

ЭЛЕКТРОЭНЕРГЕТИКА

УДК 621.314

Александр Леонидович Куликов

ФГБОУВО «Нижегородский государственный технический университет имени Р.Е. Алексеева», доктор технических наук, профессор кафедры электроэнергетики, электроснабжения и силовой электроники, Россия, Нижний Новгород, e-mail: inventor61@mail.ru

Дмитрий Игоревич Бездушный

ФГБОУВО «Нижегородский государственный технический университет имени Р.Е. Алексеева», аспирант кафедры электроэнергетики, электроснабжения и силовой электроники, Россия, Нижний Новгород, e-mail: dimok.93@mail.ru

Владислав Юрьевич Осокин

ФГБОУВО «Нижегородский государственный технический университет имени Р.Е. Алексеева», аспирант кафедры электроэнергетики, электроснабжения и силовой электроники, Россия, Нижний Новгород, e-mail: osokin-v92@mail.ru

Применение линейного дискриминантного анализа для классификации аварийных режимов электрической сети

Авторское резюме

Состояние вопроса. Развивающийся в настоящее время подход к построению алгоритмов релейной защиты на основе статистического анализа данных предлагает переход от традиционных видов защит к принципиально новым многомерным алгоритмам распознавания, способным эффективно использовать доступную информацию для выявления аварийных режимов. Используемые для распознавания параметры режима, как правило, объединяются в многомерное пространство и анализируются совместно. Это приводит к росту вычислительной сложности распознающего алгоритма. Таким образом, может быть целесообразно построение алгоритмов РЗА на основе обобщенных признаков. Указанное решение позволит снизить размерность признакового пространства и повысить быстродействие защиты.

Материалы и методы. Используется линейный дискриминантный анализ для снижения размерности признакового пространства при распознавании режимов электрической сети. Формирование обучающей выборки для статистического анализа осуществляется при помощи имитационной модели анализируемого участка электрической сети. В качестве критерия эффективности принята матрица ошибок, полученная в результате применения метода k -ближайших соседей для классификации режимов сети в исследуемом признаковом пространстве.

Результаты. Предложен способ снижения размерности признакового пространства для алгоритмов многомерной РЗА, заключающийся в переходе к обобщенным признакам срабатывания, полученным

в результате применения линейного дискриминантного анализа. Способ позволяет сформировать уставочную плоскость, наиболее эффективную для классификации различных видов коротких замыканий и обеспечивающую безошибочную классификацию, в то время как уставочные плоскости, образованные парами исходных признаков, демонстрируют погрешность, достигающую 50 %. Поставленная цель по снижению размерности признакового пространства при сохранении исходной информативности достигнута.

Выводы. Линейный дискриминантный анализ может применяться производителями РЗА для снижения размерности признакового пространства при разработке алгоритмов многомерной релейной защиты, в том числе в задачах множественной классификации.

Ключевые слова: линейный дискриминантный анализ, информационный подход, имитационное моделирование, метод Монте-Карло

Aleksandr Leonidovich Kulikov

State Technical University of Nizhny Novgorod, Doctor of Engineering Sciences, Professor of the Power Industry, Electric Supply and Power Electronics department, Russia, Nizhny Novgorod, e-mail: inventor61@mail.ru

Dmitriy Igorevich Bezdushnii

State Technical University of Nizhny Novgorod, Post graduate student of the Power Industry, Electric Supply and Power Electronics department, Russia, Nizhny Novgorod, e-mail: dimok.93@mail.ru

Vladislav Yurievich Osokin

State Technical University of Nizhny Novgorod, Post graduate student of the Power Industry, Electric Supply and Power Electronics department, Russia, Nizhny Novgorod, e-mail: osokin-v92@mail.ru

Application of linear discriminant analysis for classification of electrical grid emergency states

Abstract

Background. The approach of development of relay protection algorithms based on statistical data analysis suggests a transition from traditional protections to innovative multidimensional algorithms of recognition which are able to use the available information for emergency states identification more efficiently. The state parameters used for recognition are usually combined into a single multidimensional space. This results in computational complexity of the recognition algorithm. Thus, it is advisable to implement relay protection algorithms based on generalized features. This decision will make it possible to decrease the dimension of the feature space and increase response time of the protection.

Materials and methods. The linear discriminant analysis is used to reduce the feature space dimension when recognizing the states of an electrical grid section. The training sampling for statistical analysis is formed using a simulation model of the analyzed network section. The error matrix obtained as a result of the application of k-nearest neighbors algorithm for the state classification in the developed feature space was taken as a criterion for the feature space efficiency.

Results. A method of feature space dimension reduction for the task of multidimensional relay protection was proposed. It involves generalized features of response obtained via linear discriminant analysis. The method allowed formation of a setting plane which is the most effective for different types of short-circuit classification. This setting plane provides an error-free state classification while the setting planes obtained by pairs of initial features demonstrate an error of up to 50 %. The set goal of the feature space dimension reduction was achieved.

Conclusions. Linear discriminant analysis is advisable to use by relay protection manufacturers for reduction of the feature space dimension in multidimensional relay protection problems, including multiple classification problems.

Key words: linear discriminant analysis, information approach, simulation modeling, Monte Carlo method

DOI: 10.17588/2072-2672.2020.5.038-047

Введение. Подход к формированию алгоритмов релейной защиты и автоматики (РЗА), основанный на элементах матема-

тической статистики, получает распространение как в России [1, 2], так и за рубежом [3]. Реализация указанного подхода стала

возможной благодаря широкому распространению цифровых устройств интеллектуальной релейной защиты, способных выполнять сложные алгоритмы обработки информации в реальном времени, внедрению внутривыпускных коммуникационных протоколов, обеспечивающих цифровым устройствам доступ к большому объему данных о текущем режиме, а также благодаря развитию средств имитационного моделирования электроэнергетических систем (ЭЭС), позволяющих формировать статистические выборки о режимах защищаемого участка сети для обучения алгоритмов РЗА.

Статистический анализ данных по характерным режимам защищаемого объекта позволяет исследовать множество возможных режимов функционирования защищаемого объекта, включая маловероятные схемные варианты. Кроме того, при формировании правил и условий срабатывания устройств РЗА статистический подход обеспечивает учет не только диапазонов изменения наблюдаемых величин, но и неоднородностей распределений их вероятностей, а также корреляцию между переменными. Таким образом, возможен выбор параметров срабатывания защиты, минимизирующий вероятность ее ошибочного действия в тех случаях, когда абсолютное распознавание режимов не представляется возможным. Однако количество контролируемых величин, на основе которых осуществляется классификация режимов, как правило, выбирается произвольно и приводит к формированию многомерных признаков пространств.

Увеличение размерности признаков пространства способствует повышению качества распознавания [4], однако несет в себе ряд недостатков, связанных с ростом вычислительных затрат, необходимых для классификации режимов устройствами РЗА, а также времени обучения алгоритма защиты. Кроме того, распределения режимов и границы области срабатывания алгоритма РЗА в признаковом пространстве высокой размерности не могут быть наглядно визуализированы. В связи с этим целесообразно применение алгоритмов РЗА на основе обобщенных признаков, являющихся взвешенными комбинациями наблюдаемых величин [5]. Использование таких обобщенных признаков позволяет осуществить переход от исходного признакового пространства к

искусственно сформированному пространству существенно меньшей размерности, обладающему, однако, сопоставимой эффективностью с точки зрения распознавания режимов электрической сети.

Сжатие признакового пространства, получившее наибольшее распространение в задачах анализа данных, как правило, связывают с методом главных компонент [5]. Такой метод позволяет сформировать признаковое пространство меньшей размерности, образованное некоррелированными величинами, максимально сохраняющими дисперсию исходного признакового пространства. Пример применения метода главных компонент в бинарной задаче классификации режимов ЭЭС рассматривается в [6]. Тем не менее можно показать, что преобразование, обеспечивающее максимальную дисперсию, не всегда является оптимальным с точки зрения разделения классов. В качестве примера рассмотрим распределения двух случайных величин (рис. 1).

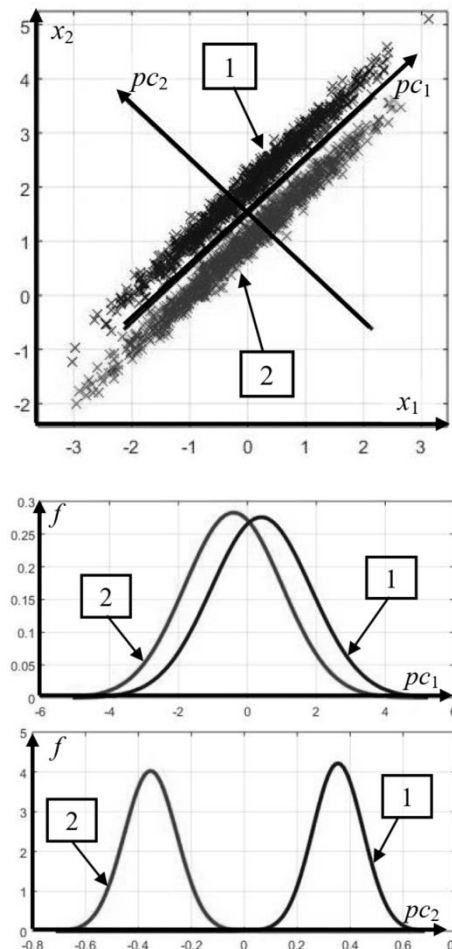


Рис. 1. Распределения точек обучающей выборки

Направление оси, проекция на которую обеспечивает максимальную диспер-

сию, обозначена как pc_1 , а ортогональная ей ось, соответствующая наименьшей дисперсии, – как pc_2 . Несмотря на максимальное значение дисперсии, проекции обучающей выборки на ось pc_1 не позволяют эффективно различить объекты обучающей выборки, относящиеся к разным классам. Вместе с тем использование в качестве информационного признака проекций данных на pc_2 позволяет безошибочно разделить обучающие выборки на классы.

Для получения обобщенных признаков, оптимальных для решения задач классификации режимов электрической сети, целесообразно учитывать статистические характеристики не только обучающей выборки в целом, но и классов в отдельности. Указанного эффекта позволяет достичь линейный дискриминантный анализ [7].

Аналогично методу главных компонент, линейный дискриминантный анализ обеспечивает выделение из исходного признакового пространства подпространства меньшей размерности, а его выбор реализуется исходя из следующих принципов:

- максимизация расстояния между математическими ожиданиями признаков различных классов;
- минимизация дисперсии признаков внутри каждого класса.

Линейный дискриминантный анализ направлен на снижение размерности пространства признаков таким образом, чтобы проекции классов представляли собой области с минимальным разбросом параметров и находились на наибольшем удалении друг от друга. Формирование решающего правила, разделяющего классы, в таком подпространстве потенциально более эффективно, чем в подпространстве, полученном с применением метода главных компонент.

Преобразование, позволяющее осуществить переход к обобщенным признакам в соответствии с линейным дискриминантным анализом, может быть получено в результате решения задачи оптимизации:

$$\max_{\mathbf{w}} \frac{\sum_{i=1}^c d_i^2}{\sum_{i=1}^c s_i^2}, \quad (1)$$

где \mathbf{w} – матрица, осуществляющая преобразование исходного признакового пространства в искомое; d_i – суммарное евклидово расстояние между математиче-

ским ожиданием i -го класса и математическим ожиданием всей выборки в искомом пространстве; s_i – дисперсия i -го класса в искомом пространстве; c – количество различных классов в выборке.

Согласно [7], (1) может быть выражено через статистические характеристики исходной обучающей выборки:

$$\max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}; \quad (2)$$

$$\begin{aligned} \mathbf{S}_B &= \sum_{i=1}^c (n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})^2) = \\ &= \sum_{i=1}^c (n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})^T (\boldsymbol{\mu}_i - \boldsymbol{\mu})); \end{aligned} \quad (3)$$

$$\begin{aligned} \mathbf{S}_W &= \sum_{i=1}^c n_i \text{cov}(\mathbf{x}_i) = \sum_{i=1}^c \sum_{j=1}^{n_i} ((\mathbf{x}_{i,j} - \boldsymbol{\mu}_i)^2) = \\ &= \sum_{i=1}^c \sum_{j=1}^{n_i} ((\mathbf{x}_{i,j} - \boldsymbol{\mu}_i)^T (\mathbf{x}_{i,j} - \boldsymbol{\mu}_i)), \end{aligned} \quad (4)$$

где \mathbf{S}_B – ковариационная матрица различий между классами; \mathbf{S}_W – взвешенная сумма ковариационных матриц классов; n_i – длина i -го класса; $\boldsymbol{\mu}_i, \boldsymbol{\mu}$ – векторы-строки математических ожиданий i -го класса и всей выборки соответственно; $\mathbf{x}_{i,j}$ – вектор-строка, содержащий координаты j -го экземпляра i -го класса.

Для решения поставленной задачи оптимизации (2) требуется отыскание собственных векторов $\mathbf{V}_1 - \mathbf{V}_m$ и собственных значений $\lambda_1 - \lambda_m$ матрицы $\mathbf{r} = \mathbf{S}_W^{-1} \mathbf{S}_B$. Аналогично методу главных компонент, результирующая матрица \mathbf{w} , осуществляющая переход в пространство обобщенных признаков, формируется путем объединения собственных векторов, соответствующих наибольшему собственному значению: $\mathbf{w} = [\mathbf{V}_1 \mathbf{V}_2 \dots \mathbf{V}_k]$, (5) где k – размерность формируемого пространства.

Тогда координаты в новом признаковом пространстве могут быть получены путем матричного преобразования:

$$\mathbf{x}^{LDA} = \mathbf{x} \mathbf{w}. \quad (6)$$

Материалы и методы. Проиллюстрируем применение линейного дискриминантного анализа для снижения размерности признакового пространства на примере разработки селектора поврежденных фаз, устанавливаемого для линии ω_1 участка электрической сети, изображенного на рис. 2.

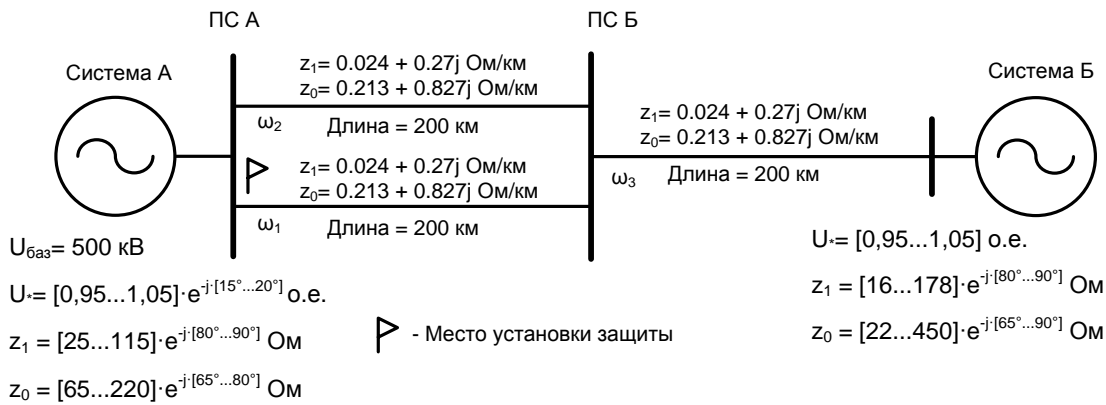


Рис. 2. Анализируемый участок электрической сети

В отличие от типовых решающих правил устройств РЗА, связанных с отключением или неотключением защищаемого объекта, задача селекции поврежденных фаз не является бинарной, а предполагает множественную классификацию. Перечень классов разрабатываемого селектора поврежденных фаз защиты (рис. 2) приведен в табл. 1.

Таблица 1. Перечень классов разрабатываемого селектора поврежденных фаз

Обозначение класса	Описание
K^{ABC}	Трехфазное короткое замыкание
K^{AB}	Двухфазное короткое замыкание фаз А и В (в т.ч., на землю)
K^{BC}	Двухфазное короткое замыкание фаз В и С (в т.ч., на землю)
K^{CA}	Двухфазное короткое замыкание фаз С и А (в т.ч., на землю)
K^{A0}	Однофазное короткое замыкание фазы А
K^{B0}	Однофазное короткое замыкание фазы В
K^{C0}	Однофазное короткое замыкание фазы С

Сформируем обучающую выборку путем многократного имитационного моделирования коротких замыканий (КЗ), перечисленных в табл. 1, на защищаемой линии ω_1 . Постоянные и переменные параметры схемы замещения представлены на рис. 2. Для каждого вида КЗ (табл. 1) реализовалось 2000 имитационных экспериментов с переменными параметрами, взятыми случайно из заданных диапазонов (рис. 2). При этом двухфазные КЗ и двухфазные КЗ на

землю моделировались разными сериями экспериментов. Таким образом, по результатам имитационного моделирования получены выборки размером 4000 элементов для классов K^{AB} , K^{BC} , K^{CA} и 2000 элементов для остальных классов. В качестве исходных информационных признаков выберем действующие значения токов трех фаз I_a , I_b , I_c линии ω_1 , замеренных со стороны ПС А. Таким образом, размерность m исходного признакового пространства равна 3.

Рис. 3 иллюстрирует расположение результатов имитационного моделирования рассматриваемых видов КЗ в исходном трехмерном признаковом пространстве, а также их проекции на плоскости, образованные парами признаков.

Анализ полученных результатов показывает, что плоскости, образованные парами исходных признаков, не способны обеспечить требуемое качество распознавания режимов, так как различные режимы накладываются друг на друга, образуя области взаимного пересечения.

Сформируем уставочную плоскость, образованную двумя обобщенными признаками, полученными в результате сжатия исходного трехмерного пространства признаков до двухмерного методом линейного дискриминантного анализа.

Для того чтобы исключить влияние различий в динамических диапазонах изменения переменных, выполним нормировку результатов имитационного моделирования с использованием выражения

$$x_{i,j} = \frac{\hat{x}_{i,j} - \bar{x}_j}{\sigma_j}, \quad (7)$$

где i – номер текущего эксперимента; j – номер текущего признака; $\hat{x}_{i,j}$ – значение j -го признака в i -м модельном эксперимен-

те; \bar{x}_j – математическое ожидание значений признака j по всем модельным экспериментам; σ_j – среднее квадратичное отклонение признака j по всем модельным экспериментам.

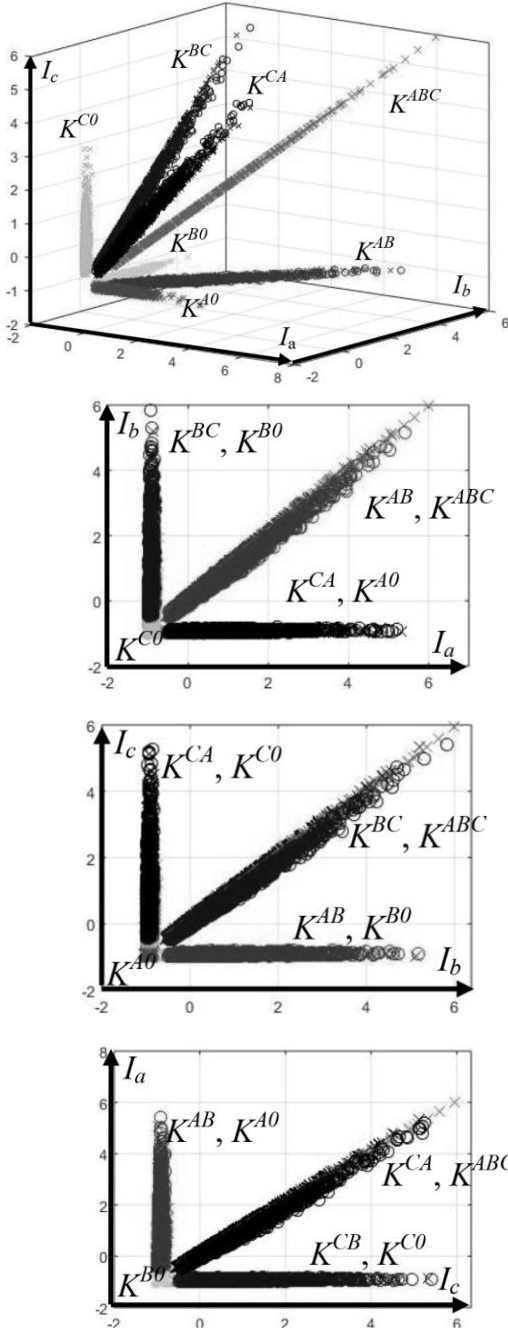


Рис. 3. Расположение результатов имитационного моделирования в исходном признаком пространстве

В результате нормировки все признаки (токи) будут иметь нулевое математическое ожидание и единичную дисперсию: $\mathbf{x} = [x_{i,j}]$, где \mathbf{x} – нормированная обучающая выборка.

Расчеты коэффициентов матрицы преобразования в соответствии с представленным выше алгоритмом требуют определения векторов математических ожиданий и ковариационных матриц каждого класса в отдельности, а также всей выборки в целом (табл. 2).

Таблица 2. Статистические характеристики выборки и отдельных классов в исходном пространстве признаков

Класс	Статистические характеристики
Выборка в целом	$\mu = (0 \ 0 \ 0)$ $\text{cov}(\mathbf{x}) = \begin{pmatrix} 1 & 0,184 & 0,199 \\ 0,184 & 1 & 0,202 \\ 0,199 & 0,202 & 1 \end{pmatrix}$
K^{ABC}	$\mu_1 = (0,896 \ 0,889 \ 0,88)$ $\text{cov}(\mathbf{x}_1) = \begin{pmatrix} 0,896 & 0,867 & 0,867 \\ 0,867 & 0,864 & 0,858 \\ 0,861 & 0,858 & 0,852 \end{pmatrix}$
K^{AB}	$\mu_2 = (0,647 \ 0,649 \ -0,932)$ $\text{cov}(\mathbf{x}_2) = \begin{pmatrix} 0,691 & 0,662 & 0,008 \\ 0,662 & 0,645 & 0,007 \\ 0,008 & 0,007 & 0,001 \end{pmatrix}$
K^{BC}	$\mu_3 = (-0,934 \ 0,667 \ -0,671)$ $\text{cov}(\mathbf{x}_3) = \begin{pmatrix} 0,001 & 0,008 & 0,007 \\ 0,008 & 0,737 & 0,704 \\ 0,007 & 0,704 & 0,686 \end{pmatrix}$
K^{CA}	$\mu_4 = (0,667 \ -0,937 \ -0,643)$ $\text{cov}(\mathbf{x}_4) = \begin{pmatrix} 0,711 & 0,008 & 0,715 \\ 0,008 & 0,001 & 0,008 \\ 0,715 & 0,008 & 0,732 \end{pmatrix}$
K^{A0}	$\mu_5 = (0,208 \ -0,944 \ -0,92)$ $\text{cov}(\mathbf{x}_5) = \begin{pmatrix} 0,269 & 0,004 & 0,002 \\ 0,004 & 0,002 & 0 \\ 0,002 & 0 & 0,003 \end{pmatrix}$
K^{B0}	$\mu_6 = (-0,92 \ -0,225 \ -0,94)$ $\text{cov}(\mathbf{x}_6) = \begin{pmatrix} 0,003 & 0,003 & 0 \\ 0,003 & 0,281 & 0,004 \\ 0 & 0,004 & 0,002 \end{pmatrix}$
K^{C0}	$\mu_7 = (-0,943 \ -0,928 \ -0,214)$ $\text{cov}(\mathbf{x}_7) = \begin{pmatrix} -0,002 & 0 & 0,004 \\ 0 & 0,003 & 0,003 \\ 0,004 & 0,003 & 0,291 \end{pmatrix}$

Выполним расчет ковариационной матрицы различий между классами \mathbf{S}_B по выражению (3), а также взвешенной суммы ковариационных матриц отдельных классов \mathbf{S}_W по выражению (4):

$$\mathbf{S}_B = \sum_{i=1}^c (n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})^T (\boldsymbol{\mu}_i - \boldsymbol{\mu})) = \begin{pmatrix} 1,211 & -0,078 & -0,068 \\ -0,078 & 1,216 & -0,056 \\ -0,068 & -0,056 & 1,203 \end{pmatrix} \cdot 10^4;$$

$$\mathbf{S}_W = \sum_{i=1}^c n_i \cdot \text{cov}(\mathbf{x}_i) = \begin{pmatrix} 7,89 & 4,46 & 4,65 \\ 4,46 & 7,84 & 4,6 \\ 4,65 & 4,6 & 7,968 \end{pmatrix} \cdot 10^3.$$

Тогда получаем

$$\mathbf{r} = \mathbf{S}_W^{-1} \mathbf{S}_B = \begin{pmatrix} 2,76 & -1,03 & -1,12 \\ -1,02 & 2,76 & -1,07 \\ -1,11 & -1,07 & 2,78 \end{pmatrix}.$$

Получим матрицу \mathbf{w} , объединив два собственных вектора матрицы \mathbf{r} , обладающие наибольшими собственными значениями:

$$\mathbf{V}_1 = \begin{pmatrix} 0,583 \\ 0,213 \\ -0,784 \end{pmatrix}; \quad \lambda_1 = 3,9;$$

$$\mathbf{V}_2 = \begin{pmatrix} 0,567 \\ -0,795 \\ 0,217 \end{pmatrix}; \quad \lambda_2 = 3,77;$$

$$\mathbf{V}_3 = \begin{pmatrix} 0,581 \\ 0,57 \\ 0,582 \end{pmatrix}; \quad \lambda_3 = 0,631;$$

$$\mathbf{w} = [\mathbf{V}_1 \ \mathbf{V}_2] = \begin{pmatrix} 0,583 & 0,567 \\ 0,213 & -0,795 \\ -0,784 & 0,217 \end{pmatrix}.$$

Проекция обучающей выборки на искомую уставочную плоскость, рассчитанные в соответствии с выражением (6), представлены на рис. 4, где области, соответствующие различным видам КЗ на ЛЭП

ω_1 , расположились на достаточном расстоянии друг от друга для реализации их корректной классификации.

Сравним полученную уставочную плоскость и плоскость, полученную в результате применения метода главных компонент на тех же модельных данных. Матрицу перехода к пространству главных компонент сформируем из двух собственных векторов корреляционной матрицы обучающей выборки \mathbf{x} , соответствующих наибольшему собственному значению.

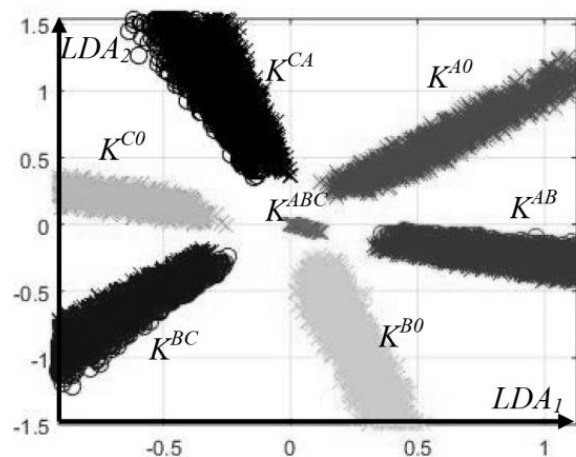
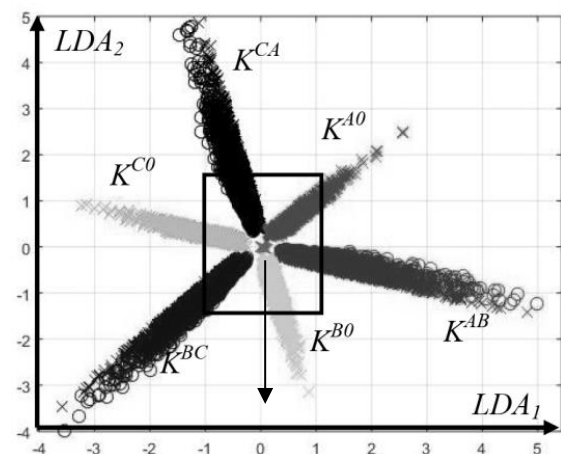


Рис. 4. Проекция обучающей выборки на уставочную плоскость, полученную методом линейного дискриминантного анализа

Результаты проецирования обучающей выборки на оси главных компонент показаны на рис. 5. Несмотря на максимальное возможное значение дисперсии обобщенных признаков, области, соответствующие различным классам на полученной уставочной плоскости, расположены более сгруппировано, чем на рис. 4, и образуют участки взаимного пересечения.

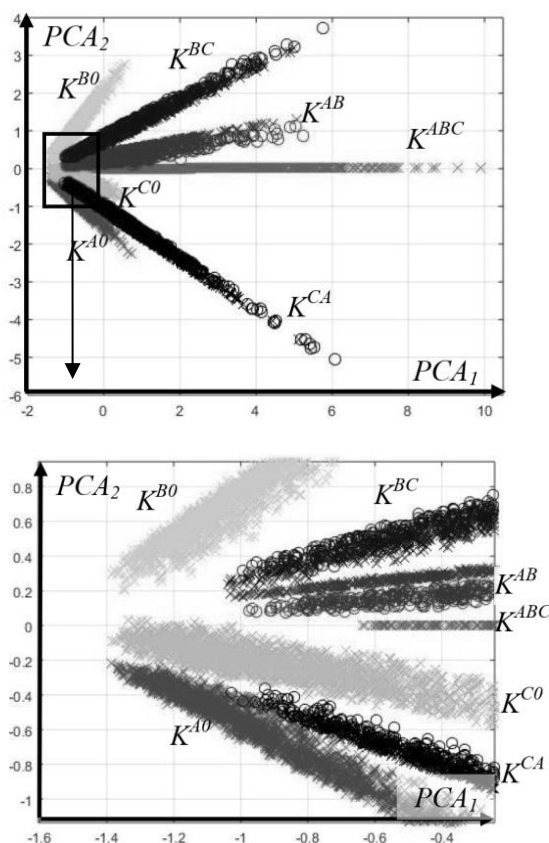


Рис. 5. Проекция обучающей выборки на уставочную плоскость, полученную методом главных компонент

Результаты. Оценим эффект от применения линейного дискриминантного анализа в задаче классификации электрических режимов, сопоставив информативность полученных уставочных плоскостей и уставочных плоскостей, образованных различными парными комбинациями исходных признаков. Критерием информативности признакового пространства может выступать вероятность правильной классификации заданного алгоритма, прошедшего обучение применительно к выбранному признаковому пространству. При этом результаты имитационного моделирования, характеризующие режимы участка электрической сети (см. рис. 2), случайным образом разделим на обучающую выборку, использующуюся для обучения алгоритма классификации, и тестовую выборку для оценки вероятности распознаваемости режимов. По результатам реализации процедуры классификации тестовой выборки обосновывается вывод об эффективности функционирования алгоритма защиты. Выберем в качестве алгоритма распознавания метод k ближайших соседей [8, 9], согласно которому анализируемый режим будет от-

несен к тому классу, представителей которого больше среди k элементов обучающей выборки, расположенных ближе всего к анализируемому электрическому режиму. Параметр k примем равным 100. Поскольку решаемая задача классификации является множественной, конечные результаты функционирования алгоритма распознавания удобнее всего отображать в виде матрицы ошибок (англ. *confusion matrix*) [10]. Матрицы ошибок, полученные с использованием уставочных плоскостей, образованных исходными признаками (рис. 3), сведены в табл. 3.

Таблица 3. Матрицы ошибок двумерных классификаторов, полученные на основе комбинаций пар исходных признаков (%)

Уставочная плоскость Ia-Ib							
	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}
K^{ABC}	97	3	0	0	0	0	0
K^{AB}	51	49,1	0	0	0	0	0
K^{BC}	0	0	79,1	0	0	20,9	0
K^{CA}	0	0	0	72,2	27,9	0	0
K^{A0}	0	0	0	40,2	59,8	0	0
K^{B0}	0	0	45,4	0	0	54,6	0,1
K^{C0}	0	0	0	0	0	0	100
Уставочная плоскость Ib-Ic							
	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}
K^{ABC}	97,45	0	2,55	0	0	0	0
K^{AB}	0	76,9	0	0	0	23,1	0
K^{BC}	50,25	0	49,75	0	0	0	0
K^{CA}	0	0	0	78,5	0	0	21,5
K^{A0}	0	0	0	0	100	0	0
K^{B0}	0	44,5	0	0	0	55,5	0
K^{C0}	0	0	0	45,9	0	0	54,1
Уставочная плоскость Ic-Ia							
	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}
K^{ABC}	96,4	0	0	3,6	0	0	0
K^{AB}	0	79,9	0	0	20,2	0	0
K^{BC}	0	0	79,9	0	0	0	20,1
K^{CA}	52,3	0	0	47,7	0	0	0
K^{A0}	0	48	0	0	52	0	0
K^{B0}	0	0	0	0	0	100	0
K^{C0}	0	0	47,75	0	0	0	52,25

Анализ матриц ошибок (табл. 3) показывает, что уставочные плоскости, образованные исходными признаками, в целом демонстрируют довольно низкую распознающую способность. Наименьшая вероятность правильной классификации наблюдается при междуфазных КЗ, когда в качестве информационных признаков используют токи соответствующих фаз. Наибольшая вероятность характерна для

трехфазных КЗ, а также однофазных КЗ, ток фазы которых не является частью уставочной плоскости.

Оценим распознающую способность алгоритмов классификации режимов с применением обобщенных признаков, полученных на основе методов главных компонент (см. рис. 5) и линейного дискриминантного анализа (см. рис. 4). Результирующие матрицы ошибок сведены в табл. 4.

Таблица 4. Матрицы ошибок двухмерных классификаторов, полученные на основе обобщенных признаков (%)

Признаки, полученные методом главных компонент							
	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}
K^{ABC}	99,9	0,1	0	0	0	0	0
K^{AB}	0,6	99,4	0	0	0	0	0
K^{BC}	0	0,35	99,65	0	0	0	0
K^{CA}	0	0	0	98,8	1,15	0	0,05
K^{A0}	0	0	0	1,25	98,6	0	0,15
K^{B0}	0	0	0	0	0	100	0
K^{C0}	0	0	0	0,05	0	0	99,95

Признаки, полученные линейным дискриминантным анализом							
	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}	K^{ABC}
K^{ABC}	100	0	0	0	0	0	0
K^{AB}	0	100	0	0	0	0	0
K^{BC}	0	0	100	0	0	0	0
K^{CA}	0	0	0	100	0	0	0
K^{A0}	0	0	0	0	100	0	0
K^{B0}	0	0	0	0	0	100	0
K^{C0}	0	0	0	0	0	0	100

Анализ данных табл. 4 показывает, что переход к обобщенным признакам обеспечил существенное увеличение вероятности правильного распознавания электрических режимов. На уставочной плоскости, образованной обобщенными признаками, полученными методом главных компонент, распознающий алгоритм позволяет корректно классифицировать любой режим с вероятностью не менее 98,6 %. Обобщенные признаки, сформированные с применением линейного дискриминантного анализа, обеспечивают реализацию уставочной плоскости, на которой безошибочно идентифицируются все виды коротких замыканий.

Выводы. Снижение размерности признакового пространства является важной задачей при использовании большого количества параметров режима электрической сети в качестве информационных признаков. Оно позволяет снизить вычисли-

тельную нагрузку на реализацию алгоритма распознавания без существенной потери его эффективности.

Линейный дискриминантный анализ обеспечивает переход от исходного признакового пространства к пространству обобщенных признаков. Эффективность метода обуславливается учетом статистических характеристик каждого из классов в отдельности при формировании матрицы преобразования в искомое признаковое пространство.

На примере построения простейшего фазового селектора устройства релейной защиты, использующего в качестве исходных признаков действующие значения токов фаз, проиллюстрирован эффект «сжатия информации» при понижении размерности признакового пространства с 3 до 2 с одновременным увеличением вероятности правильного распознавания режимов участка электрической сети.

Список литературы

1. Куликов А.Л., Шарыгин М.В. Применение статистических критериев распознавания режима релейной защиты сетей электроснабжения // Электротехника. – 2019. – № 2. – С. 58–64.
2. Куликов А.Л., Бездушный Д.И. Анализ реализаций информационного подхода в релейной защите // Релейщик. – 2016. – № 2. – С. 28–33.
3. Hasan, A.N., Pouabe-Eboule P.S., Twala B. The Use of Machine Learning Techniques to Classify Power Transmission Line Fault Types // International Conference on Optimization of Electrical and Electronic Equipment. – Brasov, Romania: IEEE, 2017. – P. 221–226.
4. Лямец Ю.Я., Нудельман Г.С., Зиновьев Д.В. Многомерная релейная защита. Ч. 2. Анализ распознающей способности реле // Электричество. – 2009. – № 11. – С. 9–15.
5. Гладков, Б.В., Роговина И.В., Трофимова Н.А. Метод главных компонент. – М.: МИЭМ, 2010. – 28 с.
6. Куликов А.Л., Бездушный Д.И. Формирование обобщенных признаков срабатывания релейной защиты на основе метода главных компонент // Релейная защита и автоматизация. – 2019. – Т. 34, № 1. – С. 20–27.
7. Linear discriminant analysis: A detailed tutorial / A. Tharwat, T. Gaber, A. Ibrahim, A.E. Hassanien // AI Commun. – 2017. – No. 30. – P. 169–190.
8. Bishop C.M. Pattern Recognition and Machine Learning. – Berlin: Springer, 2006. – 738 p.

9. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. – Second Edition. – Berlin: Springer, 2001. – 745 p.

10. Michie D., Spiegelhalter D., Taylor C. Machine Learning, Neural and Statistical Classification. – Ellis Horwood, 1994. – 290 p.

11. Гантмахер Ф.Р. Теория матриц. – М.: ФИЗМАТЛИТ, 2010. – 560 с.

12. Шнеерсон Э.М. Цифровая релейная защита. – М.: Энергоатомиздат, 2007. – 549 с.

13. Методика оценки информационных признаков, характеризующих режимы работы электрических сетей, по критерию распознаваемости релейной защитой / И.В. Нагай, В.И. Нагай, С.В. Сарры, С.Ю. Березкина // Известия вузов. Северо-Кавказский регион. Технические науки. – 2017. – № 4. – С. 27–33.

14. Нагай И.В. Обеспечение функций дальнего резервирования релейной защиты трансформаторов в условиях продольно-поперечной несимметрии // Известия вузов. Северо-Кавказский регион. Технические науки. – 2011. – № 5. – С. 19–24.

15. Измайлов А.Ф., Солодов М.В. Численные методы оптимизации. – 2-е изд., перераб. и доп. – М.: ФИЗМАТЛИТ, 2008. – 320 с.

16. Папков Б.В. Токи короткого замыкания в электрических системах. – Н. Новгород: Типография НГТУ, 2005. – 276 с.

References

1. Kulikov, A.L., Sharygin, M.V. Primenenie statisticheskikh kriteriev raspoznavaniya rezhima releynoy zashchity setey elektrosnabzheniya [Application of statistic criteria of state parameters recognition of relay protection of electrical grid]. *Elektrotehnika*, 2019, no. 2, pp. 58–64.

2. Kulikov, A.L., Bezdushnyy, D.I. Analiz realizatsiy informatsionnogo podkhoda v releynoy zashchite [Analysis of information approach application]. *Releyshchik*, 2016, no. 2, pp. 28–33.

3. Hasan, A.N., Pouabe-Eboule, P.S., Twala, B. The Use of Machine Learning Techniques to Classify Power Transmission Line Fault Types. International Conference on Optimization of Electrical and Electronic Equipment. Brasov, Romania: IEEE, 2017, pp. 221–226.

4. Lyamets, Yu.Ya., Nudel'man, G.S., Zinov'ev, D.V. Mnogomernaya reley'naya zashchita. Ch. 2. Analiz raspoznavayushchej sposobnosti rele [Multidimensional relay protection. Part 2. Analysis

of recognition property of relay]. *Elektrichestvo*, 2009, no. 11, pp. 9–15.

5. Gladkov, B.V., Rogovina, I.V., Trofimova, N.A. *Metod glavnyh component* [Principal components method]. Moscow: MIEM, 2010. 28 p.

6. Kulikov, A.L., Bezdushnyy, D.I. Formirovanie obobshchennykh priznakov srabatyvaniya releynoj zashchity na osnove metoda glavnyh component [Development of generalized features of actuation of relay protection based on principal components method]. *Reley'naya zashchita i avtomatizatsiya*, 2019, vol. 34, no. 1, pp. 20–27.

7. Tharwat, A., Gaber, T., Ibrahim, A., Hassani, A.E. Linear discriminant analysis: A detailed tutorial. *AI Commun*, 2017, no. 30, pp. 169–190.

8. Bishop, C.M. *Pattern Recognition and Machine Learning*. Berlin: Springer, 2006. 738 p.

9. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning. Berlin: Springer, 2001. 745 p.

10. Michie, D., Spiegelhalter, D., Taylor, C. Machine Learning, Neural and Statistical Classification. Ellis Horwood, 1994. 290 p.

11. Gantmakher, F.R. *Teoriya matrits* [Theory of matrices]. Moscow: FIZMATLIT, 2010. 560 p.

12. Shneerson, E.M. *Tsifrovaya reley'naya zashchita* [Digital relay protection]. Moscow: Energoatomizdat, 2007. 549 p.

13. Nagay, I.V., Nagay, V.I., Sarry, S.V., Bezezina, S.Yu. Metodika otsenki informatsionnykh priznakov, kharakterizuyushchikh rezhimy raboty elektricheskikh setey, po kriteriyu raspoznavаемости releynoy zashchity [Evaluation method of informing features of operation modes of electrical grid in a context of recognition of relay protection]. *Izvestiya vuzov. Severo-Kavkazskiy region. Tekhnicheskie nauki*, 2017, no. 4, pp. 27–33.

14. Nagay, I.V. Obespechenie funktsiy dal'nego rezervirovaniya releynoy zashchity transformatorov v usloviyakh prodol'no-poperechnoy nesimmetrii [Providing function of distant backup protection of transformers in a context of direct and quadrature asymmetry]. *Izvestiya vuzov. Severo-Kavkazskiy region. Tekhnicheskie nauki*, 2011, no. 5, pp. 19–24.

15. Izmaylov, A.F., Solodov, M.V. *Chislennye metody optimizatsii* [Numerical methods of optimization]. Moscow: FIZMATLIT, 2008. 320 p.

16. Papkov, B.V. *Toki korotkogo zamykaniya v elektricheskikh sistemakh* [Short-circuit current in electrical grids]. Nizhniy Novgorod: Tipografiya NGTU, 2005. 276 p.