

零件参数设计的数学模型

黄 杲 陈旭东 邵 伟

指导教师：数模组

(浙江大学, 杭州 310027)

编者按 本文先将粒子分离器参数 y 进行局部线性化处理, 并将 $\frac{\Delta y}{\sigma_y}$ 化归为标准正态, 其中 $\sigma_y^2 = \sum_{i=1}^7 (\frac{\partial y}{\partial x_i} \sigma_i)^2$, 最后把总费用的目标函数归纳为一个标准正态的式子 (文中 (4) 式). 以上结果合理、简洁明确. 采用网格法及蒙特卡罗法分别计算, 结果吻合、满意, 得到的解较优, 其中用蒙特卡罗法 (取二万个点) 计算, 在维数不变的情况下, 不失为一种速而有效的算法.

摘 要 本文建立了一个关于零件参数设计的数学模型. 本文首先利用概率的理论, 假设各零件产品的参数服从正态分布, 推出粒子分离器某参数 (y) 偏差的分布函数, 进而可得一批产品总费用的目标函数, 运用龙贝格数值积分将其转化为计算机可求值的函数, 然后运用网格搜索法和蒙特卡罗法求出目标函数的全局最优解.

本文将两种方法的结果精度、算法复杂度等进行比较, 重点讨论了效果较好的蒙特卡罗法. 本文最后分析了模型误差, 并对模型进行了评价和推广.

本模型最终得出产品总费用为 42.146 万元 / 千件, 其设定的零件参数为 $X^T [0.075, 0.375, 0.123, 0.115, 1.273, 12, 0.771]^T$, 其容差等级为 $G^T = [B, B, B, C, C, B, B]^T$.

一、问题的分析

要求解的问题是使总费用最低, 而总费用包括各零件成本及次、废品损失费, 综合考虑两种因素, 问题可归纳为总费用的非线性优化问题.

由于待优化的目标函数复杂, 无法利用其解析性质求最优解, 故可考虑用直接全局搜索法或随机试验点法.

从生产实际考虑, 本问题对解的精确度要求很高, 但是对求解算法的实时性无明确要求. 我们认为, 只要求解时间不是太长, 都是可接受的.

二、模型的假设及说明

1. 假设各零件参数服从参数 μ_i, σ_i^2 为正态分布, 且不同零件的参数相互独立.
2. 假设各零件容差的等级与其标定值的比为定值, 分别为: A 级 $\pm 1\%$, B 级 $\pm 5\%$, C 级 $\pm 10\%$.

说明 根据概率论知识和工程实际生产的一些测量数据可知, 成批生产的零件的参数服从参数为 μ_i, σ_i 的正态分布, 其中 μ_i 为各零件参数的期望值 (标定值), σ_i 为均方差 (即容差的 $1/3$), 可推知各零件的偏差 Δx_i 服从参数为 $0, \sigma_i$ 的正态分布.

三、文中用到符号及说明

y :	产品某参数	x_i :	各零件参数
y_0 :	y 的目标值($y_0 = 1.5$)	μ_i :	各零件参数的标定值
y_x :	各零件标定值确定的零件参数	C_i :	各零件成本
g_i :	各零件容差等级比	Δy :	产品的参数的偏差
x^T :	x_i 标定值向量($i=1,2,\dots,7$)	x_i :	各零件参数偏差
$x^{(T)}$:	x^T 的取值空间	p_i :	各零件参数均方差
G^T :	等级取值向量	p_1 :	次品概率
σ_y :	产品参数的均方差	p_2 :	废品概率

四、模型的建立和求解

本模型的建立基于概率论与误差的有关理论.

各零件偏差 Δx_i 相对于其标定值较小, 根据公式 (1), y 在 y_x 附近可以表示为:

$$y \approx y_x + \sum_{i=1}^7 \frac{\partial y}{\partial x_i} \cdot \Delta x_i \quad (2)$$

由于 Δx_i 较小, 则可得 $\Delta x_i \approx \Delta x_i$, 由于 $\Delta y = y - y_x$, 则

$$\Delta y = \sum_{i=1}^7 \frac{\partial y}{\partial x_i} \cdot \Delta x_i \quad (3)$$

在此, 我们不加证明的引入:

引理 1 x_i 服从参数为 μ_i, σ_i 的正态分布, 且彼此相互独立, α_i 为不全为零的常数, 若 $X = \sum_{i=1}^7 x_i$,

则

$$X \sim \left(\sum_{i=1}^7 \alpha_i^* \mu_i, \sum_{i=1}^7 \alpha_i^{2*} \sigma_i^2 \right)$$

根据公式 (3) $\frac{\partial y}{\partial x_i}$ 对应一组 x_i 为一定值, 而与 Δx_i 无关, 则由引理 1 可得

$$\Delta y \sim N(0, \sigma_y^2), \quad \left(\sigma_y^2 \triangleq \sum_{i=1}^7 \left(\frac{\partial y}{\partial x_i} \cdot \sigma_i \right)^2 \right)$$

(以上结论也可由方差合成定理推得, 见文献 [2] p93).

由概率论知识可得 $\frac{\Delta y}{\sigma_y} \sim N(0, 1)$

目标函数的建立

产品总费用 = 零件总成本 + 次品的损失费 + 废品损失费

即 $W = \sum_{i=1}^7 C_i + 1000 \times p_1 + 9000 \times p_2$ 可得

$$W = \sum_{i=1}^7 C_i + 9000 + 1000 \times \left[\Phi \left(\frac{1.4 - y_x}{\sigma_y} \right) - \Phi \left(\frac{1.6 - y_x}{\sigma_y} \right) \right] - 8000 \times \left[\Phi \left(\frac{1.8 - y_x}{\sigma_y} \right) - \Phi \left(\frac{1.2 - y_x}{\sigma_y} \right) \right] \quad (4)$$

目标函数为

$$\min W, \quad \text{s.t. } X^T \in X^{(7)}$$

(一) 对于原设计数据的求解

1. 在 Mathematics 2.1 下运行, 代入数据 $X^T = [0.1, 0.3, 0.1, 0.1, 1.5, 16.0, 75]$, $G = [B, C, C, C, C, C, B]$ 可得结果如下:

$$y = 1.72559, \quad p_1 = 0.6240, \quad p_2 = 0.2505, \quad W = 307.85 \text{ 万元.}$$

从结果可以看出, $p_1 + p_2 = 0.8745$, 即产品大部分为次废品, 这是 y_x 偏离 y_0 过大的结果, 此时产品总费用主要由次废品损失费决定, 由此可知, 在进行参数设计时, 应尽量先使 y_x 靠近 y_0 , 同时降低均方差. 这也是本模型降低算法复杂度的一个方向.

(二) 对目标函数 $\min W$ 的求解及参数的重新设计

1. 将目标函数转化为计算机可求解模型

由于原目标函数中的积分部分中被积函数为正态分布函数, 且其积分限为含有 7 维变量的复杂函数, 无法直接求解, 所以需将其转化为计算机可求解模型, 我们考虑用二种方法进行转化.

对于标准正态分布函数可采用最小二乘拟合法逼近, 将其转化为多项式表达, 但从其结果来看, 误差较大, 故不可取. 所以我们采用精度较高的龙贝格数值积分法来转化目标函数^[4]. 此方法为本模型高精度求解的出发点.

2. 用直接搜索法求最优解 (网格法)

原目标函数为 7 维多峰函数, 无法用解析法精确求解, 故考虑用直接搜索法. 常用的算法对于一般的多峰函数极值问题只能求出局部最优解, 而网格法为求解多峰函数全局最优解的一种较适宜的方法, 所以我们首先考虑用网格法求解最优目标.

我们对于每一个 x_i 在其取值范围内均取 6 个步长, 分为 6^7 个网格, 结合可能的容差等级组合, 在 Pentium 120 计算机上运行, 搜索约二十分钟后, 得到一个最优解, 结果是 $X^T = [0.075, 0.345, 0.115, 0.115, 1.275, 12, 0.7875]^T$, $G^T = [B, B, B, C, C, B, B]^T$, $Y_x = 1.497145$, $\sigma_y = 0.069220$, $w = 42.49146$ 万元. 从结果可以看出, 在一定精度内已求得一个较好的最优解, 我们可以通过降低算法复杂度, 使模型得到更好的应用.

但是根据网格法基本原理, 其循环次数由步长所决定, 而步长又由模型精度所决定, 故在一定精度要求下, 其算法复杂度不能大幅度降低. 我们可考虑采用蒙特卡罗法进行求解.

3. 蒙特卡罗法

蒙特卡罗法, 也就是随机实验点法. 它的基本思想是: 在函数的可行域内随机地选取实验点, 由于随机取得的点在区域中分配比较均匀, 所以对函数的大致形态能较好地体现^[3].

模型中的随机点是用以下方法产生的.

设 $m = 2^{16}$, $r_0 = 5$, 由迭代式可得 $(0, 1)$ 间的 p_i

$$\begin{cases} r_i = \text{mod}(2053r_{i-1} + 13849; m), & i = 1, 2, \dots \\ p_i = r_i/m & p_i \text{ 为第 } i \text{ 个随机数} \end{cases}$$

设 q_i 为 a 到 b 间的随机数, 则 $q_i = a + (b-a)p_i$. 我们编制了蒙特卡罗算法的程序 (略). 程序的运行时间为 2 分钟, 最终的计算结果为 $X^T = [0.0778, 0.374, 0.0979, 0.1067, 1.200, 12.526, 0.636]^T$, $G^T = [B, B, B, C, C, B, B]^T$; $y_x = 1.499136$, $\sigma_y = 0.069125$; $W = 42.3174$ 万元. 为了降低蒙特卡罗法网格法的复杂度, 采用了一些优化的方法:

(1) 程序的优化

必要的判断语句应尽量提早进行, 以减小循环次数或计算量. (如: y_x 应尽量靠近 y_0 , 因此可将偏离 y_0 过大的 y_x 判断排除, 以减少计算次数).

(2) 数据排除技巧

从粗略的直接搜索法结果可以看出, 不同容差等级组合对应的局部最优角相差较大, 而 G_1^T 和 G_2^T 因为总成本过大, 是不可能取到的.

4. 为进一步比较内最优解的大小, 可分别对优化目标在 G_1^T 和 G_2^T 内的最优解附近作探测性移动 (类似于直接搜索方法中的 Hobke=Jeeves 方法^[3]), 进行高精度计算, 分别得到 $W'_1 = 43.14865$, $W'_2 = 42.68475$, 因此, 能够求得具有一定精度的最优解 42.14865 万元. 此时所对应的零件参数为

$$X^T = [0.075, 0.345, 0.123, 0.115, 1.273, 12, 0.771]^T,$$

$$G^T = [B, B, B, C, C, B, B]^T,$$

$$y_x = 1.49683, \quad \sigma_y = 0.068777, \quad W = 42.1463 \text{ 万元}.$$

五、模型的检验及误差分析

(一) 模型的检验

由于我们没有一个确定的标准对本模型的解进行判断检验, 只能采取不同的方法, 对结果进行比较.

检验模型

在对题目中给定数据的计算结果分析时, 我们曾说过, 在设计参数时, 应尽量使 y_x 靠过 y_0 , 由此可建立如下模型. 在同一容差等级组合下, 令 $y_x = y_0$ 将 x_s 视为

$$174.42 \times \left(\frac{x_1}{y_x} \right) \times (x_3 / (x_2 - x_1))^{0.85} \\ \times \left(\left(1 - 2.62 \times \left(1 - 0.36 \times \left(\frac{x_4}{x_2} \right)^{-0.56} \right)^{1/16} \right) \right) / (x_1 \times x_7)^{0.5}$$

函数, 由概率论知识, 当 $y_x = y_0$ 时, 若 σ_y , 次废品概率越小, 总费用越低, 因此, 可将目标函数变为 $\min \sigma_y$ s.t. $y_x = 1.5$, 运行最后结果为 $X^T = [0.075, 0.375, 0.075, 0.12, 0.934, 12.8, 0.5625]^T$; $y_x = 1.5000$; $\sigma_y = 0.068901$; $W = 42.1949$ 万元. 比较原模型和检验模型的结果可知, 产品的最低总费用应在 42.1~42.2 万元之间, 而且对应最低总费用的零件参数值较合理, 这说明原模型的算法是可靠的.

另外, 网格法和蒙特卡罗方法所求得的最优解也相互吻合, 这也说明此结果是值得信赖的.

对于目标函数本身而言, 它虽然不是一个初等函数, 却是对于七个自变量的连续函数, 由连续函数的介值定理, 在全局最优解附近存在一个区域, 其内任何一点的值均小于其它鞍点的值, 即使这块区域只有整个容许区域的万分之一, 在蒙特卡罗方法取二万个点的情况下, 落入该区域的概率仍接近 60% (点与点之间相互独立), 在多次运用蒙特卡罗求解的情况下, 一般是能够求出全局最优解的, 当然, 我们并不排除得到的只是局部最优解而非全局最优解的可能性.

(二) 误差分析

1. 建模误差

根据实际情况分析, 可知公式 (2) 的推导带有一定误差, 它实际上舍弃了 Taylor 展开时的七个变量的一阶无穷小量, 其依据是 $\frac{\Delta x_i}{x_i} \leq 0.1$, 虽然较小, 但仍代有一定误差.

2. 计算机截断误差

计算机在进行求解时位长有限, 有一部分数值会被舍弃, 但对本模型基本上可乎略.

3. 受网格法的步长限制不能搜索到准确数值, 该误差可能较大, 但通过模型求解部分的探测性移动方法可使该误差减小;

同时, 龙贝格数值积分也会有误差, 程序中的积分精度为 10^{-5} .

六、模型的评价 (略)

参 考 文 献

- [1] 范大茵、陈永华, 概率论与数理统计, 浙江大学出版社, 杭州, 1996.
- [2] 肖明耀, 误差理论与应用, 计量出版社, 北京, 1985.
- [3] 席少霖、赵凤治, 最优化计算方法, 上海科学技术出版社, 上海, 1983.
- [4] 徐士良, C 常用算法程序集 (第二版), 清化大学出版社, 北京, 1996.