Метод Жордана решения линейной системы с выбором главного элемента по столбцу. Параллельная реализация

Мещеряков Вадим

30 декабря 2024 г.

1 Постановка задачи

Решаем линейную систему с матрицей A вида $Ax = b$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

2 Описание хранения

Представим матрицу A в виде:

$$\hat{A} = \left(\begin{array}{ccccc|c} A_{11}^{m\times m} & A_{12}^{m\times m} & \cdots & A_{1,k}^{m\times m} & A_{1,k+1}^{m\times l} & B_1^{m\times 1} \\ A_{21}^{m\times m} & A_{22}^{m\times m} & \cdots & A_{2,k}^{m\times m} & A_{2,k+1}^{m\times l} & B_2^{m\times 1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_{k,1}^{m\times m} & A_{k,2}^{m\times m} & \cdots & A_{k,k}^{m\times m} & A_{k,k+1}^{m\times l} & B_k^{m\times 1} \\ A_{k+1,1}^{l\times m} & A_{k+1,2}^{l\times m} & \cdots & A_{k+1,k}^{l\times m} & A_{k+1,k+1}^{l\times l} & B_{k+1}^{l\times 1} \end{array}\right)$$

Матрицу A храним в памяти в виде линейного массива следующим образом:

$$A = \{a_{11}, a_{12}, \dots, a_{1m}, a_{21}, a_{22} \dots, a_{2m}, \dots \dots, a_{nn}, b_1, \dots, b_n\}$$

Функции $get\_block()$ и $put\_block()$ остаются прежними.

3 Описание многопоточного алгоритма

3.1 Первый шаг

Предположим, что программа использует p потоков (потоки нумеруются начиная с 1). Блочных строк $k+1$, поэтому у каждого потока $d = \lceil \frac{k+1}{p} \rceil$, и возможно некоторые потоки будут иметь на одну строку меньше (таких потоков $W_1 = d \cdot p - (k+1)$). Значит потоков работающих ровно с d строками $(k+1) + p \cdot (1-d) = W_0$ штук. В данном случае точное количество строк у каждого потока не важно. Поток с номером i работает с блочными строками i, $i+p, \dots, i+p\cdot\tilde{d}$ (помним, что $\tilde{d}$ равно d (для потока $i \leqslant W_0$) или $d-1$ (для потока $i > W_0$)). Размер блочной строки не указывается так как только один поток с номером W_0 будет с ней работать, и работа с неквадратными блоками ведется аналогично тому, как это происходило и в последовательной версии алгоритма.

Так выглядит матрица для потока с номером $i \neq 1$. Прямоугольная матрица P_i содержит только «свои» строки потока с номером i.

$$P_i = \left(\begin{array}{ccccc|c} A_{i,1} & A_{i,2} & \cdots & A_{i,k} & A_{i,k+1} & B_i \\ A_{i+p,1} & A_{i+p,2} & \cdots & A_{i+p,k} & A_{i+p,k+1} & B_{i+p} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ A_{i+\tilde{d}-p,1} & A_{i+\tilde{d}-p,2} & \cdots & A_{i+\tilde{d}-p,k} & A_{i+\tilde{d}-p,k+1} & B_{i+\tilde{d}-p} \\ A_{i+\tilde{d},1} & A_{i+\tilde{d},2} & \cdots & A_{i+\tilde{d},k} & A_{i+\tilde{d},k+1} & B_{i+\tilde{d}} \end{array} \right)$$

В этой матрице находим главный элемент. Далее идет точка синхронизации и все потоки синхронизируют номер строки q с минимальной нормой из p главных элементов матриц P_i. Далее, если требуется, меняем полученную строку с номером q со строкой номер 1 (этот процесс происходит параллельно (например поток с номером a меняет блоки с номерами a, $a+p$, $a+2p$, $\ldots$, поэтому необходимо поставить точку синхронизации). Далее так же параллельно умножаем эту первую строку на обратную к главному элементу. В результате этих действий получим матрицу P_i:

$$A = \left(\begin{array}{ccccc|c} E_{1,1} & \hat{A}_{1,2} & \cdots & \hat{A}_{1,k} & \hat{A}_{1,k+1} & \hat{B}_i \\ A_{2,1} & A_{2,2} & \cdots & A_{2,k} & A_{2,k+1} & B_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ A_{n-1,1} & A_{n-1,2} & \cdots & A_{n-1,k} & A_{n-1,k+1} & B_{n-1} \\ A_{n,1} & A_{n,2} & \cdots & A_{n,k} & A_{n,k+1} & B_n \end{array} \right)$$

Поток $i = 1$ приводит матрицу к такому виду:

$$\hat{P}_i = \left(\begin{array}{ccccc|c} E_{i,1} & \hat{A}_{i,2} & \cdots & \hat{A}_{i,k} & \hat{A}_{i,k+1} & \hat{B}_i \\ 0 & \hat{A}_{i+p,2} & \cdots & \hat{A}_{i+p,k} & \hat{A}_{i+p,k+1} & \hat{B}_{i+p} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & \hat{A}_{i+\tilde{d}-p,2} & \cdots & \hat{A}_{i+\tilde{d}-p,k} & \hat{A}_{i+\tilde{d}-p,k+1} & \hat{B}_{i+\tilde{d}-p} \\ 0 & \hat{A}_{i+\tilde{d},2} & \cdots & \hat{A}_{i+\tilde{d},k} & \hat{A}_{i+\tilde{d},k+1} & \hat{B}_{i+\tilde{d}} \end{array} \right)$$

Остальные потоки при $i \neq 1$:

$$\hat{P}_i = \left(\begin{array}{ccccc|c} 0 & \hat{A}_{i,2} & \cdots & \hat{A}_{i,k} & \hat{A}_{i,k+1} & \hat{B}_i \\ 0 & \hat{A}_{i+p,2} & \cdots & \hat{A}_{i+p,k} & \hat{A}_{i+p,k+1} & \hat{B}_{i+p} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & \hat{A}_{i+\tilde{d}-p,2} & \cdots & \hat{A}_{i+\tilde{d}-p,k} & \hat{A}_{i+\tilde{d}-p,k+1} & \hat{B}_{i+\tilde{d}-p} \\ 0 & \hat{A}_{i+\tilde{d},2} & \cdots & \hat{A}_{i+\tilde{d},k} & \hat{A}_{i+\tilde{d},k+1} & \hat{B}_{i+\tilde{d}} \end{array} \right)$$

Формулы используются точно такие же как в последовательной реализации. После этого шага, исходная блочная матрица будет иметь вид:

$$A = \left(\begin{array}{ccccc|c} E_{1,1} & A_{1,2} & \cdots & A_{1,k} & A_{1,k+1} & B_1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & A_{p,2} & \cdots & A_{p,k} & A_{p,k+1} & B_p \\ 0 & A_{p+1,2} & \cdots & A_{p+1,k} & A_{p+1,k+1} & B_{p+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & A_{2p,2} & \cdots & A_{2p,k} & A_{2p,k+1} & B_{2p} \\ 0 & A_{2p+1,2} & \cdots & A_{2p+1,k} & A_{2p+1,k+1} & B_{2p+1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & A_{k,2} & \cdots & A_{k,k} & A_{k,k+1} & B_k \\ 0 & A_{k+1,2} & \cdots & A_{k+1,k} & A_{k+1,k+1} & B_{k+1} \end{array} \right)$$

Матрица $E_{1,1}$ — это единичные матрицы. Необходимо поставить точку синхронизации, чтобы потоки приостановили работу пока не получится матрица A.
Необходимо поставить точку синхронизации.
Действуем аналогично (расставляя после каждого вычитания точку синхронизации), пока не останется только первая строка с ненулевым первым блоком.
После этих преобразований, матрица имеет вид:

$$A = \left(\begin{array}{ccccc|c} E_{1,1} & A_{1,2} & \cdots & A_{1,k} & A_{1,k+1} & B_1 \\ 0 & A_{2,2} & \cdots & A_{2,k} & A_{2,k+1} & B_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & A_{k+1,2} & \cdots & A_{k+1,k} & A_{k+1,k+1} & B_{k+1} \end{array}\right)$$

3.2 Шаг номер s

$$A = \left(\begin{array}{cccccc|c} E_{1,1} & \cdots & 0 & A_{1,s} & \cdots & A_{1,k+1} & B_1 \\ 0 & \ddots & \vdots & \vdots & & \vdots & \\ \vdots & \ddots & E_{s-1,s-1} & A_{s-1,s} & \cdots & A_{s-1,k+1} & B_{s-1} \\ 0 & & 0 & A_{s,s} & \cdots & A_{s,k+1} & B_s \\ 0 & & 0 & A_{s+1,s} & \cdots & A_{s+1,k+1} & B_{s+1} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & A_{k,s} & \cdots & A_{k,k+1} & B_k \\ 0 & \cdots & 0 & A_{k+1,s} & \cdots & A_{k+1,k+1} & B_{k+1} \end{array}\right)$$

Строка $s+i$ является «своей» для потока $j = (s+i) \mod p$, $i \in \{0, \ldots, p-1\}$. Поток j умножает всю строку $s+i$ на $(A_{s+i,s})^{-1}$. Далее поток j с помощью строки $s+i$ обнуляет блоки $A_{j,s}$, $A_{j+p,s}, \ldots$. Поток доходит до точки синхронизации, и когда все потоки обнулили нижележащие строки матрица имеет вид:

$$A = \left(\begin{array}{ccccccc|c} E_{1,1} & \cdots & 0 & 0 & A_{1,s+1} & \cdots & A_{1,k+1} & B_1 \\ 0 & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \ddots & E_{s-1,s-1} & 0 & A_{s-1,s+1} & \cdots & A_{s-1,k+1} & B_{s-1} \\ 0 & & 0 & E_{s,s} & A_{s,s+1} & \cdots & A_{s,k+1} & B_s \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & A_{s+p,s+1} & \cdots & A_{s+p,s+1} & B_{s+p} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & A_{k+1,s+1} & \cdots & A_{k+1,k+1} & B_{k+1} \end{array}\right)$$

На последних шагах будет меньше W_0 потоков с ненулевыми строками, поэтому только несколько последних потоков будут работать и обнулять все вышележащие строки.

4 Оценка сложности для функции потока

4.1 Сложность на шаге s

Для нахождения главного элемента, $\frac{k-s}{p}$ обращений матрицы: $2m^3 + O(m^2)$.
Умножение обратной матрицы на строку: $2\frac{k-s}{p}m^3 + O(m^2)$.
Умножений строки на первый блок другой строки: $(d-1) \cdot (2\frac{k-s}{p}m^3 + O(m^2))$.
Вычитание строки из другой строки: $(k-s)m^2 + lm$.

$d = \frac{k}{p}$.

4.2 Общая сложность

$$\sum_{s=1}^{k}\left(2\frac{k-s}{p}m^3+2\frac{(k-s)k}{p}m^3+\frac{k(k-s)}{p}m^2+O(mn)\right)=\frac{k(k-1)}{p}m^3+\frac{(k-1)k^2}{p}m^3+$$
$$+\frac{k^2(k-1)}{2p}m^2+O(n^2)=\frac{nm^2}{p}+\frac{n^3}{p}+\frac{n^2(k-1)}{2p}+O(n^2)=\frac{n^3}{p}+\frac{nm^2}{p}+O(n^2).$$

4.3 Точки синхронизации

Список точек синхронизации для каждой итерации, для каждого потока.

1. В потоках ищем главный элемент. Далее стоит точка синхронизации.

2. Две строки меняются местами (если это необходимо). Блоки двух строк распределяются по потокам и перезаписываются (меняются местами) независимо в каждом потоке. Далее стоит точка синхронизации.

3. Потоки умножают блоки главной строки на обратную к главному элементу. Далее стоит точка синхронизации.

4. Потоки считывают блоки из строки с главным элементом и вычитают ее из «своих» строк. После этого ставим последнюю точку синхронизации.

4.4 Проверка оценки

$S(n,m,1)=n^3+nm^2+O(n^2)$
$S(n,m,p)=\frac{1}{p}n^3+\frac{1}{p}nm^2+O(n^2)$
$S(n,1,1)=n^3+O(n^2)$
$S(n,n,1)=2n^3+O(n^2)$
$S(n,1,p)=\frac{1}{p}n^3+O(n^2)$
$S(n,n,p)=\frac{2}{p}n^3+O(n^2)$