

ПРИМЕНЕНИЕ ФАКТОРНОГО АНАЛИЗА В ИССЛЕДОВАНИИ ЭНЕРГЕТИЧЕСКИХ СПЕКТРОВ РУД РАЗЛИЧНЫХ МЕТАЛЛОВ

Садреева Ф.А.

Научный руководитель — доцент Сиротин Э.Е.

Сибирский федеральный университет

В разведывании и добыче полезных ископаемых актуальна проблема определения состава породы. Чем быстрее, надежнее и дешевле способ анализа руды или какого-либо другого полезного ископаемого, тем эффективнее будет его исследование и добыча. Существует достаточно много способов определения химического состава. Среди них выделяется рентгенорадиометрический способ, основанный на регистрации вторичного излучения исследуемого материала. Достоинства этого метода заключаются в скорости анализа и бесконтактной методике.

При облучении рентгеном различные химические элементы испускают разное по интенсивности вторичное излучение, на основании чего выдается гистограмма (рис.1), представляющая собой набор спектров распределения квантов от их энергий, зарегистрированных за единицу времени. В зависимости от разрешения принимающего датчика энергетический спектр может включать в себя 256, 512 или 1024 каналов, в зависимости от значений которых определяется состав руды. На данный момент каналы, наиболее тесно связанные с величиной содержания интересующих нас химических элементов, определяются высококвалифицированными специалистами ООО «Радос». Это предприятие, которое разрабатывает, изготавливает и внедряет рудосепараторное оборудование на основе рентгенорадиометрического способа анализа.

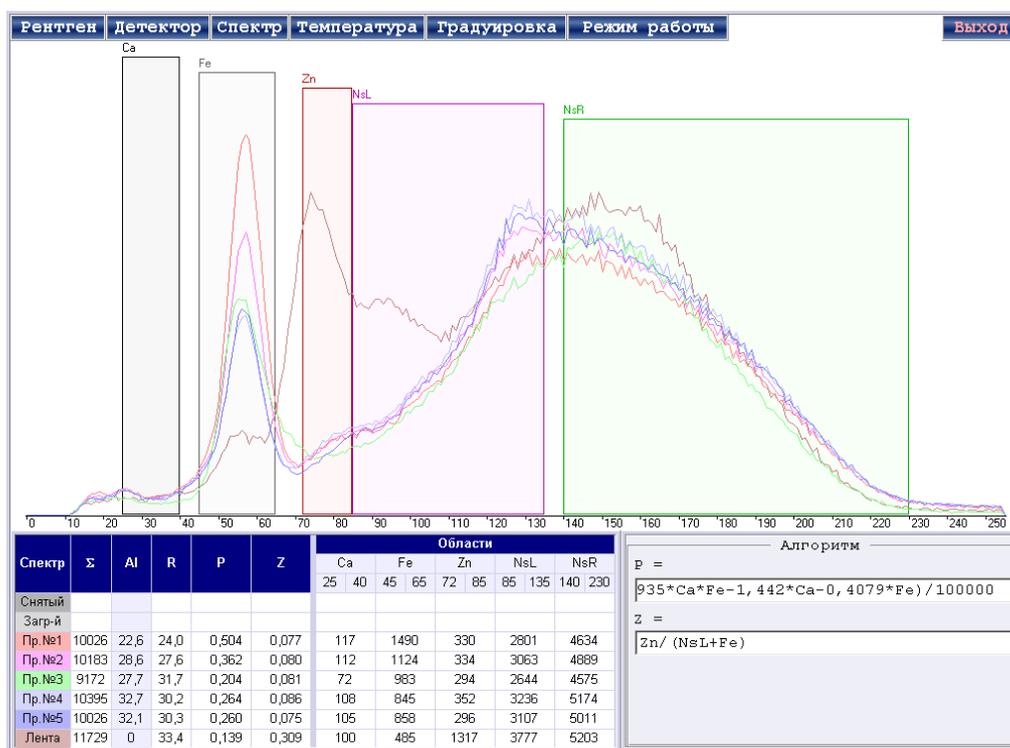


Рис. 1. Энергетические спектры нифелиновых руд АГК с различным содержанием Al_2O_3

Применение факторного анализа (ФА) для сокращения числа каналов энергетических спектров до нескольких значимых переменных позволит упростить работу специалиста или вовсе автоматизировать процесс определения состава.

ФА – раздел статистического многомерного анализа, объединяющий методы оценки размерности множества наблюдаемых переменных посредством исследования структуры ковариационных или корреляционных матриц. Основное предположение ФА заключается в том, что корреляционные связи между большим числом наблюдаемых переменных определяются существованием меньшего числа гипотетических ненаблюдаемых переменных или факторов. В терминах случайных величин – результатов наблюдений X_1, \dots, X_n – общей моделью ФА служит следующая линейная модель:

$$X_i = \sum_{j=1}^k a_{ij} f_j + b_i U_i + \varepsilon_i$$

$i = 1, \dots, n,$

где случайные величины f_j суть общие факторы, случайные величины U_i суть факторы, специфические для величин X_i и не коррелированные с f_j , а ε_i суть случайные ошибки. Предполагается, что $k < n$, случайные величины ε_i независимы между собой и с величинами f_j и U_i имеют $E\varepsilon_i = 0$, $D\varepsilon_i = \sigma_i^2$.

Главными целями ФА являются сокращение числа переменных (редукция данных) и определение структуры взаимосвязей между переменными, т.е. классификация переменных.

В зависимости от задач исследователя следует воспользоваться либо разведочным (эксплораторным), либо конфирматорным ФА. Разведочный факторный ФА осуществляется при исследовании скрытой факторной структуры без предположения о числе факторов и их нагрузках. Конфирматорный предназначен для проверки гипотез о числе факторов и их нагрузках. Так как изначально количество существенных признаков неизвестно, то будет применяться эксплораторный ФА.

В обоих случаях существуют три основных этапа:

1. подготовка соответствующей матрицы ковариаций;
2. выделение первоначальных факторов;
3. вращение с целью получения окончательного решения.

Матрица ковариаций отражает численные значения линейных взаимосвязей между элементами матрицы исходных данных. Существуют известные алгоритмы отыскания матрицы ковариаций, а функция ее отыскания вложена во многие программные пакеты обработки статистических и математических данных.

Основная цель выделения первичных факторов в разведочном ФА заключается в определении минимального числа общих факторов, которые удовлетворительно воспроизводят корреляции между наблюдаемыми переменными.

Существует несколько способов выделения первоначальных факторов. Метод главных компонент возник из анализа главных компонент применительно к редуцированной матрице корреляций. Характеристическое уравнение имеет вид:

$$\det (RI - \lambda I) = 0,$$

где RI – редуцированная корреляционная матрица с оценками общностей на главной диагонали, λ – собственное число. На сегодняшний день этот метод является единственным методом, полностью обоснованным математически и наиболее частотным в применении.

Метод наименьших квадратов, как и метод наименьших остатков, состоит из итерационной процедуры оценки общностей факторов до выделения такого количества

факторов, для которых вычисляемые коэффициенты корреляции наилучшим образом приближают наблюдаемые корреляции.

Согласно методу максимального правдоподобия, необходимо предположить количество значимых факторов и оценивать значения нагрузок на латентные переменные до минимальных.

После выделения первоначальных факторов возникнет упорядоченный набор ортогональных факторов. Необходимо найти легко интерпретируемые факторы, применив последний этап ФА – вращение. Целью всех вращений является получение наиболее простой факторной структуры.

Вращение может быть ортогональным или косоугольным. Косоугольные вращения в свою очередь подразделяется на те, которые основаны на прямом упрощении матрицы коэффициентов факторного отображения, и на те, которые используют упрощение матрицы нагрузок на вторичные оси.

Хотя косоугольное вращение является более общим в сравнении с ортогональным и позволяет выявлять структуру вторичных факторов, в рамках задачи сокращения числа факторов удобнее воспользоваться ортогональным вращением. Оно подразделяется на квартимакс и варимакс. Хотя алгоритмически метод квартимакс проще, чем варимакс, последний дает качественное разделение факторов. Эксперименты показывают, что факторная матрица, получаемая с помощью метода вращения варимакс, в большей степени инвариантна по отношению к выбору различных множеств переменных.

ФА широко применяется в обработке статистических данных в психологии, социологии, экономике. Также имеется ряд работ, в которых описано применение ФА в геологии. Он служит для определения некоторых характеристик месторождений полезных ископаемых, форм и процессов формирования рудных тел.

Геологические объекты, как правило, являются очень сложными, многообразными, так как их состав и поведение обычно обусловлены действием множества разнообразных факторов (причин). Поэтому для более полной характеристики геологических объектов их обычно характеризуют набором разнообразных признаков (параметров), а результаты измерений совокупности этих признаков представляют в виде многомерных случайных величин. При исследовании таких многопараметрических объектов возможность сокращения параметров или замены их меньшим числом каких-либо функций от них может упростить задачу. Методы ФА позволяют это сделать. В исследовании сложных геологических объектов ФА позволяет глубже понять сущность геологического объекта.

ФА применялся в электронной оже-спектроскопии для изучения поведения отдельных компонентов реакции. ФА позволяет определять число независимых компонентов в системе и изменение концентрации каждого компонента в ходе исследуемого процесса. В электронной оже-спектроскопии простое разложение спектральных линий на компоненты в предположении неизменности их формы невозможно, поскольку в оже-спектрах, как правило, помимо химического сдвига происходит и изменение формы линий. В связи с этим, для выделения отдельных химических соединений используется ФА оже-спектров, позволяющий существенно улучшить эффективное спектральное разрешение метода электронной оже-спектроскопии и проследить за поведением отдельных химических компонентов в процессе реакции.

В статье «Применение ФА геолого-геохимических и геофизических параметров для прогноза оруденения на Бобриковском золоторудном месторождении» дано описание применения комбинированного факторного и кластерного анализа.

С помощью матрицы парных корреляций исходных данных проводилось разложение всех участвующих переменных на главные компоненты. По полученной матрице факторных нагрузок были выделены группы переменных, характеризующихся минимальными внутригрупповыми и максимальными межгрупповыми дисперсиями. Это вполне закономерно, согласно исходным данным. Далее группы были проанализированы как несвязанные между собой кластеры.

Методом ФА (методом главных компонент) были выделены более информативные признаки в виде линейных комбинаций тех исходных признаков, для которых характерна синхронная изменчивость и максимальная дисперсия. В результате этого, в небольшом числе факторов была заключена информация о результатах наблюдений в более компактной форме. Таким образом, была получена количественная оценка связи между геофизическими и структурно-вещественными характеристиками месторождения.

Итак, ФА – метод статистики, позволяющий сократить количество значимых переменных в исходных наборах данных.

Существует несколько методов ФА, каждый из которых обладает своими достоинствами и недостатками. Наиболее изученным и математически обоснованным является метод главных компонент с последующим вращением факторов.

ФА широко применяется при геологических исследованиях, в том числе и при анализе спектральных характеристик полезных ископаемых. Однако, в литературе не найдено примеров использования ФА для рентгеноспектрального анализа руд полезных ископаемых. Применение ФА на первом этапе исследования руд металлов может привести к упрощению задачи рентгеноспектрального анализа.