© «Вестник ИГЭУ» Вып. 1 2011 г.

УДК 621.311.18

СОВЕРШЕНСТВОВАНИЕ МЕТОДА ХРОМАТОГРАФИЧЕСКОЙ ДИАГНОСТИКИ РАЗВИВАЮЩИХСЯ ДЕФЕКТОВ В ЭЛЕКТРООБОРУДОВАНИИ НА ОСНОВЕ ТЕОРИИ ФАКТОРНОГО АНАЛИЗА

А.Н. НАЗАРЫЧЕВ, д-р техн. наук, И.Ю. ЗЕЛЕНЦОВ, инж.

Разработан подход по совершенствованию метода хроматографической диагностики развивающихся дефектов в электрооборудовании на основе теории факторного анализа, позволяющий делать более глубокий и достоверный прогноз о наличии комбинированных дефектов без проведения дополнительных испытаний на основе уже имеющихся данных. Для силовых трансформаторов приведен пример применения факторного анализа для оценки его технического состояния.

Ключевые слова: комбинированные дефекты, матрица факторных нагрузок, силовой трансформатор, матрица численных значений факторов, коэффициент корреляции, общая дисперсия переменных.

IMPROVEMENT OF CHROMATOGRAPHIC DIAGNOSTICS METHOD OF DEVELOPING DEFECTS IN ELECTRIC EQUIPMENT ON BASIS OF FACTORIAL ANALYSIS THEORY

A.N. NAZARYCHEV, Doctor of Engineering, I.Yu. ZELENTSOV, Engineer

The authors developed the approach to improve the chromatographic diagnostics of developing defects in electric equipment on the basis of the factorial analysis theory. This approach allows to make deeper and more reliable forecast about presence of multifunction defects without undertaking the additional test based on available data. The authors give the example of using the factorial analysis for evaluation its technical state for power transformer.

Keywords: multifunction defects, matrix of factorial loads, power transformer, matrix of factor numerical values, correlation coefficient, general dispersion of variables.

Метод хроматографического анализа растворенных газов (ХАРГ) в масле маслонаполненного (трансформаторного) электрооборудования используется для диагностики развивающихся дефектов. Он основан на определении концентраций, растворенных в масле газов [1]. При помощи этого метода выявляют две группы дефектов: первая группа связана с перегревом токоведущих соединений и элементов конструкции остова; вторая − с электрическими разрядами в масле.

Наиболее сложной задачей при диагностике технического состояния трансформаторного оборудования, как и любого другого маслонаполненного оборудования, является выявление комбинированных дефектов, например сочетания дугового разряда в обмотке и перегрева сердечника трансформатора [2]. Эта задачи является актуальной для обеспечения более глубокой и достоверной оценки технического состояния электрооборудования на основе ХАРГ. Решение этой задачи предлагается искать на основе совершенствования метода хроматографической диагностики развивающихся дефектов в электрооборудовании на основе теории факторного анализа.

Факторный анализ представляет собой раздел математической статистики, относящийся к многомерному статистическому анализу. Среди задач, решаемых с помощью факторного анализа, наиболее распространенной является задача, связанная с уменьшением размерности описания системы путем замены совокупности непосредственно определяемых частных показателей качества меньшим числом более информативных комплексных показателей, отражающих существенные свойства системы.

Пусть получена совокупность из n наблюдений над состоянием изучаемой системы (объекта), характеризуемой набором m непосредственно измеряемых переменных (частных показателей качества) X_i, $i = 1…m$. Эти данные могут быть представлены в виде двумерной таблицы $X = (x_{ij})$ чисел x_{ij} размерностью $m\times n$. Каждый элемент x_{ij}, $i = 1…m$, $j = 1…n$, матрицы X есть конкретное значение i-й измеряемой (вычисляемой) переменной (i-го частного показателя качества) X_i для j-го наблюдения. Информационный массив, представленный описанной двумерной таблицей, назовем матрицей данных.

Чтобы сделать возможным сопоставление переменных X_i как величин различной размерности и устранить влияние различных единиц масштаба, матрицу исходных данных $X = (x_{ij})$ следует нормировать, вводя единый для всех переменных масштаб. Самый распространенный вид нормировки матрицы данных – переход от значений x_{ij} к значениям Z [3, 4]:

$$Z = \frac{x_{ij} - \bar{x}_i}{S_i}, \qquad (1)$$

где $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$; $S_i = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2}$.

Полученная в результате нормировки матрица $Z = (z_{ij})$ удовлетворяет следующим условиям:

$$\frac{1}{n}\sum_{j=1}^{n} z_{ij} = 0;\ \frac{1}{n-1}\sum_{i=0}^{n} z_{ij}^2 = 1;\ i = 1...m.$$

Выборочный коэффициент корреляции r_{ik} между переменными с номерами i и k может быть вычислен по формуле

$$r_{ik} = \frac{\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)\left(x_{kj} - \bar{x}_k\right)}{\sqrt{\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2 \sum_{j=1}^{n}\left(x_{kj} - \bar{x}_k\right)^2}} = \frac{\sum_{j=1}^{n} z_{ij}\, z_{kj}}{\sqrt{\sum_{j=1}^{n} z_{ij}^2 \ \sum_{j=1}^{n} z_{kj}^2}} =$$

$$= S_{ik} = \frac{1}{n-1}\sum_{j=1}^{n} z_{ij} \cdot z_{kj}\,, \qquad (2)$$

Далее рассмотрим применение факторного анализа для совершенствования метода хроматографической диагностики развивающихся дефектов на примере силовых трансформаторов.

Пусть для силового трансформатора получена следующая серия экспериментальных данных по ХАРГ:

1-й анализ: CO_2=0,17; CO=0,02; CH_4=0,0045; C_2H_4=0,005; C_2H_2-отсутствует; C_2H_6=0,002; H_2=0,008;

2-й анализ: CO_2=0,16; CO=0,02; CH_4=0,017; C_2H_4=0,05; C_2H_2=0,003; C_2H_6=0,0048; H_2=0,0075;

3-й анализ: CO_2=0,15; CO=0,02; CH_4=0,016; C_2H_4=0,048; C_2H_2=0,003; C_2H_6=0,0047; H_2=0,01.

Сформируем матрицу исходных данных $X = (x_{ij})$, которая примет вид матрицы размерностью 7×3, число строк которой соответствует числу наблюдаемых переменных, т.е. числу измеренных газов, а число столбцов соответствует количеству проведенных хроматографических анализов:

$$X = \begin{pmatrix} 0{,}17 & 0{,}16 & 0{,}15 \\ 0{,}02 & 0{,}02 & 0{,}02 \\ 0{,}0045 & 0{,}017 & 0{,}016 \\ 0{,}005 & 0{,}05 & 0{,}048 \\ 0 & 0{,}003 & 0{,}003 \\ 0{,}002 & 0{,}048 & 0{,}0047 \\ 0{,}008 & 0{,}0075 & 0{,}01 \end{pmatrix}.$$

Все m значений переменных X подвергнем нормировке по формуле (1), получим матрицу $Z = (z_{ij})$ нормированных (стандартизированных) значений:

$$Z = \begin{pmatrix} 1 & 0 & -1 \\ 0{,}8165 & 0{,}8165 & 0{,}8165 \\ -1{,}1517 & 0{,}6478 & 0{,}5039 \\ -1{,}1538 & 0{,}6162 & 0{,}5376 \\ -1{,}1547 & 0{,}5774 & 0{,}5774 \\ -1{,}1541 & 0{,}6085 & 0{,}5456 \\ -0{,}378 & -0{,}7559 & 1{,}1339 \end{pmatrix}.$$

Выборочные коэффициенты корреляции между переменными с номерами i и k вычислим по формуле (2) и сформируем корреляционную матрицу R:

$$R = \begin{pmatrix} 1 & 0 & -0{,}8278 & -0{,}8457 & -0{,}866 & -0{,}8499 & -0{,}7559 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -0{,}8278 & 0 & 1 & 0{,}9995 & 0{,}9974 & 0{,}9992 & 0{,}2585 \\ -0{,}8457 & 0 & 0{,}9995 & 1 & 0{,}9992 & 1 & 0{,}2899 \\ -0{,}866 & 0 & 0{,}9974 & 0{,}9992 & 1 & 0{,}9995 & 0{,}3273 \\ -0{,}8499 & 0 & 0{,}9992 & 1 & 0{,}9995 & 1 & 0{,}2974 \\ -0{,}7559 & 0 & 0{,}2585 & 0{,}2899 & 0{,}3273 & 0{,}2974 & 1 \end{pmatrix}$$

Для выборочной корреляционной матрицы $R = (r_{ik})$ и выборочной ковариационной матрицы $S = (s_{ik})$ имеет место соотношение [3, 4]

$$R = \frac{1}{n-1} Z\, Z^T = S. \qquad (3)$$

Считается, что каждая из переменных Z_i, i = 1…m, заданная в стандартной форме, может быть представлена как функция небольшого числа общих факторов P_1, P_2,…, P_r и характерного фактора P_{r+i}: $Z_i = f(P_1, P_2, \ldots, P_r, P_{r+i})$.

Каждый фактор P_l, l = 1…r, из числа общих факторов имеет существенное значение для анализа всех измеряемых переменных Z_i. Общие факторы обусловливают корреляцию между переменными Z_i и представляющими собой скрытые ненаблюдаемые непосредственно случайные переменные, которые проявляются только через связи между наблюдаемыми переменными. В отличие от общих факторов, изменение в характерном факторе P_{r+i} приводит к изменению значения только соответствующей переменной Z_i.

При факторном анализе, исходя из линейной модели, имеет место выражение

$$z_i = a_{i1}p_1 + a_{i2}p_2 + \ldots + a_{ir}p_r + u_i p_{r+i}\,,\ i = 1\ldots m. \qquad (4)$$

Такую модель называют основной моделью факторного анализа. Коэффициенты a_{il}, l = 1…r и u_i в этом выражении характеризуют значимость соответствующего фактора для описания j-й переменной и называются факторными нагрузками.

Допущение о линейности взаимосвязей в факторных моделях не всегда полно отражает реальные явления: переменные могут взаимодействовать и более сложным образом. Поэтому линейную модель факторного анализа следует считать первым приближением к отражению реальных процессов, происходящих при развитии дефектов.

В соответствии с моделью (4) для j-го наблюдаемого значения i-й переменной Z_i можно записать выражение

$$z_{ij} = a_{i1}p_{1j} + \ldots + a_{1l}p_{lj} + \ldots + a_{ir}p_{rj} + u_i p_{r+i,j}\,,$$
$$i = 1\ldots m,\ j = 1\ldots n,\ r \le m. \qquad (5)$$

В матричной форме выражение (4) выглядит следующим образом:

$$Z = F\,P. \qquad (6)$$

Здесь $Z = (z_{ij})$ – матрица исходных данных, записанных в стандартизированной форме; F – матрица размера $m\times(r + m)$, содержащая нагрузки характерных факторов и называемая факторным отображением; P – матрица размера $(r + m)\times n$, содержащая значения общих и характерных факторов отдельных наблюдений. Матрица F является блочной матрицей $F = (AU)$ с блоками в виде матриц

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mr} \end{pmatrix},\quad U = \begin{pmatrix} u_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & u_m \end{pmatrix}.$$

Используя выражение (6) для матрицы Z, для выборочной корреляционной матрицы R наблюдаемых переменных (3) получим

$$R = \frac{1}{n-1} Z Z^T = \frac{1}{n-1}(FP)(FP)^T =$$

$$F\left(\frac{1}{n-1}PP^T\right)F^T = FCF^T;$$

$$R = FCF^T; \qquad (7)$$

$$C = \frac{1}{n-1}PP^T. \qquad (8)$$

Матрица C, определяемая формулой (8), является выборочной корреляционной матрицей, отражающей связи между факторами.

Если наложить условие некоррелированности факторов, т.е. принять $C = I$, где I – единичная матрица, то получим [3, 4]

$$R = FF^T = AA^T + UA^T + AU^T + UU^T. \qquad (9)$$

Легко показать, что

$$UA^T = AU^T = 0. \qquad (10)$$

Следовательно,

$$R = AA^T + UU^T. \qquad (11)$$

Из выражения (11) получим

$$R_h = R - UU^T = AA^T. \qquad (12)$$

Матрица R_h называется редуцированной корреляционной матрицей. Поскольку UU^T является диагональной матрицей, то редуцированная корреляционная матрица R_h отличается от корреляционной матрицы R только элементами главной диагонали. Уравнение (12) выражает основную зависимость, служащую для нахождения факторных нагрузок a_{il}, $i = 1…m$, $l = 1…r$.

Задачей факторного анализа является определение матрицы факторных нагрузок A путем нахождения решения уравнения (12). Это уравнение не имеет однозначного решения. Действительно, матрица $B = AT$, где T – ортогональная матрица ($T^T = T^{-1}$), тоже удовлетворяет уравнению (12), что проверяется непосредственной подстановкой в него вместо матрицы A матрицы B. Получение однозначного решения уравнения становится возможным при введении ограничений на матрицу факторных нагрузок [4, 5]. Обычно устанавливаются такие дополнительные требования к матрице A, которые полезны с содержательной точки зрения.

После определения факторных нагрузок могут быть рассчитаны элементы матрицы численных значений факторов, т.е. значения факторов для отдельных наблюдений.

Процедура факторного анализа начинается с получения матрицы исходных данных $X = (x_{ij})$, $i = 1…m$, $j = 1…n$, путем измерения значений непосредственно наблюдаемых переменных X_i. По ней вычисляется корреляционная матрица $R = (r_{ik})$, $l,k = 1…m$, которая преобразуется в редуцированную корреляционную матрицу $R_h = (r_{ik}^h)$, отличающуюся от корреляционной матрицы только элементами главной диагонали. Далее находится матрица факторных нагрузок $A = (a_{il})$, $i = 1…m$, $l = 1…r$, удовлетворяющая рассмотренному уравнению (12). После определения матрицы факторных нагрузок рассчитываются оценки значений факторов для отдельных наблюдений, которые являются элементами матрицы значений факторов $P = (p_{lj})$, $l = 1….r$, $j = 1…n$.

Если сравнить между собой матрицы X и P, то станет очевидным упрощение, которое достигается в результате выполнения процедуры факторного анализа. Матрица X имеет размер $m\times n$, а матрица P – размер $r\times h$, т.е. m переменных, измеренных в n опытах, сводится к $r < m$ факторам.

Процедура факторного анализа может рассматриваться как процедура решения своеобразной задачи аппроксимации большого числа наблюдаемых переменных меньшим числом факторов или как процедура аппроксимации матрицы исходных данных $m\times n$ матрицей значений факторов $r\times n$. При $r\times n$ происходит сжатие объема исходного информационного массива с некоторой потерей в точности объекта анализа.

Чтобы определить элементы главной диагонали редуцированной корреляционной матрицы, рассмотрим, из каких составляющих складывается выборочная дисперсия i-й наблюдаемой переменной:

$$S_i^2 = \frac{1}{n-1}\sum_{i=1}^{n} z_{ij}^2, \quad i = 1...m. \qquad (13)$$

Производя замену переменных, получим

$$S_i^2 = a_{i1}^2 \frac{1}{n-1}\sum_{j=1}^{n} p_{1j}^2 + a_{i2}^2 \frac{1}{n-1}\sum_{j=1}^{n} p_{2j}^2 + ... +$$

$$+ u_i^2 \frac{1}{n-1}\sum_{j=1}^{n} p_{r+i,j}^2 + 2\left(a_{i1}a_{i2}\frac{1}{n-1}\sum_{j=1}^{n} p_{1j}p_{2j} + ... \right).$$

Учитывая, что характерные факторы не связаны между собой и с общими факторами, а также полагая некоррелированность между собой общих факторов, установим, что все суммы в скобках последнего выражения равны нулю [3, 4].

Если факторы стандартизированы (имеют нулевые средние и единичные дисперсии), то все величины (14) равны единице:

$$\frac{1}{n-1}\sum_{j=1}^{n} p_{qj}^2, \quad q = 1, 2, ..., r, r+i. \qquad (14)$$

Поэтому при стандартизированных наблюдаемых переменных и факторах, а также некоррелированных факторах

$$S_i^2 = 1 = \sum_{l=1}^{r} a_{il}^2 + u_i^2, \quad i = 1...m. \qquad (15)$$

Сумма квадратов нагрузок общих факторов дает долю единичной дисперсии переменной Z_i, приходящуюся на все r общих факторов. Она называется общностью:

$$h_i^2 = \sum_{l=1}^{r} a_{il}^2. \qquad (16)$$

Оставшаяся часть единичной дисперсии, которая не связана с общими факторами, называется характерностью:

$$u_i^2 = 1 - h_i^2. \qquad (17)$$

Поскольку

$$UU^T = \begin{pmatrix} u_1^2 & ... & 0 \\ ... & ... & ... \\ 0 & ... & u_m^2 \end{pmatrix}, \qquad (18)$$

то редуцированная корреляционная матрица с учетом выражения (11) представляется следующим образом [3, 4, 5]:

$$R_h = \begin{pmatrix} h_1^2 & r_{12} & ... & r_{1m} \\ r_{21} & h_2^2 & ... & r_{2m} \\ ... & ... & ... & ... \\ r_{m1} & r_{m2} & ... & h_m^2 \end{pmatrix}. \qquad (19)$$

Согласно (19), на главной диагонали редуцированной корреляционной матрицы расположены общности, а остальные элементы совпадают с соответствующими коэффициентами корреляции между наблюдаемыми переменными.

В настоящее время разработано достаточно много методов оценки общностей, удовлетворительных с практической точки зрения. Наиболее простым является метод, использующий в качестве оценок общностей максимальные коэффициенты корреляции соответствующих переменных со всеми остальными [3, 4]:

$$h_i^2 = \max_s r_{is}, \quad s = 1...m, \quad s \neq i. \qquad (20)$$

Этот метод особенно эффективен для корреляционных матриц большой размерности. В этом случае редуцированная корреляционная матрица примет следующий вид (значения общностей взяты в скобки для наглядности):

$$R_h = \begin{pmatrix} (0{,}866) & 0 & -0{,}8278 & -0{,}8457 & -0{,}866 & -0{,}8499 & -0{,}7559 \\ 0 & (0) & 0 & 0 & 0 & 0 & 0 \\ -0{,}8278 & 0 & (0{,}9995) & 0{,}9995 & 0{,}9974 & 0{,}9992 & 0{,}2585 \\ -0{,}8457 & 0 & 0{,}9995 & (1) & 0{,}9992 & 1 & 0{,}2899 \\ -0{,}866 & 0 & 0{,}9974 & 0{,}9992 & (0{,}9995) & 0{,}9995 & 0{,}3273 \\ -0{,}8499 & 0 & 0{,}9992 & 1 & 0{,}9995 & (1) & 0{,}2974 \\ -0{,}7559 & 0 & 0{,}2585 & 0{,}2899 & 0{,}3273 & 0{,}2974 & (0{,}7559) \end{pmatrix}.$$

Второй метод оценивания общности связан с использованием триад [3, 4, 5]:

$$h_i^2 = \frac{r_{ik}\, r_{il}}{r_{kl}}, \quad i = 1...m, \quad k \neq l \neq i, \qquad (21)$$

где k и l – номера переменных, коэффициенты корреляции которых с переменной Z_i превышают остальные ее коэффициенты корреляции.

Формула (21) имеет тенденцию завышения максимальных общностей. Это может привести к тому, что за счет малых значений r_{kl} значение h^2_i станет больше единицы, что противоречит смыслу общности $h_i^2 \leq 1$. В нашем примере для трансформатора редуцированная корреляционная матрица примет вид

$$R_h = \begin{pmatrix} (0{,}7364) & 0 & -0{,}8278 & -0{,}8457 & -0{,}866 & -0{,}8499 & -0{,}7559 \\ 0 & (0) & 0 & 0 & 0 & 0 & 0 \\ -0{,}8278 & 0 & (0{,}9987) & 0{,}9995 & 0{,}9974 & 0{,}9992 & 0{,}2585 \\ -0{,}8457 & 0 & 0{,}9995 & (1{,}0003) & 0{,}9992 & 1 & 0{,}2899 \\ -0{,}866 & 0 & 0{,}9974 & 0{,}9992 & (0{,}9987) & 0{,}9995 & 0{,}3273 \\ -0{,}8499 & 0 & 0{,}9992 & 1 & 0{,}9995 & (1{,}0003) & 0{,}2974 \\ -0{,}7559 & 0 & 0{,}2585 & 0{,}2899 & 0{,}3273 & 0{,}2974 & (0{,}2911) \end{pmatrix}.$$

Иногда в качестве оценки общности берут среднее арифметическое значение коэффициентов корреляции соответствующей переменной с остальными переменными [3, 4, 5]:

$$h_i^2 = \frac{1}{m-1}\sum_{k=1}^{m-1} r_{ik}, \quad i \neq k. \qquad (22)$$

Последняя оценка склонна к выравниванию значений общности для различных переменных, т.е. она, как правило, дает заниженные значения высоких общностей и завышенные значения низких общностей. Она также не исключает возможности получения значений общности больше единицы. В данном случае редуцированная корреляционная матрица примет вид

$$R_h = \begin{pmatrix} (0{,}6909) & 0 & -0{,}8278 & -0{,}8457 & -0{,}866 & -0{,}8499 & -0{,}7559 \\ 0 & (0) & 0 & 0 & 0 & 0 & 0 \\ -0{,}8278 & 0 & (0{,}4045) & 0{,}9995 & 0{,}9974 & 0{,}9992 & 0{,}2585 \\ -0{,}8457 & 0 & 0{,}9995 & (0{,}4071) & 0{,}9992 & 1 & 0{,}2899 \\ -0{,}866 & 0 & 0{,}9974 & 0{,}9992 & (0{,}4096) & 0{,}9995 & 0{,}3273 \\ -0{,}8499 & 0 & 0{,}9992 & 1 & 0{,}9995 & (0{,}4077) & 0{,}2974 \\ -0{,}7559 & 0 & 0{,}2585 & 0{,}2899 & 0{,}3273 & 0{,}2974 & (0{,}0695) \end{pmatrix}.$$

При большой размерности корреляционной матрицы, т.е. при большом числе наблюдаемых переменных, точность приближения оценок общностей к истинным значениям слабо влияет на конечные результаты факторного анализа [3, 4, 5]. Это обусловлено тем, что с увеличением числа переменных удельный вес диагонального элемента среди всех элементов данного столбца редуцированной корреляционной матрицы уменьшается, благодаря чему он оказывает все меньшее влияние на определение факторных нагрузок. Поэтому с точки зрения практики при большом числе переменных не так уж важно иметь очень точные оценки общностей. При двадцати и более переменных уже не приходится опасаться погрешностей в оценках общностей. В тех случаях, когда размерность корреляционной матрицы невелика, качество оценок общностей оказывает заметное влияние на точность определения факторных нагрузок [3, 4, 5].

Самым распространенным способом вычисления факторных нагрузок является метод главных факторов (метод главных компонент).

Метод главных факторов позволяет выделить упорядоченную последовательность ортогональных факторов, причем каждый последующий фактор P_l дает меньший общий вклад W_l в дисперсию всех наблюдаемых переменных, чем предыдущий:

$$W_l = \sum_{i=1}^{m} a_{il}^2. \qquad (23)$$

Дополнительные условия, обеспечивающие получение однозначного решения уравнения (12), в методе главных факторов приобретают следующую форму: сумма квадратов нагрузок первого фактора должна составлять максимум суммарной дисперсии наблюдаемых переменных, сумма квадратов нагрузок второго фактора должна составлять максимум от оставшейся дисперсии и т.д. При таких условиях первым фактором является тот, который дает максимальный вклад в общую дисперсию переменных, вторым – фактор, дающий максимальный вклад во вновь найденную в результате исключения первого фактора общую дисперсию, и т.д.

Метод главных факторов строится как метод решения упорядоченной цепочки связанных между собой оптимизационных задач. Первая задача этой цепочки формулируется следующим образом: найти максимум вклада W_1 фактора P_1 как функции факторных нагрузок a_{i1} этого фактора на все наблюдаемые переменные при условии, что матрица факторных нагрузок удовлетворяет соотношению (11), рассматриваемому как ограничение оптимизационной задачи:

$$W_1 = \sum_{i=1}^{m} a_{i1}^2 \to \max. \qquad (24)$$

После решения, т.е. нахождения вектора факторных нагрузок фактора P_1 $\vec{a}_1 = \left(a_{11}, a_{21}, ..., a_{m1}\right)$, соответствующего максимуму W_1 при заданных ограничениях, строится матрица

$$A_1 = a_1 a_1^T \qquad (25)$$

и с ее помощью вычисляется матрица остаточных коэффициентов корреляции:

$$R_h^1 = R_h - A_1. \qquad (26)$$

Далее решается следующая оптимизационная задача: найти максимум вклада W_2 фактора P_2 при новых ограничениях:

$$W_2 = \sum_{i=1}^{m} a_{i2}^2 \to \max, \text{ при } AA^T = R_h^1. \qquad (27)$$

Пусть процедура последовательного нахождения факторных нагрузок выполнена до l-го фактора включительно, т.е. оказались найденными все векторы $\vec{a}_1, \vec{a}_2, ..., \vec{a}_l$. После этого строится матрица

$$R_h^l = R_h^{l-1} A_l , \qquad (28)$$

где $A_l = \vec{a}_l \, \vec{a}_l^T$, $\vec{a}_l = \left(a_{1l}, a_{2l}, ..., a_{ml}\right)^T$, – вектор-столбец факторных нагрузок фактора P_l, формулируется и решается (l+1)-я оптимизационная задача.

При применении метода главных факторов возникает вопрос о том, когда должен быть закончен процесс выделения факторов или, иначе, каким числом факторов можно удовлетвориться. Обычно придерживаются стратегии, состоящей в том, что вначале выделяется на один фактор больше, а затем он отбрасывается, либо оставляется на основании результатов дальнейшего анализа и выбранного критерия прекращения выделения факторов. На практике выделение факторов производится до момента распределения по факторам 95 % общности. При определении числа r общих факторов можно руководствоваться также и другими критериями.

Рассматривая приведенную выше последовательность оптимизационных задач, можно показать, что метод главных факторов сводится к решению проблемы собственных значений действительной симметрической матрицы R_h [3, 4, 5]. Эта проблема представляется как проблема поиска собственных векторов $\vec{\alpha}_l$ и соответствующих им собственных значений λ_l матрицы R_h путем решения следующего уравнения:

$$R_h \, \vec{\alpha}_l = \lambda_l \, \vec{\alpha}_l \quad \text{или} \quad \left(R_h - \lambda_l \, I\right)\vec{\alpha}_l = 0. \qquad (29)$$

Искомые значения a_{il} элементов матрицы факторных нагрузок А получаются по компонентам $б_{il}$ собственных векторов матрицы R_h нормированием:

$$a_{il} = \frac{\alpha_{il}\sqrt{\lambda_l}}{\sqrt{\alpha_{1l}^2 + \alpha_{2l}^2 + ... + \alpha_{ml}^2}}. \qquad (30)$$

После того как будут определены факторные нагрузки, можно рассчитать значения факторов. Задача получения текущих значений факторов для каждого опыта формулируется как задача нахождения наилучших в том или ином смысле оценок значений факторов при известных факторных нагрузках и матрице стандартизированных наблюдаемых переменных. Оценки значений факторов можно получить с помощью метода наименьших квадратов, используя технику регрессивного анализа. Применение стандартизованных переменных, среднее значение которых равны нулю, а дисперсии – единице, позволяет упростить соответствующие расчетные формулы, так как при этом исчезает свободный член в уравнении регрессии. Для решения поставленной задачи можно использовать равенство [3]

$$P = A^{-1} Z. \qquad (31)$$

Так как Z известна, то для получения P нужно определить лишь A^{-1}. Для обращения матрицы необходимо, чтобы она была квадратная. Матрица A квадратная только тогда, года выделяются все m главных компонент (31). По известным матрицам Z и A определим значения P. Эта процедура приводит к точным и однозначным результатам.

В нашем примере в качестве оценки общностей возьмем средние арифметические значения коэффициентов корреляции соответствующей переменной с остальными. Имеется много способов вычисления собственных значений и собственных векторов, прежде всего, это метод Якоби [3, 4] и способ диагонализации корреляционной матрицы Хаусхолдера [3, 4]. Целесообразнее всего использовать имеющиеся в распоряжении математического обеспечения ЭВМ пакеты прикладных программ, включающие в себя программу вычисления собственных значений и собственных векторов, такие как MathLab, MathCad и др. Воспользовавшись одним из математических пакетов, получим собственные значения редуцированной корреляционной матрицы и соответствующие им собственные вектора редуцированной корреляционной матрицы:

$$\lambda = \begin{pmatrix} 4,39 \\ 0,3096 \\ -0,5284 \\ -0,5973 \\ -0,5928 \\ -0,5919 \\ 0 \end{pmatrix};$$

$$\alpha = \begin{pmatrix} 0,4413 & -0,5598 & 0,69 & -0,1156 & 0,0265 & 0,0413 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -0,4344 & -0,2579 & -0,0565 & -0,8353 & -0,1324 & -0,1623 & 0 \\ -0,4377 & -0,2237 & 0,0808 & 0,1329 & 0,848 & 0,1227 & 0 \\ -0,4409 & -0,1817 & 0,2583 & 0,4398 & -0,2733 & -0,6616 & 0 \\ -0,4384 & -0,2153 & 0,1153 & 0,2155 & -0,4317 & 0,7179 & 0 \\ -0,1961 & 0,7005 & 0,6589 & -0,177 & 0,0396 & 0,0615 & 0 \end{pmatrix}.$$

Факторные нагрузки определим по формуле (30), получим факторное отображение

$$A = \begin{pmatrix} 0,9246 & -0,3115 & 0,5016i & -0,0893i & 0,0204i & 0,0318i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0,9101 & -0,1435 & -0,0411i & -0,6456i & -0,1019i & -0,1249i & 0 \\ -0,917 & -0,1244 & 0,0587i & 0,1027i & 0,6529i & 0,0944i & 0 \\ -0,9238 & -0,1011 & 0,1878i & 0,3399i & -0,2104i & -0,509i & 0 \\ -0,9185 & -0,1198 & 0,0838i & 0,1666i & -0,3324i & 0,5523i & 0 \\ -0,411 & 0,3897 & 0,4789i & -0,1368i & -0,0305i & 0,0473i & 0 \end{pmatrix}.$$

По формуле (31) вычислим значения факторов. В результате получим

$$P = \begin{pmatrix} 1,2102 & -0,4411 & -0,7692 \\ 0,3395 & -1,9239 & 1,5844 \\ 0,0255i & 0,3653i & -0,3908i \\ -0,0045i & -0,0772i & 0,0817i \\ 0,00095i & 0,0176i & -0,0186i \\ 0,0015i & 0,0275i & -0,0289i \\ 3,25\cdot 10^7 & 3,25\cdot 10^7 & 3,25\cdot 10^7 \end{pmatrix}.$$

Факторные нагрузки принимают значения от –1 до +1 (при ортогональных факторах), каждый фактор характеризуется столбцом, каждая переменная – строкой матрицы A. С практической точки зрения не обязательно рассматривать столько общих факторов, сколько их дает редуцированная корреляционная матрица, т.е. сколько отличных от нуля собственных значений имеет эта матрица. Набор полученных в процессе нахождения факторных нагрузок вес собственных значений матрицы R_h позволяет построить много различных правил определения существенных общих факторов. Таким образом, нас интересуют первые два фактора, а факторное отображение примет вид

$$A = \begin{pmatrix} 0,9246 & -0,3115 \\ 0 & 0 \\ -0,9101 & -0,1435 \\ -0,917 & -0,1244 \\ -0,9238 & -0,1011 \\ -0,9185 & -0,1198 \\ -0,411 & 0,3897 \end{pmatrix}.$$

И соответствующие значения факторов:

$$P = \begin{pmatrix} 1,2102 & -0,4411 & -0,7692 \\ 0,3395 & -1,9239 & 1,5844 \end{pmatrix}.$$

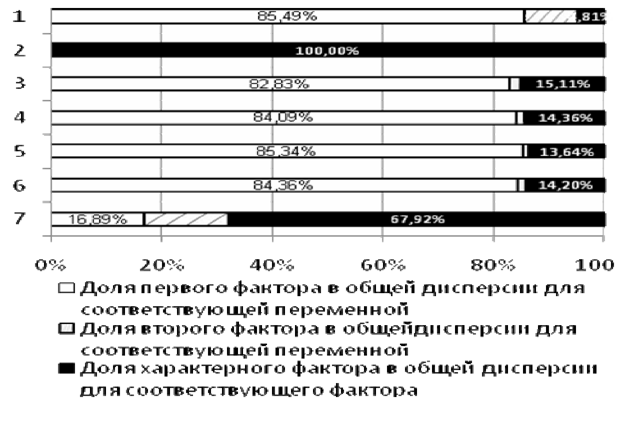
По известным факторным нагрузкам определим при помощи соответствующих выражений значения дисперсии первого и второго факторов каждой переменной, а также общность и характерность. Результаты представлены в табл. 1. Графически данные табл. 1 представлены на рисунке в процентах от полной дисперсии. В табл. 2 представлены суммарные значения показателей.

Таблица 1. **Дисперсии первого и второго факторов, общность и характерность**

Факторные нагрузки первого фактора	Факторные нагрузки второго фактора	Дисперсия первого фактора	Дисперсия второго фактора	Общность	Характерность
0,9246	-0,3115	0,854885	0,097032	0,951917	0,048083
0	0	0	0	0	1
-0,9101	-0,1435	0,828282	0,020592	0,848874	0,151126
-0,917	-0,1244	0,840889	0,015475	0,856364	0,143636
-0,9238	-0,1011	0,853406	0,010221	0,863628	0,136372
-0,9185	-0,1198	0,843642	0,014352	0,857994	0,142006
-0,411	0,3897	0,168921	0,151866	0,320787	0,679213

Таблица 2. **Суммарные значения дисперсий, общности и характерности**

Показатель	Значение	%
Полная дисперсия	7	100
Дисперсия первого фактора	4,390026	62,71466
Дисперсия второго фактора	0,309539	4,421989
Суммарная общность	4,699565	67,13664
Суммарная характерная дисперсия	2,300435	32,86336



Влияние факторов на содержание соответствующих газов в масле: □ – доля первого фактора в общей дисперсии для соответствующей переменной; ☐ – доля второго фактора в общей дисперсии для соответствующей переменной; ■ – доля характерного фактора в общей дисперсии для соответствующего фактора

Анализ полученных данных показывает, что на переменные основное влияние оказывает первый фактор (дисперсия первого фактора 62,7 % полной дисперсии), и лишь незначительное влияние оказывает второй фактор. Первый фактор оказывает основное влияние на первую, третью, четвертую, пятую, и шестую переменную, т.е. на изменение содержания следующих газов: CO_2, CH_4, C_2H_4, C_2H_2, C_2H_6. При этом содержание CO определяется полностью характерным фактором, а содержание H_2 определяется на 67,92% характерным фактором, не зависящим от общих факторов (главных) и влияющим только на данные переменные. Наличие второго главного фактора, хотя его влияние и незначительное, говорит о возможности наличия начальной стадии развития второго дефекта трансформатора, помимо основного, явно выраженного дефекта, или наличия какого-то дополнительного фактора, вызвавшего изменение содержания газов в масле, что указывает на необходимость детального анализа условий и режимов работы трансформатора. Согласно [1], такими факторами могут быть: остаточные концентрации газов от устраненного дефекта во время ремонта трансформатора (если не была проведена дегазация масла); увеличение нагрузки трансформатора; перемешивание свежего масла с остатками старого, насыщенного газами, находящегося в системе охлаждения, баках РПН, расширителе; доливка маслом, бывшим в эксплуатации и содержащим растворенные газы; проведение сварочных работ на баке; повреждения масляных насосов с неэкранированным статором; перегревы из-за дефектов системы охлаждения (засорение наружной поверхности охладителей, отключение части масляных насосов и др.); перегрев масла теплоэлектронагревателями при его обработке в дегазационных и других установках; переток газов из бака расширителя контактора РПН в бак расширителя трансформатора, имеющего РПН типа РС-3 или РС-4; сезонные изменения интенсивности процесса старения; воздействие токов короткого замыкания. Также это могут быть факторы, снижающие содержание газов, такие как: продувка азотом в трансформаторах с азотной защитой масла; уменьшение нагрузки трансформатора; замена силикагеля; длительное отключение; дегазация масла; доливка дегазированным маслом; частичная или полная замена масла в баке трансформатора; заливка маслом под вакуумом, в том числе частичным вакуумом; замена масла в маслопроводах, навесных баках, расширителе, избирателе устройств РПН и т.д.

Таким образом, путем применения факторного анализа произошло сжатие первичной информации до двух факторов. Очевидно, что исходная матрица 7×3 преобразовалась в матрицу 2×3. Теперь два комплексных показателя описывают состояние семи измеренных переменных (содержание семи газов в масле). Факторный анализ позволяет выявить факторы, влияющие на состояние объекта, но непосредственно не измеряемые.

Заключение

Применение факторного анализа позволяет усовершенствовать методику ХАРГ путем выявления количества причин (числа факторов) появления или изменения концентраций растворенных в масле газов, включая эксплуатационные факторы (причины), а следовательно, помогает оценить техническое состояние электрооборудования.

Выявление двух и более факторов может указать на наличие комбинированных дефектов в трансформаторе без проведения дополнительных испытаний на основе уже имеющихся данных хроматографического анализа газов. Выявление комбинированных развивающихся дефектов позволяет снизить ремонтные затраты путем устранения обоих дефектов при ремонте трансформатора.

Действующая методика ХАРГ определяет наличие одного дефекта. При этом не исключается возможность существования другого дефекта, наличие которого выявится только при дальнейшей эксплуатации или в процессе ремонта. Применение факторного анализа для оценки результатов ХАРГ дает возможность установить факт наличия (отсутствия) второго (комбинированного) дефекта.

Список литературы

1. РД 153-34.0-46.302-00. Методические указания по диагностике развивающихся дефектов трансформаторного оборудования по результатам хроматографического анализа газов, растворенных в масле. – М., 2001.

2. Алексеев Б.А. Обследование состояния силовых трансформаторов. СИГРЭ-2002 // Электрические станции. – 2003. – № 6. – С. 74–80.

3. Иберла К. Факторный анализ / пер. с нем. В.М. Ивановой. – М.: Статистика, 1980.

4. Лоули Д., Максвелл А. Факторный анализ как статистический метод / пер. с англ. Ю.Н. Благовещенского. – М.: Мир, 1967.

5. Харман Г. Современный факторный анализ / пер. с англ. В.Я. Лумельского; под ред. Э.М. Бравермана. – М.: Статистика, 1972.

Назарычев Александр Николаевич,
ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина»,
доктор технических наук, профессор, зав. кафедрой электрических станций, подстанций и диагностики электрооборудования,
телефон (4932) 26-99-43,
e-mail: kafedra@esde.ispu.ru

Зеленцов Игорь Юрьевич,
ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина»,
инженер кафедры электрических станций, подстанций и диагностики электрооборудования,
телефон (4932) 26-99-43,
e-mail: kafedra@esde.ispu.ru