

## ПРИМЕНЕНИЕ МЕТОДОВ МНОГОКОМПОНЕНТНОГО АНАЛИЗА ДЛЯ ОПРЕДЕЛЕНИЯ СОСТАВА ВЕЩЕСТВА ПО ЦВЕТОВОЙ ШКАЛЕ

Т.А. Выймова, А.С. Спиридонова  
Томский политехнический университет  
[vymovat@mail.ru](mailto:vymovat@mail.ru)

### Введение

Информация о составе и строении сложных смесей (веществ), определение их компонент является одной из наиболее актуальных задач аналитической химии. Она необходима для обеспечения надлежащего качества промышленного сырья и продукции, повышения эффективности сельскохозяйственного производства, совершенствования здравоохранения, решения экологических проблем. С каждым годом потребность в подобной информации возрастает в связи с увеличением количества выполняемых анализов и исследуемых веществ и расширением диапазона определяемых содержаний.

Целью работы является разработка метода многокомпонентного определения состава вещества по цветовой шкале

Наиболее распространенными методами химического анализа для определения качественного и количественного состава веществ являются следующие:

- спектральный анализ;
- хроматографический метод;
- спектрофотометрический метод;
- цветометрический анализ [1].

Использование цветометрии в качестве получения информации в химическом анализе началось несколько десятков лет назад. Любой цвет можно представить как сумму трех его составляющих – основных цветов, в соответствии с трехкомпонентной теорией зрения; в этой системе цвет может быть изображен с тремя координатами цвета – тремя числами, где последние соответствуют количествам основных цветов в данном цвете при стандартных условиях его наблюдения [2].

В последние десятилетия развивается цифровой цветометрический анализ (ЦЦА). Суть данного метода заключается в том, что он позволяет определить качественные и количественные характеристики состава вещества, при взаимодействии реагентов с веществом, образец меняет цвет, что указывает на присутствие в нем определенного компонента [3]. Данный метод также относится к аналитической химии, где в свою очередь используются хемометрические методы получения информации.

### Анализ многомерных данных

Анализ многомерных данных (АМД) является предметом хемометрики и применяется для моделирования многомерных (многофакторных) данных. Он основан на применении проекционных

математических методов, которые позволяют выделить в больших массивах данных скрытые переменные и проанализировать связи, которые существуют в изучаемой системе [4]. В электронном ресурсе Российского хемометрического общества [5] рассматривают такие методы исследования, обработки и интерпретации накопленных данных, как:

- метод главных компонент (МГК);
- множественная калибровка;
- множественная линейная регрессия (МЛР);
- регрессия на главные компоненты (РГК)
- проекция на латентные структуры (ПЛС1, ПЛС2).

В ходе предварительных исследований были рассмотрены и выявлены преимущества следующих методов: метод главных компонент; множественная линейная регрессия и регрессия на главные компоненты. Исследование методов проводилось с применением цветовой шкалы. Данная шкала представляет собой набор оптодов. Оптоды вырабатываются из полимерных прозрачных материалов по специальной технологии. Это прозрачный индикатор чувствительного материала (ИЧМ), который позволяет упростить визуальную и фотометрическую оценку изменения его окраски после контакта с анализируемым объектом [6]. Объектом являлась смесь из трех веществ  $Co$ ,  $Cu$ ,  $Ni$ , эксперимент повторили семь раз.

### Реализация методов

Целью метода главных компонент (ГК) является преобразование исходного описания образцов с помощью специальных переменных в новую форму, допускающую извлечение из данных необходимой информации [5].

В ходе реализации алгоритма был построен график зависимости собственных значений, которые характеризуют важность каждой компоненты, от общего числа компонент, график представлен на рисунке 1.

Общее число ГК  $a$  определяется точкой пересечения графика с осью  $n$ .

МЛР является классическим подходом или, как ее еще называют, классической калибровкой, в которой участвуют несколько каналов (в нашем случае три канала:  $R$ ,  $G$ ,  $B$ ). Соответствующая модель строится с помощью множественной линейной регрессии (МЛР). При большом числе канал отбирается такое число каналов, которые будут меньше числа веществ, входящих в смесь.

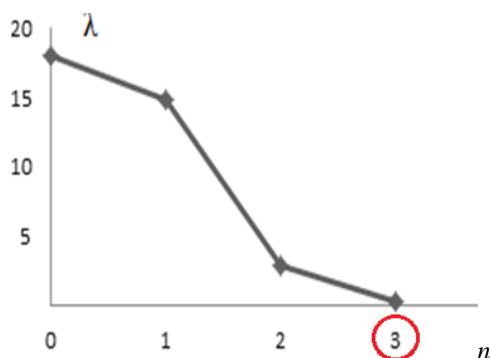


Рис. 1 – График зависимости собственных значений –  $\lambda$  от общего числа компонент  $n$ .

Так как смесей у нас семь, а каналов всего три берем все три канала и строим по ним модель МЛР, данная модель отражает предсказанное количества вещества и измеренное (концентрацию).

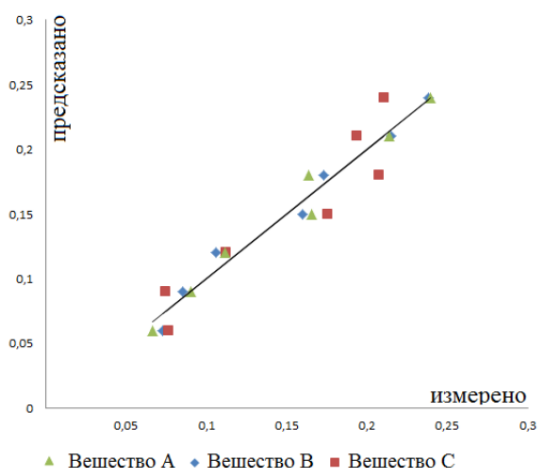


Рис. 2 – Измеренное и предсказанное значение концентрации (мг/л) веществ в смесях

Регрессия на главные компоненты рассматривается, как двухэтапная процедура: с помощью МГК преобразуется матрица  $X$ , а затем получившаяся матрица счетов  $T$  непосредственно используется в модели многомерной регрессии.

Таблица 1 – Результаты реализации РГК

№	Канал	Уравнение	$R^2$
1	R	$y = 0,861x - 0,021$	0,861
	G	$y = 0,297x - 0,106$	0,297
	B	$y = 0,975x - 0,004$	0,975
2	R	$y = 0,660x - 0,099$	0,66
	G	$y = 0,978x - 0,147$	0,978
	B	$y = 0,978x - 0,147$	0,978
3	R	$y = 0,603x - 0,090$	0,603
	G	$y = 0,789x - 0,118$	0,789
	B	$y = 0,876x - 0,131$	0,876

В данном методе в отличие от МЛР строятся три модели, для каждого вещества в смеси определяется свой рабочий канал. Данные приведены в таблице 1.

Если сравнивать два последних метода (РГК и МЛР), то по результатам видно, что на модели МЛР все линии тренда совпали и коэффициенты корреляции равны и стремятся к единице. Т.е. была получена «слишком хорошо» описанная модель. Однако, при излишней сложности ( $n = 4, 5$ ), модель все хуже работает. А вот РГК модель показывает рабочие каналы для каждого вещества отдельно. Были получены высокие значения коэффициентов корреляции для рабочих каналов.

### Заключение

В данной работе были реализованы следующие методы: метод главных компонент, множественная линейная регрессия и регрессия на главные компоненты.

МГК является качественным методом, с его помощью мы определили число главных компонент в смесях. При сравнении МЛР и РГК выявили, что регрессия на главные компоненты имеет явные преимущества в сравнении с методами классической калибровки. Этот способ моделирования точнее и имеет меньшее смещение. Это объясняется тем, что в многомерной калибровке используются все имеющиеся экспериментальные данные, что позволяет избежать переоценки данных.

Дальнейшая работа будет состоять в исследовании других методов, таких как: проекция на латентные структуры (ПЛС1 и ПЛС 2) и множественная калибровка. Будет разработан алгоритм обработки многомерных данных с использованием математического пакета MatLab.

### Литература

1. Шаевич А.Б. Аналитическая служба как система. – М.: Химия, 1981.
2. Синтез цвета // Фотокинетика: Энциклопедия. – М.: Советская энциклопедия, 1981.
3. Муравьев С.В., Гавриленко Н.А., Силушкин С.В., Овчинников П.Г. Мобильный цветометрический комплекс для измерения состава вещества на основе полимерных оптодов // Известия Томского политехнического университета. – 2011. – Т. 318. – № 4. – С. 68-73.
4. Martens H., Martens M., Multivariate Analysis of Quality, NY, Willey, 1998.
5. Российское хемометрическое общество [Электронный ресурс]. – Режим доступа: <http://rcs.chph.ras.ru/>. – (дата обращения 15.02.2014)
6. Индикаторный чувствительный материал для определения микроколичеств веществ: пат. 272284 Рос. Федерация No 2004125304/04; заявл. 18.08.04; опублик. 20.03.06, Бюл. No 8 – 9 с