



2016 湖南省研究生数学建模竞赛参赛承诺书

我们仔细阅读了湖南省研究生数学建模竞赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们授权湖南省研究生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号是（从组委会提供的试题中选择一项填写）：A 题

我们的参赛报名号为（如果组委会设置报名号的话）：A201518001019

所属学校（请填写完整的全名）：国防科学技术大学

参赛队员（打印并签名）：

1. 张逢哲
2. 张帆
3. 张佩俊

指导教师或指导教师组负责人(打印并签名):

日期： 2016 年 4 月 19 日

评阅编号（由组委会评阅前进行编号）：

2016 湖南省研究生数学建模竞赛

编号专用页

评阅编号（由组委会评阅前进行编号）：

评阅记录（可供评阅时使用）：

[illegible]

湖南省第二届研究生数学建模竞赛

题 目

国家电网电容器投标报价策略

摘

要：

投标竞价是市场经济条件下工程活动中主要的竞争形式，它将竞争机制引入工程项目中，促进了企业之间的竞争。随着竞争机制的完善，企业投标中标难度变得越来越大，即使公司中标，其利润空间也在逐步缩小，因此研究采购方的评价标准、预测其他企业的投标价格以提高其中标率显得越来越重要。

本文从投标企业的角度出发，结合国家电网电容器投标报价的评标方法，首先分析了招标方的区间平均下浮双边曲线算法。评分算法体现了采购方所期望的投标价格与参与的投标公司所提供的价格之间的关系，研究了在总体报价已知的情况下，个体得分随着报价的关系，得出了算法的价格下浮比例（或称下浮系数）以及减分速率指数对个体报价得分的影响，并且对投标企业给出几点分析意见。

题目给出了从 2013 年第 5 批至 2014 年第 3 批共 5 批数据，数据包中含有大量的待处理数据。为了简化数据处理，在不降低计算精度的前提下假设：在同一批次，同一公司相同产品的价格相等；从而极大地降低了数据处理难度。由于产品的价格随时间处于不断变化之中，所以求解了所有批次的数据，得到所有批次中所有产品的价格。之后采用了灰色模型、Logistic 预测模型，预测出 2014 年第四批产品的价格，继而得到该批次中包的报价。再用熵权组合模型，使用信息熵思想对各种预测算法得到的预测结果进行合理加权组合，得到最精确的预测结果。

利用文中建立的数学模型，分别计算出前 5 批次各个公司的报价结果。从得到的计算结果可以发现，所建立的数学模型得到的结果与真实的报价数据之间的偏差一般在 5% 以内，从而可以得到这样的结论：文中提出的模型的计算精度比较高，可以很好地吻合已知数据。在采用多种预测算法并进行加权计算后，得到的预测结果精度很高，可以用于预测实际的报价结果。

在得到各个公司的预测报价后，为了计算出使得特定公司的中标概率最大、报价最高的最优报价，采用了遗传算法对优化变量进行寻优计算。优化计算中的优化变量为特定公司的报价，性能指标为该公司报价的得分最高，从而得到公司对于某包的最优报价。计算得到的最优报价均可以使得该公司的报价得分在 59.99 分左右，其它公司的报价得分均小于该得分，因此采用优化算法后，公司的中标概率得到大幅提升。因此，可以将文中所建立的数学模型应用于实际报价活动中，使得公司的中标概率达到较高的水平，并使其报价最高，这对于企业的良好运营起到至关重要的作用。

最后我们总结了本次建模过程中遇到的问题，分析了本模型的优点和不足，之后对这个问题做出了进一步探索与讨论。

关键字：最小二乘，熵权组合预测模型，灰色预测，Logistic 预测，遗传算法

一、问题重述

1.1 问题背景

随着社会的发展，工程项目的投竞标环境不断的规范化，所面对的经济对象也越来越复杂。投标报价过程中，报价过高很有可能失去中标机会；报价过低，则会造成中标后企业的亏损，因此有效地投标报价策略对于厂家的经济效益显得尤为重要。

国家电网 2013 至 2015 年每年举行 6 批电容器类货物的招标，主要的投标厂家有 17 家左右。前两年，国家电网在综合评标的基础上进行招标，该方法主要对厂家的价格因素进行打分，厂家所得的评分最高则中标。

从卖方即投标厂家的角度来说，实际投标过程是随着市场环境和自身条件的变化而变化的，即同一包同一个厂家在不同批次报价可能不同，故厂家竞标过程中需要一个合理的报价模型以提高中标率。但目前尚未找到一种行之有效的报价预测方法，如能较准确的预测下一批各个厂家对不同包的报价，综合分析国家电网的招标策略，提高投标报价的合理性，将对厂家的生存发展具有非常重要的现实意义。

1.2 问题提出

附件中提供了货物发包清单，价格具体得分和国家电网电容器规定的最高限价等信息，问题要求根据题目国家电网电容器投竞标模式，了解相关背景知识，收集资料，建立合理数学模型求解以下四个问题：

一，对国家电网评标时采用的区间平均下浮双边曲线算法做出全面研究分析，给出对投标方有价值的研究结论。

二，分析合容公司在 2013 年第 5 批至 2014 年第 3 批共 5 批投标相关数据，建立该公司合理的报价模型以提高中标率。

三，根据所建立的模型，给出合容电气对 2014 年第 4 批以下指定各包的具体报价：包 24，包 29，包 41，包 42，包 57，包 62，包 66，包 71，包 74，包 76，包 84，包 87。

四，在分析研究的基础上给出关于国家电网招投标方法改革的合理化建议。

二、问题分析

针对问题一，我们准备从众多包中任选取一组数据，任选一组数据具有普适性，代入区间平均下浮双边曲线算法模型之中，做出投资方的分数随着算法模型的三个参数的变化的图，再改变某一个投标方的报价，分析所得分数随投标方报价的变化，做出总结。

针对问题二，我们做出如下分析：

1：分析某一批次，每一个包的物资组成，其组成成分如下图所示：

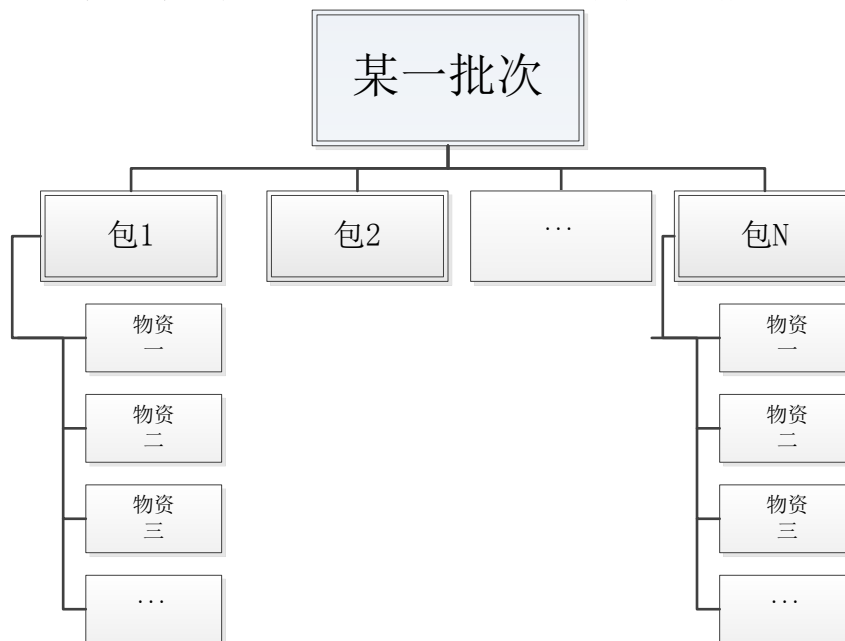


图 1 某一批次包及物资组成结构

2：每一个包的价格都不相同，其原因归根结底为每一个包内的所需物资不同，为此我们做如下合理假设：合容电气公司每一个季度某件商品的价格保持稳定不变。

3：采用最小二乘估计的方式，物资规格是一定的，假设每一种不同的规格的商品价格不同，根据这些包可以最小二乘估计出每一个批次采购期间，不同规格的商品的价格。

4：不同批次采购的物资规格是一样的，根据不同的批次最小二乘估计出的单个商品价格，再用熵权法预测出所需要估计的批次的商品价格，根据那些单价乘以所需估计的包内的物资数量，得出其估计价格。

针对问题三：根据问题一和问题二可以得出如下结论：

1：若已知其他人的出价，可以搜索出一个最好的报价，可以让投资方满分中标。

2：统计参与竞标频率最多的厂商，取前 15 名，假设需要预测的几个包他们会参与竞争，其他的投标方的出价可以通过问题二的模型计算得出。

针对问题四：我们可以通过问题一的分析过程和问题二建模过程，给出这个问题的意见。

三、问题假设

1. 假设对于同一种产品，在同批次中，其期望价格也是相同的；
2. 假设对于同一种产品，在不同批次中，其期望价格是不一定相同的；
3. 假设从 2013 年第五批次至 2014 年第四批次，每次的投标时间间隔基本均匀；
4. 假设在每次投标过程中，虽然同一厂家对同一个包在不同批次的报价不同，但是在基于已有的数据和经验基础上进行报价。

四、符号说明

符号	定义
p	投标人的评标总价，即报价
a	价格下浮比例（或称下浮指数）
m	减分速率指数 m
n	减分速率指数 n
$p_n, p_{n+1}, p_{n+2} \cdots p_m$	竞标报价
$N_{n,1}, N_{n,2}, N_{n,3} \cdots N_{n,var}$	产品的数量

五、模型的建立与求解

5.1 问题一的分析与解答

为了对国家电网评标时采用的区间平均下浮双边曲线算法做全面分析，我们选取以下一组数据包来讨论。故我们选取 2014 年第二批包 106 号的具体得分，该组数据包多个商家的报价在有效报价区间内，对分析算法具有代表性。

对于国家电网采用的区间平均下浮双边曲线算法而言，总共有四个输入变量：所有厂家的报价 P ，下浮系数 a ，减分速率指数 m 、 n 。下面分别研究这几个量的变化关系。

5.1.1 限制变量 P ，依次改变 a, m, n 的值

此时，可以做出如下所有人报价 P 不变的情况下， a, m, n 对得分变化趋势图。

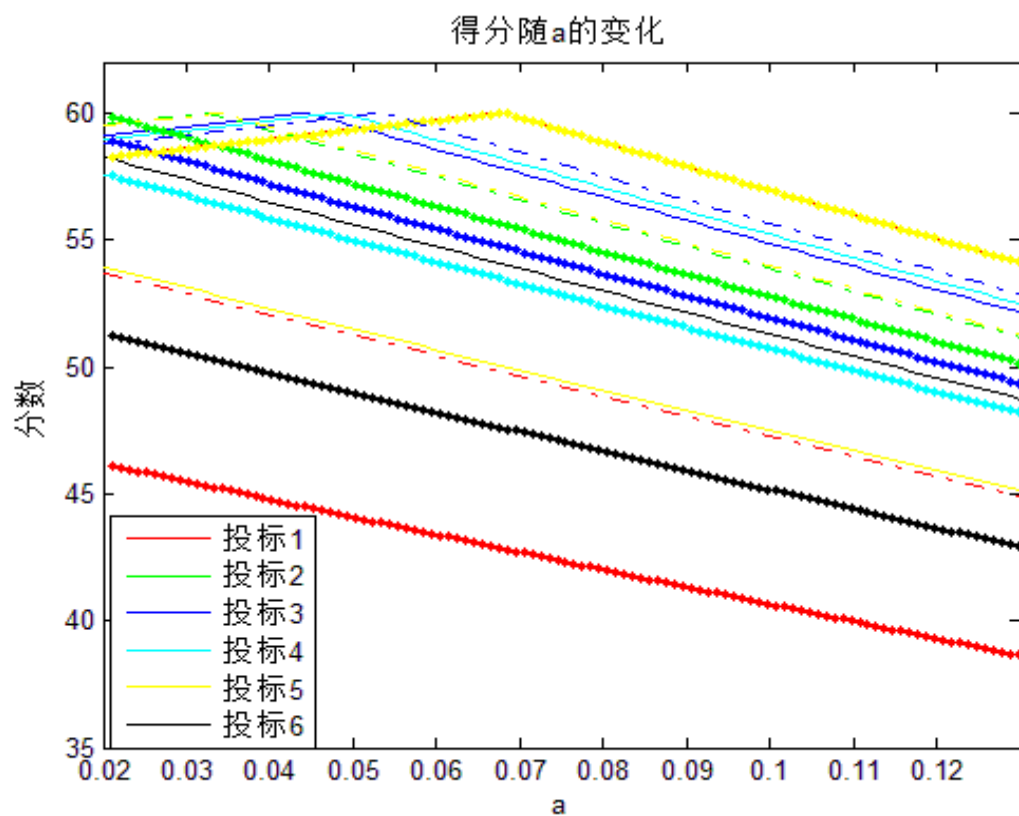


图2 得分随 a 的变化

分析上图2可以看出，改变 a 的值，不同的报价会有不同的得分，总体而言，a 较大的时候，高报价得分下降趋势较大。

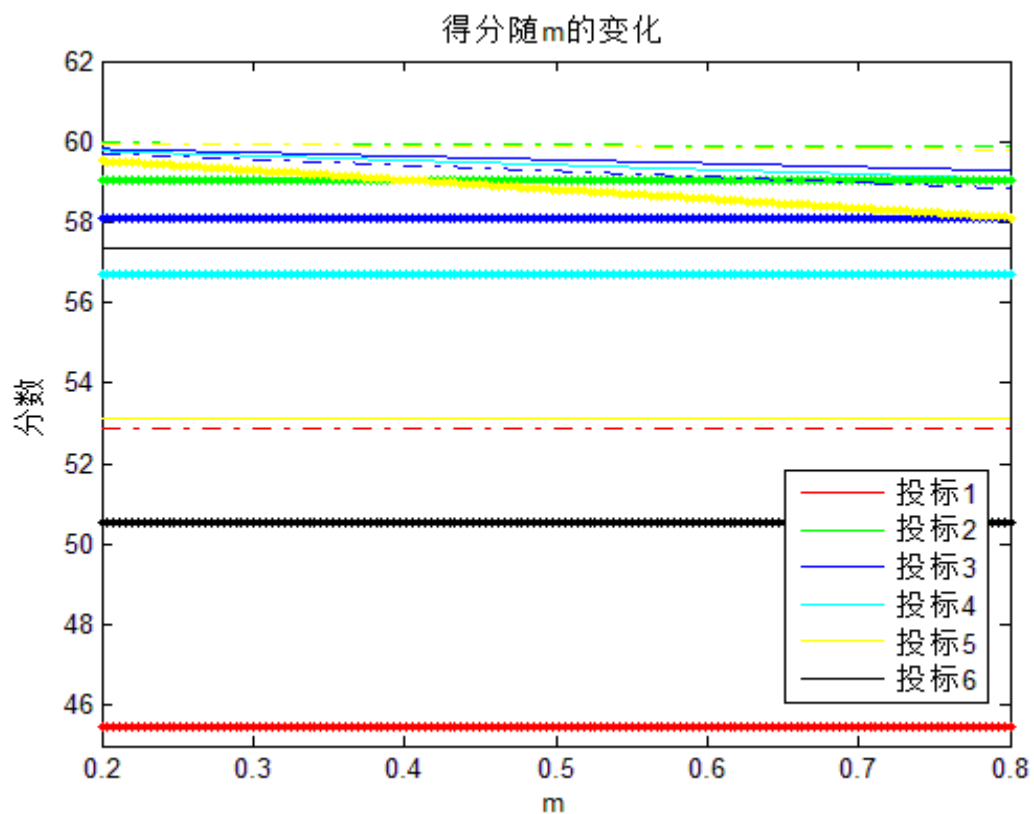


图3 得分随 m 的变化

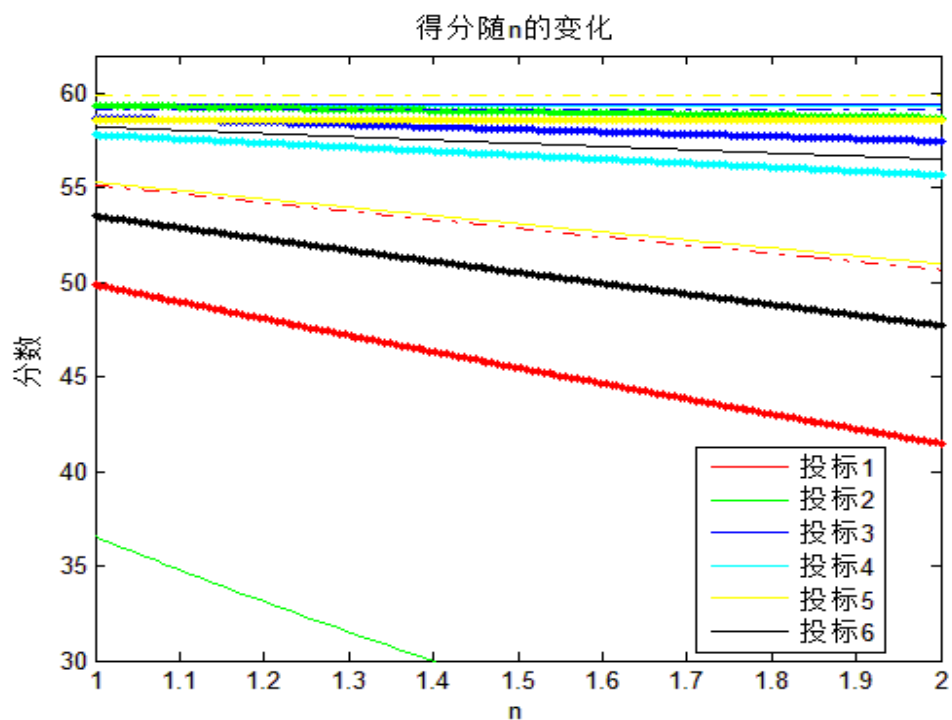


图 4 得分随 n 的变化

上图 3 和图 4 可以看出：部分商品的价格受到 n 值的影响而改变，随着 n 变大，分数降低。

5.2.2 限制其他厂家报价不变，改变某一投标商报价 P_i

当其他厂家报价不变的情况下，有效报价区间、某一投标商得分随其报价的变化如下图。

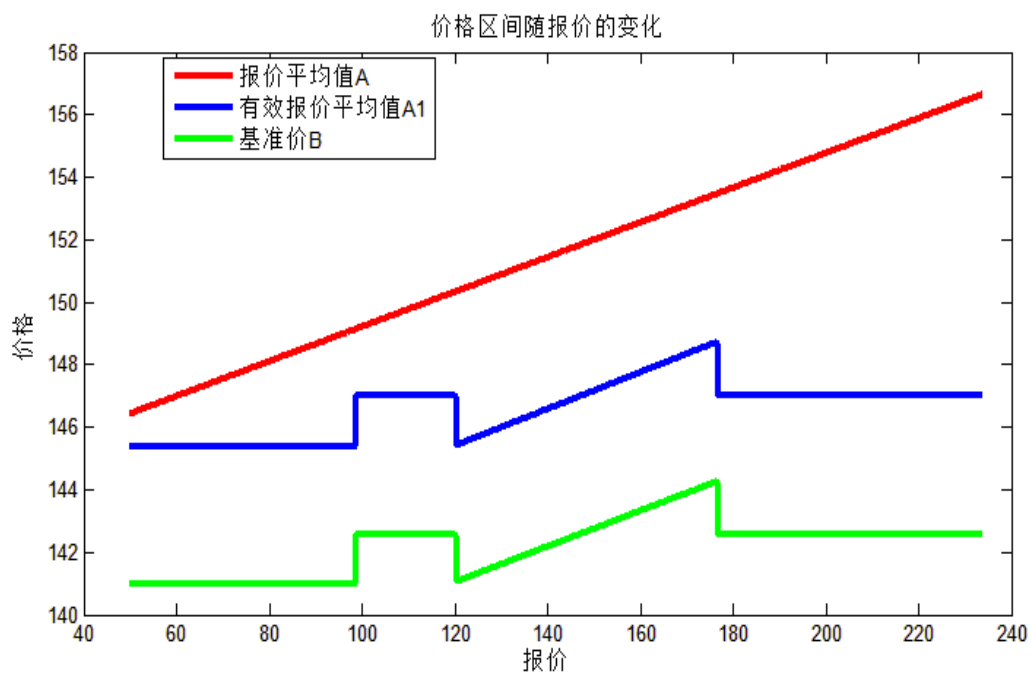


图 5 有效报价区间随某一投标商报价 P_i 的变化

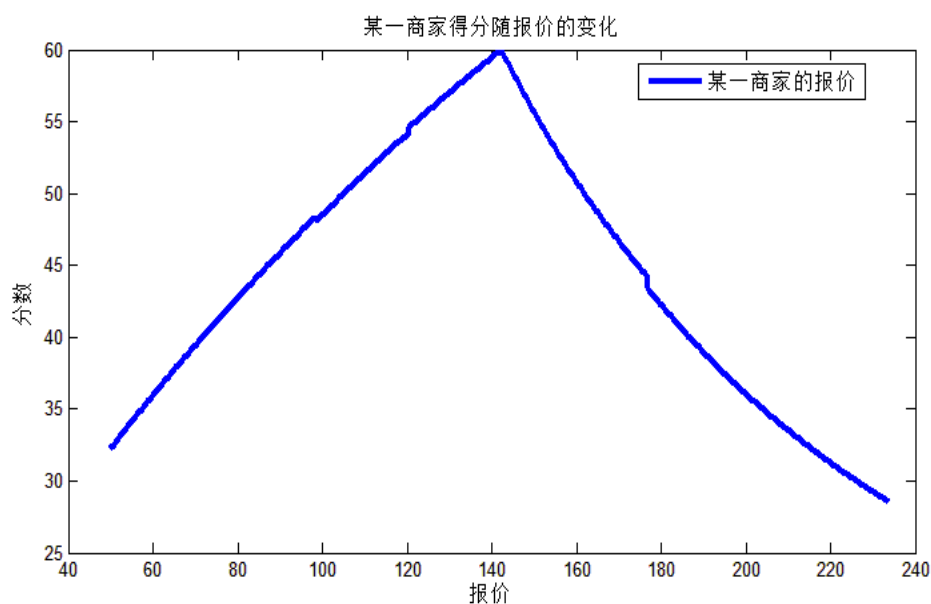


图 6 某一投标商得分随报价的变化

通过图 5 和图 6 可以看出，若假设其他厂家报价已知而且不变的情况下，某一投标商自己的报价得分在下浮系数及减分速率指数已知为 0.03；0.6；1.5 的情况下：

1. 当得分在低于基准价 B ，若报价提高的话，则得分上升，而且上升曲线为凸曲线；
2. 当价格高于基准价 B 的时候，报价上升导致得分下降，而且下降为凹函数。

这表明提出项目方有如下心理：希望自己购买的产品价格不能太低以免买到质量不高的商品，同时又希望自己购买的产品质量达标的情况下价格尽量低。由图可以看出，当商家价格低于基准价 B 的时候的曲线是上凸函数说明增加相同价格，提高分数的比例在下降。当商家价格高于基准价 B 的时候的曲线是下凹函数说明，提高价格导致分数下降的比例在增大。意味着买家不希望卖家漫天要价。

5.2.3 某一厂家得分随 a , m , n 的变化

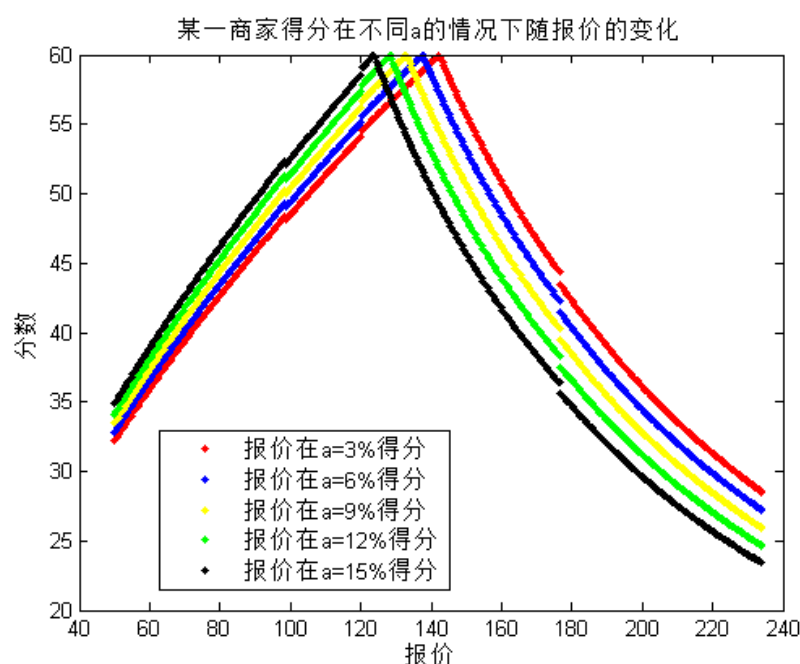


图 7 某一商家得分在不同 a 的情况下随报价的变化

本商家自己的报价得分在减分速率指数保持为 0.6; 1.5 的情况下, 改变下浮系数可以得到上图。上图中可以看出下浮系数越高, 得到满分 60 分的价格越低, 这代表项目方希望购买的商品价格越低越好。

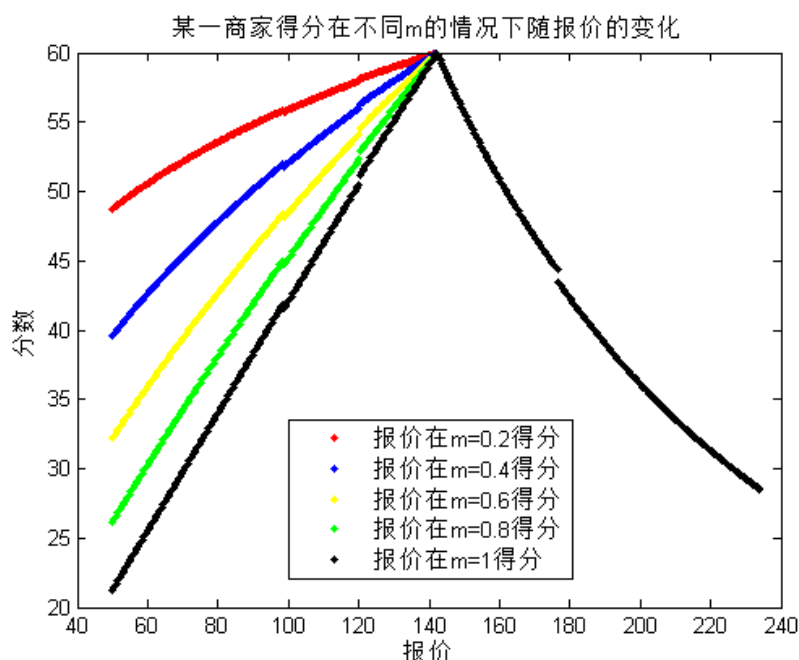


图 8 某一商家得分在不同 m 的情况下随报价的变化

本商家自己的报价得分在下浮系数 a 保持为 0.03; 减分速率 n 保持为 1.5 的情况下改变减分速率 m 可以得到上图。上图可以看出减分速率 m 不会影响商家报价高于 B 的时候的得分下降速率, 而是影响商家价格低于 B 时候的得分上升速率。 m 越大, 得分随

着报价上升的速率越大，但是相同报价得分低于 m 值小的情况。

可以做出如下判断， m 越小，商家在低于 B 的价格情况下提高报价得分提升意义不大，当 m 越大的时候，商家低于 B 的价格情况下提高报价得分提升快，意味着买家希望价格不要太低。

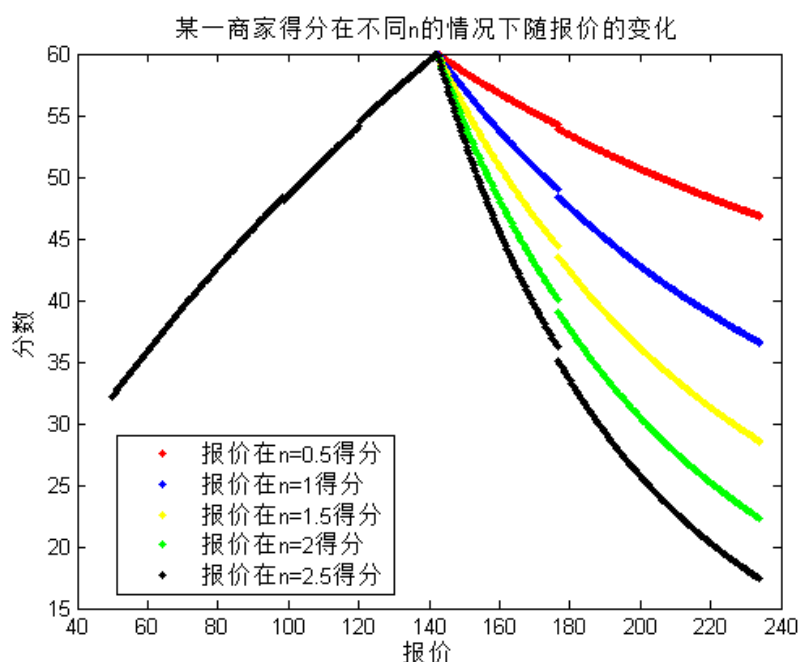


图 9 某一商家得分在不同 n 的情况下随报价的变化

本商家自己的报价得分在下浮系数 a 保持为 0.03；减分速率 m 保持为 0.6 的情况下改变减分速率 n 可以得到上图。上图可以看出减分速率 n 不会影响商家报价低于 B 的时候的得分上升速率，而是影响商家价格高于 B 时候的得分下降速率。 n 越大，得分随着报价下降的速率越大，相同报价得分低于 n 值小的情况。

可以做出如下判断， m 越小，商家在高于 B 的价格情况下提高报价得分降低速率不大，当 m 越大的时候，商家高于 B 的价格情况下提高报价得分下降很快，意味着买家希望价格要低。

5.2.4 结果总结

从买家角度而言：国家电网评标时采用的区间平均下浮双边曲线算法既能保证购买的商品的质量不会太差，也压低了差不多质量的商品的价格。

从投标方角度而言：若能根据以往的其他商家报价，推测出某次某个包他们的报价的，并且已知下浮系数及减分速率指数的情况下，可以得出一个价格相对较高，评分也相对较高的价格作为自己的报价参与竞标。

具体而言：投标方的价格需要预测其他投标方的出价，给出合适的投标价格。投标价格若超出有效投标报价区间的话，则不会影响到基准价 B ，基准 B 是作为得分计算公式的评价指标的，己方价格超出有效投标价格的情况下：投标人的评标总价 $P \geq$ 基准价 B ，得分 = (基准价 B / 评标总价 P) $n \times 100$ ；当投标人的评标总价 $P <$ 基准价 B ，得分 = (评标总价 P / 基准价 B) $m \times 100$ 。得分都会比较低。若己方价格在有效投标报价区间内时， B 会向自己的报价靠拢，则根据这个评分计算公式得分会上升。

投标方需要根据买方公布的下浮系数及减分速率指数来调整自己的报价策略。当买方公布的下浮系数较高时，代表买方希望自己购买的商品价格较低，此时投标方出价应

该要在预测大家出价的基础上,进行下调,以免于基准价 B 相差过大;当买家公布的减分速率指数中的 m 较高时,代表投标方低价得分较低,此时应该适量向上提价;当买家公布的减分速率指数中的 n 较高时,价格高于 B 的情况下分数降低很快。此时应当出价保守。

5.2 问题二的分析与解答

以合容电器公司的厂家报价模型为研究对象,详细的给出了以下模型的建立与求解。

在问题描述中,不同公司所处地理位置不同,其自身管理状况、技术水平、任务饱满量与原材料成本方面等均具有很大的差异性,从而对于同一种产品,不同公司的报价必然存在差异。除此之外,不同公司对于特定项目的利润期望值均有所不同。所以在设计算法时,必须考虑到这些因素的影响情况。事实上,这些参数直接影响着在第四批次不同公司报价预测的精度。在算法设计与模型建立时,必须将这些因素都考虑进来,这样的模型才是合理真实的。

5.2.1 算法主要思想

由于已有公司在五个批次,每个批次有 100 个左右包的报价数据,该数据提供了预测第四批次各个公司报价的原始材料。依据已有的数据,对第四批次其余各个公司的报价进行合理的预测。在得到其余公司的报价数据后,采用优化算法对合容公司的报价进行优化计算,从而得到使得合容公司中标概率最大且报价最高的报价值。

分析题目要求可以发现,在每个批次中,不同包里面的产品类型或者数量均存在差异,从而若直接以包作为单位,对第四批次特定包的报价进行预测具有很大的误差。因此算法必须考虑到在特定批次中,公司对每种产品的期望价格。算法中假设在一个批次中,所生产的同一种产品的成本与附加值相等,从而同一种产品的期望价格为不变量。除此之外,即使是对于同一种产品,在不同批次中,其期望价格也是不相同的,这主要是由于随着时间变化,生产该种产品的劳动力价格、原材料、市场价格、企业经营状况等均存在波动。因此必须计算出不同批次中同一种产品的期望价格,利用期望价格随时间的变化趋势,得到下一次该种产品的期望价格数值。之后利用该期望价格数值乘以目标项目中该种产品的数量,进行求和计算,即可得到其余企业在特定批次中的报价值。这样计算的好处在于所计算出的报价更加精确与合理:

1. 充分利用各个批次中所有包的数据(仅包含符合题目要求的包,不符合题目要求的包已被去掉);

2. 考虑到每个包里面产品的价值属性,对每种产品的期望价值进行建模,从而使得对于所含不同产品的包的价格估计更加可信。

3. 算法中考虑了每种产品的期望价格随时间变化这一因素,采用合理的预测算法对特定批次中该种产品的期望价格进行预测,使得算法更加合理。

5.2.2 算法流程

1. 某一批次中 A 公司的对于某一种产品的期望价格的算法:

假设在该批次中, A 公司共参与包 n 至包 m ($m \geq n$) 的竞标活动(该区间需要满足题目的计算要求),共产生 $m-n+1$ 个竞标报价,分别记为: $p_n, p_{n+1}, p_{n+2} \cdots p_m$ 。对于包 n ,其中每种产品的数量分别计为 $N_{n,1}, N_{n,2}, N_{n,3} \cdots N_{n,var}$, var 为在该批次中,所含有的产品

种类总数，每种产品的期望价格 $x_1, x_2, x_3 \dots x_{\text{var}}$ 为待求量。从而可以得到含有 var 个未知变量，共 $m-n$ 个等式的方程组。分析已有数据可以发现，对于各个批次中，所含有的未知变量个数小于所建立的等式的个数($m-n+1 > \text{var}$)，从而一下方程组为超定方程组：

$$\begin{aligned} N_{n,1}x_1 + N_{n,2}x_2 + N_{n,3}x_3 + \dots + N_{n,\text{var}}x_{\text{var}} &= p_n \\ N_{n+1,1}x_1 + N_{n+1,2}x_2 + N_{n+1,3}x_3 + \dots + N_{n+1,\text{var}}x_{\text{var}} &= p_{n+1} \\ N_{n+2,1}x_1 + N_{n+2,2}x_2 + N_{n+2,3}x_3 + \dots + N_{n+2,\text{var}}x_{\text{var}} &= p_{n+2} \\ &\vdots \\ N_{m,1}x_1 + N_{m,2}x_2 + N_{m,3}x_3 + \dots + N_{m,\text{var}}x_{\text{var}} &= p_m \end{aligned} \quad (1)$$

写为矩阵形式为：

$$\begin{bmatrix} N_{n,1} & N_{n,2} & \dots & N_{n,\text{var}} \\ N_{n+1,1} & N_{n+1,2} & \dots & N_{n+1,\text{var}} \\ \vdots & \vdots & \ddots & \vdots \\ N_{m,1} & N_{m,2} & \dots & N_{m,\text{var}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\text{var}} \end{bmatrix} = \begin{bmatrix} p_n \\ p_{n+1} \\ p_{n+2} \\ \vdots \\ p_m \end{bmatrix} \quad (2)$$

简记为：

$$A_i x_i = y_i \quad (3)$$

其中 i 为该批次的顺序，其余符号说明如下：

$$A_i = \begin{bmatrix} N_{n,1} & N_{n,2} & \dots & N_{n,\text{var}} \\ N_{n+1,1} & N_{n+1,2} & \dots & N_{n+1,\text{var}} \\ \vdots & \vdots & \ddots & \vdots \\ N_{m,1} & N_{m,2} & \dots & N_{m,\text{var}} \end{bmatrix} \quad (4)$$

$$x_i = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\text{var}} \end{bmatrix}, y_i = \begin{bmatrix} p_n \\ p_{n+1} \\ p_{n+2} \\ \vdots \\ p_m \end{bmatrix} \quad (5)$$

对该超定方程组进行求解即可得到在该批次中不同种类产品的期望价格，但由于无法直接将系数矩阵 A_i 求逆得到未知变量的结果，为了更加准确地得到未知变量的值，采用最小二乘法进行求解：

$$x_i = (A_i^T A_i)^{-1} A_i^T y_i \quad (6)$$

至此可以得到在批次 i 中， A 公司所有产品的期望价格 x_i ，重复此步骤直到将参与该批次的所有公司的所有产品的期望价格计算出来。

在得到批次 i 中所有公司的所有种类产品的期望价格后，重复以上步骤，直到所有批次中所有公司的所有种类产品的期望价格

$$x_i^A, i=1, 2, \dots, 5 \quad (7)$$

从 2013 年第 5 批次至 2014 年第三批次共 5 个批次，其中 A 表示 A 公司， i 表示 i 批次。

在得到 $x_i^A, i=1, 2, \dots, 5$ 后，对第 2014 年第 4 批次的 x_i^A 进行预测。在得到 x_i^A 后，将该批次中特定包中产品的数量 N_i 乘以 x_i^A 即可得到所有公司对于该包的报价数据。在进行报价预测时，采用多种算法进行预测，然后计算出不同算法的拟合方差，依据最小方差无偏估计方法计算出最优的估计结果。该结果即为各个公司的最终报价结果。本文使用的预测算法有灰色预测模型和 Logistic 预测模型模型。

在得到各个公司对于 2014 年第 4 批次的预测报价后，优化合容公司的报价数值，使得其报价最高且得分最高，从而使得其中标的概率最大。该步骤中采用的优化算法有遗传算法与序列二次规划算法，优化变量为合容公司的报价，性能指标函数为合容公司的报价得分。至此即可得到最优的报价数值，该结果既能使得合容公司的获利最大，中标概率最大，符合题目要求。

该模型的计算流程图如下所示：

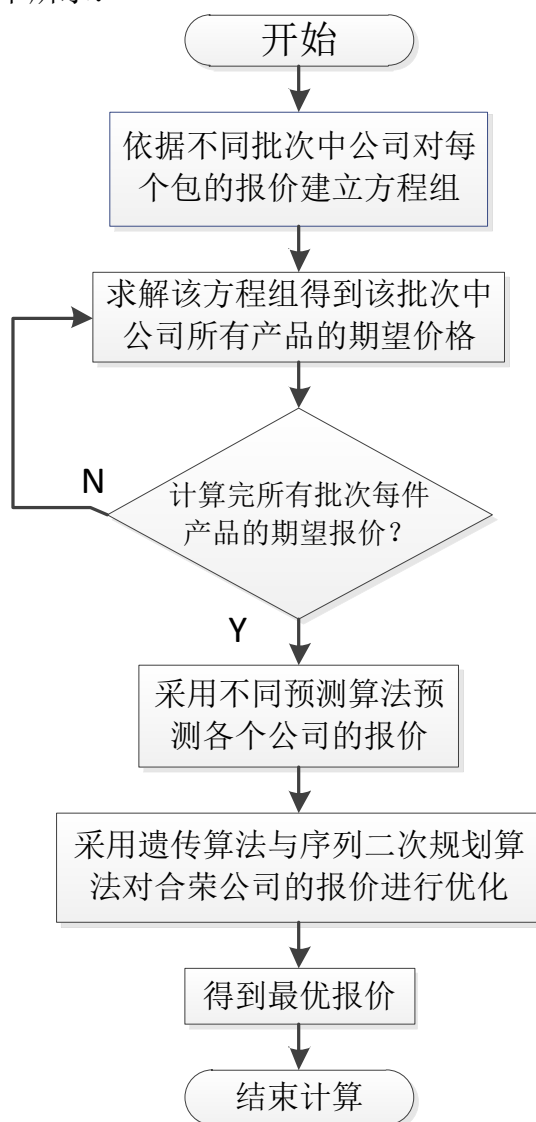


图 10 模型计算流程

5.2.3 熵权组合预测模型

考虑合容公司投标，在 2013 年第 5 批至 2014 年第 3 批共 5 批数据进行分析的基础上，对于下一批报价进行预测。由于对考虑问题的出发点和层次不同，可以建立多个单项预测的模型，而每个模型都有其相应的适用性，熵权组合预测综合利用了各个单项模型所提供的信息，是一种既不否定单项预测也区别于传统预测模型的新型预测方法，可以减少避免指标所确定的经验化。

熵权组合预测模型的关键在于确立各个预测方法的权重。本文将从信息论的角度，

根据个体预测方法误差指标的信息熵，确定各个模型所占的权重。

熵权组合预测模型建立

1、熵权法的概念及基本步骤

熵权法是一种决定指标的方法，我们知道，综合指标取决于单个指标数的确定，一般情况下的权重是根据经验来确定的，但是这种确定权重的方法缺少科学根据，也不能保证确立的综合指标能反映原始指标的大部分信息，且权重确立因人而异，所以其应用受到了限制，而熵权法就能够避免这些问题，使权重确立具有科学的根据，具有说服力。熵权法的步骤如下所示：

1) 第 j 项指标下第 i 个方案的指标比重

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^m y_{ij}} \quad (8)$$

2) 计算指标 j 的熵值

$$e_j = -k \sum_{i=1}^m p_{ij} \ln p_{ij} \quad (k = \frac{1}{\ln m}) \quad (9)$$

3) 计算第 j 项指标的差异系数

$$g_j = 1 - e_j \quad (10)$$

4) 定义权重

$$w_{ij} = \frac{g_j}{\sum_{i=1}^m g_j} \quad (11)$$

2、误差指标的选举

为了能全面的各个预测方法以及组合预测的预测效果，必须制定一套切实可行的误差指标。

① 平方和误差

$$SSE = \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (12)$$

② 平均绝对值误差

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (13)$$

③ 均方误差

$$MSE = \frac{1}{n} \sqrt{\sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (14)$$

3、组合模型权重的确定

假定有 m 个体预测方法， n 个误差指标， m 种个体预测方法对应 n 个误差指标构成了评价指标值矩阵；

$$R = (r_{ij})_{m \times n} \quad (15)$$

第 j 个指标下第 i 种个体方法的指标比重值 P_{ij} 为

$$P_{ij} = r_{ij} / \sum_{t=1}^m r_{tj} \quad (16)$$

第 j 个指标的熵值为:

$$E_j = -\sum_{i=1}^m P_{ij} \ln P_{ij} \quad (17)$$

记

$$e_j = \ln E_j \quad (18)$$

第 i 个指标的权重为:

$$\theta_j = (1 - e_j) / \sum_{i=1}^m (1 - e_j) \quad (20)$$

记矩阵 R 中每列最优值为 r_j^* , 对该矩阵所有元素做标准化处理, 可得:

$$d_{ij} = \begin{cases} r_{ij} / r_j^* & \text{指标 } j \text{ 的指标值越大越好} \\ r_j^* / r_{ij} & \text{指标 } j \text{ 的指标值越小越好} \end{cases} \quad (21)$$

这样, 各个体预测方法的熵权评价值 λ_i , 可以表示为:

$$\lambda_i = \sum_{j=1}^m \theta_j d_{ij} \quad (i = 0, 1, 2, \dots, m) \quad (22)$$

将上式进行归一化处理, 即可以得到各个个体的权重。

以下是两个预测算法的核心思想。

一、灰色预测模型

灰色预测法是一种对含有不确定因素的系统进行预测的方法。灰色系统内的一部分信息是已知的, 另一部分信息是未知的, 系统内各因素间具有不确定的关系。灰色预测通过鉴别系统因素之间发展趋势的相异程度, 即进行关联分析, 并对原始数据进行生成处理来寻找系统变动的规律, 生成有较强规律性的数据序列, 然后建立相应的微分方程模型, 从而预测事物未来发展趋势的状况。

建模的实质是求解微分方程的系数

设原有的数据序列 $X^{(0)}$ 有 n 个观察值, $X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$, 通过累加生成新序列 $X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)\}$, 则 GM (1, 1) 模型相应的微分方程为:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu \quad (23)$$

其中 a 和 μ 为待识别的灰参数, a 称为发展灰数, μ 称为内生控制灰数。

如果设 $\hat{\alpha}$ 为待估参数向量, $\hat{\alpha} = \begin{pmatrix} a \\ \mu \end{pmatrix}$, 可利用最小二乘法求解。解得:

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_n \quad (24)$$

求解微分方程, 即可得 GM (1, 1) 预测模型的时间响应序列为

$$\hat{X}^{(1)}(k+1) = \left[X^{(0)}(1) - \frac{\mu}{a} \right] e^{-ak} + \frac{\mu}{a}, \quad k = 0, 1, 2, \dots, n \quad (25)$$

二、Logistic 预测模型

Logistic 预测模型, 是一种概率型的非线性回归模型, 主要研究分类观察结果 y 与一些影响因素 x 之间的关系的一种多变量分析方法。该方法的定义是: 设一事件发生概率为 p , 则 $1-p$ 为事件不发生的概率, 将两者之比 $p/(1-p)$ 取自然对数 $\ln(p/(1-p))$, 这叫做 Logit 转换, 记做 $\text{Logit}(p)$, 其取值范围为 $(-\infty, +\infty)$, 把 $\text{Logit}(p)$ 记做自变量, 可

建立线性回归方程

$$\text{Logit}(p) = \alpha + \beta_1 x_1 + \cdots + \beta_n x_n \quad (26)$$

反解 p 可以得到

$$p = \frac{\exp(\alpha + \beta_1 x_1 + \cdots + \beta_n x_n)}{1 + \exp(\alpha + \beta_1 x_1 + \cdots + \beta_n x_n)} \quad (27)$$

用最大似然估计可以求解模型中参数。

基本原理：用一组观察的数据来拟合 Logit 模型，揭示出若干个 x 与一个因变量的关心，反映 y 对 x 的依存关系。

5.2.4 寻优算法

一、遗传算法

基本原理：遗传算法是一种典型的启发式算法，属于非数值算法范畴。它是模拟达尔文的自然选择学说和自然界的生物进化过程的一种计算模型。它是采用简单的编码技术来表示各种复杂的结构，并通过对一组编码表示进行简单的遗传操作和优胜劣汰的自然选择来指导学习和确定搜索的方向。遗传算法的操作对象是一群二进制串（称为染色体、个体），即种群，每一个染色体都对应问题的一个解。从初始种群出发，采用基于适应度函数的选择策略在当前种群中选择个体，使用杂交和变异来产生下一代种群。如此模仿生命的进化进行不断演化，直到满足期望的终止条件。

运算流程：

- Step 1: 对遗传算法的运行参数进行赋值。参数包括种群规模、变量个数、交叉概率、变异概率以及遗传运算的终止进化代数。
- Step 2: 建立区域描述器。根据轨道交通与常规公交运营协调模型的求解变量的约束条件，设置变量的取值范围。
- Step 3: 在 Step 2 的变量取值范围内，随机产生初始群体，代入适应度函数计算其适应度值。
- Step 4: 执行比例选择算子进行选择操作。
- Step 5: 按交叉概率对交叉算子执行交叉操作。
- Step 6: 按变异概率执行离散变异操作。
- Step 7: 计算 Step 6 得到局部最优解中每个个体的适应值，并执行最优个体保存策略。
- Step 8: 判断是否满足遗传运算的终止进化代数，不满足则返回 Step 4，满足则输出运算结果。

二、序列二次规划（SQP）

基本原理：序列二次规划算法是一种非线性规划问题，该方法可以将一些约束添加到某些变量中，如果初始值不满足约束，那么优化算法迭代后，同样可以生成满足约束的新的值。

Matlab 中调用此类非线性问题的数学模型可以写成如下形式

$$\begin{aligned} \min & f(x) \\ \begin{cases} Ax \leq B \\ Aeq \cdot x = Beq \\ C(x) \leq 0 \\ Ceq(x) = 0 \end{cases} \end{aligned} \quad (28)$$

其中 $f(x)$ 是标量函数， A, B, Aeq, Beq 是相应维度的矩阵和向量， $C(x), Ceq(x)$ 是非线性

向量函数。在求解过程中，我们用到 `Fmincon` 命令直接调用序列二次规划算法，直接将非线性优化问题转化为二次规划问题求解。

5.2.5 模型精确性验证

首先对题目所给的数据进行处理，提取出有用的数据，建立起各个公司在各个批次中每件产品的价格与其报价数据之间的等式关系，建立多组等式方程，构成产品价格与报价之间的等式方程组。该组方程为超定方程，为了更加逼近地求得各个产品的价格，采用最小二乘法进行求解计算。对于每个公司而言，根据已有的 5 批数据可以建立起 5 个超定方程组，对该五个超定方程组进行分别求解，从而得到该五个批次时刻对应的每件产品的价格数据。为了验证求解的产品单价的准确性，将计算得到的产品单价分别乘以相应产品的数据，得到公司的报价结果。并将该计算结果与实际的报价数据进行对比，查看计算结果与真实值之间的偏差情况。为了节省篇幅，仅画出 2014 年第一至第三批次的计算结果偏差图像，如下所示。

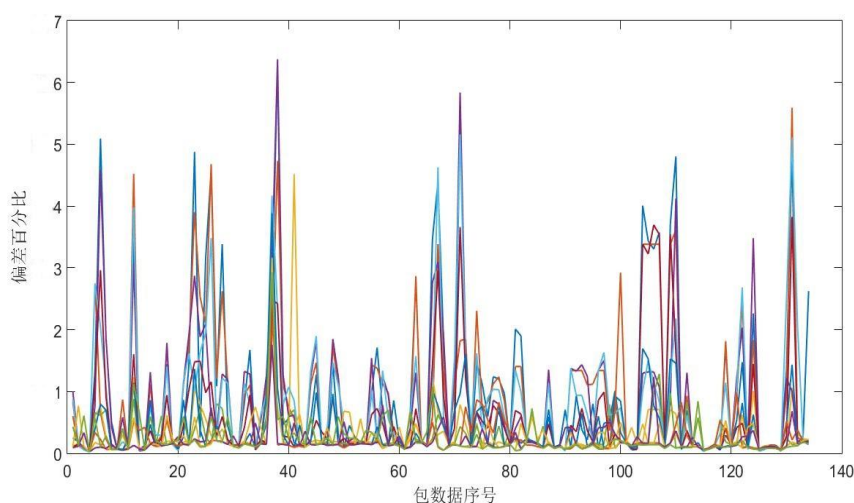


图 11 2014 年第一批次各个公司的实际报价与模型计算结果偏差

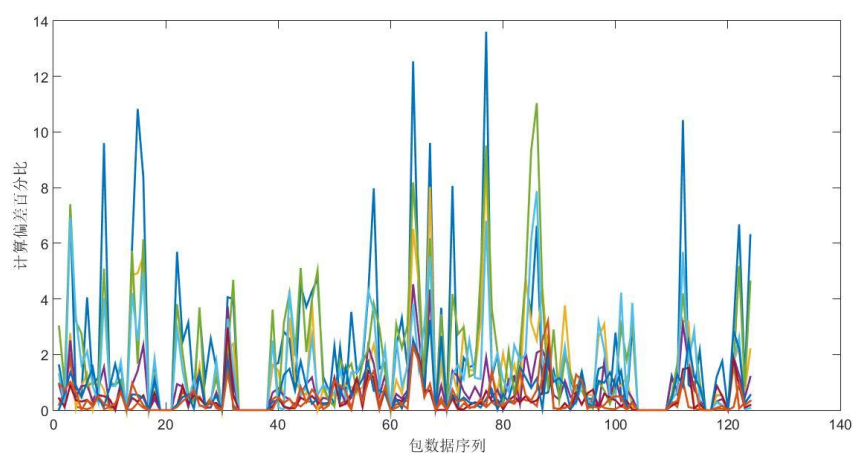


图 12 2014 年第二批次各个公司的实际报价与模型计算结果偏差

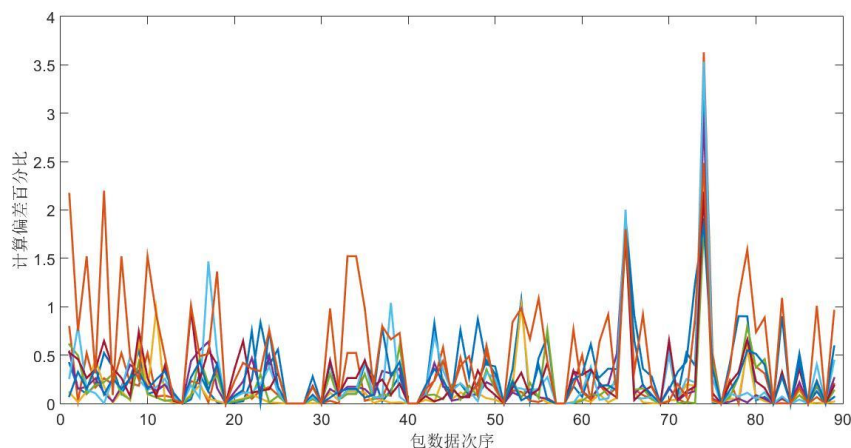


图 13 2014 年第一批次各个公司的实际报价与模型计算结果偏差

图中横轴为各个批次中包数据的序列号，纵轴为各个公司的实际报价与模型计算的结果偏差的百分比。从图中可以看出，所提出的模型的计算结果与实际计算的结果偏差绝大多数在 4% 以内，因此所提出的模型的拟合精度很高，能够很好地描述公司的报价与产品之间的关系，可以应用该模型对此问题进行求解，并且计算精度足够。

在完成以上工作后，基本建立起在每个批次中，每个公司的所有种类产品的价格计算模型，并且该计算模型具有很高的可靠性。接下来要对五个批次中所计算出的每件产品的价格，通过合理的算法推测出 2014 年第 4 批次时各个厂商每件产品的价格。而对于各种预测算法而言，其对不同问题的处理能力存在一定的差异，因此可以采用多种预测算法，对 2014 年第 4 批次中各个厂商每件产品的价格进行预测，从而得到多个预测结果。为了更好地预测出产品的价格，需要对各种预测算法计算的结果进行加权综合。其加权系数采用信息熵的思想进行计算，从而使得计算出的加权系数更合理有效。在完成以上工作后，可以预测出 2014 年第 4 批次时各个厂商每件产品的价格，再读取该批次清单中每件产品的数量，从而得到该批次中各个公司对每包报价的预测结果。根据对其它公司的报价进行合理预测，再适用遗传算法对合荣公司的报价进行优化计算，此时性能指标为合荣公司报价的得分，优化变量为合荣公司的报价。至此，可以得到使得合荣公司中标概率最大的最高报价，该报价既能使得该公司的中标率达到最大，又能使其获利最大。当获得的历史数据更多时，最各个公司报价预测值更加准确，从而该模型更加准确，合荣公司的中标概率也更大。

5.3 问题三的分析与解答

根据第三问中建立的模型，研究合容电气公司对以下 12 个包的报价。

其中由于恒顺和桂容公司对这 12 个包内竞标的产品竞争不多，故舍去这两个公司的计算。

采用上述计算方法，计算得到在 2014 年第 4 批次时各个厂商对包 24，包 29，包 41，包 42，包 57，包 62，包 66，包 71，包 74，包 76，包 84，包 87 的报价结果，如下表格：

表1 各个主要厂家对2014年以下各包的报价预测

厂商 包号	合容	日新	永锦	西电	赛晶	思源	新东 北	泰开	上虞	ABB	顺容	迪生	中原
包 24	98.15	101.30	97.23	107.46	96.02	102.40	100.87	97.20	98.87	104.50	109.92	113.91	101.31
包 29	113.63	109.19	72.07	104.77	79.09	147.77	78.87	51.69	46.90	83.30	88.66	107.66	82.92
包 41	104.42	102.18	80.27	82.67	86.85	87.17	87.81	82.99	94.67	84.37	92.58	92.76	83.10
包 42	84.93	71.28	77.16	79.16	77.94	132.09	102.72	76.26	80.91	99.42	111.92	161.15	106.67
包 57	182.56	196.96	134.86	188.78	148.59	196.66	178.42	133.28	148.85	149.85	154.94	189.63	140.65
包 62	77.36	82.31	69.84	82.30	73.81	101.84	71.43	72.06	64.70	84.94	88.80	127.71	77.28
包 66	136.23	135.63	116.41	133.15	92.42	137.95	100.13	106.62	129.62	131.01	129.14	146.06	118.01
包 71	102.58	111.25	84.34	117.48	88.58	130.18	93.85	87.23	80.41	94.52	103.99	128.65	96.36
包 74	26.96	28.64	21.25	28.46	24.43	39.41	23.15	22.80	18.40	26.77	28.49	47.25	24.80
包 76	123.92	126.46	114.43	129.92	116.19	125.43	118.55	114.71	115.31	119.22	127.20	129.50	116.53
包 84	83.08	69.70	103.44	66.64	84.73	13.51	76.95	91.31	124.79	137.21	150.10	147.42	147.62
包 87	94.04	45.27	109.46	22.97	75.61	115.71	45.93	82.52	143.29	176.47	186.30	187.21	168.92

已知电网公司给出的每个产品的单价上限，可以得到每包产品的报价上限，表格中的数据均满足该基本条件。采用不同预测算法计算出的报价之间的差别在 3%以内，因此采用不同预测算法的预测效果相差不大。之后再利用信息熵的思想，对多个算法的预测结果进行加权计算，使得所得到的结果更加准确。

得到其余公司的报价结果后，使用优化算法对合荣公司的报价进行优化，所得到的优化结果如下所示：

表 2 合容电气对每批包的具体报价

包号	报价（万元）
包24	91.5619
包29	71.5416
包41	77.4471
包42	80.6688
包57	132.1522
包62	69.4985
包66	110.2802
包71	83.1242
包74	23.0522
包76	108.1068
包84	81.1328
包87	96.4965

此时合荣公司的报价得分均在 59.99 分以上，从而保证其中标的概率最大，获利最大，满足题目要求。

5.4 问题四的分析与解答

为了促进我国电力设备厂家的公平竞争，我们需要进一步建立和完善我国招投标法律法规。我国电力行业，包括国家电网采用的招投标方式起步较晚，在招投标的管理制度和评标方法上存在一些不足。因此国家电网需要在招投标方法上做出改革。

结合本文对国家电网电容器设备投标报价策略的研究，我们提取了以下建议：

在招投标的管理制度层面，国家电网需要明确要求对一定规模的电力设备的采购进行统一招标，集中采购。这是一种集中管理和精益管理的重要措施，国家电网招标中心，对主要设备和主要材料的招标进行汇总，统一对外发布招标信息，实行公开招标。这样，可以带来明显的规模效益，降低设备和材料的报价以及招标过程的管理成本。另外，对于评分良好的中标供应商采用框架招标，框架招标是指国家电网对中标设备供应商实行时间区间管理，即对于各个时期安排的信用良好中标供应商及其设备的价格、预估的数量等条件进行详细的规定。

在评标方法上，国家电网在 2013 年-2014 年采用综合评标法进行招标，其评标打分使用区间平均下浮双边曲线算法。这种评标方法，综合考虑技术、价格和商务所占的比例，打分过程中主要考虑价格的影响。由于过于注重价格对于得分的影响，容易忽视投标单位的技术水平和整体实力，可能影响评标结果的客观性，有些质量优秀但价格偏高的产品可能不能入选。因此，综合评标法虽然操作性相对简单但是还存在本身的不足。

随着现代管理科学和数学在现代决策系统领域的发展，我们建议引入一些新的方法到电网的评标上来。借鉴此类新方法，例如层次分析法、模糊综合评价法等，还可以将这些方法的优势结合起来，实行两阶段评价理论，基于决策者偏好的数据包络分析法等。

六、模型的评价

6.1 模型优点

1、题目中已知大量数据，为了在不降低模型的精确度的情况下，减少计算量，载模型中假设在同一批次同一种产品的价格不变，从而极大地减少计算量。

2、算法利用了批次中的所有数据，采用最小二乘算法求解某一个批次中的某个商品的价格，计算结果比较可信。

3、对于问题二的求解过程中，运用了熵权组合模型，综合利用各种预测算法得到的预测结果，与传统单一预测模型相比，具有较高的预测精度和较好的稳定性，可以减少和避免指标确定的经验化，使指标权重的确立过程具有科学的依据，具有说服力。

4、针对第三个问题，我们首先统计不同商家参与投标竞价的频率，频率高的商家假设参与到所需预报的批次的竞争之中。避免了模型的不精确和不科学。

6.2 模型缺点

1、由于要处理大量数据，影响批次中产品价格的因素很多，模型采用最小二乘法可能会带来一定的误差。

2、在使用最小二法计算时，数据提取十分复杂与耗时，计算量大。

3、模型假设在批次内同一件产品的价格是不变的，这样虽然简化了计算过程，但也可能带来一定的误差。

七、问题的进一步拓展

整个模型计算量十分巨大，商品数目繁多，可以做如下假设进一步简化算法过程

1、由附件三可以看出相同的大规格下的小的不同技术指标最高报价相同，可以近似认为这一下小的技术规格不同的商品价格一致，从而匹配大的技术规格来确定矩阵A，减少变量。

2、不同商家报价是有博弈心理的，虽然同一包报价双方都不知晓。然而模型并没有考虑博弈论的思想。假设投标方对对方的竞争毫不在意。

3、对于投标方而言，尽量保证中标的同时最求价格最高，以保证利益最大化，这个思想并没有在模型之中提现，只是寻优约束条件为得分最高。如何平衡出价过高导致不中标和出价较低中标利益不大的情况？

4、不同的项目地域因素也存在，靠项目点较近的企业交货速率，问题反馈速率也快，所以有一些地方小公司也参与了项目竞标。这个因素是否也要加入模型预测？

八、参考文献

- 【1】吴孟达，成礼智，吴翊，毛紫阳，王丹. 数学建模教程，长沙：国防科学技术大学出版社，2013
- 【2】姜启源 谢金星 叶俊，数学模型（第三版），北京：高等教育出版社，2003
- 【3】肖华勇，实用数学建模与软件应用，西安：西北工业大学出版社，2008
- 【4】王雪青，国际工程投标报价决策系统研究：[博士学位论文]. 天津大学，2003
- 【4】李潇. 基于支持向量回归的电力采购评标方法研究：[硕士学位论文]. 天津大学，2008
- 【5】张星. 非线性组合预测模型及其在期货中的应用研究：[硕士学位论文]. 西安理工大学，2014
- 【6】姚卓宏. 基于价格预测的动态投标报价研究：[硕士学位论文]. 长沙理工大学，2012
- 【7】王岩 隋思涟，试验设计与 MATLAB 数据分析，北京：清华大学出版社，2012
- 【8】徐根玖，刘洋，刘一鸣，张泽卿. 基于熵权法的高等教育生源危机预测分析[J]. 统计与决策，2014，21： 92-95.

九、附录

附录 1: Matlab 批量提取提取 Excel 信息

```
function [ A ] = ExcelManageA( filepath,filename )
%%excel 处理函数，识别产品类型和产品数量，
%%输入：filepath（所需处理的文件夹目录）
%%输出最小二乘矩阵的  $AX=B$  中的，A,B
%%读取一堆表格的方法
%%电容器_货物清单-国网 2014 年第一，三批可算
%%
filename = '电容器_货物清单-国网 2014 年第一批.xls';
filepath = 'C:\Users\zhangfan\Desktop\ProblemA\附件 2-价格具体得分\国网 2014 年第
一批得分情况\';%定义字符串
pathtemp = [filepath,filename];
[data,str,row]=xlsread(pathtemp);

typemark = cell(1,1);%产品标记
packagemark = cell(1,1);%包标记
[m,n] = size(row);%m :行数  n:列
marksize = 1;
packagesize = 1;
for i=1:1:m

    if i ==1
        typemark(1,1) = row(i,2);
    else
        for j=1:1:marksize
            logical1 = strcmp(typemark,row(i,2));
            if logical1==0
                marksize = marksize+1;
                typemark(marksize,1) = row(i,2);
            end
        end
    end
end

if i==1
    packagemark(1,1) = row(i,1);
else
    for k=1:1:packagesize
        logical2 = strcmp(packagemark,row(i,1));
        if logical2==0
            packagesize = packagesize+1;
            packagemark(packagesize,1) = row(i,1);
        end
    end
end
```



```

        end
    end
end
end
%%%%%%%%%%
%%%%%%%%%%
[typex,typex] = size(typemark);
[packy,packx] = size(packagemark);
A=zeros(packy,typex);

for i=1:1:packy
    asd = double(strcmp(raw,packagemark(i,1)));
    for j=1:1:typex
        ads = double(strcmp(raw,typemark(j,1)));
        decide = asd+ads;
        for k =1:1:m
            if norm(decide(k,1:2))==norm([1,1])
                A(i,j) =A(i,j)+ cell2mat(raw(k,4));
            end
        end
    end
end
end
end

```

附录 2：求某批次产品报价

```

%%excel 处理函数，识别产品类型和产品数量，
%%输入：filepath（所需处理的文件夹目录）
%%输出最小二乘矩阵的 AX=B 中的，A,B
%%读取一堆表格的方法
%%电容器_货物清单-国网 2014 年第一，三批可算
%%
clear
clc
filename = '电容器_货物清单-国网 2014 年第二批.xls';
[data,str,row]=xlsread(filename);

typemark = cell(1,1);%产品标记
packagemark = cell(1,1);%包标记
[m,n] = size(row);%m :行数  n:列
marksize = 1;
packagesize = 1;
for i=1:1:m

    if i ==1
        typemark(1,1) = row(i,2);
    end
end

```

```

else
    for j=1:1:marksize
        logical1 = strcmp(typemark,raw(i,2));
        if logical1==0
            marksize = marksize+1;
            typemark(marksize,1) = raw(i,2);
        end
    end
end

if i==1
    packagemark(1,1) = raw(i,1);
else
    for k=1:1:packagesize
        logical2 = strcmp(packagemark,raw(i,1));
        if logical2==0
            packagesize = packagesize+1;
            packagemark(packagesize,1) = raw(i,1);
        end
    end
end

end
end
%%%%%%%%%%%%%%
%%%%%%%%%%%%%%
[typey,typex] = size(typemark);
[packy,packx] = size(packagemark);
A=zeros(packy,typey);

%
% for i=1:1:packy
%     asd = double(strcmp(raw,packagemark(i,1)));
%     for j=1:1:typey
%         ads = double(strcmp(raw,typemark(j,1)));
%         decide = asd+ads;
%         for k =1:1:m
%             if norm(decide(k,1:2))==norm([1,1])
%                 A(i,j) =A(i,j)+ cell2mat(raw(k,4));
%             end
%         end
%     end
% end

i=1;
flag1=1;
for i=1:packagesize

    while strcmp( str(flag1,1), packagemark(i) ) %还在同一个包

        for j=1:typey

```

```

        if strcmp( str(flag1,2),typemark(j) )
            A(i,j)=A(i,j)+data(flag1);
            flag1=flag1+1;
            break;
        end %结束判断字符串

    end%对所有字符变量进行比较
    if flag1>length(data)
        break;
    end

end%结束 while

end%结束所有的包
%% 计算 Y 矩阵
global company
company = {'合容';'日新';'永锦';'西电';'赛晶';'思源';'新东北';'泰开';'上虞';'恒顺';'ABB';
'顺容';'迪生';'中原';'桂容'};
B = zeros( packagesize,length(company) );
%filepath1 = 'C:\Users\Desktop\数模竞赛试题数据处理\ProblemA\附件 2-价格具体得分\国网 2014 年第一批得分情况\';%定义字符串
%filepath2 = 'C:\Users\Desktop\数模竞赛试题数据处理\ProblemA\附件 2-价格具体得分\国网 2014 年第二批得分情况\';%定义字符串
%filepath3 = 'C:\Users\Desktop\数模竞赛试题数据处理\ProblemA\附件 2-价格具体得分\国网 2014 年第三批得分情况\';%定义字符串
for j=1:length(company) %j 为列

    comStr=company(j); %公司的字符串

    for file=1:length(packagemark)
        [score,strName]=xlsread( packagemark{file} );

        for filecontent=1:length(strName)
            if strcmp( comStr,strName(filecontent,1) )
                B( file,j )=score( filecontent-1,1 );
                break;
            end
        end

    end%顺序读取文件内容，寻特定公司的报价数值
end

end%结束公司字符串循环
X2=pinv(A)*B;
save('secondseason')

```

附录 3：区间平均下浮双边曲线算法寻优

```
function [ m_Score ] = AverageInterval ( x )
%区间平均下浮双边曲线算法寻优
%输入其他人的报价，寻找自己的最合适的报价让分数最高
% global Datatest
% x = [Datatest(:,1);m_P];

a = 0.03;
m = 0.6;
n = 1.5;
count = size(x);
A = nanmean(x);
Aup = A*1.15;
Alow = A*0.8;
Peff = [];
j = 0;

for i=1:1:count(1)
    if isnan(x(i,1))==0
        if x(i,1)>Alow && x(i,1)<Aup
            j = j+1;
            Peff(j) = x(i,1);
        end
    end
end

if j==0
    B = A*(1-a);
    A1 = 0;
else
    A1 = mean(Peff);
    B = A1*(1-a);
end

Score=[];
for i=1:1:count(1)
    if isnan(x(i,1))==0
        if x(i,1)>=B
            Score(i) = ((B/x(i,1))^n)*60;
        else
            Score(i) = ((x(i,1)/B)^m)*60;
        end
    else
        Score(i)=0;
    end
end
m_Score = -Score(end);

end
```