

Кривенко Д.Г.

Науч. руководитель д.т.н. Марасанов В.В.

Херсонский национальный технический университет

Аппроксимационные методы классификации веществ в масс-спектрометрических методах

Результаты масс-спектрометрии представляют набор дискретных данных, расположенных на различных (непериодических) временных интервалах. Для сравнения образцовых и эталонных веществ с результатами экспериментальных исследований апроксимируем как эталонные данные, так и экспериментальные полиномом n -степени. Порядок полинома выбираем исходя из объема (длины дискретного ряда) эталонных данных. Процесс идентификации (близости) экспериментальных данных к эталону находится путем сравнения (нахождения) минимального расстояния между коэффициентами эталонного и экспериментального полинома[1].

Постановка задачи. Проверка на контрольной выборке данных масс-спектрометрических исследований пригодности аппроксимационных методов распознавания веществ. Сравнительная оценка с другими методами.

Решение задачи. Для набора данных

$$(x_i, y_i) = 1, 2, \dots, N \quad (1)$$

где $N=65$ в нашем случае, требуется найти такой полином степени n ,

$$p^{(n)}(x) = p_1 x^n + p_2 x^{n-1} + \dots + p_n x + p_{n+1} \quad (2)$$

который наименее уклоняется от заданных данных в том смысле, что сумма квадратов расстояний от заданных точек (x_i, y_i) до $(x_i, p^{(n)}(x_i))$ будет минимальной[2]. Степень полинома должна быть минимальной, в нашем случае степень полинома равна 23 (минимальная погрешность).

Затем находим минимальное расстояние(S_i), между коэффициентами эталонного и экспериментального полинома.

$$S_i = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (P_{i_{обр}} - P_{i_{эта}})^2} \quad (3)$$

Минимальное отклонение расстояния коэффициентов полинома будет свидетельствовать о большой степени подобия данного образца и эталона.

Моделирование осуществлялось программой MathLab, текст программы представлен ниже:

```
for i=1:1:4
x=[1:1:65];
p= polyfit(x,A(i,:),23) ;
p1= polyfit(x,A(1,:),23) ;
v(i)=abs(sqrt(sum(p.^2-p1.^2)));
xx = linspace(x(1), x(end), 100);
yy = polyval(p, xx);
G(i,:)=yy;
subplot(2, 2, i),plot(x,A(i,:), '*', xx,
yy); switch (i)
    case 1
        title('Образцовое вещество'
); case 2
        title('1 эталон'
); case 3
        title('2 эталон'
); case 4
        title('3 эталон' );
end
xlabel('Время (мин.)')
ylabel('Концентрация соединения (C)')
grid
on end
```

На рисунке 1 показан график полинома от заданных точек. В таблице 1 представлены коэффициенты близости экспериментальных данных к эталонным значениям

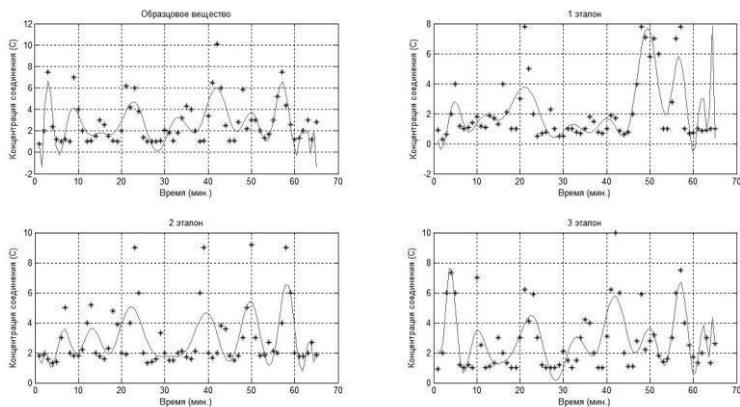


Таблица 1. Коэффициенты близости полиномов

1 эталон	2 эталон	3 эталон
423.2911	352.7262	185.3795

Рис.1. Вид полиномов

Выводы. Метод дал совпадение по правильной идентификации. Преимущество – простая алгоритмическая реализация по сравнению с методами главных компонент[3]. Данный метод целесообразно использовать для оптимизации и отсеивания вариантов исследования.

Литература.

1. Елисева И.И., Юзбашев М.М. Общая теория статистики: Учебник / Под ред. И.И. Елисеевой. — 4-е издание, переработанное и дополненное. — Москва: Финансы и Статистика, 2002. — 480 с.

2. Обзор средств MATLAB и ToolBox'ов для приближения данных [Электронный ресурс] – Режим доступа:<http://matlab.exponenta.ru/spline/book1/7.php#4>

3. Метод главных компонент [Электронный ресурс] – Режим доступа:<http://ru.wikipedia.org/> метод главных компонент