

УДК 543.4

В. С. МУХА

ПРОВЕРКА ГИПОТЕЗ В ПАРАЛЛЕЛЬНОМ ФАКТОРНОМ АНАЛИЗЕ

Белорусский государственный университет информатики и радиоэлектроники

(Поступила в редакцию 25.11.2014)

Введение. Параллельный факторный анализ (PARAFAC) достаточно активно развивается в хеометрике. Математическая модель измерений в PARAFAC имеет следующий вид [1, 2]:

$$g_{a,i,j,k} = \sum_{f=1}^{n_f} a_{i,f} b_{j,f} c_{k,f} + v_{i,j,k}, \quad i = \overline{1, n_a}, \quad j = \overline{1, n_b}, \quad k = \overline{1, n_c}. \quad (1)$$

Эта модель используется в задаче флуоресценции растворов веществ [1, 2]. Предполагается, что раствор (образец) содержит n_f веществ и анализируется с помощью спектрофлуорометра. Переменная $b_{j,f}$ в (1) интерпретируется как относительная эмиссия света единицей f -го вещества на j -й длине волны излучения (emission wavelenght), $c_{k,f}$ – как относительная эмиссия света единицей f -го вещества на k -й длине волны возбуждения (excitation wavelenght), и $a_{i,f}$ означает концентрацию f -го вещества в i -м растворе. Переменная $g_{a,i,j,k}$ означает интенсивность эмиссии света i -м раствором на j -й длине волны при его возбуждении на k -й длине волны, а переменная $v_{i,j,k}$ – шум при измерениях. Переменные модели (1) можно считать элементами матриц: $A = (a_{i,f})$, $B = (b_{j,f})$, $C = (c_{k,f})$, $G_a = (g_{a,i,j,k})$, $V = (v_{i,j,k})$, $i = \overline{1, n_a}$, $j = \overline{1, n_b}$, $k = \overline{1, n_c}$, $f = \overline{1, n_f}$, f -й столбец матрицы B как функция длины волны излучения называется спектром эмиссии вещества f , а f -й столбец матрицы C как функция длины волны возбуждения – спектром возбуждения вещества f .

Реалистичной и практически важной представляется задача получения оценки матрицы концентраций A по имеющимся матрицам G_a, B, C . Матрицы B, C могут быть известны, так как можно создать банк спектров различных веществ. В таком виде PARAFAC-задача является регрессионной с неизвестным матричным параметром A . Она сформулирована и решена в работах [3, 4]. Модель измерений (1) представляется в многомерно-матричной форме [5], для чего формируется трехмерная матрица M_a ,

$$M_a = (b_{j,f} c_{k,f}) = (m_{a,f,j,k}), \quad f = \overline{1, n_f}, \quad j = \overline{1, n_b}, \quad k = \overline{1, n_c},$$

и модель (1) записывается в виде

$$G_a = (g_{a,i,j,k}) = \left(\sum_{f=1}^{n_f} a_{i,f} m_{a,f,j,k} \right) + (v_{i,j,k}) = {}^{0,1}(AM_a) + V, \quad i = \overline{1, n_a}, \quad j = \overline{1, n_b}, \quad k = \overline{1, n_c}. \quad (2)$$

Оценка \hat{A} матрицы концентраций веществ A модели (2) отыскивается методом наименьших квадратов, т. е. из критерия

$$F = {}^{0,3}(G_a - {}^{0,1}(AM_a))^2 \rightarrow \min_A,$$

и определяется выражением

$$\hat{A} = {}^{0,1}({}^{0,2}(G_a M_a^{H_{3,1}}) {}^{0,1}({}^{0,2}(M_a M_a^{H_{3,1}}))^{-1}). \quad (3)$$

Здесь $M_a^{H_{3,1}}$ означает матрицу M_a , транспонированную соответственно подстановке $H_{3,1}$ [5], а $a^{0,1} ({}^{0,2}(M_a M_a^{H_{3,1}}))^{-1}$ – матрицу, (0,1)-обратную к матрице ${}^{0,2}(M_a M_a^{H_{3,1}})$ [5].

Дальнейший анализ модели измерений (2) может быть выполнен стандартными для регрессионного анализа приемами: проверка значимости параметров (концентраций) и отсеивание незначимых, проверка адекватности модели и ее замена в случае неадекватности путем включения в нее новых веществ [6, 7]. Указанные процедуры могут быть выполнены при известных законах распределения интересующих нас статистик. Этот вопрос также разработан в работе [3]. Ниже предлагается иной подход – рассмотрение задачи как частного случая задачи многомерно-матричного линейного регрессионного анализа [8]. При этом будут получены эквивалентные по содержанию, но иные по форме результаты.

Распределения статистик и проверка гипотез в PARAFAC. Модель измерений (2) допускает иное представление. Введем в рассмотрение одномерные матрицы $g_a = (g_{a,i})$, $v = (v_i)$, $i = \overline{1, n_a}$, $m_a = (m_{a,f})$, $f = \overline{1, n_f}$, зафиксировав мультииндекс $\mu = (j, k)$ в матрицах G_a , V , M_a . Матрицы g_a , v и m_a являются, таким образом, сечениями ориентации $\mu = (j, k)$ матриц G_a , V , M_a соответственно, т. е. $G_a = ((g_{a,i})_\mu) = (g_{a,\mu})$, $V = ((v_i)_\mu) = (v_\mu)$, $M_a = ((m_{a,f})_\mu) = (m_{a,\mu})$. Здесь и далее $\mu = (j, k)$ – мультииндекс, пробегающий значения от 1 до $n = n_b n_c$. В этих новых обозначениях модель измерений (2) приобретает вид

$$g_{a,\mu} = {}^{0,1}(Am_{a,\mu}) + v_\mu, \mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c}. \quad (4)$$

При этом, конечно, предполагается наличие гипотетической модели данных вида

$$g = {}^{0,1}(Am_a) = (g_i), i = \overline{1, n_a}, \quad (5)$$

для заданных пар данных $(m_{a,\mu}, g_\mu)$, т. е.

$$g_\mu = {}^{0,1}(Am_{a,\mu}), \mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c}.$$

Получение матрицы данных $G_a = (g_{a,\mu})$ будем называть одним опытом, а фиксированное значение мультииндекса $\mu = (j, k)$ – отдельной точкой опыта.

Модель вида (4) представляет собой частный случай модели измерений (p, q) -мерно-матричного линейного регрессионного анализа [8] с p -мерной выходной и q -мерной входной матрицами соответственно, а именно, модель (1,1)-мерно-матричного линейного регрессионного анализа.

В соответствии с работой [8], оценка \hat{A} матрицы A модели измерений (4) определяется выражением

$$\hat{A} = (\hat{a}_{i,f}) = {}^{0,1}(s_{gm} s_{m^2}^{-1}), i = \overline{1, n_a}, f = \overline{1, n_f}, \quad (6)$$

где

$$s_{gm} = \frac{1}{n_b n_c} \sum_{\mu=1}^{n_b n_c} {}^{0,0}(g_{a,\mu} m_{a,\mu}),$$

$$s_{m^2} = \frac{1}{n_b n_c} \sum_{\mu=1}^{n_b n_c} {}^{0,0}(m_{a,\mu} m_{a,\mu}),$$

и $s_{m^2}^{-1}$ – матрица, (0,1)-обратная к матрице s_{m^2} . Легко убедиться в том, что $ns_{gm} = {}^{0,2}(G_a M_a^{H_{3,1}})$ и $ns_{m^2} = {}^{0,2}(M_a M_a^{H_{3,1}})$, так что оценка (6) совпадает с оценкой (3). Оценка \hat{g} отклика g в (5) определяется выражением

$$\hat{g} = {}^{0,1}(\hat{A}m_a) = (\hat{g}_i), i = \overline{1, n_a}. \quad (7)$$

В связи со сказанным дальнейшие результаты являются частным случаем и интерпретацией результатов работы [8] и приводятся без доказательств.

Теорема 1. Если матрицы ошибок измерений v_μ в PARAFAC-модели измерений (4) имеют математическое ожидание $E(v_\mu) = 0$ и дисперсионную матрицу $D(v_\mu) = \sigma^2 E(0,1)$, где $E(0,1)$ – $(0,1)$ -единичная $(n_a \times n_a)$ -матрица (обычная единичная матрица), σ^2 – скалярный параметр, имеющий физический смысл дисперсии ошибок измерений, и независимы по μ , то математическое ожидание оценки \hat{A} определяется выражением

$$E(\hat{A}) = A, \quad (8)$$

а дисперсионная матрица – выражением

$$D(\hat{A}) = E^{(0,0)}((\hat{A} - E(\hat{A}))(\hat{A} - E(\hat{A}))) = \frac{\sigma^2}{n_b n_c} D^{(a)}, \quad (9)$$

где

$$D^{(a)} = (d_{i,j,i',j'}^{(a)}) = D^{(H_{2,1}, E_2)}, \quad i, i' = \overline{1, n_a}, \quad j, j' = \overline{1, n_f},$$

$$D = (d_{j,i,i',j'}) = \begin{cases} s^{-1} & i = i'; \\ m^2 & i, i' = \overline{1, n_a}, \quad j, j' = \overline{1, n_f}; \\ 0 & i \neq i'; \end{cases}$$

и 0 – нулевая $(n_f \times n_f)$ -матрица. Математическое ожидание оценки отклика (7) определяется выражением

$$E(\hat{g}) = {}^{0,1}(Am_a) = g, \quad (10)$$

а дисперсионная матрица – выражением

$$D(\hat{g}) = E^{(0,0)}((\hat{g} - E(\hat{g}))(\hat{g} - E(\hat{g}))) = \frac{\sigma^2}{n_b n_c} D^{(g)}, \quad (11)$$

где

$$D^{(g)} = {}^{0,1}(m_a \quad {}^{0,1}(Dm_a)) = (d_{i,i'}^{(g)}), \quad i, i' = \overline{1, n_a}.$$

Если, кроме того, ошибки измерений v_μ нормально распределены с математическим ожиданием $E(v_\mu) = 0$ и дисперсионной матрицей $D(v_\mu) = \sigma^2 E(0,1)$, то оценка \hat{A} имеет нормальное распределение с математическим ожиданием (8) и дисперсионной матрицей (9), и оценка \hat{g}_μ имеет нормальное распределение с математическим ожиданием (10) и дисперсионной матрицей (11).

Дисперсия σ^2 ошибок измерений v_μ в модели измерений (4) чаще всего неизвестна, и по результатам измерений требуется получить ее оценку $\hat{\sigma}^2$. Такой оценкой может служить статистика

$$\hat{\sigma}^2 = \frac{1}{n_a n_b n_c} \text{tr} \left(\sum_{\mu=1}^{n_b n_c} {}^{0,0}(g_{a,\mu} - \hat{g}_{a,\mu})^2 \right). \quad (12)$$

Теорема 2. В условиях теоремы 1 (кроме условия нормальности ошибок измерений) оценка $\hat{\sigma}^2$ (12) параметра σ^2 является асимптотически несмещенной, а оценка

$$\hat{\sigma}_1^2 = \frac{n_a n_b n_c \hat{\sigma}^2}{n_a n_b n_c - n_a n_f} = \frac{1}{n_a n_b n_c - n_a n_f} \text{tr} \left(\sum_{\mu=1}^{n_b n_c} {}^{0,0}(g_{a,\mu} - \hat{g}_{a,\mu})^2 \right)$$

несмещенной. Если, кроме того, ошибки измерений v_μ распределены по нормальному закону $N(0, \sigma^2 E(0,1))$, то статистика

$$v = \frac{n_a n_b n_c \hat{\sigma}^2}{\sigma^2} = \frac{(n_a n_b n_c - n_a n_f) \hat{\sigma}_1^2}{\sigma^2} \quad (13)$$

имеет распределение хи-квадрат с $n_a n_b n_c - n_a n_f$ степенями свободы. Оценки $\hat{\sigma}^2$ и \hat{A} , а также $\hat{\sigma}_1^2$ и \hat{A} независимы.

З а м е ч а н и е 1. В правой части равенства (13) работы [8], на основе которого записано равенство (13) данной работы, множитель n должен отсутствовать.

Рассмотрим теперь некоторые статистики, связанные с полученными оценками. На основании теорем 1 и 2 можно сделать вывод, что статистики

$$u_{i,f}^{(a)} = \frac{\widehat{a}_{i,f} - a_{i,f}}{\sigma \sqrt{d_{i,j,i,j}^{(a)}}} \sqrt{n_b n_c}, \quad i = \overline{1, n_a}, \quad f = \overline{1, n_f},$$

$$u_i^{(g)} = \frac{\widehat{g}_i - g_i}{\sigma \sqrt{d_{i,i}^{(g)}}} \sqrt{n_b n_c}, \quad i = \overline{1, n_a},$$

распределены по нормальному закону $N(0,1)$, а статистики

$$t_{i,f}^{(a)} = \frac{\widehat{a}_{i,f} - a_{i,f}}{\widehat{\sigma}_1 \sqrt{d_{i,f,i,f}^{(a)}}} \sqrt{n_b n_c}, \quad i = \overline{1, n_a}, \quad f = \overline{1, n_f}, \quad (14)$$

$$t_i^{(g)} = \frac{\widehat{g}_i - g_i}{\widehat{\sigma}_1 \sqrt{d_{i,i}^{(g)}}} \sqrt{n_b n_c}, \quad i = \overline{1, n_a}, \quad (15)$$

имеют распределение Стьюдента с $n_a n_b n_c - n_a n_f$ степенями свободы. Следует иметь в виду, что статистики $u_{i,f}^{(a)}$ и $t_{i,f}^{(g)}$ рассчитываются для любых заданных значений $m_{a,\mu}$ вектора m_a , $\mu = (j, k)$, $j = \overline{1, n_b}$, $k = \overline{1, n_c}$.

З а м е ч а н и е 2. Правая часть равенства (15) работы [8], на основе которого записано равенство (15) данной работы, должна быть умножена на \sqrt{n} .

Полученные выше статистики и их распределения позволяют строить доверительные интервалы и проверять гипотезы для PARAFAC-модели аналогично тому, как это делается в классическом регрессионном анализе [6, 7]. В частности, статистики $t_{i,f}^{(a)}$ (14) позволяют проверять гипотезы вида

$$\{H_0 : a_{i,f} = 0; H_1 : a_{i,f} \neq 0\}, \quad (16)$$

которые называются гипотезами о значимости элементов $a_{i,f}$ матрицы концентраций веществ $A = (a_{i,f})$. Если эмпирическое значение $t_{i,f,\alpha}^{(a)}$ статистики $t_{i,f}^{(a)}$ удовлетворяет условию

$$|t_{i,f,\alpha}^{(a)}| \leq t_{i,f,\alpha/2}^{(a)},$$

где $t_{i,f,\alpha/2}^{(a)} - 100\alpha / 2$ -процентная точка распределения Стьюдента с $n_a n_b n_c - n_a n_f$ степенями свободы, то гипотеза $H_0 : a_{i,f} = 0$ принимается, т. е. считается, что вещество f в i -м растворе отсутствует.

Статистики $t_{i,f}^{(a)}$ позволяют строить доверительные интервалы для элементов $a_{i,f}$ матрицы концентраций веществ $A = (a_{i,f})$. Эти интервалы с доверительной вероятностью γ имеют следующий вид:

$$\widehat{a}_{i,f} - t_{i,f,(1-\gamma)/2}^{(a)} \frac{\widehat{\sigma}_1 \sqrt{d_{i,f,i,f}^{(a)}}}{\sqrt{n_b n_c}} \leq a_{i,f} \leq \widehat{a}_{i,f} + t_{i,f,(1-\gamma)/2}^{(a)} \frac{\widehat{\sigma}_1 \sqrt{d_{i,f,i,f}^{(a)}}}{\sqrt{n_b n_c}}, \quad i = \overline{1, n_a}, \quad f = \overline{1, n_f},$$

где $t_{i,f,(1-\gamma)/2}^{(a)} - 100(1-\gamma) / 2$ -процентная точка распределения Стьюдента с $n_a n_b n_c - n_a n_f$ степенями свободы.

Доверительный интервал для значений функции отклика g_i строится на основе статистики $t_i^{(g)}$ и имеет вид

$$\widehat{g}_i - t_{i,(1-\gamma)/2}^{(g)} \frac{\widehat{\sigma}_1 \sqrt{d_{i,i}^{(g)}}}{\sqrt{n_b n_c}} \leq g_i \leq \widehat{g}_i + t_{i,(1-\gamma)/2}^{(g)} \frac{\widehat{\sigma}_1 \sqrt{d_{i,i}^{(g)}}}{\sqrt{n_b n_c}}, \quad i = \overline{1, n_a},$$

где $t_{i,(1-\gamma)/2}^{(g)}$ – 100(1 – γ) / 2-процентная точка распределения Стьюдента с $n_a n_b n_c - n_a n_f$ степенями свободы, γ – доверительная вероятность.

Важным вопросом является проверка адекватности применяемой математической модели. Если модель не адекватна, то полученные выводы об оценках концентраций веществ будут ошибочными. Выполним $n_\delta \geq 2$ независимых дополнительных опытов, т. е. получим еще n_δ матриц данных G_a . По этим данным можно получить матрицу несмещенных оценок $\hat{\sigma}_{1,\delta}^2 = (\hat{\sigma}_{1,i,\mu,\delta}^2)$ дисперсии σ_0^2 ошибок объекта (прибора с растворами) по формуле

$$\hat{\sigma}_{1,i,\mu,\delta}^2 = \frac{1}{n_\delta - 1} \sum_{k=1}^{n_\delta} (g_{a,i,\mu} - \bar{g}_{a,i,\mu})^2, \quad i = \overline{1, n_a}, \mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c},$$

где $\bar{g}_{a,i,\mu} = \frac{1}{n_\delta} \sum_{k=1}^{n_\delta} g_{a,i,\mu,k}$. Понятно, что статистика

$$w_{i,\mu} = \frac{(n_\delta - 1) \hat{\sigma}_{1,i,\mu,\delta}^2}{\sigma_0^2}, \quad i = \overline{1, n_a}, \mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c},$$

имеет распределение хи-квадрат с $n_\delta - 1$ степенями свободы. Статистики v и $w_{i,\mu}$ независимы.

Гипотеза об адекватности модели (однородности дисперсии) имеет вид $H_0 : \sigma^2 = \sigma_0^2$ (против альтернативы $H_1 : \sigma^2 > \sigma_0^2$) [7]. Если H_0 верна, то статистика

$$F_{i,\mu} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_{1,i,\mu,\delta}^2}, \quad i = \overline{1, n_a}, \mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c},$$

имеет распределение Фишера с $n_a n_b n_c - n_a n_f, n_\delta - 1$ степенями свободы. Гипотезу $H_0 : \sigma^2 = \sigma_0^2$ можно проверить следующим образом. Если эмпирическое значение $F_{\alpha,i,\mu}$ статистики $F_{i,\mu}$ удовлетворяет неравенству $F_{\alpha,i,\mu} \geq F_\alpha$, где F_α – 100 α -процентная точка распределения Фишера с $n_a n_b n_c - n_a n_f, n_\delta - 1$ степенями свободы, то гипотеза об адекватности модели отклоняется (модель признается не адекватной) на уровне α для одного раствора $i = \overline{1, n_a}$ и одной точки опыта $\mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c}$. Модель признается не адекватной в целом, если она не адекватна хотя бы для одного раствора. Можно также посчитать число отклонений гипотезы согласно условию $F_{\alpha,i,\mu} \geq F_\alpha$ для всех растворов $i = \overline{1, n_a}$ и всех точек опыта $\mu = (j, k), j = \overline{1, n_b}, k = \overline{1, n_c}$. Если процент этих отклонений (от общего числа проверок) будет меньше некоторого малого числа (например, 5 %), то модель можно признать адекватной.

Компьютерное моделирование. Продолжим анализ примера из работы [4], где моделировались данные, подобные флуоресценции растворов аминокислот в [1, 2], полученной на спектрофлуорометре PE LS50B. Рассмотрим четыре вещества, спектры флуоресценции которых представлены на рис. 1.

В приборе анализируются четыре раствора (образца) ($n_a = 4$), концентрации веществ в которых определяются матрицей концентраций

$$A = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 4 & 4 \\ 4 & 1 & 4 \\ 4 & 3 & 5 \end{pmatrix}. \quad (17)$$

Столбцы матрицы соответствуют первым трем веществам рис. 1, а строки – растворам. Мы видим, что в растворах 1 и 2 вещество 1 отсутствует, а в растворах 3 и 4 присутствуют все три вещества. Изображения флуоресценции четырех растворов при дисперсии ошибок прибора $\sigma_0^2 = 1,0e - 08$ ($\sigma_0 = 0,0001$) представлены на рис. 2.

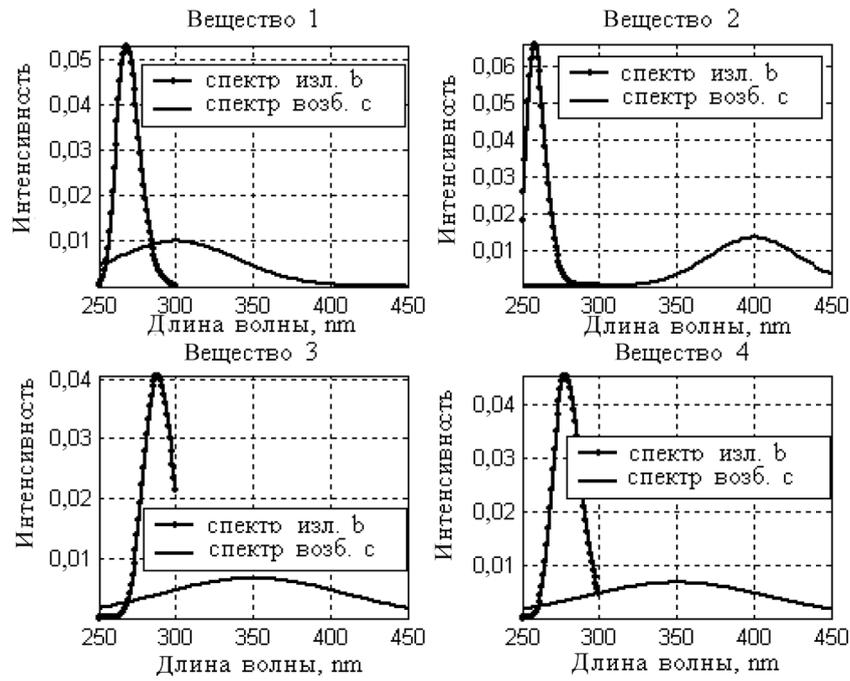


Рис. 1. Спектры флуоресценции четырех веществ

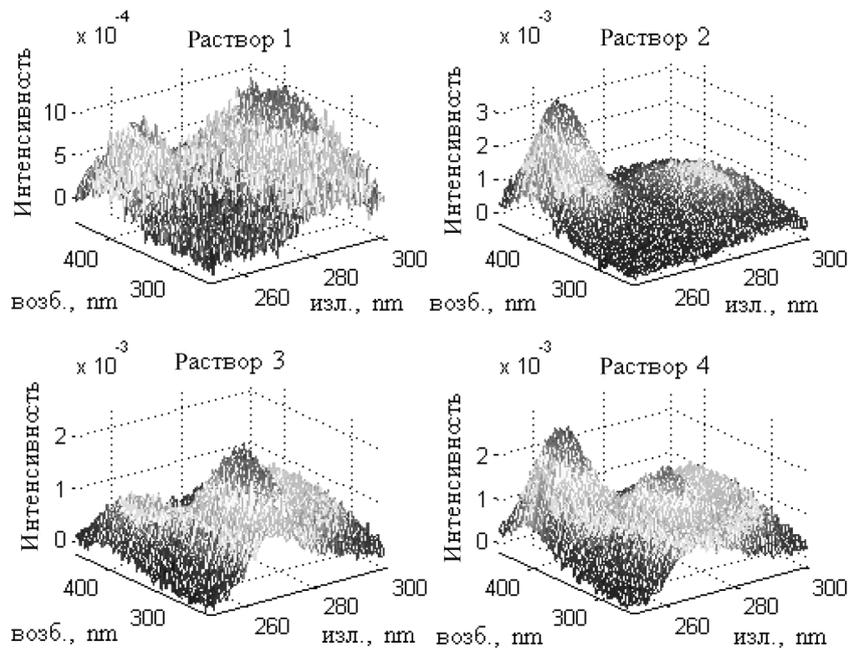


Рис. 2. Изображения флуоресценции четырех растворов

Данные, приведенные на рис. 2, содержатся в массиве G_a размером $4 \times 51 \times 101$. В математической модели (1) использовались спектры тех же веществ, что и в растворах. Один из опытов дает оценку \hat{A} (3) матрицы концентраций веществ в растворах:

$$\hat{A} = \begin{pmatrix} 0,1328 & 0,9785 & 3,9520 \\ -0,2058 & 4,0191 & 4,0789 \\ 3,9581 & 1,2736 & 4,1066 \\ 3,7225 & 3,2155 & 4,8616 \end{pmatrix}, \quad \hat{A}_1 = \begin{pmatrix} 0,0000 & 0,9785 & 3,9520 \\ 0,0000 & 4,0191 & 4,0789 \\ 3,9581 & 1,2736 & 4,1066 \\ 3,7225 & 3,2155 & 4,8616 \end{pmatrix}. \quad (18)$$

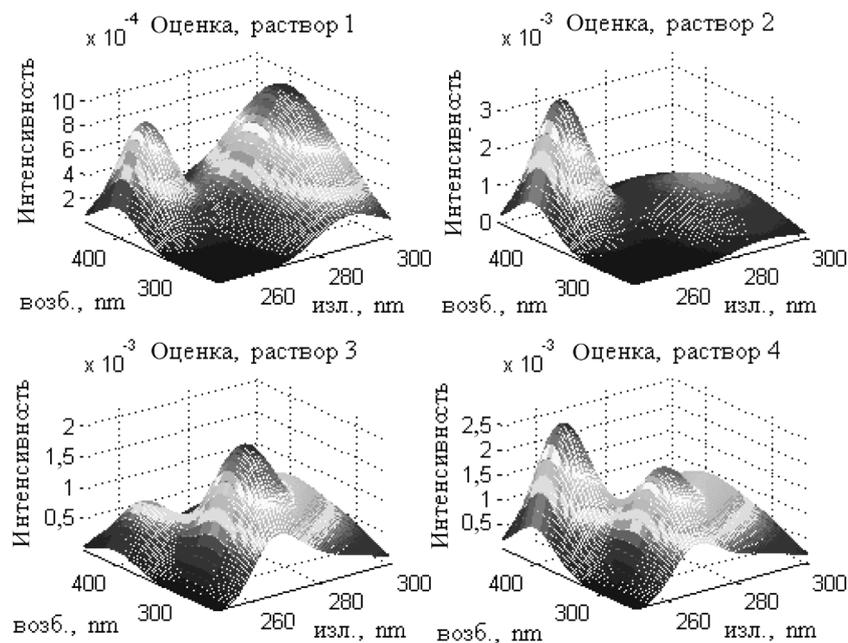


Рис. 3. Изображения оценок флуоресценции четырех растворов

Оценка \hat{A} свидетельствует, что каждое вещество в определенной концентрации присутствует в каждом из растворов, хотя понятно, что отрицательных концентраций быть не может. Однако проверка гипотез (16) о значимости концентраций на основе статистик $t_{i,f}^{(a)}$ (14) на уровне значимости $\alpha = 0,05$ показывает, что концентрации $\hat{a}_{1,1} = 0,1328$ и $\hat{a}_{2,1} = -0,2058$ не значимы, т. е. в растворах 1 и 2 вещество 1 отсутствует. Проверка гипотезы об адекватности математической модели на основе предложенной выше процедуры при двух дополнительных опытах ($n_d = 2$) показывает, что модель адекватна. В результате мы можем считать, что концентрации веществ в растворах определяются матрицей \hat{A}_1 (18), которая близка к матрице истинных концентраций A (17). На рис. 3 изображены оценки (7) флуоресценции, полученные с использованием оценки \hat{A} (18). По числу максимумов можно заключить, что растворы 1 и 2 содержат по два вещества, а растворы 3, 4 – по три.

Выполнялись также расчеты для измененной математической модели, в которой вместо спектров веществ 1, 2, 3 использовались спектры веществ 1, 2, 4 (см. рис. 1). В одном из опытов получена оценка \hat{A} матрицы концентраций A , представленная ниже:

$$\hat{A} = \begin{pmatrix} -0,7262 & 0,8700 & 3,3343 \\ -0,7212 & 3,8790 & 3,3153 \\ 3,2782 & 0,8786 & 3,3122 \\ 3,1060 & 2,8250 & 4,1602 \end{pmatrix}, \quad \hat{A}_2 = \begin{pmatrix} 0 & 0,8700 & 3,3343 \\ 0 & 3,8790 & 3,3153 \\ 3,2782 & 0,8786 & 3,3122 \\ 3,1060 & 2,8250 & 4,1602 \end{pmatrix}. \quad (19)$$

Проверка гипотез о значимости концентраций приводит к тому, что концентрации $a_{1,1}$ и $a_{2,1}$ признаются не значимыми, так что мы получаем оценку \hat{A}_2 (19). Однако проверка гипотезы об адекватности данной математической модели показала, что она не адекватна. Это значит, что матрице оценок \hat{A}_2 (19) доверять не следует. Действительно, сравнив матрицы A (17) и \hat{A}_2 (19), мы обнаружим в них значительные отличия. В случае неадекватности в математическую модель можно включить новый набор веществ и выполнить расчеты и анализ повторно.

Заключение. Представленные в работе результаты (статистики, их распределения, критерии) могут быть использованы на практике либо непосредственно, либо в адаптированном под конкретные оптические приборы виде. Они являются также дополнительным подтверждением правильности и работоспособности результатов работы [8], а также примером использования многомерного регрессионного анализа в практических приложениях.

Літэратура

1. Bro R. // Chemometrics and Intelligent Laboratory Systems. 1997. Vol. 38. P. 149–171.
2. Ji Ji R. D., Andersson G. G., Booksh K. S. // J. of Chemometrics. 2000. Vol. 14. P. 171–185.
3. Муха В. С. // Проблемы управления и информатики. 2008. № 4. С. 100–114.
4. Муха В. С. // Весці НАН Беларусі. Сер. фіз.-мат. навук. 2007. № 3. С. 131–134.
5. Муха В. С. Анализ многомерных данных. Минск, 2004.
6. Вучков И. Н., Бояджиева Л., Солаков Е. Прикладной регрессионный анализ. М., 1987.
7. Муха В. С. Статистические методы обработки данных. Минск, 2009.
8. Муха В. С. // Весці НАН Беларусі. Сер. фіз.-мат. навук. 2014. № 2. С. 71–81.

V. S. MUKHA

CHECKING THE HYPOTHESES IN THE PARALLEL FACTOR ANALYSIS

Summary

Statistics, their distributions and criteria for building confidence intervals and checking the hypotheses in parallel factor analysis are presented. Their use is illustrated by the example of computer simulation.