

基于多级代理模型的优化算法

王琦^{1,2} 丁运亮¹ 陈昊²

(1. 南京航空航天大学飞行器先进设计技术国防重点学科实验室, 南京, 210016; 2. 南昌航空大学, 南昌, 330063)

摘要: 在仿真优化中高精度的仿真模型大多难以实现, 本文提出一种由全局和局部代理模型共同作用的多级代理模型, 并与仿真优化相结合提出基于多级代理模型的仿真优化算法。运用Kriging近似理论和RBF神经网络分别构建全局代理模型和局部代理模型, 并在仿真优化的过程中在线更新代理模型。通过算例对算法进行验证, 结果表明多级代理模型具有良好的逼近能力, 基于多级代理模型的仿真优化方法具有良好的鲁棒性和寻优性能。

关键词: 代理模型; 遗传算法; 神经网络; 仿真优化

中图分类号: O 232 **文献标识码:** A **文章编号:** 1005-2615(2008)04-0501-06

Simulation Optimization Based on Multilevel-Surrogate Models

Wang Qi^{1,2}, Ding Yunliang¹, Chen Hao²

(1. Key Laboratory of Fundamental Science for National Defense-Advanced Design Technology of Flight Vehicle,

Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, China;

2 Nanchang Hangkong University, Nanchang, 330063, China)

Abstract: The high precision simulated models are difficult to be actualized during simulation optimization. In this paper, a new multilevel-surrogate model constituted by global and local surrogate model is proposed, and a new simulation optimization based on multilevel-surrogate model is given. The global and the local surrogate models by using Kriging model and RBF neural network are constructed and are renewed in time during the optimization process. The computation results indicate that multi-level surrogate model has the capacity of good approximations, and the simulation optimization based on multi-level-surrogate models has the performances of good robustness and optimization.

Key words: surrogate models; genetic algorithm; neural networks; simulation optimization

随着科学的发展, 仿真优化已成为优化设计、系统仿真等领域所共同关注的前沿课题, 如何降低计算成本、提高优化的寻优性能和鲁棒性是其中的关键。应用代理模型来解决这一问题是当前研究的热点。所谓代理模型是指计算量小但其计算结果与高精度模型的计算结果相近的分析模型。代理模型是一种包含试验设计和近似算法的建模技术, 在设计优化过程中用代理模型替代原有的高精度分析模型, 可以提高仿真优化的寻优效率, 降低算法的计算成本。

目前, 国内外许多专家学者致力于基于代理模型的仿真优化理论的研究。文献[1]探讨了如何利用双精度仿真模型构建代理模型的方法, 首先通过低精度仿真模型构建代理模型, 然后利用少量的高精度模型数据样本对其进行修正。文献[2]用基于代理模型的进化算法进行飞行器飞行动力学设计, 取得不错的优化效果。文献[3]运用高斯近似模型进行全局近似, 并通过多个数值算例进行算法性能验证。文献[4]提出基于序列响应面的协同优化算法, 并讨论了响应面更新的问题。文献[5]利用双代

理模型来提高仿真优化的统计性能,并用其进行压力管设计。

本文提出一种由全局和局部代理模型共同作用的具有良好逼近能力的多级代理模型,并将其应用到仿真优化算法中,把多级代理模型与仿真优化算法相结合,进一步提出一种基于多级代理模型的仿真优化算法。选用Kriging近似模型作为全局代理模型,RBF神经网络作为局部代理模型。与其他近似模拟技术相比,Kriging是一种更具统计性的近似技术,它的有效性不依赖于随机误差的存在,具有良好的全局统计特性,因此适合作为全局代理模型进行全局近似^[6]。RBF网络是以函数逼近理论为基础构造的一类前向网络,具有结构自适应确定、输出与初始权值无关等优良特性,适合作为局部代理模型进行局部的修正^[7]。

1 基于多级代理模型优化算法的原理及实现

1.1 全局代理模型

选取Kriging代理模型作为全局代理模型,采取均匀设计方法选取样本点^[8]。通常Kriging模型包括两部分,具体模型为

$$y(x) = f(x) + z(x) \quad (1)$$

式中: $f(x)$ 是对全部设计空间的全局模拟,可看作一个常数 β , β 值可由已知响应值进行估计; $z(x)$ 是期望为0,方差为 σ^2 的高斯随机函数,表示全局模拟的插值。

$z(x)$ 的协方差矩阵可表示为

$$\text{cov}[z(x^i), z(x^j)] = \sigma^2 R(x^i, x^j) \quad (2)$$

式中: R 为相关矩阵; $R(x^i, x^j)$ 为相关函数, $i, j = 1, 2, \dots, n$, n 为已知的样本中数据的个数。 $R(x^i, x^j)$ 的函数形式是可以选择的,在这里选择的表达式为

$$R(x^i, x^j) = \exp \left(- \sum_{k=1}^{n_k} \theta_k |x_k^i - x_k^j|^2 \right) \quad (3)$$

式中: n_k 为设计变量的个数; θ 为未知的相关参数矢量。一般地可用标量 θ 来代替 θ 。

根据Kriging理论,未知点 x 处的响应值 y 的估计值 \hat{y} 可通过下式给出

$$\hat{y} = \hat{\beta} + r^T(x) R^{-1} (y - f \hat{\beta}) \quad (4)$$

式中: y 为样本的响应值; f 为长度为 n 的单位列向量; $r^T(x)$ 为未知向量 x 与样本输入数据 $[x_1, x_2, \dots, x_n]$ 之间的相关向量,表达式为

$$r^T(x) = [R(x, x_1), R(x, x_2), \dots, R(x, x_n)]^T \quad (5)$$

相关系数 θ 可以由极大似然估计给出,即在 $\theta > 0$ 时使得下式取最大值

$$- [n \ln(\hat{\sigma}^2) + \ln |R|] / 2 \quad (6)$$

$\hat{\beta}, \hat{\sigma}^2$ 的估计值由下式给出

$$\hat{\beta} = (f^T R^{-1} f)^{-1} f^T R^{-1} y \quad (7)$$

$$\hat{\sigma}^2 = (y - f \hat{\beta})^T R^{-1} (y - f \hat{\beta}) / n \quad (8)$$

式中 R 和 $\hat{\sigma}^2$ 都是由 θ 表示的函数,由求解式(6)得到 θ 值构成的Kriging模型就是性能最好的代理模型。

由于在Kriging模型中全局模拟的插值用随机函数 $z(x)$ 表示,这使得Kriging模型的有效性不依赖于随机误差的存在,从而更具统计特性,适合进行全局近似。

1.2 局部代理模型

由于全局代理模型是由少量的真实数据构建的,不能完全真实模拟出实际的仿真模型,因此根据全局代理模型进行寻优操作往往无法收敛至真实的全局最优解。而且,由代理模型确定的最优解是近似值,代理模型在最优解附近的局部范围内近似性能的好坏与整个寻优精度之间有着至关重要的联系。因此,有必要建立局部代理模型对全局代理模型的修正。

设 x^* 为全局寻优得到的最优点,在 x^* 点周围采取中心组合设计^[8]的方法选取少量试验点,利用得到的样本构建局部代理模型。本文采用RBF神经网络构建局部代理模型进行局部修正,具体步骤可参考文献[9]。

1.3 代理模型更新

当前仿真优化中常用的代理模型多为全局代理模型,即用多项式模型、径向基模型和Kriging模型等一次性构建代理模型,然后引入仿真优化算法中代替真实模型进行评价。这种方式将代理模型的构建从优化算法中独立出来,忽略了代理模型的更新问题。由于单一的代理模型无法进行高精度近似,所以如何合理地进行代理模型更新是其能否正确反映真实模型的关键所在,这也直接影响到寻优的精度与效率。代理模型的统计性能与样本点的选取有关,样本点越集中代理模型的平均绝对值误差越小。若随着进化算法寻优的深入,根据当前最优解、某种规则不断地合理缩小搜索区域,对代理模型进行序列更新,可以在一定程度上对代理模型进行修正,进一步提高代理模型的逼近能力。

代理模型更新的实质是对样本的更新,即样本中心和样本范围的更新,本文根据置信域理论对样本进行更新。

1.4 基于多级代理模型的仿真优化算法

在仿真优化中,目标函数的确定常常需要进行

有限元仿真、模型参数估算等计算成本高的环节,而且存在各种各样的噪声干扰,从而导致仿真优化算法计算成本过高、寻优精度和可靠性低。本文提出一种基于多级代理模型的仿真优化算法来解决这一困难,利用 Kriging 模型构造全局代理模型、RBF 神经网络构造局部代理模型、置信域理论进行样本更新,构建多级代理模型框架。算法的流程图如图1所示。图中 N 表示当全局寻优当连续 N 步收敛至同一值时进行局部代理模型的构建, M 表示局部搜索 M 步后进行搜索空间的更新。

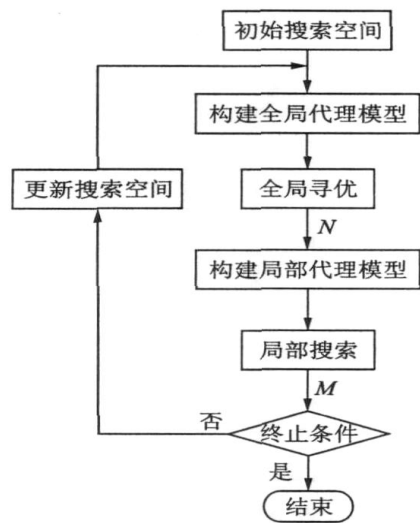


图1 基于多级代理模型的仿真优化算法流程图

在新算法中,对代理模型构建与寻优操作顺序进行的常规方式进行改进。采用遗传算法进行全局寻优,利用当前最优解的位置根据置信域理论进行代理模型更新,同时根据置信区间改变全局寻优的搜索区间,在提高代理模型近似能力的同时合理地缩小搜索空间从而提高优化精度。在局部环节,对于局部代理模型采用最速下降法进行局部寻优,从而提高算法的局部搜索能力,加速收敛速度。

2 算例分析

2.1 代理模型性能的分析

2.1.1 近似性能分析

Rastrigin 函数是一个典型的数值测试函数,它的表达式为

$$f(x_1, x_2) = \sum_{i=1}^2 (x_i^2 - 10 \cos(2\pi x_i) + 10) - 10 < x_i < 10$$
 (9)

依据均匀试验设计表 $U_{31}(31^2)$ 选取样本,分别用 Kriging 模型(I)、带更新的 Kriging 模型(II)、多级代理模型(III)进行代理模型的构建,II与III的置信系数都取0.05。随机选取 $N_c = 20$ 个点作为校验样本,比较如下的误差项目:

(1) 最大绝对值误差

$$\max\text{-error} = \max\{|f_i - \tilde{f}_i|\} \quad i = 1, \dots, N_c$$
 (10)

(2) 平均绝对值误差

$$\text{Ave. error} = \frac{1}{N_c} \sum_{i=1}^N |f_i - \tilde{f}_i|$$
 (11)

(3) 均方根误差

$$\text{RMSE} = \sqrt{\frac{1}{N_c} \sum_{i=1}^N (f_i - \tilde{f}_i)^2}$$
 (12)

相关数据如表1所示。通过表中的数据对比可以看出,与I相比,II与III的误差值随着序列更新的深入而逐渐减小,这说明代理模型的更新环节能够逐步改善代理模型的近似能力;与II相比,随着序列更新的深入模型III的各项误差指标迅速下降,这说明局部代理模型的修正在进一步提高模型逼近能力的同时,提高了代理模型的构建效率。

2.1.2 鲁棒性分析

为了验证多级代理模型的鲁棒性能,在Rastrigin 函数中加入白噪声: $F = f + k \cdot \text{causs}(0, 1)$, k 为增幅系数,这里取 $k = 1$ 。仍依据 $U_{31}(31^2)$ 选取样本,用两种方法构建代理模型,模型I用本文提出的多级代理模型,模型II用二次多项式模型代替Kriging 模型作为全局代理模型,其他同模型I,置信系数都取0.05,代理模型的序列图如图2所示。

通过图2中的图形对比可以看出,对于带噪声干扰的数学模型,由模型I构建的代理模型曲面光滑,且随着更新逐步精确(图2(a)),而由模型II构建的代理模型曲面呈凹凸状(图2(b))。这说明采用不受随机噪声干扰的Kriging 模型作为全局代理模型的多级代理模型更具鲁棒性,抗干扰能力强。

表1 带延时优化问题的优化结果比较

代理模型	I			II			III		
序列号		5	10	15	5	10	15		
Max-error	34.123 5	24.814 0	15.724 7	13.638 0	14.039 4	6.742 8	0.527 2		
Ave-error	16.578 8	5.618 3	4.192 0	3.831 9	3.879 3	0.829 3	0.052 8		
RMSE	13.470 3	4.218 3	3.183 1	2.651 6	2.432 5	0.362 5	0.019 6		

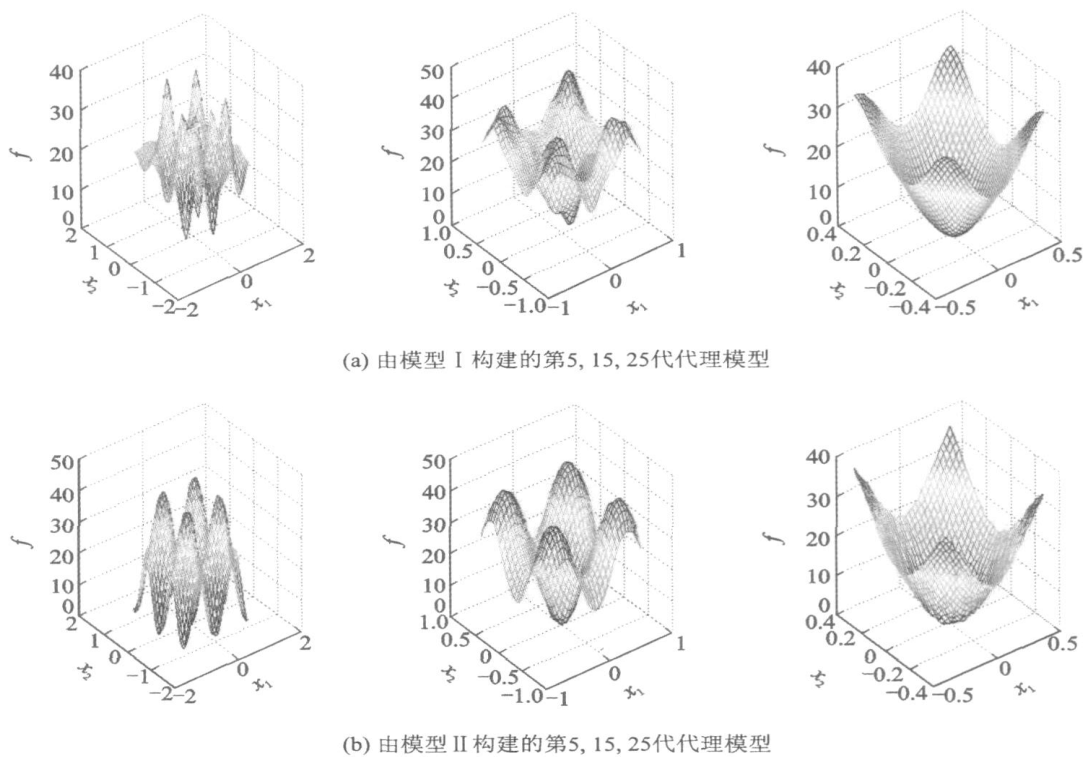


图 2 两种代理模型序列图对比

2.2 仿真优化算法的性能分析

在实际的应用中, 目标函数的确定通常需要进行参数估算等低成本计算环节, 在带噪声干扰的 Rastrigin 函数中引入延时环节: $F_d = f + k \cdot \text{causs}(0, 1) + \text{delay}(l)$ 对这一特性进行模拟, 其中 l 为延时时间, 本文取 $l = 0.1 \text{ s}$ 。为了对比仿真优化算法的寻优性能, 方法 I, II, III 均用遗传算法进行寻优操作, 取相同的遗传操作系数, 分别由原始模型 F_d 、Kriging 模型、带更新的 Kriging 模型进行目标函数的评价, 方法 IV 采用本文提出的基于多级代理模型的优化算法, 全局寻优算法与前 3 种采用相同的遗传操作系数, 取 $N = 5, M = 10$ 。寻优结果对比如表 2 所示, 其中真实值是指将最优解反代回不带噪声的函数 f 的真实值, t 为运行时间, β 为平均绝对值误差, Num 为到达稳定值时代理模型更新次数。

由表中的数据对比可以看出, 对于带噪声干扰、延时环节的仿真优化问题, 方法 I 无法进行有效的寻优操作且计算成本过大、优化效率低; 方法 II 能够大幅地降低计算成本却无法保证良好的寻优精度, 误差过大; 由于更新环节的引入, 方法 III 具有良好的寻优精度, 但是精度的提高是以多次的代理模型更新为代价的, 计算成本仍然很高; 在方法 IV 中, 不仅进行代理模型的更新而且还引入了局部代理模型的构建与寻优, 使其在保证寻优精度的同时提高了寻优精度。图 3 为方法 III 与方法 IV 的迭代曲线对比图, 从图中可以进一步地看出方法 III 需要 20 次左右的代理模型更新, 而方法 IV 仅需 10 次左右即可趋于稳定。因此, 基于多级代理模型的仿真优化方法具有良好的寻优精度、寻优效率及鲁棒性。

表 2 仿真优化结果对比

方法	最优解		近似值	真实值	t/s	β	Num
I	- 6.674e- 2	- 0.937 8		2.456 3	71.43		
II	- 0.940 0	0.896 3	5.713 6	4.351 6	3.27	10.254	
III	1.940e- 2	1.553e- 3	0.215 4	0.075 8	42.65	4.372	20
IV	1.816e- 4	1.463e- 5	2.105e- 4	8.526e- 6	18.95	0.026 5	9

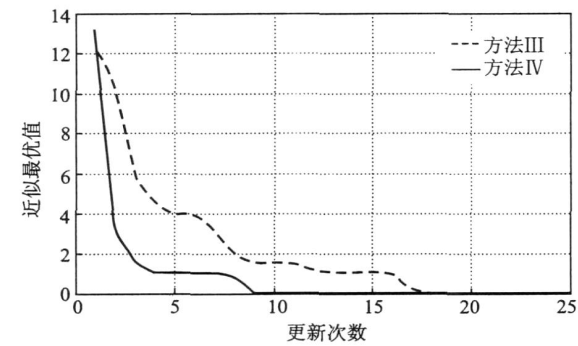


图3 方法Ⅲ与方法Ⅳ的迭代曲线对比

2.3 中孔悬臂梁的优化设计问题^[10]

具有矩形横截面的中孔悬臂梁承受着水平、垂直两个方向载荷的作用, 其结构图如图 4 所示, 设计目标是在满足屈服极限 α 的条件下使得梁体总重量最小。

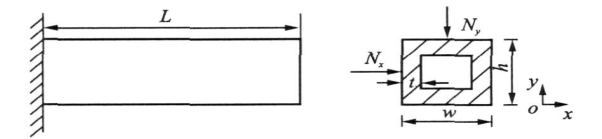


图4 中孔悬臂梁结构图

设计变量为梁的横截面尺寸 w, h, t , 梁的长度 L 为 3 m, 载荷 N_x, N_y 与许应弯曲应力 R 均值固定, 但假定为服从正态分布的随机变量, 如表 3 所示。假定材料的密度为常数, 可直接取梁的截面积 A 最小作为目标函数, 优化设计的数学模型为

$$\begin{aligned} \min \quad & A = wh - (w - 2t)(h - 2t) \\ \text{s.t.} \quad & g = R - \frac{Lw}{2I_y} N_x + \frac{Lh}{2I_x} N_y > 0 \end{aligned}$$

(13)

表3 载荷与屈服极限的概率分布

取值	N_x/kN	N_y/kN	α/MPa
均值	2.0	4.0	280
标准差	0.5	0.2	10

注: α_s 与 R 满足系数关系 $\alpha_s = R\mu$, 一般 μ 的取值范围为 1.2~2.5, 本文取 $\mu = 1.5$ 。

分别采用 3 种优化方案对中孔悬臂梁进行优化设计, 每种方案运行 30 次, 设计变量 w, h, t 的取值范围分别为: 50~500 mm, 50~500 mm, 2~5 mm。设计 I: 基于真实数学模型, 取 N_x, N_y, R 的均值进行约束适应度函数的计算, 采用确定性优化; 设计 II: 基于全局代理模型(Kriging 模型), 取置信度为 90% 进行代理模型更新; 设计 III: 基于本文提

出的多级代理模型, 其他条件同设计 II。在对目标值 A 进行对比的同时, 引入可靠优化设计中可靠度的概念对寻优精度进行进一步的表征, 可靠度 β 的表达式为^[9]

$$\beta = -\Phi^{-1}(1 - P(g > 0))$$

(14)

式中: Φ 为标准正态分布函数; P 为事件 $g > 0$ 的概率; g 为优化数学模型中的约束函数。3 种优化方案的优化结果如表 4 所示。

表4 三种优化方案的结果对比

方案	w/mm	h/mm	t/mm	A^*/mm^2	A/mm^2	可靠度/%
I	154.925 7	154.234 4	2.034 8		1 241.6	50
II	130.384 3	144.056 4	2.012 6	1 183.4	1 088.5	99.43
III	93.098 6	110.621 0	2.000 6	902.3	799.09	99.99

注: A^* 为代理模型输出的近似值; A 为将设计变量 w, h, t 反代入真实模型的真实值。

对比表 4 中的数据可以看出, 与设计 I 相比采用代理模型的设计 II 和设计 III 所得的截面积分别小了 12.33%, 35.64%, 表明代理模型的引入克服了常规计算中取均值造成的计算偏差, 在一定程度上提高了寻优的精度。设计 III 寻优结果大大优于设计 II, 表明采用全局代理模型和局部代理模型相结合的近似方法能够更进一步地提高代理模型的逼近能力, 基于多级代理模型的仿真优化算法有着更好的寻优精度。采用确定性优化设计 I 的可靠度仅为 50%, 而基于代理模型的设计 II 和设计 III 将可靠度提高到 99% 以上, 且设计 III 要略优于设计 II, 这说明基于代理模型的优化算法在优化目标函数的同时, 降低了响应随随机因素和设计变量的敏感性, 提高了优化设计的寻优精度和稳健性。

3 结束语

进行仿真优化时, 由于高精度的仿真模型大多难以实现, 这直接影响到优化的精度、效率与可靠性。本文给出一种将全局代理模型和局部代理模型相结合的多级代理模型, 选择 Kriging 模型和 RBF 神经网络分别进行全局和局部的代理模型构建, 讨论了进化搜索过程中样本点的选取与更新的问题并给出更新的详细步骤。将多级代理模型和优化算法相结合, 提出一种基于多级代理模型的仿真优化算法, 选用遗传算法进行全局寻优, 在全局代理模型更新的同时更新全局寻优的搜索空间, 并对局部代理模型用梯度下降法进行局部寻优。算例验证表明, 多级代理模型具有良好的近似能力, 基于多级代理模型的仿真优化算法具有良好的鲁棒性和寻优性能。

参考文献:

- [1] Zhou Z, Ong Y S, Nair P B, et al. Combining global and local surrogate models to accelerate evolutionary optimization [J]. IEEE Transactions On Systems, Man and Cybernetics, 2007, 37(1): 1-27.
- [2] Qian Z, Seepersad C C, Joseph V R, et al. Building surrogate models based on detailed and approximate simulations [J]. ASME Journal of Mechanical Design, 2006, 128: 668-677.
- [3] Bauche D, Schraudolph N N, Koumoutsakos P. Accelerating evolutionary algorithms with gaussian process fitness function models [J]. IEEE Transactions on Systems, Man and Cybernetics, 2004, 35(2): 183-194.
- [4] 李响, 李为吉. 基于序列响应面方法的协同优化算法 [J]. 西北工业大学学报, 2003, 21(1): 79-82.
- [5] 王凌, 吉利军, 郑大钟. 基于代理模型和遗传算法的仿真优化研究 [J]. 控制与决策, 2004, 19(6): 626-630.
- [6] Martin J D, Simpson T W. Use of Kriging models to approximate deterministic computer models [J]. American Institute of Aeronautics and Astronautics, 2005, 43(4): 853-863.
- [7] Arakawa M, Nakayama H, Ishikawa H. Approximate optimization using radial basis function networks and genetic range genetic algorithms [C]. 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2004, 30(1): 853-863.
- [8] 方开泰, 马长兴. 正交与均匀试验设计 [M]. 北京: 科学出版社, 2001.
- [9] 苏金明, 阮沈勇, 王永利. MATLAB 工程数学 [M]. 北京: 电子工业出版社, 2005.
- [10] 李玉强, 崔振山. 基于响应面模型的 6σ 稳健设计方法 [J]. 上海交通大学学报, 2006, 40(2): 201-205.