

基于回归模型和优化模型的 C4 烯烃制备问题

摘要

本文研究的是不同催化剂组合和温度对 C4 烯烃收率和 C4 烯烃选择性的影响,通过回归模型、方差分析和优化模型,得出较优的催化剂组合和温度使得 C4 烯烃收率尽可能高, C4 烯烃选择性尽可能大。

首先,对数据进行了预处理。计算出 C4 烯烃收率并记录。通过控制单一变量将 A, B 分成 12 组。

针对问题一: 为确定温度对乙醇转化率和 C4 烯烃选择性的关系。首先,通过附件 1 数据画出其随温度的散点图,通过观察散点图建立一元线性回归模型和一元二次回归模型,通过判断系数 R^2 确定使用一元二次回归模型,得出乙醇转化率随着温度的升高而升高, C4 烯烃选择性随温度的升高而升高。

对于附件 2 数据,在给定 350℃温度下,通过观察散点图,对有线性关系的变量建立一元线性回归模型,对没有线性关系数据验证其随机性。回归模型根据拟合结果可以看出,乙醇转化率随时间的增加而减少,乙烯、乙醛选择性随时间的增大而增大,碳数为 4-12 脂肪醇选择性随时间的推移呈递减趋势。而 C4 烯烃、甲基苯甲醛不遵循线性回归。通过 3σ 原则验证 C4 烯烃、甲基苯甲醛选择性均为随机变量,说明 C4 烯烃的选择性等随着时间的增大在均值附近随机震荡。

针对问题二: 由于原始数据中温度对乙醇转化率和 C4 烯烃选择性有部分缺失,因此利用问题一中的一元二次回归模型将缺失值补齐。附件 1 中影响因素有催化剂组合和温度两种,利用无重复多因素方差分析和假设检验,可知催化剂组合和温度对乙醇转化率和 C4 烯烃选择性均有显著影响。将数据预处理得到的每个分组,进行单因素方差分析,可以得出在催化剂组合中,装料方式,乙醇浓度对于乙醇转化率有显著影响。催化剂总质量,Co 负载量,乙醇浓度和是否加入 HAP 对 C4 烯烃选择性有显著影响。

针对问题三: 选择催化剂组合和温度使 C4 烯烃收率尽可能的大,对 C4 烯烃收率建立关于催化剂组合中的 4 个变量和温度的五元二次回归方程。以所建立的五元二次回归方程为目标函数,以 4 个变量和温度范围约束条件,建立优化模型,利用 Lingo 求解出最优解。在 A 种装料方式下,使 C4 烯烃收率尽可能高应选择 200mg 2wt%Co/SiO₂-200mg HAP-乙醇浓度 1.68ml/min 的催化剂组合和 400℃的温度。在 B 种装料方式下则应选择 124.3902mg 2.840285wt%Co/SiO₂-124.3902mg HAP-乙醇浓度 0.3ml/min 的催化剂组合和 400℃的温度。当温度在 350℃下时,同理求出最优解。A 种装料方式下,选择 200mg 3.176584wt%Co/SiO₂-99.11717mg HAP-乙醇浓度 2.1ml/min 的催化剂组合和 350℃的温度。B 种装料方式下,99.11717mg 2wt%Co/SiO₂-99.11717mg HAP-乙醇浓度 2.1ml/min 的催化剂组合和 350℃的温度。

针对问题四: 我们设计了 5 次实验,解决现有数据难以分析出的问题,分别为:①催化剂组合和温度的更优值 ②C4 烯烃收率在 400℃以上随温度变化 ③改变催化物研究 C4 烯烃收率随温度变化 ④制备时间的最优长度 ⑤副产物随时间变化的生成,以其为例:在 350℃时,取附件 2 中给定的催化剂组合,在时间区间[0, 280]上,增加观测频率,每隔 10min 观测一次,分别记录乙醇转化率、乙烯选择性、C4 烯烃选择性大小随时间的变化情况。

关键词: 一元回归模型 无重复多因素方差分析 单因素方差分析 优化模型

一、问题的背景与重述

1.1 问题的背景

随着人们对化工产品和医用药品需求的日益增长。如何生产这些产品成为我们需要思考的问题。其中 C4 烯烃的应用范围最为广泛，最常用的制备方法是用乙醇偶合成 C4 烯烃。随着时间的推移探究 C4 烯烃的制备问题具有重大意义。

1.2 问题的重述

针对温度和催化剂组合对乙醇转化率的影响和 C4 烯烃选择性的影响的问题，在实验室进行多次实验。通过附录 1 和附录 2 给出的数据，运用合适的方法，建立数学模型解决下列问题。

1、在固定的催化剂组合下，温度对乙醇转化率的影响和对 C4 烯烃选择性的影响。之后把温度固定在 350℃，探究在相同温度下，分析出不同时间得出的结果。

2、探究温度和催化剂组合这两种因素对 C4 烯烃选择性的影响。

3、在无关变量相同的情况下，探究如何选择温度和催化剂组合使得 C4 烯烃的收率最高。如果温度低于 350℃，此时应该如何选择温度和催化剂组合使得 C4 烯烃的收率最高。

4、增加设计出 5 次实验，并说明理由。

二、问题的分析

对于问题一，①附件 1 中给出了 21 种催化剂组合。画出每一个催化剂组合的散点图，借助图像，同时用一元回归模型和一元二次回归模型进行拟合，通过比较判断系数 R^2 得出较优的回归模型，并确定温度对乙醇的转化率和 C4 烯烃选择性的关系。②在给定的 350℃温度下，基于一元回归模型和 3σ 原则，分析乙醇转化率和各生成物选择性随时间变化的情况。

对于问题二，探究催化剂组合和温度两个因素对乙醇转化率和 C4 烯烃选择性大小的影响，采用无重复双因素方差分析结合假设检验进行求解。首先，分析温度和催化剂组合这两种因素是否对乙醇转化率和 C4 烯烃选择性产生影响。其次，通过对数据处理得到的对照实验分组进行单因素方差分析，得出在不同催化剂组合因素下的结果。

对于问题三，要使 C4 烯烃收率尽可能的高，需要综合考虑催化剂组合中各种因素和温度对 C4 烯烃收率的影响。此时，易想到利用附件 1 所给数据建立多元回归方程，然后以该方程作为目标函数，将问题 2 中得出的关系以及文献^[2]确定约束条件建立优化模型确定 C4 烯烃收率较大的值。当温度低于 350℃时，改变约束条件得到新的优化模型，确定较优的 C4 烯烃收率。

对于问题四，首先根据附件 2 可以考虑缩短实验时观测时间间隔，进一步探究产物随时间变化情况。还可以考虑问题三中得到的较优解，通过在较优解附近细微划分，设置不同的实验，在保证其他因素不变时，改变一个条件，进而找到最优解。

三、模型的假设

1、根据实际情况曲线应该为 S 型，但由于所给数据限制，因此假设我们所得到的回归方程均为部分 S 型曲线。

2、假设除了题目所给的变量因素外，其他试验因素就保持不变。

四、符号说明

符号	表示含义	单位
R^2	判断系数	/
σ	标准差	/
α	显著性水平	/
k	因素水平个数	/
n	全部观测值个数	/
w	C4 烯烃收率	/

五、数据的预处理

为了探究影响乙醇转化率和 C4 烯烃选择性的因素，首先需要对附件 1 的催化剂组合进行分类。（运用**控制变量法**，分出催化剂组合中具体因素对乙醇转化率和 C4 烯烃选择性的影响。）

- ①第一次分类按照投料方式的不同进行分类；
- ②第二次 A 中按照催化剂的总质量进行分类；B 中不进行分类；
- ③第三次在 A,B 两组中通过控制单一变量进行分类。

具体分类如下表所示：

表 1：数据分类表

第一次分类	第二次分类	第三次分类	实验考察的影响因素
A	A1~A6 (总质量： 400mg)	A1、A2、A4、A6	Co 负载量 (总质量)
		A1、A3	乙醇浓度 (1wt%)
		A2、A5	乙醇浓度 (2wt%)
	A7~A14 (总质量： 100mg)	A7、A8、A9、A12	乙醇浓度 (总质量)
		A9、A10	Co 负载量 (总质量)
		A11、A12	是否加入 HAP
		A12、A13、A14	Co/SiO ₂ 和 HAP 装料比
		A1、A12	A 组总质量
B		B1、B2、B3、B4、 B6	B 中 Co/SiO ₂ 和 HAP 的 总质量
		B1、B5	乙醇浓度 (100mg)
		B2、B7	乙醇浓度 (200mg)
		B1、A12	装料方式

其次，将所有分组中每个温度对应的收率全部算出，并记入表格，详细数据见附录 3。

六、模型的建立和求解

6.1 问题一模型的建立与求解

6.1.1 确定温度与乙醇转化率和 C4 烯烃选择性关系

根据附件 1 所给信息，分别作出乙醇转化率和 C4 烯烃选择性随温度的散点图。

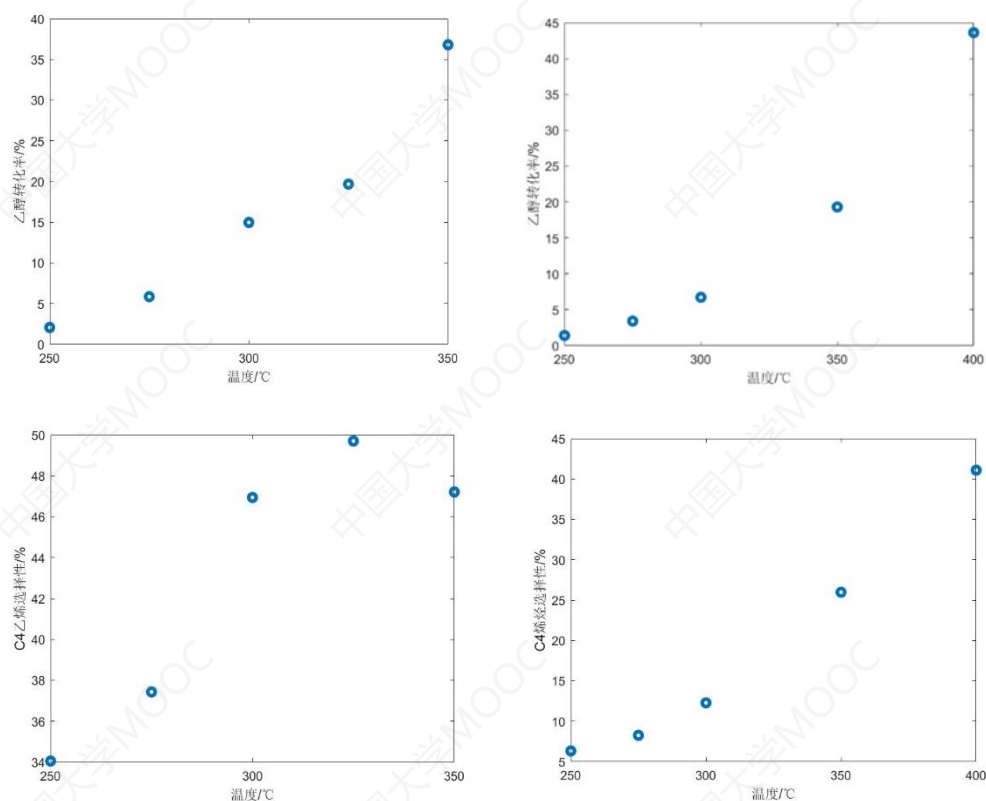


图 1: A,B 两种装料方式的散点图

通过观察上述散点图可以确定 A,B 两种方式的乙醇转化率和 C4 烯烃选择性具有一定的线性或非线性关系，为了判断具体关系，我们分别建立回归模型进行对比。

首先，研究在不同催化剂组合下，温度与乙醇转化率、温度与 C4 烯烃选择性的关系，利用附录中的数据，建立一元线性回归模型和一元二次回归模型。

$$y_i = ax_i + b(i = 1, 2) \quad (1)$$

$$y_i = b_0 + b_1x_i^2 + b_2x_i + \varepsilon_i(i = 1, 2) \quad (2)$$

其中： x_i 是温度；

y_1 是乙醇转化率；

y_2 是 C4 烯烃选择性；

b_0, b_1, b_2 是回归系数；

ε_i 是误差项。

为了将拟合效果量化，我们引入拟合优度：判断系数 R^2 ，通过 R^2 的取值来判断乙醇转化率随温度的变化和 C4 烯烃选择性随温度的变化关系的拟合效果的优劣。参考文献^[1]计算 R^2 的公式如下：

$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (3)$$

其中： \bar{y} 平均值； \hat{y} 预测值； y_i 平均值

SST 为总平方和，SSR 为回归平方和，

总平方和 (SST) = 回归平方和 (SSR) + 残差平方和 (SSE)

由统计学知识可知， $R^2 \in [0,1]$ ， R^2 越接近于 1 时，表明拟合程度越好， R^2 越接近于 0，表明拟合程度越差。

结合所给数据，利用 Matlab，拟合出每种催化组合，在不同投料方式下，温度对乙醇转化率和 C4 烯烃选择性的回归方程，方程的系数和拟合程度 R^2 如表 2、表 3 所示

表 2：A 投料方式下乙醇转化率的回归系数和拟合程度 R^2

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
A1	141.7531	0.002545	-1.19359	0.965491306	0.979748
A2	-227.393	-0.00074	1.105786	0.994990619	0.99108
A3	-134.344	-0.00033	0.647554	0.981975074	0.96632
A4	-106.943	0.000365	0.34454	0.997512098	0.995911
A5	231.943	0.003196	-1.6694	0.934547947	0.994027
A6	51.94621	0.001665	-0.58361	0.983590011	0.986032
A7	-102.598	-0.00027	0.55655	0.999392951	0.999708
...

(注：具体数据数据见附录 1)

表 3：B 投料方式下乙醇转化率的回归系数和拟合程度 R^2

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
B1	121.238	0.001884	-0.94845	0.9620627	0.998876
B2	188.2547	0.002512	-1.36477	0.929265231	0.9912
B3	106.9536	0.001388	-0.77071	0.890145194	0.991475
B4	155.2747	0.002057	-1.12904	0.899288998	0.987028
B5	185.7407	0.002502	-1.35467	0.912407316	0.99144
B6	190.0828	0.002813	-1.44471	0.939907357	0.990249
B7	228.7107	0.003278	-1.71105	0.936120534	0.996598

①从上述两张表中可以比较出一元二次回归模型的 R^2 值接近 1 的个数大于一元一次回归模型的 R^2 值接近 1 的个数，所以比较之下，一元二次模型拟合效果更好。

②得出的结果拟合程度 R^2 均都近于 1，所以在每种催化剂组合中，乙醇转化率随着温度的升高而升高。

表 4: A 种投料方式, C4 烯烃选择性的回归系数和拟合程度 R^2

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
A1	-190.783	-0.00211	1.422246	0.887083496	0.915957
A2	234.7454	0.003113	-1.64629	0.91429111	0.980284
A3	-171.076	-0.00095	0.924416	0.955413385	0.955075
A4	72.77336	0.001214	-0.56242	0.957758122	0.976519
A5	57.88371	0.001122	-0.49953	0.969582876	0.990547
A6	176.6157	0.002204	-1.23355	0.885412623	0.945403
A7	74.78099	0.001154	-0.56512	0.968183363	0.999651
...

(注: 具体数据数据见附录 1)

表 5: B 投料方式, C4 烯烃选择性的回归系数和拟合程度 R^2

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
B1	39.06776	0.000929	-0.36544	0.985781681	0.997177
B2	41.32949	0.000993	-0.40276	0.984782436	0.997736
B3	20.96986	0.000478	-0.19034	0.970967565	0.976297
B4	81.074	0.001002	-0.54866	0.895001515	0.973684
B5	32.40236	0.00065	-0.27614	0.977619294	0.998222
B6	-12.1287	0.000326	-0.02167	0.982118433	0.97093
B7	-9.85843	0.000452	-0.06002	0.994396671	0.997144

从上述两张表中可以得出的结果是所有 R^2 的拟合程度都接近于 1, 所以在每种催化剂组合中, **C4 烯烃选择性随温度的升高而升高**。

为了更加直观的看出乙醇转化率和 C4 烯烃选择性随温度的变化, 分别做出下列图像。

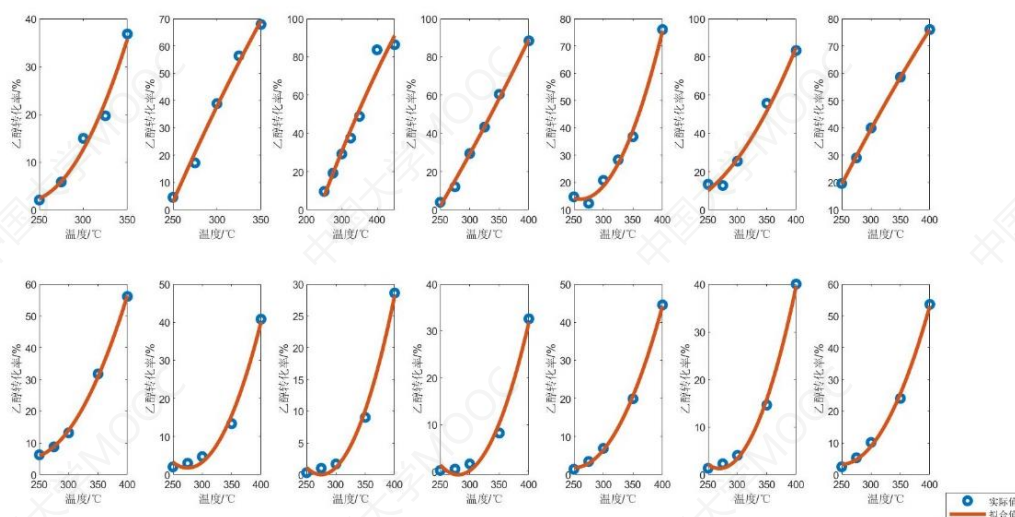


图 2: A 装料方式温度与乙醇转化率的关系

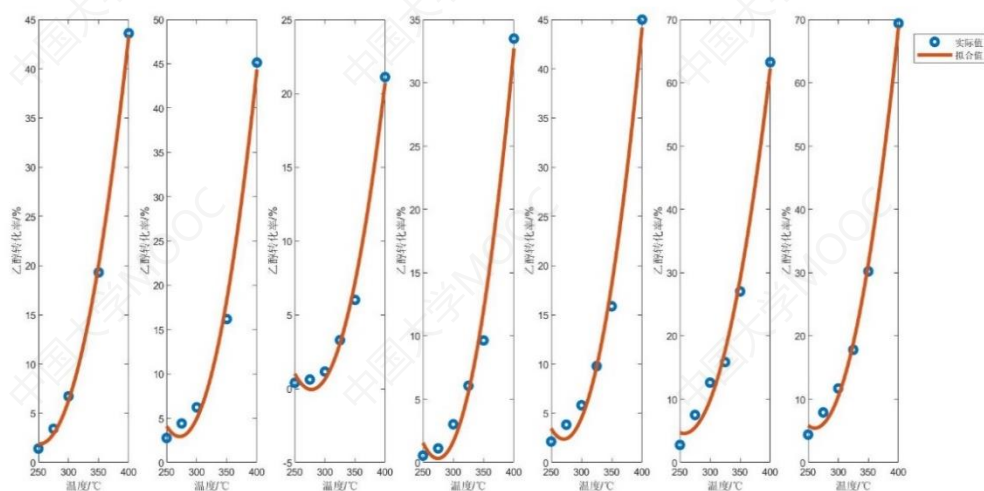


图 3: B 装料方式温度与乙醇转化率的关系

由于催化剂组合不同, 乙醇转化率随温度变化曲线有所不同, 但曲线呈现出随温度升高乙醇转化率逐渐增大的情况。

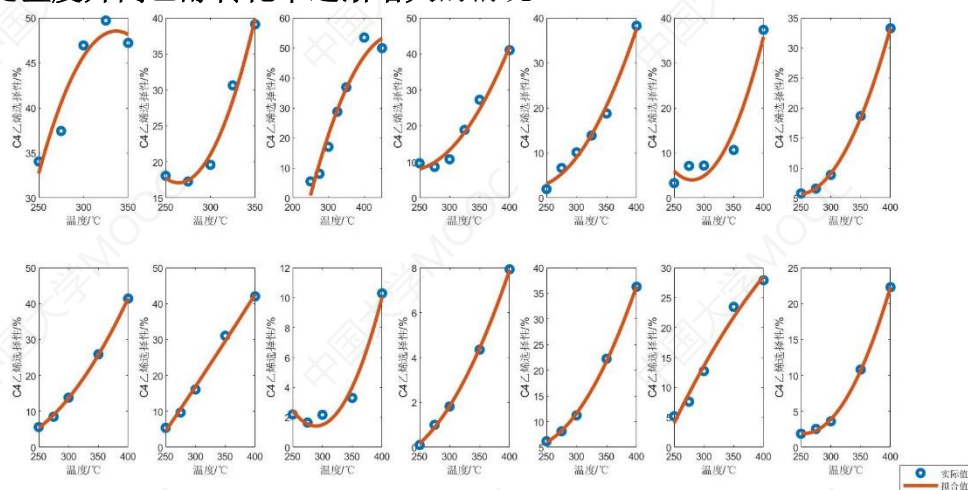


图 4: A 装料方式温度与 C4 烯烃选择性的关系

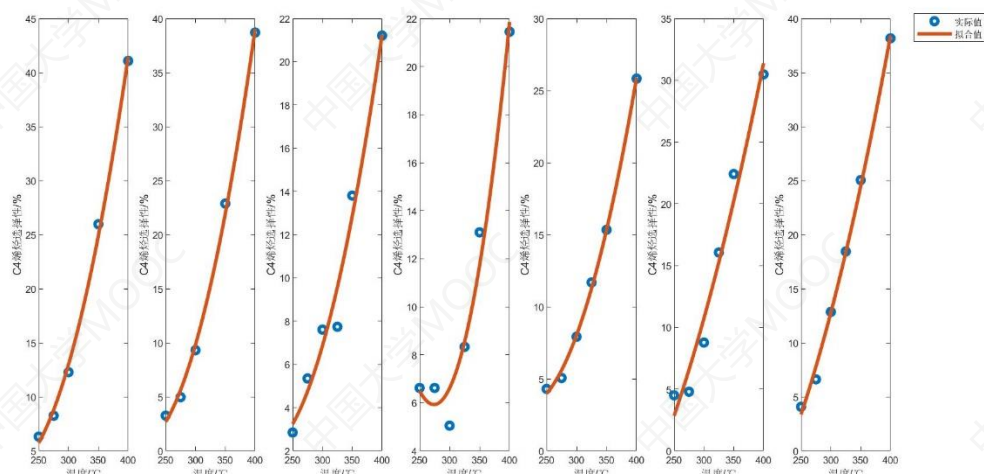


图 5: B 装料方式温度与 C4 烯烃选择性的关系

由于催化剂组合不同，C4 烯烃选择性随温度变化曲线有所不同，但曲线均呈现出随温度升高 C4 烯烃选择性逐渐增大的情况。

6.1.2 探究确定温度下各种因素随时间的变化

根据附件 2 中所给数据，并绘制散点图求相关系数，得到乙醇转化率、乙烯选择性、乙醛选择性和脂肪醇选择性分别于时间有线性关系。其余没有。

因此，在温度稳定在 350℃，催化剂组合不变的情况下，乙醇转化率、乙烯选择性、乙醛选择性和碳数为 4-12 脂肪醇选择性能够满一元线性回归模型。

通过 Matlab，建立一元线性回归模型，得到的图像如下图所示：

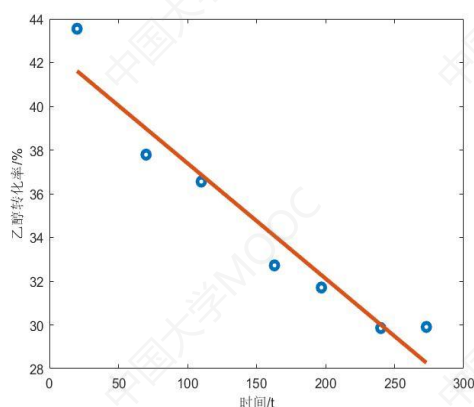


图 6：乙醇转化率时间的变化

从图 6 中可以看出乙醇转化率随时间增加而减少，并满足一元一次回归方程。根据已有数据看，在超过 240min 后维持在较为稳定的状态，这种趋势很好的反应了勒夏特列原理，表明在一定温度和催化剂组合条件下，反应物浓度越高反应速率越快，随着反应物浓度的减小，反应速率也相应减小。

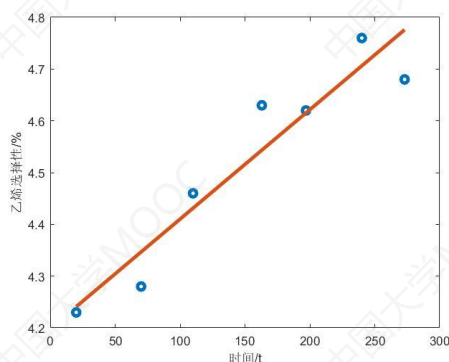


图 7：乙烯选择性随时间变化

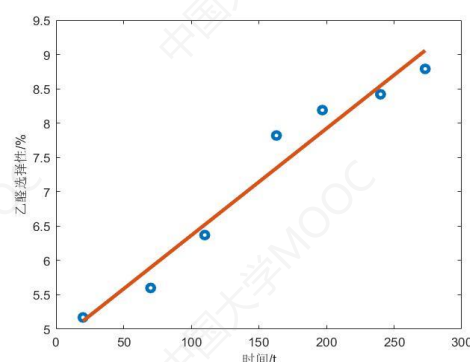


图 8：乙醛选择性随时间变化

上述两幅图中可以看出乙烯、乙醛的选择性数据也能满足一元一次回归方程，由此可以判断乙烯、乙醛的选择性随时间增加而增大，可以认为随着 C4 烯烃的制备，乙烯、乙醛作为副产物产生并积累。

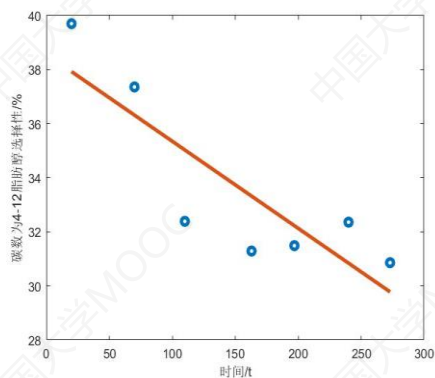


图 9：脂肪醇选择性随时间变化

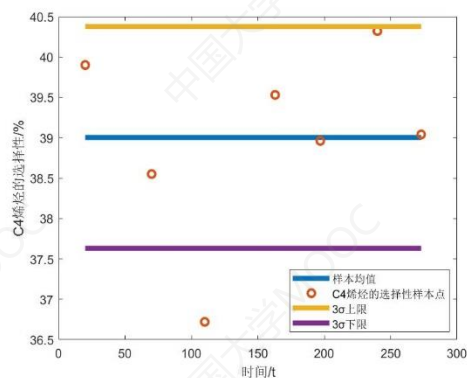


图 10：C4 烯烃选择性与时间的关系

①图 9 可以看出，碳数为 4-12 脂肪醇选择性同样满足一元一次回归方程，但是与时间呈负相关关系，该类物质可能作为副产物生成，但是随着反应的进行，此类物质又参与了 C4 烯烃、乙烯、乙醇等物质的形成过程，因此**碳数为 4-12 脂肪醇选择性随时间的推移呈递减趋势**。

②图 10 中可以看出，C4 烯烃选择性并未随时间变化出现单调递增或递减趋势，而是在区间[36.72, 40.32]内发生震荡。我们考虑其可能为随机变量，因此求出其均值和标准差，并求出其相应的 3σ 区间，大部分数据点在 3σ 区间内，仅有一点在区间外，因此可以认为 **C4 烯烃选择性随时间的分布是一个随机变量**。

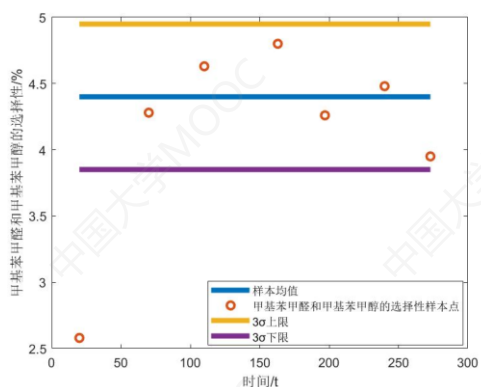


图 11：甲基苯选择性随时间变化

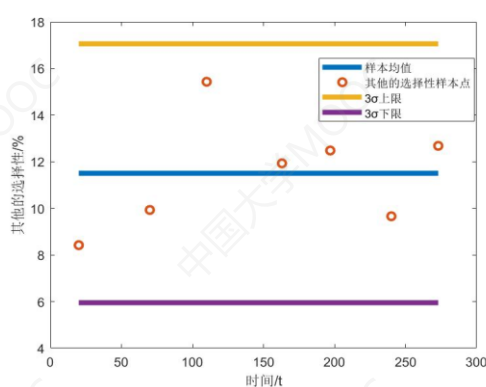


图 12：其他选择性随时间变化

我们同样做出了甲基苯甲醛和甲基苯甲醇的选择性和其他物质的选择性的 3σ 图像，结果数据点基本都能够在 3σ 区间内，因此**甲基苯甲醛和甲基苯甲醇选择性看作随机变量**。

在制备 C4 烯烃过程中，随着各类物质的产生与消耗，看作随机变量的物质的选择性震荡但在一定数值上下浮动，此类物质在反应中的产生与消耗接近平衡，能够稳定产出。

6.2 模型的建立与求解

6.2.1 模型的建立

根据题目要求，主要分析不同催化剂组合及温度对乙醇转化率以及 C4 烯烃选择性大小的影响。因此可以将影响因子分为两种因素：①不同催化剂组合；②温度。

为了分析因素影响，采用**无重复双因素方差分析结合假设检验**进行分析。催化剂组合及温度对乙醇转化率以及 C4 烯烃选择性是否有影响的假设检验，

但是由于附件 1 所给的乙醇转化率与 C4 烯烃的选择性部分数据不全，因此需要数据处理，根据第一问中建立的回归模型，将缺失数据补充，进而达到进一步分析。（补充数据见附录 3、4）

1、提出假设

对行因素： $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ vs $H_1: \mu_1, \mu_2 \dots \mu_k$ 不全相等；

对列因素： $H_0: \mu_1 = \mu_2 = \dots = \mu_r$ vs $H_1: \mu_1, \mu_2 \dots \mu_r$ 不全相等；

m, n 分别为行，列因素的个数

2、构造F检验统计量

行因素均方：

$$MSR = \frac{SSR}{k-1} \quad (4)$$

列因素均方：

$$MSC = \frac{SSC}{r-1} \quad (5)$$

随机误差项的均方

$$MSE = \frac{SSE}{(r-1)(k-1)} \quad (6)$$

检验行因素的检验统计量：

$$F_R = \frac{MSR}{MSE} \sim F(k-1, (r-1)(k-1)) \quad (7)$$

检验列因素的检验统计量：

$$F_C = \frac{MSC}{MSE} \sim F(r-1, (r-1)(k-1)) \quad (8)$$

其中：SSR行因素产生的误差平方和；

SSC列因素产生的误差平方和

SSE除去行因素与列因素外的误差平方和；

MSR行因素均方；

MSC列因素均方

MSE随机误差项均方；

k 为行因素个数

r 为列因素个数

3、做出统计决策

将根据题目所给数据计算出的检验统计量 F 与事先确定的显著性水平下 α 的临界值 F_α 进行比较，从而做出对原假设 H_0 的决策。也可以对比显著性水平 α 与 p 。这里取 $\alpha = 0.05$ 。

①若 $F > F_\alpha$ 或 $p < \alpha$ ，则拒绝原假设，认为催化剂组合或温度对于乙醇转化率或 C4 烯烃选择性有影响。

②若 $F < F_\alpha$ 或 $p > \alpha$ ，则不拒绝原假设，认为催化剂组合或温度对于乙醇转化率或 C4 烯烃选择性没有影响。

6.2.2 模型的求解

一、催化剂组合及温度对乙醇转化率的影响

提出假设：

催化剂（行）因素：

H_0 ：催化剂组合对乙醇转化率没有影响。

H_1 : 催化剂组合对乙醇转化率有影响。

温度（列）因素

H_0 : 温度对乙醇转化率没有影响。

H_1 : 温度对乙醇转化率有影响。

利用 Excle 将处理完的数据进行方差分析，利用 P 值来判断温度和催化剂组合对乙醇转化率以及 C4 烯烃选择性大小的影响。

◆ A 组中对乙醇转化率的影响

方差分析如下表：

表 6: A 组催化剂组合和温度对乙醇转化率的方差分析

差异源	SS	df	MS	F	P-value
催化剂	20811.03	13	1600.848	24.6012	2.39×10^{-22}
温度	35722.31	6	5953.718	91.49438	3.17×10^{-33}
误差	5075.612	78	65.07195		
总计	61608.95	97			

此处用 p 值做决策，由：

$$p_{A_{\text{催}}} = 2.39 \times 10^{-22} \ll \alpha = 0.05 \quad (9)$$

$$p_{A_{\text{温}}} = 3.17 \times 10^{-33} \ll \alpha = 0.05 \quad (10)$$

由 p 值远远小于置信度 α ，因此做出决策：

①对于催化剂组合影响，**拒绝原假设**，即对于 A 组认为催化剂组合对乙醇转化率**有影响**。

②对于温度影响，**拒绝原假设**，即对于 A 组认为温度对乙醇转化率**有影响**。

◆ B 组中对乙醇转化率的影响

同理：对于 B 组

催化剂（行）因素：

H_0 : 催化剂组合对乙醇转化率没有影响。

H_1 : 催化剂组合对乙醇转化率有影响。

温度（列）因素

H_0 : 温度对乙醇转化率没有影响。

H_1 : 温度对乙醇转化率有影响。

方差分析如下表：

表 7: B 组催化剂组合和温度对乙醇转化率的方差分析

差异源	SS	df	MS	F	P-value
催化剂	2179.884	6	363.314	11.85603	2.63×10^{-7}
温度	10912.77	6	1818.795	59.35277	3.4×10^{-17}
误差	1103.177	36	20.64381		
总计	14195.83	48			

此处用 p 值作决策，由：

$$p_{B_{\text{催}}} = 2.63 \times 10^{-7} \ll \alpha = 0.05 \quad (11)$$

$$p_{B_{\text{温}}} = 3.4 \times 10^{-17} \ll \alpha = 0.05 \quad (12)$$

由 p 值远远小于置信度 α ，因此做出决策：

①对于催化剂组合影响，拒绝原假设，即对于 B 组认为催化剂组合对乙醇转化率**有影响**

②对于温度影响，拒绝原假设，即对于 B 组认为温度对乙醇转化率**有影响**。

二、催化剂组合及温度对 C4 烯烃选择性大小的影响

◆ A 组中对 C4 烯烃选择性的影响

提出假设：

催化剂（行）因素：

H_0 ：催化剂组合对 C4 烯烃选择性大小没有影响。

H_1 ：催化剂组合对乙 C4 烯烃选择性大小有影响。

温度（列）因素

H_0 ：温度对 C4 烯烃选择性大小没有影响。

H_1 ：温度对 C4 烯烃选择性大小有影响。

方差分析表如下：

表 8：A 组催化剂组合和温度对 C4 烯烃选择性的方差分析

差异源	SS	Df	MS	F	P-value
催化剂	11034.05	13	848.7727	22.39366	4.01×10^{-21}
温度	9350.578	6	1558.43	41.11694	3.36×10^{-22}
误差	2956.385	78	37.90237		
总计	23341.01	97			

由 p 值作决策：

$$p_{A_{\text{催}}} = 4.01 \times 10^{-21} \ll \alpha = 0.05 \quad (13)$$

$$p_{A_{\text{温}}} = 3.36 \times 10^{-22} \ll \alpha = 0.05 \quad (14)$$

由 p 值远远小于置信度 α ，因此做出决策：

①对于催化剂组合影响，拒绝原假设，即认为催化剂组合对乙 C4 烯烃选择性大小**有影响**。

②对于温度影响，拒绝原假设，即认为温度对乙 C4 烯烃选择性大小**有影响**。

◆ B 组中对 C4 烯烃选择性的影响

提出假设：

催化剂（行）因素：

H_0 ：催化剂组合对 C4 烯烃选择性大小没有影响

H_1 ：催化剂组合对乙 C4 烯烃选择性大小有影响

温度（列）因素

H_0 ：温度对 C4 烯烃选择性大小没有影响

H_1 ：温度对 C4 烯烃选择性大小有影响

方差分析表如下所示：

表 9: B 组催化剂组合和温度对 C4 烯烃选择性的方差分析

差异源	SS	df	MS	F	P-value
行	680.9959	6	113.4993	10.14451	1.46×10^{-6}
列	4175.329	6	695.8881	62.19814	1.59×10^{-17}
误差	402.7769	36	11.18825		
总计	5259.101	48			

由 p 值作决策:

$$p_{A_{\text{催}}} = 1.46 \times 10^{-6} \ll \alpha = 0.05 \quad (15)$$

$$p_{A_{\text{温}}} = 1.59 \times 10^{-17} \ll \alpha = 0.05 \quad (16)$$

由 p 值远远小于置信度 α , 因此做出决策:

①对于催化剂组合影响, 拒绝原假设, 即认为催化剂组合对乙 C4 烯烃选择性大小有影响。

②对于温度影响, 拒绝原假设, 即认为温度对乙 C4 烯烃选择性大小有影响。

◆ 对每个分组进行单因素方差分析

根据上述的方差分析结果, 可以得出温度和催化剂组合对于乙醇转化率以及 C4 烯烃选择性分别具有显著影响, 对于温度的研究, 第一问中已经体现;

对于不同催化剂组合对乙醇转化率以及 C4 烯烃选择性的影响, 根据对数据预处理中实验用控制变量法分组的结果:

对 13 组对照试验组分别探究对应催化剂组合中的因素对于乙醇转化率以及 C4 烯烃选择性的影响, 对此我们分别针对 13 个实验组进行单因素方差分析, 得到了每组因素差别是否显著, 结果如下:

表 10: 分类及单因素方差 p 值表

第一次分类	第二次分类	第三次分类	实验考察的影响因素	乙醇转化率 p-value	C4 烯烃选择性 p-value
A	A1~A6 (总质量: 400mg)	A1、A2、A4、A6	Co 负载量 (总质量)	0.5411	0.0030
		A1、A3	乙醇浓度 (1wt%)	0.3762	0.0490
		A2、A5	乙醇浓度 (2wt%)	0.2891	0.0628
	A7~A14 (总质量: 100mg)	A7、A8、A9、A12	乙醇浓度 (总质量)	0.0048	0.6872
		A9、A10	Co 负载量 (总质量)	0.4761	0.0035
		A11、A12	是否加入 HAP	0.3357	0.0050
		A12、A13、A14	Co/SiO ₂ 和 HAP 装料比	0.6984	0.1610

		A1、A12	A 组总质量	0.3190	0.0002
B		B1、B2、 B3、B4、 B6	B 中 Co/SiO ₂ 和 HAP 的总 质量	0.2555	0.3265
		B1、B5	乙醇浓度 (100mg)	0.9335	0.2039
		B2、B7	乙醇浓度 (200mg)	0.34277	0.8073
		B1、A12	装料方式	0.0024	0.7292

我们设置显著性水平 α 为 0.05，因此 $p - value$ 小于 0.05 时则差异性显著，否则该因素影响不显著。

从上述表格可以得出在不同催化剂组合因素下可得出的结果：

①对于乙醇转换率有显著影响的有装料方式，A 装料方式 100mg 下乙醇浓度。

②对于 C4 烯烃选择性有显著影响的有，催化剂总质量，Co 负载量，400mg1wt% 下乙醇浓度，是否加入 HAP。

综上所述：

- 温度对乙醇浓度转化率和 C4 烯烃选择性均有影响。
- 催化剂组合对乙醇转换率有影响的有①装料方式，②A 装料方式 100mg 下乙醇浓度。
- 催化剂组合对 C4 烯烃选择性有影响的有，①催化剂总质量，②Co 负载量，③400mg1wt%下乙醇浓度，④是否加入 HAP。

6.3 模型的建立与求解

6.3.1 优化模型的建立

每种催化剂组合都存在 4 个维度的数值，即 Co/SiO₂ 质量，Co 负载量，HAP 质量和乙醇浓度。在挑选催化剂组合和温度时，则共需要挑选出包含温度在内的 5 个维度的具体值。

所给数据中，各维度的取值都是离散的，而实验间对比所得的结论也并不能说明在五个维度下催化剂组合和温度的优劣。

采用拟合的方法使未给出的因素取值尽可能连贯，从而方便分析因素对收率的影响。而又考虑到五个维度可能同时影响 C4 烯烃收率变化，且除各因素单独产生作用外，还可能会出现两两因素相互影响的情况。因此我们将所有已知数据拟合成一元二次函数：

$$w = \beta_1 x_1^2 + \beta_2 x_2^2 + \beta_3 x_3^2 + \beta_4 x_4^2 + \beta_5 x_5^2 + \cdots + \beta_{21} x_4 x_5 + \beta_0 + \varepsilon \quad (17)$$

其中： $x_i (i = 1, 2 \dots 5)$ 分别为 Co/SiO₂ 质量，Co 负载量，HAP 质量，乙醇浓度，

温度；

$\beta_1 - \beta_{21}$ 是系数

w 是 C4 烯烃的收率；

β_0 是常数项；

ε 是随机变量。

利用所给数据进行逐步回归分析，得到回归方程为：

A 装料方式的函数为：

$$w_A = 76.855 - 0.224x_3 - 0.521x_5 + 0.000887x_5^2 - 0.00279x_2x_5 + 0.00837x_3x_4 + 0.000823x_3x_5 \quad (18)$$

其中 A 装料方式的判断系数 $R^2 = 0.863506$ 。

B 装料方式的函数为：

$$w_B = 0.806 - 0.00627x_5 - 0.0000369x_1^2 + 0.0000124x_5^2 - 0.00140x_1x_4 + 0.0000240x_1x_5 \quad (19)$$

其中 B 装料方式的判断系数 $R^2 = 0.935021$ 。

两种投料方式的判断系数都接近 1，所以这 5 种因素对 C4 烯烃收率有显著的影响。

结合文献^[2]，可以得知要想使目标函数值最大，催化剂各指标均要有一定的区间和限制，具体的限制条件如下所示：

温度(HAP 活性限制)： $x_5 \in (0,450)$

Co/SiO₂ 质量与 HAP 质量： $x_1 = x_3$

Co 负载量： $x_2 \in [2,5]$

乙醇浓度： $x_4 \in [0.3,2.1]$

由目标函数和确定的约束条件，建立优化模型：

目标函数：

$$\max w$$

约束条件：

$$\begin{cases} x_1 = x_3 \\ 0 \leq x_1 \leq 200 \\ 2 \leq x_2 \leq 5 \\ 0.3 \leq x_4 \leq 2.1 \\ 0 < x_5 \leq 450 \end{cases}$$

6.3.2 优化模型的求解

通过 Lingo，求解优化模型，得出如下表格：

表 11：A,B 两种装料方式的最优解

	Co/SiO ₂ 质量	Co 负载量	HAP 质量	乙醇浓度	温度	C4 烯烃收率 (%)
A	200	2	200	2.1	400	32.6984
B	124.3902	2.840285	124.3902	0.3	400	0.859351

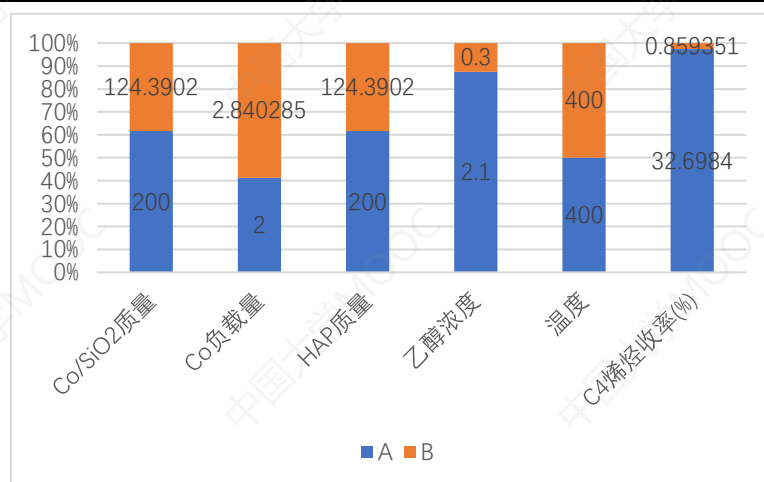


图 13：350℃以上 A、B 两种催化剂组合与温度对比

综上所述：

①实验条件相同，在 A 种装料方式下，为了使 C4 烯烃收率尽可能的高，应选择 **200mg 2wt%Co/SiO2-200mg HAP-乙醇浓度 1.68ml/min** 的催化剂组合和 **400℃** 的温度。

②实验条件相同，在 B 种装料方式下，为了使 C4 烯烃收率尽可能的高，则应选择 **124.3902mg 2.840285wt%Co/SiO2-124.3902mg HAP-乙醇浓度 0.3ml/min** 的催化剂组合和 **400℃** 的温度。

6.3.3 给定温度下使 C4 烯烃收率较高的选择

同上述原理相同，仅仅改变温度范围。首先，求出温度在 350℃ 以下的五元二次回归方程

A 装料方式的函数为：

$$w_A = 75.908 - 0.188x_3 - 7.524x_4 - 0.494x_5 - 1.405x_2^2 + 0.000888x_5^2 + 0.0108x_1x_2 + 6.372x_2x_4 - 0.0189x_2x_5 + 0.000717x_3x_5 \quad (20)$$

其中 A 装料方式的判断系数 $R^2 = 0.850213$ 。

B 装料方式的函数为：

$$w_B = 1.700 - 0.00458x_1 - 0.0121x_5 - 0.00002492x_1^2 + 0.000022x_5^2 + 0.0000272x_1x_5 \quad (21)$$

其中 B 装料方式的判断系数的 $R^2 = 0.914698$ 。

两种投料方式的判断系数都接近 1，所以催化剂组合的各个因素和温度对 C4 烯烃收率有显著的影响。

与上述不同之处是，温度的约束条件变成 $x_5 \in (0,350)$ ，则此时的优化模型变成：

目标函数：

$$\max w$$

约束条件：

$$\begin{cases} x_1 = x_3 \\ 0 \leq x_1 \leq 200 \\ 2 \leq x_2 \leq 5 \\ 0.3 \leq x_4 \leq 2.1 \\ 0 < x_5 \leq 350 \end{cases}$$

通过 Lingo 求解，得出表格：

表 12：350℃ 以下 A,B 两种装料方式的最优解

	Co/SiO2 质量	Co 负载量	HAP 质量	乙醇浓度	温度	C4 烯烃收率 (%)
A	200	3.176584	99.11717	2.1	350	22.75501
B	99.11717	2	99.11717	2.1	350	0.404819

为了更加直观的观察出 A, B 两种装料方式的对比，将表格中数据呈现在图像中，如下图所示：

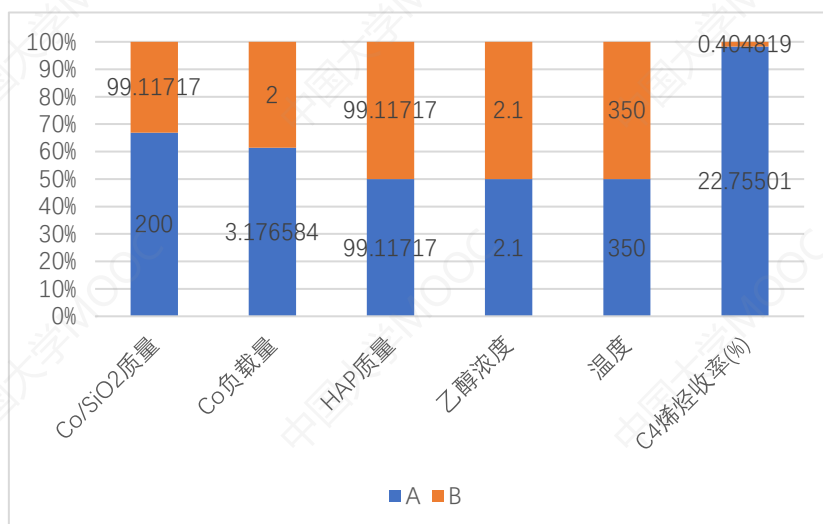


图 14: 350℃以下 A、B 两种催化剂组合与温度对比

综上所述:

①实验条件相同, 在 A 种装料方式下, 为了使 C4 相同收率尽可能的高, 则应选择 **200mg 3.176584wt%Co/SiO2-99.11717mg HAP-乙醇浓度 2.1ml/min** 的催化剂组合和 **350℃** 的温度。

②实验条件相同, 在 B 种装料方式下, 为了使 C4 烯烃收率尽可能的高, 则应选择 **99.11717mg 2wt%Co/SiO2-99.11717mg HAP-乙醇浓度 2.1ml/min** 的催化剂组合和 **350℃** 的温度。

6.4 设计实验

6.4.1 找出催化剂组合和温度的更优情况

实验目的及理由:

问题三中, 我们用来求出最优情况的目标函数是通过拟合得出的, 该函数不可避免地会和实际有误差, 但在一定程度上该函数能够较好符合变化趋势。因此我们可以猜测, 实际的最优情况与我们得到的会较为接近。为此我们设计了下述实验, 来证明我们的猜测。由于催化剂组成有四个维度, 我们仅给出其中一个维度得到最优值的实验方法, 其他维度最优值得出方法一致。

实验设计:

在 Co/SiO2 负载量 2wt%、乙醇浓度 2.1ml/min 和温度 400℃的条件下, 保持 Co/SiO2 和 HAP 装料比为 1: 1, 使 Co/SiO2 和 HAP 的装料质量在区间[190, 210]上以 1 为步长逐渐增加, 测得每个点对应的 C4 烯烃收率, 找到装料质量的更优值。

6.4.2 研究温度在 400℃以上, C4 烯烃收率随温度变化情况

实验目的及理由:

通过对于附件 1 中 A3 组实验数据的观察, 我们发现在温度小于 400℃时, 收率随温度增大而增大, 并在 400℃时达到最大值 44.72%, 但是在 450℃时收率开始下降值为 43.11%。由于其他所给实验并未研究 450℃的情况, 因此我们有理由怀疑所有催化剂组合都可能在 400℃时使 C4 烯烃收率达到最大。我们做实验来验证这一想法。

实验设计:

在附件 1 中给出的 20 种催化剂组合中 (除 A3 组) 随机抽选 5 种, 观察此 5

种催化剂组合在温度区间[350, 500]上 C4 烯烃收率的变化情况。

6.4.3 用其他物质做催化剂代替 Co/SiO₂-HAP, 观察 C4 烯烃收率随温度变化情况

实验目的及理由:

通过查询文献资料^[2], 知除 Co/SiO₂-HAP 可做催化剂外,沸石、金属氧化物也可以作为乙醇制备 C4 烯烃反应的催化物, 因此我们可以尝试做改变催化剂类型的实验观察在相同温度区间中不同催化物对 C4 烯烃收率的影响。

实验设计:

我们分别以下述三种组合做催化剂:

- ① 200mg 1wt%Co/SiO₂-200mg HAP-乙醇浓度 1.68ml/min
- ② 200mg 1wt%沸石-200mg HAP-乙醇浓度 1.68ml/min
- ③ 200mg 1wt%某种金属氧化物-200mg HAP-乙醇浓度 1.68ml/min

除催化剂外其余条件均相同, 观察在温度区间[250, 400]上, C4 烯烃随温度变化情况。

6.4.4 找出制备时间的最优长度

实验目的及理由:

根据附件 2 中数据, 我们能够发现 C4 烯烃收率随制备时间延长而降低。因此研究在不同的时间范围内, 收率下降速率大小, 可以帮助我们给出制备 C4 烯烃时间的最优长度。我们可以做如下实验, 得到所需数据。

实验设计:

在附件 1 中给出的 21 种催化剂组合中随机抽取 5 种, 固定温度为 350°C, 在时间区间[0, 280]上观察 5 种情况的 C4 烯烃收率下降情况。

6.4.5 副产物随时间变化的生成情况

实验目的及理由:

在乙醇偶合制备 C4 烯烃的过程中, 还会产生许多副产物。这些副产物有的能在生活中的不同领域发挥重要作用。因此在保证 C4 烯烃产量稳定时, 研究副产物中各物质比例很有价值。从附件 2 中所给数据, 我们能够发现随着时间的, 有部分副产物能够呈正相关或负相关变化, 但是, 观测间隔不均匀且跨度大。因此, 我们增加实验, 研究副产物选择性大小随时间的变化情况。

实验设计:

在 350°C 时, 取附件 2 中给定的催化剂组合, 在时间区间[0, 280]上, 增加观测频率, 每隔 10min 观测一次, 分别记录乙醇转化率、乙烯选择性、C4 烯烃选择性大小随时间的变化情况。

七、模型的评价与改进

7.1 模型的评价

7.1.1 模型的优点

①所给实验组合根据实验探究的因素进行分类, 并对各因素分别进行方差分析, 严谨科学。

②在同种催化剂条件下, 对于第一题分别建立一元一次和一元二次模型, 两模型对比, 选择拟合优者, 考虑全面, 拟合效果更显著。

③催化剂组合及温度与 C4 烯烃收率的关系采用逐步回归的方法, 并结合文献^[2], 用化学反应机理对回归函数进行约束, 避免了仅从函数本身出发而产生的问

题。

④从反应中的副产物，时间等方面增加实验，考虑到反应中的经济因素，考虑周全。

7.1.2 模型的缺点

①仅从所给数据出发，未考虑实际中模型的真实性和完整性，比如 C4 烯烃选择性与温度实际上应该是 S 型曲线，本文基于局部仅作出一元二次模型。

②在建立催化剂与温度这 5 种因素对 C4 烯烃收率的回归模型时，在局部出现了负值，拟合效果并不理想。

7.2 模型的改进

①对于温度与 C4 烯烃选择性的模型，可以收集更多数据，进而能够拟合出 Compertz 曲线，进而更加接近真实值。

②对于问题三中五元函数回归模型，实际效果并不是很好，应该采取更多更加合适的模型，比如采用 1-0 分类模型，通过各个变量系数表现对 C4 烯烃收率的影响。

八、 模型的应用与推广

本文主要运用了统计学知识，运用了一元回归模型，多元回归模型，多因素方差分析，单因素方差分析等方法。随着社会的发展，多元回归模型被广泛推广在经济，环境，生态，等各个领域的现实生活中。

九、 参考文献

- [1]贾俊平，何晓群，金勇进，《统计学》（第 7 版）[M].中国人民大学出版社.2018。
- [2]吕绍沛，乙醇偶合制备丁醇及 C4 烯烃[D].大连.理工大学.2015。
- [3]姜启源，谢金星，叶俊，《数学模型》[M].高等教育出版社.2003。
- [4]张俊旺，乙醇制混合烃的研究[D].兰州，兰州大学，2012。

附录：

附录 1：温度与乙醇转化率关系的拟合程度

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
A1	141.7531	0.002545	-1.19359	0.965491306	0.979748
A2	-227.393	-0.00074	1.105786	0.994990619	0.99108
A3	-134.344	-0.00033	0.647554	0.981975074	0.96632
A4	-106.943	0.000365	0.34454	0.997512098	0.995911
A5	231.943	0.003196	-1.6694	0.934547947	0.994027
A6	51.94621	0.001665	-0.58361	0.983590011	0.986032
A7	-102.598	-0.00027	0.55655	0.999392951	0.999708
A8	96.86278	0.001751	-0.80148	0.977146393	0.999039
A9	186.2871	0.002442	-1.34255	0.920969156	0.990262
A10	135.6512	0.001796	-0.98715	0.922864351	0.994009
A11	177.9177	0.002272	-1.27408	0.90305177	0.987127
A12	121.5259	0.001903	-0.95436	0.963057942	0.99908
A13	164.3022	0.002246	-1.21	0.936280636	0.996508
A14	137.9912	0.002178	-1.0834	0.963901057	0.997162
B1	121.238	0.001884	-0.94845	0.9620627	0.998876
B2	188.2547	0.002512	-1.36477	0.929265231	0.9912
B3	106.9536	0.001388	-0.77071	0.890145194	0.991475
B4	155.2747	0.002057	-1.12904	0.899288998	0.987028
B5	185.7407	0.002502	-1.35467	0.912407316	0.99144
B6	190.0828	0.002813	-1.44471	0.939907357	0.990249
B7	228.7107	0.003278	-1.71105	0.936120534	0.996598

附录 2：温度与 C4 选择性关系的拟合程度

催化剂组合	常数项系数	二次项系数	一次项系数	一次拟合 R^2	二次拟合 R^2
A1	-190.783	-0.00211	1.422246	0.887083496	0.915957
A2	234.7454	0.003113	-1.64629	0.91429111	0.980284
A3	-171.076	-0.00095	0.924416	0.955413385	0.955075
A4	72.77336	0.001214	-0.56242	0.957758122	0.976519
A5	57.88371	0.001122	-0.49953	0.969582876	0.990547
A6	176.6157	0.002204	-1.23355	0.885412623	0.945403
A7	74.78099	0.001154	-0.56512	0.968183363	0.999651
A8	19.81904	0.000747	-0.24463	0.99157993	0.999526
A9	-53.4302	5.49E-05	0.218015	0.997399552	0.994887
A10	59.71128	0.000699	-0.40359	0.861457779	0.978469
A11	5.585768	0.000183	-0.06745	0.989048175	0.999444
A12	45.32475	0.000902	-0.38293	0.983164824	0.999258

A13	-64.3329	-0.00028	0.342454	0.988318725	0.981657
A14	64.86287	0.000969	-0.49395	0.959150411	0.999463
B1	39.06776	0.000929	-0.36544	0.985781681	0.997177
B2	41.32949	0.000993	-0.40276	0.984782436	0.997736
B3	20.96986	0.000478	-0.19034	0.970967565	0.976297
B4	81.074	0.001002	-0.54866	0.895001515	0.973684
B5	32.40236	0.00065	-0.27614	0.977619294	0.998222
B6	-12.1287	0.000326	-0.02167	0.982118433	0.97093
B7	-9.85843	0.000452	-0.06002	0.994396671	0.997144

附录 3：所有分组每个温度对应的收率

温度 组别	250.00	275.00	300.00	325.00	350.00	375.00	400.00
A1	0.7039	2.1903	7.0264	9.7816	17.3738	23.6251	28.592
A2	0.8318	2.9714	7.6288	17.2643	26.5408	46.0645	71.9809
A3	0.5318	1.5467	4.9755	10.7927	18.0333	26.5888	44.7281
A4	0.3867	1.0428	3.1599	8.1842	16.4777	23.9643	36.2778
A5	0.2898	0.8262	2.1058	3.9292	6.9022	15.6726	29.0624
A6	0.4419	0.9075	1.8294	3.2382	5.9379	16.1007	31.1137
A7	1.1302	1.9053	3.5346	6.4025	10.924	16.9548	25.2791
A8	0.3528	0.7465	1.8213	4.0991	8.2125	14.0949	23.2434
A9	0.1109	0.2906	0.7614	1.8387	4.1571	9.4686	17.1511
A10	0.007	0.0168	0.0366	0.1058	0.2976	1.1909	2.9424
A11	0.0002	0.0052	0.0293	0.1155	0.357	1.1884	2.5847
A12	0.089	0.2817	0.7758	1.9969	4.4325	8.929	16.1662
A13	0.0699	0.1757	0.5181	1.4704	3.4343	6.6672	11.1772
A14	0.0471	0.136	0.3665	1.0643	2.6027	6.0399	11.9562
B1	0.089	0.282	0.8257	2.2052	5.0146	9.9509	17.909
B2	0.0901	0.2188	0.58	1.5407	3.7036	8.8988	17.4673
B3	0.0113	0.0329	0.0874	0.2543	0.83	2.1947	4.4753
B4	0.036	0.0737	0.1523	0.5044	1.2633	3.4099	7.1835
B5	0.0912	0.194	0.4603	1.1438	2.4334	5.9794	11.6188
B6	0.1246	0.3597	1.1067	2.547	6.0576	11.2364	19.2772
B7	0.1795	0.523	1.5046	3.2841	7.5651	14.9762	26.49

附件:

Matlab R2018b

1、求解每种催化剂下，乙醇转化率与温度一元二次回归方程；C4 烯烃的选择性与温度的一元二次回归方程以及各自的拟合优度和图象。

```
clear
clc
num_A=[5 5 7 6 6 5 5 5 5 5 5 5];%A 组每次实验下温度数据个数
num_B=[5 5 6 6 6 6 6];%B 组每次实验下温度数据个数
%% 计算相关性及回归模型
data_A=xlsread('A_B 组 催化剂组合 温度 乙醇转化率和 C4 烯烃选择性.xlsx','A','A2:C75');
data_B=xlsread('A_B 组 催化剂组合 温度 乙醇转化率和 C4 烯烃选择性.xlsx','B','A2:C41');
Temp_A=data_A(:,1);%A 的温度序列
Temp_B=data_B(:,1);%B 的温度序列
C2H6Orate_A=data_A(:,2);%A 的乙醇转化率序列
C2H6Orate_B=data_B(:,2);%B 的乙醇转化率序列
C2H4ele_A=data_A(:,3);%A 的乙烯选择性序列
C2H4ele_B=data_B(:,3);%A 的乙烯选择性序列
%% 温度与乙醇转化率
%% A
%分别计算每次实验的温度与乙醇转化率的相关性
R1=[];%存放每次实验的温度与乙醇转化率的相关系数（假设）
k=0;
for i=1:length(num_A)
    j=num_A(i);
    temp=Temp_A(k+1:k+j);
    C2H6O=C2H6Orate_A(k+1:k+j);
    r=corrcoef(temp,C2H6O);
    r=r(1,2);
    R1=[R1 r];
    k=k+j;
end
%全部大与 0.90，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型
SSE_A1=[];%存放每次实验的温度与乙醇转化率的每个模型的总平方和
SSA_A1=[];%存放每次实验的温度与乙醇转化率的每个模型的回归平方和
B_A_1=[];%存放每次实验的温度与乙醇转化率的系数
t_p_A_1=[];%存放每次实验的温度与乙醇转化率的系数的 t 检验和 P 值
k=0;
figure(1)
% title('温度与乙醇转化率一元二次回归')
m=1;
for i=1:length(num_A)
```

```

j=num_A(i);
temp=Temp_A(k+1:k+j);
C2H6O=C2H6Orate_A(k+1:k+j);
% X=[ones(j,1) temp];
% [b, bint,r,rint,stats]=regress(C2H6O,X);%线性模型
% X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)]];
X=[ones(j,1) temp.^2 temp];
[b, bint,r,rint,stats]=regress(C2H6O,X);%非线性模型
X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)].^2'
[temp(1):(temp(end)-temp(1))/999:temp(end)]];
subplot(2,7,m)
Y=X*b;
plot(temp,C2H6O,'o','LineWidth',3)
hold on
plot(temp(1):(temp(end)-temp(1))/999:temp(end),Y,'-','LineWidth',3)
xlabel('温度/°C')
ylabel('乙醇转化率/%')
x=[ones(j,1) temp.^2 temp];
Y=x*b;
ssa=(Y-C2H6O).^2;
SSA_A1=[SSA_A1 sum(ssa)];
C2H6Orate_Abar=mean(C2H6O);
sse=sum((C2H6O-C2H6Orate_Abar).^2);
SSE_A1=[SSE_A1 sse];
B_A_1=[B_A_1 b];
t_p_A_1=[t_p_A_1;stats];
k=k+j;
m=m+1;
end
legend('实际值','拟合值');
%% B
R2=[];
k=0;
for i=1:length(num_B)
j=num_B(i);
temp=Temp_B(k+1:k+j);
C2H6O=C2H6Orate_B(k+1:k+j);
r=corrcoef(temp,C2H6O);
r=r(1,2);
R2=[R2 r];
k=k+j;
end
%全部大与 0.90，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型

```

```

%全部大与 0.90，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型
figure(2)
m=1;
SSE_B1=[];
SSA_B1=[];
B_B_1=[];%存放每次实验的温度与乙醇转化率的系数
t_p_B_1=[];%存放每次实验的温度与乙醇转化率的系数的 t 检验和 P 值
k=0;
% figure(2)
for i=1:length(num_B)
    j=num_B(i);
    temp=Temp_B(k+1:k+j);
    C2H6O=C2H6Orate_B(k+1:k+j);
%    X=[ones(j,1) temp];
%    [b, bint,r,rint,stats]=regress(C2H6O,X);%线性模型
%    X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)]'];
    X=[ones(j,1) temp.^2 temp];
    [b, bint,r,rint,stats]=regress(C2H6O,X);%非线性模型
    X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)].^2
[temp(1):(temp(end)-temp(1))/999:temp(end)]'];
    subplot(1,7,m)
    Y=X*b;
    plot(temp,C2H6O,'o','LineWidth',3)
    hold on
    plot(temp(1):(temp(end)-temp(1))/999:temp(end),Y,'-','LineWidth',3)
    xlabel('温度/°C')
    ylabel('乙醇转化率/%')
    x=[ones(j,1) temp.^2 temp];
    Y=x*b;
    ssa=(Y-C2H6O).^2;
    SSA_B1=[SSA_B1 sum(ssa)];
    C2H6Orate_Bbar=mean(C2H6O);
    sse=sum((C2H6O-C2H6Orate_Bbar).^2);
    SSE_B1=[SSE_B1 sse];
    B_B_1=[B_B_1 b];
    t_p_B_1=[t_p_B_1;stats];
    k=k+j;
    m=m+1;
end
legend('实际值','拟合值');

%% 温度与 C4 乙烯选择性
%% A

```

```

%分别计算每次实验的温度与 C4 乙烯选择性的相关性
R3=[];%存放每次实验的温度与 C4 乙烯选择性的相关系数（假设）
k=0;
n=1;
for i=1:length(num_A)
    j=num_A(i);
    temp=Temp_A(k+1:k+j);
    C2H4=C2H4ele_A(k+1:k+j);
    r=corrcoef(temp,C2H4);
    r=r(1,2);
    R3=[R3 r];
    k=k+j;
end
%全部大与 0.80，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型
SSE_A2=[];
SSA_A2=[];
B_A_2=[];%存放每次实验的温度与 C4 乙烯选择性的系数
t_p_A_2=[];%存放每次实验的温度与 C4 乙烯选择性的系数的 t 检验和 P 值
k=0;
n=1;
figure(3)
for i=1:length(num_A)
    j=num_A(i);
    temp=Temp_A(k+1:k+j);
    C2H4=C2H4ele_A(k+1:k+j);
    %    X=[ones(j,1) temp];
    %    [b, bint,r,rint,stats]=regress(C2H4,X);%线性模型
    %    X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)]'];
    X=[ones(j,1) temp.^2 temp];
    [b, bint,r,rint,stats]=regress(C2H4,X);%非线性模型
    X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)].^2
[temp(1):(temp(end)-temp(1))/999:temp(end)]'];
    subplot(2,7,n)
    Y=X*b;
    plot(temp,C2H4,'o','LineWidth',3)
    hold on
    plot(temp(1):(temp(end)-temp(1))/999:temp(end),Y,'-','LineWidth',3)
    xlabel('温度/°C')
    ylabel('C4 乙烯选择性/%')
    x=[ones(j,1) temp.^2 temp];
    Y=x*b;
    ssa=(Y-C2H4).^2;
    SSA_A2=[SSA_A2 sum(ssa)];

```

```

C2H6Orate_Abar=mean(C2H4);
sse=sum((C2H4-C2H6Orate_Abar).^2);
SSE_A2=[SSE_A2 sse];
B_A_2=[B_A_2 b];
t_p_A_2=[t_p_A_2;stats];
k=k+j;
n=n+1;
end
legend('实际值','拟合值');
%% B
R4=[];
k=0;
for i=1:length(num_B)
    j=num_B(i);
    temp=Temp_B(k+1:k+j);
    C2H4=C2H4ele_B(k+1:k+j);
    r=corrcoef(temp,C2H4);
    r=r(1,2);
    R4=[R4 r];
    k=k+j;
end
%全部大与 0.90，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型
%全部大与 0.90，因此线性相关性显著，可以用一元线性回归模型
%一元线性回归模型
SSE_B2=[];
SSA_B2=[];
B_B_2=[];%存放每次实验的温度与乙醇转化率的系数
t_p_B_2=[];%存放每次实验的温度与乙醇转化率的系数的 t 检验和 P 值
k=0;
n=1;
figure(4)
for i=1:length(num_B)
    j=num_B(i);
    temp=Temp_B(k+1:k+j);
    C2H4=C2H4ele_B(k+1:k+j);
    % X=[ones(j,1) temp];
    % [b, bint,r,rint,stats]=regress(C2H4,X);%线性模型
    % X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)]]';
    X=[ones(j,1) temp.^2 temp];
    [b, bint,r,rint,stats]=regress(C2H4,X);%非线性模型
    X=[ones(1000,1) [temp(1):(temp(end)-temp(1))/999:temp(end)].^2
[temp(1):(temp(end)-temp(1))/999:temp(end)]]';
    subplot(1,7,n)

```



```

Y=X*b;
plot(temp,C2H4,'o','LineWidth',3)
hold on
plot(temp(1):(temp(end)-temp(1))/999:temp(end),Y,'-','LineWidth',3)
xlabel('温度/°C')
ylabel('C4 烯烃选择性/%')
x=[ones(j,1) temp.^2 temp];
Y=x*b;
ssa=(Y-C2H4).^2;
SSA_B2=[SSA_B2 sum(ssa)];
C2H6Orate_Bbar=mean(C2H4);
sse=sum((C2H4-C2H6Orate_Bbar).^2);
SSE_B2=[SSE_B2 sse];
B_B_2=[B_B_2 b];
t_p_B_2=[t_p_B_2;stats];
k=k+j;
n=n+1;
end
legend('实际值','拟合值');
%% 拟合优度
R_A1=1-SSA_A1./SSE_A1;
R_B1=1-SSA_B1./SSE_B1;
R_A2=1-SSA_A2./SSE_A2;
R_B2=1-SSA_B2./SSE_B2;
2、求解附件二中各种生成物，反应物与反应时间的相关性，求解具有显著相关关系的生成物反应物与时间的回归方程，回归方程图象，不具有相关性变量的散点图，均值，σ图
clear
clc
%% 计算相关性及回归模型
data_350=xlsread('附件 2','A4:H10');
Time=data_350(:,1);
%% 线性相关性分析
r=corrcoef(data_350);%
disp(r)
%根据线性相关系数 1 分别与 2 3 5 6 线性相关系数大与 0.8 因此，对其建立线性模型
C2H6Orate=data_350(:,2);%乙醇转化率序列
C2H4ele=data_350(:,3);%乙烯选择性序列
C4ele=data_350(:,4);%C4 烯烃选择性序列
C2H4Oele=data_350(:,5);%乙醛选择性序列
C412=data_350(:,6);%碳数为 4-12 脂肪醇
CC7H9=data_350(:,7);%甲基苯甲醛和甲基苯甲醇
qita=data_350(:,8);%其他

```

```

figure(1)
%% 乙醇转化率序列
X=[ones(7,1) Time];
[b, bint,r,rint,stats]=regress(C2H6Orate,X);%线性模型
X=[ones(1000,1) [Time(1):(Time(end)-Time(1))/999:Time(end)]];
Y=X*b;
subplot(2,2,1)
plot(Time,C2H6Orate,'o','LineWidth',3)
hold on
plot(Time(1):(Time(end)-Time(1))/999:Time(end),Y,'-','LineWidth',3)
xlabel('时间/t')
ylabel('乙醇转化率/%')
%% 乙烯选择性序列
X=[ones(7,1) Time];
[b, bint,r,rint,stats]=regress(C2H4ele,X);%线性模型
X=[ones(1000,1) [Time(1):(Time(end)-Time(1))/999:Time(end)]];
Y=X*b;
subplot(2,2,2)
plot(Time,C2H4ele,'o','LineWidth',3)
hold on
plot(Time(1):(Time(end)-Time(1))/999:Time(end),Y,'-','LineWidth',3)
xlabel('时间/t')
ylabel('乙烯选择性/%')
%% 乙醛选择性序列
X=[ones(7,1) Time];
[b, bint,r,rint,stats]=regress(C2H4Oele,X);%线性模型
X=[ones(1000,1) [Time(1):(Time(end)-Time(1))/999:Time(end)]];
Y=X*b;
subplot(2,2,3)
plot(Time,C2H4Oele,'o','LineWidth',3)
hold on
plot(Time(1):(Time(end)-Time(1))/999:Time(end),Y,'-','LineWidth',3)
xlabel('时间/t')
ylabel('乙醛选择性/%')
%% 碳数为 4-12 脂肪醇选择性序列
X=[ones(7,1) Time];
[b, bint,r,rint,stats]=regress(C412,X);%线性模型
X=[ones(1000,1) [Time(1):(Time(end)-Time(1))/999:Time(end)]];
Y=X*b;
subplot(2,2,4)
plot(Time,C412,'o','LineWidth',3)
hold on
plot(Time(1):(Time(end)-Time(1))/999:Time(end),Y,'-','LineWidth',3)
xlabel('时间/t')

```

```

ylabel('碳数为 4-12 脂肪醇选择性/%')
figure(2)
%% C4 烯烃的选择性
C4elebar=mean(C4ele);
C4elevar=var(C4ele);
plot(Time,ones(1,7).*C4elebar,'LineWidth',4)
hold on
plot(Time,C4ele,'o','LineWidth',2)
hold on
plot(Time,C4elebar+ones(1,7).*C4elevar,'LineWidth',4)
hold on
xlabel('时间/t')
ylabel('C4 烯烃的选择性/%')
plot(Time,C4elebar-ones(1,7).*C4elevar,'LineWidth',4)
legend('样本均值','C4 烯烃的选择性样本点','3 $\sigma$  上限','3 $\sigma$  下限')
%% 甲基苯甲醛和甲基苯甲醇
figure(3)
CC7H9bar=mean(CC7H9(2:7));
CC7H9var=var(CC7H9);
plot(Time,ones(1,7).*CC7H9bar,'LineWidth',4)
hold on
plot(Time,CC7H9,'o','LineWidth',2)
hold on
plot(Time,CC7H9bar+ones(1,7).*CC7H9var,'LineWidth',4)
hold on
xlabel('时间/t')
ylabel('甲基苯甲醛和甲基苯甲醇的选择性/%')
plot(Time,CC7H9bar-ones(1,7).*CC7H9var,'LineWidth',4)
legend('样本均值','甲基苯甲醛和甲基苯甲醇的选择性样本点','3 $\sigma$  上限','3 $\sigma$  下限')
%% 其他
figure(4)
qitabar=mean(qita);
qitavar=var(qita);
plot(Time,ones(1,7).*qitabar,'LineWidth',4)
hold on
plot(Time,qita,'o','LineWidth',2)
hold on
plot(Time,qitabar+ones(1,7).*qitavar,'LineWidth',4)
hold on
xlabel('时间/t')
ylabel('其他的选择性/%')
plot(Time,qitabar-ones(1,7).*qitavar,'LineWidth',4)
legend('样本均值','其他的选择性样本点','3 $\sigma$  上限','3 $\sigma$  下限')

```

3、逐步回归求解问题三五元二次回归方程并画图比较

```
clear
clc
%% 逐步回归求 5 元二次回归方程
data=xlsread('T_3_A 数据整理.xlsx','B2:I75');
x1=data(1:end,1);
x2=data(1:end,2);
x3=data(1:end,3);
x4=data(1:end,4);
T=data(1:end,5);
Y=data(1:end,8);
X=[ones(74,1),x1,x2,x3,x4,T,x1.^2,x2.^2,x3.^2,x4.^2,T.^2,x1.*x2,x1.*x3,x1.*x4,x1.*
T,x2.*x3,x2.*x4,x2.*T,x3.*x4,x3.*T,x4.*T];
stepwise(X,Y)
%调整后 R^2Adj R^2=0.86188
%所得方程
y=74.455+0.165.*x1-0.373.*x3-0.514*T+0.000890.*T.^2-0.000537.*x1.*T-
0.00270.*x2.*T+0.00820.*x3.*x4+0.00130.*x3.*T;
%对比图
plot(Y)
hold on
plot(y)
xlabel('第几组')
ylabel('C4 烯烃收率/%')
clear
clc
XX=[];
YY=[];
for x1=10:50:200
    x3=x1;
    for x2=0.5:0.1:5
        for x4=0.1:0.1:2
            for T=200:25:450
                y=74.455+0.165.*x1-0.373.*x3-0.514*T+0.000890.*T.^2-
0.000537.*x1.*T-0.00270.*x2.*T+0.00820.*x3.*x4+0.00130.*x3.*T;
                XX=[XX;x1 x2 x3 x4 T];
                YY=[YY y];
            end
        end
    end
end
plot(YY)
[num,i]=max(YY);
XL=XX(i,:)
```

```

XX=[];
YY=[];
for x1=150:1:170
    x3=x1;
    for x2=0.4:0.01:0.6
        for x4=1.5:0.01:2
            for T=370:10:450
                y=74.455+0.165.*x1-0.373.*x3-0.514*T+0.000890.*T.^2-
0.000537.*x1.*T-0.00270.*x2.*T+0.00820.*x3.*x4+0.00130.*x3.*T;
                XX=[XX;x1 x2 x3 x4 T];
                YY=[YY y];
            end
        end
    end
end
end
plot(YY)
[num,i]=max(YY);
XL=XX(i,:);

```