

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. Ломоносова

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ и КИБЕРНЕТИКИ

П Р О Г Р А М М А

VI Международной научно-практической конференции
«Современные информационные технологии и
ИТ-образование»

12 декабря – 14 декабря 2011 г.
Москва

Многоцелевые алгоритмы

Исламов Г.Г., УдГУ, профессор, ggislamov@gmail.com

Введение

Высокопроизводительные вычисления, как правило, ассоциируются с применением высокопроизводительной вычислительной техники. Однако, это не единственно возможный путь ускорения расчётов. Мы считаем, что огромные резервы увеличения производительности вычислений можно обнаружить при правильном выборе алгоритма для конкретных типов задач, когда в процессе вычислений одновременно удаётся решить весь класс проблем, относящихся к данной задаче. Такой алгоритм мы называем многоцелевым. Ярким примером многоцелевого алгоритма является алгоритм Горнера деления многочлена от одной переменной на линейный многочлен. В ходе его исполнения для многочлена, заданного координатами в степенной базе, можно одновременно получить переразложение по базисам Тейлора, Ньютона и Эрмита. Менее тривиальные примеры многоцелевых алгоритмов будут приведены ниже.

Быстродействие современных компьютеров и развитые средства межпроцессорной обработки данных предъявляют повышенные требования к форме представления арифметических данных и точности арифметических вычислений. Некоторая опасность таится в попытке организации высокопроизводительных вычислений на базе машинной арифметики, реализованной в стандартных системах программирования C++ и CUDA [1]. Причина в том, что принятая в них форма представления отрицательных целых чисел сужает диапазон целых положительных чисел, представимых фиксированным числом двоичных разрядов. Кроме того, лежащий в основе представления вещественных чисел стандарт IEEE 754 по точности вычислений значительно уступает машинной арифметике с 40 двоичными разрядами под мантиссу при одинарной точности и 80 двоичными разрядами при двойной точности вычислений, реализованной в отечественной ЭВМ БЭСМ-6 ([2], с. 24). Большой объём вычислений, как правило, сопровождается значительной абсолютной и относительной погрешностью конечного результата. Поэтому, эффективные высокопроизводительные вычисления можно организовать лишь на базе специализированных вычислительных процессоров [3]. Эмуляция таких процессоров программными и микропрограммными средствами возможна и представляет собой типичное решение, однако более перспективной

нам представляется разработка соответствующей интегральной схемы на СБИС.

В настоящей работе рассматриваются основы математической теории спецпроцессора высокопроизводительных вычислений, реализующего многоцелевые алгоритмы для определённого класса практических задач.

1 Фундаментальная структура числовой информации

Теория представления алгебраических структур, обладающих важными для приложений свойствами, приводит нас к пониманию фундаментальной роли кольца матриц над заданным числовым полем, в общем случае телом (конечномерной алгебры с делением). Как известно ([4], с. 464), n - мерная ассоциативная алгебра A над полем P изоморфна некоторой подалгебре в $M_k(P)$, где $k \leq n+1$. Здесь

$M_k(P)$ есть кольцо матриц порядка k с элементами из поля P .

Теорема Веддербарна-Артина ([5], с. 122) доставляет примеры колец с единицей, изоморфных прямой сумме конечного числа полных колец матриц над некоторыми телами. Знаменитая теорема Фробениуса утверждает, что над полем вещественных чисел R существует лишь три конечномерных тела: R , C и H ([5], с. 202). Здесь C - поле комплексных чисел, изоморфное алгебре квадратных матриц вида

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \approx a + bi, \text{ где } a, b \in R; \quad H - \text{ алгебра кватернионов}$$

Гамильтона, изоморфная алгебре квадратных матриц вида

$$\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} \approx \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k, \text{ где } u = \alpha_0 + \alpha_1 i, v = \alpha_2 + \alpha_3 i \in C$$

([6], с. 137) Теория линейных представлений конечных групп ([4], с. 347) указывает на важную роль кольца матриц над полем вещественных чисел при изучении свойств симметрии окружающего нас мира. В теории формальных языков ([7], с. 30) показывается эквивалентность для языка L следующих четырёх условий: L представим; L матрично представим; L представим конечным детерминированным автоматом; L является языком типа 3 в иерархии Хомского. Здесь используется морфизм, отображающий совокупность всех слов над заданным алфавитом Σ в мультипликативный моноид квадратных матриц. Другой пример применения матричной техники к изучению конечных распознавателей и регулярных алгебр приводится в курсе лекций ([8], с. 339) и монографии ([9], с. 127).

Сказанное позволяет нам при описании структурной организации числовой информации вычислительных моделей ограничиться агрегатами из матричных структур над полем вещественных чисел.

2 Универсальная операция над матричными структурами

Стандартные бинарные операции в алгебре матриц $M_n(R)$ порядка n над полем R вещественных чисел дополним унарной операцией, которая в работе [10] названа *универсальной операцией над матричными структурами*. Рассмотрим две группы переменных $(x_1, \dots, x_n)$ и $(y_1, \dots, y_n)$. Линейное отображение $y_i = \sum_{j=1}^n a_{ij} x_j, i=1, \dots, n$ и ненулевой элемент a_{kl} порождают новое линейное отображение группы переменных $(x_1, \dots, x_{l-1}, y_k, x_{l+1}, \dots, x_n)$ в группу переменных $(y_1, \dots, y_{k-1}, x_l, y_{k+1}, \dots, y_n)$ по формулам

$$\begin{cases} x_l = \frac{1}{a_{kl}} y_k + \sum_{j=1, j \neq l}^n \frac{-a_{kj}}{a_{kl}} x_j, \\ y_i = \frac{a_{il}}{a_{kl}} y_k + \sum_{j=1, j \neq l}^n \left(a_{ij} - \frac{a_{il} \cdot a_{kj}}{a_{kl}} \right) x_j, i \neq k. \end{cases}$$

Обозначая через b_{ij} элементы матрицы нового линейного отображения, получим формулы универсальной операции над матричными структурами

$$b_{ij} = \frac{1}{a_{kl}} \begin{cases} 1, & \text{если } i=k \text{ и } j=l, \\ a_{il}, & \text{если } j=l, \\ -a_{kj}, & \text{если } i=k, \\ a_{ij} \cdot a_{kl} - a_{il} \cdot a_{kj}, & \text{если } i \neq k \text{ и } j \neq l. \end{cases}$$

Ряд свойств этой операции рассмотрен в работе [11]. Ниже будут представлены многоцелевые алгоритмы, в основе которых лежит эта универсальная операция.

3 Тетрада как форма представления вещественного числа

Вид формулы универсальной операции над матричными структурами указывает на целесообразный путь её программной реализации. Как известно, операция деления является наиболее затратной при выполнении машинной арифметики [12]. Значит, необходима такая форма представления числа, которая исключала бы деление при получении промежуточных результатов. Этого можно достичь за счёт введения упорядоченной пары (α, β) ([13], с. 119),

которая представляет отношение вещественных чисел $\frac{\alpha}{\beta}$. С другой

стороны, операция вычитания близких чисел приводит к значительному повышению относительной погрешности вычислений. Поэтому вычитание также следует исключить при получении промежуточных результатов за счёт введения дополнительной упорядоченной пары $\langle \alpha, \beta \rangle$ ([13], с. 120), обозначающей разность $\alpha - \beta$. В любой реализации множество компьютерных чисел конечно, арифметические операции коммутативны, но не ассоциативны и не дистрибутивны. Этот недостаток устраняется за счёт введения класса *приближённых чисел* как совокупности пар вида $\{a, \delta_a\}$, где a есть компьютерное число, а δ_a - переменная с известной границей изменения. Каждая такая пара представляет вещественное число вида $A = a(1 + \delta_a)$. Следует помнить, что результатом вычислений на компьютере является не компьютерное число, а приближённое число с указанной границей изменения переменной относительной погрешности δ_a . Для положительных чисел $A = a(1 + \delta_a)$ и $B = b(1 + \delta_b)$, их суммы $C = A + B$ и произведения $D = A \cdot B$ имеем $C = a \oplus b(1 + \delta_c)$ и $D = a \otimes b(1 + \delta_d)$, где $a \oplus b$ и $a \otimes b$ есть компьютерные числа и можно указать вполне определённые

границы для переменных $\delta_c = \delta_a \frac{a}{a+b} + \delta_b \frac{b}{a+b} + \delta_+$ и $\delta_d = \delta_a + \delta_b + \delta_*$.

Всё это указывает на то, что при программной реализации универсальной операции следует ограничиться машинным представлением только положительных чисел. Структура

$(\langle \alpha_1, \beta_1 \rangle, \langle \alpha_2, \beta_2 \rangle)$, обозначающая число вида $\frac{\alpha_1 - \beta_1}{\alpha_2 - \beta_2}$, позволяет

при работе с компьютерными числами применять только две операции: сложение и умножение положительных чисел.

Воспользуемся встроенным векторным типом `float4` CUDA-расширения языка C [1] для программной реализации введённой тетрады представления чисел и арифметических операций над этими тетрадами. Ниже приводится функция операции вычитания. Аналогичные функции могут быть выписаны для операций сложения, умножения и деления тетрад.

```
__device__ float4 operator – (const float4 & a , const float4 & b );  
{ float4 c;  
  
  c.x=a.x·b.z+a.y·b.w+a.z·b.y+a.w·b.x;  
  c.y=a.y·b.z+a.x·b.w+a.z·b.x+a.w·b.y;  
  c.z=a.z·b.z+a.w·b.w;  
  c.w=a.w·b.z+a.z·b.w;  
  return c;  
}
```

4 Примеры многоцелевых алгоритмов

Здесь будет представлен класс алгоритмов, которые в ходе своего исполнения позволяют достичь нескольких целей, относящихся к определённой проблематике. В основе всех алгоритмов класса лежит универсальная операция над матричными структурами. Она применяется к агрегату из матриц, отвечающих решаемой проблеме, согласно правилу выбора *разрешающего элемента*. Представленные в докладе алгоритмы завершаются через конечное число шагов применения универсальной операции. В качестве примера рассмотрены симметричная задача линейного программирования, балансовая модель экономики, управление спектром дискретных процессов методом минимальной обратной связи, одномерное и двумерное сглаживание экспериментальных данных. Эти результаты служат прямым подтверждением тезиса проф. В.А. Вышинского, что основой вычислительных машин будущего поколения должна стать *универсальная алгоритмическая матрично-алгебраическая система* [14]. В новой машинной технологии обработка сложных структур данных сведётся к обработке матриц, при этом будет использоваться математический аппарат регулярного матричного представления для реализации отвечающей решаемой проблеме числовой алгебры. Аппа-

ратурная реализация принципиально нового языка высокого уровня позволит существенно упростить и удешевить программное обеспечение создаваемой машины будущего поколения. При этом, удастся заменить маломощную машинную арифметику машинной алгеброй и, тем самым, добиться устранения противоречия, возникшего между интеграцией вычислительных схем в аппаратуре вычислительных систем и отсутствием адекватной интеграции в вычислительном технологическом процессе. Кроме того, откроется прямой путь повышения производительности вычислений не за счёт повышения тактовой частоты генератора и последующего преодоления «теплового барьера», а за счёт извлечения преимущества из сверхбольшой интеграции.

Программная реализация многоцелевых алгоритмов указанного класса представлена для гибридных вычислительных систем, поддерживающих технологии многопоточного параллельного программирования OpenCL и CUDA. Один из этих алгоритмов одновременно находит решение прямой и двойственной задачи линейного программирования, заданной в симметричной форме, а также области устойчивости решений этих задач. Другой алгоритм для балансовой модели материальных и финансовых ресурсов находит одновременно объёмы и цены производимой продукции, матрицу полных затрат и отвечает на вопрос о продуктивности балансовой модели. Третий алгоритм для матриц простой структуры одновременно находит подобную матрицу Фробениуса, преобразующую матрицу и обратную к ней, что позволяет строить одно-ранговые возмущения матриц простой структуры с заданными спектральными свойствами.

Литература

1. NVIDIA CUDA Programming Guide Version 4.0 // www.nvidia.com.
2. Годунов С.К. Решение систем линейных уравнений. – Новосибирск: Наука, 1980. – 177 с.
3. Байков В.Д., Смолов В.Б. Специализированные процессоры: Итерационные алгоритмы и структуры. – М.: Радио и связь, 1985. – 288 с.
4. Кострикин А.И. Введение в алгебру. – М.: Наука, 1977. – 496 с.
5. Скорняков Л.А. Элементы общей алгебры. – М.: Наука, 1983. – 272 с.
6. Демидов И.Т. Основания арифметики. – М.: Учпедгиз, 1963. – 160 с.
7. Саломеа А. Жемчужины теории формальных языков. – М.: Мир, 1986. – 159 с.
8. Кук Д., Бейз Г. Компьютерная математика. – М.: Наука, 1990. – 384 с.
9. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. Т. 1. Синтаксический анализ. – М.: Мир, 1978. – 616 с.

10. Исламов Г.Г. Универсальная операция над матричными структурами // Современные проблемы вычислительной математики и математической физики: Международная конференция, Москва, МГУ имени М.В. Ломоносова, 16-18 июня 2009 г. Тезисы докладов. – М.: Издательский отдел факультета ВМК МГУ имени М.В. Ломоносова; Макс Пресс, 2009. – 396 с.
11. Исламов Г.Г., Коган Ю.В. Об одном алгоритме поиска базисного минора матрицы // Вестник удмуртского университета. – 2006. - № 1 : Математика. – С. 63 – 70.
12. Шауман А. М. Основы машинной арифметики. – Л.: Изд-во ЛГУ, 1979. – 312 с.
13. Арнольд И. В. Теоретическая арифметика. – М.: Учпедгиз, 1939. – 400 с.
14. Вышинский В.А. Решение фундаментальной проблемы современного развития вычислительной техники // Труды Первой международной конференции «Трёхмерная визуализация научной, технической и социальной реальности. Кластерные технологии моделирования, Т. 1. Ижевск, УдГУ (4-6 февраля 2009 г.), С. 27-36.