

УДК 543.51:543.86

ИСПОЛЬЗОВАНИЕ КОЭФФИЦИЕНТА МНОЖЕСТВЕННОЙ КОРРЕЛЯЦИИ ПРИ КОЛИЧЕСТВЕННОМ МАСС-СПЕКТРОМЕТРИЧЕСКОМ АНАЛИЗЕ

Е. С. БРОДСКИЙ, А. И. НОВИКОВ, Р. А. ХМЕЛЬНИЦКИЙ

(Кафедра физической и коллоидной химии)

Проведение количественного масс-спектрометрического анализа основано на аддитивности вкладов компонентов и линейной зависимости аналитических характеристик последних, что отражается системой линейных уравнений вида

$$AX=Y, \quad (1)$$

Здесь A — матрица (m, n), элементы (a_{ij}) которой представляют собой аналитические характеристики (интенсивности пиков или групп ионов) в масс-спектрах чистых компонентов; $Y=(Y_1, \dots, Y_m)$ — вектор аналитических характеристик анализируемого образца; $X=(X_1, \dots, X_n)^T$ — вектор вкладов каждого компонента в масс-спектр смеси. Вектор неизвестных X обычно находится путем решения квадратной системы линейных уравнений, выделяемой из (1), или решением переопределенной системы (1) методом наименьших квадратов [1].

Во многих случаях, особенно при анализе сложных смесей, могут возникнуть трудности, обусловленные недостаточно точным определением матрицы калибровочных коэффициентов A или неполным знанием вида компонентов, присутствующих в анализируемой смеси. Поэтому желательно, чтобы расчетная модель позволяла оценивать не только значения неизвестных и ошибки их

определения, но и полноту учета всех неизвестных, адекватность линейной модели и существенность вклада каждого из переменных. Это можно осуществить с помощью коэффициента множественной корреляции.

Пусть

$$AX=\hat{Y}=Y+\varepsilon, \quad (2)$$

где ε — невязки, обусловленные несоответствием экспериментально получаемых значений Y и определяемых с помощью модели. Все измерения представляются в стандартной форме с нулевым математическим ожиданием и единичной дисперсией. Приведя систему (2) к виду нормальных уравнений, получим

$$MX=S, \quad (3)$$

где $M=A^TWA$ — корреляционная матрица аналитических характеристик компонентов; $S=A^TWY$ — вектор коэффициентов корреляции между аналитическими характеристиками чистых компонентов и анализируемого образца; W — весовая матрица.

Степень линейной связи между Y и набором X_j в целом характеризуется коэффициентом множественной корреляции:

$$R = \frac{Y^T W \hat{Y}}{(\hat{Y})}. \quad (4)$$

Значения неизвестных и их вклады в коэффициент детерминированности для легкой углеводородной фракции (т. кип. 180—260°)

X_j	σ_{X_j}	K_j	V_j	Матрица вкладов			
				X_1	X_2	X_3	X_4
0,267	0,0073	0,517	0,2323	0,2016	0,0775	0,0370	0,0099
0,482	0,0085	0,851	0,5591		0,4319	0,0475	0,0821
0,229	0,0109	0,478	0,1919			0,1593	0,0211
0,067	0,0200	0,268	0,0176				0,0043
			$R^2=0,9999$				
			$R=0,9999$				

Примечание. X_j — значения неизвестных; σ_{X_j} — ошибка определения неизвестных; K_j — коэффициенты корреляции столбцов матрицы A_j и вектора аналитических характеристик образца Y ; V_j — вклад j -го компонента в R^2 .

Величина R^2 (коэффициент детерминированности) оценивает часть Y обусловленную существованием линейной связи (2).

Из соотношения $(Y - \hat{Y})^T Y = 0$ получаем

$$R\sigma(\hat{Y}) = \tilde{\sigma}^2(\hat{Y}).$$

Отсюда, используя (3) и (4), будем иметь

$$R^2 = Y^T W A X = S^T X = X^T M X. \quad (5)$$

Вклад i -го компонента в модель есть i -й член скалярного произведения

$$(S^T X)_i = S_i X_i = \sum_{j=1}^n M_{ij} X_i X_j.$$

Этот вклад, в свою очередь, можно разделить на непосредственный вклад i -го компонента, равный $m_{ii} X_i^2$ и опосредствованный, получающийся за счет корреляции i -го компонента и других компонентов — $\sum_{j \neq i}^n m_{ij} X_i X_j$; элементы $m_{ij} X_i X_j$ характеризуют взаимное влияние компонентов.

По физическому смыслу значения неизвестных и их вклады в коэффициент де-

терминированности R^2 не могут быть отрицательными. Поэтому наличие больших отрицательных элементов в матрице вкладов может служить признаком недостаточной определенности соответствующих параметров модели.

В таблице приведены результаты определения состава легкой углеводородной фракции с использованием калибровочных коэффициентов [2]. В ней даны значения неизвестных, полученные методом наименьших квадратов, оценка ошибки их определения, коэффициенты корреляции столбцов матрицы и вектора аналитических характеристик образца, а также общий вклад каждого неизвестного в R^2 и матрица вкладов. Как показывают данные таблицы, расчетная модель (2) достаточно хорошо определяет анализируемый образец: коэффициент детерминированности равен 0,9999, коэффициент множественной корреляции — 0,9999. Для всех неизвестных, кроме X_4 , собственный вклад в R^2 значительно превышает вклад за счет корреляции. Матрица вкладов позволяет оценить влияние каждого неизвестного на определение других неизвестных.

ЛИТЕРАТУРА

1. Бернارد Дж. Современная масс-спектрометрия. М.: ИЛ, 1957. — 2. Бродский Е. С., Лукашенко И. М., Волков Ю. А. Детализация масс-спектрометрического анализа группового состава на-

сыщенных углеводородов нефти. — Химия и технол. топлив и масел, 1976, № 2, с. 56.

Статья поступила 18 марта 1981 г.

SUMMARY

A method is suggested for the use of multiple correlation coefficient in calculation model for quantitative mass spectrometric analysis of complex mixtures. This method allows to determinate along side with unknown values, a contribution of each unknown into linear model.