
31	0.666 7	-0.666 7	9.687 9	0.619 3	-0.756 9	9.846 4	$C_{sv_{11}}$	78.62	0.555 6	0.444 4	9.600 1	0.631 4	0.631 5	10.631 8
27	0.333 3	0.333 3	9.502 9	-0.061 6	-0.062 8	12.863 8								
34	0.666 7	0.333 3	8.659 0	0.631 4	0.631 5	10.631 8								
35	0.666 7	0.666 7	10.638 3	0.631 4	0.631 5	10.631 8	$C_{sv_{12}}$	83.96	0.333 3	0.833 3	8.041 0	0.631 4	0.631 5	10.631 8
28	0.333 3	0.666 7	8.649 9	0.631 4	0.631 5	10.631 8								
29	0.333 3	1.000 0	7.432 1	0.631 4	0.631 5	10.631 8	$C_{sv_{13}}$	86.62	0.777 8	-0.222 2	6.627 8	0.619 3	-0.757 0	9.846 4
32	0.666 7	-0.333 3	7.144 6	0.619 3	-0.756 9	9.846 4								
33	0.666 7	0.000 0	6.316 3	-0.061 6	-0.062 8	12.863 8								
38	1.000 0	-0.333 3	6.422 6	0.619 3	-0.756 9	9.846 4	$C_{sv_{14}}$	87.09	1.000 0	-0.833 3	7.569 6	-0.061 6	-0.062 8	12.863 8
36	1.000 0	-1.000 0	7.264 1	0.619 3	-0.756 9	9.846 4								
37	1.000 0	-0.666 7	7.875 1	0.619 3	-0.757 0	9.846 4	$C_{sv_{15}}$	75.67	1.000 0	0.666 7	7.048 6	-0.061 6	-0.062 8	12.863 8
39	1.000 0	0.333 3	7.428 4	0.631 4	0.631 5	10.631 8								
40	1.000 0	1.000 0	6.668 7	0.631 4	0.631 5	10.631 8								

表4 各类寻优算法对于多极值的遍历性对比

Table 4 Ergodicity about multi-extreme of different optimizing approaches

[θ_1, θ_2]	n	过程极值数	基准迭代次数 (支持向量聚类数)	寻优所发现的极值数(以100轮重复计算的概率)		
				基于支持向量聚类的SQP	随机选取初始点的SQP	遗传算法
[2, 2]	36	2	11	2(100%)	1(13%), 2(87%)	1(100%)
	49	2	13	2(100%)	1(7%), 2(93%)	1(100%)
	64	2	18	2(100%)	1(2%), 2(98%)	1(100%)
[3, 2.5]	49	5	15	5(100%)	3(1%), 4(30%), 5(69%)	1(97%), 2(3%)
	64	5	13	5(100%)	3(1%), 4(21%), 5(78%)	1(89%), 2(11%)
	81	5	23	5(100%)	3(1%), 4(14%), 5(85%)	1(88%), 2(12%)
[4, 3]	64	8	22	8(100%)	6(16%), 7(59%), 8(25%)	1(11%), 2(89%)
	81	8	22	8(100%)	5(1%), 6(25%), 7(55%), 8(19%)	1(10%), 2(84%), 3(6%)
	100	8	30	8(100%)	6(3%), 7(53%), 8(44%)	1(3%), 2(93%), 3(4%)

注: 各类算法的各轮寻优结果中均包含全局最优点.

表5 d_{xy} 、 D_y 、 P_x 、 P_y 与过程极值数目(n_{opt}) 的方差分析表

Table 5 ANOVA of d_{xy} 、 D_y 、 P_x 、 P_y versus n_{opt}

来源	自由度	平方和				均方				F 值				P 值			
		d_{xy}	D_y	P_x	P_y												
n_{opt}	2	0.008 89	1.235	908	908	0.004 44	0.617	454	454	1.19	2.51	1.83	1.83	0.368	0.161	0.239	0.239
误差	6	0.022 45	1.475	1 488	1 488	0.003 74	0.246	248	248								
总和	8	0.031 34	2.71	2 396	2 396												

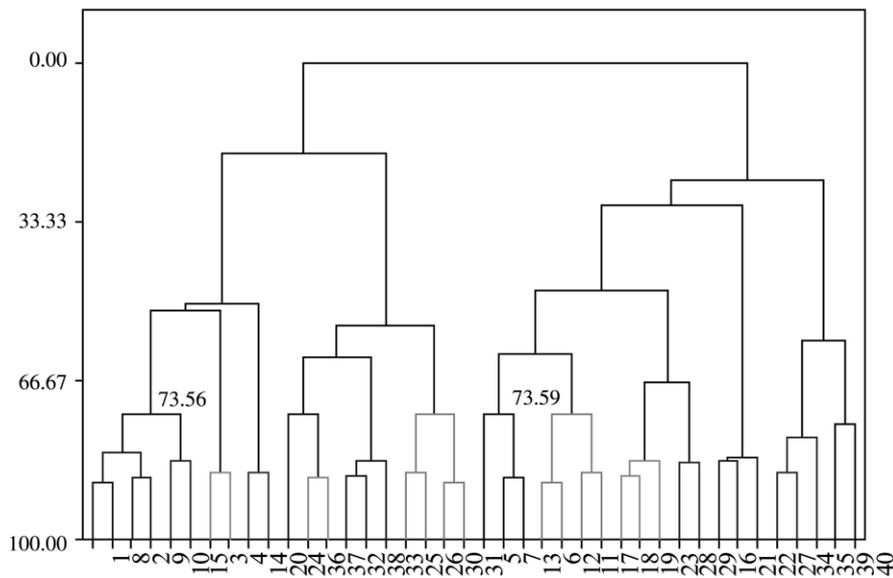


图3 40个支持向量聚类过程的谱系图

Fig. 3 Clustering dendrogram of 40 support vectors

4 结果讨论

4.1 基于支持向量聚类的SQP算法寻优结果分析

1) 表1和表2说明, 在不同样本量 n 、不同 θ_1 、

θ_2 的情况下, 绝对偏差 d_{xy} (总均值为0.15) 与相对偏差 D_y 均较小(总均值为1.28%), 且 d_{xy} 与预测误差 SEP 之间具有较强的正线性相关关系, 其 Pearson 相关系数为0.784 (P 值 = 0.012). 即 SEP 越小, d_{xy} 越小. 为考察不同的过程极值数对于 d_{xy}

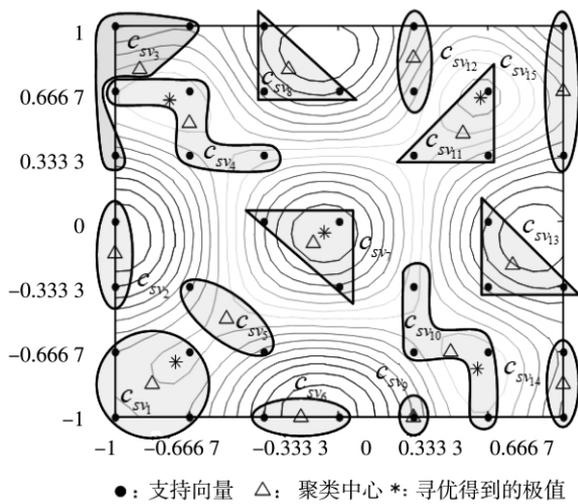


图 4 15 个聚类类别的分布及寻优结果

Fig. 4 Distributions and optimization results of 15 clusters

的影响, 将 d_{xy} 与“过程极值数 (n_{opt})”作方差分析 (如表 5 所示). 由于 P 值 = 0.368 (> 0.05), 说明在不同的过程极值数情况下, d_{xy} 没有显著性的差别; 同样 D_y (P 值 = 0.161)、 P_x (P 值 = 0.161)、 P_y (P 值 = 0.161) 也没有显著性的差别. 或者说, d_{xy} 、 D_y 、 P_x 、 P_y 不随过程极值数的增大而增大, 说明了方法的稳定性;

2) 表 1 和表 2 说明, 寻优结果与过程实际极值之间的偏差主要是由于 y 维度上的偏差 P_y (总均值为 82.08%) 造成的, 而在 x 维度 (即参数维度) 上的偏差 P_x (总均值为 17.92%) 则较小. 说明寻优可以得到较为准确的参数设置, 而这进一步保证了能够得到最优的质量特性, 因此对于生产实践具有较大的实际意义;

3) 表 3 说明, 经过支持向量聚类后再进行 SQP 寻优与未经支持向量聚类直接进行 SQP 寻优, 所发现的极值是一致的, 都寻找到了过程的全部 5 个极值; 进一步从表 2 可以看出, 经过聚类, SQP 并行寻优的次数大幅下降, 平均降低 60% 以上, 且至少降低 50% 以上. 因此, 经过支持向量聚类以后, 不仅能够发现全部极值, 而且在较大程度上提高了寻优效率.

4.2 聚类相似度水平的确定及其对寻优结果的影响

1) 对于由图 2(b) 所示的过程, 其聚类的相似度水平的确定如下: 由于参数的数目 $m = 2$, 则每个聚类中支持向量的数目不应超过 $2^m = 4$; 考察

图 3 相似度水平最初为 100% 40 个支持向量形成 40 个聚类. 随着相似度水平的减小, 聚类数目在减少, 各聚类所包含支持向量数目则在增加; 相似度水平减小至 73.56% 时, 首次出现了样本数目超过 4 的聚类 (例如编号为“1”、“8”、“2”、“9”、“10”、“15”的支持向量组成的聚类), 因此, 相似度水平应大于 73.56%. 当相似度水平增加至 73.59% 以上时, 原有的某些聚类将被打破 (例如编号为“5”、“7”、“13”的支持向量组成的聚类将会被打散为两个聚类), 因此相似度水平也不宜大于 73.59%, 否则会增加 SQP 寻优次数. 综合判断, 最小相似度水平应在 (73.56%, 73.59%], 而此范围内的任意相似度水平, 其聚类结果都一致, 本例将聚类的最小相似度水平选为 73.57%;

2) 图 4 显示, 在最小相似度水平为 73.57% 时, 40 个支持向量形成了 15 个聚类. 这 15 个聚类可以分为两种类型: 第 1 种类型是极值点附近的聚类. 过程 5 个极值点的邻域内的支持向量均聚为了一类 (例如聚类 C_{sv1} 、 C_{sv4} 、 C_{sv7} 、 C_{sv10} 、 C_{sv11}), 由其聚类中心进行 SQP 寻优, 收敛至对应的极值点. 这说明在该最小相似度水平下进行聚类, 能够保证寻优发现过程全部的极值, 而不至于错过某些极值; 第 2 种类型是远离极值点的其他支持向量形成的聚类 (例如聚类 C_{sv2} 、 C_{sv3} 、 C_{sv5} 、 C_{sv6} 、 C_{sv8} 、 C_{sv9} 、 C_{sv12} 、 C_{sv13} 、 C_{sv14} 、 C_{sv15}). 由其聚类中心出发运行 SQP 也收敛到了相应的 5 个极值点. 因此, 最小相似度水平的选择是合适的.

4.3 各类寻优算法对于多极值的遍历性对比分析

1) 表 4 说明, 针对不同的 θ_1 、 θ_2 以及不同的样本量 n , 在基准迭代次数下, 各类算法的各轮寻优均发现了过程的全局最优点, 但是并非都能够发现过程的全部极值. 具体来说, 基于支持向量聚类的 SQP 算法均能够以 100% 的概率发现过程所有的极值点, 遗传算法均不能够发现过程的全部极值点, 而随机选取初始点的 SQP 的性能介于二者之间;

2) 表 4 说明, 过程极值数显著影响算法发现过程全部极值的能力. 当过程极值数较少 ($n_{opt} = 2$) 时, 遗传算法只能收敛于过程的单个极值; 随机选取初始点的 SQP 算法, 虽然发现过程全部极值

的概率较高(其均值为 92.7%),但是仍不能确保 100% 的概率.随着过程极值数的增多($n_{opt} = 5$ 和 $n_{opt} = 8$)遗传算法发现过程多个极值的能力增加(最多时,可以 12% 的概率发现过程所有 5 个极值中的 2 个),但是仍不能发现过程的全部极值;而随机选取初始点的 SQP 算法发现过程全部极值的概率则在显著下降(其均值分别为 77.3% 和 29.3%).因此,基于支持向量聚类的 SQP 算法的性能对于过程极值的数不敏感,而遗传算法、随机选取初始点的 SQP 算法的性能则显著受到过程极值数的影响;

3) 上述对比说明,遗传算法这一随机性的寻优算法虽然可以在可行域全局范围内寻找极值,但是由于搜索的随机性,不论初始种群如何选择,算法总倾向于跳出局部最优而达到某个特定的最优值点,因此并不利于发现过程的全部极值;另一方面, SQP 这一确定性的寻优算法受初始点位置影响较大.然而随机选取初始点并不能使算法必然收敛于不同的极值;但是,通过有目的的采用支持向量聚类的方式来选择初始点,可使 SQP 算法

必然收敛于过程的所有极值.因此,通过合理地选择 SQP 算法的初始迭代点,有效地提高了 SQP 算法发现过程全部极值的能力.

5 结束语

本文研究了多极值复杂作用关系过程的参数全局性优化问题,在建立 SVM 拟合模型的基础上,选择合适的最小相似度水平,对支持向量进行聚类,再利用 SQP 算法进行并行寻优.仿真分析表明,所提方法能够有效地寻找到过程的全部极值;而且通过支持向量聚类,在有效地降低 SQP 寻优次数的同时,也保证了寻优的遍历性.同时,寻优所得极值与实际极值的偏差较小,其中在参数维度上的偏差的比例也较小,而且偏差与过程极值数无关.说明了此方法的有效性与稳定性.本文的研究也表明,远离极值点的支持向量形成的聚类对于最终的寻优结果不是必备的.因此,根据这一特点,如何在保证寻优性能的前提下,降低建模及寻优所需样本量,将是本文进一步的研究方向.

参 考 文 献:

- [1] 夏国洪,李伯虎,唐晓青,等.复杂系统(产品)集成制造工程的技术研究与应用[J].中国工程科学,2005,7(9): 49-55.
Xia Guohong, Li Bohu, Tang Xiaqing, et al. Research and application of technologies for complex system (product) integrated manufacturing engineering[J]. Engineering Science, 2005, 7(9): 49-55. (in Chinese)
- [2] Ilzarbe L, Alvarez M J, Viles E, et al. Practical applications of design of experiments in the field of engineering: A bibliographical review[J]. Quality and Reliability Engineering International, 2008, 24(4): 417-428.
- [3] Mariam I, Nagra S A, Haq I, et al. Application of 2-factorial design on the enhanced production of calcium gluconate by a mutant strain of aspergillus niger[J]. Bioresource Technology, 2010, 101(11): 4075-4080.
- [4] Myers R H, Montgomery D C, Anderson-Cook C M. Response Surface Methodology: Process and Product Optimization Using Experiments[M]. New York: John Wiley and Sons, 2009.
- [5] Gan C C, Latiff A A. Optimisation of the solvent extraction of bioactive compounds from Parkia speciosa pod using response surface methodology[J]. Food Chemistry, 2011, 124(3): 1277-1283.
- [6] Fang K T, Lin D K. Uniform Design in Computer and Physical Experiments[M]//Tsubaki H, Nishina K, Yamada S. The Grammar of Technology Development, Berlin: Springer, 2008: 105-125.
- [7] Wang J, Liu J, Zhang G, et al. Orthogonal design process optimization and single factor analysis for bimodal acoustic agglomeration[J]. Powder Technology, 2011, 210(3): 315-322.
- [8] Wang G G, Shan S. Review of metamodeling techniques in support of engineering design optimization[J]. Journal of Mechanical Design, 2007, 129(4): 370-380.
- [9] Muller H W, Sperlich M, Werwatz S. Nonparametric and Semiparametric Models[M]. Berlin: Springer, 2004.
- [10] Robinson T J, Birch J B, Alden S B. A semi-parametric approach to dual modeling when no replication exists[J]. Journal of Statistical Planning and Inference, 2010, 140(10): 2860-2869.
- [11] Chen W C, Tai P H, Wang M W, et al. A neural network-based approach for dynamic quality prediction in a plastic injection molding process[J]. Journal of Intelligent and Manufacturing, 2011, 22(4): 417-428.

- tion molding process [J]. *Expert Systems with Applications*, 2008, 35(3): 843–849.
- [12] Sathyanarayananurthy H, Chinnam R B. Metamodels for variable importance decomposition with applications to probabilistic engineering design [J]. *Computers & Industrial Engineering*, 2009, 57(3): 996–1007.
- [13] Alessandri A, Cassettari L, Mosca R. Nonparametric nonlinear regression using polynomial and neural approximators: A numerical comparison [J]. *Computational Management Science*, 2009, 6(1): 5–24.
- [14] Kleijnen J P. Kriging metamodeling in simulation: A review [J]. *European Journal of Operational Research*, 2009, 192(3): 707–716.
- [15] 黄章俊, 王成恩. 基于 Kriging 模型的涡轮盘优化设计方法 [J]. *计算机集成制造系统*, 2010, 16(5): 905–911. Huang Zhangjun, Wang Chengen. Turbine discs optimization design based on Kriging model [J]. *Computer Integrated Manufacturing Systems*, 2010, 16(5): 905–911. (in Chinese)
- [16] 崔庆安, 何 桢, 崔 楠. 基于 SVM 的 RSM 模型拟合方法研究 [J]. *管理科学学报*, 2008, 11(1): 31–41. Cui Qingan, He Zhen, Cui Nan. SVM-based RSM model fitting approach [J]. *Journal of Management Sciences in China*, 2008, 11(1): 31–41. (in Chinese)
- [17] Cui Q A. An SVR and Single Response Modeling Strategy Based RPD Approach for Multi-Extreme Processes [C]// *Proceedings of the IEEE 17th International Conference on Industrial Engineering and Engineering Management*, Vol A, Xiamen: 2010.
- [18] Andrieu L, Cohen G, Vazquez-Abad F J. Gradient-based simulation optimization under probability constraints [J]. *European Journal of Operational Research*, 2011, 212(2): 345–351.
- [19] 王 勇, 蔡自兴, 周育人, 等. 约束优化进化算法 [J]. *软件学报*, 2009, 20(1): 11–29. Wang Yong, Cai Zixing, Zhou Yuren, et al. Constrained optimization evolutionary algorithms [J], *Journal of Software*, 2009, 20(1): 11–29. (in Chinese)
- [20] Liu H, Maghsoodloo S. Simulation optimization based on Taylor Kriging and evolutionary algorithm [J]. *Applied Soft Computing*, 2011, 11(4): 3451–3462.
- [21] Yaochu J. Surrogate-assisted evolutionary computation: Recent advances and future challenges [J]. *Swarm and Evolutionary Computation*, 2011, 1(2): 61–70.
- [22] Vapnik V N. *The Nature of Statistical Learning Theory* (2nd Edition) [M]. Berlin: Springer, 1999.
- [23] Barakat N, Bradley A P. Rule extraction from support vector machines: A review [J]. *Neurocomputing*, 2010, 74(1/3): 178–190.
- [24] 姚宝珍, 杨成永, 于 滨. 动态公交车辆运行时间预测模型 [J]. *系统工程学报*, 2010, 25(3): 365–370. Yao Baozhen, Yang Chengyong, Yu Bin. Dynamic bus travel-time prediction model [J]. *Journal of Systems Engineering*, 2010, 25(3): 365–370. (in Chinese)
- [25] Jian J B, Chao M T. A sequential quadratically constrained quadratic programming method for unconstrained minimax problems [J]. *Journal of Mathematical Analysis and Applications*, 2010, 362(1): 34–45.
- [26] Byrd R H, Curtis F E, Noceda J. Infeasibility detection and sqp methods for nonlinear optimization [J]. *SIAM Journal on Optimization*, 2010, 20(11): 2281–2299.
- [27] Agarwal P, Alam M A, Biswas R. A hierarchical clustering algorithm for categorical attributes [C]// *Computer Engineering and Applications (ICCEA) 2010 Second International Conference on*, Vol 2, Bali Island: 365–368.
- [28] 刘向丽, 程 刚, 成思危, 等. 中国期货市场价格久期波动聚类特征研究 [J]. *管理科学学报*, 2010, 13(5): 72–81. Liu Xiangli, Cheng Gang, Cheng Siwei, et al. Research on China futures market price duration volatility cluster effect [J]. *Journal of Management Sciences in China*, 2010, 13(5): 72–81. (in Chinese)

Global optimization of process parameters for multi-extreme quality characteristics

CUI Qing-an

Institute of Management Engineering, Zhengzhou University, Zhengzhou 450001, China

Abstract: For manufacturing processes whose output quality characteristics have multi-extremes and whose

(下转第 73 页)

$$U_c(c, H, I) = J_w(W, X) - bJ_H(W, X)$$

所以

$$J_{ww} = U_{cc} \frac{\partial c}{\partial W} + J_{HW}$$

从而由式(10)得

$$D_i^{k'} = - \frac{U_c^{k'} - bJ_H}{U_{cc}c_w + bJ_{HW}} \sum_{j=1}^n v_{ij}(\alpha_i - r) \quad \text{证毕.}$$

C 命题 1 的证明:

可以构造三“基金”: 一个是无风险资产, 一个是与外在性消费习惯完全负相关的组合, 一个是全部由无风险资产构成的, 且比例为下式的组合

$$\omega_i^k = \frac{\omega_i^k W^k}{W_r^k} = \frac{\sum_{j=1}^n v_{ij}(\alpha_i - r)}{\sum_{i=1}^n \sum_{j=1}^n v_{ij}(\alpha_i - r)} \quad \text{证毕.}$$

D 命题 2 的证明:

由方程(10)易证投资者 k 投资于第 i 个风险资产的财富为

$$\omega_i^k W^k = A^k \sum_{j=1}^n v_{ij}(\alpha_i - r)$$

投资于所有风险资产的财富为

$$W_r^k = \sum_{i=1}^n \omega_i^k W^k = A^k \sum_{i=1}^n \sum_{j=1}^n v_{ij}(\alpha_i - r)$$

所以, 投资于第 i 个风险资产的比例为

$$\omega_i^k = \frac{\omega_i^k W^k}{W_r^k} = \frac{\sum_{j=1}^n v_{ij}(\alpha_i - r)}{\sum_{i=1}^n \sum_{j=1}^n v_{ij}(\alpha_i - r)}$$

该式表明, 所有投资者投资于风险资产上的财富的分配方案完全相同, 这种方案等于市场投资组合. 证毕.

E 命题 4 的证明:

命题 4 的模型中对应的 Hamilton-Jacobi-Bellman 方程为

$$0 = \max \left\{ U(c, X) + J_t + J_w \{ W[\omega(\alpha - r) + r - c] \} + J_{XX} \mu_X X + \frac{1}{2} J_{ww} \omega^2 \sigma^2 + J_{wX} W X \omega \sigma_{1X} + \frac{1}{2} J_{XX} X^2 \sigma_X^2 \right\}$$

由一阶条件解得

$$\omega_i^* = \frac{1}{\gamma} \frac{\alpha - r}{\sigma^2} + \frac{\gamma - 1}{\gamma} \frac{\sigma_{1X}}{\sigma^2}$$

和

$$c_i^* = \eta W_i^*$$

代入 Hamilton-Jacobi-Bellman 方程得 η . 证毕.

(上接第 57 页)

input-output relationships are complex, existing approaches of quality improvement can only obtain a local optimization of the process parameters. Therefore, there is still great potential for product quality improvement. By adopting support vector machines (SVM) to approximately model the complex relationship processes, this article proposes a support vector clustering and sequential quadratic programming (SQP) based approach for global optimization of process parameters. Firstly, a SVM based approximation model for the complex processes is set up. Then, based on ε -tube theory, the minimal similarity level and hence the appropriate number of clusters are determined through analyzing the clustering dendrogram, and the support vectors in the neighborhood of each quality characteristics' extremes are clustered into one group. Lastly, by using the geometrical centers of each cluster as initial points, the multi-extreme values of the quality characteristics are found through concurrent SQP optimization. The simulation study shows that the proposed approach can effectively reflect the distributions of quality characteristics' extremes and achieve global optimization of the process parameters; the average absolute deviation and the average relative deviation of the optimization results from the actual extreme values are 0.15 and 1.28%, respectively, and these deviations are independent of the number of the quality characteristics' extremes, as demonstrates the high accuracy and stability of the approach. Moreover, after clustering the support vectors, not only the ergodicity of SQP optimization results for all of the process extremes has been ensured, but also the number of SQP optimization has been decreased by at least 50%, as increases the optimization efficiency of the approach.

Key words: multi-extreme quality characteristics; global optimization; support vector machines; cluster analysis; quality improvement