

模式识别在甲醛生产优化控制中的应用

邵惠鹤 赵维刚

(华东化工学院自动化研究所)

摘 要

本文试用模式识别方法研究甲醛生产过程,并通过按PCA方法选择稳定子集,建立了主分量 y_1 和 y_2 的平面而获得了该系统的控制模型,求得了在满足产品质量和产量约束条件下使原料单耗最低的优化操作条件.由实际运行表明,本文提供的最优条件可显著地降低原料单耗,提高产品质量,延长催化剂寿命.

关键词: 甲醛,计算机控制,最优化,模式识别.

模式识别是用一组表示被研究对象的特征变量构成模式空间,按“物以类聚”的观点分析数据结构,划分出具有特定属性模式类的空间聚集区域,并辨认每一个模式的类别.由计算机按模式识别原则处理大量信息,选择决定分类的特征变量,并作出最优决策.

本文讨论模式识别在甲醇氧化脱氢反应制取甲醛的工业生产装置优化控制中的应用.从生产操作数据,找出影响反应的主要因素,并获得使原料甲醇单耗为最低的优化控制区.

一、特征变量的选择和样本数据的预处理

甲醛生产过程是以甲醇、空气和水蒸汽为原料,在装有电解银催化剂的绝热反应器中,进行氧化、脱氢反应而生成37% (重量比)浓度的甲醛水溶液.通过系统分析^[1]可知,在催化剂及其装填状况一定的条件下,影响反应的主要因素有反应温度 T [°C]、氧醇分子比 E 、蒸汽添加量 G_s 、空速 ν ,过热器出口温度 T_1 [°C]、吸收塔软水添加量 G_w .而尾气中CO和CO₂含量,产品中甲醇含量 A_m 和酸度等是反映化学反应的转化率 x 和选择性 β 等的指标.由于本装置已实现微机实时检测,因此上述数据都可用微机在线测得或计算得到.

根据反应机理和实际操作经验,在众多的因素中,氧醇比 E 、反应温度 T 、产品中醇含量 A_m 、尾气中CO₂%含量等四个变量作为描述单耗的特征变量.按单耗 s 高低,把已知样本集分成两类, $s \leq 445$ [kg/T]为第一类,共十个样本点; $s > 446$ [kg/T]为第二类,共二十一个样本点.全部数据都由微机打印制表得到.

设 x_1 为产品中醇含量, x_2 为尾气中CO₂含量, x_3 为反应温度, x_4 为氧醇比.

1. 数据标准化处理

为了避免过分突出某些数值较大(或较小)的变量的作用,可用标准化方法处理,使各特征变量所处的地位大致相同. 现按总体标准化方法,即将样本数据 x'_{ij} ($i = 12 \dots 31$, $j = 1, 2, 3, 4$) 按下列公式标准化. 经标准化后的数据为 x_{ij} .

$$x_{ij} = \frac{x'_{ij} - \bar{x}_j}{s_j}; \quad \bar{x}_j = \frac{1}{31} \sum_{i=1}^{31} x'_{ij}; \quad s_j = \frac{1}{31} \sqrt{\sum_{i=1}^{31} (x'_{ij} - \bar{x}_j)^2}. \quad (1.1)$$

标准化处理后的数据集均值为零,并消除了物理量纲的影响.

2. 加权处理

变量的加权是增强那些对分类比较有效的变量的作用,使类的分布相对明确,增加模式空间的可分性. 设各特征变量相应的权因子为 w_1, w_2, w_3 和 w_4 , 则可按下式求得 w_j , $j = 1, 2, 3, 4$.

$$w_j = \frac{\sum'_{i,k} (x_{ij} - x_{kj})^2}{\sum''_{i,k} (x_{ij} - x_{kj})^2}. \quad (1.2)$$

式中 \sum' 表示属于不同类的第 i 和第 k 样本点求方差和; \sum'' 表示属于相同类的第 i 和第 k 个样本点求方差和. 由计算机求得

$$w_1 = 1.2700, \quad w_2 = 0.8212, \quad w_3 = 2.0785, \quad w_4 = 10.8272.$$

可见 $w_4 > w_3 > w_1 > w_2$.

二、模式分析

用主成份分析 (PCA) 法进行模式分析^[2], 可得相应的特征值和特征向量, 如表 1 所示.

表 1 特征值和对应的特征向量

序 号	特 征 值	系 数 (特 征 向 量)			
		x_1	x_2	x_3	x_4
1	$\lambda_3 = 11.1185$	0.93685	0.1642	-0.30147	0.06678
2	$\lambda_4 = 2.8233$	-0.24759	0.9336	-0.25886	0.00933
3	$\lambda_2 = 20.255$	0.24482	0.31805	0.90881	-0.11389
4	$\lambda_1 = 894.568$	-0.03266	0.016685	0.12717	0.99121

由表 1 可知,目前的分类由第 4 个特征值所控制,它构成 96% 以上的变异. 第一主分量为

$$y_1 = -0.03266x_1 + 0.016685x_2 + 0.12717x_3 + 0.99121x_4. \quad (2.1)$$

这个分量好像是一个简单的物量,与氧醇比和温度有关,和产品中醇含量 x_1 , 尾气中 CO_2 含量 x_2 无关,称为“氧醇比温度”.

第二主分量为

$$y_2 = 0.24482x_1 + 0.31805x_2 + 0.90881x_3 - 0.11389x_4. \quad (2.2)$$

它构成的变异很小, 只有总变异的 2.2% 左右. 此主分量可看成反应温度 x_3 所控制.

以此两个主分量为坐标轴构成一个模式空间, 此时主成分数 $L = 2$. 计算出分类信息图如图 1 所示.

由图 1 可见, 以原料甲醇单耗最低为优化目标, 获得了显著的分类效果, 并清楚地显示了原料单耗高低的操作区.

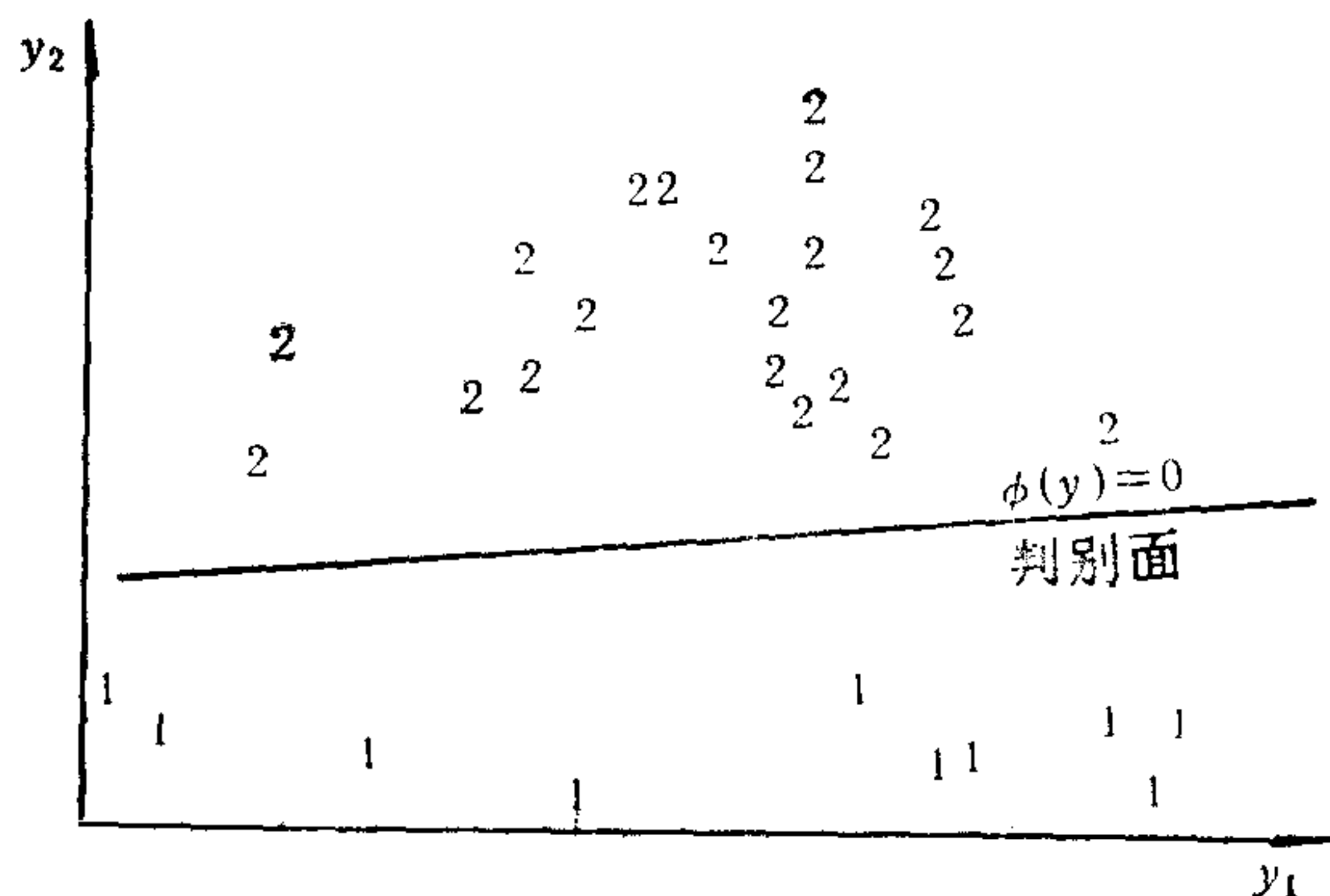


图 1 甲醛生产过程的模式分析

三、模式分类器的建立

由模式分析结果可知, 以主分量 y_1 和 y_2 构成模式空间, 以单耗大小为指标, 可将两个不同类的区域分割得比较明显. 若以对反应影响较大的氧醇比 x_4 和反应温度 x_3 构成模式空间时, 则两个区域分割得较差. 因此, 仍采用主分量 y_1 和 y_2 构成模式空间

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

并由此建立两类线性分类器. 现采用类重心距离分类法分别求出两个不同类重心坐标. 第一类重心坐标为

$$c_{11} = -14.514, \quad c_{12} = 0.0628, \quad \text{即 } \mathbf{c}_1^T = \begin{bmatrix} c_{11} \\ c_{12} \end{bmatrix}.$$

第二类重心坐标为

$$c_{21} = 6.911, \quad c_{22} = -0.033, \quad \text{即 } \mathbf{c}_2^T = \begin{bmatrix} c_{21} \\ c_{22} \end{bmatrix}.$$

采用线性判别函数 $\phi(\mathbf{y})$, 则有

$$\begin{aligned} \phi(\mathbf{y}) &= (\mathbf{c}_1 - \mathbf{c}_2) \cdot \mathbf{y} - \frac{1}{2} (|\mathbf{c}_1|^2 - |\mathbf{c}_2|^2) \\ &= -21.425y_1 + 0.0958y_2 - 81.499. \end{aligned} \quad (3.1)$$

令 $\phi(\mathbf{y}) = 0$ 就是一个判别面, 则有

$$21.425y_1 - 0.0958y_2 + 81.499 = 0, \quad (3.2)$$

在 y_1, y_2 平面上是一条分割两类不同属性的直线。对于一个未知类别的样本点 x , 可先经预处理及加权后化为对应的 y 。然后, 用判别函数 $\phi(y)$ 判别其属性。当 $\phi(y) > 0$ 时, 则 y 属于第一类; 当 $\phi(y) < 0$ 时, 则 y 属于第二类。

当用于优化操作和控制时, 是寻找单耗较低的第一类区域 y_1^* 和 y_2^* 。然后再求得相应的 $x_3^* = T$ 和 $x_4^* = E$ 的适宜值, 从而使生产运行在低单耗区域内, 可获得较好的经济效益。根据现场运行可知, 按所求得的优化条件运行, 可使单耗降低 7[kg/T], 达到该厂历史上最好水平, 并使催化剂的寿命, 由原来 30 天延长到 50 天, 按年产 7.5 万吨甲醛计算, 每年可获经济收益 47.25 万元。

参 考 文 献

- [1] 竺建敏、邵惠鹤、蒋慰孙, 甲醛生产过程的数学模型与优化, 华东化工学院学报, 第十二卷 (1986 年) 增刊, 1—13。
[2] 肯德尔, M., 多元分析, 科学出版社 (1983 年)。

THE APPLICATION OF PATTERN RECOGNITION TO THE OPTIMUM CONTROL OF FORMALDEHYDE PRODUCTION

SHAO HUIHE ZHAO WEIGANG

(East China University of Chemical Technology)

ABSTRACT

Pattern recognition is applied to the research of formaldehyde production process. By using PCA method, a stable subset is selected and a plane with the principal components y_1 and y_2 is established, thus a control model of the process and an optimum operation condition of making the lowest raw material consumptions and satisfying the constrained conditions of the product quality and output are obtained. Practical operations show that the optimum conditions presented in this paper can greatly reduce the raw material consumptions, increase product quality and prolong catalyst life.

Key words —— Formaldehyde; computer control; optimization; pattern recognition.