

Нижегородский государственный университет им. Н.И.Лобачевского

Факультет Вычислительной математики и кибернетики

Параллельные численные методы

Базовые итерационные методы. Чебышевское ускорение

При поддержке компании Intel

Баркалов К.А.,
Кафедра математического обеспечения ЭВМ

Содержание

- Постановка задачи
- Свойства итерационных методов
- Метод простой итерации
  - Последовательный алгоритм и его свойства
  - Способы распараллеливания
- Методы Якоби и Зейделя
- Метод верхней релаксации
  - Последовательный алгоритм и его свойства
  - Параллельный алгоритм
- Чебышевское ускорение итерационных методов
- Результаты экспериментов



Постановка задачи

- Рассмотрим систему из n линейных алгебраических уравнений вида

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 &&+ \ldots + a_{1n}x_n &&= b_1 \\ a_{21}x_1 &+ a_{22}x_2 &&+ \ldots + a_{2n}x_n &&= b_2 \\ &\ldots \\ a_{n1}x_1 &+ a_{n2}x_2 &&+ \ldots + a_{nn}x_n &&= b_n \end{aligned}$$

- В матричном виде система может быть представлена как

$$Ax=b$$

- $A=(a_{ij})$ есть вещественная матрица размера $n \times n$; b и x – вектора из n элементов; точное решение системы обозначим x^*.
- *Итерационный метод* генерирует последовательность векторов $x^{(s)} \in R^m$, $s=0,1,2,\ldots$, где $x^{(s)}$ – приближенное решение системы.



Свойства итерационных методов

- Итерационный метод называется сходящимся, если

$$\forall\ x^{(0)} \in R^m \quad \lim_{s\to\infty} \left\| x^{(s)} - x^* \right\| = 0$$

- Критерии остановки итерационных методов: остановка по точности и остановку по числу итераций.
  - Стоп, если $\|x^{(s)} - x^{(s-1)}\| \le \varepsilon_1$. При этом $\varepsilon_2 = \|x^{(s)} - x^{(s-1)}\|$ – достигнутая точность метода.
  - Стоп, если $\|r^{(s)}\| \le \varepsilon_1$. При этом $\varepsilon_2 = \|r^{(s)}\|$ – достигнутая точность метода.
  - Стоп, если $s = N$. При этом $x^{(N)}$ трактуется как полученное решение. Максимальное число итераций N задается.
- Ниже будем предполагать, что матрица A – симметричная положительно определенная матрица.



Метод простой итерации (МПИ)

- Решаем систему $Ax=b$
- Преобразуем ее к виду $x=Gx+c$

  Преобразование можно сделать разными способами

  Пусть $A=M-N$, тогда $(M-N)x=b$, $Mx = Nx+b$

  $$x=M^{-1}Nx+M^{-1}b, \qquad \text{т.е. } G=M^{-1}N,\ c=M^{-1}b.$$

- Итерационный процесс

  $$x^{(s+1)} = Gx^{(s)} + c$$

- Сходимость
  - $\|G\|<1$ (необходимо); $\rho(G) < 1$ (достаточно).
  - $\left\|x^* - x^{(s)}\right\| \le \|G\|^s \left\|x^* - x^{(0)}\right\|$



Метод простой итерации (МПИ)

- Частный случай метода $\frac{x^{(s+1)} - x^{(s)}}{\tau} + Ax^{(s)} = b$

  где $\tau \neq 0$ – параметр метода.

- Вычисление следующего приближения

$$x^{(s+1)} = -\tau(Ax^{(s)} - b) + x^{(s)} = -\tau \cdot r^{(s)} + x^{(s)}$$

  где $r^{(s)}$ – невязка s-го приближения к решению.

- Покомпонентная форма записи метода

$$x_i^{(s+1)} = -\tau\left(\sum_{j=1}^{n} a_{ij} x_j^{(s)} - b_i\right) + x_i^{(s)}$$

- Оценка трудоемкости L итераций метода

$$T_1 = L(2n^2 + 2n).$$

МПИ – сходимость

- Если матрица A симметрична и положительно определена и $\tau \in (0, \lambda_{\max})$, метод сходится к точному решению системы с любого начального приближения.
- Оптимальное значением параметра $$\tau_{opt} = \frac{2}{\lambda_{\min} + \lambda_{\max}}$$
- Для МПИ с оптимальным параметром справедлива оценка

$$\left\| z^{(s+1)} \right\|_2 \le \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| z^{(0)} \right\|_2$$

где μ_A – число обусловленности матрицы A,

$z^{(s)} = x^{(s)} - x^*$ – погрешность очередного приближения,

$\mu_A = \lambda_{max} / \lambda_{min}$ – спектральное число обусловленности.



МПИ – параллельный алгоритм

- Выполнение итераций метода осуществляется последовательно
- Распараллелим вычисления, реализуемые в ходе выполнения одной итерации:
  - Основные вычисления, выполняемые в соответствии с методом, состоят в умножении матрицы A на вектор $x^{(s)}$,
  - Дополнительные вычисления (умножение на скаляр и сложение векторов) имеют меньший порядок сложности.
- Можно применить известные алгоритмы параллельного умножения матрицы на вектор



МПИ – параллельный алгоритм

- Оценка трудоемкости параллельной операции $Ax^{(s)}$ при использовании схемы ленточного горизонтального разделения матрицы A составляет

$$2n^2/p+\delta$$

n – длина вектора, p – число потоков, δ – накладные расходы

- Вычисления, имеющие меньший порядок сложности, выполняются в однопоточном режиме
- Общая оценка трудоемкости параллельного МПИ

$$T_p = L\left(\frac{2n^2}{p}+2n+\delta\right)$$

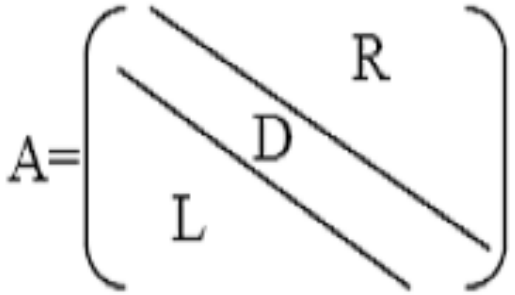
где L – число итераций метода.

Методы Якоби и Зейделя

- Вернемся к решению системы $Ax=b$

  с симметричной положительно определенной матрицей A.
- Представим матрицу в виде $A=L+D+R$,

  D, L, R – диагональная,
  строго нижняя треугольная
  строго верхняя треугольная
  части матрицы A.



- Метод Якоби

$$Dx^{(s+1)} = (-L-R)x^{(s)} + b$$

- Метод Зейделя

$$(D+L)x^{(s+1)} = (-R)x^{(s)} + b$$

Методы Якоби и Зейделя – сходимость

- Запишем методы в покомпонентной форме
  - Метод Якоби $$x_i^{(s+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(s)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(s)} \right) \Big/ a_{ii}$$
  - Метод Зейделя $$x_i^{(s+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(s+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(s)} \right) \Big/ a_{ii}$$
- Сходимость методов
  - Якоби: A>0, строгое диагональное преобладание
  - Зейделя: A>0
- Матрицы перехода методов
  - Якоби: $G_{Jac} = -D^{-1}(L + R) = D^{-1}A - E$
  - Зейделя: $G_{GS} = -(D+L)^{-1} R = (D+L)^{-1} A - E$



Метод верхней релаксации (SOR)

- Метод верхней релаксации (МВР) записывается в виде

$$\frac{(D+\omega L)(x^{(s+1)}-x^{(s)})}{\omega}+Ax^{(s)}=b$$

где ω – параметр метода.

- Сходимость: $\omega\in(0,2)$ (необходимо), если $A>0$ – то и достаточно
- При численном решении задач матфизики выбирают

$$\omega_{opt}\approx 2-O(h)$$

- Требуемое число итераций при $\omega=\omega_{opt}$: $O(h^{-1})$
  при $\omega=1$ (МВР совпадает с методом Зейделя): $O(h^{-2})$
- Более точная оценка

$$\omega_{opt}=\frac{2}{1+\sqrt{1-\rho^2\left(D^{-1}(R+L)\right)}}$$



SOR – алгоритм

- С учетом $A–L=R+D$, преобразуем в удобный вид

$$Dx^{(s+1)} = -\omega L x^{(s+1)} + (1-\omega)Dx^{(s)} - \omega R x^{(s)} + \omega b$$

- Компоненты нового приближения вычисляются как

$$a_{ii}x_i^{(s+1)} = -\omega\sum_{j=1}^{i-1} a_{ij}x_j^{(s+1)} + (1-\omega)a_{ii}x_i^{(s)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(s)} + \omega b_i$$

- Матрица перехода $G_{SOR} = (D+\omega L)^{-1}((1-\omega)D - \omega R)$
  - несимметричная!
- Общая трудоемкость одной итерации

$$t_1 = 2n^2 + n$$

- Выполнение L итераций метода

$$T_1 = L(2n^2 + n).$$

SSOR – симметричный метод

- Шаг SSOR состоит из:

1. Шаг SOR, вычисляем компоненты $x^{(s+1/2)}$ прямом порядке;
2. Шаг SOR, вычисляем компоненты $x^{(s+1)}$ обратном порядке.

- Шаг SSOR в матричной форме

1. $(D+\omega L)x^{(s+1/2)} = (1-\omega)Dx^{(s)} - \omega R x^{(s)} + \omega b$
2. $(D+\omega U)x^{(s+1)} = (1-\omega)Dx^{(s+1/2)} - \omega L x^{(s+1/2)} + \omega b$

- Матрица перехода

$$G_{SSOR} = (D+\omega U)^{-1}((1-\omega)D-\omega L)(D+\omega L)^{-1}((1-\omega)D-\omega R)$$

  - обычно больше итераций, чем SOR с ω_{opt}
  - G_{SSOR} – симметричная, используется в чебышевском ускорении.



Чебышевское ускорение

- Пусть найдены приближения $x^{(0)}, x^{(1)}, ..., x^{(m)}$
- Найдем $y^{(m)} = \sum_{i=0}^{m} \alpha_i x^{(i)}$, которое лучше, чем $x^{(m)}$
- Запишем погрешность $y^{(m)}$

$$y^{(m)} - x^* = \sum_{i=0}^{m} \alpha_i x^{(i)} - x^* = \sum_{i=0}^{m} \alpha_i (x^{(i)} - x^*) = \sum_{i=0}^{m} \alpha_i G^i (x^{(0)} - x^*) = p_m(G)(x^{(0)} - x^*),$$

здесь $p_m(G) = \sum_{i=0}^{m} \alpha_i G^i$ – полином от матрицы G, $p_m(1) = \sum_{i=0}^{m} \alpha_i = 1$

- $\rho(p_m(G)) \to \min$ – минимизируем спектральный радиус.
- $p_m(G)$ можно получить, используя полиномы Чебышева $T_m(x)$



Чебышевское ускорение

- $p_m(G) = \mu_m T_m(G/\rho)$, где $\mu_m \equiv 1/T_m(1/\rho)$, $T_m(x)$ – полином Чебышева, ρ – спектральный радиус матрицы G.
- Полиномы Чебышева

  $$T_0(x) = 1 \quad T_1(x) = x \quad T_m(x) = 2xT_{m-1}(x) - T_{m-2}(x)$$

- Трехчленное соотношение позволяет использовать только три вектора $y^{(m)}$, $y^{(m-1)}$, $y^{(m-2)}$, а не все вектора $x^{(m)}$, $0 \le i \le m$.
- Можно вывести соотношения

  $$y^{(m)} = \frac{2\mu_m}{\mu_{m-1}} \frac{G}{\rho} y^{(m-1)} - \frac{\mu_m}{\mu_{m-2}} y^{(m-2)} + \frac{2\mu_m}{\rho\mu_{m-1}} c \qquad \mu_m = \left(\frac{2}{\rho\mu_{m-1}} - \frac{1}{\mu_{m-2}} \right)^{-1}$$

- Требования к G: $\lambda_i \in [-\rho, \rho]$
  - SOR не подходит, но можно использовать SSOR



Чебышевское ускорение

- Итак, чебышевское ускорение метода $x^{(s+1)} = Gx^{(s)} + c$ состоит в:
  - положить $\mu_0 = 1,\ \mu_1 = \rho, y^{(0)} = x^{(0)}, y^{(1)} = Gx^{(0)} + c.$
  - Для m=2, 3, … вычислить

$$\mu_m = \left(\frac{2}{\rho \mu_{m-1}} - \frac{1}{\mu_{m-2}} \right)^{-1} \qquad y^{(m)} = \frac{2\mu_m}{\rho \mu_{m-1}} \left(G y^{(m-1)} + c \right) - \frac{\mu_m}{\mu_{m-2}} y^{(m-2)}$$

- Не надо вычислять вычислять G и c явно, итерация проводится в два этапа:

1) $\bar{y} = G y^{(m-1)} + c$ 2) $y^{(m)} = \frac{2\mu_m}{\rho \mu_{m-1}} \bar{y} - \frac{\mu_m}{\mu_{m-2}} y^{(m-2)}$



Результаты – разреженные матрицы

- Матрицы из коллекции университета Флориды
  http://www.cise.ufl.edu/research/sparse/matrices/
- Характеристики используемых матриц

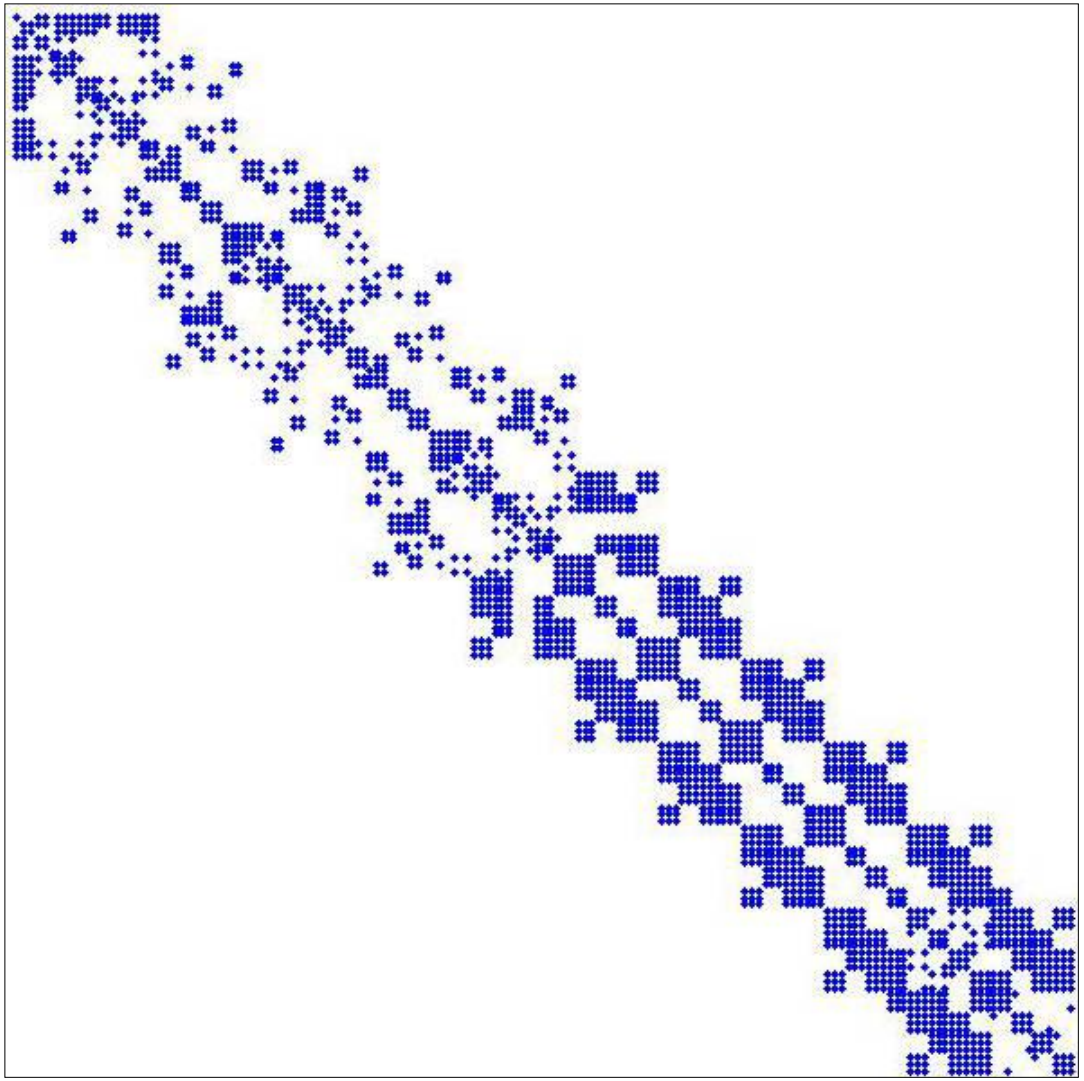
Название	n	nz	μ_A
mesh1em6	48	306	6.1
bcsstk04	132	3648	$2.3*10^6$
bcsstk05	153	2423	$1.4*10^4$
bcsstk09	1083	18437	$9.5*10^3$
chem97ZtZ	2541	7361	$2.5*10^2$

- Все матрицы – симметричные положительно определенные.



Результаты – разреженные матрицы

- Портрет матрицы bcsstk05
- Точное решение $x^* = \{x_i = 1, 1 \le i \le n\}$
- Правая часть $b = Ax^*$
- Система уравнений $Ax=b$



Результаты – разреженные матрицы

- Оптимальные ω и ρ – для матрицы общего вида аналитически вычислить сложно, определялись экспериментально.
- Точность метода $\varepsilon=10^{-6}$.
- В таблице указано число итераций s.

Название задачи	SOR		SOR-Cheb		
	ω	s	ω	ρ	s
mesh1em6	1.9	146	1.9	0.9	35
bcsstk04	1.9	341	1.09	0.99	229
bcsstk05	1.87	986	1.0	0.998	251
bcsstk09	1.95	885	1.11	0.999	537
chem97ZtZ	1.9	144	1.9	0.9	125



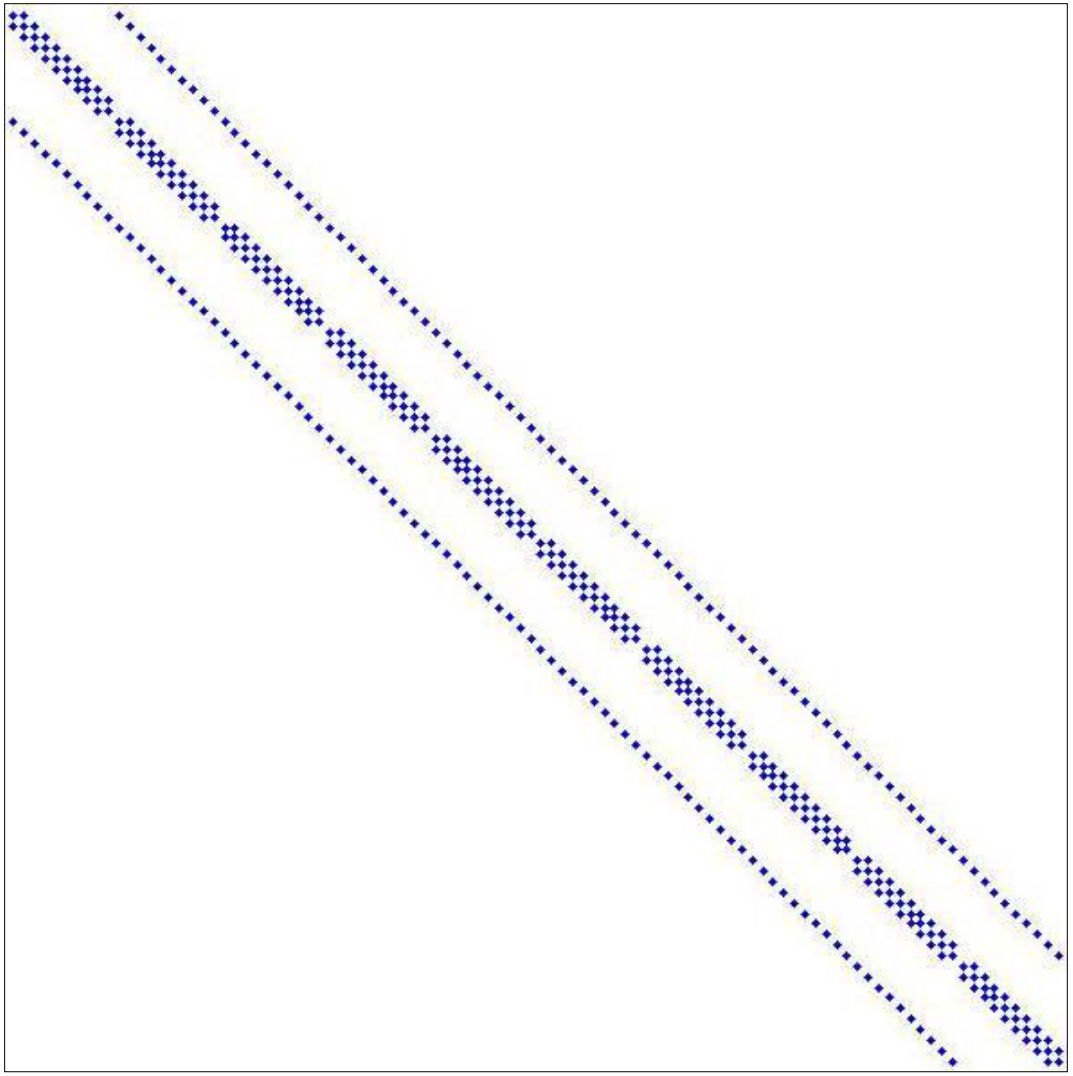
Результаты – уравнение Пуассона

- СЛАУ возникает при дискретизациии ДУЧП
- можно сформировать тестовую задачу с заранее известным точным решением
- для данного типа задач известно $\omega_{opt} = \frac{2}{1 + 2\sin(\pi h/2)}$
- для данного типа задач известно $\rho_{opt} = 1 - \frac{\pi h}{2}$
- матрица СЛАУ имеет пятидиагональный портрет



Результаты – уравнение Пуассона

- Портрет СЛАУ для уравнения Пуассона



Результаты – уравнение Пуассона

- Параметры метода: ρ=0.99, ε=10^{-6}.

n	nz/n	ω	s		
			SOR	SSOR	SSOR-Cheb
10000	$4{,}9\cdot10^{-6}$	1.9397	286	342	53
22500	$9{,}8\cdot10^{-6}$	1.9592	428	512	65
40000	$3{,}1\cdot10^{-7}$	1.9692	569	682	72
62500	$1{,}2\cdot10^{-7}$	1.9753	711	852	123
90000	$6{,}1\cdot10^{-8}$	1.9793	853	1022	91
122500	$3{,}3\cdot10^{-8}$	1.9823	995	1192	85
160000	$1{,}9\cdot10^{-8}$	1.9845	1137	1362	97
202500	$1{,}2\cdot10^{-8}$	1.9862	1278	1532	143
250000	$7{,}9\cdot10^{-9}$	1.9875	1420	1702	276



SOR – параллельный алгоритм

- Выполнение итераций метода осуществляется последовательно
- Вычисление компонент очередного приближения также осуществляется последовательно
- Распараллелить можно вычисление отдельных компонент очередного приближения
  - Основные вычисления состоят подсчете $\sum_{j=1}^{i-1} a_{ij} x_j^{(s+1)}$ и $\sum_{j=i+1}^{n} a_{ij} x_j^{(s)}$
- Используем для подсчета сумм известные алгоритмы параллельного суммирования (напр., *каскадная схема*).



SOR – параллельный алгоритм

- Оценка трудоемкости параллельной суммирования составляет

$$2n/p+\delta$$

n – длина суммы, p – число потоков, δ – накладные расходы

- Оценка трудоемкости одной итерации составит

$$t_p = n\left(2n/p+\delta\right)+n$$

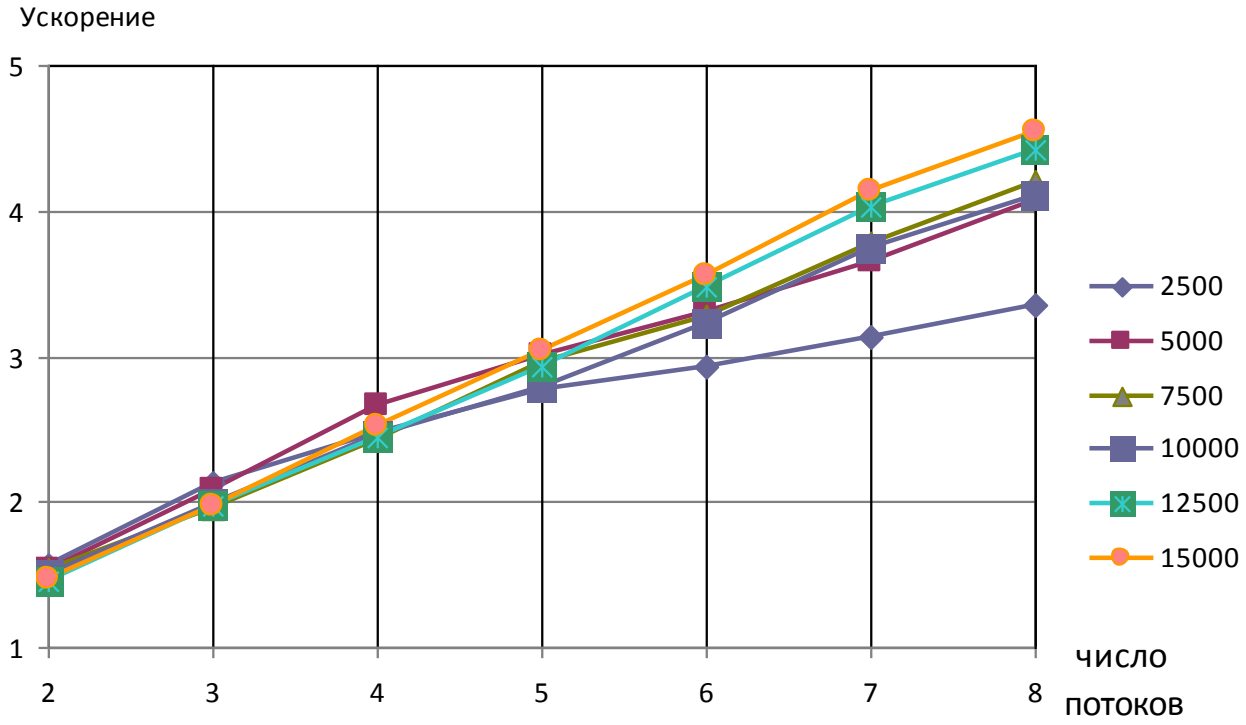
- Общая оценка трудоемкости параллельного МВР

$$T_p = L\left(2n^2/p+n+\delta n\right)$$



Результаты – SOR, плотные матрицы

- Ускорение по отношению к однопоточной версии



Заключение

- На лекции рассмотрено:
  - Понятие итерационных методов
  - Метод простой итерации
    - Последовательный алгоритм и его свойства
    - Способы распараллеливания
  - Методы Якоби и Зейделя
  - Метод верхней релаксации
    - Последовательный алгоритм и его свойства
    - Параллельный алгоритм
  - Чебышевское ускорение итерационных методов
  - Результаты экспериментов



Литература

1. Самарский А.А., Гулин А.В. Численные методы. – М.: Наука, 1989.
2. Белов С.А., Золотых Н.Ю. Численные методы линейной алгебры. – Н.Новгород, Изд-во ННГУ, 2005.
3. Голуб Дж., Ван Лоун Ч. Матричные вычисления. – М.: Мир, 1999.
4. J. Dongarra et al. Templates for the solution of linear systems: building blocks for iterative methods. SIAM, 1994.
5. Y. Saad. Iterative Methods for Sparse Linear Systems. SIAM, 2003.



Ресурсы сети Интернет

5. Intel Math Kernel Library Reference Manual.
[http://software.intel.com/sites/products/documentation/hpc/mkl/mklman.pdf].



Авторский коллектив

- Баркалов Константин Александрович,
  к.ф.-м.н., доцент кафедры
  математического обеспечения ЭВМ факультета ВМК ННГУ.
  barkalov@fup.unn.ru

- Коды учебных программ разработаны Роговой Дарьей

