Об использующем суперпозиции способе эффективного вычисления систем линейных полиномов с целыми коэффициентами

А. И. Мамонтов (НИУ «МЭИ», Москва)

Рассматривается система линейных полиномов с целыми коэффициентами. Предлагается схема эффективного вычисления этой системы. Схема вычислений позволяет устранить избыточность представления исходных данных, а также упростить вычисления устранением многократности вычислений одних и тех же значений при той же функциональности, что и без использования схемы. Приводится пример применения предлагаемой схемы вычислений.

Ключевые слова: линейный полином, сложность вычислений, классификация.

Введение

Данная статья посвящена вычислению значений систем линейных полиномов с целыми коэффициентами. В ней приводится основанная на применении суперпозиций схема эффективного вычисления и хранения таких систем. Также рассматриваются примеры применения этой схемы при классификации данных.

Следует отметить, что методы организации вычислений значений полиномов, а также методы удобного для таких вычислений представления знаний, активно развиваются. В качестве примера можно привести работы по схемам вычислений вещественных полиномов. Э. Г. Белага доказал невозможность построения схемы требующей определенного числа операций [1]. В. Е. Пан построил схемы близкие к оптимальным [2]. Рассматриваются вычисления на равномерных сетках, позволяющие получить лучшие оценки [3]. Значительно количество работ по определяющей скорость вычисления сложности полиномиальных представлений k-значных функций [4, 5]. При организации вычислений и представлении знаний часто возникают различные алгоритмы выделения общих подформул [6]. Представления полиномов с целыми коэффициентами рассматриваются, например, в [7, 8]. Есть и исследования о представлении с помощью линейных и нелинейных полиномов нейронных сетей — популярных конструкций, которые используются при классификации [9, 10].

Постановка задачи

Дана система линейных полиномов:

$$P_1(x_1, \dots, x_n) = a_{01} + \sum_{i=1}^{n} a_{i1} x_i,$$

$$\dots$$

$$P_m(x_1, \dots, x_n) = a_{0m} + \sum_{i=1}^{n} a_{im} x_i,$$

где a_{ij} — целые числа.

При этом $|a_{ij}| < A$, где A — некоторое положительное целое число.

Требуется многократно вычислять эту систему полиномов при различных значениях $(x_1, \dots, x_n)$.

Предлагается следующая схема выслений. **1.** Вычисляются все умножения $a_{ij}x_i$. При этом, процедура вычислений организовывается таким образом, что если при некоторых j_1, j_2 выполняется $a_{ij_1} = a_{ij_2}$, то $a_{ij_1}x_i = a_{ij_2}x_i$ должно вычисляться только один раз.

2. Выполняются все сложения.

При такой схеме вычислений на первом этапе производится не более $2An$ умножений, так как число различных коэффициентов a_{ij} не более $2A$. На втором этапе производится mn сложений.

Следует отметить, что оценка числа умножений не зависит от количества полиномов, что дает возможность получить существенный выигрыш при большом числе полиномов. Поскольку при традиционной схеме проводится mn умножений, а не $2An$, как в пунктах 1,2 предлагаемой нами схемы.

Дальнейшая оптимизация. Пусть у двух полиномов есть совпадающая часть, то есть при некоторых j_1, $j_2 \in \{1, \dots, m\}$ и некоторых

$i_1, \ldots, i_k \in \{1, \ldots, n\}$ выполняются равенства $a_{i_1 j_1} = a_{i_1 j_2}, \ldots a_{i_n j_1} = a_{i_n j_2}$. Обозначим

$$P(x_1, \ldots, x_n) = \sum_{t=1}^{k} a_{i_t} x_{i_t},$$
$$R_{i_1}(x_1, \ldots, x_n) = P_{i_1}(x_1, \ldots, x_n) - P(x_1, \ldots, x_n),$$
$$R_{i_2}(x_1, \ldots, x_n) = P_{i_2}(x_1, \ldots, x_n) - P(x_1, \ldots, x_n).$$

Следует отметить, что $R_{i_1}(x_1, \ldots, x_n)$ и $R_{i_2}(x_1, \ldots, x_n)$ — по построению линейные полиномы, тогда

$$P_{i_1}(x_1, \ldots, x_n) = R_{i_1}(x_1, \ldots, x_n) + P(x_1, \ldots, x_n),$$
$$P_{i_2}(x_1, \ldots, x_n) = R_{i_2}(x_1, \ldots, x_n) + P(x_1, \ldots, x_n).$$

Предлагается сначала вычислять с помощью сложений $P(x_1, \ldots, x_n)$, а затем $R_{i_1}(x_1, \ldots, x_n)$, $R_{i_2}(x_1, \ldots, x_n)$ и $P_{i_1}(x_1, \ldots, x_n)$, $P_{i_2}(x_1, \ldots, x_n)$. В этом случае объем вычислений сокращается на количество переменных в полиноме P, то есть на $(k-1)$.

По нашему мнению, вместо хранения в базе данных всех коэффициентов полиномов P_{i1}, P_{i2}, рациональнее хранить информацию только о том, что полиномы P_{i1}, P_{i2} есть суммы других полиномов, а также коэффициенты полиномов P, R_{i_1}, R_{i_2}. Существенного уменьшения количества сложений и экономии памяти в базе данных можно достичь, если у многих полиномов есть совпадающие части.

Поиск совпадающих частей можно выполнять:

1) с помощью переборных алгоритмов,
2) используя предварительные знания, пришедшие из практической задачи о возникновении системы и планируемых совпадениях в конкретных полиномах,
3) с помощью приближенных методов.

В третьем случае задачу поиска повторяющихся частей в полиномах можно сводить к задаче кластеризации векторов, использовать метрику Хемминга — процент совпадающих элементов в полиномах. Вообще, задача получения одних полиномов из других довольно тяжелая [7].

Примеры применения

Предложенная схема вычислений может применяться, например, при классификации данных, а именно, при классификации текстов и изображений. При классификации часто используются линейные полиномы с вещественными коэффициентами. В некоторых случаях можно перейти к целым.

Пример. Пусть имеются изображения, содержащие пейзажи без промышленных объектов (природные изображения), изображения, содержащие промышленные объекты (промышленные изображения), а также изображения, содержащие дорожные знаки (подкласс промышленных изображений). Допустим, что есть некоторая обучающая выборка. При обучении получены следующие результаты:

- Линии встретились в 0,5 изображений, не относящихся к промышленным;
- Углы встретились в 0,01 изображений, не относящихся к промышленным;
- Треугольники встретились в 0,01 изображений, не относящихся к промышленным;
- Линии встретились в 0,75 изображений, не относящихся к дорожным знакам;
- Углы встретились в 0,01 изображений, не относящихся к дорожным знакам;
- Треугольники встретились в 0,01 изображений, не относящихся к дорожным знакам.

Полученные вероятности используются при работе байесовского классификатора, при этом вектор коэффициентов будут иметь следующий вид:

$$(0,5; 0,01; 0,01), (0,75; 0,01; 0,01).$$

Далее рассмотрим логарифмы

$$\ln 0,5 = -0.693, \ln 0,01 = -4,605, \ln 0,75 = -0,288.$$

В результате округления получаем вектора коэффициентов полиномов $(1; 5; 5)$, $(0; 5; 5)$ и полиномы

$$f_1(x_1, x_2, x_3) = x_1 + 5x_2 + 5x_3, f_2(x_1, x_2, x_3) = 5x_2 + 5x_3.$$

У обозначенных полиномов совпадают коэффициенты при некоторых переменных, потому введем полиномы следующего вида:

$$s(x, y) = x + y, f_{12}(x, y) = 5x + 5y, f_{11}(x) = x, f_{22}(x) = 0,$$

тогда

$$f_1(x_1, x_2, x_3) = s(f_{11}(x_1), f_{12}(x_2, x_3)),$$
$$f_2(x_1, x_2, x_3) = s(f_{22}(x_1), f_{12}(x_2, x_3)).$$

Для хранения коэффициентов полиномов f_{11}, f_{12}, f_{22} в памяти ЭВМ нужно использовать четыре совокупности ячеек памяти, а для хранения коэффициентов байесовского классификатора — шесть совокупностей ячеек, таким образом используется меньше памяти ЭВМ.

Замечание про классификацию текстов. При классификации текстов обычно требуется узнать количество ключевых слов, встречаемых в исследуемых текстах. При этом возникает задача поиска массива слов в тексте, которая эффективно решается с помощью алгоритма Ахо — Корасик. На следующем этапе, для эффективных вычислений линейных сумм можно использовать схемы вычислений, приведенные в данной статье.

Преимущества

Итак, использование приведенной выше схемы, позволяет устранить избыточность представления исходных данных, а также упростить вычисления устранением многократности вычислений одних и тех же значений при той же функциональности, что и без использования схемы.

Работа выполнена при финансовой поддержке РФФИ (грант 14–01–00671а).

Список литературы

[1] Белага Э. Г. О вычислении значений многочлена от одного переменного с предварительной обработкой коэффициентов // Проблемы кибернетики. — Вып. 5. — М.: Физматгиз, 1961. — С. 7–15.

[2] Пан В. Я. Некоторые схемы для вычисления значений полиномов с вещественными коэффициентами. // Проблемы кибернетики. — Вып. 5. — М.: Наука, 1961. — С. 17–29.

[3] Bachmann O., Wang P. S., Zima E. V. Chains of Recurrences a method to expedite the evaluation of closed-form functions. // Proc. of the ISSAC'94. — Oxford, UK: ACM Press, July 1994. — P. 242–249.

[4] Селезнева С. Н. Сложность систем функций алгебры логики и систем функций трехзначной логики в классах поляризованных полиномиальных форм // Дискретная математика. — Т. 27. Вып. 1. — 2015. — С. 111–122.

[5] Маркелов Н. К. Нижняя оценка сложности функций трехзначной логики в классе поляризованных полиномов // Вестник Моск. унив-та. Сер. 15. Вычисл. математика и кибернетика. — Вып. 3. — 2012. — С. 40–45.

[6] Косовская Т. М. Самообучающаяся сеть с ячейками, реализующими предикативные формулы // Труды СПИИРАН. — Вып. 6. — 2015. — С. 94–113.

[7] Алексиадис Н. Ф. Алгоритмическая неразрешимость проблемы полноты для полиномов с целыми коэффициентами // Вестник МЭИ. — Вып. 3. — 2015. — С. 110–117.

[8] Мамонтов А. И., Мещанинов Д. Г. Алгоритм распознавания полноты в функциональной системе $L(Z)$ // Дискретная математика. — Т. 26. Вып. 1. — 2014. — С. 85–95.

[9] Gorban A. N., Wunsch D. C. II. The General Approximation Theorem // In Proceedings IJCNN'98, IEEE. — 1998. — P. 1271–1274.

[10] Половников В. С. О нелинейных характеристиках нейронных схем в произвольных базисах // Интеллектуальные системы. — Т. 17. Вып. 1–4. — 2013. — С. 87–90.