

НОВЫЕ ПОДХОДЫ К ПРИМЕНЕНИЮ МЕТОДА МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ В СПЕКТРОФОТОМЕТРИЧЕСКОМ АНАЛИЗЕ МНОГОКОМПОНЕНТНЫХ СМЕСЕЙ

И.В. Власова, А.С. Шелпакова, С.М. Добровольский, А.В. Фисенко

*ГОУ ВПО Омский государственный университет им. Ф.М. Достоевского
644077 Омск, пр. Мира, 55а
vlaso-iri@yandex.ru*

Поступила в редакцию 31 июля 2009 г.

Метод множественной линейной регрессии применен для вычисления коэффициентов поглощения компонентов по спектрам смесей. Использование новых коэффициентов позволяет повысить точность спектрофотометрического анализа неразделенных смесей, а также сократить время на проведение анализа, поскольку не требуется проводить дополнительную оптимизацию спектральных диапазонов для каждого компонента.

Ключевые слова: спектрофотометрический анализ, множественная линейная регрессия, анализ многокомпонентных смесей.

Власова Ирина Васильевна, доцент кафедры аналитической химии Омского государственного университета им. Ф.М. Достоевского, к.х.н.

Область научных интересов – спектроскопический анализ неразделенных смесей органических веществ.

Опубликовано 58 работ, в том числе 3 авторских свидетельства.

Шелпакова Анна Сергеевна – аспирант кафедры аналитической химии Омского государственного университета им. Ф.М. Достоевского.

Область научных интересов – хемометрика, анализ многомерных данных, спектроскопический анализ неразделенных смесей органических веществ.

Опубликовано 9 работ.

Добровольский Сергей Михайлович, доцент кафедры математического анализа Омского государственного университета им. Ф.М. Достоевского.

Область научных интересов – стохастические динамические системы, прикладная статистика.

Опубликовано 35 работ, в том числе 2 авторских свидетельства.

Фисенко Анна Владимировна – студентка химического факультета Омского государственного университета им. Ф.М. Достоевского.

Метод множественной линейной регрессии (**МЛР**) позволяет расширить возможности спектрофотометрии и анализировать сложные смеси без их разделения даже в тех случаях, когда спектры компонентов полностью перекрываются [1-3]. Расчеты ведут с применением предварительно вычисленных коэффициентов поглощения. Однако при реализации данного способа нахождения концентраций компонентов смеси возникает ряд трудностей. Во-первых, алгоритм МЛР предполагает соблюдение условия аддитивности оптической плотности, и если в спектрах смесей есть участки с достоверными отклонениями от аддитивности, их необходимо

исключать. Во-вторых, для снижения погрешностей определения требуется оптимизация спектральных диапазонов и выбор для каждого компонента своего участка спектра [3, 4]. Это увеличивает время на разработку методики анализа, однако не всегда приводит к желаемому результату [4].

Но выбор спектрального диапазона не является единственным способом оптимизации регрессионного анализа. Возможен и другой путь, который основан на вычислении коэффициентов не по спектрам растворов индивидуальных соединений, а по спектрам их смесей точно известного состава [5, 6]. Это позволяет

приблизить условия вычисления коэффициентов к условиям анализа реальных смесей, а, следовательно, нивелировать погрешности, связанные с возможным взаимодействием компонентов между собой и появлением неаддитивных участков в спектрах.

Цель работы – повышение точности спектрофотометрического анализа многокомпонентных смесей. Для достижения поставленной цели предлагается вычислять коэффициенты поглощения компонентов по спектрам смесей известного состава, используя метод МЛР.

Объектами исследования были бинарные смеси лекарственных веществ – папаверина гидрохлорида (ПГ) и дибазола (Дб), а также 4-х компонентные смеси водорастворимых витаминов группы В. Исходные растворы индивидуальных веществ готовили растворением точных навесок в дистиллированной воде. Рабочие растворы ПГ, Дб и их смесей готовили в среде аммиачного буферного раствора, растворы витаминов – в среде 0,01 М раствора соляной кислоты. В общей сложности были приготовлены 23 смеси ПГ и Дб и 25 смесей витаминов, в которых меняли концентрации и соотношения компонентов. Часть смесей в дальнейшем была использована для вычисления коэффициентов регрессии, а оставшиеся – для проверки результатов анализа, выполненных с применением новых коэффициентов.

Спектры растворов индивидуальных соединений и их смесей регистрировали на спектрофотометре СФ-2000 в кварцевых кюветках толщиной 1 см, в спектральном диапазоне от 220 до 300 нм с шагом 0,2 нм для смесей витаминов, и с шагом 5 нм – для смесей ПГ и Дб. По спектрам растворов индивидуальных компонентов вычисляли удельные коэффициенты поглощения k_{y_0} так, как изложено в [4].

Полученные спектральные данные обрабатывали методом МЛР при помощи программного пакета Microsoft Excel, используя меню Сервис → Анализ данных → Регрессия (или Поиск решения). В основе вычислений МЛР лежит метод наименьших квадратов (МНК). В данном случае, согласно закону светопоглощения и свойству аддитивности оптической плотности, осуществляли поиск решения, которое удовлетворяло условию:

$$\sum_{i=1}^{i=n} (A_i - \sum_{j=1}^m k_{ij} \cdot l \cdot C_j)^2 \rightarrow \min \quad (1)$$

где A_i – оптическая плотность смеси при i -ой длине волны; k_{ij} – удельный коэффициент поглощения j -го компонента смеси при i -ой длине волны; l – толщина поглощающего слоя, равная 1 см; C_j – концентрация j -го компонента, мкг/мл.

Решать систему (1) можно как относительно коэффициентов, так и относительно концен-

траций. Оба подхода реализованы в рамках данной работы. В первом случае задавали A_i , концентрации компонентов в смесях и решали систему, в которой число уравнений равно числу модельных смесей. Вычисленные таким образом для всех компонентов смеси коэффициенты k_{ij} относились к i -ой длине волны. Расчеты повторяли столько раз, сколько длин волн содержал спектр смеси.

Во втором случае задавали значения оптической плотности, коэффициентов поглощения и решали переопределенную систему относительно концентраций, в которой число уравнений равно числу длин волн, при которых измерены значения A .

Для каждого j -го компонента смеси вычисляли относительную погрешность его определения (δ_{jq}), а полученные данные усредняли по n смесям:

$$\delta_j = n^{-1} \sum_{q=1}^{q=n} |\delta_{jq}| \quad (2)$$

Кроме того, рассчитывали обобщенную по m компонентам смеси характеристику погрешностей анализа ξ :

$$\xi = \sqrt{\sum_{i=1}^{i=m} (\delta_i)^2} \quad (3)$$

Разумеется, усредненные значения δ_j , а также величина ξ не являются погрешностями в строгом метрологическом смысле этого слова, однако использование этих условных характеристик удобно для сопоставления точности анализа одних и тех же смесей с применением разных коэффициентов.

Рассмотрим результаты, полученные для системы папаверин-дибазол. Как известно, при анализе бинарных смесей достаточно точные результаты можно получать, не прибегая к МЛР, а используя метод Фирордта (МФ) [7]. Ранее нами было установлено, что в спектрах смесей ПГ и Дб имеются участки с достоверными отклонениями от аддитивности ΔA , причем абсолютные значения отклонений различны при разных соотношениях компонентов [8]. Традиционно считается, что МФ можно применять только при условии аддитивности оптической плотности при аналитических длинах волн. Однако, как показано в работе [9], МФ может давать точные результаты и при наличии достоверных отклонений от аддитивности ΔA , если только отклонения не превышают некоторых допустимых значений. Так, например, при соотношении ПГ : Дб равным 1 : 1, ΔA оказались невелики и позволяли одновременно определять оба компонента на одном наборе аналитических длин волн – 240 и 280 нм, с погрешностями менее 5 % отн. При увеличении доли Дб в смеси, до 1 : 10, ΔA возрастали настолько, что приводили к систе-

матическим погрешностям, вплоть до 10-15% отн. [9]. Иными словами, МФ с применением коэффициентов, вычисленных традиционным способом, оказался пригодным для анализа данной системы только в узком интервале соотношений ПГ и Дб – от 2 : 1 до 1 : 2.

Переход к вычислению концентраций ПГ и Дб методом МЛР с применением всего спектрального диапазона не дал хороших результатов – хотя погрешности определения ПГ и составляли 0,5–5 % отн., но погрешности определения Дб во всех модельных смесях превышали 10-15 % отн. Исключение неаддитивного участка спектра 235-250 нм, приходящегося на область максимального поглощения ПГ, позволило снизить погрешности определения дибазола, до 2-7 % отн., но при этом в несколько раз, до 10 % отн., возросли погрешности определения ПГ, особенно в смесях с большим избытком Дб (от 1 : 6 до 1:20).

Таким образом, переход к анализу смесей методом МЛР с применением тех же коэффициентов позволил расширить интервал соотношений, при которых возможно одновременно определять ПГ и Дб с погрешностями, не превышающими 5 % отн., до соотношения 1 : 5, но дальнейшее увеличение избытка как ПГ, так и Дб, по-прежнему приводило к резкому возрастанию погрешностей.

Следующим этапом стало вычисление коэффициентов поглощения ПГ и Дб методом МЛР по спектрам их смесей. Для этого использовали пять модельных смесей, в которых был представлен достаточно широкий размах соотношений ПГ : Дб, от 3 : 1 до 1 : 10. Остальные 18 смесей с такими же соотношениями, но другими концентрациями, выступали в качестве проверочных.

На рис. 1 приведены спектры поглощения ПГ и Дб, а на рис. 2 - зависимости коэффициентов поглощения, вычисленных методом МЛР, от длины волны. Как видно, в обоих случаях характер кривых идентичен, но в области 235-250 нм, где отмечается неаддитивность оптической плотности в спектрах смесей, коэффициенты поглощения дибазола, вычисленные разными способами, отличаются между собой.

Повторный анализ смесей ПГ и Дб с применением новых коэффициентов привел к следующим положительным результатам: во-первых, снизились погрешности определения компонентов; во-вторых, одновременное определение компонентов с погрешностью не более 5 % отн. стало возможным при любых соотношениях ПГ и Дб, от 3 : 1 до 1 : 20; в-третьих, при анализе методом МЛР не потребовалась дополнительная оптимизация спектрального диапазона путем исключения неаддитивного участка спектра. В качестве примера результаты анализа 6 смесей

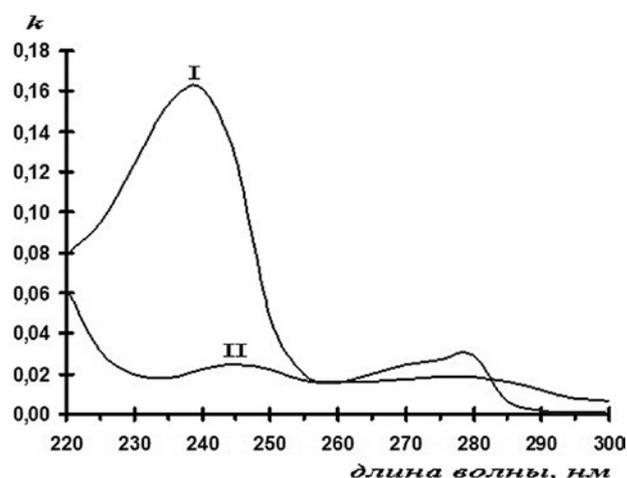


Рис. 1. Спектры поглощения папаверина (I) и дибазола (II)

ПГ и Дб из числа проверочных (их состав указан в табл. 1), полученные с применением традиционных и новых коэффициентов, приведены в табл. 2. Относительные погрешности здесь и далее приведены по модулю.

Как видно, предложенный способ вычисления коэффициентов позволяет не только расширить диапазон определяемых концентраций, но и повысить точность анализа смесей ПГ и Дб как методом Фирордта, так и методом МЛР: обобщенная погрешность ζ снижается в том и другом случае.

Аналогичным образом методом МЛР были проанализированы четырехкомпонентные смеси витаминов, сначала с применением коэффициентов, вычисленных по спектрам индивидуальных соединений. Спектры смесей витаминов были аддитивны, но для снижения погрешностей определения требовалась дополнительная оптимизация спектральных диапазонов: для каждого витамина был выбран свой участок спектра. Однако даже после этого погрешности определения витамина В₁ в отдельных смесях

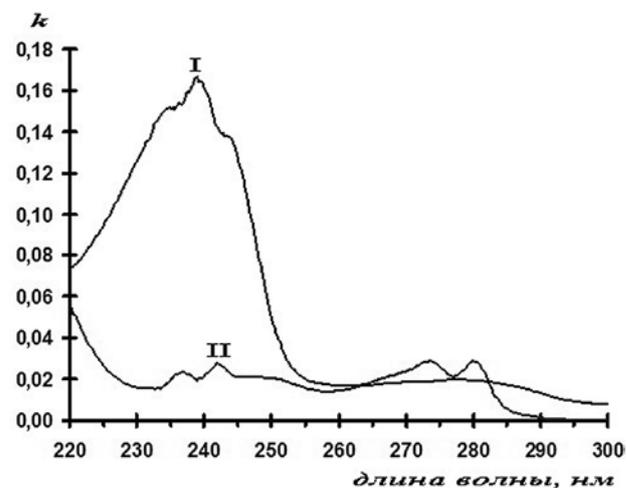


Рис. 2. Зависимость от длины волны коэффициентов поглощения папаверина (I) и дибазола (II), вычисленных методом МЛР

Таблица 1

Составы некоторых модельных смесей, мкг/мл

2-компонентные смеси			4-компонентные смеси				
№ смеси	ПГ	Дб	№ смеси	B ₁	B ₂	B ₅	B ₆
1	7,2	2,4	7	1,0	1,2	7,0	1,7
2	4,8	2,4	8	1,2	2,3	7,1	1,0
3	3,6	3,6	9	1,8	4,0	5,6	2,0
4	2,4	4,8	10	1,4	1,6	7,1	1,0
5	1,2	4,8	11	1,2	3,9	14,2	2,9
6	2,4	24,0	12	1,0	3,2	8,0	2,8

Таблица 2

Относительные погрешности (δ , % отн.) определения папаверина и дибазола в смесях с применением коэффициентов поглощения, вычисленных двумя способами

Параметр		1 способ*				2 способ**			
		МФ		МЛР		МФ		МЛР	
		$\delta_{ПГ}$	$\delta_{Дб}$	$\delta_{ПГ}$	$\delta_{Дб}$	$\delta_{ПГ}$	$\delta_{Дб}$	$\delta_{ПГ}$	$\delta_{Дб}$
№ смеси	1	0,1	7,9	2,0	7,1	0,2	1,9	0,8	0,4
	2	1,0	9,0	0,6	5,1	0,2	2,7	0,7	1,0
	3	0,6	2,8	3,8	5,1	0,2	2,5	0,6	0,8
	4	4,0	1,3	2,1	0,7	1,9	3,0	1,5	0,5
	5	5,4	1,0	1,7	0,3	1,9	3,3	1,2	1,6
	6	1,1	5,6	8,5	2,5	1,5	1,6	2,4	1,1
δ %, средние по 6 смесям		2,0	4,6	3,1	3,5	1,0	2,5	1,2	0,9
ξ		5,0		4,7		2,7		1,5	

Примечание: *1 способ - коэффициенты вычислены по спектрам индивидуальных соединений, **2 способ - коэффициенты вычислены по спектрам смесей методом МЛР

Таблица 3

Относительные погрешности (δ , % отн.) определения витаминов в смесях методом МЛР

Параметр		1 способ*				2 способ**			
		δ_{B1}	δ_{B2}	δ_{B5}	δ_{B6}	δ_{B1}	δ_{B2}	δ_{B5}	δ_{B6}
№ смеси	7	16,5	6,5	4,3	17,8	5,6	2,0	3,5	2,2
	8	5,5	11,0	8,4	0,9	0,8	7,1	1,3	6,6
	9	6,8	7,0	6,3	2,2	2,8	0,5	4,8	2,6
	10	9,0	8,2	5,0	10,7	1,4	2,7	0,5	2,5
	11	10,2	2,4	6,9	5,7	9,2	1,7	0,0	0,6
	12	7,2	4,9	6,6	0,7	8,0	4,9	1,6	6,8
δ %, средние по 6 смесям		9,2	6,7	6,2	6,3	4,6	3,2	2,0	3,6
ξ		14,4				6,9			

Примечание: *1 способ - коэффициенты вычислены по спектрам индивидуальных соединений, **2 способ - коэффициенты вычислены по спектрам смесей методом МЛР

превышали 10 % отн. Для вычисления новых коэффициентов витаминов методом МЛР использовали 9 модельных смесей (назовем их обучающей выборкой), остальные 16 смесей выступали в качестве проверочных, состав некоторых из них приведен в табл. 1, а результаты их анализа – в табл.3. Как и в случае бинарных смесей, при использовании новых коэффициентов отмечается снижение погрешностей определения для всех компонентов. Приведенные в табл. 3 значения средних по 6 смесям погрешностей определения каждого витамина, а также обобщенная погрешность убедительно свидетельствуют о повышении точности анализа в случае использования новых коэффициентов. Все расчеты при этом выполняли в одном спектральном диапазоне, без дополнительной его оптимизации для каждого компонента. Это существенно сократило время анализа.

Таким образом, применение метода МЛР для вычисления коэффициентов является перспективным направлением развития спектрофотометрического анализа многокомпонентных смесей. Заметим, однако, что вопрос, какой при этом должна быть обучающая выборка – по числу смесей и их составу – является не простым и требует дальнейшего серьезного изучения.

Литература

1. Сутормина Е.Ф. Спектрофотометрическое определение аммиака и оксидов азота в продуктах реакции окисления аммиака // Ж. аналит. химии. 2004. Т. 9, № 4. С. 373-376.
2. Erdal D. Linear regression analysis and its application to the multivariate spectral calibrations for the multiresolution of a ternary mixture of caffeine,

paracetamol and metamizol in tablets // J. of Pharmaceutical and Biomedical Analysis. 2003. V. 33. P. 605-615.

3. Масыкова Е.Н., Шелпакова А.С., Феллер А.В., Власова И.В. Определение водорастворимых витаминов группы В в неразделенных смесях // Разработка, исследование и маркетинг новой фармацевтической продукции. Сб. научн. тр. Вып. 63. Пятигорск: Изд-во Пятигорская фармацевтическая академия. 2008. С. 303-304.
4. Власова И.В., Шелпакова А.С., Масыкова Е.Н. Спектрофотометрический анализ смесей витаминов с применением метода множественной линейной регрессии // Аналитика и контроль. 2009. Т. 13, № 2. С. 86-90.
5. Шараф М.А., Иллман Д.Л., Ковальски Б.Р. Хемометрика. Л.: Химия, 1989. 270 с.
6. Отто М. Современные методы аналитической химии (в 2-х томах). Т. 2. М.: Техносфера. 2004. 281 с.
7. Берштейн И.Я., Каминский Ю.А. Спектрофотометрический анализ в органической химии. Л.: Химия, 1986. 200 с.
8. Власова И.В., Шилова А.В. Новые подходы к спектрофотометрическому анализу многокомпонентных смесей // Вісник Харківського національного університету. 2007. № 770. Хімія. Вип. 15(38). С.14.
9. Власова И.В., Вершинин В.И. Возможность определения компонентов бинарных смесей методом Фирордта с погрешностями, не превышающими заданный предел // Ж. аналит. химии. 2009. Т. 64. № 6. С.571.

NEW APPROACH TO THE APPLICATION OF MULTIPLE LINEAR REGRESSIONS IN THE SPECTROPHOTOMETRIC ANALYSIS OF MULTICOMPONENT MIXTURES

I.V. Vlasova, A.S. Shelpakova, S.M. Dobrovol'skiy, A.V. Fisenko

*Omsk state university it. F.M.Dostoevskogo
vlaso-iri@yandex.ru*

The method of multiple linear regressions is applied for the calculation of absorption factors on the spectrums of mixtures. The use of new coefficients allows to promote adequacy of the spectrophotometric analysis of undivided mixtures, and to shorten time on the realization of analysis, as it is not required to conduct additional optimization of spectral ranges for every component.

Keywords: spectrophotometric analysis, multiple linear regression, analysis of multicomponent mixtures.