

# Параллельные алгоритмы решения систем линейных уравнений методом Монте-Карло

Тросиненко Анатолий Владимирович, гр. 522

Санкт-Петербургский государственный университет  
Математико-механический факультет  
Кафедра статистического моделирования

Научный руководитель: д.ф.-м.н., проф. Ермаков С. М.  
Рецензент: к. ф.-м. н., доц. Коробейников А. И.



Санкт-Петербург  
2014г.

Система линейных алгебраических уравнений

$$X = \mathcal{A}X + F, \quad \rho(\mathcal{A}) < 1 \quad (1)$$

решается на вычислительной системе с распределённой памятью.

Матрица  $\mathcal{A}$  не помещается в оперативной памяти одного узла и разделена между  $m$  узлами на блоки строк  $(A_{i1}, \dots, A_{im})$ , таким же образом разделены вектора:

$$\mathcal{A} = \begin{pmatrix} A_{11} & \dots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \dots & A_{mm} \end{pmatrix}, \quad X = \begin{pmatrix} X^{(1)} \\ \vdots \\ X^{(m)} \end{pmatrix}, \quad F = \begin{pmatrix} F^{(1)} \\ \vdots \\ F^{(m)} \end{pmatrix}.$$

**Требуется** исследовать способы нахождения решения уравнения (1) с использованием метода Монте-Карло на вычислительной системе с распределённой памятью с учётом времени сетевых обменов.

План доклада:

- ❶ Модификация метода простой итерации
  - скорость сходимости при неотрицательной матрице
  - зависимость сходимости от разбиения индексов
- ❷ Использование метода Монте-Карло
  - выбор допустимых погрешностей
- ❸ Асинхронный вариант обменов
  - свойства
  - сходимость к решению

Систему

$$X = \mathcal{A}X + F$$

можно решать методом простой итерации:

$$X_{k+1} = \mathcal{A}X_k + F.$$

$X_k$  сходится к решению при  $\rho(\mathcal{A}) < 1$  при любом  $X_0$ .

При  $\rho(|\mathcal{A}|) < 1$  для решения также можно использовать методы Монте-Карло на марковских цепях:

- оценка по поглощению
- оценка по столкновениям
- ...
- Sequential Monte Carlo [Halton, 94]

Уменьшение количества сетевых обменов:

- Выбрать начальное приближение  $X_0 = \text{col}(X_0^{(1)}, \dots, X_0^{(m)})$ .
- Внешние итерации по  $k$ :
  - На  $i$ -ом из  $m$  вычислителей:
  - Собрать части вектора  $X_k$  с остальных вычислителей.
  - Выбрать начальное приближение  $X_{k+1,0}^{(i)}$ .
  - Внутренние итерации по  $l$ :

$$X_{k+1,l+1}^{(i)} = A_{ii}X_{k+1,l}^{(i)} + \left[ \sum_{\substack{j=1,\dots,m \\ j \neq i}} A_{ij}X_k^{(j)} + F^{(i)} \right].$$

- Разослать значение  $X_{k+1,l_0}^{(i)}$  в качестве  $X_{k+1}^{(i)}$ .
- Взять  $X_{k_0}$  в качестве ответа на задачу.

$X_{k+1,l}^{(i)}$  при  $\rho(A_{ii}) < 1$ ,  $l \rightarrow \infty$  сходятся к решению системы уравнений внутренних итераций

$$X_{k+1}^{(i)} = A_{ii}X_{k+1}^{(i)} + \left[ \sum_{\substack{j=1, \dots, m \\ j \neq i}} A_{ij}X_k^{(j)} + F^{(i)} \right],$$

Если решение данной системы находится точно, то внешние итерации переписутся как

$$X_{k+1} = \mathcal{B}X_k + \mathcal{F},$$

где

$$\mathcal{B} = \begin{pmatrix} \mathbf{0} & \widetilde{A_{11}}A_{12} & \dots & \widetilde{A_{11}}A_{1m} \\ \widetilde{A_{22}}A_{21} & \mathbf{0} & \dots & \widetilde{A_{22}}A_{2m} \\ \dots & \dots & \dots & \dots \\ \widetilde{A_{mm}}A_{m1} & \widetilde{A_{mm}}A_{m2} & \dots & \mathbf{0} \end{pmatrix}, \quad \mathcal{F} = \begin{pmatrix} \widetilde{A_{11}}F^{(1)} \\ \widetilde{A_{22}}F^{(2)} \\ \vdots \\ \widetilde{A_{mm}}F^{(m)} \end{pmatrix},$$

$$\widetilde{A_{ii}} = (\mathbf{I} - A_{ii})^{-1}$$

$$X = \mathcal{A}X + F \quad (2)$$

$$X_{k+1} = \mathcal{B}X_k + \mathcal{F} \quad (3)$$

## Теорема

*Если  $\rho(\mathcal{A}) < 1$  и  $\rho(\mathcal{B}) < 1$ , то итерационный процесс (3) сходится к решению исходного уравнения (2).*

*Если элементы матрицы  $\mathcal{A}$  неотрицательны и  $\rho(A_{ii}) < 1$  при всех  $i$ , то*

- *если  $\rho(\mathcal{A}) < 1$ , то  $\rho(\mathcal{B}) \leq \rho(\mathcal{A}) < 1$ ,*
- *если  $\rho(\mathcal{A}) > 1$ , то  $\rho(\mathcal{B}) \geq \rho(\mathcal{A}) > 1$  или  $\rho(\mathcal{B}) = 0$ .*

## Пример:

Модифицированный метод может расходиться при сходимости исходного:

$$\mathcal{A} = \begin{pmatrix} 2a & -a \\ a & 0 \end{pmatrix} \text{ — матрица Холтона}$$

$$\rho(\mathcal{A}) = a, \quad \rho(\mathcal{B}) = \frac{a}{\sqrt{1-2a}} \rightarrow \infty \text{ при } a \rightarrow 1/2 - 0$$

## Пример:

На компьютере была подобрана матрица  $4 \times 4$ , такая что модифицированный процесс при  $m = 2$

- сходится при разбиении индексов  $\{1, 2\}, \{3, 4\}$
- расходится при разбиении индексов  $\{1, 3\}, \{2, 4\}$



На каждом внешнем шаге  $i$ -ый вычислитель находит несмещённую оценку решения системы уравнений внутренних итераций методом Монте-Карло.

Последовательность случайных векторов  $Y_k$ :

- Начальное приближение  $Y_0$ .
- Итерации по  $k$ :
  - На  $i$ -ом вычислителе из  $m$ :

$$Y_{k+1}^{(i)} = \mathcal{B}^{(i)} Y_k + \mathcal{F}^{(i)} + \mathcal{E}_k^{(i)}, \quad \mathbb{E} \|\mathcal{E}_k^{(i)}\| \leq \bar{\varepsilon}_{k,i}$$

- $Y_{k_0}$  берётся в качестве ответа.

Требуется сходимость ошибки к нулю:

- $\mathbb{E} \|Y_k - \bar{X}\| \rightarrow 0$ , где  $\bar{X}$  — точное решение  $X = \mathcal{A}X + F$ , или
- $\|Y_k - \mathcal{B}Y_k - \mathcal{F}\| \rightarrow 0$

Требуется подобрать оптимальные  $\bar{\varepsilon}_k = (\bar{\varepsilon}_{k,1}, \dots, \bar{\varepsilon}_{k,m})^\top$ .

Подходы к выбору  $\bar{\varepsilon}_k$ :

- минимизация выражения

$$T(k, \bar{\varepsilon}_0, \dots, \bar{\varepsilon}_{k_0-1}) = \alpha \times \left[ \frac{1}{\|\bar{\varepsilon}_0\|^2} + \dots + \frac{1}{\|\bar{\varepsilon}_{k_0-1}\|^2} \right] + k_0 t,$$

при условии достижения достаточной точности,

где  $\alpha$  зависит от дисперсии и выбранной схемы метода Монте-Карло,

$t$  — продолжительность обмена

- обеспечение убывания ошибки на каждой внешней итерации в  $\|\mathcal{B}\|(1 + \beta)$  раз,  $\beta > 0$ .

Для данных подходов были получены выражения для допустимых погрешностей.

**Проблема:** разделение фазы вычислений и сетевых обменов.

**Решение:** узлы должны периодически асинхронно рассылать обновлённые значения посчитанных приближений и оценки матожиданий модулей ошибок.

Для решения этой проблемы был рассмотрен аналог метода асинхронных итераций [Baudet, 78].

**Проблема:** компоненты решения находятся методом МК со случайной погрешностью, зависящей от продолжительности вычислений.

**Решение:** рассматривать выпуклую линейную комбинацию полученных оценок для приходивших приближений

$$Z_l^{(i)} = \mathcal{B}^{(i)} Y_l + \mathcal{F}^{(i)} + \mathcal{E}_l^{(i)}$$
$$\tilde{Z}_k^{(i)} = \beta_1 Z_1^{(i)} + \dots + \beta_k Z_k^{(i)}, \quad \sum \beta_l = 1, \quad \beta_l \geq 0,$$

при этом минимизировать оценку сверху величины  $E \|\tilde{Z}_k^{(i)} - \bar{X}^{(i)}\|_2$ .

При минимизации оказывается, что

## Теорема

- *Выборки метода МК для двух последовательных совпадающих приближений объединяются*
- *Приход уточнённого приближения в любой момент не ухудшает результат (при пренебрежимых накладных расходах на получение новых приближений и равенстве дисперсий)*

Также для асинхронного варианта было доказано достаточное условие сходимости.

## Полученные результаты:

- Исследован итерационный метод решения больших систем линейных уравнений на системах с распределённой памятью
  - Доказана теорема о сходимости для неотрицательной матрицы  $A$
  - Приведён пример зависимости сходимости от выбора разбиения матрицы между узлами
- Получены выражения для норм ошибок, обеспечивающих сходимость к решению при использовании метода МК в случае системы с распределённой памятью
- Сформулирован асинхронный метод
  - Доказаны некоторые естественные свойства
  - Доказано достаточное условие сходимости
- Написаны программы на R и C++ (с использованием MPI)