

ОЦЕНКА ЭФФЕКТИВНОСТИ ПАРАЛЛЕЛЬНОГО АЛГОРИТМА СОВМЕСТНОГО ГЕОСТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Шестаков В. В.

Научный руководитель: Степанов Д. Ю., к.т.н., доцент
Национальный исследовательский Томский политехнический университет
e-mail: valeriy.shestakov@inbox.ru

Введение

Примером дискретных наблюдений, заданных большими объемами, являются разрезы и кубы данных 2D и 3D метода общей глубинной точки (МОГТ). Они проходят многочисленные этапы цифровой обработки, длительность которых, в зависимости от параметров наблюдений и сложности алгоритмов, варьируется от нескольких часов до нескольких дней. Сократить данные сроки можно путем оптимизации и распараллеливания процесса обработки.

В данной работе, осуществляется оценка эффективности распараллеливания алгоритма совместного геостатистического моделирования. В качестве критерия эффективности используется коэффициент ускорения, представляющий собой отношение времени работы реализаций последовательного и параллельного алгоритмов.

Алгоритм совместного геостатистического моделирования

Назначение исследуемого алгоритма – прогноз петрофизических свойств геологической среды (пористость, проницаемость, водонасыщенность и т. д.) методом совместной интерпретации данных наземной сейсморазведки и геофизических исследований скважин.

В основе алгоритма лежит уравнение Крайгинга, в соответствии с которым прогнозное значение петрофизического параметра, в заданной точке пространства может быть представлено в виде взвешенной суммы его известных значений [1]:

$$\hat{F}(x_0, y_0, t_0) = \sum_{i=1}^N w_i(x_0, y_0, t_0) f_i(t_0) \quad (1)$$

где $w_i(x_0, y_0, t_0)$ – весовой коэффициент i -й скважины; $f_i(t)$ – известные значения петрофизического параметра; $\hat{F}(x_0, y_0, t_0)$ – прогнозное значение петрофизического параметра; N – количество скважин.

Метод поиска численных значений весовых функций в выражении (1) сводится к решению системы линейных алгебраических уравнений (СЛАУ) вида:

$$C(t_0) \cdot W^0(x_0, y_0, t_0) = C^0(x_0, y_0, t_0) \quad (2)$$

где $C(t_0)$ – матрица значений ковариограмм сейсмического атрибута в точках скважин; $C^0(t_0)$ – матрица значений ковариограмм сейсмического атрибута в точке прогноза. При фиксированных

x_0, y_0 решение СЛАУ последовательно ищется для каждого t , принадлежащего $[t_0, t_{max}]$. Общая схема алгоритма совместного геостатистического моделирования приведена на рисунке 1.

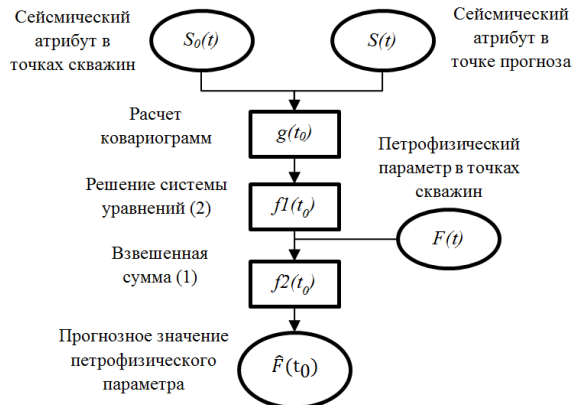


Рисунок 11. Общая схема последовательного алгоритма

Параллельный алгоритм совместного геостатистического моделирования

Идея параллельного алгоритма заключается в том, что при фиксированных x_0, y_0 , каждому отдельному потоку в соответствии ставится момент времени $t \in [t_0, t_{max}]$, для которого по аналогии с рисунком 1 осуществляется оценка петрофизического параметра.

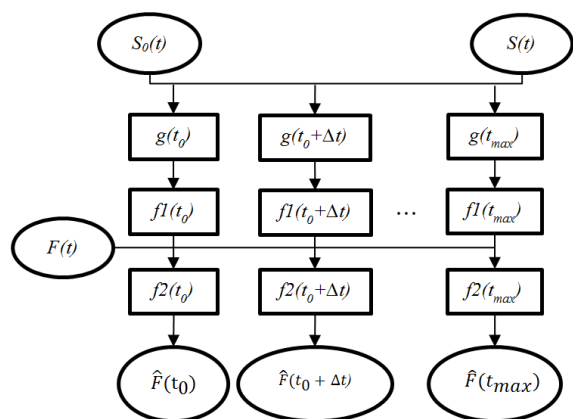


Рисунок 12. Общая схема параллельного алгоритма

Оценка эффективности параллельного алгоритма

Доля последовательных операций в разработанной программной реализации составляет примерно 15%. В неё входят операции считывания и подготовки первичных данных, а

также запись результата. Остальные 85% приходится на параллельный алгоритм, схема которого изображена на рисунке 2. Согласно второму закону Амдала [2] предположительное минимальное время работы параллельного алгоритма можно рассчитать в виде:

$$T_p = \alpha T_1 + \frac{(1-\alpha)T_1}{p} \quad (3)$$

где T_1 – время работы последовательной реализации; α – доля последовательных операций в параллельном алгоритме; p – количество параллельных процессов.

При обработке данных наземной сейсморазведки, с общим объемом 1.6 Гб, и данных шести скважин (размерность СЛАУ - 7), время работы последовательного алгоритма составляет 8.75 минуты. Предполагая, что параллельная реализация будет запущена на восьми-ядерном процессоре, в соответствии с формулой (3) минимальное время работы составит:

$$T_p = 0.15 \cdot 8.75 + \frac{(1-0.15) \cdot 8.75}{8} = 2.24 [\text{мин}]$$

Максимальное ускорение распараллеливания соответственно равно:

$$S = \frac{8.75}{2.24} = 3.90$$

Влияние размера СЛАУ на загрузку центрального процессора и время работы алгоритма

На практике время работы распараллеленного алгоритма составило 3.25 минуты, а загрузка центрального процессора – 79%. Очевидно, что данные показатели будут варьироваться в зависимости от размера СЛАУ. Из рисунка 3 видно, что загрузка ЦП уменьшается при решении систем с малым числом уравнений.

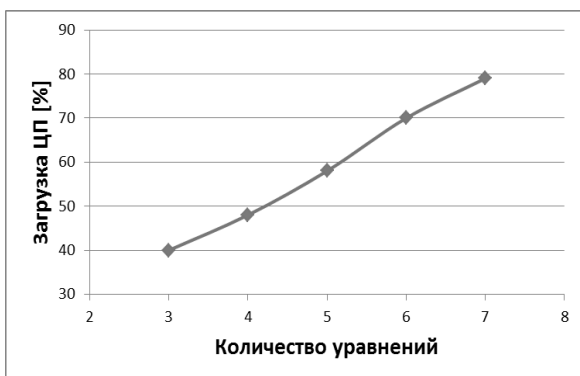


Рисунок 13. Зависимость загрузки ЦП от количества уравнений в системе

Из рисунка 4 можно судить, что с увеличением размера системы, коэффициент ускорения стремится к своему максимальному значению. Так, в случае системы с двенадцатью

уравнениями, средняя загрузка ЦП равна 89%, время работы последовательного алгоритма - 29.33 минуты, время работы параллельного алгоритма – 9.02 минуты, а коэффициент ускорения равен 3.25, что сопоставимо с теоретически максимальным ускорением, равным 3.90.

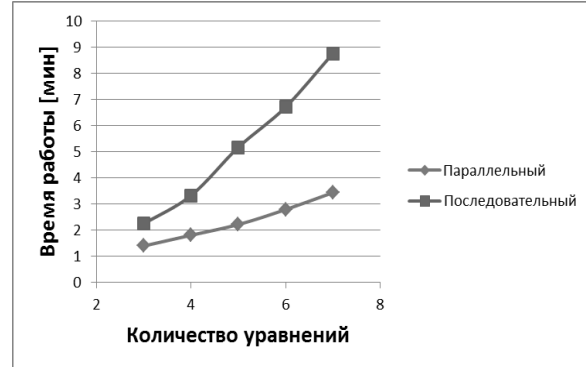


Рисунок 14. Зависимость времени работы алгоритма от количества уравнений

Заключение

Разработанный параллельный алгоритм оценки петрофизических параметров максимально эффективен при размерности СЛАУ более 10 (коэффициент ускорения в данном случае составляет 3.25 и более). В случае минимально допустимого размера СЛАУ, равного трем уравнениям, эффективность использования ресурсов ЦП резко падает (до 41%), однако, не смотря на это, эффект от распараллеливания положительный - коэффициент ускорения равен 1.61.

Литература

- Шестаков В.В., Степанов Д.Ю., Сысолятина Г.А. Построение трехмерных моделей параметров геологических сред // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции, Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 1 - С. 87-91
- Гергель В.П., Стронгин Р.Г. Основы параллельных вычислений для многопроцессорных вычислительных систем / Издательство Нижегородского государственного госуниверситета - Нижний Новгород, 2003 г., 184 с.