

Использование методов Монте-Карло и латинского квадрата при планировании градуировочного эксперимента

А.С. Мананков

Самарский государственный технический университет, Самара, Россия

Обоснование. В химическом спектральном анализе часто исследуются сложные смеси, включающие два и более компонентов. На основании этих спектральных данных строится градуировочная модель, с помощью которой можно количественно определять концентрации веществ в образце. Модель строится на основании обучающего набора. Имея обученные модели, можно определить концентрации всех анализируемых компонентов из единичного спектра.

Планирование эксперимента является, таким образом, оптимизационной задачей, и для ее решения необходимы параметры, которые могут численно охарактеризовать качество набора [1] с точки зрения его конечной цели. В нашем случае это создание достаточно точных и устойчивых градуировочных моделей на спектральных данных обучающих и проверочных образцов.

Одним из распространенных методов создания градуировочного набора для многокомпонентной смеси является метод Монте-Карло, т. е. случайное заполнение пространства эксперимента (ПЭ) — квадрата (в случае двухкомпонентной смеси), стороны которого образованы концентрациями анализируемых компонентов.

В отличие от методов систематического или простого случайного заполнения ПЭ, методы Монте-Карло предполагают многократное повторение случайного заполнения для поиска оптимальных или приемлемых решений. При этом чем меньше образцов включено в градуировочный набор (то есть чем меньше точек расположено в ПЭ), тем значительно возрастает число итераций, необходимых для нахождения качественных наборов. Основным недостатком данных методов — необходимость многократных повторений для достижения статистической достоверности результатов, что связано с длительным временем расчета.

Цель — с помощью ограничений латинского квадрата сформировать градуировочный набор и достигнуть допустимых критериев качества минимальным числом обучающих образцов.

Методы. Рассматривался набор от 9 до 25 образцов, где при каждом формировании набора высчитывались три критерия качества [2], отражающие корреляцию концентрации компонентов, равномерность (функция размаха) и полноту заполнения ПЭ (незаполненность) с фиксацией минимальных значений [1]. Диапазоны критериев отражены в табл. 1.

Таблица 1. Границы критериев качества

Критерий	Наилучший диапазон	Средний диапазон	Допустимый диапазон	Худший диапазон
Корреляция	0,0–0,15	0,15–0,3	–	>0,3
Функция размаха	0,0–0,2	0,2–0,4	0,4–0,8	>0,8
Незаполненность	0,0–0,15	0,15–0,3	–	>0,3

Результаты. Достигнутые минимальные значения критериев качества отражены в табл. 2.

Таблица 2. Результаты моделирования

Число образцов	Корреляция (MIN)	Функция размаха (MIN)	Незаполненность пространства (MIN)
9	< 0,02	0,64	0,22
12	< 0,05	0,5	0,18
15	< 0,1	0,39	0,14

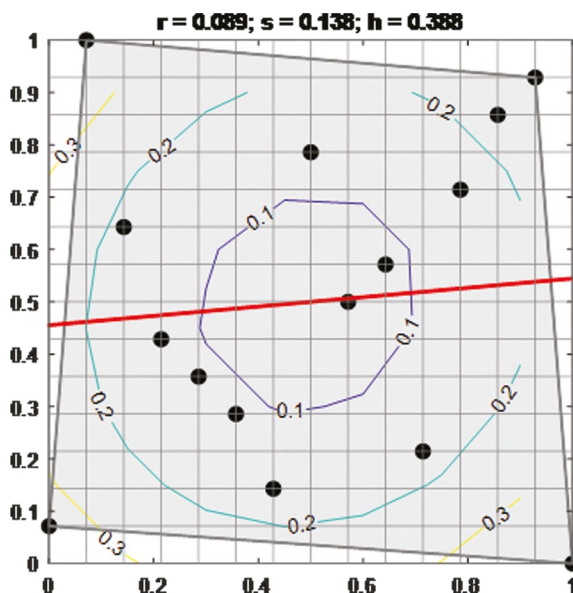


Рис. 1. Один из найденных наборов из 15 точек-образцов

На рис. 1 точка определяется концентрациями двух компонентов смеси, наклон красной линии регрессии по точкам показывает величину коэффициента корреляции между концентрациями компонентов (r), контурный график описывает уровни значений функции размаха (h), а площадь закрашенного выпуклого многоугольника обозначает область заполнения ПЗ, так что незаполненность (s) является отношением незаполненной площади ПЗ к ее общей площади.

Выводы. Использование латинского квадрата позволило сократить объем обучающих данных. Например, среднее качество градуировочного набора было достигнуто для 15 образцов (меньше чем в работе [3]).

Ключевые слова: планирование эксперимента; латинский квадрат; критерии качества; математическое моделирование.

Список литературы

1. Bogomolov A. Diagonal designs for a multi-component calibration experiment // *Analytica Chimica Acta*. 2017. Vol. 951. P. 46–57. doi: 10.1016/j.aca.2016.11.042 EDN: YUVUCH
2. Мананков А.С., Богомолов А.Ю. Критерии оценки обучающего набора образцов при планировании многокомпонентного градуировочного эксперимента // Сборник научных трудов Международной научно-практической конференции молодых ученых, аспирантов и студентов. Кинель, 2024. С. 61–64. EDN: SYSQZP
3. Мананков А.С., Богомолов А.Ю. Применение квазислучайных множеств при планировании многокомпонентного градуировочного эксперимента // L Самарская областная студенческая научная конференция: тезисы докладов. Санкт-Петербург: Эко-Вектор, 2024. С. 356–358. EDN: SJCBKU

Сведения об авторе:

Александр Сергеевич Мананков — аспирант, группа 2-УПНК-1.4.2, химико-технологический факультет; Самарский государственный технический университет, Самара, Россия. E-mail: s90w23.14@mail.ru

Сведения о научном руководителе:

Андрей Юрьевич Богомолов — доктор химических наук, доцент; Самарский государственный технический университет, Самара, Россия. E-mail: c11b0f5057f5@mail.ru