

УДК 681.513

ПРИМЕНЕНИЕ МЕТОДА ИДЕНТИФИКАЦИИ ДИСПЕРСИОННОЙ МОДЕЛИ ДЛЯ КОНТРОЛЯ КАЧЕСТВА ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА ВЫДЕРЖКИ ВИНМАТЕРИАЛА

Филатова Д.В.

Решается проблема оценки разброса отклонения химических показателей виноматериала относительно установленных ГОСТом с целью осуществления корректировки технологического процесса выдержки и дальнейшего прогнозирования поведения готовой продукции путем использования метода идентификации дисперсионной модели.

Марочное вино изготавливается путем выдержки виноматериалов в бутах. Срок выдержки виноматериалов зависит от особенностей технологического процесса производства. Для контроля протекания технологического процесса выдержки один раз в год измеряют химические показатели и проводят органолептические тесты для каждого отдельного бута. Значения химических показателей сравниваются со значениями, установленными ГОСТом, и в случае отклонения проводится корректировка. Целью корректировки является такое изменение параметров, которое дает наименьший разброс средних значений химических показателей каждого бута внутри партии. В этом случае встает проблема оценки этого разброса и дальнейшего прогнозирования поведения готовой продукции.

Представим технологический процесс изготовления вина в виде некоторого объекта управления. В качестве компонент вектора входных переменных $\vec{x}$ будем использовать следующие показатели: время выдержки (x_1), усредненные по партии значения объемного содержания спирта (x_2) и массовой концентрации сахаров (x_3). Под выходной переменной $v(\vec{x})$ будем понимать значение разброса средних значений химических показателей по всей партии вина. Значения $v(\vec{x})$ могут быть измерены косвенно путем преобразования вектора состояний. Вектор состояния $\vec{y}(\vec{x})$ описывается такими показателями, как плотность, массовая концентрация титруемых кислот, летучих кислот, сернистой кислоты, железа, меди, свинца и экстракта. Количество реализаций переменных $\vec{y}(\vec{x})$ определяется количеством бутов, используемых для изготовления вина, а количество реализаций переменных $\vec{x}$ зависит от требований технологического процесса производства определенной марки вина.

Пусть количество бутов N , а технологический процесс требует K лет выдержки, тогда всю информацию о состоянии технологического процесса для i -ой реализации вектора $\vec{x}$ ($i=1,2,K$) можно представить в виде следующей панели [1]

$$\mathbf{Y}(\bar{\mathbf{x}}_i) = \begin{pmatrix} y_{11}(\bar{\mathbf{x}}_i) & y_{21}(\bar{\mathbf{x}}_i) & \dots & y_{r1}(\bar{\mathbf{x}}_i) \\ y_{12}(\bar{\mathbf{x}}_i) & y_{22}(\bar{\mathbf{x}}_i) & \dots & y_{r2}(\bar{\mathbf{x}}_i) \\ \dots & \dots & \dots & \dots \\ y_{1N}(\bar{\mathbf{x}}_i) & y_{2N}(\bar{\mathbf{x}}_i) & \dots & y_{rN}(\bar{\mathbf{x}}_i) \end{pmatrix}, \quad (1)$$

где r – количество переменных вектора состояний.

Основными характеристиками панели (1) считаются вектор средних значений переменных $\bar{y}_l(\bar{\mathbf{x}}_i)$ ($l = 1, 2, \dots, r$, $i = 1, 2, \dots, K$) и разброс наблюдений относительно центра распределения, который сопоставляется с максимальным собственным числом ковариационной матрицы

$$\mathbf{D}\{\mathbf{Y}(\bar{\mathbf{x}}_i)\} = \frac{1}{N-1}(\mathbf{Y}(\bar{\mathbf{x}}_i) - \bar{\mathbf{Y}}(\bar{\mathbf{x}}_i))(\mathbf{Y}(\bar{\mathbf{x}}_i) - \bar{\mathbf{Y}}(\bar{\mathbf{x}}_i))^T, \quad (2)$$

где $\bar{\mathbf{Y}}(\bar{\mathbf{x}}_i)$ – матрица размером $N \times r$, с элементами $\bar{y}_{lk}(\bar{\mathbf{x}}_i) = \bar{y}_l(\bar{\mathbf{x}}_i)$, $l = 1, 2, \dots, r$, $k = 1, 2, \dots, N$.

Максимальное собственное число матрицы (2) находится путем вычисления характеристического уравнения вида $(\mathbf{D}\{\mathbf{Y}(\bar{\mathbf{x}}_i)\} - \bar{\lambda}(\bar{\mathbf{x}}_i)\mathbf{I})\bar{\beta}(\bar{\mathbf{x}}_i) = 0$, где $\bar{\lambda}(\bar{\mathbf{x}}_i)$ – вектор собственных чисел, $\bar{\beta}(\bar{\mathbf{x}}_i)$ – собственный вектор, $\mathbf{I}$ – единичная матрица размера $r \times r$. В нашем случае под выходной переменной $v(\bar{\mathbf{x}})$, характеризующей разброс химических показателей по всей партии, будем понимать значение максимального собственного числа $\lambda_{\max}(\bar{\mathbf{x}}_i)$ матрицы (2). Пусть зависимость между выходной переменной $\lambda_{\max}$ и вектором входных переменных $\bar{\mathbf{x}}$ определяется соотношением

$$\lambda_{\max} = \eta(\bar{\mathbf{x}})\xi, \quad (3)$$

где $\eta(\bar{\mathbf{x}})$ – истинная модель, вид которой неизвестен; и ξ – независимая

случайная ошибка, распределенная по закону $\frac{\chi^2(\gamma)}{\gamma}$; γ – число степеней

свободы случайной величины ξ , распределенной по закону χ^2 Пирсона.

Модель (3) будем называть дисперсионной моделью объекта управления. Целью идентификации модели (3) является построение модели с оптимальной прогнозирующей способностью. Для этого будем использовать следующий прием. Пусть на классе моделей задана структура, позволяющая ввести частичный порядок. При таком порядке классы моделей как бы вложены один в другой: $S_1 \subset S_2 \subset \dots \subset S_q$. Такого эффекта можно достичь, если модели представляют собой полиномы возрастающей степени. Для линейной по параметрам модели можно задать структуру в зависимости от членов модели. В этом случае каждый класс S_j задается моделью

$$\eta(\bar{\mathbf{x}}, \bar{\mathbf{a}}_j) = \mathbf{f}^T(\bar{\mathbf{x}}) \bar{\mathbf{a}}_j \quad (j = 1, 2, \dots, q), \quad (4)$$

где $\mathbf{f}_j^T(\bar{\mathbf{x}}) = [f_1(\bar{\mathbf{x}}), f_2(\bar{\mathbf{x}}), \dots, f_{n_j}(\bar{\mathbf{x}})]$ – вектор известных функций от вектора $\bar{\mathbf{x}}$; n_j – число параметров модели j ; q – максимально возможный порядок модели; $\bar{\mathbf{a}}_j^T = [\alpha_1, \alpha_2, \dots, \alpha_{n_j}]$ – вектор неизвестных параметров, требующий определения.

Идентификация структуры моделей осуществляется с помощью специальных критериев, которые учитывают сложность модели и позволяют выбирать модель с наилучшей прогнозирующей способностью. С точки зрения этого приема метод идентификации дисперсионной модели объекта управления может быть представлен в виде следующего алгоритма.

- 1 Для каждой панели данных вида (1) оценить ковариационную матрицу (2) и путем решения характеристического уравнения определить максимальное собственное число матрицы (2). В результате проведенных вычислений должен быть получен следующий ряд

$$\lambda_{\max}(\bar{\mathbf{x}}_1), \lambda_{\max}(\bar{\mathbf{x}}_2), \dots, \lambda_{\max}(\bar{\mathbf{x}}_i), \dots, \lambda_{\max}(\bar{\mathbf{x}}_K). \quad (5)$$
- 2 Для ряда (5) выбирается структура класса линейных по параметрам моделей, при помощи которых описывается зависимость вида (3). Параметры каждой модели класса (4) оцениваются итерационным взвешенным методом наименьших квадратов. Выбор модели, обладающей наилучшей прогнозирующей способностью на заданной структуре класса моделей, осуществляется при помощи определения критерия скользящего контроля для всех моделей исследуемого множества [2].
- 3 Для прогнозирования выбирается модель с минимальным значением критерия селекции.

Рассмотрим применение предложенного метода для контроля качества технологического процесса выдержки виноматериала. Данные о процессе выдержки вина марки «Старый Нектар» предоставлены заведующей производственной лабораторией Инкерманского завода марочных вин Л.П. Шимчук. Для производства партии вина было использовано сырье, полученное от восьми производителей. Этот факт был учтен в исследовании для разбиения партии на подпартии. Значения переменных x_2 и x_3 усреднялись для каждой подпартии. Без перемешивания сырье было разлито в бутылки, где проходило его настаивание в течение четырех лет. Для каждой подпартии было использовано по 32 бутылки. На основании предоставленных данных были сформированы панели данных вида (1), на основе которых были оценены ковариационные матрицы (2) и определены их максимальные собственные числа. Результаты вычислений показаны на рисунке 1.

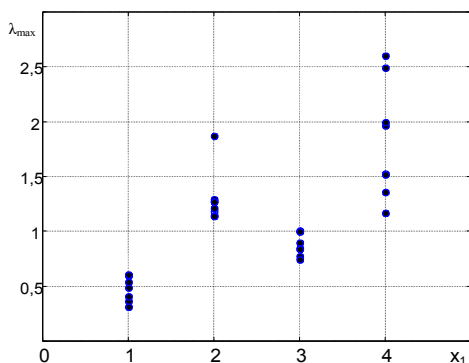


Рисунок 1 – Экспериментальные значения разброса химических показателей по подпартиям

ленных данных были сформированы панели данных вида (1), на основе которых были оценены ковариационные матрицы (2) и определены их максимальные собственные числа. Результаты вычислений показаны на рисунке 1.

Для описания зависимости значения максимального собственного числа от значений входных переменных использована следующая структура множества моделей

$$\begin{aligned} 1 \quad \lambda_{\max} &= \alpha_0 + \alpha_1 x_2 + \alpha_2 x_3 \\ 2 \quad \lambda_{\max} &= \alpha_0 + \alpha_1 x_2 + \alpha_2 x_3 + \alpha_3 x_1 \\ 3 \quad \lambda_{\max} &= \alpha_0 + \alpha_1 x_2 + \alpha_2 x_3 + \alpha_3 x_1 + \alpha_4 x_1^2, \end{aligned} \quad (6)$$

где α_0 , α_1 , α_2 , α_3 и α_4 – неизвестные параметры, требующие определения.

Согласно алгоритму метода идентификации дисперсионной модели объекта управления оценены параметры моделей множества (6) и вычислены значения критерия селекции. Анализ результатов вычисления (таблица 1) показывает, что наилучшей прогнозирующей способностью на заданной структуре класса моделей обладает вторая модель. Таким образом, зависимость максимального собственного числа от входных переменных можно описать как

$$\lambda_{\max} = -17,3460 + 0,2477x_2 + 0,8159x_3 + 0,4107x_1. \quad (7)$$

Таблица 1 – Результаты вычислений

Номер модели	Параметры модели					Значение критерия селекции
	α_0	α_1	α_2	α_3	α_4	
1	-1,499	-1,482	1,677			0,254
2	-17,346	0,248	0,816	0,411		0,202

сырье, полученное от восьми производителей. Этот факт был учтен в исследовании для разбиения партии на подпартии. Значения переменных x_2 и x_3 усреднялись для каждой подпартии. Без перемешивания сырье было разлито в бутылки, где проходило его настаивание в течение четырех лет. Для каждой подпартии было использовано по 32 бутылки.

На основании предостав-

3	-20,227	0,677	0,566	0,093	0,079	0,219
---	---------	-------	-------	-------	-------	-------

Учитывая то, что переменные x_2 и x_3 ограничены требованием ГОСТ, и осуществляется прогноз на следующий год жизни вина, функцию (7) можно оптимизировать, используя методы линейного программирования. Значения переменных x_2 и x_3 , которые будут приводить функцию (7) к минимуму, могут быть использованы технологом как эталонные значения для корректировки показателей каждого бута в партии.

Библиография

1 Baltagi B.H. Econometric Analysis of Panel Data / B.H. Baltagi. — N.Y.: Wiley, 1995. — 398 p.

2 Tsukanov A.V. The Monte-Carlo comparison of two criteria for the selection of the models / A.V. Tsukanov, A.M. Herzberg // J. of Statist. Computation and Simulation. — GB, 1985. — № 22. — P. 113–126.

Поступила в редакцию 10.10.2000 г.