

Секция «Математика и механика»

Построение корректной комбинации алгоритмов вычисления оценок,
настроенных методом скользящего контроля.

Игнатъев Олег Анатольевич

Аспирант

*Московский государственный университет имени М.В. Ломоносова, Факультет
вычислительной математики и кибернетики, Москва, Россия*

E-mail: kcdoleg@gmail.com

В данной работе рассматривается алгебраический подход к алгоритмам распознавания, предложенный Ю.И. Журавлевым и его учениками. Алгебраический подход предполагает введение нескольких базовых операций над алгоритмами для того, чтобы исследовать возможность решения задачи распознавания при помощи различных комбинаций этих алгоритмов. В некоторых случаях класс используемых алгоритмов может быть достаточно бедным, но замыкание этого класса с помощью некоторых операций может приводить к классу алгоритмов, решающих достаточно сложные задачи. Используемый подход предоставляет прежде всего информацию о том, правда ли, что выбранный класс используемых алгоритмов позволяет нам решить поставленную задачу. Чаще всего проверяется, есть ли в замыкании корректный алгоритм, то есть алгоритм, не допускающий ни одной ошибки на контрольной выборке. В качестве базового класса алгоритмов обычно рассматривается класс алгоритмов вычисления оценок (далее АВО), так как алгоритмы из этого класса просто описываются и в то же время являются достаточно гибкими. Алгоритм из класса АВО позволяет регулировать веса обучающих объектов, веса опорных множеств признаков и использует оператор близости между двумя объектами. Результатом работы алгоритма из класса АВО является матрица оценок Γ (Γ_{ij} — оценка принадлежности объекта i к классу j). В качестве операций над алгоритмами (или над матрицами оценок, что тоже самое) обычно используются следующие:

1. Сложение.
2. Умножение на число.
3. Покомпонентное умножение.

Используя эти операции над классом алгоритмов мы получаем линейное пространство. Класс алгоритмов, которые получаются с помощью первых двух операций называется линейным замыканием. Класс алгоритмов, которые получаются с помощью всех трех операций называется алгебраическим замыканием. Ю.И. Журавлевым и его учениками доказано, что при определенных условиях, которые называются условиями регулярности, можно построить корректный алгоритм для любой контрольной выборки. В данной работе рассматривается следующая задача. Рассмотрим метод скользящего контроля leave one out применительно к алгоритмам из класса АВО для некоторой фиксированной обучающей выборки. Ставится вопрос, можно ли комбинируя алгоритмы, которые получаются при обучении этим методом построить алгоритм, корректный на всей обучающей выборке. Таким образом мы рассматриваем теоретическую возможность построить алгоритм, который с одной стороны не должен быть склонен к переобучению,

так как любой из алгоритмов не использует в обучении объект, для которого мы хотим получить правильную классификацию. С другой стороны, если эта комбинация корректна на всей выборке разумно предполагать, что сама комбинация достаточно хорошо решает задачу распознавания. Для полученной комбинации необходимо оценить максимальную степень, входящего в неё алгоритма, а также исследовать устойчивость комбинации. Устойчивость комбинации — важное свойство, позволяющее сказать, что алгоритмы могут быть получены не единственным методом, так как существует некоторая окрестность в пространстве начальных информаций, внутри которой результат распознавания остается неизменным. В работы получены следующие результаты:

- Построена комбинация алгоритмов, которая решает поставленную задачу. Для построения алгоритмов используются операторы разметки из работ А.Г. Дьяконова, а также их некоторые модификации.
- Получены достаточные условия для построения корректной комбинации.
- Дана оценка на максимальную степень используемых алгоритмов, достаточную для решения задачи.
- Доказана устойчивость комбинации алгоритмов в некоторой окрестности начальных информаций. Также получены достаточные условия для устойчивости.

Литература

1. Журавлев Ю.И. Корректные алгебры над множествами некорректных (эвристических алгоритмов). I. // "Кибернетика"К., 1977, № 4.
2. Журавлев Ю.И. Корректные алгебры над множествами некорректных (эвристических алгоритмов). II. // "Кибернетика"К., 1977, № 6.
3. Журавлев Ю.И. Корректные алгебры над множествами некорректных (эвристических алгоритмов). III. // "Кибернетика"К., 1978, № 2.
4. Дьяконов А.Г. Исследование алгебраических замыканий алгоритмов распознавания: операторы разметки // Таврический вестник информатики и математики, 2008, №1, С.199-203.

Слова благодарности

Работа поддержана грантом РФФИ, номер проекта 14-07-00965.