Курс “Методы решения СЛАУ”

ФОНД ЦЕЛЕВОГО КАПИТАЛА МФТИ

Кузнецов А.А.
Петров Д.А.

Методы решения систем линейных алгебраических уравнений

Разреженные матрицы. Простейшие итерационные солверы

Кузнецов Александр
kuznetsov.aa@mipt.ru

Разреженные матрицы

Разреженные матрицы - матрицы, с большим количеством нулей

$$\begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & a_{n-1\,n-2} & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & 0 & 0 & a_{n\,n-1} & a_{nn} \end{pmatrix}$$

Трехдиагональная матрица - пример разреженной

Применимость:
Одномерные параболические уравнения
Одномерные краевые задачи

Разреженные матрицы

A	$\frac{1}{h^2}$	0	0	$\frac{1}{k^2}$	0	0	0				
$\frac{1}{h^2}$	A	$\frac{1}{h^2}$	0	0	$\frac{1}{k^2}$	0	0				
0	$\frac{1}{h^2}$	A	$\frac{1}{h^2}$	0	0	$\frac{1}{k^2}$	0				
0	0	$\frac{1}{h^2}$	A	0	0	0	$\frac{1}{k^2}$				
$\frac{1}{k^2}$				A	$\frac{1}{h^2}$			$\frac{1}{k^2}$			
	$\frac{1}{k^2}$			$\frac{1}{h^2}$	A	$\frac{1}{h^2}$			$\frac{1}{k^2}$		
		$\frac{1}{k^2}$			$\frac{1}{h^2}$	A	$\frac{1}{h^2}$			$\frac{1}{k^2}$	
			$\frac{1}{k^2}$			$\frac{1}{h^2}$	A				$\frac{1}{k^2}$
				$\frac{1}{k^2}$				A	$\frac{1}{h^2}$		
					$\frac{1}{k^2}$			$\frac{1}{h^2}$	A	$\frac{1}{h^2}$	
						$\frac{1}{k^2}$			$\frac{1}{h^2}$	A	$\frac{1}{h^2}$
							$\frac{1}{k^2}$			$\frac{1}{h^2}$	A

Пятидиагональная матрица - пример разреженной

Применимость: Двумерные параболические уравнения

диффузия, теплопроводность, сильно вязкие течения жидкости

Разреженные матрицы

Идеология разреженных матриц:

Зачем хранить множество нулей?

Зачем умножать что либо на нуль?

Разреженные матрицы

Пример

Всего элементов в матрице n^2, ненулевых элементов $2n - 1$

При $n = 10^5$ (современные задачи) занимает в памяти:

Плотная матрица $\sim 10^{10} * 8$ *байт* ~ ***80 Гб***

Разреженная матрица ~ 2 * 10^5 * 8 байт ~ 1.5 Мб

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} + \begin{pmatrix} a_2 \\ a_3 \\ \dots \\ a_n \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 \\ a_2 + a_3 \\ \dots \\ a_{n-1} + a_n \\ a_n \end{pmatrix}$$

Разреженные матрицы

Пример

Если вы все же поместили ее в оперативную память ...

Количество арифметических операций:

Плотная матрица ~ $2 * n^2$ ~ $2 * 10^{10}$ ~ ***0.2*** **секунды** (Core i7 3770)

Разреженная матрица ~ 3 * n ~ 3 * 10^5 ~ **10^{-6} секунды** (Core i7 3770)

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} + \begin{pmatrix} a_2 \\ a_3 \\ \dots \\ a_n \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 \\ a_2 + a_3 \\ \dots \\ a_{n-1} + a_n \\ a_n \end{pmatrix}$$

Разреженные матрицы. Виды

Разреженные матрицы

DOK
(Dictionary of Keys)

LIL
(List of Lists)

CSR
(compressed sparse row)

CRS
(compressed row storage)

CUSTOM

Разреженные матрицы. DOK

DOK - словарь по ключам
Ключ: номер строки, номер столбца
Значение: Элемент матрицы с заданными номерами строки и столбца
Применение: Для заполнения матриц

key (0, 0) value 5

key (0, 1) value 4

key (0, 3) value 7

key (1, 2) value 2

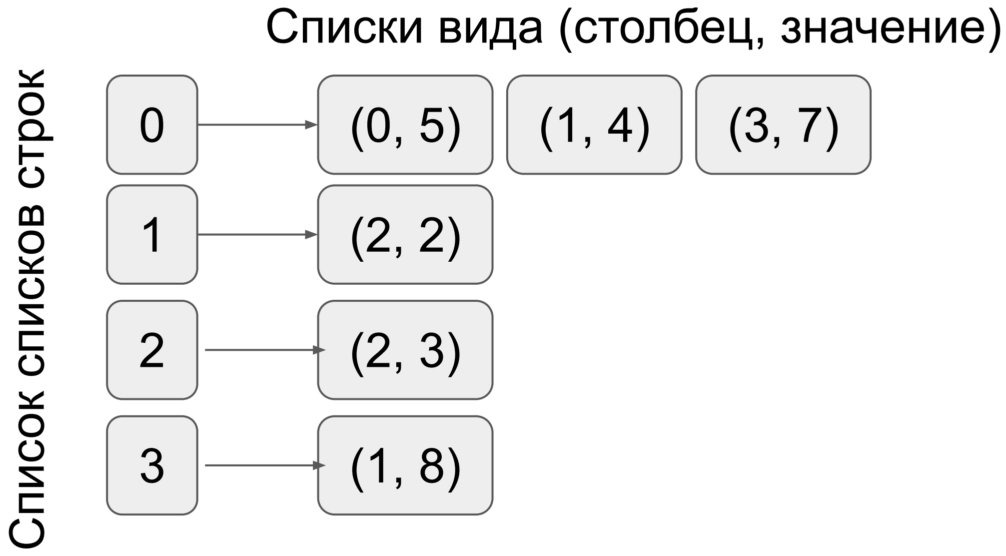
key (2, 2) value 3

key (3, 1) value 8

$$\begin{pmatrix} 5 & 4 & 0 & 7 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 8 & 0 & 0 \end{pmatrix}$$

Разреженные матрицы. LIL

LIL - список списков
Первый список содержит списки строк
Список строки содержит список элементов (номер столбца, значение)
Применение: транспонирование (списки столбцов)



$$\begin{pmatrix} 5 & 4 & 0 & 7 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 8 & 0 & 0 \end{pmatrix}$$

Разреженные матрицы. CSR

CSR - compressed sparse row

Три списка:

1. Значения ненулевых элементов (при обходе по строкам)
2. Столбцы ненулевых элементов (при обходе по строкам)
3. Количество ненулевых элементов с начала матрицы

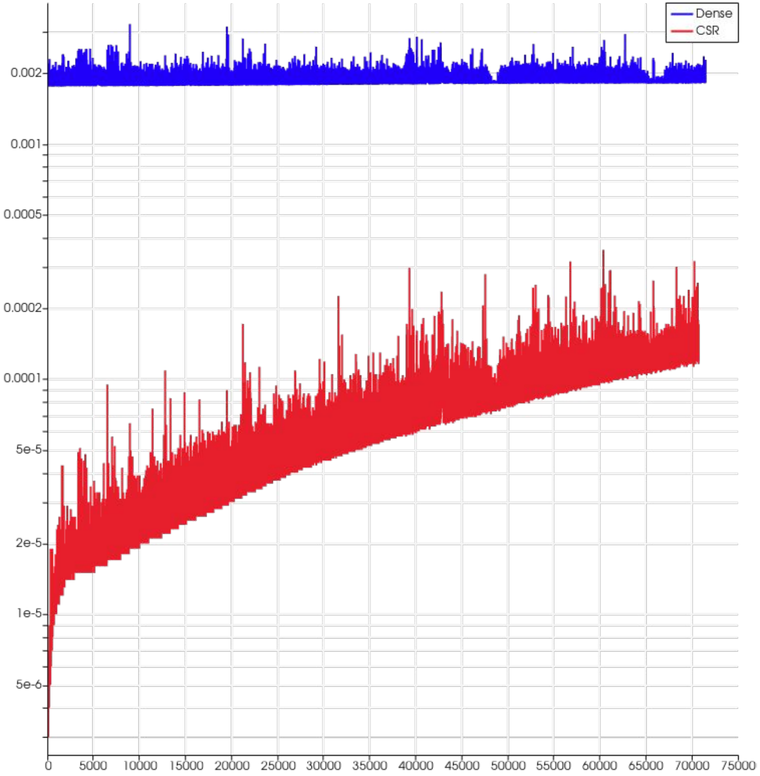
Применение: Умножение матриц

5	4	7	2	3	8
0	1	3	2	2	1
0	3	4	5	6	

нуль хранится для удобства !!!

$$\begin{pmatrix} 5 & 4 & 0 & 7 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 8 & 0 & 0 \end{pmatrix}$$

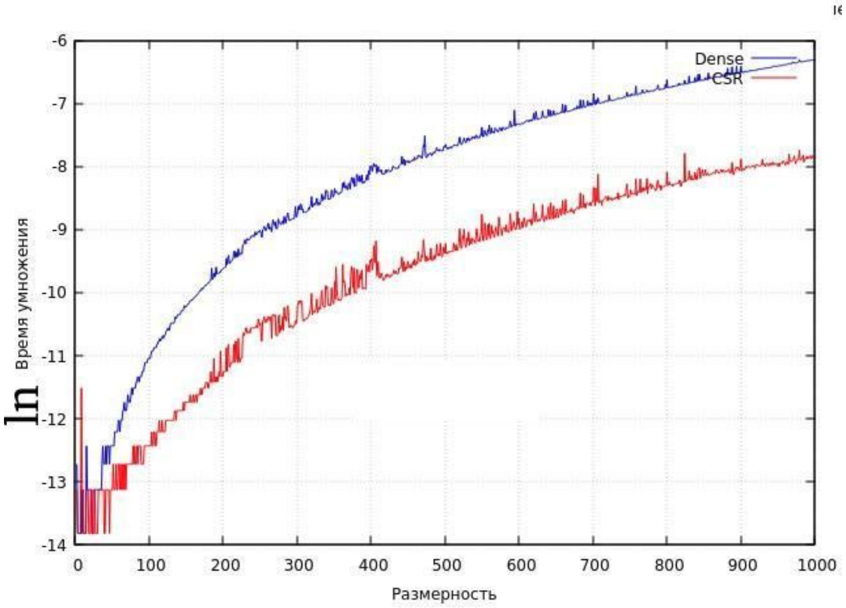
Разреженные матрицы. CSR



Зависимость времени умножения матрицы на столбец от количества ненулевых элементов

Вертикальная ось - время в секундах
Горизонтальная - количество ненулевых элементов

Разреженные матрицы. CSR



Зависимость времени умножения матрицы на столбец от размера при одинаковой заполненности (20%)

Вертикальная ось - ln (время в секундах)
Горизонтальная - размер матрицы

Разреженные матрицы. Custom

Конкретная задача - конкретная матрица !!!

Универсальные типы разреженных матриц не использую специфику задачи

$$\begin{pmatrix} b_1 & c_1 & 0 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ \dots & \dots & \dots & \dots \\ a_{n-1} & b_{n-1} & c_{n-1} & d_{n-1} \\ 0 & a_n & b_n & d_n \end{pmatrix}$$

Треугольная матрица для метода прогонки и столбец правой части

Норма вектора

Опр. Нормой вектора $\mathbf{x}$ называется правило, по которому каждому вектору соответствует число $||\mathbf{x}||$, причем:

1. $||\mathbf{x}|| = 0$ тогда и только тогда, когда $\mathbf{x} = \mathbf{0}$
2. $||\mathbf{x}|| > 0$, $\mathbf{x} \mathrel{!}= \mathbf{0}$
3. $||\lambda\, \mathbf{x}|| = |\lambda|\, ||\mathbf{x}||$, где λ - число
4. $||\mathbf{x} + \mathbf{y}|| \le ||\mathbf{x}|| + ||\mathbf{y}||$

Норма вектора

Распространенные нормы:

1. “Бесконечная” норма (на физтехе принято называть “первой”)

$$\| \boldsymbol{x} \|_{\infty} = max\, |x_i|$$

2. “Первая” норма (на физтехе принято называть “второй”)

$$\| \boldsymbol{x} \|_{1} = \sum_{i=1}^{n} |x_i|$$

3. “Вторая” норма (на физтехе принято называть “третьей”)

$$\| \boldsymbol{x} \|_{2} = \left(\sum_{i=1}^{n} x_i^{\,2} \right)^{0.5}$$

Норма вектора

Докажем, что “Вторая” норма является нормой:

$$\| \boldsymbol{x} \|_2 = \left(\sum_{i=1}^{n} x_i^{\,2} \right)^{0.5} = 0 \quad \leftrightarrow \quad x_i = 0, \; i = 1, \ldots, n \quad \leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}$$

$$\| \boldsymbol{x} \|_2 = \left(\sum_{i=1}^{n} x_i^{\,2} \right)^{0.5} > 0 \quad при \quad \boldsymbol{x} \; != \boldsymbol{0}$$

$$\| \lambda \boldsymbol{x} \|_2 = \left(\sum_{i=1}^{n} \lambda^2 \; x_i^{\,2} \right)^{0.5} = |\lambda| \left(\sum_{i=1}^{n} x_i^{\,2} \right)^{0.5} = \lambda \; \| \boldsymbol{x} \|_2$$

$$\| \boldsymbol{x} + \boldsymbol{y} \|_2 = \left(\sum_{i=1}^{n} \left(x_i + y_i \right)^2 \right)^{0.5} \overset{неравенство\ К.\,Б.}{\leq} \left(\sum_{i=1}^{n} x_i^{\,2} \right)^{0.5} + \left(\sum_{i=1}^{n} y_i^{\,2} \right)^{0.5} = \| \boldsymbol{x} \|_2 + \| \boldsymbol{y} \|_2$$

Подчиненная норма матрицы

Опр. Нормой матрицы ||*A*||, подчиненной норме вектора || * || назовем число:

$$|| A || = \sup_{x \,!= 0} \frac{|| A x ||}{|| x ||} = \sup_{|| x || = 1} || A x ||$$

Утв. Подчиненные нормы к вышеизложенным нормам векторов выражаются:

$$|| \boldsymbol{A} ||_{\infty} = \max_{i} \sum_{j=1}^{n} | A_{ij} |$$

Максимальная по строкам сумма модулей элементов в строке

Подчиненная норма матрицы

Опр. Нормой матрицы ||*A*||, подчиненной норме вектора || * || назовем число:

$$|| A || = \sup_{x \, != \, 0} \frac{|| A x ||}{|| x ||} = \sup_{|| x || = 1} || A x ||$$

Утв. Подчиненные нормы к вышеизложенным нормам векторов выражаются:

$$|| A ||_1 = \max_j \sum_{i=1}^{n} | A_{ij} |$$

Максимальная по столбцам сумма модулей элементов в столбце

Подчиненная норма матрицы

Опр. Нормой матрицы ||*A*||, подчиненной норме вектора || * || назовем число:

$$\| A \| = \sup_{x \,!= 0} \frac{\| A x \|}{\| x \|} = \sup_{\| x \| = 1} \| A x \|$$

Утв. Подчиненные нормы к вышеизложенным нормам векторов выражаются:

$$\| \boldsymbol{A} \|_2 = \sqrt{\lambda_{max}(A^T A)}$$

??? - непонятная вещь

Подчиненная норма матрицы

Докажем для второй нормы:

$$\| A \|_2 = \sqrt{\lambda_{max}(A^T A)}$$

$$\| A \|_2 = \sup_{\|x\| = 1} \|Ax\|_2 = \sup_{\|x\| = 1} \sqrt{(Ax)^T (Ax)} = \sup_{\|x\| = 1} \sqrt{x^T A^T A x}$$

Матрица $A^T A$ симметрична, поэтому существует ОНБ из собственных векторов $\{e\}_{i=1}^{N}$

$$\sup_{\|x\| = 1} \sqrt{\boldsymbol{x}^{\mathrm{T}} A^T A \boldsymbol{x}} = \sup_{\|x\| = 1} \sqrt{\boldsymbol{x}^{\mathrm{T}} A^T A \sum_i \alpha_i \boldsymbol{e}_i} = \sup_{\|x\| = 1} \sqrt{\boldsymbol{x}^{\mathrm{T}} \sum_i \lambda_i \alpha_i \boldsymbol{e}_i} = \sup_{\|x\| = 1} \sqrt{\sum_{\mathrm{i}} \boldsymbol{\alpha}_i \boldsymbol{e}_i^{\mathrm{T}} \sum_i \lambda_i \alpha_i \boldsymbol{e}_i}$$

Подчиненная норма матрицы

Докажем для второй нормы:

$$|| A ||_2 = \sqrt{\lambda_{max}(A^T A)}$$

$$\sup_{||x|| = 1} \sqrt{\sum_{\mathrm{i}} \alpha_i \boldsymbol{e}_i^{\mathrm{T}} \sum_i \lambda_i \alpha_i \boldsymbol{e}_i} = \sup_{||x|| = 1} \sqrt{\sum_i \lambda_i \alpha_i^2 \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_i} \le \sqrt{\lambda_{max}} \sup_{||x|| = 1} \sqrt{\sum_i \alpha_i^2 \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_i} = \sqrt{\lambda_{max}}$$

λ_{max} является именно верхней гранью, так как она достигается на собственном векторе с максимальным собственным числом матрицы A^TA

$\lambda_{max} > 0$, так как собственные числа матрицы A^TA не отрицательны, а в случае невырожденности матрицы А положительны

Подчиненная норма матрицы

Свойства нормы:

$$\| \lambda A \| = \sup_{\|x\| = 1} \| \lambda A x \| = |\lambda| \sup_{\|x\| = 1} \| A x \| = |\lambda| \, \| A \|$$

$$\frac{\|A y\|}{\|y\|} \leq \sup_{\|x\| = 1} \|A x\| = \|A\| \quad \rightarrow \quad \|A y\| \leq \|A\| \, \|y\|$$

Простейшие итерационные методы

Опр. Пусть дана система уравнений:

$$A\,x = b$$

Запишем ее в эквивалентной форме каким-либо образом:

$$\boldsymbol{x} = P\boldsymbol{x} + c$$

Простейшим итерационным методом назовем процесс с начальным приближением $\boldsymbol{x}_0$:

$$\boldsymbol{x}_{i+1} = P\boldsymbol{x}_i + \boldsymbol{c},\ i = 0,\ \ldots$$

Простейшие итерационные методы. Сходимость

Опр. Простейший итерационный метод сходится, если последовательность $\boldsymbol{x}_{i}$, имеет конечный предел $\boldsymbol{x}^*$

$$x_i \xrightarrow{i \to \infty} x^*$$

Утв. Если простейший итерационный процесс сходится, то он сходится к решению:

$$x^* = Px^* + c \leftrightarrow Ax^* = b$$

Простейшие итерационные методы. Сходимость

Суть итерационных методов:

Продолжать итерационный процесс до тех пор, пока решение на i-ой итерации $\boldsymbol{x}_{i,}$ не будет достаточно билзко к “честному” решению $\boldsymbol{x}^*$.

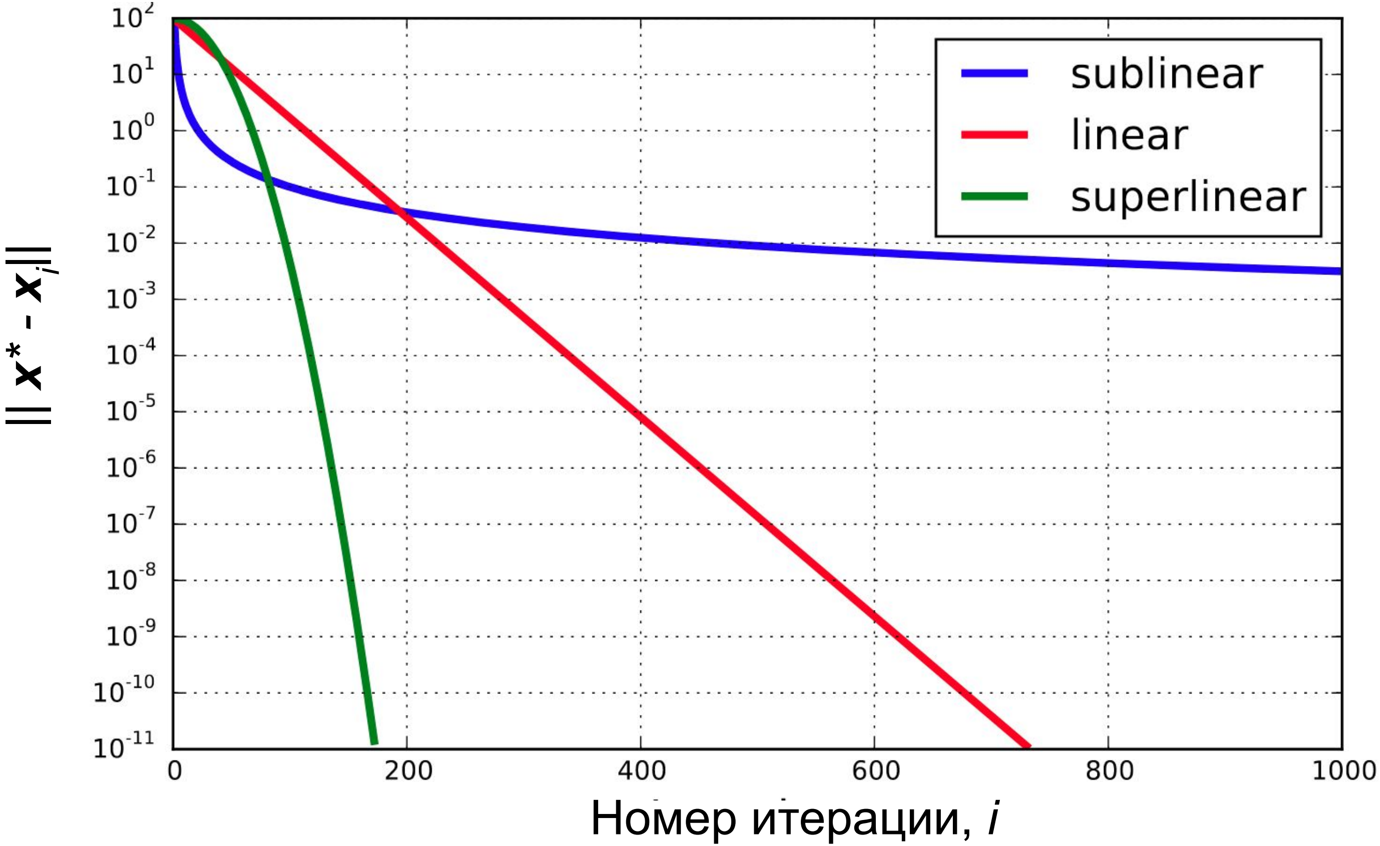
Как установить близость $\boldsymbol{x}_{i,}$ к неизвестному $\boldsymbol{x}^*$? - рассмотрим далее

Простейшие итерационные методы. Сходимость

Виды сходимости:

1. Линейная $\| \boldsymbol{x}^* - \boldsymbol{x}_i \| < C q^i$, где $0 < q < 1$ (как убывающая геом. прогрессия)
2. Сверхлинейная $\| \boldsymbol{x}^* - \boldsymbol{x}_i \| < C q^i$ для любого $0 < q < 1$ *(быстрее линейной)*
3. Сублинейная - медленнее линейной ($\| \boldsymbol{x}^* - \boldsymbol{x}_i \| < q^i$, где $0 < q < 1$ не выполняется)

Простейшие итерационные методы. Сходимость



Простейшие итерационные методы

Примеры простейших итерационных методов:

1. Метод **Якоби.** *D* - матрица диагональных элементов, *L* - матрица поддиагональных элементов, *U* - наддиагональных:

$$Ax = b \quad \leftrightarrow \quad (L + D + U)x = b$$

$$\leftrightarrow \quad x = D^{-1}(b - (L + U)\,x)$$

$$x_{i+1} = D^{-1}\left(b - (L + U)x_i\right),\ P = -\,D^{-1}(L + U)$$

Простейшие итерационные методы

Примеры простейших итерационных методов:

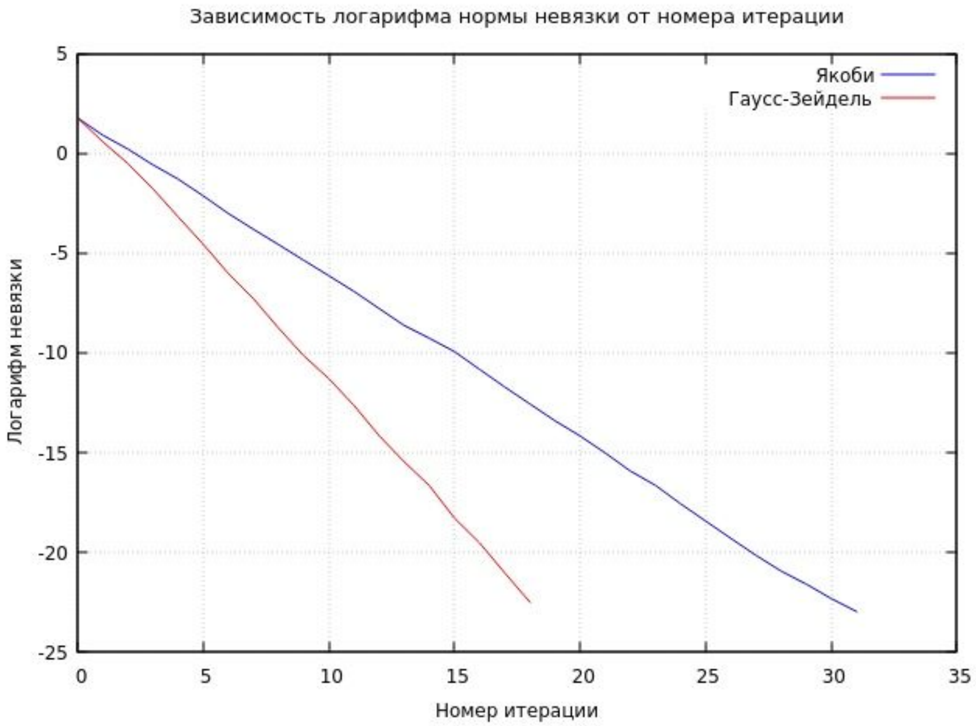
2. Метод **Гаусса - Зейделя.** *D* - матрица диагональных элементов, *L* - матрица поддиагональных элементов, *U* - наддиагональных:

$$A x = b \quad \leftrightarrow \quad (L + D + U) x = b$$

$$\leftrightarrow \quad x = (L + D)^{-1}(b - U x)$$

$$x_{i+1} = (L + D)^{-1}\left(b - Ux_i\right), \ P = -(L + D)^{-1} U$$

Простейшие итерационные методы



Зависимость невязки от числа итераций для методов Якоби и Гаусса-Зейделя

Вертикальная ось - ln(невязки)
Горизонтальная - число итераций

Простейшие итерационные методы

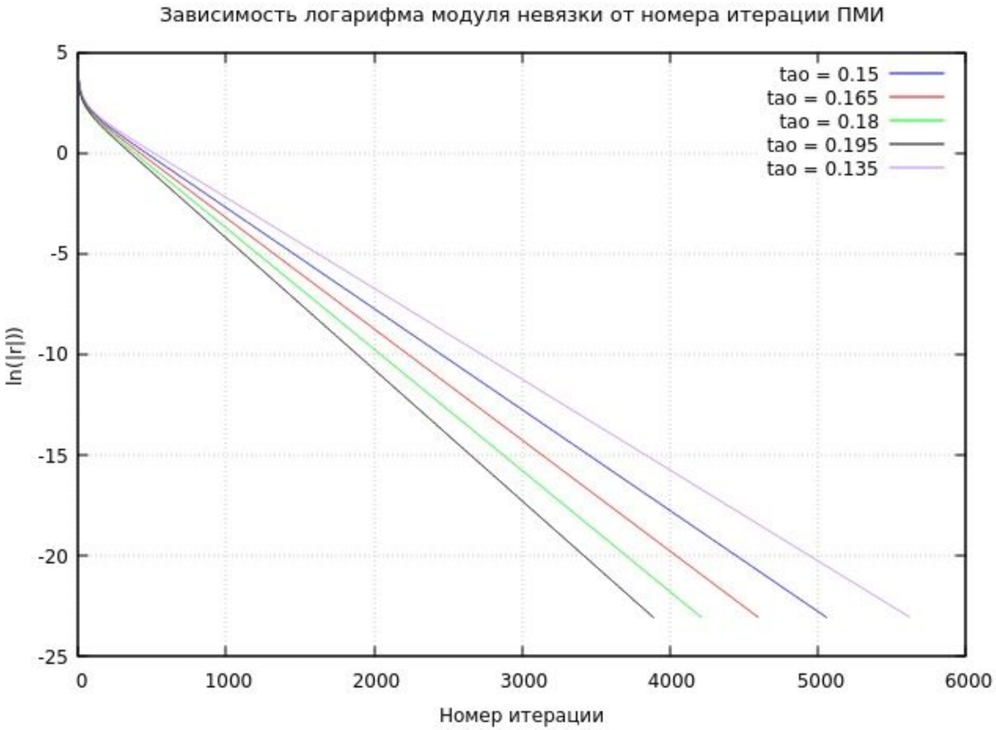
Примеры простейших итерационных методов:

3. Метод **Простой итерации**

$$A\,x = b \quad \leftrightarrow \quad x = x + \tau(b - Ax)\,,\ \tau > 0$$

$$x_{i+1} = (E - \tau A)\,x_i + \tau b, \quad P = (E - \tau A)$$

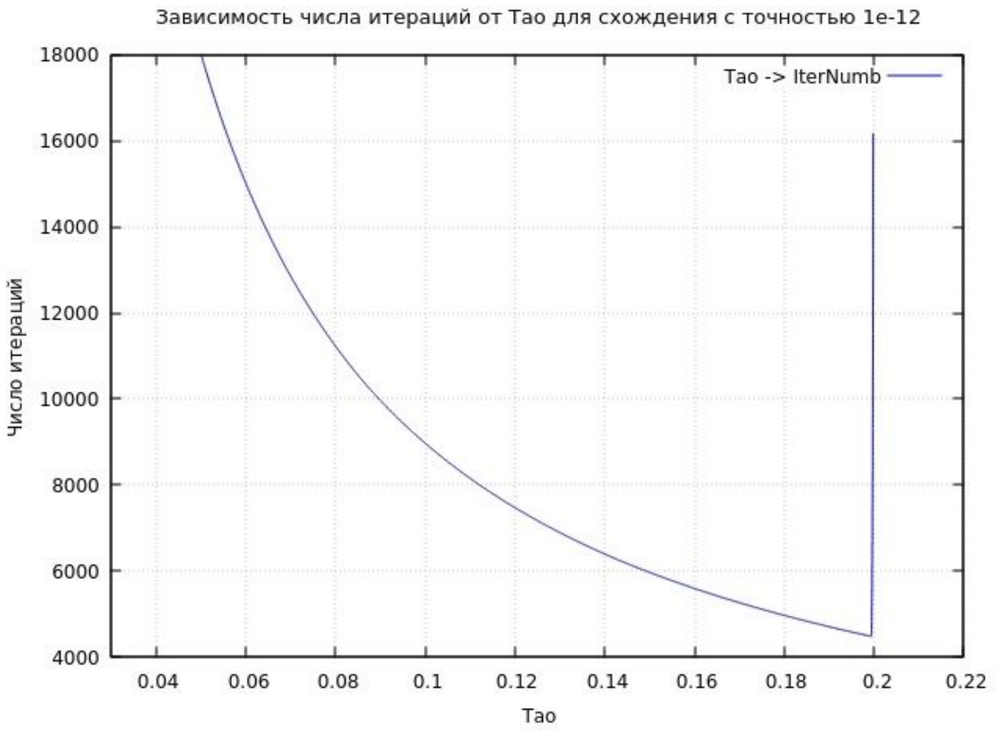
Простейшие итерационные методы



Зависимость невязки от числа итераций при разных параметрах тау

Вертикальная ось - ln(невязки)
Горизонтальная - число итераций

Простейшие итерационные методы



Зависимость числа итераций для достижения определенной невязки от параметра тау

Вертикальная ось - кол-во итераций
Горизонтальная - параметр τ

При τ ≥ 0.2 метод расходится!!!

Простейшие итерационные методы

Примеры простейших итерационных методов:

3. Метод **Простой итерации**

$$x_{i+1} = (E - \tau A)\, x_i + \tau b, \quad P = (E - \tau A)$$

Замечания:

1) Как выбирать параметр *τ*?
2) Можно ли выбирать *τ* разным на разных итерациях?

Простейшие итерационные методы

Теорема. Простейший итерационный метод сходится из любого начального приближения, если || P || < 1. При этом сходимость линейная

$$||x^* - x_i|| = ||x^* - Px_{i-1} - \boldsymbol{c}||$$

$$||Px^* - \boldsymbol{c} - Px_{i-1} + \boldsymbol{c}|| = ||P\left(x^* - x_{i-1}\right)||$$

$$||x^* - x_i|| = \leq q\ ||x^* - x_{i-1}||,\ q = ||P|| < 1$$

Простейшие итерационные методы

Теорема. Простейший итерационный метод сходится из любого начального приближения, если || P || < 1. При этом сходимость линейная

$$||x^* - x_i|| = \leq q^i \ ||x^* - x_0||, \ q = ||P|| < 1$$

Теорема. Простейший итерационный метод сходится из любого начального приближения тогда и только тогда, когда все собственные числа матрицы P мо модулю меньше 1

Простейшие итерационные методы

Теорема. Если матрица исходной системы положительно определена и симметрична, то метод **Гаусса - Зейделя** сходится из любого начального приближения

Теорема. Если матрица исходной системы имеет диагональное преобладание (см. Лекцию 1), то метод **Якоби** сходится из любого начального приближения

Простейшие итерационные методы

Интерпретация теоремы.

Пусть $\boldsymbol{e}_i = \boldsymbol{x}_i - \boldsymbol{x}^*$ - ошибка решения.

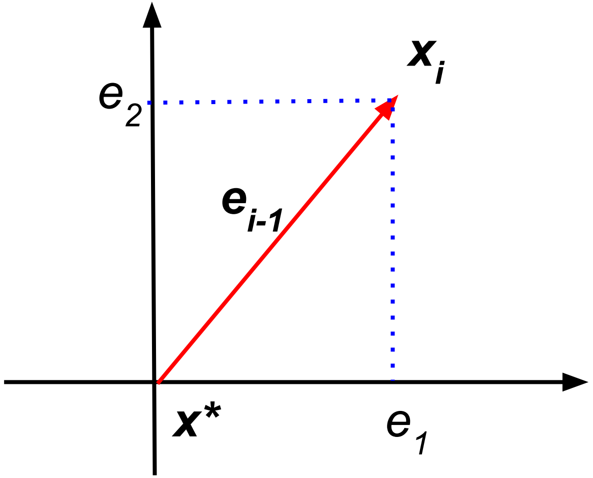
$$e_i = x_i - x^* = Px_{i-1} - c - Px^* + c = P\left(x_{i-1} - x^*\right) = P\,e_{i-1}$$

$$\boxed{\boldsymbol{e}_i = P\,\boldsymbol{e}_{i-1}}$$

Простейшие итерационные методы

Интерпретация теоремы.

$$\boxed{e_i = P\, e_{i-1}}$$



Пусть λ_1 и λ_2 - собственные векторы P, тогда:

$$e_i = P\, e_{i-1} = \lambda_1 e_1 + \lambda_2 e_2$$

Если хотя бы одно собственное число не меньше 1 по модулю, то компонента ошибки не убывает с ростом итераций

Если все собственные числа меньше 1 по модулю, то ошибка убывает с ростом итераций

Критерий остановки

Когда заканчивать итерационный процесс?

1. После фиксированного числа итераций
2. После уменьшения невязки $\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}_i$ до определенного уровня
3. По оценке ошибки $\boldsymbol{e}_i = \boldsymbol{x}_i - \boldsymbol{x}^*$

$$||\boldsymbol{x}^* - \boldsymbol{x}_{\mathbf{i}}|| = q^N \, || \boldsymbol{e}_{\mathbf{0}}||$$

- За N итераций можно гарантировать убыть ошибки в q^N раз

q - характеристика метода (матрицы P)
- можно оценить по виду матрицы P

Численная ошибка

Влияет ли численная ошибка на результат ?

Пусть на итерации i решение $\boldsymbol{x}_i$ зашумлено на ошибку $\boldsymbol{\varepsilon}$. Значит на (i+1)-ой итерации зашумление уменьшится в q раз!

Ошибка округления не так страшна, как в прямых методах!

В следующий раз ...

1) Метод SOR
2) Полиномы Чебышева
3) Выбор параметра для метода простой итерации
4) Чебышевское ускорение и его устойчивость