

# Отчет о вычислительном эксперименте.

## Введение

Алгоритм нахождения изоморфных деревьев применяется для упрощения суперпозиции элементарных функций. Суперпозиция строится, как приближение функциональной зависимости данных по биржевым опционам.

## Унифицирование графа

Имеющееся дерево суперпозиции может содержать изоморфные ветки. Мы приравняем параметры в соответствующих вершинах веток между собой, чтобы получить упрощенную модель. Таким образом, алгоритм унифицирования заключается в уменьшении размерности вектора параметров регрессионной модели.

В эксперименте использовался следующий функционал ошибки:

$$E = \sum_{i=1}^N (f(x_i) - y_i) + C \cdot \dim(\Omega);$$

где  $\Omega$  - вектор параметров, а  $f$  - нелинейная суперпозиция.

Было положено, что  $C = 0.01$ . На генерируемых в эксперименте моделях  $\dim(\Omega)$  почти всегда не превышает 15. Ожидалось уменьшение функционала ошибки преобразованной модели.

В следующей таблице приведены результаты для упрощаемых моделей.

Начальное число параметров	Ошибка	Итоговое число параметров	Ошибка
9	6.821454e+22	6	1.404598e+21
8	1.810071e+09	6	2.600142e-01
13	3.627707e-01	11	3.427702e-01
15	2.919851e-01	12	2.765169e-01
4	2.400142e-01	2	2.200142e-01
9	2.369950e-01	6	2.369241e-01
4	2.166848e-01	2	1.900595e-01
4	2.593220e-01	2	2.499030e-01
6	8.985552e-01	4	8.785552e-01
8	1.829372e-01	6	1.629372e-01

Результат оправдал ожидания - функционал ошибки уменьшается с упрощением модели.

Пример упрощенной алгоритмом модели:

$plus2([], plus2([], plus2([], normal(w_{1,3}, x_1, normal(w_{4,6}, x_2), normal(w_{7,9}, x_1))), x_1)$

Упрощается к виду:

$plus2([], plus2([], plus2([], normal(w_{1,3}, x_1, normal(w_{4,6}, x_2), normal(w_{1,3}, x_1))), x_1)$

## Упрощение по шаблону

Развивается предложенная в [1] идея упрощения графа по заданным шаблонам. В [1] рассматривается случай упрощения только для поддеревьев, для которых ребра из корня ведут только в листья. С помощью алгоритма нахождения изоморфных поддеревьев

возможно применять правила замены для произвольной вершины. Было выбрано два правила и показано, как алгоритм работает для каждого из них

$$\cos(f) * \sin(f) = 0.5 \cdot \sin(2 \cdot f)$$

$$plus(f, f) = 2 \cdot f$$

где  $f$  - произвольная функция.

Упрощаемые строки генерируются вручную. В этой части алгоритма для простоты не использовалась структура *model*, а потому вместо использования реализованных ранее *ExtractSubModel*, *ReplaceSubModel* использовался упрощенный вариант реализации алгоритма замены подстрок.

Результат работы алгоритма приведен ниже:

$$plus2(expl(times2(sin(expl(x1)), cos(expl(x1)))), plus(x1)) \longrightarrow plus2(expl(0.5 * sin(2 * expl(x1))), plus(x1))$$

$$plus2(expl(sin(expl(x1))), expl(sin(expl(x1)))) \longrightarrow 2 * expl(sin(expl(x1)))$$

$$plus2(expl(times2(sin(expl(x2)), cos(expl(x1)))), plus(x1)) \longrightarrow plus2(expl(times2(sin(expl(x2)), cos(expl(x1))))$$

Таким образом, для сложных  $f$  данные правила дают возможность вычислять на одну ветку меньше, что существенно упрощает модель - в том числе и в рамках уменьшения размерности вектора параметров.

## Литература

- [1] Г. И. Рудой and В. В. Стрижов. Упрощение суперпозиций элементарных функций при помощи преобразований графов по правилам. In *Интеллектуализация обработки информации. Доклады 9-й международной конференции*, pages 140–143, 2012.