

О.П. Бардачева
Криворожский государственный
педагогический университет
IV курс, группа И-05
Научный руководитель:
к.пед.н., доц. С.А. Семериков

ВОССТАНОВЛЕНИЯ ЗАВИСИМОСТЕЙ ПО ВЫБОРКАМ ОГРАНИЧЕННОГО ОБЪЕМА

Основная проблема, возникающая из-за ограниченности объема эмпирических данных, состоит в правильном соотнесении «сложности» приближающей функции с количеством имеющихся данных.

Рассмотрены три задачи восстановления зависимости:

- задача обучения распознавания образов;
- задача восстановления регрессии;
- задача интерпретации результатов косвенных экспериментов.

Для ограниченного объема выборки учет «сложности» приближающей функции приводит к установлению наилучшего подпространства признаков при распознавании образов, оптимального числа членов ряда при восстановлении регрессии, наилучшего регуляризатора при интерпретации результатов косвенных экспериментов.

Для ограниченного объема информации целесообразно различать в задаче восстановления зависимостей две постановки:

- восстановление функциональной зависимости;
- восстановление значений функции в заданных точках.

При восстановлении функциональных зависимостей могут быть выделены различные постановки задач, которые сводятся к одной и той же математической схеме – минимизации среднего риска по эмпирическим данным. Эти постановки различаются тем, в каком классе функций ведется восстановление искомой зависимости: в или самом простом классе – классе индикаторных функций (задача распознавания образов); или в классе функций более сложной природы, что принадлежат множеству функций, которые интегрируются с квадратом (задача восстановления регрессии); или, наконец, в классе функций, которые являются образом некоторого другого класса функций (задача интерпретации результатов не прямых экспериментов).

Однако эти особенности задач восстановления зависимостей являются частными, и поэтому основное направление исследований в каждой из них связано с изучением общей схемы минимизации среднего риска по эмпирическим данным.

Все три задачи восстановления зависимостей сводятся к одной и той же схеме – схеме минимизации среднего риска, возможно лишь, приближенное решение задачи минимизации среднего риска по эмпирическим данным. Обеспечит ли приближенное решение этой задачи нужную близость найденной зависимости к истинной?

Ответ на этот вопрос различен для разных задач восстановления зависимостей. Для задачи обучения распознаванию образов ответ безусловный – да, обеспечит просто по определению (ведь согласно постановке требуется найти функцию, доставляющую функционалу величину, близкую к минимальной).

В задаче восстановления регрессии ответ не столь определенный. Легко можно сказать, что если близость функций понимать в смысле метрики, то из близости функционала к минимальному следует близость найденной функции к регрессии.

В задаче интерпретации результатов косвенных экспериментов близость функционала к минимальному не гарантирует близость восстанавливаемой функции к искомой

Таким образом, несмотря на то, что для всех задач восстановления зависимостей функции, доставляющие точный минимум функционалу, определяют решение, приближенная минимизация не всегда приводит к цели. Поэтому, прежде чем применить конкретный метод минимизации среднего риска по эмпирическим данным, необходимо убедиться, что этот метод минимизации обеспечит приближение к искомому решению.