

混合 Kriging 代理模型的高维参数估计优化算法

王 红^{*1,2}, 王希诚^{1,3}, 李克秋¹

- (1. 大连理工大学 计算机科学与技术学院, 辽宁 大连 116024;
2. 大连东软信息学院 计算机系, 辽宁 大连 116023;
3. 大连理工大学 工业装备结构分析国家重点实验室, 辽宁 大连 116024)

摘要: 基于 Kriging 代理模型的优化算法对于解决函数计算昂贵的优化问题非常有效, 但不适用于高维参数的优化. 针对该问题, 提出了一个混合 Kriging 代理模型和多种优化技术的算法. 该算法在 Kriging 模型选择新样例点时使用单维参数独立优化以克服维度灾难并提高收敛速度, 同时基于已构建的 Kriging 代理模型信息提出一种新的动态坐标扰动策略, 并将该策略用于高维参数优化以得到更好的目标函数值. 为了保证不丢失全局最优解, 在使用一般期望提高加点策略作为选点原则时, 在期望函数的多个峰值同时选点. 为了验证算法的有效性, 将该算法应用于具有 41 维参数的人类白细胞代谢网络参数估计问题. 实验结果表明, 在有限的迭代次数下, 该算法能产生较小的目标函数值, 以及和实验拟合较好的参数估计结果.

关键词: Kriging 代理模型; 高维参数估计; 有限的计算资源; 一般期望提高

中图分类号: TP391.9 **文献标识码:** A **doi:** 10.7511/dllgxb201502015

0 引言

参数估计问题通常被当作优化问题来解决, 该优化问题致力于找到一组参数 p , 使用最小二乘法作为目标函数, 使得该组参数能对实验数据给出最好的拟合结果. 在此优化过程中每给定一组参数就要计算对应的函数值, 再和实验数据进行比较确定目标函数值.

在许多实际的工程参数估计问题中, 和参数对应的函数计算常常极其昂贵, 需要耗费大量的计算时间, 特别当优化问题含有多个局部最优值且待估计参数又是高维时, 要找到全局最优解就变得尤为困难. 此时比较现实的做法应该是, 在有限的计算资源条件(如确定的迭代次数或函数计算次数)下找到近似全局最优解, 本文提出的优化算法是服从于这个大前提的.

使用基于梯度的优化算法进行局部优化, 然

后通过多初始值方法找到全局最优解是参数估计最常采用的算法^[1]. 但当梯度需要使用数值差分计算时, 不仅会带来计算时间延长的问題, 还会由于函数不光滑使得计算得到的梯度不可信, 从而影响参数估计的准确性. 启发式的优化算法是进行参数估计的另一策略, 如模拟退火算法、遗传算法、粒子群算法等^[2-3]都曾用于参数估计领域. 此外分散搜索^[4]也常常用于全局优化以找到拟合最好的参数集合.

对于参数估计中含有昂贵的函数计算问题, 最自然的想法就是使用代理模型来代替昂贵的函数计算以期减轻计算负担. 常用的代理模型有响应面模型, 径向基函数、神经网络和 Kriging 模型. 在这些基于代理模型的优化算法中, 基于 Kriging 模型进行优化研究是应用最广泛的, 这主要归结于 Kriging 模型用于优化问题时不仅能给出优化结果, 而且能够给出优化的不确定性估计,

收稿日期: 2014-07-05; 修回日期: 2015-01-08.

基金项目: 国家自然科学基金资助项目(11072048); “九七三”国家重点基础研究发展计划资助项目(2009CB918501); 辽宁省教育厅科学研究一般项目(L2013519).

作者简介: 王 红^{*}(1974-), 女, 博士生, E-mail: wanghong@neusoft.edu.cn; 王希诚(1947-2013), 男, 博士, 教授, 博士生导师; 李克秋(1972-), 男, 博士, 教授, 博士生导师.

这种不确定性等价于 Kriging 代理模型的不确定性,同时该模型特别适用于解决非线性问题^[5].但是将 Kriging 代理模型直接用于高维参数优化领域却并不现实,因为参数的高维度意味着选择样例点要多,而样例点的增多会增加 Kriging 建模的计算时间.而且,高维参数之间复杂的相互作用也会使得各维度对目标函数的影响被削弱,表现为目标函数值不会随着优化次数的增多而有明显的下降.本文针对含有昂贵函数计算的高维参数估计问题,在 Kriging 优化框架下,使用一般期望提高(GEI)^[5]加点准则在期望提高(EI)^[6]函数的多个峰值搜索时,对每个维度参数独立优化以克服各维参数之间的相互影响及由此带来的维度灾难;同时基于已构建的 Kriging 模型信息提出一种新的动态坐标扰动策略,并应用该策略微扰样例点中目标值较小点的部分坐标以得到更好的样例点.考虑到 Kriging 建模时间会随着样例点数目增多而有所增加,因此使用李征^[7]提出的区域分割法以节省建模时间.

1 算法描述

1.1 基于 Kriging 的优化算法

一般的基于 Kriging 的优化算法流程如图 1 所示.

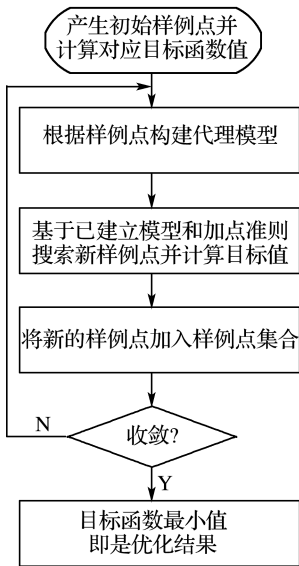


图 1 基于 Kriging 的优化算法流程图
Fig. 1 The flow chart of Kriging-based optimization algorithm

在该流程中,初始样例点的生成可以基于任

何取样技术,关键是如何保证所选样例点在所有参数空间均匀分布.常用的采样技术有拉丁超立方取样^[8]、极大熵取样^[9]、均匀取样^[10]等.

Kriging 代理模型通常表示为如下形式^[11]:

$$\begin{aligned}\hat{f}(x) &= \hat{\mu} + \mathbf{r}^T(x)\mathbf{R}^{-1}(\mathbf{f} - \mathbf{1}\hat{\mu}) \\ \hat{\mu} &= \mathbf{1}^T\mathbf{R}^{-1}\mathbf{f}/\mathbf{1}^T\mathbf{R}^{-1}\mathbf{1}\end{aligned}\tag{1}$$

式中: $\hat{f}(x)$ 是使用 Kriging 模型计算出来的预测值; $\hat{\mu}$ 是 Kriging 中值; $\mathbf{R}$ 是 $n \times n$ 矩阵,它在 (i, j) 处的元素为 $k(x(i), x(j))$,其中 k 是核函数; $\mathbf{r}$ 是一个 n 维向量,它的第 i 个元素是 $k(x, x(i))$; $\mathbf{f} = (f(x(1)) \ f(x(2)) \ \cdots \ f(x(n)))$,其中 $f(x(i))$ 表示点 $x(i)$ 处的响应值.

加点准则就是如何选择新的样例点的准则,该准则一般会作为选点过程的目标函数而存在. Kriging 是一种基于高斯过程的模型,因此可以估计出模型的不确定性.这一特征有助于确定在设计空间的哪些区域增加新样例点以提高模型的精度.常用的加点准则有期望提高准则,该加点准则不仅给出了可能提高的信息而且给出了这种提高尺度的度量,该准则如下:

$$E(I_n(x)) = \begin{cases} (y_{\min} - \hat{y}(x))\Phi((y_{\min} - \hat{y}(x))/s_n(x)) + s_n(x)\phi((y_{\min} - \hat{y}(x))/s_n(x)) & s_n > 0 \\ 0 & s_n = 0 \end{cases}\tag{2}$$

式中: $\Phi(\cdot)$ 是标准分布函数, $\phi(\cdot)$ 是可能密度函数, $y_{\min}$ 是样例点中的最小目标函数值, $\hat{y}(x)$ 是在 x 点的预测值, $s_n(x)$ 是在 x 点的均方误差.式(2)的第一部分表示目标函数值小且估计精确的区域,在该区域进行搜索即是在当前最小目标函数值点附近进行搜索,是一种局部搜索;式(2)的第二部分表示高度不确定区域,在此部分的搜索被认为是一种全局搜索. EI 最大加点准则就是搜索这样的点:要么预测值比当前最小目标函数值小的点,要么预测值有较大不确定性的点.可见,期望提高搜索能自动平衡局部与全局搜索两种趋势,因此 EI 最大化常常作为选择新样例点的目标函数.

1998 年 Jones 等^[12]指出,由于 Kriging 的常用工具箱 DACE 在其代码实现时低估了方差,加之初始建立的 Kriging 模型通常不精确,可能会使期望提高(EI)最大化搜索徘徊在已经探索过的局部极值区域.因此需要一种在优化开始就能进

行更广泛全局探索的优化算法,一般期望提高 (GEI) 准则应运而生. 该准则为期望提高加上了一个指数权, 其定义如下^[5]:

$$E(I^g) = s^g \sum_{k=0}^g (-1)^k \binom{g}{k} u^{g-k} T_k$$
$$T_0 = \Phi(x), T_1 = -\phi(x)$$
$$T_k = -x^{k-1} \phi(x) + (k-1) T_{k-2}$$

(3)

同 EI 公式定义一样, $\Phi(\cdot)$ 是标准分布函数, 而 $\phi(\cdot)$ 是可能密度函数. u 表示在 x 点的期望提高值, 被定义为 $u = \frac{y_{\min} - \hat{y}(x)}{s}$. 式(3) 中 g 因子起到在局部搜索和全局搜索之间平衡的作用: 大的 g 值暗示搜索会在更全局的区域进行, 而小的 g 值则表示该优化算法会探索当前最优值的附近区域.

基于 Kriging 的优化算法通常采用的收敛准则为: 在下一个待加点处期望提高的绝对值小于一个足够小的数(如 1%) 或是达到预先设定的最大迭代次数.

1.2 混合 Kriging 代理模型的高维参数估计优化算法

本文提出的优化算法是在 Kriging 的优化算法框架下提出的改进, 这种改进主要包括如下 3 点: (1) 以 GEI 为加点原则(即 GEI 最大为目标函数), 针对 GEI 定义设定多个 g 值, 在一次迭代中同时兼顾局部与全局搜索; (2) 单参数独立优化克服维度灾难并提高收敛速度, 使用序列二次规划方法针对每维参数单独优化; (3) 基于已建立的 Kriging 代理模型提出新的动态坐标扰动策略. 对所有样例点的目标函数值升序排列, 选取前 m 个点, 根据本文提出的动态坐标扰动策略对其坐标进行扰动. 此时流程图 1 中的第三部分变为图 2 所示. 图 2 中的 n 在算法中取 4, 即 g 分别设为 $\{0.25, 0.50, 0.75, 1\}$, 一轮完毕后 $g = g/2$, m 设为 2.

动态坐标扰动策略的提出基于这样一种思想: 最优值应该在当前具有最好目标值的样例点附近(当前最好目标值不是最优值时). 因此, 对当前最好样例点的部分坐标进行微扰既保证了新点不会离原来点太远, 又为其探索多个方向提供了可能. 虽然扰动最好样例点的所有坐标会带来更大的探索空间, 但也会使得扰动点偏离当前最优点更远, 从而浪费后续的迭代次数. 本文提出的动态扰动策略是在 DYCORDS^[13] 策略启发下设计的. 这两种策略的主要区别在于 DYCORDS 采用

严格的递减函数达到从全局向局部的过渡, 而本文提出的动态扰动策略充分利用了已建立的 Kriging 代理模型, 将模型中描述参数相关性的 θ 作为选择扰动坐标的标准. θ 越大说明该变量更重要, 即该维度的一个微小变动就会对相关函数有大的影响, 因而对函数目标值也有大的影响. 在模型不精确情况下, 具有大 θ 的参数较多, 随着模型精度的不断提高, 大 θ 对应的参数不断减少, 从而达到扰动坐标数目的逐步递减. 动态坐标扰动过程如图 3 所示, 其中 $\epsilon = 103$.

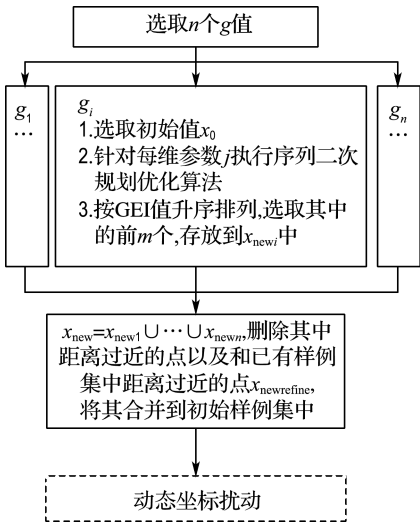


图 2 修订基于 Kriging 优化算法第三部分
Fig. 2 The third modified part of Kriging-based optimization algorithm

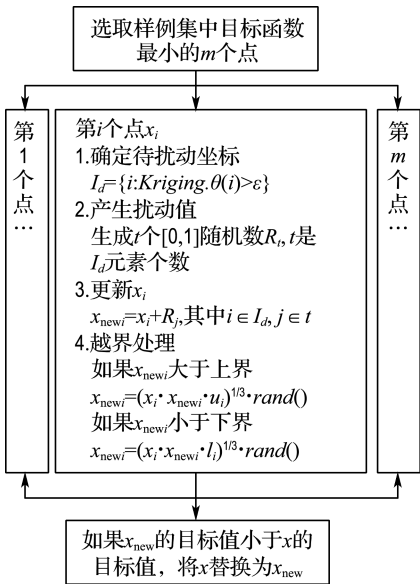


图 3 动态坐标扰动算法
Fig. 3 The algorithm of dynamic coordinate disturbance

考虑到 Kriging 建模时间会随着样本数目的增多而延长,本文采用李征提出的区域分割法即基于主成分分析方法对样本进行空间划分,在多个子空间并行建模以节省计算时间,具体区间划分方法参考文献[7].引入并行化后的改进优化算法流程如图 4 所示.

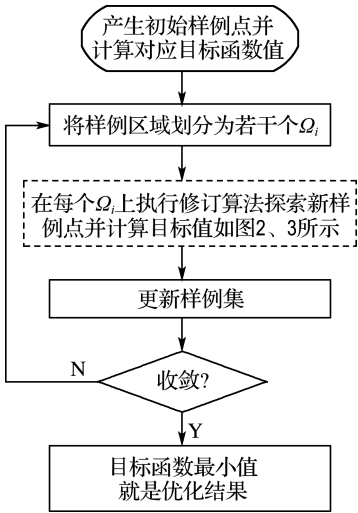


图 4 引入并行化后的改进优化算法流程图
Fig. 4 The flow chart of parallel improved optimization algorithm

2 计算结果

2.1 算例介绍

为了检测本文提出算法的有效性,将该算法应用于人类白细胞代谢网络的参数估计问题.人类白细胞代谢网络使用 13 个米氏方程表示的酶促反应来描述其动力学特征,米氏方程使用常微分方程描述,为了估计该网络的参数常常需要进行积分计算. Voit 等^[14]认为,对于带有积分计算的优化问题,积分所耗费的时间占整个优化时间的 95% 以上,因此是一种带有昂贵计算的优化问题.

该模型中总计有 41 个参数不能从实验中获取,必须依赖于优化算法进行估计.根据这 41 个参数在代谢过程中承担的角色和它们的取值范围,可以将其分为 8 类^[2],如表 1 所示,其中参数的最大变动范围跨越 6 个数量级,最小的跨越 1 个数量级.因此该问题的参数估计不仅是一个含有昂贵计算的高维参数估计问题,还是一个参数变动范围大的参数估计问题.由于米氏方程的高度非线性,此问题还同时具有高度非线性特征.

表 1 待估计参数信息
Tab. 1 The information of parameters to be estimated

类名	变化下限	变化上限	参数个数
Kcat	100	1×10^4	5
Km	0.1	1×10^3	5
Kd	1×10^{-5}	10	10
Kicom	0.1	1×10^2	7
Kilr	0.01	10	4
KI	100	1×10^3	4
Kilrb5lox	0.01	1	1
Enzyme	1×10^{-3}	10	5

2.2 计算条件设定及结果说明

针对人类白细胞代谢网络的参数特征,在初始样例点生成时采用对数拉丁超立方取样和均匀取样相结合的方式.因为初始样例构建的模型一定不准确,所以初始样例的数目应该尽量少,在节省计算时间的同时为后续样例保留机会.本文初始为该网络生成 4 组样例,每组样例 400 个.每组样例中目标函数最小值分别为 $1.630\ 0 \times 10^4$ 、 $1.524\ 1 \times 10^4$ 、 $1.337\ 2 \times 10^4$ 、 $1.466\ 3 \times 10^4$.本文算法的收敛准则定义为两条.其中第一条如下所示:

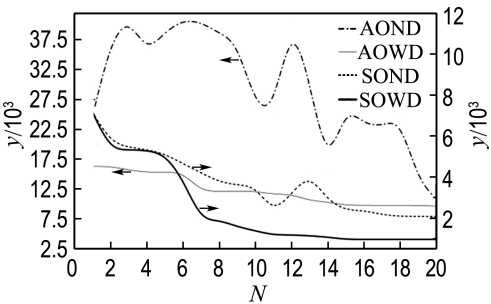
$$\max E(I(\mathbf{x})) \leqslant \epsilon(Y_{\max} - Y_{\min}); \epsilon < 1 \quad (4)$$

$$\mu(\mathbf{x}) = |\hat{y}_{\text{predict}} - y_{\text{true}}| / \hat{\sigma}(\mathbf{x}) < 3$$

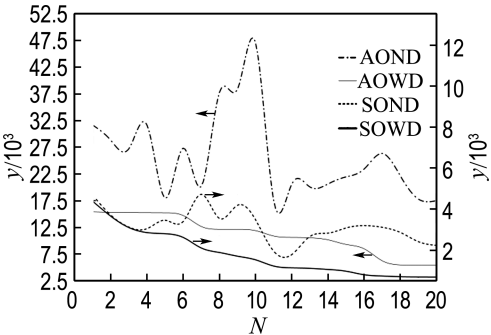
$\mu(\mathbf{x})$ 是标准差.第二条是迭代次数少于 21 次,以此体现的是资源受限优化.该算例的积分计算使用 Sundials 工具箱,软件的开发平台是 Visual C++ 6.0,运行环境是 Intel(R) Core (TM) 4 CPU 2.8 GHz.

为了证明该算法的有效性,特别是单参数优化和动态坐标扰动策略的有效性,首先给出人类白细胞代谢网络参数估计在 4 种优化策略作用下的结果比较,其中 4 种优化策略分别是:(1)所有坐标同时优化没有动态坐标扰动(AOND);(2)所有坐标同时优化有动态坐标扰动(AOWD);(3)单坐标分别优化没有动态坐标扰动(SOND);(4)单坐标分别优化有动态坐标扰动(SOWD).4 种策略的优化过程如图 5 所示,其中 AOND 和 AOWD 对应左侧坐标,SOND 和 SOWD 对应右侧坐标.

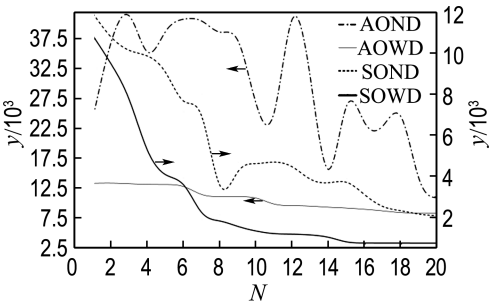
从图 5 可以看出,所有带有扰动的迭代过程都呈现出递减的趋势,这主要是因为动态扰动都是在数据集中具有最小目标值的样例点上扰动完成的,且只有当扰动后的目标值小于原目标值才会发生替换,但因为在某些点扰动不会带来目标函数值的减小,因此这些曲线并不表现为单调递



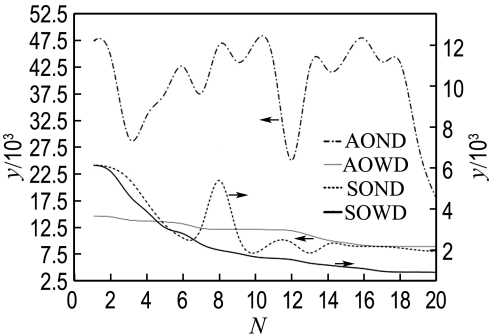
(a) 数据集 1



(b) 数据集 2



(c) 数据集 3



(d) 数据集 4

图 5 4 种策略优化结果比较

Fig. 5 Comparison of optimization results by four strategies

为了比较不同加点准则对优化结果的影响,使用相同的 4 组初始样例集,分别以基于 EI 最大和 GEI 最大作为选择新样例点的目标函数,在 Kriging 框架下执行 SOND 策略,执行结果如表 2 所示.

表 2 EI 和 GEI 两种加点准则下得到的最优目标函数值比较

Tab. 2 The comparison of optimal objective value for EI and GEI sampling fill criterion

数据集	初始目标函数最小值	最优目标函数值	
		EI	GEI
1	$1.630\ 0\times10^4$	2 693.64	2 074.47
2	$1.524\ 1\times10^4$	3 230.57	2 225.83
3	$1.337\ 2\times10^4$	2 197.91	2 086.35
4	$1.466\ 3\times10^4$	1 940.64	1 924.07

由表 2 可以看出,基于 GEI 加点优化得到的最优目标函数值普遍优于基于 EI 加点优化得到的最优目标函数值.

图 6 给出了不同处理器数目对 4 个数据集优化过程的影响.由图 6 可以看出,优化结果会随着处理器数目的增加而变好,这是因为随着处理器数目的增加,样例空间区域划分的范围变小,建模精确度提高.但这并不意味着不断增加处理器的数目就会使得优化结果持续变好,因为处理器的数目对应着样例区间的划分,区间划分得太细会造成该区域样例点的过分稀疏从而影响建模的精度和优化的结果.由图 6 也可看出当处理器数目不同时,同一数据集优化过程选择的新样例点会有交叉.

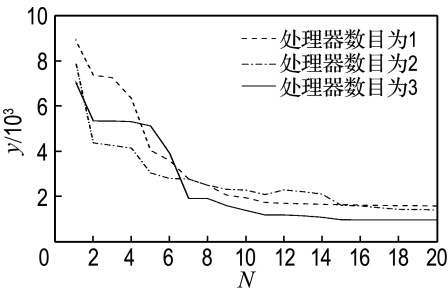
为了显示本文算法的有效性,将该算法(4 个处理器 20 次迭代 4 组集合中最优值)和其他 3 个可用于此类问题的参数优化算法进行比较,参与比较的算法为 TS-VFBRH^[15]、MCA^[2]、SSMGO^[4].优化结果如表 3 所示,由表可以看出,本文算法在较少时间内得到了和 3 个算法中最好算法接近的目标函数值.这些优化算法估计得到的参数值和实验数据拟合的效果如图 7 所示.

基于 Kriging 代理模型的优化算法有一项重要的评价指标就是计算 Kriging 代理模型的近似误差.这种近似误差可以作为衡量参数估计结果不确定性的标准.一般可以使用两种误差定义来估计^[16].第一种是平均相对误差,定义如下:

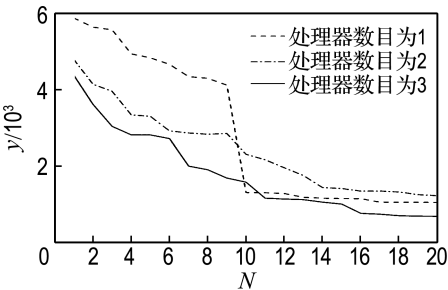
$$\bar{e} = \frac{1}{n} \sum_{i=1}^{n_t} e^{(i)}, \quad e^{(i)} = \|(\hat{y}_i^{(i)} - y_i^{(i)})/y_i^{(i)}\| \quad (5)$$

其中 n_t 是检测样例点的个数, $\hat{y}_i^{(i)}$ 和 $y_i^{(i)}$ 分别是在

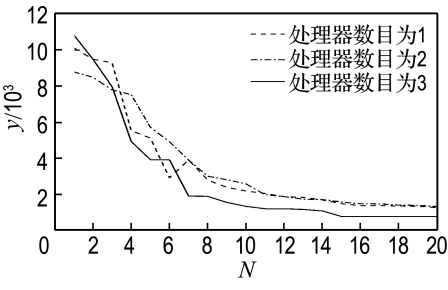
减趋势.同时还可以看出单维优化效果要明显好于多维同时优化效果,这主要归因于多维同时优化时彼此之间复杂的相互作用,可能抵消单维独立优化对目标值的影响.



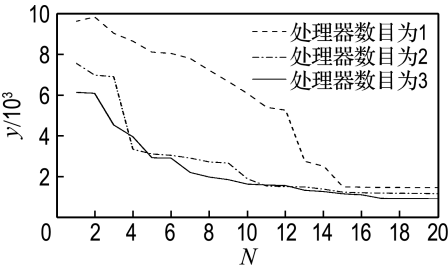
(a) 数据集 1



(b) 数据集 2



(c) 数据集 3



(d) 数据集 4

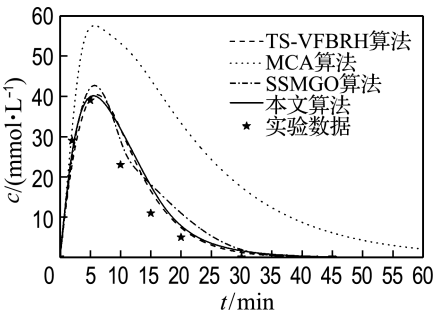
图 6 不同处理器数目对优化结果影响

Fig. 6 Effect of different number of processors on optimization results

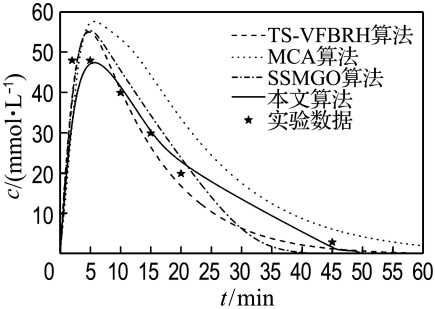
表 3 不同优化算法的比较

Tab. 3 Comparison of different optimization algorithms

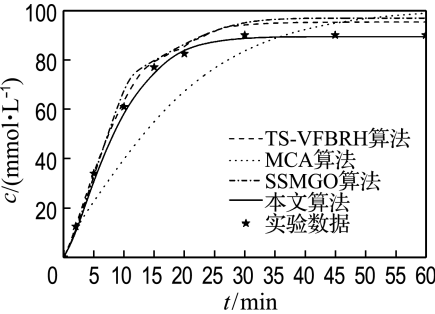
算法	最优目标函数值	计算时间/min
本文算法	667.994 0	23.189 5
SSMGO 算法	1 135.576 0	28.595 5
TS-VFBRH 算法	627.205 4	32.229 6
MCA 算法	6 876.421 0	



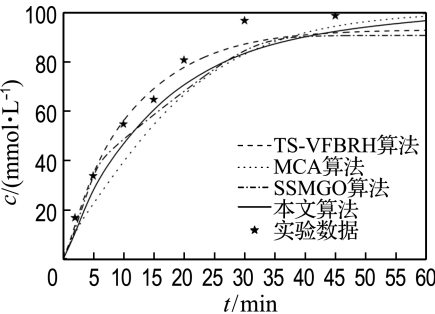
(a) ltb4_noAA



(b) ltb4_AA



(c) 20-OH-LTB4_noAA



(d) 20-OH-LTB4_AA

图 7 不同算法与实验数据拟合效果比较

Fig. 7 Effect comparison of different algorithms fitting the experimental data

测试点 i 的预测值和计算值. 第二种是均方误差, 定义如下:

$$\sigma_e = \sqrt{\sum_{i=1}^{n_t} (e^{(i)})^2 / n_t} \tag{6}$$

由于本文算法具有参数维度高,参数变化范围大这两个特点,在有限的计算资源条件下(固定的迭代次数),要求构建的 Kriging 模型在所有参数、所有维度的整个变化范围都具有高精度是不现实的,因此本文在对模型精度进行估计时采用两种策略:(1)在最优值的邻近区域随机取样计算近似误差;(2)在所有参数的整个变化范围随机取样(全局取样)计算近似误差.两种策略下都分别取 20 个点进行估计,比较结果如表 4 所示.在最

优值的邻近区域随机取样计算近似误差均小于 3%,说明此时模型是准确的.而全局取样计算误差的最小误差也是大于 40%的,说明建立的模型在全局范围内精度不高,这主要归咎于参数维度过高,变化范围过大,在有限样例个数和有限的迭代次数下得不到全局范围内高精度的模型.这也符合本文提出这个算法的初衷,即在有限的计算资源条件下,找到全局近似最优解.

表 4 Kriging 模型的近似误差
Tab. 4 Approximation errors of Kriging model

策略	平均相对误差				均方误差			
	数据集 1	数据集 2	数据集 3	数据集 4	数据集 1	数据集 2	数据集 3	数据集 4
最优值附近	0.020 676 0	0.021 320 0	0.026 300 0	0.019 676 0	0.023 983 7	0.021 510 0	0.029 200 0	0.020 837 0
全局范围	0.542 0	0.440 0	0.601 1	0.572 0	0.613 0	0.480 0	0.659 3	0.613 0

3 结 语

本文针对带有昂贵计算的高维参数估计优化问题,提出了一种基于 Kriging 代理模型同时引入单参数优化、动态坐标扰动,及多 GEI 加点的混合优化算法.为了检测该算法的有效性,将该算法应用于一个带有 41 维参数的人类白细胞代谢网络的参数估计问题,结果表明,在有限的计算资源下该算法能在较少的时间内得到和其他算法具有可比性的优化结果,而且估计的参数和实验数据拟合效果较好.下一步的工作是改进加点准则,使得新增加的点具有最大的信息量.

参考文献:

[1] Conn A R, Scheinberg K, Vicente L N. **Introduction to Derivative-Free Optimization** [M]. Philadelphia: Society for Industrial and Applied Mathematics, 2009.

[2] YANG Kun, BAI Hong-jun, OUYANG Qi, *et al.* Finding multiple target optimal intervention in disease-related molecular network [J]. **Molecular Systems Biology**, 2008, **4**:228.

[3] Auliac C, Frouin V, Gidrol X, *et al.* Evolutionary approaches for the reverse-engineering of gene regulatory networks: A study on a biologically realistic dataset [J]. **BMC Bioinformatics**, 2008, **9**: 91.

[4] Egea J A, Rodriguez-Fernandez M, Banga J R, *et al.* Scatter search for chemical and bioprocess

optimization [J]. **Journal of Global Optimization**, 2007, **37**(3):481-503.

[5] Ponweiser W, Wagner T, Vincze M. Clustered multiple generalized expected improvement: A novel infill sampling criterion for surrogate models [C] // **2008 IEEE Congress on Evolutionary Computation, CEC 2008**. Piscataway: IEEE Computer Society, 2008:3515-3522.

[6] Schonlau M, Welch W J, Jones D R. Global versus local search in constrained optimization of computer models [J]. **Lecture Notes-Monograph Series**, 1998, **34**:11-25.

[7] 李 征. 高聚物成型工艺的系统优化设计及其并行计算[D]. 大连:大连理工大学, 2013.

LI Zheng. The system-based optimization of the injection molding process and its parallel computing [D]. Dalian: Dalian University of Technology, 2013. (in Chinese)

[8] McKay M D, Beckman R J, Conover W J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code [J]. **Technometrics**, 2000, **42**(1): 55-61.

[9] Shewry M C, Wynn H P. Maximum entropy sampling [J]. **Journal of Applied Statistics**, 1987, **14**(2):165-170.

[10] FANG Kai-tai, Lin D K J, Winker P, *et al.* Uniform design: Theory and application [J]. **Technometrics**, 2000, **42**(3):237-248.

[11] Shimoyama K, Jeong S, Obayashi S. Kriging-

- surrogate-based optimization considering expected hypervolume improvement in non-constrained many-objective test problems [C] // **2013 IEEE Congress on Evolutionary Computation, CEC 2013**. Washington D C: IEEE Computer Society, 2013: 658-665.
- [12] Jones D R, Schonlau M, Welch W J. Efficient global optimization of expensive black-box functions [J]. **Journal of Global Optimization**, 1998, **13**(4): 455-492.
- [13] Regis R G, Shoemaker C A. Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization [J]. **Engineering Optimization**, 2013, **45**(5): 529-555.
- [14] Voit E O, Almeida J. Decoupling dynamical systems for pathway identification from metabolic profiles [J]. **Bioinformatics**, 2004, **20**(11): 1670-1681.
- [15] WANG Hong, WANG Xi-cheng. Parameter estimation for metabolic networks with two stage Bregman regularization homotopy inversion algorithm [J]. **Journal of Theoretical Biology**, 2014, **343**: 199-207.
- [16] HAN Zhong-hua, ZHANG Ke-shi. Surrogate-based optimization [M] // **Real-World Applications of Genetic Algorithms**. New York: Springer International Publishing, 2012: 343-362.

Hybrid Kriging surrogate model optimization algorithm for high-dimension parameter estimation

WANG Hong^{*1,2}, WANG Xi-cheng^{1,3}, LI Ke-qiu¹

(1. School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China;

2. Department of Computer Science, Dalian Neusoft University of Information, Dalian 116023, China;

3. State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian 116024, China)

Abstract: Kriging surrogate-based model optimization algorithm is an effective algorithm in solving optimization problems with expensive computation. However, it is not feasible to deal with the parameter estimation of high-dimension. Aiming at this problem, a new optimization algorithm, i. e. hybrid Kriging surrogate model and other optimization technologies for high-dimension parameter estimation is proposed. In this algorithm, single parameter estimation is adopted for Kriging model infill sampling to conquer the curse of dimensionality and improve the convergence rate. Meanwhile, based on the information of the built Kriging surrogate model a new dynamic coordinate disturbance strategy is presented and used for high-dimension parameter estimation to refine objective values. To avoid missing the global optimal value, multi-modal searching based on sampling fill criterion. i. e. generalized expected improvement is introduced. Its effectiveness is verified by estimating parameters of a human polymorphonuclear leukocyte metabolic network with 41 dimension parameters. The experimental results show that the algorithm can produce better parameters' estimation results agreeing with experimental data and small objective values under limited number of iterations.

Key words: Kriging surrogate model; parameter estimation of high-dimension; limited computing resource; generalized expected improvement