04烯焼制备分析与试验设计

摘 要

C；烯短可用于化学工业生产与医药制造，是重要的基础化工原料，是石油化 工产业的基础。近些年，随君国内化工产业的不断进步，Q烯姪的综合利用越来 越受到重视.乙酔催化偶合制备0烯姪逐渐进入人们的视野。因此,选择合适 的催化生产工艺,实现稳定、高效生产的目标，将有助于相关企业经济效益的提 升，同时也能促进我国的化工产业的发展。

针对问题•.首先，经过皮尔逊相关系数分析.计算每神催化剤组合下乙醵 转化率、C4烯燈选择性与温度的相关性：根据计算结果所得的相关性系数表可 知数据之间存在较强的相关性。以乙醇转化率、C4烯燈选择性为因变量・温度 作为自变锻进行非线性回归拟合，所有组合的相关性系数R2的均值在95%以上， 满足二项式非线性回归函数形式。再佃i出乙醇转化率以及烯経选择性与温度的图 像，对图像进行描述。其次对于第二小问.同时通过聚类分析对数据进行降维处 理，将5个附加产物变量降维成2组,在给定温度和催化削组合下对结果进行了 定居分析,发现随着时间的增加.在一定温度和相同催化剂下，随着时间的增加， 反应物乙酔转化率持续下降，逐渐趋于稳定；C4烯炷选择性増加，并趋于稳定， 与时间呈现出弱负相关性。

针对问题二.根据附件1中数据分析，得出“石英砂''对乙酔转化率以及C4 烯烧选择性的影响可以忽略，因此剔除相关该数据，同时利用问题一中的聚类情 况.对附件一中的数据进彳.降维,利用多元线性回归建模，并旦通过偏最小二乘 回归建模完善改进多元线性回归建模无法分析因変量之间的关系的飮点，得到两 种模型对应的残差平方和都在10左右波动，因此两模型差异很小，得出乙醇转 化率、C4烯煙选择性这两个因変量之间呈现低相关性。并且得出温度在解释这 两个因变量时,具有极为重要的作用.装料比的解释能力均为•般。

针对问题三,探充多个自变量与在相同的实验条件下，如何组合从而使得*C4* 墻炷的收率最大。通过题中所给的C4烯燈收率计算式.写出基飴单目标最优化 模型，同时求出C4烯姪收率的多元回归函数对模型进行优化，引入方差分析探 充多自变量之间的交互作用，从而史好地拟合数据最终求得温度在450C时C4 烯焼的最大收率为47.1%。当温度限制在35(TC时C4烯煙的最大收率为28.2%。

针对向题四，为了能够在5次实验内更加充分的了解如何提高C4烯燈的收 率，本文釆取均匀设计实验法•以更少的实验次数获取更多的信息。本文根据问 題二備最小二乘模型所得相关系数判断各个催化剂组合各组分与C4烯燈收率的 关系，以此来设置各因素水平及排除相关性极低可视为常量的因变位•满足均匀 设计5此实羚对自变堡的要求，并根据参数水平表设置实验方案。为了化学实验 操作安全，本文剔除了均匀设计使用表的最后一行，并以问题三所得最优结果代 替。最终，根据均匀设计实验法排列出了再増加的5次实验。

关键词 多元回归模型偏最小二乗回归分析最优化方差分析均匀设计

1问题提出

1.1 背景分析

Q烯蜂可用于生产聚丙烯等多项化工产品，是重要的基础化工原料.是优质 的汽油调和组分，是石油化工产业的基础。近些年,随若国内煤化工业的不断发 展.g烯姪的综合利用越来越受到重视，乙醇催化偶合制备a烯姪逐渐进入人 们的视野。⑴

因此.选择合适的催化生产工艺，达到相定、高效生产的目的.提高a烯 燃的选择性和收率.不仅能为企业带来明显的经济效益，同时也能促进我国相关 化工企业的发展“ 1.2 问题重述

某化工实验室针对不同催化剂在不同温度下做了一组实验，结果如附件1和 附件2所示。语通过数学建模完成下列问题：

问题-：针对附件1所给的催化剂组合，研究其乙醇转化率、0烯燈的选择 性与温度的关系；并分析附件2所给定的催化剂组合在一次实验中350C下不同 时间的实验结果进行分析。

问题二：研究不同催化剂组合和温度对乙醇转化率和G烯炷选择性的影响。

问题三：如何选择催化剂组合与温度，使得在相同实验条件下C4烯燈收率 尽诃能高，若使温度低于350度・又如何选择催化剂组合与温度，使得C4烯姪 收率尽可能高。

问题四：如果允许再增加五次实羚.应如何设计,并给岀详细理由。

2问题分析

2.1 问题一的分析

问题-要求探讨对毎种催化剂组合分别研究乙醇转化率、C4烯煙的选择性 与温度的关系并对附件2中确定的催化剂组合和温度下，分析反应物生成物随时 间变化趋势。在第一小问中，观察附件1中的数据，发现温度与两项指标有着某 种关系，利用皮尔逊相关系数判断数据足否存在相关性。通过对数据进行折线图 的絵制以及非线性和线性拟合，最终确定温度对乙睥转化率以及C4烯炷选择性 大小存在二次函数的关系。同时得出各个催化剂组合下的折线图，得岀其中不同 催化剤组合条件下确切的影响关系。其次•我们进行了第二部分，对附件：中的 数据进行处理。因为给定了温度和催化剂组合，数据两项指标变为了与时间有关 的变常，因为产生的无关产物也会对两项指标产生影响.对无关产物进行了聚类 处理，就可以实现変欧的降维。从而建立二次拟合方程对変量进行拟合，得到二 次函数形式的回归方程.从而得到两项指标与时间的具体函数关系。

2.2 问题二的分析

问题二要求探讨不同催化剂组合及温度对乙醇转化率以及C4烯燈选择性大 小的影响。在问题-的基础上，我们发现数据之间都具有较强的相关性。不妨对 附件•中不同组合的催化剂、温度、以及生成物的比率设置赋予变量值，对其进 行多元线性回归。在进行相关性检验时，发现变量之间具有很强的相关性，并由 此得到多元线性方程.得到的回归方程拟合程度较为理想。为了脸i正因变量之间 也存在影响，同时又进行了偏最小二乘的回归分析进行对比,判断得到的回归方 程拟合度与多元线性回归所得结果差别大小，若相差不大说明因変量:之间对两项 指标的影响不大，若相差较大，，浅需对比谁的显著性水平更高从而进行选择回归 方程。得到的回归方程就可以作为不同催化剂组合和溫度对两项指标的影响指标 进行技化分析。

2.3 问题三的分析

问题三要探究-组催化剂组合与温度在相同实验条件下的最佳组合，以C4 烯焼收率为目标函数.寻找使得目标函数最大值时的催化削、温度配比。根据C4 烯燈收率的求解式，因约束条件不充分，无法求解最优值.加入C4烯烛收率的 多元线性回归分析进行模型优化.而C4烯経收率不仅受到来自乙酵转化率与C4 烯焼的回归系数的限制.还有可能受到其他变量的交互影响。因此,对多自变址 进行方差分析.得到了温度与Co负栽量以及乙带浓度的组合交互时对因变鼠也 产生-•定的影响.将之作为注意力自变加入自変敬中，进而再利用SPSS进行 拟合从而完成，得到了一个对温度信息更敏感的方差分析模型。

2.4 问题四的分析

问题四要求再增加5组实验来探索各个自变量（催化剂各组分）与因变量

（C4嬌炷收率）之间的关系。为了在有限次的实验内更加充分的了解如何自变 量与因变量之间的关系.本文釆取均匀设计实脸法。本文根据问题二偏最小二乘 模型所得结果首先剔除了与因变量相关性极低的自变量，以满足均勾设计5次实 验对自变量个数的要求.并根据化学实验安全原则剔除了均匀设计使用表中可能 使化学反应过于剧烈而导致危险发生的组合.并以问题三所得最佳最优结果代替。 核终，根据均匀设计实验法设计出了所求的5种实验。

3模型假设

为了构建更为精确的数学模型，本文根据实际情况做出以下合理的假设或条 件约束：

◊ 假设一：乙醇催化偶合制备履烯姪的过程中，除给定催化剂组合以及 温度和反应时间外.其它条件对实验结果的影响可忽略不计。

◊假设二：在反应进行时，实验室内环境温度和气压不会对反应造成影响。

◊ 假设三：实验设备完好且在实验过程中不会岀现漏，、破损等情况导致 实脸失败或数据出现差错。

4符号申明

本文中涉及到的主要变量符号说明：

|  |  |  |
| --- | --- | --- |
| 符号 | 符号含义及说明 |  |
| *Vi* | 多元线性回归方程的因变量，*i* = 1,2 |  |
| 灼 | 多元线性回归方程的自变量，*i = l,2* |  |
| & | 偏最小二乘回归分析的原始因变景，*j* | =1.2,3,4 |
| *xi* | 偏最小二乘回归分析的原始自变量，;=1,2,3.4,5 |
| *y^j* | 偏最小二乘回归分析的标准化因变燉, | *3* = 123,4 |
| g | 值最小二乘回归分析的标准化自变量, | >=1.2,3,4.5 |
| 丄 | 皮尔逊相关系数 |  |
| u, | 二项式回归函数，i = 1,2 |  |
| 0m | 偏回归系数，*m* = 1,2,3,4,5 |  |
| 月 | 第J•个自变量X,样本均值 |  |
| *8i* | 第j个自变改勺样本标准差 |  |

注：未申明的变畳以其在符号出现处的具体说明为准。

设豹表示乙擀转化率，洗表示C4烯炷选择性.旳表示Co/SiO2质暈，巳表 示C。负栽量，％表示HAP质量，勺表示乙醇浓度，工5表示温度。

5模型建立

5.1 问题一的求解与分析

根据题意将问题分为两个小问。第一小问对实验类型分组.通过在数据的可 视化以及非线性函数拟合来分析每种催化剂组合、温度对乙酔转化率、C4嬌姪 的选择性影响分析。第二小问通过对数据聚合，可视化数据变化以及函数拟合关 撻变量从而对结果进行分析。

5.1.1数据预处理

根据各个催化剂组合的组分不同.本文将所给的催化剂组合分成4类，如下 表所示：

表1:催化剤组合分类表

|  |  |
| --- | --- |
| 组别 催化剂组合 | 分类依据 |
| 1 A1-A6 | Co负载餃 |
| 2 A7-A12 | 乙醇浓度 |
| 3 Bl—B6 | Co/SiO2和HAP装彩比 |
| 4 A13—A14A7-A8 | 特殊对照实验 |



*Class* 1：

在Al至A6中，各组的装料比率相同，因此本文将其作为一类实验进行分 析。

且A4.温度400-CT.乙斡转化率在实验组中达到最大，而在催化剂组合 A3.温度400C下，C4烯姪选择性最高。

*Class 2t*

在A7至A9及A12中，催化剂组分中的Co负我量、装料比率都相同.因 此本文同样将其作为一类实验进行分析。

在这些实验组中可以分析得出.随若乙酔加入浓度速率上升时・乙醇转化率 逐渐降低。

*Class 3：*

在Bl至B6中，Bl、B2、B3、B4、B6中以催化剂质量比为变量进行研究， B5为其他变量:的对照试验。因此本文同样将其作为一类实验进行分析。

实验组B3、B4、Bl、B6、B2中质量比逐渐增加，由试脸结果分析可知， 随着温度的升高质量比越大C4烯燈选挿性就越高。B1与B2以乙醇浓度为变 匱.当升高乙醇浓度时，可以得知随着温度升高，乙博转化率与C4烯燈选择性 都略有下降。

*Class 4-*

在A13、A14的对比下，易知在该条件下.“乙醇转化率偶）”变化不大，但

“C4烯短选择性”波动。

A8与B7之间仅存在若装料方式的不同，有结果分析，两种装袋方式对乙 辞转化率、C4选择性的变化趋势的影响很小，但是经过第(罗马数字二)装袋 方式的效果总体优于第一种。

5.1.2问题一第(1)问模型的建立与求解



图1：数据分析流程囲

*Step*厶相关性分析

本文首先通过计算皮尔逊相关系数.每种催化剂组合下乙醇转化率、C4烯 烷选择性与温度的相关性。

在公式(54冲.c代表温度.y分别代表乙醇转化率和C4烯统选择性。

g) \_ 剧(工一气)(；一认)1

(5-1)

计算结果详见目录。在21组数据中有19组的Pearsg相关系数大于0.9,

因此，不同催化剂组合的条件下，温度和乙醇转化率、C,烯姪选择性数据极强相 关。

亥2:乙酩轩化率' 5爆炫选择性与温度的相关系敏亥

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 催化剂组合 | 乙辭转化率(%) | C4烯燃选择性 (%) | 催化剂组合 | 乙辭转化率(%) | a烯姪选择性(%) |
| A1 | 0.965 | 0.887 | A12 | 0.963 | 0.983 |
| A2 | 0.995 | 0.914 | A13 | 0.937 | 0.988 |
| A3 | 0.982 | 0.955 | A14 | 0.964 | 0.959 |
| A4 | 0.998 | 0.958 | Bl | 0.962 | 0.986 |
| A& | 0.934 | 0.97 | B2 | 0.962 | 0.9&5 |
| A6 | 0.S84 | 0.885 | B3 | 0.92 | 0.971 |
| A7 | 0.999 | 0.968 | B4 | 0.91 | 0.895 |
| A8 | 0.977 | 0.992 | B5 | 0.913 | 0.978 |
| A9 | 0.921 | 0.997 | B6 | 0.94 | 0.982 |

A10 0.923 0.861 B7 0.936 0.994

All 0.903 0.989

*Step* 2：非线性回归曲线拟合

由上述分析可知.温度和乙薛转化率、0烯焼选择性数据存在强相关性，且 数据呈现正相关的趋势，故本文釆用非线性回归的方式进行曲线拟合，分别利用 线性方程、"S“型曲线和二次曲线进行拟合，以催化剂组合分组，可得出21张拟 合图像：

注：因售橘所限.此是仅晨示2幅有代表性的拟合囹像.箕它拟合厨像相关性系数矽 相差不大.均在93%以上.故不在此处罗列.洋见附录.



观察所有拟台图像可知，三种拟合方式均冇较强的相关性。

*Step 3：*相关性检验

表3:不同拟合方式的相关性系数表

|  |  |  |  |
| --- | --- | --- | --- |
| 方程 | *R2* | 方程 | 依 |
| 綫性 | 0.932 | 线性 | 0.943 |
| 二次 | 0.980 | 二次 | 0.976 |
| S型 | 0.976 | S型 | 0.967 |
| 注：因篇幅所限.此处仅展示上述2悟有代表性的拟合图像的相关性系数R2计算结果. 其它拟合图像的相关性系数洋见附录. |

由完整的相关性系数表可知・所有图像的相关性系数&，的均值在95%以上， 且始终大于线性和”S”型曲线的相关性系数。因此本文认为，乙醇转化率、C4烯 焼的透择性与温度的具有二次相关性，且为正相关。

5.1.3问题一第(2)问模型的建立与求解

Step 1：数据预处理

首先，本文为获得时间与乙酔转化率和反应产物的量之间的关系，考虑到该 反应副产物选择性之间的数值关系，但是变量过于冗余，可以将相关性高的分为一类•对反应产物进行了降维处理，也就是对各个反应产物进行了分类合并处理， 利用SPSS软件进行系统Q型聚类，做出了所有副反应生成物数据谓系图（如下 图所示），并针对反应产生的副产物进行分类。



图3：各反应物聚类谱系图

分析系谱图可以发现，我们可以将乙烯选择性、苯甲基甲醪和苯甲基甲醇的 选择性、乙醛选择性以及其他生成物的选择性分为一类称作次要附加产物：脂肪 醇的选择性分为一类称作主要附加产物。从而实现数据的降推，因此在六种副生 成物中.我们只需要考虑两类生成物对乙醇转化率和C4烯經选择性的影响即可。

分析附件二中数据得岀下图乙醇转化率与各产物选择性图随时间变化图（如 图4所示）.



50%

40%

30%

20%

10%

0%

— ^nr

—主耍前加产物 一次妥附加产物 —乙醇转化二 —C4烧經选择性一收宰

图4：乙酔转化率与各产物选择性图随时间变化图

图5:乙醇转化率、C4烯炫选择性、C4臓收率随时间变化图

由图4可知.目标生成物中.0烯燈的选择性在40%附近波动，且趋向于程 定，其方差为1.1768 x】0T：主要附加产物的含量:在时冋区间［20^,1105）之间呈 减少的趋势.次要附加产物在该时间区间内呈现増加的趋势，且増加和滅少的变 化量:相比0烯燈选择性变化较大，而在110m之后.两附加产物的选择性趋于和 定：此外，乙盼转化率在整个实验中均呈减小的趋势。

假设乙辞转化率以及C4烯姪收率粉时间变化的函数关系式分别为：

| 表a：乙醇转化率 |
| --- |
| R | *R3* | 调墜后胆 | 估算标准误差 | 系gim |
| .994 | .988 | .982 | .679 | -0.105. 0.00017. 45.258 |
| 表5：烯燈选择性 |
| R | *R3* | 调堅后R2 | 估算惊准误差 | 糸数8m |
| .968 | 937 | .906 | .063 | 0.004、-0.6-10-. 1.112 |

\* =a11f2 + a12t + a13 *U2* = ciojt2 + 崩 + 经过spss软件的分析，海到.

表4：相关系致表

其中R；达到*J* 9&8%和93.7%非常接近1.假设成立，T以说乙酔转化率以 及C4烯燈收率与时间具有二次函数的关系，满足下列函数关系式

Ui = -0.105t2 + 0.00017t + 45.258

U2 = 0.004 - 0.6 x lO-t + 4.112

在一定温度和相同催化剂下，随着时间的増加.反应物乙醇转化率持续下降, 逐渐趋于襁定：C4烯姪选择性増加,并趋于稳定，与时间呈现出弱负相关性。

5.2 问题二的求解与分析



5.2.1数据预处理

1. 考虑到乙醇在催化偶含制条('1烯炷的过程中，所给定的催化剂组合毎 •种都占相应的对照实验.因此，本文选取所仃类型的实验，通过Python编程

软件中的pandas库对附件1中的催化剂组合按组分进行拆分，其中装料比是比 值的形式在数据中存在，不利于进行数据处理，因而将装料比拆分成Ce/Si02质 最和IIAP质景,得到单独的两个自变量:。最终我们得到5个自变暈。

1. 石英砂在实验中充当防暴沸的效果、在实验中起对照作用.并FL在加 入石英砂的实验组中乙邮转化率和('1烯烧选择性并效率没有明显増大.我们认 为石英砂不具仃催化作用。因此我们舍弃石英砂这•变既組，详细变或数据见附 录。
2. 同时，根据向題-中対反应产物的聚类结果，我们仍将杂质分为主要測 产物和次要副产物，并得到如下表所示的变量表。

表5:因変量自変量的选择表

|  |  |
| --- | --- |
| 自变隨 | 因变散 |
| 目惊产物 | 主要副产物 | 次要制产物 |
| 编码 | 变fit | 编码 | 变砒 | 編码 变觉 | 编码 | 变敬 |
|  | Co/Si02 |  | C1烯经选 |  |  | 甲基苯甲酸和甲基 |
| *xl* | 质址 | 成 | 择性 | 碳数为1- |  | 苯甲醛选择性 |
| *x2* | Co负载依 | 1/1 | 乙辭转化串 | 12 *HuVjnf* | 护 | 乙醪选择性 |
| *x3* | IIAP质战 |  |  |  |  | 其他生成物选择性 |
| 1\*4 | *乙醇浓度* |  |  |  |  | 乙烯选择性 |

*温度* 5.2.2模型建立及误差分析

根据附件1中的拆解数据作为自変覺.进行多元回归分析・从而得到标准回 归系数，进而分析对乙醇转化率、C4烯姪选择性大小的主要影响成分。针对本 題中两个因变豪，我们通过倡最小二乘法来改进原冇的分析方法，探究多因变景: 的分析.最终通过回归系数来分析乙酔转化率以及*C4*烯姪转化率的主要影响因 素。

模型1：多元线性回归模型

Part 1：方差分析

根据问题一，可知各个催化剂组合对最终反应生成物的影响存在某种线性关 系。此时则对不同的催化物和温度进行分析•判断两者对乙醇转化率以及C4烯 燈选择性大小的影响。以勺,％口3淄4,%作为自变量.以*y^y2*作为因变量进 行多元线性回归分析。

首先对变做进行方差分析。利用SPSS软件进行方差分析，得到如下表所示 方差齐次检验的显著性指标：

表6：方差齐次检蹬表

|  |
| --- |
| 乙醇转化率(yl) |
| 自变量 | Co/SiO2 质 敬(X1) | Co负栽憐(x2) | HAP质量(x3) | 乙辞浓度 (y4) | 温度(y5) |
| 显著性 | 0.125 | 0.043 | 0.041 | 0.05 | 0.08 |
| C4烯燈选择性(y2) |
| 自变量 | Co/SiO2 质 % <xl) | Co负栽量:(x2) | HAP质量(x3) | 乙酔浓度(y4) | 温度(y5) |
| 显著性 | 0.125 | 0.043 | 0.041 | 0.05 | 0.08 |

分析图表可知.方差的显著性均大于0.05,可以认为本次方差分析是齐的。 因此进一步进行単因素方差分析得到如表所示単因素方差分析表：

|  |
| --- |
| 表7：主体间效应检脸 |
| 海 | III类平方和 | 自由度 | 均方 | F | 显著性 |
| 修正模IH | 98469877.8" | 103 | 956018.231 | 51.598 | .000 |
| 雌 | 20101179.490 | 1 | 20101179.490 | 1084.889 | .000 |
| *xi* | .000 | 0 |  |  |  |
| \*2 | 188H19.655 | 3 | 627039.885 | 33.842 | .000 |
| \*3 | 531008. 332 | 1 | 531008.332 | 28.659 | .000 |
| X， | 861063. 569 | 3 | 287021.190 | 15.491 | .000 |
| *Xi* | 38683517.781 | 6 | 6447252.964 | 347.967 | .000 |

2014792.493

993214.612

13606<M.082

185283.350

137924965.473

98655161. 190

154984.038

248303.653

97192.434

18528.335

8.365

13.401

5.246

关于多个控制变量的独立作用部分.可以得出对因变量影响排序为弥，％, 必，旳，叼.说明温度的贡献量最大，即对C,烯姪收率的大小影响最大，Co的负 栽量贡献量最小，即对其大小影响最小。因此我们可以进行回归分析模型的建立。

Part 2：建立多元回归模型

多元线性回归分析的模型为：

八/ =禺+片旳+ 02旳+ "/亦& + %\_ .9 „

(5 — 2)

L~N（0,〃）， 2=1,2,...,”

其中，际缶,％…，。，/是偏回归系数,与％叼产3，...，％无相关性。€为 随机误差项。

假设，因变量与各自变量之间存在线性关系，那他们之间的线性总体回归模 型可以表示为：

pl =^0 + 01^11 + 32^12 + 83=13 + 8^14 + M15 + £1

(5-3)

= &> + 31=21 +。2叼2 + ^3^23 +。4处4 + 但刀25 +《2

其中，£为随机误差项€~N（0,°2）。

Part 3：求解回归系数及分析

下表是回归系数分析表。

表8:回归系数分斬表

|  |  |  |  |
| --- | --- | --- | --- |
| 模型 | 未标准化系数 | 标准化系数 | t 显著性 |
| B | 标准误差 | Beta |
| (\*\*) | -80.509 | 7.231 |  | -11.134 | .000 |
|  | -.034 | .071 | -.104 | -.476 | .635 |
|  | .134 | .898 | 007 | 149 | .882 |
| “ ％ | .141 | ,069 | .448 | 2.045 | .043 |
|  | -8.765 | 2.065 | -.199 | -4.244 | .000 |
| % | .333 | .019 | .763 | 17.530 | .000 |
| （常量） | •48.732 | 5.112 |  | .9.532 | .000 |
|  | .003 | .050 | .015 | .059 | .953 |
|  | -3.164 | .635 | -.271 | -4983 | .000 |
| 1,2 / | .086 | .049 | .462 | 1 769 | .080 |
|  | 2.673 | 1.460 | .102 | 1.831 | .070 |
|  | .181 | .013 | .701 | 13.489 | 001) |

并且得到的多元线性回归模型可以表示为：

% = -80.509 - 0.034^1 + 0.134x2 + 0.14 lx3 一 8.765x4 + 0.333i5

洗=—48.732 + 0.003X1 - 3.164x2 + O.O86X3 + 2.673x4 + 0.181X5

Part 4：误差分析

利用SPSS算变愤之间的相关系数.得表4.叩％与旳,％,3和為之间 显著相关.

|  |  |
| --- | --- |
|  | 表9:回归方程的显著性指标 |
| 模型 | R R方 | 调整后R方 | 标准估算的误差總宾-沃様 |
|  | y】 | .892 .796 |  | .786 | 10.55848 | .937 |  |
|  | 。2 | .842 70S |  | .696 | 7.46480 | .580 |  |
| 从表中姑果可以得出，经过逐步线性回归分析后拟合出的模型饥和s的R2值 分别为0.796和0.709,其值接近1,可以认为模型拟合程度较好. |
|  |  |  | 表10: | 残差统计表 | 乙障转化率残差统计 |
|  |  | 最小值 | 最大依 |  | 平均值 | 标准偏差 | 个案数 |
| 預測值 | -13.5101 | 83.1577 | 21.9846 | 20.36060 | 114 |
| 残差 |  | -18.60947 | 28.80237 | .00000 | 10.32224 | 114 |
| 标准預測值 | -1.743 | 3.004 | .000 | 1.000 | 114 |
| 标准残差 | -1.763 | 2.728 | .000 | .978 | 114 |
| 炳燈选择性残茏统计 |
|  |  | 最小值 | 最大值 |  | 平均值 | 标准偏差 | 个案数 |
| 预测值 |  | -9.1596 | 49.9478 | 16.4508 | 11.39512 | 114 |
| 残差 |  | -17.30414 | 22.09034 | .00000 | 7.29778 | 114 |
| 标准預測值 | -2.247 | 2.940 | .000 | 1.000 | 114 |
| 标准残差 | •2.318 | 2.959 | .000 | .978 | 114 |
|  | MCI \*)1化成蒸的壬态P妒用WftB.乙師”化率 |  |  | h«i际忘化我如的11 i p-p m |  |  |
|  |  |  | / | : |  |  | / |  |
|  |  | */* |  |  |  |  |  |  |
|  |  |  |  |  | *A* |  |  |  |
|  | / |  |  |  | */* |  |  |  |
|  | 林■累》«奉 |  |  |  |

图7:回归标准化残差的正态D-P图

由表9和图6的残差分布表图.残差最大值均在25左右，且标准化残差都 集中分布在直线附近.可以认为标准化残差满足正态分布・误差通过.建立的回 归方程合理且误差较小。

模型2：偏最小二乘法回归模型

Part 1：模型优化机理

上文在利用多元线性回归建立模型时，分别建立了gl,y2所对应的多元线性 回归模型，而对于在化学背景下的多因变时问题中，因变鼠之间很有可能产生相 互影响,针对本题中两个因变量，我们通过偏最小二乗法来改进原有的分析方法， 探究多因变量:的分析.在原有的基础上探究因变量之间是否存在影响。

Part 2：构造偏最小二乘法模型

依据节5.2.1中因变量:和自变最的分类.存储在下面两个矩阵中：

自变量:的观测数据矩阵，4 = (%,)｛唤必｝,因变饿的观測数据矩阵为8 = (^•J )108x4 •

Step 1：数据标准化

将各指标％转化成指标值弓，有

时,2=1,2,…,108/= 1,2,3,4.

(5-4)

其中，岬=点£隼,，罗=届理殍可"1，2,3,4,即

相气9羿为第顶个自变量勺的样本均值和样本标准差。对应地，称

(5-5)

同理，有标准化指数值

b — (2)

b，= ~~9%\*7~~=1，2，・“，108/一 1,2,3,4. (5-6)

其中，噂=点迁X，始=\縞绪(寸助'，j = W,4.即

卩?),$羿为笫顶个自变量气的样本均值和样本标准差。对应地，称

9；=~~七［"~~。=1，2,3,4,

*si*

(5 — 7)

Step 2：求相关系数矩阵 下表给出了这9个変量的简单相关系数矩阵。

表11：简单相关系敬矩阵

*xl x2* x3 *x4 x5* yl *yl y3* y4

*xl* 1 0.2089 0.瞻00 -0.2992 -0.0018 0.3946 0.3799 -0.1748 -0.0351

*X2* 0.2089 I 0.2136 0.1413 -0.0142 0.0418 -0.1650 0.0979 -0.0137

*x3* 0.9800 0.2136 1 -0.3017 -0.0014 0.4065 0.3876 -0.0799 -0.1653

*x4* -0.2992 0.1413 -0.3017 I -0.0385 -0.3312 -0.1071 0.0237 0.0435

-0.0018 -0.0142 -0.0014 -0.0385 1 0.7706 0.6998 -0.7032 0.4370

jrl 0.3946 0.0418 0.4065 -0.3312 0.7706 1 0.7353 -0.5935 0.2679

*y2* 0.3799 -0.1650 0.3876 -0.1071 0.699S 0.7353 1 -0.6634 0.1752

jf3 -0.1748 0.0979 -0.0799 0.0237 -0.7032 -0.5935 -0.6634 I -0.8529

*y4* -0.0351 -0.0137 -0.1653 0.0435 0.4370 0.2679 0.1752 -0.8529 1

从相关系数矩阵可以看出Co/SiO2质量与Co负栽量正相关.Co负栽量与 HAP质量:、乙醇浓度正相关，HAP与乙醇浓度负相关，乙醇转化率与Co/SiO2 质量、Co负栽量、HAP质量和温度呈正相关，与乙靜浓度呈负相关：C4烯炷选 择性与Co/SiO2质量、HAP质量:、温度呈正相关,与Co负栽量、乙酔浓度呈 负相关.

Step 3：分别提岀自变量组和因变暈组的成分

由Matlab程序（详见附录），可以得到三对成分[ul,vl],[u2,v2],|u3,v3], |u4,v4], [u5,v5]。

C

 = —0.0296ai + 0.0063% - 0.0265x3 + 0.015% - O.OTOSXj

=-9.22912/j 一 11.8003% - 7.799切

C

 = -0.0218Xj + 0.0158% - 0.037x3 + 0.0229^ + 0.0602x5

=-2.56530 - 2.27591/2 + 6.7408勿

*fu3 =* 一0.025旳 + 0.082lx2 一 0.0325% - 0.068x4 + 0.0106x5

*\v3* = 1.8275^1 — 2.3454% + 0.4012^

C

 = —0.1090跖 + 0.051 旗2 + 0・082奴3 — 0.0681 工一 0.0106x5

=0.9458s - 0.2164% - 2.5390%

ru5 = -0.4577Xj + 0.0083x2 + 0.463x3 + 0.0159x4 + 0.0034%

1% = 2.3562s - 0.1258% - 8.218豹

而前四个成分解释自变紙的的比率为98.69%,因此只要选取前四对成分即 可满足分析需求。

Step 4：求两个成分对时标准化指标变量与成分变量之间的回归方程

求得自变量:和因变紙组与％ ％. 吟，之间的回归方程分别为

*(X,* = -6.6226% - 7.7166u2 - 2.1872u3 - 2.0490k,

*x2* = -0.2756U! - 2.9268% + 6.7982u3 一 7.4630u4

*x3* = -6.6197% - 7.76331/2 - 2.1411 % - 1.6233%

x4 = 4.006% + 4.066侦2 - 4.4631% - 7.6109%

« x5 = —8.0602% + 6.7369电 + 1.3742% — 0.8111u4 (5 \_ 幻

= 一9.2291% + 1.6398临 + 1.2505% + 0.3285%

\*2 = -8.3245% + *2.2226u2 -* 2.3624% - 0.1029u4

*y3* = 6.5551U1 一 4.1319u2 + 0.6684% + 1.495u4

O = -2.8172u. + 3.886叱 + 0.7688u3 - 1.8953为

Step 5：求因变量组与自变最组之间的回归方程。

把步骤2.5中成分％代入Step 3中y；的回归方程，得到标准化指标变量:之间 的回归方程

将标准化变缺况,耳(项=1,2,3,4)分别还原成原始变量:孙匂・得到回归方程 *y1* = -79.8445 + 0.0554% + 0.0351x2 - 0.0538旳-9.18S6& + 0.3323rs *y2* = -48.3595 + 0.0517Xj - 3.2176x2 + 0.0387x3 + 2.4439x., + 0.1807% *y3* = 184.2684 一 0.103% - 1.8552x2 + 0.0287x3 - 7.066x4 一 0.3518xs 3  *yA* = -35.909 + 0.0515旳 + 1.3624 - 0.0674x3 + 4.622lx4 + 0.1710

并且根据分别计算出R2为。.812和0.765.其值接近于1.可以认为拟 合情况较好。

模型的对比与检验

Part 1：模型对比

当比较两个模型在同一组数据上的拟合效果时.更倾向于比较他们的残差平 方和来确定，当残差平方和越小时.说明该模型的拟合效果越好。因此利用 Python程序分别计算两模型所对应回归函数的残差平方和.进行对比。如下表 所示：

表12：均方误差模型参数对比表

|  |
| --- |
| 多元线性回归 |
| 因变量y | 砂 | 均方误差(MSE) |
| 乙酢转化率＜yl) | 0.796 | 10.348 |
| C4烯姪选择性(y2) | 0.709 | 9.474 |
| 帳最小二乘回归分析 |
| 乙酵转化率(yl) | 0.812 | 10.404 |
| C4烯炷选择性(y2) | 0.765 | 7.057 |

本文认为，在本题中可以宜接用多元线性回归分别对因变量进行建模’可以 减少考虑本文中多个因变堡之间的联系。因为由图中数据可知，两模型所对应的 妲,猝的残差平方和相差很小，R2的差ff［也很小，说明无论是通过多元线性回归 建模，还是通过偏最小二乘法建模得到的回归系数都可以很好的解释多个自变砒

与因变攻之冋的相关性。

Part 2：模型的解释

利用偏最小二乘法建模时.根据预测值和实际值做出关于中心轴线的离散图， 得到乙醇转化率预测图以及烯烧转化率预测图，根据下图数据分布情况得知拟合 效果良好。



图8：乙醉转化率' 烯経蒋化率预测图与偏最小二柔回归系数百方图

从冋归系数图中可以观察到.温度变貝:在解释两个乙衅转化率、C1烯燈网归 方程中起到J'极为重要的作用，当温度和其它门变W相组合时，对因变賃的结果 也会产生重要的影响。H Co/Si()2与IIAP的装料比的解释能力在两个冋归方程 中差别不大，均为正相关。而C。负救最在乙醇转化率中琏本没冇解释性.但是 在C1烯姪选择性中冇较强的负相关解释性。而乙醇浓度则对乙醇转化率起负相 关解释性，而对C1烯烧选择性很少斉解释性。

5.3 问题三的求解与分析

为了求解組催化剤组合与温度在相同实验条件下的最佳配比.我们首先根 据化学反应关系得出C1烯烧收率的直接影响因素，又雄于何题：中的结论，推 导岀C4烯烧收率与多个自变遺之间的影响关系,建立了单目标最优化模型，并 H对多个自变早:进行方差分析.对原有模型进行改进，从而在两种温度条件情况 下进行全局最优的求解，以获得使得C1烯烧收率最大的自变敬配比。



图9：问题三解题流程图

5.3.1 C4烯鰻收率的模型建立

5.3.1.1烯恨收率的基础模型

Step 1：数据处理

继续参照何題 得到的聚类结果，将生成物分为C4烯姪、主要阳加产物(脂 肪醇)、次要附加产物，进行数据分析。

由收率公式：C4烯炷收率=乙醇转化率xC4烯煌选择性，可求得任意 顼催 化剂组合所对应的C4烯炷收率，因此后两项为C4烯炷收率的直接影响因素。考 虑到本題中需要拟合优良数据.从而得到C4烯烧最大收率.而实验中存在部分 拟合效果欠佳的数据，因此通过对】09组C4烯炷收率进行排序，确定C4烯燈收 率值阈值为1%，以阈值作为筛选条件对数据进行初步筛选。同时对于…些效果 不佳的对照实验，也进行数据剔除.最终得到61组积极数据。

Step 2：基础模型建立

设*p林C4煽炷*收率，可以得到下列关系式：

P = 1/1 X j/2 (5 - 10)

5.3.1.2最大C4烯姪收率的单目标优化模型

Part 1：目标函数确定

*maxp = yx x y2* (5-11)

Part 2：约束条件确定

①约束条件1：所有生成物选择性之和等于1()()%，则有

=100

(5-12)

■ -1

②约束条件2：所有自变量满足己有实验设计的取值范围。根据问题二中对 附件一所进行的数据拆分，分析得到每一种自变量的实稣取值范围，因此一个约 束条件是：

X。£气 MX" = 1,2,...,5 (5-13)

其中，*Xlt,Xlr*分别表示自变量的左右区间。

Part 3：模型整合

综上所述.我能得到了最大C4烯煌收率的单目标优化模型：

*maxp = y1xy2*

詹"。。2

*(XI, <X,< Xi,*

1,2, ...5

(5-14)

(5-15)

5.3.2基于多元线性回归的模型优化

Part 1：模型改进机理分析

根据题目中的描述以及附件1中数据的特点，可知本数据为化学反应中的记 录值，无法提过足够有效的约束条件.经过测试无法直接利用上述的最优化模型 进行最优化求解。因此我们提出基于多元线性回归的C4燔姪最大收率模型。由 于C4烯煌选择率来自乙铮转化率与C4烯姪选择性的乘枳，其两者的函数表达 式是通过多元线性回归来拟合，因此可以直接用拆分得到的多个自变量对C4烯 燃收率来做多元线性回归分析，从而对原有的模型进行有效的改进。

Part 2：基于多元线性回归拟合C4烯炷收率

参照何题二中的多个自变钺参数.使用SPSS软件对C4烯焼收率进行同等 方法的数据拟合，可得到以下的分析结果：

依=0.751

|  |  |
| --- | --- |
| p，= —3619.647 4 | -1.236X, - 69.431XJ + 4.121 蜜-1276.736% + 11.995% |

由于*R-squeare >* 0.7,则我们认为该多元线性回归模型具有较高的可信性。

Part 3：获得多元线性回归改进的C4烯経最大收率的最优化模型

因此基于多元线性回归拟合的改进下，由上述分析可得C4烯烷最大收率的

|  |  |
| --- | --- |
| 单目标最优化模型： | max p'E翁M我i=u…5 |

5.3.3基于方差分析的C4烧鰻收率模型优化

Part 1：优化机理分析

由于C4烯燈收率的是由乙醇转化率与C4烯炷选择性相乘所得，因此对于 C4培炷收率来说.受到了乙醇转化率与*C4烯雑*转化率的系数的直接限制，因 此，我们这此引入方差分析，来看这些变量之间是否含有相互作用关系，从而来 对C4烯炷的单目标优化模型做优化。

Part 2：通过方差分析研究C4烯燈收率

根据问题二中的方差分析结果，x2 \* %、工4 • %与叼\*%之间存在交互作用

同时,根据两者的F值与均方差值可知其对C4烯燈收率的大小产生显:著影 响。因此我们将该三个交互作用部分作为相互作用因子，令x6 = x2x5, xT = %% % = 通过附件1中数据的合并处理.我们得到了 8个自变粮，进而 通过SPSS软件分析计算多元线性回归模型如下所示：

*p* =总(-3362.4S9 + 5.23&E] + 67.621% - 27.568x3 + 206.59&r4 + 11.56乌 -0.529死 + 0.086x7 一 1.15%)

Part 3：获得方差分析改进的C4烯燈最大收率的最优化模型

因此当考虑温度因素时，则通过方差分析对模型进行改进，由上述分析可得 C4烯燈最大收率的敏目标最优化模型：

ax

X X X • ===<X 貫r<」

X

5.3.3.1 C4烯经最大收率时的催化剤最佳配比模型的求解

Case 1：当在相同的实验条件时，使用多元线性回归优化的模型进行求解

在相同实验条件下时.所有的自变量的取值与附件1中的数据取值限制相 同.说明对于温度等主要影响C4嬌炷收率的变量没有受到特殊限制.因此我们 宜接选取基于多元线性回归改进的C4烯燈最大收率的最优化模型进行求解。

经过Ling。程序求得最优債为：47.1%

Case 2：当在限制温度的条件时，使用经方差分析二次优化模型进行求解 由上分析同理可得・此时题目中对温度进行了限制・因此我们釆用对温度等 主要影响系数有増强作用的方差分析改进的C4烯燈最大收率的最优化模型进行 求解。

经过Ling。程序(见附录)求得最优值为：28.2%

表13：主体间效应检验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 力1 |  | % | *X4* | 工3 |
| Case 1 | 200 | 0.5 | 200 | 0.3 | 450 |
| Case 2 | 200 | 0.5 | 200 | 0.3 | 350 |

5-4 问题四的求解与分析

为了能够在有限的实验次数内更加充分地了解如何提高乙醇选择性和C4烯 焼转化率，本文在实验设计时选择均匀设计的方法。

相比于正交设计，均勾设计只考虑试验点在实验范围内的均匀散布。因此， 相比于正交设计.均匀设计减少了进行实验的次数，且更加适合在较少的实验中 获取更多信息，能够在相同的实验次数下更迅速的找到最优解[8, 9]0

在乙醇催化偶合制备c4烯燈的实验中.对其产物的造成主要影响的因素有 xk工2、丁3、工4、工5等五种因素。

由于均匀设计的设计原理约束了亿最多只能包含4个自变紙，故本文参考 了前面的问题中做岀的偏最小二乘回归系数直方图，将相关性足够低且在直方图 中回归系数最接近。的自变量g乙醇浓度定为常量，在设计时若重研究其它4 种自变量与C4烯烧收率的关系。

根据附件1所给实验结果，本文在4种因素内选取5种水平，表示为"尸

均勾设计爲(54)表如下所示：



务因素水平表如下所示，波浪线表示该数据来自问题三所求最终结果。在设 计各因素水平的过程中.观察到偏最小二乗回归系数直方图中,反应温度的相关 系数远大于其他几项.因此在设定“温度”因索的水平时，超过附件1所给最大 值后増加的步进值较小.但通过研究“温度”因素更能够更好地确定使收率更高 的因素，以提高反应速率：相关性最小的因素方在超过问题三所得最佳结果后的 步进值较大：而旳、“3的相关性在花与乌之间，但总体更接近％，其因素水平取 值也比较粮近实验1中己存在的数据。

表8S：均匀设计中各■■水平



75

100

3.5

300

400

450

475

500

因化学试验存在特殊性，若将最高水平组合在一起可能会导致化学反应过于 剧烈，从而使反应容器龟裂，基至发生爆炸等危险.故在均匀设计使用表中，移 除最后一行.得到更加均匀、安全的使用表，以防止事故的发生。此时.设计方 案变为

为了补全笫5次实验，现有2种方法可以选择：①将问题三使用多元线性回 归优化的模型所待结果作为第5次实验：②将#6的均匀设计使用表的最后一行 删除，化作区使用。但②因均匀设计原理决定了进行6次实验最多只能存在2 个自变量，故舍弃该方法。

表16：使用均匀设计増加的5蛆新实验表

实验号 \*1 妃 屛 M

x3

"300-

50

400

100

200

xl

~10~

50

75

100

200

500-

475

450

350

380

小结设计的实验方案如上表所示。 第一次实验条件： 第二次实验条件： 第三次实验条件： 第四次实验条件：

500C 下.

470C 下，

450C 下,

350C 下.

380C 下.

lOing 2wt% Co/SiOz. 300ing HAP。 50mg 5wt% Co/SiOa. 50mg HAP。 75mg Iwt% Co/SiO：. 400mg HAP。 lOOmg 3.5wt% Co/SiO2, lOOing HAP。 200mg 0.5wt% Co/SiOa\* 200mg HAP。

第五次实验条件：

灵敏度分析

根据问题二中的分析可知，温度等自变紙对因变量:的变化冇着强烈的影响， 因此我们通过对C4烯炷收率的回归方程中温度的系数进行灵敏度分析.使其值 经过上下5%的数据波动.絵制出C4烯炷收率的变化情况，如下图所示：

灵敏度分析



图10：灵敬度分析囹

温度的回归系数在10%的变化区间内引起了 C4嬌炷收率18%的变化，比较 稳定，因此表明模型具有较好的灵敏性。

表17:灵敏度数据变化表

|  |  |  |  |
| --- | --- | --- | --- |
| C4焰経收率 | 温度回归系数 | CI燃紹收率 | 温度回归系数 |
| 19.176759 | 10.982 | 21.604359 | 11.6756 |
| 19.662279 | 11.12072 | 22.089879 | 11.81432 |
| 20.147799 | 11.25944 | 22.575399 | 11.95304 |
| 20.633319 | 11.39816 | 23.060919 | 12.09176 |
| 21.118839 | 11.53688 | 23.546439 | 12.23048 |

7模型的评价与推广

7.1 模型的优点

L方差分析将有助于探究自变缺之间的相互作用，从而探充更深层次自变 貲之间的联系，以此将存在相互作用关系的毎一组变最存为新的一个変 贵.反过来作为多元函数中的一员，提供更好的拟合效果。

1. 偏最小二乘法和多元线性.回归两者相结合进行对比分析，可以探究多个 因变量之间是否存在着更深层次的联系.更有利于化学实验型数据的拟 合，使其得到更为准确的拟合方程，探究实验物之间的联系，对实验结 果进行预测。

7.2 模型的缺点

1. 基于多元线性回归与方差分析的C4炳姪最大收率模型中.各个组成部 分的权重的确定仍然是按照变技贡献率特点进行分析.并无一个完整的 体系，含有一定的主观性。
2. 通过偏最小二乘法进行研究时，只是对应线性多元回归中的乙酔转化率 和C4烯短选择性这两个因变量进行分析，得出了因变量之间相关性不 大，但模型没有分析其他因变玷与只是对应线性多元冋归中的乙醇转化 率和C4烯燈选择性是否存在若隐含关系。存在一定的偶然性。
3. 在聚类分析时，只是弟纯的晴化考虑了数値上的聚类.而忽略了其内部 化学结构之间联系.使结果偏离于真实情况。

7.3 模型的推广

本文问题二中主要应用了多元线性回归以及偏最小二乘回归分析模型。

从横向来看，在加入了偏最小二乗法回归分析对于多个因变量，有利于研究多因 変景之间的关系，更深层的关注多个因变量:之间的潜在相关性，有助于我们分析 在化学反应时，更全面的分析各个生成物之间的关系，此时可以再对该模型进行 延伸，在医学诊断、化学实验预测中发挥更大的作用，做到••未卜先知

8参考文献

III任丽阵赵国良，滕加伟,王仰东，谢在库；La修饰ZSM-5分子筛催化剂用于S 烯燈催化裂解制丙烯卩］;工业催化;2007年03期.

［2］ 吕绍沛.乙酔偶合制备丁醉及U烯姪［D］.大连理工大学,2018.

［3］ 朱水东.化学实验中的暴沸与沸石［J).高考(综合版),2012(11):127.

141杨忠赞，迟凤琴，隋虹均,匡恩俊，张久明,宿庆瑞，张一安，刘亦丹.基于多元线性 回归研究有机肥替代对土壤并分及产匱的影响卩］.东北农业科学，2021, 46(02):37-42+102.

1. 王赫然，尉佳，范忠仁.基于多元线性回归的高新技术企业发展影响因素分析［J］. 科技通报,2020,36(12):112-115.
2. 石彦国，单彤彤，曾剑华,刘琳琳，朱秀清.基于主成分分析和偏最小二乘法的蒸 意大豆食味品质评价卩］.中国食品学报,2019,19(10):265-277.
3. 李会芳，刘静婷，郎霞，路晓娟.基于偏最小二乗法关联分析梔子不同炮制品化 学成分与肝肾毒性卩］.药物评价研究,2021,44(09):1890-1896.

181浦仕彪，王钺涵,杨竹雅，刘录，周志宏.均匀设计法对临床方剂胃肠炎宁颗粒组 方优化研究［J］.云南民族大学学报(自然科学版),2021,30(04):343-346.

191均匀设计及其应用卩］・方开泰.数理统计与管理.1994(01)

A.问题一第(1)问结论

|  |
| --- |
| 催化剂组合分类表 |
| 组别 | 催化剂组合 | 分类依拥 |
| 1 | Al—A6 | Co负毂量 |
| 2 | A7-A12 | 乙酹浓度 |
| 3 | Bl—B6 | Co/SiO2和HAP装料比 |
| 4 | A13—A14 *A7—A8* | 特殊对照实验 |

*Class* 1：

在Al到A6的实验组中，催化剂组合各分组的质量比为1.因此本文将其 作为一类实验进行分析。



图11：催化剂組合A1-A7T相关性图 IM乙辟转化率的・・C4烯燈选择性

*(%)*

根据催化剂组合特点，可以将Al、A2、A4、A6作为一组的对照实验，对 照变量为Co/SiO2的含量，根据图像变化特点可以分析得出.当Co/Si02的负 我量为Iwt%且在温度区间为(250C, 300-C)时，*C4烯燈*的选择性远高于其他含 屈的情况，但随若温度不断升高，其他组呈上升趋势・A1组虽趋于平缓.但仍 多余其他组。

而A5中的Co/SiO2含含量比A3中的多一倍.但乙酔相对其他组过少.导 致最终C4选择性少于大部分对照实验.但与Co/SiO2过多的A6组的选择性类似。

*Class 2：*

在催化剂组合A7至A9以及A12中.催化剂组分中的Co负裁景:均为Iwt%, 且lwt%Co/SiO2与HAP含蛍均为50mg,因此本文同样将其作为一类实验进行 分析。



根据附件1易知・从A7、A8、A12与A9这四种催化剂组合中，其乙醇每 分钟加入速度从0.3ml/min递增至2.1ml/min0由上图可以得出.在乙酔加入速 度越快的情况下，乙醇转化率逐渐降低，在温度区间（254TC, 350C］上，C,选择 性基本不变.在温度区间（350C, 400\*0］呈现先升高后降低的趋势。

*Class* 3：



图13:催化剂组含A13、A14T相关性图

I乙辞转化寧(%)^C.烯炷选择性

在第II种装袋方式.实验组Bl至B6中，Bl、B2、B3、B4、B6中以催化 剂质量比为变量:进行研究，B5为其他变量的对照试验。因此此本文同样将其作 为一类实验进行分析。



实验组B3、B4、Bl、B6、B2中质量比逐渐增加.由试验结果分析可知， 随着温度的升高质量:比越大C4嬌炷选择性就越高。B1与B2以乙醇浓度为变 量.当升高乙博浓度时，可以得知随希温度升高，乙醇转化率与C4烯姪选择性 都略有下降。



*Class* 4：

在催化剂组合A13, A14的条件下.易知其组合的变鼠lwt%Co/SiO2与 HAP-乙辭浓度的含量不同.其含蛍“互换）在该条件下，"乙醇转化率（勁”变 化不大，但“C4烯烷选择性”波动。

A8与B7之间仅存在着装料方式的不同.冇结果分析.两种装袋方式对乙 醉转化率、C4选择性的变化趋势的影响很小.但是经过第〈罗马数字二）装袋 方式的效果总体优于第一种。

B.问题一拟合图像与相关性系数

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 模型摘要自由度1 | 自由度*2* | 显著性 | 善数估算（ft |
| 常量 bl | b2 |
| 线性 | *.932* | *U2M* | 1 | 3 | 0C8 | ■84.074 | .333 |  |
| 二次 | .980 | 48.340 | 2 | *2* | .020 | 141.807 | -1.194 | (MB |
| S | 983 | 178.598 | 1 | 3 | .001 | 10.708 | -2470.152 |  |

模型摘要和参数估算值

因变虽乙酔转化率1

自变量为S3.



模型摘要和参数估算值 因变員：財烤繰选择性I

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 模型嫡要 自由度1 | 自由度2 |  | 参数？fi算值 |
| Al bl | b2 |
| 綫性 | ,787 | 11.079 | 1 | 3 | .045 | -3.242 | .154 |  |
| 二次 | .916 | 10899 | 2 | 2 | <84 | -190.783 | 1.422 | •.002 |
| S | .848 | 16.784 | 1 | 3 | .026 | 4.898 | -338.868 |  |

自变宣为温度.



模型摘要和参数估算值

因变量：乙酔转化虜2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 自由度1 | 自由度*2* 昱若性 |  |
| «1 bl | b2 |
| 税性 | .990 | 297.825 | 1 - | 3 .000 | -161.892 | .663 |  |
| 二次 | .991 | 111.381 | 2 | 2 .009 | -227.415 | 1.106 | -.001 |
| S | .944 | 50.190 | 1 | 3 .006 | 11.258 | -2366.823 |  |
| 自变昆为温度. |



模型摘要和参数估算值

因变宣：C4墉燈选择性2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 自由度1 | 自由度2 | 斗著性 | 参数估芽值 |
| 常H | bl | b2 |
| 践性 | 836 | 15.285 | 1 | 3 | .030 | -41.546 | .222 |  |
| 二次 | .980 | 49.722 | 2 | 2 | .020 | 234.745 | -1.646 | .003 |
| S | .782 | 10.758 | 1 | 3 | (M6 | 5.559 | -709.059 |  |
| 自变St为*混度.* |



模型摘要和参数估算值

因变員乙酔转化風3



| 方程 | *R2* | F | 自由度1 | 自由度*2* 昱若性 |  |
| --- | --- | --- | --- | --- | --- |
| «1 bl | b2 |
| 税性 | ,964 | 134.968 | 1 - | 5 .000 | -95.828 | .419 |  |
| 二次 | .966 | 57371 | 2 | 4 .001 | -134.189 | *.647* | .000 |
| S | .974 | 185.577 | 1 | 5 .000 | 7.421 | -1245.588 |  |

自变量为温度A3.

模型摘要和参数估算值 因变宣：C4 *1AH&*择性3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R1* | F | 供型策要 自由度1 | 自由度2 4著性 | 参数佔H舰 |
| 常 B. bl | b2 |
| 践性 | .913 | 52349 | 1 | 5 .001 | -59.195 | .261 |  |
| 二次 | .955 | 42519 | *2* | 4 .002 | -171.076 | .924 | -.001 |
| S | .931 | 67.340 | 1 | 5 .000 | 7.333 | -1380.400 |  |

自变SL为涅度A3.



模型摘要和参数估算值

因变量：乙酔转化室4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 自由度1 | 自由度*2* 昱若性 |  |
| «1 bl | b2 |
| 税性 | .995 | 801.582 | 1 - | 4 .000 | -144.540 | 582 |  |
| 二次 | .996 | 362.876 | 2 | 3 .000 | -107.611 | .349 | .000 |
| S | .951 | 77.767 | 1 | *A* .001 | 9.953 | -2066.750 |  |
| 自变量为温度A4. |



模型摘要和参数估算值

因变宣：C4墉燈选择性4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 候型捅要自Ei度1 自由度2 | 畏著性 | 参数估尊（ft |
| bl | b2 |
| 践性 | .917 | 44 368 | 1 4 | .003 | -52.411 | .227 |  |
| 二次 |  | 62381 | 2 3 | .004 | 72.773 | -.562 | .001 |
| S | .871 | 27.095 | 1 4 | 006 | 6.317 | -1090.270 |  |

自变SL为涅度A4.



模型摘要和参数估算值

因变量：乙酔转化率5

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F |  | 模型捕耍 自由度1 | 自由度 | 2 SL若性 | «1 | 参敬估算（ft bl | b2 |
| 税性 | *873* | 27.503 | 1 |  | 4 | .006 | -97 5® | .4(8 |  |
| 二次 | *994* | 249910 | 2 |  | 3 | .000 | 232407 | -1.672 | 003 |
| S | .882 | 29.952 | 1 |  | 4 | .005 | 6.973 | -1149.781 |  |
| 自变或为温度A4・ |  |  |  | 3«4化》» |  |  |  |  |  |
|  |  |  |  |  |  | J |  |  |  |  |
|  |  |  |  | z | / |  |  |  |  |  |
|  |  |  | >2? |  | / | J |  |  |  |  |
| 因变量：方程 | C-＞爛燈选择性5*R2* F | 模型摘要和参数估算值 候型捅要自由度1 自田度2 里著性 |  | 参数估尊｛ftbl | b2 |
| 践性 | .940 | 62.768 | 1 |  | 4 | .001 | -57.813 | .230 |  |
| 二次 | .991 | 157.173 | 2 |  | 3 | .001 | 57.884 | -.500 | .001 |
| S | .963 | 103.385 | 1 |  | 4 | 001 | 8.276 | -1833200 |  |

自变SL为涅度A4.

模型摘要和参数估算值

因变量：乙酔转化室6

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 自由度1 | 自由度*2* 昱若性 | 卷故估芽值 |
| bl | b2 |
| 税性 | .968 | 89.414 | 1 - | 3 .003 | -119.731 | .501 |  |
| 二次 | .986 | 7QOI2 | 2 | 2 .014 | 5I.0M | -.577 | 002 |
| S | .943 | 49.586 | 1 | 3 .006 | 7.814 | -1362.320 |  |

自变量为温度A6.

模型摘要和参数估算值 因变宣：C4墉燈选择性6

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方程 | *R2* | F | 候型捅要自Ei度1 自由度2 | 畏著性 | 参数估尊（ft |
| bl | b2 |
| 践性 | .784 | 10.886 | 1 3 | .046 | -50.748 | .203 |  |
| 二次 | .945 | 17.316 | 2 2 | .055 | 176616 | -1.234 | .002 |
| S | .880 | 21.968 | 1 3 | 018 | 6770 | -1392.868 |  |

自变SL为*混度A6.*





模型摘要和参数估算值



因变員乙酔转化審10

模型捕耍

参数估算（ft

自变量为温度A6.



模型摘要和参数估算偵

C4炫燈冼择性14

1862218

84.314

■1 /V>Oll

自变宣为沮度A6.





模型摘要和参数估算偵









模型摘要和参数估算值

b2

•109.343

.420

135(.

228.711

-1711

Ml 7RC

-1801.328

自变量为温度B3.

-1539.007

自变宣为沮度B3.

卷故估芽值

口 I bl

因变員：C4炳统通择性B；





C.模型建立过程中所使用的源代码

1. ling。程序求得最优解：

sets:

parin'all/1. .6/:pl ,p2,]>3,p4;

endsets

data:

pl^-79.844,0.05542,0.03508,0.05383,-9.1856,0.33232;

p2=-48.3594,0.0517,-3.217,0.0386,2.4439,0.1807;

p3-184.268,-0.1032,1.85520,0.0287,-7.0659,-0.3518;

p4=-35.908,0.05148,1.3623,-0.0673,4.6220,0.17108; enddata

!max =(・3619.647+1.236\*xl・69.431\*x2+4.121\*x& 1276.736\*x4+11.995\*x5)\*a+b\*y 1 \*y2+(x5\*x2+x5\*x4)\*c;

hnax (-3619.647+1.236\*xl・69.431\*x2+4.121\*x3>

1276.736\*x4+11.995\*x5)\*0.7+0.6\*y 1 \*y2+(x5\*x2+x5\*x4)\*0.5;

max = ((-6314.470+3.910\*x 1 -154.727\*x2+3.933\*x3-

312.063\*x4+21.375\*x5)\*0.5+0.4\*y】\*y2+0.55\*(x5\*x2+x5\*x4))/100;

!max =(・&3】4.470+3.910\*xl-154.727、2+3.933\*xA

312063\*x4+21.375\*x5)\*0.5+0.5\*yl\*y2;

yl=pl(l)+pl ⑵ \*xl+pl ⑶ \*x2+pl(4)\*x3+pl ⑸、4+pl(6)\*x5; y2=p2(l)+p2(2)\*xl+p2(3)\*x2+p2(4)\*x3+p2(5)\*x4+p2(6)\*x5;

y3=p3(l)+p3 ⑵ \*xl+p3(3)\*x2+p3(4)\*x3+p3(5)\*x4+p3(6)\*x5;

y4=p4 ⑴ +p4(2)\*xl+p4 (3)\*x2+p4 (4)\*x3+p4(5)\*x4+p4(6)\*x5;

y2+y3+y4i=100;

@bnd(10,xl,200);

@bnd(0.5,x2,5);

@bnd(10,x3,200);

@bnd(0.3,x4,2.1);

@bnd(250,x5,350);

2.第二问偏最小二乘代码.m

clc,clear

abO= xhread('pz.xls，)； ％原始数据存放在纯文本文件pz.txt中

mu = mean( ab0);sig=std(al>0); % 求均值和标准差

rr= corrcoef(abO)%求相关系数矩阵

ab= zseore(abO); %数据标准化

a=ab(:,[l:5]);b=ab(：46:end]); %提出标准化后的自变量和因变量数据

[XL,YL,XS,YS,BETA,PCTVAR,MSE,stats] =plsregress(a,b) xw=a“XS %求自变量提出成分系教每列对应-•个成分，这里xw等于 stats, w

yw =b“YS %求因变量提出成分的系数

ncomp =input('ncomp =');

[XL2,YL2,XS2,YS2,BETA2,PCTVAR2,MSE2,stats2]

=plsregres3( a, b,wcomp)

n=size(a,2);m=s屁(b,2);%n是自変S：的个数,m是因变量的个数 beta3(l,:) = mu(n 4-l:end)-

mu(l:n)./Sig(l:n)\*BETA2([2:end]r).\*«g(«+l：end);% 原始数据回归方程的常数 项

beta3([2:n+l],:) =(l./8ig(lm))\*$ig(n +l:end).\*BETA2([2:end],:)% 计算原 始変量xl,…；xn的系数，每一列是一个回归方程

hold on;

%x3sir= ‘乙醉转化率？C4烯姪转化率乙醛选择性脂肪醇甲醛甲醛7 其他生成物乙烯选择性'"；％新坐标的值

%cll = categorical。；

bar(BETA2,；k,)% 画直方图

title(”'偏最小二乘回归系数宜方图”')

yhat = repmat(beta3( 1 ,:),[size(a, 1 ),1 ]) 4-abO(:,[l:n])\* beta3([2:end],:)%求 yl,, ym的预测值

ymax =max([yhat,abO(:,[n+1 ;end])]);

%求预测值和观测值的最大値

%下面両yl,y2,y3的预测图.并画直线y=X

figure, subplot( 1,2,1)

al=[O:ymax ⑴];

a2=[0:ymax ⑴];

bl = [O:ymax ⑵ j;

b2= [O:ymax ⑵];

plot(yhat(:,l),abO(:,n+l al,a2.,color,.1k,)

title(，乙醇转化率预测图，)

subplot (1,2,2)

p】ot(yhat(:,2),ab0(:,n+2),'O',bl,b2,'color','k')

title(，嬌炷转化率预测图')

%legend(，烯燃,2)

3. mse.ipynb

Mcells5': I

”celftyp矿："code”，

M execut ion count": 66,

metadata":-

"collapsed": true

"outpu仲：[],

source":[

"import pandas as pd、”，

” import numpy as np"n",

n import matplotlib"

”cell，type”： ”code”，

"execution.count”： 68,

"output矿：[],

“source”：[

"#自己把所有的XY (只有两列.乙博转化率烯燈选择性)“n”,

"X'self = pd.readexcel(. /X.xlsx UM) ,

” X = X'self.values4tnw,

” Yself = pd.readexcel(“”./Y.xlsx

"Y = Y'self. values%”，

”# print(3\*2)M

1»

"metadata":-

"collapsed": false,

"pycharm":-

”name”： ”#%% 读取数据，1n”

”cell'type”： ”code”，

,,execution'countM: 90.

"outputs":[],

“source”：[

"sum-2 = 0.0“n”，

= 0.0H,

*nsum3* = 0.0"

”sum'4 = 0.0“n”，

” for i in range。09): “n”，

” yl = -79.8445 + X[i| [0)\*0.0554 + X[i][l]\*(-0.0351) +

X[i][2]\*(0.0538) + X[i](3]\*(-9.1856) + X [训4] \*0.3323、”，

” y'2 = -48.3595 + X[i] [0)\*0.0517 + X[i|[l]\*(-3.2176) +

X[i][2]\*(0.0387) + X[i][3]\*(-2.4439) + X [训4] \*0.1807、”，

” syl = -80.218 + X[i][0]\*(-0.033) + X[i][l]\*(0.134) +

X[i][2]\*(0.141) + X(i][3]\*(-8.765) + X[i][4]\*0.333%”，

” sy2 = -54.118 + X[i][0]\*0.005 + X(i][l)\*(-3.173) + X|i][2]\*(0.091)

+ X[i][3]\*(2.670) + X[i][41\*0.181unw,

” delta-1 = (y l - Y[i][0])\*\*2KnB,

” delta'2 = (y'2 - Y[i][l])\*\*2un,,>

” delta'3 = (gy'l -

” delta'4 = (sy-2 - 丫阳1|)\*\*2%”，

” suml += delta」“n”，

” sum'2 += delta。、”，

” sum-3 += deltaS'V,

" sum,4 += delta'4 unM,

”print("”MSE.yl, MSE y2 In PLS :：■(8um」/109)\*\*(l/2), (蛔2/109)\*\*(1/2))侦，，

”print(“”MSE・yl, MSEy2 In SPSS:—, (sum 3/109)\*\*(l/2),

(sunf4/109)\*\*(l/2))”

1,

"metadata":-

"collapsed": false,

” pycharm":

”name”： ”#%% 求解 MSE侦'

Wtype”： *''code,*

“executioncount”： 72,

"outputs":[

"name”： ''stdouC,

"output'type”: "stream”， ”text”：[

”0“n”，

"3“n"

'source”： 0,

’metadata”： "collapsed'': false, "pycharm”：-

”name”： ”#%%“n”

“cell'type”： “code”，

"execution，count": null, outputs":[],

“source”：[], “metadata”： -

”collapsed”： false,

” pycharm”：-

”name”： ”#%%“n”

'metadata": -

” kernelspec” ：-

“display'name": “Python 3”,

"language": "python",

"name": "python3"

"languag/intb”:-

"codemirror'inode":-

"name": "ipython",

"version”： 2

"fildextension”： ”.py", "mimetype”： ”text/x-python”， “name”： “python”，

,,nbconvert exporter,T: "python" “ pygments" lexer": “ jpython2", “version”： ”2.7.6”

”nbformat”： 4,

"nbfbrmat'minor": 0