

中国研究生创新实践系列大赛
“华为杯”第十七届中国研究生
数学建模竞赛

学 校

中南大学

参赛队号

20105330033

队员姓名

1. 黄新杰

2. 饶家玮

3. 李明

中国研究生创新实践系列大赛

“华为杯”第十七届中国研究生 数学建模竞赛

题 目 降低汽油精制过程中的辛烷值损失模型

摘 要：

汽油是现代社会重要的石油产品之一，提高汽油中辛烷值，降低其中硫、烯烃含量，对提高汽油产品燃烧性能和降低有害气体排放有着重要的意义。我国 70% 以上的石油通过催化裂化生产，而催化裂化得到的汽油的含硫杂质和烯烃含量较高，因此需要对催化裂化汽油进行精制处理。

针对问题一，原料及产品的辛烷值是重要建模变量，但设备及测量手段原因，辛烷值的含量难以实时测量，其数据采集频次（每周 2 次）远小于操作变量的采集频次（3 分钟/次），因此难以一一对应。根据工艺流程的实际情况，可认为辛烷值的测量值是测量时刻前两小时操作变量的综合效果。本文通过附件二给出的各项规则构建两棵规则判别树（也叫决策树），利用判别树剔除了不符合附件二规则的采样样本和采集位点，最后将辛烷值的测量时刻前 2 小时内操作变量的平均值与辛烷测量值构建了对应关系，并利用此方法对 285 号样本、313 号样本进行了处理，并与厂家给定的数据做了对比，实验结果表明，利用本文所提方法处理原始数据的结果与厂家采用的方法处理结果基本一致。

针对问题二，催化裂化汽油精制过程中，辛烷值损失预测模型涉及到 13 个基础变量，以及生产过程中的 354 个操作变量，而样本仅有 325 个。导致样本在特征空间的分布变得稀疏，甚至在某一维度上根本没有样本，高斯三法则不再适用，无法直接求解。这是一个典型的高维小样本问题，容易发生“维度灾难”问题。对该问题进行建模求解，则首先需要对数据进行降维处理。本文考虑到传统的 PCA 这类特征抽取方法得到的低维特征是从其他高维特征融合而来，难以有一个具体的物理解释与含义，无法对实际工业操作过程提供调整控制策略指导。而降维的第二类方法——特征选择，不仅能选出具有完整物理含义的特征，而且不涉及复杂的数学空间映射变换，实时性更高。所以本文选择了一种混合式特征选择方法来对数据进行特征选择，以实现数据的降维，选取出对辛烷值损失有较大贡献的主要特征（主变量）。本文共选择出 23 个主要变量，并对变量进行了相关性分析，相关性图谱显示选择的主变量之间独立性较高，冗余较小，证明了我们方法的有效性。

针对问题三，辛烷是判断产品汽油优劣的一个重要指标，成品汽油的命名都是由汽油

中辛烷含量来命名的，所以在催化裂化汽油精制过程中，一个很重要的目标便是不能使成品汽油较原料的辛烷损失值太大，最理想的情况是没有损失。世界的本质是非线性的，现实生活中几乎不存在线性模型，催化裂化汽油精制是一个典型的高度非线性过程，且为了满足实际工业过程的实时性要求，模型不能太复杂。本文借鉴了“极限学习机”的思想，采用问题二得到的主要变量和因素，建立了一个以非线性为核心，同时兼顾线性模型易于求解的性质的数学模型，具体步骤为：首先对输入数据 X 做非线性变换得到中间变量 Z ，然后 Z 与输出 Y 之间建立一个线性加权模型，既保证了非线性，又能利用线性模型的性质对问题进行快速求解，提高实时性。预测结果表明，本文提出的模型确能有效的逼近真实值，且时间复杂度低，实时性较高。

针对问题四，问题可理解为以操作变量改变后硫含量值限定在 $5\mu\text{g/g}$ 以内，尽量降低辛烷值（RON）损失 30% 以上为目标的多目标优化问题，其约束为每个操作变量须在其取值范围内变化。首先运用问题三的建模思想对产品硫含量建立预测模型。然后建立多目标优化模型，将问题转化为求解 Pareto 前沿解的问题，之后采取的是带精英策略的非支配排序的遗传算法（NSGA-II）来求解出操作变量的优化策略。最终，以该优化策略回归预测得，在产品硫含量均小于 $5\mu\text{g/g}$ 的情况下，大部分样本对辛烷值（RON）损失有 40%-50% 的降幅，优化效果明显。

针对问题五，实际工业过程需平稳生产，所以每次只能有调整幅度值 Δ 的改变。问题五所求的即可转变为：根据问题四解得的操作变量优化方案，以第 133 号样本为例，在逐步改变操作变量的情况下，对硫含量和辛烷值（RON）损失的预测，并以可视化的形式展现出两者随调整幅度值 Δ 变化而同步变化，最终趋于目标范围的过程。由结果观察得，随着操作变量的逐步变化，硫含量和辛烷值的损失都在减少且最终趋于目标范围，但其减少的速率也随着调整幅度值 Δ 逐步趋于平缓。

目录

1	问题重述.....	1
1.1	问题背景.....	1
1.2	待解决的问题.....	2
1.3	问题分析.....	2
2	基本假设与符号说明.....	5
3	符号说明.....	6
4	数据预处理及主变量提取.....	7
4.1	问题一建模与求解.....	7
4.1.1	问题分析.....	7
4.1.2	问题建模.....	7
4.1.3	模型求解.....	8
4.2	问题二建模与求解.....	8
4.2.1	问题分析.....	8
4.2.2	问题建模.....	9
4.2.3	模型求解流程与结果展示.....	11
5	辛烷值损失预测模型的建立与求解.....	15
5.1	问题分析.....	15
5.2	问题建模.....	16
5.3	模型求解与结果展示.....	17
6	操作变量优化及结果可视化.....	19
6.1	问题四建模与求解.....	19
6.1.1	问题分析.....	19
6.1.2	变量优化模型建立.....	20
6.1.3	模型求解.....	20
6.1.4	结果展示与分析.....	24
6.2	模型的可视化展示.....	26

6.2.1 辛烷值损失和硫含量逐步更新	26
6.2.2 结果展示与分析	27
参考文献.....	28

1 问题重述

1.1 问题背景

汽油是一种重要的石油产品，其良好的燃烧性能以及积碳少、抗爆震的特点使其在内燃机上广泛使用。汽油中的辛烷值是衡量汽油质量最重要的指标，辛烷值越高，汽油的燃烧性能越好，其抗爆震能力也越强，从而提高发动机的压缩比，增加行车的里程数。我国的原油产量较低，对进口原油资源依存度较高，其中大部分原油是来源中东地区的含硫和高硫原油。这类原油中重油含量高，含硫杂质也高，需要通过热裂化、催化裂化等工艺来提高石油收率，得到汽油、柴油以及烯烃等产物。我国在发展和利用重油资源上取得了有效的成果，形成了以催化裂化为核心的重油轻质化工艺。由于我国 70% 以上的石油通过催化裂化生产，而汽油产品中绝大部分的含硫杂质和烯烃都来源于催化裂化汽油。汽油中的含硫杂质会导致含硫气体排放污染，因此需要对汽油进行精制处理，下图展示了一个典型的催化裂化汽油精制过程^{[1][4]}。

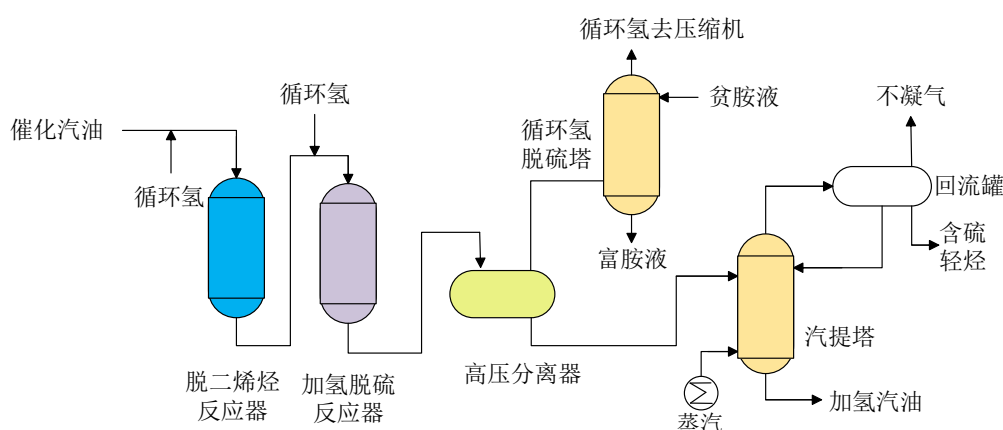


图 1.1 催化裂化汽油精制工艺流程图

烯烃是汽油中高辛烷值组分，在对催化裂化汽油进行加氢脱硫和降烯烃过程中，会使得烯烃加氢饱和，普遍会降低辛烷值，制约了汽油质量提高，造成的经济效益损失巨大。因此，需要对该生产过程中辛烷值损失进行建模分析，在保证脱硫效果的情况下，提高产品辛烷值。目前，使用机理等方法对石油裂解化工过程进行建模已经取得了一定的成果。但由于炼油过程复杂，具有强耦合、非线性等特点，而且不同企业使用的炼油设备存在差异，这一方法存在较大的局限性。依据采集的大量生产数据，通过数据挖掘的方法进行建立辛烷值损失预测模型，可以有效解决这一问题，对不同情况下炼油过程的建模也更具普适性。

1.2 待解决的问题

问题一：数据处理

依据附件二中的样本确定方法，参考附件一中近四年的工业数据，对附件三中的 285 号和 313 号样本原始数据进行处理，并将处理后的数据添加到附件一中，供下一步分析使用。

问题二：寻找建模主要变量

辛烷值损失预测模型涉及 7 个原料性质，2 个待生吸附剂性质，2 个再生吸附剂性质，2 个产品性质，以及生产过程中的 354 个操作变量，共 367 个变量。工程应用上经常使用先降维后建模的方法，先通过数据降维的方法从 367 个变量中筛选出建模主要变量，再建立模型，并说明筛选过程及其合理性。

问题三：建立辛烷值损失预测模型

利用问题二中选出的建模主要变量，利用附件一所提供的样本，通过数据挖掘技术建立辛烷值（RON）损失预测模型，并进行模型验证。

问题四：主要变量操作方案优化

在保证产品硫含量不大于 $5 \mu\text{g/g}$ 的标准下，对 325 个样本数据中对应的主要变量的操作条件进行优化（优化过程中保持原料、待生吸附剂、再生吸附剂的性质保持不变），使得辛烷值损失降幅大于 30%。

问题五：模型可视化

使用问题四中的优化操作方案，对 133 号样本进行测试，以图形展示其主要操作变量优化调整过程中对应汽油辛烷值和含硫量的变化轨迹，其中主要操作变量的调整规则依据附件四。

1.3 问题分析

本节分别对五个问题进行了完整充分的分析、阐述了基本的解决思路。是对本文内容的简要概括，下面介绍对每个问题的分析与解决思路。

1.3.1 数据预处理

辛烷是判断产品汽油优劣的一个重要指标^[4]，成品汽油的命名由汽油中辛烷含量来命名（#93、#95、#89），故原料及产品的辛烷值是重要建模变量，但设备及测量手段原因，辛烷值的含量难以实时测量，其数据采集频次为每周 2 次，远小于操作变量的采集频次（3 分钟/次），因此难以形成对应。根据工艺流程的实际情况，可认为辛烷值的测量值是测量时刻前两小时操作变量的综合效果。

附件三中提供的 285 号和 313 号样本原始数据，该原始数据来源于中石化实时数据库，其中原始数据中包含多套装置的数据，每套装置均有部分点位存在问题，部分变量只有部

分时间段数据，部分数据存在空值或全为空值。无法直接用于将每次辛烷值的测量数据与其前两小时内个操作变量的多个测量数据对应，需要先对数据进行简单的处理，本文通过附件二给出的各项规则构建两颗规则判别树（也叫决策树），利用判别树剔除了不符合附件二规则的采样样本，最后将辛烷值的测量时刻前 2 小时内操作变量的平均值与辛烷测量值构建了对应关系，并利用此方法对 285 号样本、313 号样本进行了处理。

1.3.2 寻找主要变量

催化裂化汽油精制过程中，我们的目标是降低汽油中的硫、烯烃含量，同时尽量保持其辛烷值，辛烷值损失预测模型涉及到 7 个原料性质，2 个待生吸附剂性质，2 个再生吸附剂性质，2 个产品性质等 13 个基础变量，以及生产过程中的 354 个操作变量，且样本只有 325 个。这是一个典型的高维小样本问题，容易发生“维度灾难”问题，所以对原始数据进行数据降维是必要的。在数据挖掘中，样本的变量又被称为样本的特征，本文选择了一种混合式特征选择方法来进行特征选择，以实现数据的降维，具体步骤如下：

（1）去除异常的特征。如缺失值较多无法补充的特征、操作量纲不统一的特征。

（2）移除低方差的特征。如取值单一的特征，或虽然取值不是唯一值，但取值整体变化幅度较小的特征。

（3）利用混合式特征选择方法对剩下的变量进行特征选择，该方法分为两个阶段：a、Filter 阶段，该阶段利用经典的 $relif_F$ 方法在回归问题上的扩展算法 $rrelif_F$ 方法计算出每个特征的重要程度，并剔除一些重要程度较低的特征。b、Wrapper 阶段，将特征选择问题建模成一个 0/1 整数规划问题，利用优化算法进行求解，得到一个较好的特征组合。

1.3.3 建立辛烷值损失预测模型

辛烷是判断产品汽油优劣的一个重要指标，甚至成品汽油的命名都是由汽油中辛烷含量来命名的，所以在对催化裂化汽油精制过程中，一个很重要的目标便是不能使得成品汽油较原料的辛烷损失值太大，最理想的情况是没有损失。

现实生活中基本不在线性模型，催化裂化汽油精制是一个典型的高度非线性过程，且为了满足实际工业过程的实时性要求，模型不能太复杂。本文借鉴了“极限学习机”的思想，采用问题二得到的主要变量和因素，建立了一个以非线性为核心，同时兼顾线性模型易于求解的性质的数学模型，具体步骤为：首先对输入数据 X 做非线性变换得到中间变量 Z ，然后 Z 与输出 Y 之间建立一个线性加权模型，既保证了非线性，又能利用线性模型的性质对问题进行快速求解，提高实时性；该方法的核心在于如何寻找一个好的非线性映射，具体方法实现将在“问题三建模与求解”一章进行详细阐述。

1.3.4 主要变量操作方案的优化

基于问题三所建立的辛烷值损失预测模型，在保证产品硫含量不大于 $5\mu\text{g/g}$ 的标准下，对 325 个样本数据中对应的主要变量的操作条件进行优化（优化过程中保持原料、待生吸附剂、再生吸附剂的性质保持不变），使辛烷值损失降幅大于 30%。

因为产品硫含量不大于 $5\mu\text{g/g}$ 的约束，首先需对产品硫含量建立预测模型。操作变量的优化方案的制定需限定在其取值范围内，每次主变量改变（增加或减少）所预测得到的辛烷值损失应较上一次的预测值减少 30%，这方才达到优化的标准。

1.3.5 模型可视化展示

与问题四的优化操作不同，对于汽油加工生产而言，其工业装置需平稳生产，所以对于操作变量的优化不能突然产生剧烈变化，每次只能有调整幅度值 Δ 的改变。所以对于 133 号样本而言，每调整一个 Δ 都将对应有辛烷值和含硫量的变化，以图形的形式同步展现出两者的变化轨迹，直到最终操作变量不再改变，对应的硫含量应不大于 $5\mu\text{g/g}$ ，而对应的辛烷值损失较未优化前降幅在 30% 以上。

2 基本假设与符号说明

在建模的过程中，在不影响模型可靠性和有效性的前提下，为了简化模型，降低模型建立的复杂度，也便于后续的优化操作变量，做出以下假设：

- （1）假设附件中提供的样本数据是在正常工况下采集，真实有效的数据。
- （2）假设样本数据总体分布相同。
- （3）假设不考虑除附件一所提供的变量外的影响因素对生产过程影响较小或无影响。
- （4）假设人为操作因素对生产过程不产生影响。

3 符号说明

符号	意义
x_{ij}	第 i 个样本在第 j 个特征上的取值
D	样本集合
p_k	样本集合中第 k 类样本的比例
$NH(j)$	与样本 s_j 相同类别样本中，与样本 s_j 距离最近的样本，其大小为 m_j
$NM(j, y)$	与样本 s_j 不同样本中，与样本 s_j 距离最近的样本，其大小由 h_{jy} 表示
X	输入
Z	X 做非线性变换得到中间变量
$g(\square)$	非线性映射
W	线性加权的权重
Y	输出
$f(\square)$	线性映射
f_L	隐含层输出函数
b	偏置
\hat{x}_i	更新后的操作变量
α	操作变量优化系数矩阵
A_{\max}	各操作变量的最大取值
A_{\min}	各操作变量的最小取值
y_{loss}	每步的辛烷值损失值结果
S	每步操作后的硫含量结果
Δx	每个操作变量对应的调整幅度值
K	各操作变量达到优化方案的步数
$iter$	优化操作变量改变的当前步数

4 数据预处理及主变量提取

4.1 问题一建模与求解

4.1.1 问题分析

附件一中的数据绝大部分已经处理好了，但 285 号样本和 313 号样本需要手动处理，由于测量手段与设备原因，产品辛烷值的测量频率（每周 2 次）远低于操作变量的采集测量频次（3 分钟/次），因此难以一一对应，根据过去现场工人师傅们的经验和实际工业情况，可认为辛烷值的测量值是测量时刻前两小时的所有操作变量变化构成的综合效果。所以需要每次辛烷值的测量数据与其前两小时内个操作变量的多个测量数据进行对应，数据的处理方法以附件二给出。

4.1.2 问题建模

由附表二可知，对原始数据处理标准有：

- （1）对于只含有部分时间点的位点，如果其残缺数据较多，无法补充，将此类位点删除；
- （2）删除 325 个样本中数据全部为空值的位点；
- （3）对于部分数据为空值的位点，空值处用其前后两个小时数据的平均值代替；
- （4）根据工艺要求与操作经验，总结出原始数据变量的操作范围，然后采用最大最小的限幅方法剔除一部分不在此范围的样本；
- （5）根据拉依达准则（ 3σ 准则）去除异常值。

显然，由这几条规则可以分为对特征的处理（1、2、3）和对原始数据中样本的处理（4、5），从而形成两个规则表，从而快速的建立两颗判别树^[16]（见图 4.1），判别树的节点即为规则中的统计量。

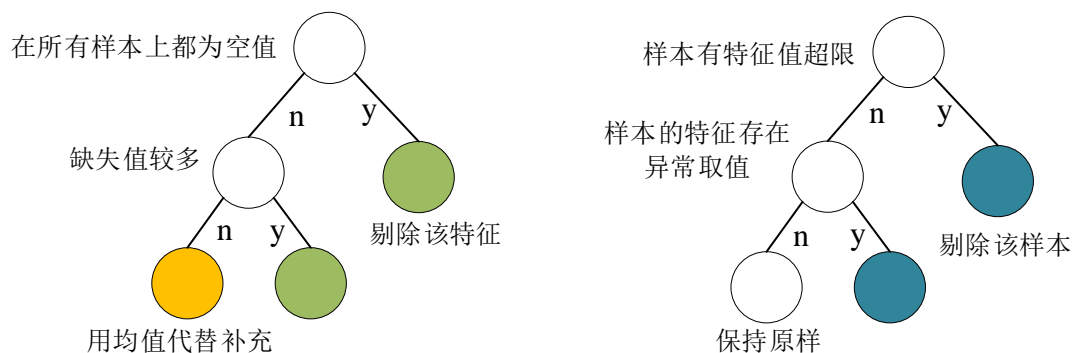


图 4.1 特征处理判别树（左）和样本处理判别树（右）

4.1.3 模型求解

对附件三中 285 号样本和 313 号样本的原始数据进行处理，剔除不合规则样本与特征后，计算辛烷值测量时刻前 2 小时内各操作变量的平均值，并与辛烷测量值建立对应关系。本文按上述方法处理后的数据与厂家给定数据的差异曲线图如图 4.2 所示。

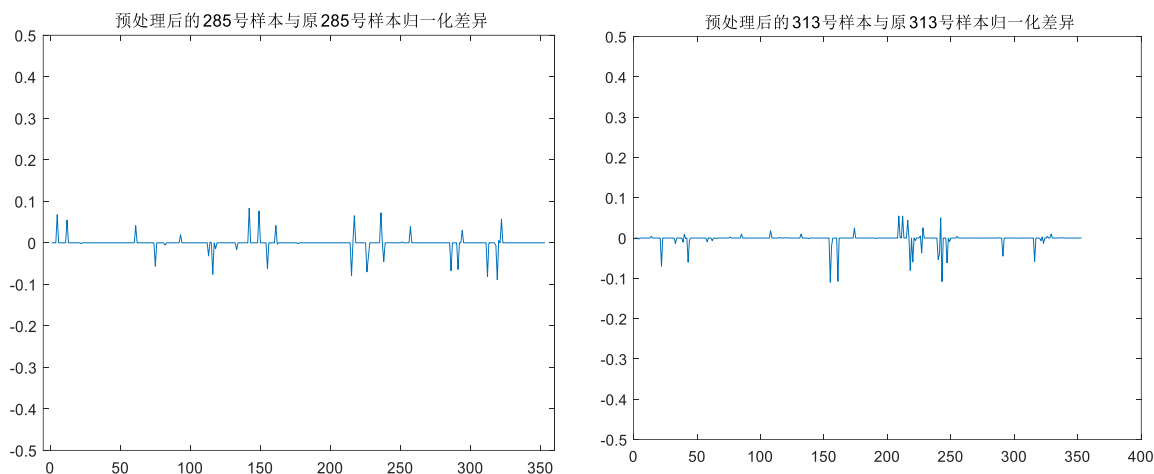


图 4.2 285 号样本（左）和 313 号样本（右）处理结果示意图

4.2 问题二建模与求解

4.2.1 问题分析

催化裂化汽油精制过程中，由于样本数量少，特征（变量）数量多，导致样本在特征空间的分布变得稀疏，甚至在某一维度上根本没有样本，高斯三法则不再适用，无法直接求解。这是一个典型的高维小样本问题，若要对问题进行建模求解，首先需要对数据进行降维处理。

目前常用的降维方法有两大类。一类是特征提取，该方法基于空间变换的方式，通过数学方法将数据从高维映射到低维实现降维，比如主成分分析（PCA）方法，然而，令人遗憾的是，这类方法得到的低维特征是由其他特征融合而来，难以有一个具体的物理解释与含义，也就无法对实际工业操作过程提供指导。

另一类便是特征选择，与特征提取方法相较而言，特征选择不改变原始特征空间，通过从原始空间中剔除无关和冗余的特征，选择出与数据挖掘任务高相关的特征子集，使学习算法在约简后的数据上得到更精确、更易理解的模型，而且特征选择的计算复杂度相对特征抽取较低，实时性更好，能更及时的对工业过程中的各种操作提供理论指导。故本文采用了一种混合式特征选择方法来进行数据降维^{[13]-[15]}。

我们首先依据题目附件 2 中给出的标准剔除了不合规则的特征，然后利用低方差滤波去除了变化极小的特征（我们认为变化波动小的操作变量对最终结果的影响较小），接着我们使用了一种混合式特征选择方法对剩下的特征进行特征选择。

4.2.2 问题建模

(1) 数据归一化

在众多的操纵变量和固有变量中，变量的取值范围不同，量纲也不同，这将导致变量各种统计量之间难以直接比较。比如一个取值在 $[0, 5]$ 之间的变量 **a** 和另一个取值范围为 $[147000, 153000]$ 变量 **b** 在统计量上会出现极大反差，如下图所示，我们可以看到属性 **a** 在其取值范围内波动较大，属性 **b** 的波动较小，但属性 **b** 的方差会远大于属性 **a**，所以我们在进行统计量比较之前，必需对数据进行归一化处理。

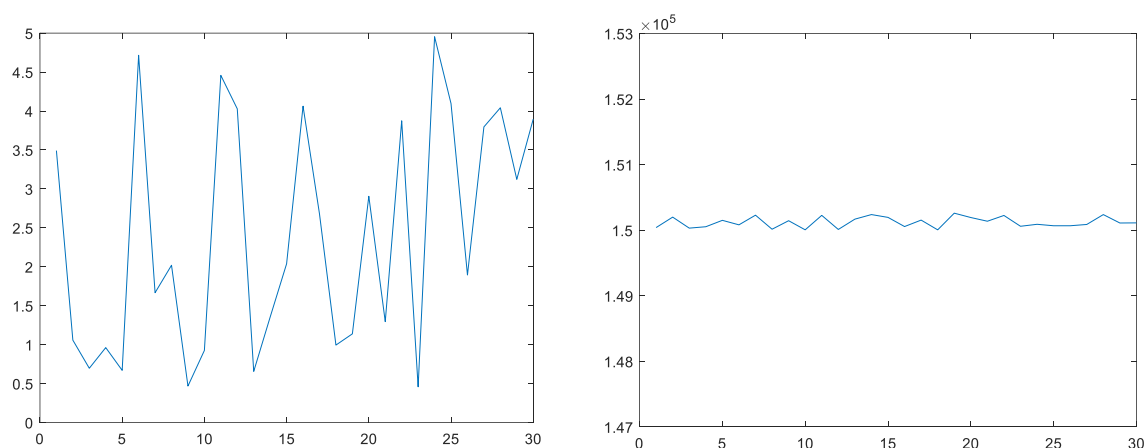


图 4.3 属性 **a** 的波动范围（左）和属性 **b** 的波动范围（右）

归一化公式如下：

$$x'_{ij} = \frac{x_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \quad (4-1)$$

其中 x_{ij} 是第 i 个样本在第 j 个特征上的取值， X_j 是第 j 个特征上的所有取值构成的向量， x'_{ij} 是归一化后的数据。

除了能消除量纲差异，让每个维度的统计量之间可以直接进行比较之外，数据归一化还能带来如下的其他优势：

- 消除特征之间单位和尺度差异的影响，可以同等看待每一维特征；
- 加快算法收敛速度；
- 提高精度。

(2) 低方差滤波

在信息论中，一个特征携带的信息的多少可以由信息熵大小来量定，信息熵即为不确定性的度量，信息熵越大，则代表该特征的不确定性越大，带给我们的信息也就越多，对于取值为离散的属性而言，信息熵的数学描述如下：

$$\text{Ent}(\mathbf{D}) = -\sum_{k=1}^{|\mathbf{y}|} p_k \log_2 p_k \quad (4-2)$$

其中 \mathbf{D} 是样本集合， p_k 是样本集合中第 k 类样本的比例。

在连续域中，我们利用方差来近似信息熵的作用，方差的值越大，代表该特征在取值上波动越大，携带信息也就越多，对我们的数据挖掘任务就越重要。方差值越小代表该特征的取值变化范围小，携带信息少，可以予以忽略。因此我们利用归一化后数据的方差信息进行了特征的初步筛选。

(3) 特征选择与 0/1 整数规划

为了方便地描述特征选择问题，我们使用一个二值向量描述候选的特征子集，将特征选择问题建模成一个 0/1 整数规划问题（也叫布尔型整数规划），其具体描述如下：

定义 1： $X = \{x_1, x_2, \dots, x_n\}$ 表示一个候选特征子集， $x_j = \{0, 1\}, j = 1, 2, \dots, n$ ，其中 x_j 表示第 j 个特征的状态， $x_j = 1$ 表示对应特征被选中， $x_j = 0$ 表示对应特征被剔除。

假设某个二值向量 X 的每个分量的取值如下所示：

1	0	0	0	1	0	1	1	0	1
---	---	---	---	---	---	---	---	---	---

其对应物理含义为： X 是一个 10 维的二值向量 $[1, 0, 0, 0, 1, 0, 1, 1, 0, 1]$ ，对应的特征选择结果为第 1 个、第 5 个、第 7 个、第 8 个、第 10 个特征被选中，其它的特征被剔除。

特征选择的核心在于如何找到预测精度较高的特征组合，目标函数为预测精度最高（即均方误差损失最小）。在定义 1 的条件下，我们就可以将特征选择问题重新用数学语言描述为如下的 0/1 整数规划问题：

$$\begin{aligned} \min \quad & \text{loss}(X) \\ \text{s.t.} \quad & X = \{x_1, x_2, \dots, x_n\}, x_j \in \{0, 1\}, i = 1, 2, \dots, M \\ & 1 \leq \|X\|_0 \leq M \end{aligned} \quad (4-3)$$

其中， X 表示一个选取的特征子集， $\text{loss}(X)$ 表示 X 对应的均方误差，而 $\|X\|_0$ 是 X 的 0 范数，表示 X 中选中特征的数目。

尽管我们在前一步利用方差信息滤去了一部分特征，但剩余特征仍有较高维度，特征与特征之间也存在冗余关系，直接利用优化算法求解式 (4-3) 极其耗时且容易陷入局部最优。因此本文采用了一种混合式特征选择方法^[13]进行特征选择，该方法分为两个部分：Filter 阶段与 Wrapper 阶段。

Filter 阶段利用 Relif 方法在回归问题上的扩展算法 RRelif_F 计算了每个特征的重要程度，我们给定了一个阈值滤掉了重要程度较低的特征，降低后续特征空间的维度。RRelif_F 算法计算特征权重的原理如下：

对于一个给定的包含 N_d 个样本、 M 维特征的数据集 $D_s \in \mathbb{R}^{N_d \times M}$ ，采用 ReliefF 法对其中的特征求取权重，首先需要从全部的 N_d 个样本中随机选出 l 个样本，而对任一特征 f_i 求取权重的公式如下：

$$R_F(f_i) = \frac{1}{c} \sum_{j=1}^l \left(-\frac{1}{m_j} \sum_{s_r \in NH(j)} M(D_s(i, j) - D_s(r, i)) \right. \\ \left. + \sum_{y \neq y_j} \frac{1}{h_{jy}} \frac{p(y)}{1 - p(y)} \sum_{s_r \in NM(j, y)} M(D_s(i, j) - D_s(r, i)) \right) \quad (4-4)$$

其中， $NH(j)$ 和 $NM(j, y)$ 分别表示在与样本 s_j 相同类别和与样本 s_j 不同样本中，与样本 s_j 距离最近的样本，其大小分别由 m_j 和 h_{jy} 表示。

Wrapper 阶段通过智能优化算法对式 (4-1) 进行求解，得到最佳特征子集。具体为对 Filter 阶段得到的特征子集进行更进一步的精确搜索，主要包括初始化、候选特征子集生成和子集的评价及选择三部分，具体步骤如下所示：

- Step 1：根据 Filter 阶段留下特征的权重进行初始化，生成一个较优的并且契合特征选择问题的初始解。并设置算法的基本参数。
- Step 2：利用状态转移算子，由当前最优特征子集产生多个候选特征子集，并分别计算其均方误差，选出其中最优的特征子集，并和历史最优解比较，更新当前的最优解；
- Step 3：重复 Step 2 直到迭代次数大于等于最大迭代次数，输出此时的当前最优解，该结果即为本章所提方法最终特征选择的结果，为一个多维的二值向量。

4.2.3 模型求解流程与结果展示

选取主要变量问题的求解流程如图 4.4 所示，其中特征选择阶段的具体算法流程图如图 4.5 所示。

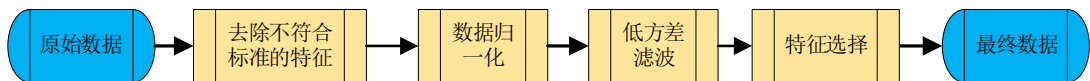


图 4.4 主要变量选取流程图

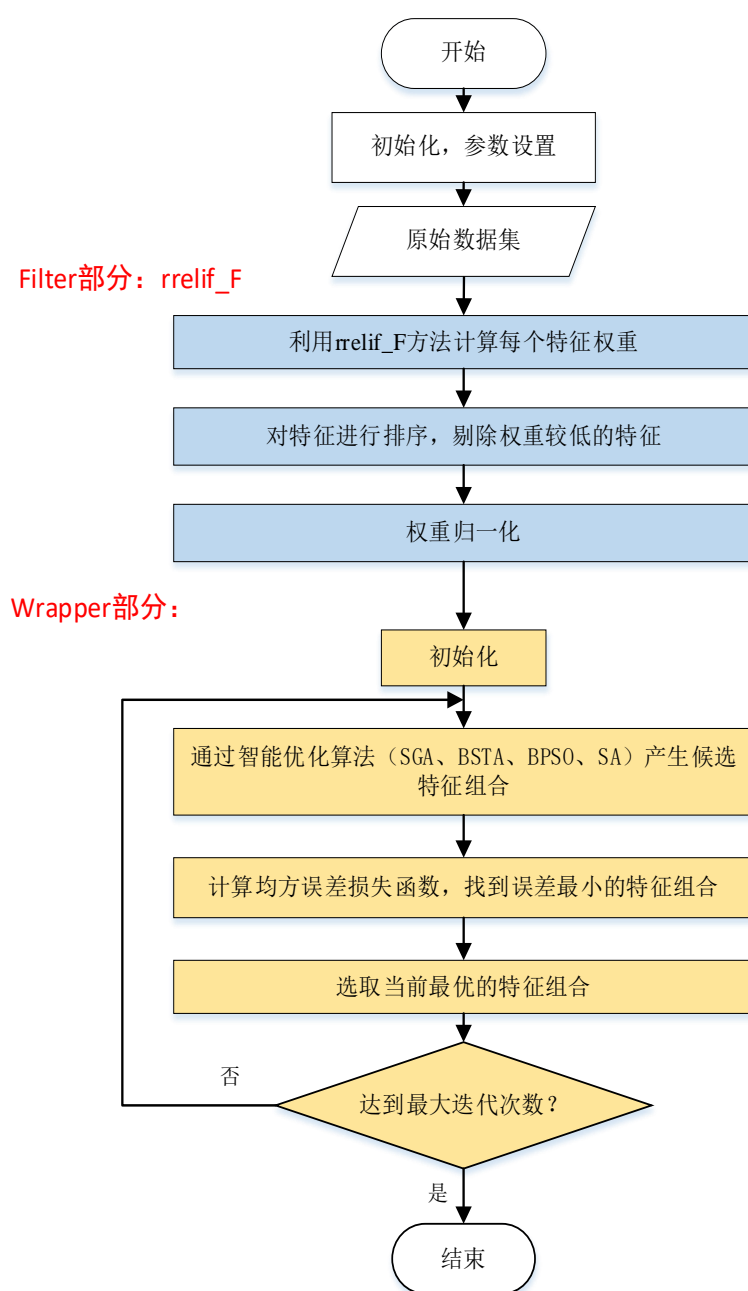


图 4.5 混合式特征选择算法流程图

为得到最佳的特征组合，我们分别用了 4 种在求解 0/1 整数规划问题上表现极佳智能优化算法 SGA、BSTA、BPSO、SA 分别对式 4-3 进行求解。四种算法分别运行 30 次，选择均方误差最小的的一次实验结果作为该算法求得的最优解，结果下表

表 4.1 四种算法求解式 3-3 的结果对比

指标	特征数目	均方误差	平均求解时间 (s)
SA	31	0.1116	1.653
SGA	27	0.1049	2.470
BPSO	23	0.0427	1.358
BSTA	23	0.0427	1.126

最终我们通过优化算法得到了 23 个主要特征（主变量），23 个特征的具体物理含义和取值范围见表 4.2，23 个主变量之间的相关关系见图 4.6。

表 4.2 最终选取的主变量

序号	变量名	变量取值范围	变量单位	Δ
1	辛烷值 RON（原料）	/	/	/
2	烯烃（原料）	/	v%	/
3	焦炭（待生吸附剂性质）	/	wt%	/
4	硫含量（待生吸附剂性质）	/	wt%	/
5	焦炭（再生吸附剂性质）	/	wt%	/
6	硫含量（再生吸附剂性质）	/	wt%	/
7	催化汽油进装置总流量	40-145	t/h	5
8	D-109 吸附剂料位	-0.05-（70）	/	5
9	闭锁料斗 N2 过滤器出口气 流量	0-60	/	5
10	R-102 底部锥段温度	100-500	°C	5
11	E-206 壳程出口管温度	50-150	°C	1
12	P-101A 入口过滤器差压	-0.5-12	/	1
13	E-106 管程出口管温度	2-50	°C	1
14	F-101 循环氢出口管温度	320-480	/	2
15	D-125 液位	-0.85-(-2.00)	mm	0.2
16	反应器质量空速	2.95-7.00	h-1	0.5
17	再生器温度	450-520	°C	1
18	还原器温度	200-350	°C	1
19	精制汽油出装置温度	30-45	°C	1
20	氢油比	0.20-0.37	/	0.01
21	反应器上部温度	410-435	°C	1
22	反应器底部温度	399-430	°C	1
23	反应器顶部压力	2.25-2.55	MPa	0.1

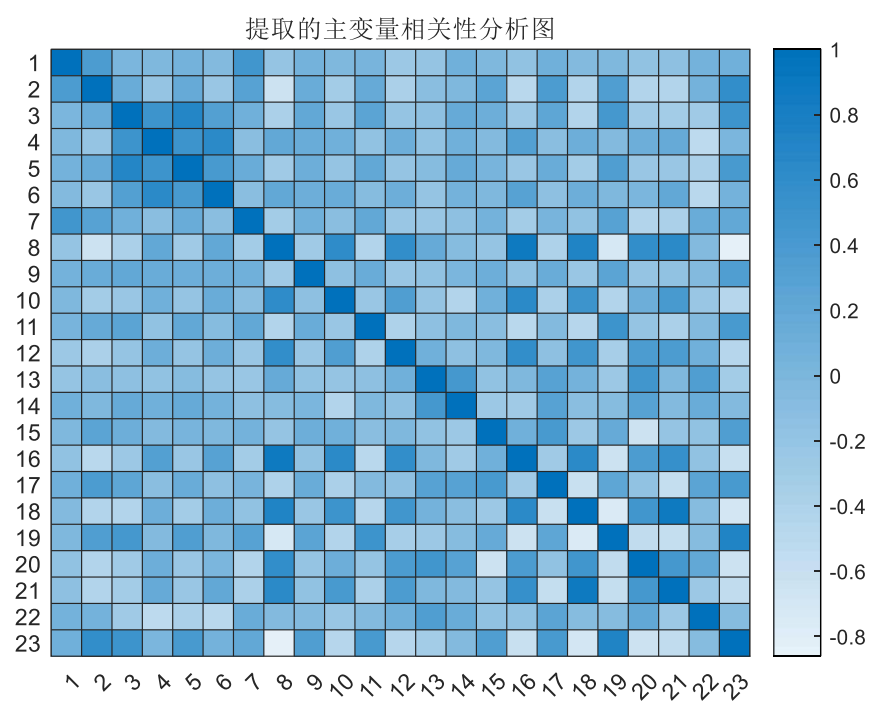


图 4.6 23 个主变量相关性图谱

5 辛烷值损失预测模型的建立与求解

虽然线性模型由于可解释性强、易于求解等在数学上的应用十分普遍，但现实生活中往往不存在真正的线性模型，可以说世界的本质是非线性的，在实际工业生产中基本不存在线性模型，管理者也不会相信一个线性模型的预测结果。催化裂化汽油精制过程就是一个典型的非线性工业过程，且车间内环境复杂，信号的检测与传输会有极大的延迟，所以对实时性的要求比较高，即模型不能太复杂。

5.1 问题分析

辛烷值损失直接关系到最终产品的优劣程度，辛烷值的损失与操作工艺密切相关，受到诸多操作步骤的影响。催化裂化汽油精制工业过程是一个高度非线性过程，直接使用线性模型显然是毫无依据的，虽然可以用一个极其复杂的深层神经网络去拟合一个复杂的复合函数去代替真实模型，但其中的参数以千万计，而实际工业过程中很多控制器都是基于单片机、PLC 来实现逻辑控制的；模型如果过于复杂、参数太多，会导致调整策略的滞后，对工厂造成极大损失。

本文希望在考虑催化裂化汽油精制过程中非线性的同时，利用线性模型的易于求解的性质来减少时间复杂度，提高控制策略的实时性。本文在借鉴了“极限学习机”的思想的基础上提出一个新的模型：首先对输入 X 做非线性变换得到中间变量 Z ，然后 Z 与输出 Y 之间建立一个线性加权模型，既保证了非线性，又能利用线性模型的性质对问题进行快速求解，提高实时性；其数学描述如下

$$\begin{aligned} Z &= g(X) & g(\cdot) &\text{是一个非线性映射} \\ Y &= WZ + B = f(Z) & f(\cdot) &\text{是一个线性映射} \end{aligned} \quad (5-1)$$

模型的核心思想在于把所有的非线性包含在 $g(\cdot)$ 里面，非线性映射的输出 Z 最后的输出 Y 之间是线性关系，便于求解。

极限学习机 (Extreme Learning Machine, ELM) [7]-[9] 在 2004 年由南洋理工大学的 Guang-Bin Huang、Qin-Yu Zhu 和 Chee-Kheong Siew 提出，并发表于当年的 IEEE 国际交互会议 (IEEE International Joint Conference) 中，目的是为了对反向传播算法 (Backward Propagation, BP) 进行改进以提升学习效率并简化学习参数的设定，传统的 ELM 具有单隐含层，其模型结构如下：

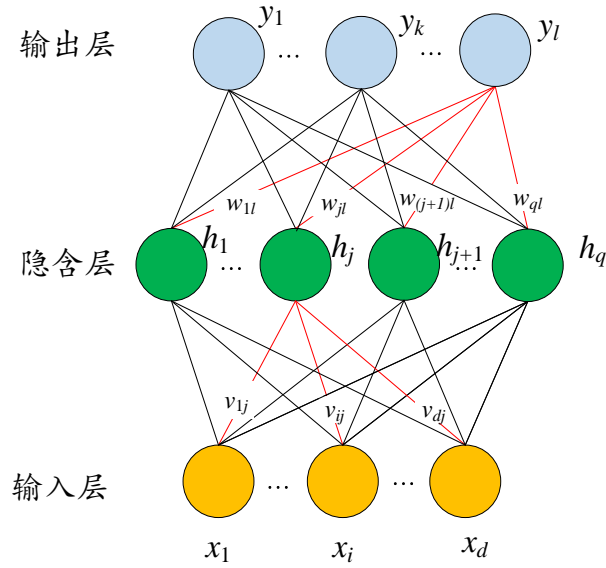


图 5.1 标准的 ELM 网络结构示意图

图 5.1 显示了标准的 ELM 使用单层前馈神经网络（Single Layer Feedforward neuron Network, SLFN）的结构。具体地，SLFN 的组成包括输入层、隐含层和输出层，其中隐含层的输出函数具有如下定义：

$$f_L = \sum_{i=1}^l w_i g_i(\mathbf{x}) = \mathbf{g}(\mathbf{x})\mathbf{W} \quad (5-2)$$

\mathbf{x} 是神经网络的输入， \mathbf{W} 是输出权重， $\mathbf{g}(\mathbf{x})$ 被称为特征映射，其作用是将输入层的数据由其原本的空间映射到 ELM 的特征空间：

$$\mathbf{g}(\mathbf{x}) = G(\mathbf{v}, \mathbf{x}) = H_{l \times q} \quad (5-3)$$

其中 \mathbf{v} 是特征映射的参数。

ELM 中输入层至隐含层的特征映射是随机的或人为给定的且不进行调整，因此 ELM 的特征映射是随机的。依据通用近似定理，ELM 无限趋近于任意连续的目标函数，特征映射可以是任意非线性的片段连续函数如 RBF 函数、sigmoid 函数、LU 函数等，ELM 算法的核心是求解输出权重使得误差函数最小。此外 ELM 具有较强的泛化能力，其原因是算法中对特征映射参数的随机初始化增强了各输入特征的相互独立性，创造了一个更大的求解空间，从而有利于找到正确的目标函数进行学习。

5.2 问题建模

依据上述分析，本文针对辛烷值损失建立了如下预测模型：

$$Y = f(g(X)) = w_1 g(x_1) + w_2 g(x_2) + \cdots + w_n g(x_n) + b \quad (5-4)$$

其中 $g(\cdot)$ 是一个非线性映射， x_i 代表样本第 i 个特征的取值， w_i 代表样本第 i 个特征

的权重， b 是偏置。

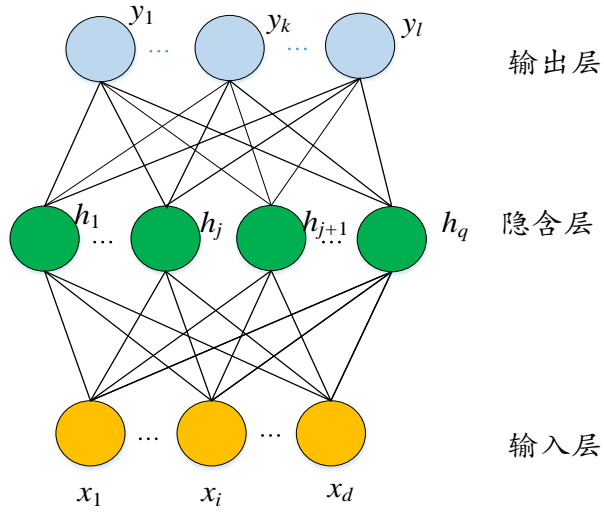


图 5.2 辛烷值损失预测模型结构

结构如上图所示，损失函数（loss function）定义为：

$$\min \|Y - T\|_2^2 \Rightarrow \min \|HW - T\|_2^2 \quad (5-5)$$

其中 H 是隐含层输出， W 是隐含层到输出层的权重矩阵， T 是真实值。为了防止过拟合，引入了 L2 正则化项，修改式 5-5 为：

$$\min \|HW - T\|_2^2 + \frac{C}{2} \|W\|_2^2 \quad (5-6)$$

C 为正则化系数，求解该损失函数等价于岭回归问题，其解有如下表示：

$$W^* = \left(H^T H + \frac{1}{C} \right)^{-1} H^T T \quad (5-7)$$

5.3 模型求解与结果展示

本文将 325 个样本通过随机抽样选取了其中 2/3 作为训练集，剩余 1/3 作为测试集，并以问题二中选取的 23 个主变量作为样本特征，用模型在测试集上的预测精度对模型进行了验证，实验结果如下。

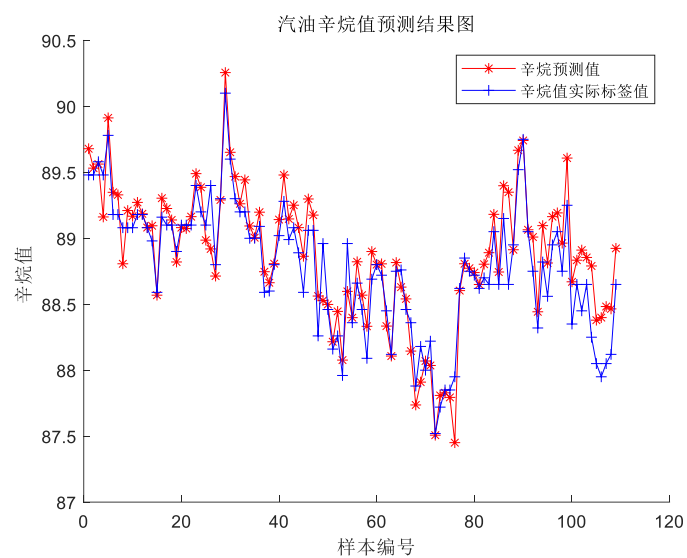


图 5.3 测试集上产品辛烷值预测结果

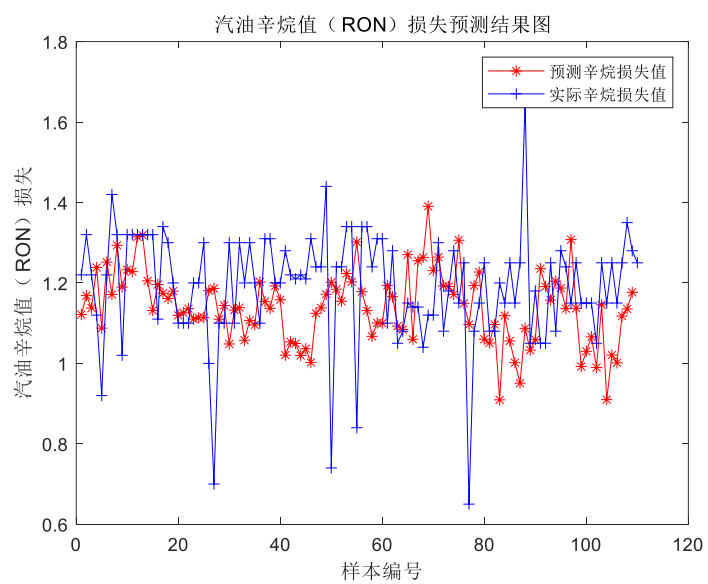


图 5.4 测试集上产品辛烷值损失预测结果

6 操作变量优化及结果可视化

6.1 问题四建模与求解

6.1.1 问题分析

本问题即为以操作变量改变后含硫量限定在 $5\mu\text{g/g}$ 以内，尽量降低辛烷值损失 30% 以上为目标的多目标优化问题，其约束为每个操作变量须在其取值范围内变化。首先运用问题三的预测算法对产品硫含量建立预测模型。然后建立多目标优化模型，将问题转化为求解 Pareto 前沿解的问题。而前沿解求解算法主要有基于数学的规划方法和基于遗传算法的两类方法。而本次采取的是带精英策略的非支配排序的遗传算法（NSGA-II）来求解^{[10][11]}。

约束优化问题就是指在一定约束条件下寻求一个（一组）最合适的数值,使目标值达到最优或者使最优性度量值得到满足。传统的约束优化算法可分为：通过罚函数将约束问题转化为无约束问题求解的拉格朗日乘子法；通过迭代点用线性函数逼近约束问题目标函数的牛顿法等；将无约束问题扩展到约束问题求解的投影梯度法等三类。而近代的模拟退火算法、遗传算法、蚁群算法、粒子群算法等智能算法的提出以其各自的优势也解决了优化问题。

多目标优化是指对两个或两个以上相互冲突的指标进行优化，最终获得多个指标权衡比较下的最优解集。求解此类问题十分困难，基于智能优化的多目标优化算法致力于解决这一难题，其实用性很强，因而得到了广泛的研究。下图为多目标优化问题求解算法的介绍。

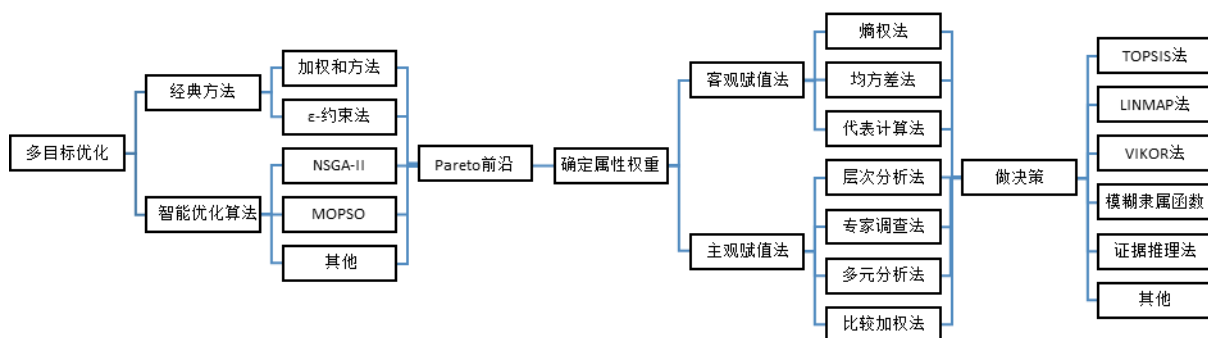


图 6.1 多目标优化问题求解算法介绍

而在实际生产加工过程中，各子指标的优化可能会存在一定的冲突，这便为这类问题的求解带来了一定的难度。此外，对于多目标优化中的涉及到的各种指标而言，很难区分任意两种设计方案的优劣，这是解决多目标优化问题的又一难点。单目标优化设计问题经过计算可以得到唯一的最优解，但是多目标优化问题中能得到的仅为非劣解。而且，通常有多个非劣解，这些非劣解组成 Pareto 最优集^{[12][13][14]}。

6.1.2 变量优化模型建立

首先根据问题要求，确定目标函数为：

$$\begin{aligned} \arg \min_{\alpha} \sum_{i=1}^n (f(g(\hat{x}_i)) - f(g(x_i)))^2 \\ \arg \min_{\alpha} \sum_{i=1}^n (h(g(\hat{x}_i)) - 5) \\ s.t. \ a_{\min}(j) \leq \hat{x}_{ij} \leq a_{\max}(j) \quad p < j \leq q \end{aligned} \quad (6-1)$$

其中， α 为操作变量优化系数矩阵：

$$\alpha = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} & \alpha_{1(p+1)} & \cdots & \alpha_{1q} \\ \vdots & & & \ddots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{np} & \alpha_{n(p+1)} & \cdots & \alpha_{nq} \end{pmatrix}_{n \times q}$$

$$\alpha_i = (\alpha_{i1} \quad \cdots \quad \alpha_{ip} \quad \alpha_{i(p+1)} \quad \cdots \quad \alpha_{iq})_{1 \times q}$$

\hat{x}_i 为更改后的操作变量：

$$\hat{x}_i = x_i + \alpha_i \quad (6-2)$$

对于非操作变量（第 1 到 p 列）不进行调整，所以

$$\alpha_i = (\alpha_{i1} \quad \cdots \quad \alpha_{ip})_{1 \times p} = (0 \quad \cdots \quad 0)_{1 \times p} \quad (6-3)$$

各操作变量的最大取值为： $A_{\max} = (a_{\max}(1) \cdots a_{\max}(p) \cdots a_{\max}(q))_{1 \times q}$ 最小取值为：

$A_{\min} = (a_{\min}(1) \cdots a_{\min}(p) \cdots a_{\min}(q))_{1 \times q}$ ，本问题待求的即为符合约束条件的 α 。

6.1.3 模型求解

(1) Pareto 的概念

在存在多个 Pareto 最优解的情况下，如果没有关于问题的更多的信息，那么很难选择哪个解更可取，因此所有的 Pareto 最优解都可以被认为是同等重要的。由此可知，对于多目标优化问题，最重要的任务是找到尽可能多的关于该优化问题的 Pareto 最优解。因而，在多目标优化中主要完成找到一组尽可能接近 Pareto 最优域的解和找到一组尽可能不同的解两个任务。

Pareto 占优和 Pareto 最优解的概念：考察两个决策向量 $a, b \in X$ 。那么 a Pareto 占优 b 记为 $a > b$ ，当且仅当：

$$\{\forall i \in \{1, 2, \dots, n\} f_i(a) \leq f_i(b)\} \wedge \{\exists j \in \{1, 2, \dots, n\} f_j(a) < f_j(b)\} \quad (6-4)$$

如果在整个参数空间内不存在任何决策向量 Pareto 占优某个决策向量，则称该决策向量即是 Pareto 最优解。所有 Pareto 最优解组成了 Pareto 最优解集合。

下图对 Pareto 占优等相关概念进行更加直观化的阐述：

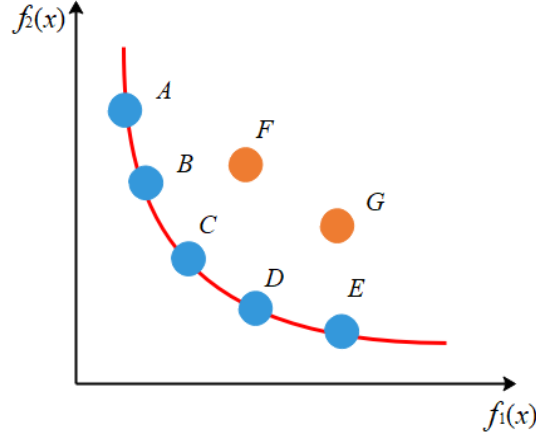


图 6.2 两个目标的 Pareto 最优前沿曲线

标记为 A-G（F 除外）的圆圈代表 6 个候选解，图中表示了这些解在目标函数空间内的位置。以候选解 A 和 B 为例，可以看出候选解 A 的第一个目标函数值小于候选解 B 的，而其另一个目标函数大于候选解 B 的，故 A,B 不能直接比较，A 和 B 是相互不占优的。针对候选解 C 和 G,可以看出候选解 C 的两个目标函数均小于候选解 G 的，满足 Pareto 占优的概念，故候选解 C 是占优候选解 G，候选解 C 为非占优解；若在整个搜索空间内，候选解 A,B,C,D,E 对于其它候选解都占优，则候选解 A,B,C,D,E 称为 Pareto 最优解，他们共同组成了 Pareto 最优解集，其目标函数的集合形成了 Pareto 前沿。

（2）NSGA-II 算法

NSGA-II 是由 Kalyanmoy Deb 等人于 2000 年提出来的一种多目标优化算法，其最主要的特点有^[15]：

- 1.为了降低时间复杂度，提出了一种快速非支配排序方法；
- 2.采用了拥挤度和拥挤度比较算子，使得解在目标空间分布更加均匀；
- 3.引入精英策略，使父代和子代个体共同参与竞争，产生下一代种群，从而获得更好的子代，减少优秀候选解的丢失。

接下来的内容将对以上 3 部分进行详细描述。

（1）快速非支配排序

快速非支配 排序方法可以概括为：对于种群 P ，其中的每一个个体 p 都具有 2 个参数，分别为 N_p 和 S_p 。 N_p 为能支配个体 p 的其他个体的数量，个体 p 能支配的其他所有个体的集合为 S_p 。首先，对种群中所有 $N_p = 0$ 的个体进行搜索，将其放入集合 L_1 中，并将相应的非支配序 p_{rank} 赋予此个体。然后对集合中 L_1 的每个个体 p ，考察它所能支配的集合 S_p 。对于每个 $q \in S_p$ ，将集合 S_p 中每个个体 q 的 N_q 减 1（因为 p 能支配 q ）。如果 $N_q = 0$ ，表

明个体 q 是 S_p 中的非支配个体，则将个体 q 放入另一个集合 Q 中，对 Q 进行分级并赋予它一个新的非支配序；重复上述操作步骤，直至 P 中的所有个体都被分级。快速非支配排序的算法流程如算法 6.1 中的伪代码所示。

算法 6.1 快速非支配排序

```

for each  $p \in P$ 
     $S_p = \emptyset$ 
     $N_p = 0$ 
    for each  $q \in P$ 
        if ( $p < q$ ) then                %如果  $p$  支配  $q$ 
             $S_p = S_p \cup \{q\}$ 
        else if ( $q < p$ ) then
             $N_p = N_p + 1$ 
    if  $N_p = 0$  then                    %如果  $p$  是非支配解
         $p_{rank} = 1$ 
         $L_1 = L_1 \cup \{p\}$ 
    i=1
    while  $L_i \neq \emptyset$ 
         $Q = \emptyset$ 
        for each  $p \in L_i$ 
            for each  $q \in S_p$ 
                 $N_q = N_q - 1$ 
            if  $N_q = 0$  then
                 $q_{rank} = i + 1$ 
                 $Q = Q \cup \{q\}$ 
         $i = i + 1$ 

```

(2) 拥挤度策略

在同一层非支配个体集合中，为了保证每个解都能够均匀地分布在 Pareto 前沿上，NSGA-II 算法采用了拥挤度策略，即计算同一非支配层级中某一特定个体周围存在的其他个体的密度。以两个目标函数的优化问题为例，如图 6.3 所示，它指出了在个体周围包含个体本身但不包含其他个体的最小的长方形。

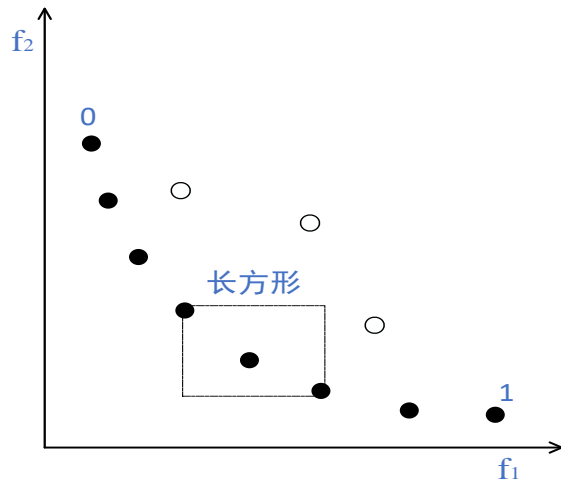


图 6.3 拥挤度策略图示

(3) 精英策略

精英策略被用来防止种群中的优良个体被剔除，其具体操作可以归纳为：当前父代种群 P_t 和其产生的子代种群 Q_t 的规模大小都为 N 。然后合并父代和子代种群，由此得到的新种群 R_t 的中候选解的数量为 $2N$ 。然后利用快速非支配排序法对种群 R_t 中的所有个体进行分层，并计算各支配层上候选解之间的拥挤距离，最后根据个体的优劣程度排序，从种群 R_t 中选择前 N 个个体，使其形成新的父代种群。如图 6.4 所示。

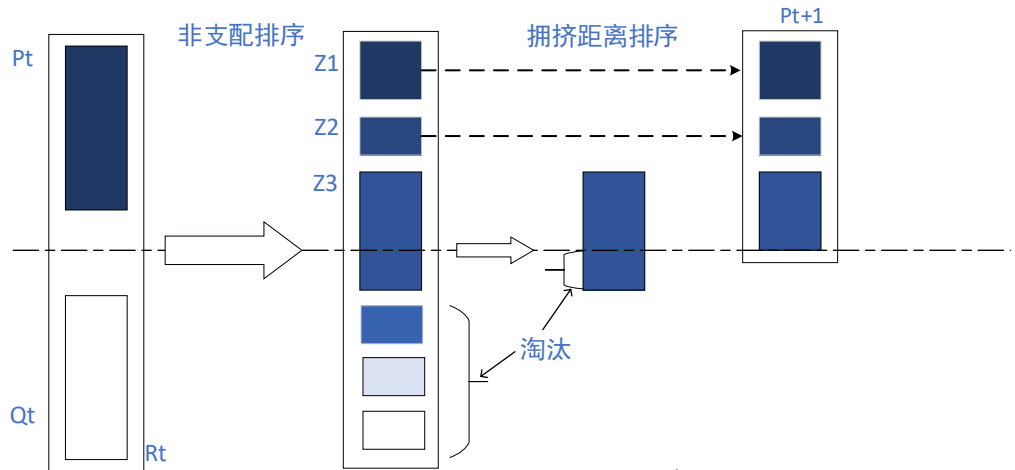


图 6.4 NSGA-II 算法基本步骤图示

综上所述，NSGA-II 算法的基本步骤可以归纳为：

- (1) 在决策空间内，随 N 个候选解组成初始种群；
- (2) 利用快速非支配排序方法对当前种群中的个体进行分层，再利用选择、交叉和变异算子产生新的候选解组成新的子代种群；
- (3) 融合父代种群与子代种群，利用非支配排序方法对种群中的所有候选解进行分层，同时根据拥挤度策略计算出每个非支配层中的个体的拥挤距离，再根据非支配关系以及个体之间的拥挤距离选择合适的个体组成新的父代种群；

(4) 利用遗传算法中的三种基本算子的基本操作产生新的候选解，组成相应的子种群，返回步骤(2)；以此类推，直到满足算法结束的条件。

由于 NSGA-II 算法复杂度较低，且得出的 Pareto 解分布较均匀，因此将 NSGA-II 算法用于本问题的可操作变量的优化方案制定，得到的是关于 α 的解集，依照 Pareto 前沿曲线的性质，选取解集中对两个问题的解相差较小的 α 作为最终优化方案。

6.1.4 结果展示与分析

图 6.5 中优化后的产品辛烷值明显介于原料辛烷值与优化前预测所得辛烷值之间，而且更接近于原料辛烷值，所以优化效果良好。

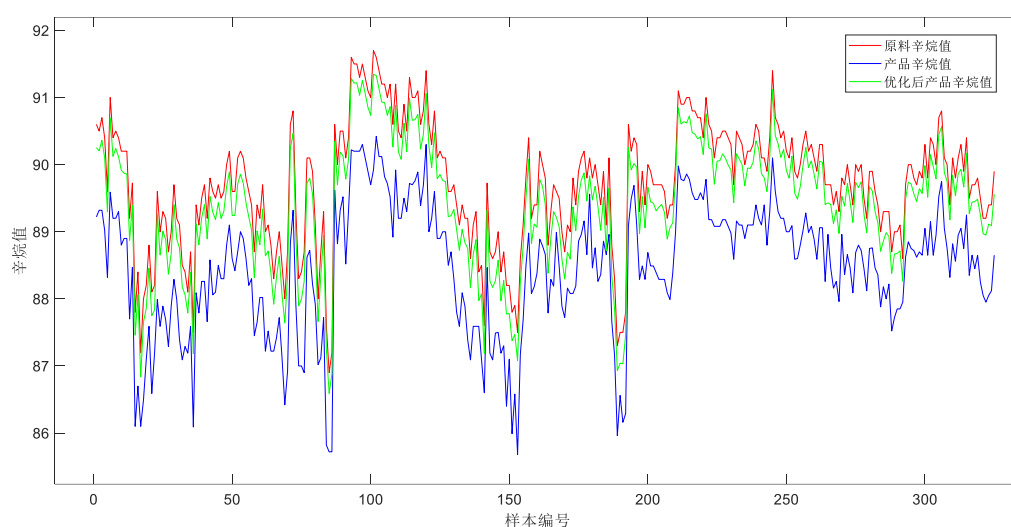


图 6.5 优化前后产品辛烷值与原料辛烷值对比图

图 6.6 中以散点图的形式展现了优化后辛烷值损失降幅，除了极个别样本降幅小于 30% 外，其它的样本所预测得辛烷值损失降幅大多集中在 50%-60% 之间，降幅效果明显，并超额完成降幅目标。

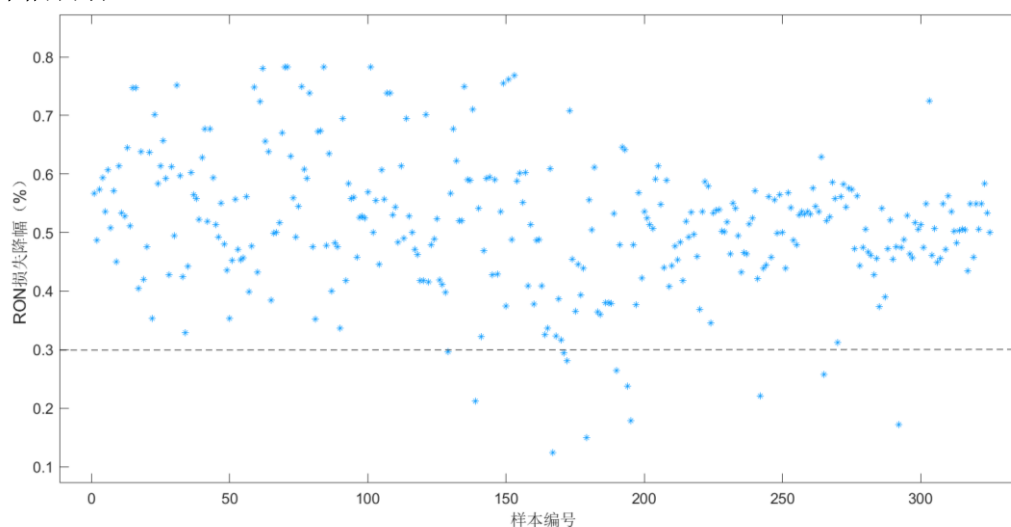


图 6.6 优化后辛烷值损失降幅散点图

表 6.1 为部分样本优化前后产品辛烷值损失对比

表 6.1 优化前后产品辛烷值损失对比

样本编号	原料辛烷值	产品辛烷值	优化后产品辛烷值	辛烷值损失降幅
1	90.6	89.22	89.96	53.80
2	90.5	89.32	89.86	45.38
3	90.7	89.32	90.05	53.08
4	90.4	89.02	89.73	51.43
5	89.6	88.32	88.95	49.23
6	91.0	89.59	90.44	60.04

在对操作变量进行优化 RON 损失值降幅分布直方图如图 6.7 所示，可以看出，大多数样本的辛烷值损失降幅都能够达到 30% 以上，只有部分样本降幅没达到问题要求。经过对该样本预测结果及操作变量的优化结果进行分析，这可能是由于预测模型在该样本上的预测误差较大，导致对操作变量的优化效果不明显。

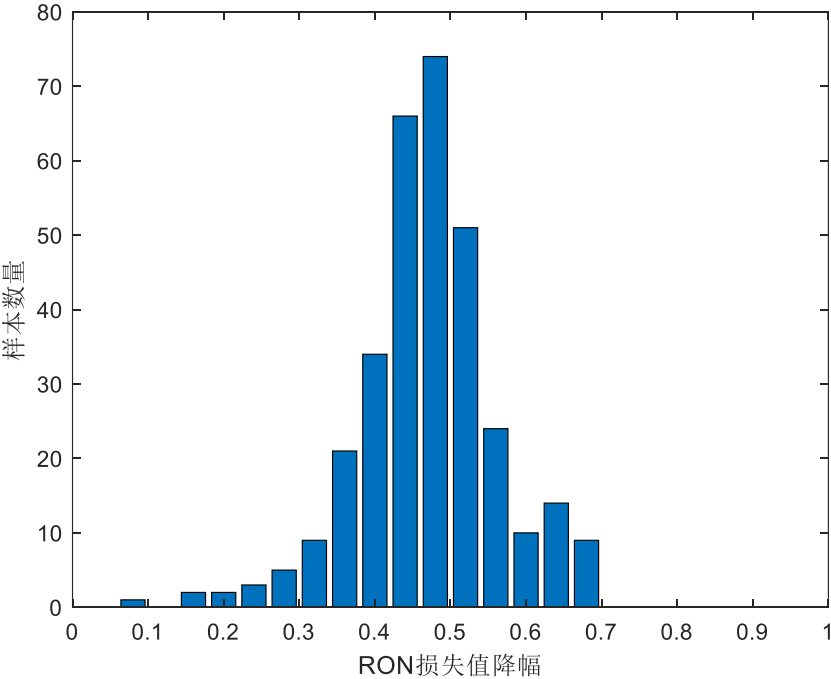


图 6.7 RON 损失降幅直方图

图 6.8 为操作变量优化前后的硫含量对比图，优化后，硫含量严格地控制在 5 $\mu\text{g/g}$ 以内，这也为优化方案实行奠定了基础。对于操作前硫含量接近最小值（3.2 $\mu\text{g/g}$ ）的样本，在变量优化后，其硫含量略微减少；而对于操作前硫含量远大于 5 $\mu\text{g/g}$ 的样本，在优化操作之后，其产品硫含量大幅减少，且均符合要求。

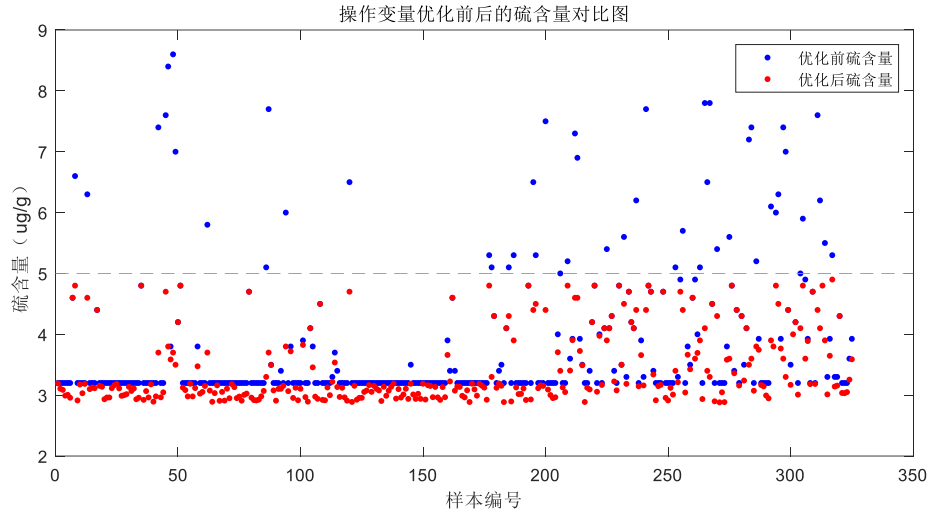


图 6.8 操作变量优化前后硫含量对比图

6.2 模型的可视化展示

在问题四解得各样本的相应操作变量的优化方案之后，选取 133 号样本的操作变量。与问题四的只需得到最终的操作变量优化方案不同，问题五的 133 号样本不能对操作变量的优化一步到位。因为对于汽油加工生产而言，其工业装置需平稳生产，所以对于操作变量的优化不能突然产生剧烈变化，每次只能有调整幅度值 Δ 的改变。所以对于 133 号样本而言，每调整一个 Δ 都将对应有辛烷值和硫含量的变化，以图形的形式同步展现出两者的变化轨迹，直到最终操作变量不再改变，对应的硫含量应不大于 $5\mu\text{g/g}$ ，而对应的辛烷值损失较未优化前降幅在 30% 以上。

6.2.1 辛烷值损失和硫含量逐步更新

其每步的辛烷值损失值可表示为： $y_{loss} = f(g(x_{133}')) - f(g(x_{133}))$ ；每步的硫含量值为： $S = h(g(x_{133}'))$ ；每步更新后的操作变量为： $x_{133}' := x_{133}' + t_{iter} \square \text{diag}(\Delta x)$ ；每个操作变量对应的调整幅度值为： $\Delta x = (\Delta x_1 \cdots \Delta x_p \Delta x_{p+1} \cdots \Delta x_q)_{1 \times q}$ ；各操作变量达到优化方案的步数为 K ： $K = (k_1 \cdots k_p k_{p+1} \cdots k_q)_{1 \times q}$ ；其中第 j 个变量的步数为： $k_j = \frac{\alpha_{133j}}{\Delta x_j} \quad 0 < j \leq q$ 。

若 $\text{iter} > k_j$ ，则其对应的第 j 列操作变量之后将不再变化，即第 t_{iter+1} 可表示为 $t_{iter+1} = (\text{iter}+1 \quad \text{iter}+1 \quad \cdots \quad k_j \quad \cdots \quad \text{iter}+1 \quad \text{iter}+1)_{1 \times q}$ ，其中，然后 iter 值自增 1。最终的 t_{iter} 应等于 $\max(K)$ 。

6.2.2 结果展示与分析

图 6.9 展现的是辛烷值（RON）损失随操作变量逐步变化所得到的预测值。图 6.10 展现的是产品硫含量随操作变量逐步变化所得到的预测值。观察发现，随着操作变量的逐步变化，虽然硫含量和辛烷值的损失都在减少，但其减少的速率也在逐步趋于平缓。

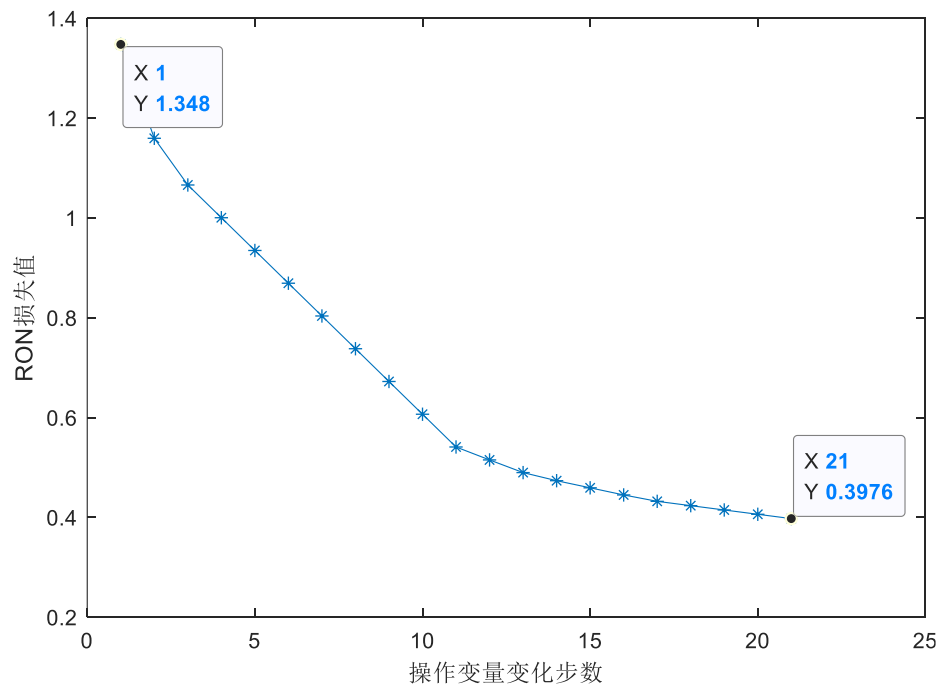


图 6.9 133 号样本在操作变量逐步变化下的辛烷损失值同步图

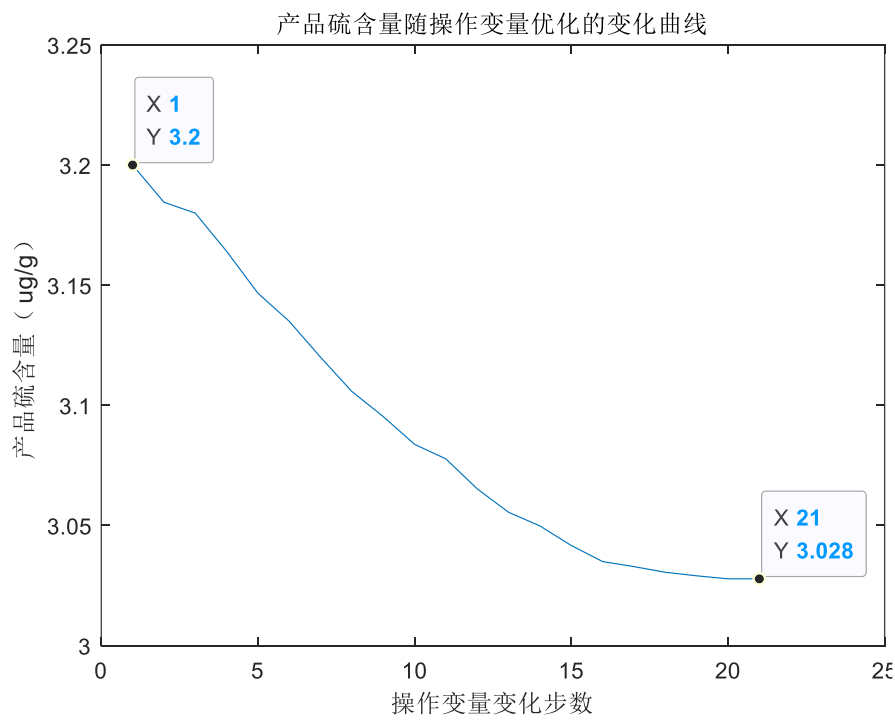


图 6.10 133 号样本在操作变量逐步变化下的产品硫含量同步图

参考文献

- [1] 袁俊, 周小伟, 杨伯伦. 基于 LM/SVM 方法的二次反应清洁汽油辛烷值预测[J]. 高校化学工程学报, 2010, 024(002):258-262.
- [2] 王瑾, 蒋书波. 汽油辛烷值 NIR 数据处理与建模仿真[J]. 计算机与应用化学, 2011, 28(007):947-950.
- [3] 高俊, 姚成, 章俊. 人工神经网络用于近红外光谱预测汽油辛烷值[J]. 分析科学学报, 2006(01):76-78.
- [4] 孙忠超, 山红红, 刘熠斌, et al. 用于 FCC 汽油辛烷值预测的非线性数学模型[J]. 炼油技术与工程, 2012, 42(2):60-64.
- [5] 刘永才, 刘传勤. S Zorb 装置汽油辛烷值损失偏大的原因分析与措施[J]. 齐鲁石油化工, 2012, 40(003):230-233.
- [6] 韩跃辉. 降低 S Zorb 装置汽油辛烷值损失优化措施探讨[J]. 中国石油和化工标准与质量, 2017, 37(006):87-89.
- [7] Miche Y , Sorjamaa A , Bas P , et al. OP-ELM: Optimally-Pruned Extreme Learning Machine[J]. IEEE Transactions on Neural Networks, 2010, 21(1):158-162.
- [8] Huang G B , Zhu Q Y , Siew C K . Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70(1/3):489-501.
- [9] Huang G B , Zhou H , Ding X , et al. Extreme Learning Machine for Regression and Multiclass Classification[J]. IEEE Transactions on Systems Man & Cybernetics Part B, 2012, 42(2):513-529.
- [10] Jing H , Tao Z , Yang G , et al. Particle Swarm Optimization Algorithm for Multi-objective Programming Problems Based on Elite Strategy[J]. Journal of Yangtze University(Natural ence Edition), 2018.
- [11] Goldberg D E. Genetic Algorithm in Search, Optimization and Machine Learning[J]. Addison Wesley, 1989, xiii(7): 2104–2116.
- [12] Such F P, Madhavan V, Conti E, et al. Deep Neuroevolution: Genetic Algorithms are a Competitive Alternative for Training Deep Neural Networks for Reinforcement

- Learning[J]. arXiv: Neural and Evolutionary Computing, 2017.
- [13] Zhou X J, Wang Q A, Zhang R D, et al. A hybrid feature selection method for production condition recognition in froth flotation with noisy labels[J]. Minerals Engineering,doi: 10.1016/j.mineng.2020.106201, 2020.
 - [14] Paul Ijjina E, Krishna Mohan C. Human action recognition using genetic algorithms and convolutional neural networks[J]. Pattern Recognition, 2016: 199-212.
 - [15] Srinivas N, Deb K. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms[J]. Evolutionary Computation, 1994, 2(3): 221-248.
 - [16] 周志华. 机器学习[M]: 清华大学出版社, 2016.
 - [17] Unler A, Murat A, Chinnam R B, et al. mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification[J]. Information Sciences, 2011, 181(20): 4625-4641.
 - [18] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
 - [19] Sahu R K, Panda S, Padhan S. A novel hybrid gravitational search and pattern search algorithm for load frequency control of nonlinear power system[J]. Applied Soft Computing, 2015, 29: 310-327.
 - [20] Yegireddy N K, Panda S. Design and performance analysis of PID controller for an AVR system using multi-objective non-dominated shorting genetic algorithm-II[C]. International Conference on Smart Electric Grid, 2015.