

МЕТОДЫ ОПТИМИЗАЦИИ ВЫПОЛНЕНИЯ ТЕНЗОРНЫХ ОПЕРАЦИЙ НА МНОГОЯДЕРНЫХ ПРОЦЕССОРАХ

Гареев Роман Альбертович

Сергеев Кирилл 17122

Актуальность темы исследования

Цели и задачи исследования

1. Разработать модель гипотетического процессора, которая позволяет сократить время выполнения матрично-векторных операций и их обобщений на замкнутые полукольца с элементами из множества вещественных чисел.

2. Разработать новые алгоритмы выполнения тензорных операций константной сложности относительно размерности тензоров, уменьшающие время выполнения таких операций.

3. Разработать программную система АОТО для автоматической оптимизации времени выполнения тензорных операций и их автоматического распараллеливания при компиляции программ для многоядерных процессоров общего назначения.

4. Провести вычислительные эксперименты, подтверждающие эффективность разработанной программной системы по сравнению с аналогами, использующими ручную настройку и автонастройку.

Результаты, полученные в работе

1. Разработана новая модель гипотетического процессора, которая позволяет сократить время выполнения матрично-векторных операций и их обобщений на замкнутые полукольца с элементами из множества вещественных чисел.

2. Разработаны новые алгоритмы выполнения тензорных операций константной сложности относительно размерности тензоров, уменьшающие время выполнения таких операций. С помощью моделирования выполнения представленных алгоритмов на гипотетическом процессоре выведены формулы, позволяющие получить значения параметров алгоритмов выполнения тензорных операций в зависимости от характеристик многоядерных процессоров общего назначения для архитектур x86-64, x86, risc64le, aarch64. Доказаны утверждения о существовании значений параметров, при которых отсутствует простой конвейера векторных инструкций гипотетического процессора для представленных алгоритмов при возможности мгновенной загрузки данных из памяти на векторные регистры.

3. Разработана программная система АОТО для автоматической оптимизации времени выполнения тензорных операций и их автоматического распараллеливания при компиляции программ для многоядерных процессоров общего назначения. Получена оценка производительности многопоточной программы.

Теоретическая и практическая значимость работы

Результаты, изложенные в диссертации, могут быть эффективно использованы при численном решении на многоядерных процессорах общего назначения задач математической физики, математической геофизики, механики, квантовой химии.

Спасибо за внимание!