Pisco Med Publishing

Original Research Article

Study on Catalyst Combination for Ethanol Coupling to C4 Olefins Based on Multivariate Ridge Regression Model Fused With Radial Basis Function

Kehan Liu[Corresponding Author], Yichi Cao, Jin Shi

School of Statistics and Mathematics, Shanghai Lixin School of Accounting and Finance, Shanghai, 200000, China

Abstract: The effects of different catalyst combinations and temperatures on ethanol conversion and C4 olefin selectivity were discussed.Firstly,the multivariate ridge regression model based on L2 penalty term is introduced,and the best estimation of ridge regression is obtained by using the least square method.However,ridge regression always regards this problem as a linear problem. In order to solve this problem,the radial basis function based on Gaussian kernel is introduced as the basic variable for model optimization.Finally,the model is solved and the results are obtained.
Keywords:Ethanol;C4 olefin;Catalyst Combination;Multiple Regression Analysis

1. Introduction

C4 olefins are widely used in the production of chemical products and medicine.Ethanol is the raw material for the production of C4 olefins.In the preparation process,the catalyst combination and temperature will have an impact on the selectivity and yield of C4 olefins.Therefore,it is of great significance and value to explore the process conditions for the preparation of C4 olefins by ethanol catalytic coupling through the combination design of catalysts.

In this paper,the effects of different catalysts and temperatures on ethanol conversion and C4 olefin selectivity were studied. The catalyst combination is only a combination of variables.Based on this,we can turn the problem into investigating the effects of CO loading,Co/SiO mass,Co/SiO2 to HAP mass ratio,ethanol pouring speed and temperature on ethanol conversion and C4 olefin selectivity.Therefore,the multivariate ridge regression model based on L2 penalty term is introduced,and the best estimation of ridge regression is obtained by the least square method.However,ridge regression always regards this problem as a linear problem.In order to solve this problem,the radial basis function based on Gaussian kernel is introduced as the basic variable for model optimization.

2. Model Establishment and Solution

2.1 Multiple Ridge Regression Model Based on L2 Penalty Term

We need to consider the effects of catalyst combination and temperature on ethanol conversion and C4 olefin selectivity.In fact,the catalyst combination is only a combination of some variables,so we can turn the problem into investigating the effects of CO loading,Co/SiO mass,Co/SiO2 to HAP mass ratio,ethanol pouring speed and temperature on ethanol conversion and C4 olefin selectivity.

1) Nonlinear
2) Dimension is too high,linear regression and simple regression have no credibility

Because we are facing high-dimensional data,in order to increase the reliability of the model and reduce the variance of the model,we decided to use the ridge regression method to model the problem.Both methods are consistent in modeling ethanol conversion and C4 olefin selectivity.Next,ethanol conversion is used as the main body of modeling.

Let x_{ij} be the value of the jth variable in the i-th sample,and y_i be the ethanol conversion rate in the i-th sample.In ridge regression,it shrinks the parameter value in the model by adding L2 penalty term,so as to achieve the purpose of parameter selection.The parameters selected by the model need to minimize the following formula.

$$\hat{\beta}_{ridge} = \arg\min_{\beta}\left(\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{m}x_{ij}\beta_j)^2+\lambda\sum_{j=1}^{p}\beta_j{}^2\right)$$

Where, $\lambda \geq 0$, we call it the adjustment coefficient,which is used to control the degree of contraction.

When $\sum_{j=1}^{p}\beta_j{}^2 = 0$ $\lambda \to \infty$,using the least square method,we can get the best estimation of ridge regression as follows:

doi: 10.18282/l-e.v10i4.2514

$$\hat{\beta}_{ridge} = \left(X'X + \lambda I \right)^{-1} X'Y$$

2.2 Radial Basis Function Model Based on Gaussian Kernel

Because the L2 penalty term is added to the ridge regression model,it can effectively control the model parameters and improve the reliability of the model.However,the ridge regression regards this problem as a linear problem.In order to solve this problem,we decide to introduce the radial basis function based on Gaussian kernel as the basic variable.

The essence of radial basis function is to transform the original nonlinear space into a linear space,and then regress on the new linear space.Let the original sample space be:

$$\mathrm{X} \subset \boldsymbol{R}^n, x = \left(x_1, x_2, ..., x_n \right)^T \in \mathrm{X},$$

The new sample space is:

$$\mathrm{Z} \subset \boldsymbol{R}^n, z = \left(z_1, z_2, ..., z_n \right)^T \in \mathrm{Z}$$

Consider a mapping from X to Z

$$\phi(x) : \mathrm{X} \to \mathrm{Z}$$

Make the function $K\left(x, z\right)$ satisfy the condition for all $x, z \in \mathrm{X}$

$$K\left(x, z\right) = \phi(x) \cdot \phi(z)$$

We do not explicitly define mapping function $\phi\left(x\right)$,but directly define function $K\left(x, z\right)$,that is,our radial basis function.

In this problem,we choose Gaussian kernel function as radial basis function,i.e

$$K\left(x, z\right) = \exp\left(-\frac{\left\| x - z \right\|^2}{2\sigma^2} \right)$$

2.3 Multivariate Ridge Regression Model Integrating Radial Basis Function

After introducing the radial basis function,the original nonlinear space will be transformed into a linear space,and the independent variable of the regression model is no longer the original x_j ,but $\phi(x_j)$,i.e

$$y = \beta_0 + \sum_{j=1}^{m} \beta_j \phi(x_j) + \varepsilon$$

We define $B = [\beta_0, \beta_1, ..., \beta_m]$, $F = [1, \phi(x_1), \phi(x_2), ..., \phi(x_m)]$.Then

$$y = B^T F$$

According to the objective function of multiple ridge regression,we have

$$\begin{aligned} B^T &= \arg\min_{B^T} \left\| B^T F - X^T \right\|^2 + \lambda \left\| B \right\|^2 \\ &= X^T F^T (FF^T + \lambda I)^{-1} \\ &= X^T G \end{aligned}$$

Among them,G completely depends on f

2.4 Model Solution

The algorithm steps for solving the radial basis function model are as follows:

Input:samples $\boldsymbol{x}$ requiring nonlinear transformation

Output:sample $\boldsymbol{x}'$ after nonlinear transformation

Step 1:map the original data point x_j of sample $\boldsymbol{x}_i$ to a new feature vector x_j' .

$$x_j' = \exp\left(-\frac{\left\| x_j - z_j \right\|^2}{2\sigma^2} \right)$$

Step 2:calculate the point multiplication of $\boldsymbol{x}_i'$ and $\boldsymbol{x}_{i+1}'$

$$\boldsymbol{x}_i' \cdot \boldsymbol{x}_{i+1}' = \boldsymbol{K}\left(\boldsymbol{x}_i, \boldsymbol{x}_{i+1} \right)$$

Step 3:stop if all samples are traversed;Otherwise, $i = i + 1$,skip to step 1

When the multiple ridge regression model is solved only when $\sum_{j=1}^{p} \beta_j^2 = 0$, $\lambda \to \infty$,the best estimation of ridge regression can be obtained by using the least square method.

$$\hat{\beta}_{ridge} = \left(X'X + \lambda I \right)^{-1} X'Y$$

At this time,the best estimate $\hat{\beta}_{ridge}$ is the model parameter with radial basis function $\phi(x_j)$ as independent variable.

3. Model Evaluation

We design a multiple ridge regression model based on L2 penalty term to deal with the high-dimensional data faced in this problem,so as to increase the reliability of the model.Considering that ridge regression is suitable for linear problems,the radial basis function based on Gaussian kernel is further introduced as the basic variable to transform the original nonlinear space into a new linear space,and then carry out regression analysis to make the data analysis results more universal and persuasive.However,due to the small amount of data,it may be accidental and uncertain,which makes the conclusion contrary to the facts.

References:

[1] LV Shaopei(2018).Preparation of butanol and C by ethanol coupling_4 olefins[D].Dalian University of technology.

[2] Zhao Dongbo(2017).Research on multicollinearity in linear regression model[D].Bohai University.

[3] Long Teng,Guo Xiaosong,Peng Lei,Liu Li(2014).Optimization strategy of dynamic radial basis function proxy model based on trust region[J]. Journal of mechanical engineering,50(07):184-190

[4] Pan Haiyang,Yang Yu,Zheng Jinde,Cheng Junsheng(2017).Pattern recognition method of variable prediction model based on radial basis function[J].Journal of Aeronautical dynamics,2017,32(02):500-506