

А. Н. Дмитриев

МАТЕМАТИЧЕСКАЯ ОБРАБОТКА ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ ДАННЫХ В ЗАДАЧАХ ПРОГНОЗНО-ПОИСКОВОГО ПРОФИЛЯ

Наиболее полно и последовательно математическая обработка данных производится специально разработанными методами. Иногда эти методы строятся по принципу "технологической цепочки", в которой методология, теория, решение задач и реализация полученных результатов находятся в тесной взаимосвязи. В данной работе излагаются сведения о специально разработанной совокупности математических методов. Эта совокупность представляет собой довольно полную цепочку правил и математических операций и предназначена для решения класса задач прогнозно-поискового характера (таблица).

Разработанная совокупность правил и приемов первичной обработки данных, а также создание математических подходов обеспечивают растущие запросы геологов в решении теоретических и производственных задач [1-5]. Полнота этого обеспечения не исчерпывающа, но все же охватывает вопросы от общерегионального прогноза [6-8] до оптимизации разведочного процесса [9, 10]. В целом, согласно перечню разработанных методов, вычленены два математических подхода в обработке информации.

1. Тестовый подход - создан на базе идейных и процедурных возможностей дискретной математики в ее логико-комбинаторных направлениях [11-13].

2. Итерационный подход - создан на базе возможностей и следствий, содержащихся в основной теореме о сингулярном разложении [5, 14, 15].

Оба подхода имеют содержательное сходство в плане класса решаемых задач и обладают сходством вычисляемых на таблицах

Технологическая цепочка	Содержание подразделений
Методология и теория геологических постановок задач прогнозно-поискового профиля	Прогнозно-поисковые геологические задачи рассмотрены как информационные проблемы; сформулированы условия постановок задач и правила доалгоритмической обработки данных
Логико-математические разработки	На основе тестового подхода и методов на итерационных основах разработан комплекс логико-математической обработки геолого-геофизических данных в задачах прогнозно-поискового профиля
Программные реализации алгоритмов на ЭВМ	Алгоритмические разработки методов обеспечены программами, организованными в тематические комплексы (на командах машин и программных языках; машины: БЭСМ-6, М-222)
Решение конкретных научных и производственных задач	Решение научно-производственных задач подразделено на два направления: "рудо-прогноз" и "нефтепрогноз"; охвачены объекты прогноза от отдельных участков и локальных структур до крупных рудных узлов и нефтяных провинций
Геологическая интерпретация результатов решения	Результаты решения включают в себя: формирование комплексов признаков, контролирующих рудо- и нефтенакпления, ранжирование объектов и признаков по целеуказанию, прогнозирование перспективных объектов, распознавание и сортировку
Производственное применение результатов решения и экономическая оценка	Применение поисковых критериев для принятия решения по очередности опробования перспективных объектов; коррекция разведочных и поисковых работ на объектах промышленного значения; оценка эффективности применения рекомендаций экономическим расчетом

данных параметров. Подходы нацелены на формализацию операций сравнительного изучения объектов по их описаниям [16, 8].

Элементаризация правил сравнительного изучения геологических объектов выявила, что геологи производят следующие основные операции сравнения:

а) операции по выявлению наличия "сходств" и "различий" с последующей оценкой полноты сходства и/или различия по "основному" признаку;

б) те же операции осуществляются для совокупности признаков с последующей оценкой полноты сходства и/или различия объектов по всему списку признаков;

в) попарное сравнение признаков пар объектов заканчивается, как правило, выявлением преимущества сходства или различия сравниваемых между собой групп описаний геологических объектов.

В соответствии с а) - в), как в тестовом так и в итерационном подходе, было предусмотрено, чтобы математические операции не только могли выявлять сходство и различие сравниваемых объектов, но и устанавливать количественную меру "подобий и разнообразий". В тестовом подходе это требование является основополагающим и соблюдено по существу во всех ведущих алгоритмах.

P-тесты (тесторы), являясь элементарными различающими единицами счета, применяются для установления мер различия объектов между собой [I, II].

Q-тесты (тесторы), являясь элементарными единицами отождествления объектов, применяются для установления мер сходства объектов между собой [I3].

H-тесты (тесторы), являясь элементарными единицами совместного "сходства-различия", применяются для установления взаимной вложенности мер различия и сходства объектов между собой [I3].

С учетом P, Q, H - единиц элементарных показателей различий, сходств, совместности объектов исследования (см. таблицу) и построена вся совокупность алгоритмов тестового подхода, нацеленного на облегчение сравнительных операций при большом числе объектов и характеризующих их признаков.

В итерационном подходе наиболее полно особенностям а) - в) отвечает метод целевой итерационной классификации [17]. Этот метод характеризуется высокой пластичностью в отношении выявления индивидуальной значимости каждого признака на его способность выявлять сходство, различие и совместность.

Часто в практических постановках задач возникает необходимость комплексной оценки групп характеристических признаков, т.е. требуется выяснить коллективную способность признаков различать или отождествлять объекты исследования. Именно для решения подобных задач (связанных в конечном итоге с таксономией объектов) и была предусмотрена алгоритмическая последовательность шагов решения по обработке данных большеформатных таблиц.

Для выяснения коллективной значимости логически сцепленных признаков, в "портретах" геологических объектов большими возможностями обладает метод табличных свойств [18], созданный на основе модификации тестовых алгоритмов. Как и предполагалось, естественное развитие тестового подхода как по Т-свойствам [18], так и по D-тестам [19], вылилось в построение конкретных правил и формализованных процедур, позволяющих "изымать" логически сцепленные признаки, по совокупности которых составляются содержательно интерпретируемые сокращенные (в соответствии с данным целеуказанием) характеристики объектов. Такие характеристики являются основой для выявления трудно вскрываемых взаимосвязей признаков исследуемых месторождений и применяются в последующих шагах решения, при количественном прогнозе полезных ископаемых [6,7,8,20]. В настоящее время этот подход рассматривается в качестве уместного для решения вопросов, связанных с машинным обнаружением закономерностей в исследуемых объектах.

Для выяснения индивидуального значения каждого признака в процессе сравнительного изучения объектов предназначены алгоритмы итерационного подхода. Конечно, эти алгоритмы применимы и для выяснения коллективных свойств признаков. Но введение "поощрений и наказаний" признаков делает этот подход более предпочтительным для задач, формализованная постановка которых требует минимизации пространства признаков для формирования поисковых критериев [8].

Метод согласованных оценок [3, 16] на базе выявления структуры таблицы данных устанавливает меру согласованности каждого столбца (нагрузки признаков) с каждой строкой таблицы (важность объекта) и наоборот. Метод согласованных оценок усиливается возможностями построения таблиц сравнения из таблиц исходных данных. Именно поэтому методы целевой итерационной классификации ("Цикл") и суммарного учета мер приуроченности и согласования ("Каскад") были нацелены на выполнение требований сравнительного изучения объектов в плане учета индивидуальных свойств признаков в их соответствии целевому признаку [8]. Выявление индивидуальной оценки признаков в их способности "различать", "отождествлять", "совмещать" объекты исследования особенно важно в операциях распознавания и в вопросах обнаружения "наиболее существенного признака", при разделении или сжатии исследуемых классов объектов (например, месторождений, рудопроявлений, потенциальных районов). Родство приемов обработки, а также и их различие в разработанных подходах позволяют охватить по существу основные операции сравнительного изучения объектов по их описаниям формализованными процедурами.

По мере практических применений и теоретических разработок геологические и формализованные постановки задач сблизились и оформились в ряд общих требований. Постановка задач обрела устойчивость при соблюдении общих условий:

1. Требуется однозначно заданное целеуказание на обработку данных, соответствующих содержательно сформулированной задаче.

2. Имеется специализированный перечень логических и алгоритмических средств, реализованных программами на ЭВМ.

3. Существует взаимосвязь между геологами и математиками на этапах возникновения задачи, ее решения и интерпретации результатов.

Наиболее типичными внутренними структурами задач прогнозно-поискового профиля являются, например, следующие:

а) на этапе обучения в заданном пространстве признаков, которыми характеризуются эталоны (от нескольких - до сотен), производится поиск информативной системы признаков для проверки однородности перечня объектов по заданному критерию;

б) распознавание объектов-проб (после этапа обучения или без него) на предмет их принадлежности к "продуктивному классу" или отдельному эталону, имеющему промышленное значение;

в) выделение комплекса характеристических признаков, упорядочивающих исследуемые месторождения по целевому признаку (запасы);

г) поиск и выделение информативного комплекса признаков, контролирующих рудо- или нефтепроявления на обширных площадях прогноза;

д) прогнозирование масштаба перспективности объектов-проб, обнаруживающих принадлежность к классу промышленных месторождений.

Вполне закономерен также при постановке и решении конкретных задач учет доалгоритмического этапа обработки данных, включающего в себя мобилизацию и первичную обработку исходной информации [8, 13, 16].

Как уже упоминалось, специализация общей последовательности обработки данных происходит благодаря учету характера обычных геологических исследований. Исследование геологических объектов и явлений производится последовательностью операций сравнения в пространстве характеристических признаков. Отсюда и происходят содержательные требования к математическим средствам [8, 16].

1. Методы должны работать в условиях разной информационной обеспеченности и разнообразных по природе геологических данных.

2. Совокупность методов должна содержать средства для быстрой и эффективной обработки мало- и крупноформатных таблиц в задачах, требующих ориентирующих результатов решения.

3. Необходима достаточно высокая пластичность и взаимная дополнительность методов для охвата многоцелевых постановок задач.

4. Необходимо достаточное разнообразие методов и их возможностей для решения трудноформализуемых задач в их естественном разнообразии.

5. Методы должны давать результаты, удобные для их профессиональной интерпретации как в научном, так и в производственном отношениях, особенно в сфере принятия решений.

Естественно, что эти требования могут выполняться только в условиях специализации разрабатываемых методов, хотя ряд вопросов может успешно решаться готовыми методами прикладной математики. Математическое решение проблем теоретической и производственной геологии в два последних десятилетия обрело повсеместный характер, завершив собой период "болезней роста и сомнений".

В целом, разработанная схема была апробирована по широкому спектру ее звеньев от региональных прогнозов [6, 8, 20] и прогноза продуктивности локальных структур [10, 13] до оптимизации разведочного бурения и прогноза зон развития пород-коллекторов [9].

В плане дальнейшего развития подходов усматриваются следующие возможности.

По тестовому подходу в направлении совершенствования работ по математическому обнаружению закономерностей следует изучить и разработать вопрос ассоциаций признаков и последующей таксономии объектов, для которых возможен поиск закономерностей, содержащихся в совокупности описаний родственных объектов.

По итерационному подходу намечается направление, в котором в качестве объекта распознавания выступают целые таблицы (каждой таблице может соответствовать либо совокупность объектов, либо единичный объект). Важен вопрос о принципах построения классификаций с учетом свойств распознающих алгоритмов.

Л и т е р а т у р а

1. Дмитриев А.Н. и др. Организация и обработка геологической информации с помощью ЭВМ на основе построения тупиковых тестов. - В кн.: Логико-информационные решения геологических задач. М., 1975, с. 83-128.
2. Логико-математическая обработка геологической информации. (Теория и математический аппарат). - Новосибирск: ИГиГ СО АН СССР, 1976, - 151 с.
3. Логико-информационные исследования в геологии. (Теория и математический аппарат). - Новосибирск: ИГиГ СО АН СССР, 1977, - 146 с.

4. Математические методы решения прогнозных задач нефтяной геологии.-Новосибирск: ИГиГ СО АН СССР, 1978.-151 с.
5. Дмитриев А.Н. и др. Метод согласованных оценок.-Новосибирск: ИГиГ СО АН СССР, 1982.-132 с.
6. Трофимук А.А. и др. Распознавание образов гигантских нефтяных месторождений. - В кн.: Проблемы нефтеносности Сибири. Новосибирск, 1971, с. 34-50.
7. Трофимук А.А. и др. Методы прогнозной оценки нефтегазоносных районов Сибири: Препринт № 48. - Новосибирск, 1982. - 51 с. - В надзаг.: ИГиГ СО АН СССР.
8. Федосеев Г.С. и др. Распознавание образов в задачах качественного прогноза рудных месторождений. - Новосибирск: Наука, 1980. - 201 с.
9. Афанасьев С.А. и др. Опыт применения логико-информационных методов для решения некоторых вопросов разведки литолого-стратиграфических залежей. - В кн.: Проблемы литолого-стратиграфических и структурных исследований. Новосибирск, 1980, с. 73-91.
10. Дмитриев А.Н. Логико-математические средства в задачах прогнозно-поискового профиля. - В кн.: Геолого-геохимические критерии нефтегазоносности. Новосибирск, 1976, с. 14-32.
11. Дмитриев А.Н., Журавлев Ю.И., Кренделев Ф.П. Об одном принципе классификации и прогноза геологических объектов и явлений. - Геология и геофизика, 1968, № 5, с. 50-64.
12. Дмитриев А.Н., Красавчиков В.О. Тестовый подход в решении проблем обработки геологической информации. - В кн.: Логико-информационные исследования в геологии. Новосибирск, 1977, с. 3-47.
13. Дмитриев А.Н. Новые тестовые разработки в задачах прогнозирования рудоносности. - В кн.: Математические методы при прогнозе рудоносности. М., 1977, с. 104-163.
14. Васильев Ю.Л., Дмитриев А.Н. Спектральный подход к сравнению объектов, охарактеризованных набором признаков. - Докл.АН СССР, 1972, т. 206, № 6, с. 1309-1312.
15. Дмитриев А.Н. Описание метода согласованных оценок. - В кн.: Метод согласованных оценок. Новосибирск, 1982, с. 10-30.

16. Дмитриев А.Н. Вопросы формализованных постановок геологических задач прогнозно-поискового профиля. - В кн.: Логико-математическая обработка геологической информации. Новосибирск, 1976, с. 3-22.
17. Бышаев А.А. Метод "Целевая итерационная классификация". - В кн.: Логико-математическая обработка геологической информации. Новосибирск, 1976, с. 70-92.
18. Красавчиков В.О. Комплекс алгоритмов для программы распознавания по Т-свойствам. - В кн.: Программные комплексы для целевой обработки информации. Новосибирск, 1977, с. 151-162.
19. Смертин Е.А. Вопросы теории и алгоритмы на базе построения Д-тестов. - В кн.: Логико-информационные исследования в геологии. Новосибирск, 1977, с. 48-67.
20. Дмитриев А.Н. и др. Примеры конкретных решений. - В кн.: Метод согласованных оценок. Новосибирск, 1982, с.96-112.