Вычислительная математика. Итерационные методы решения систем линейных уравнений.

МФТИ

Недостатки прямых методов

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n;\ J = \frac{\partial F}{\partial x} = \left[\frac{\partial F_i}{\partial x_j}(x_0)\right]$$

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n; \ J = \frac{\partial F}{\partial x} = \left[\frac{\partial F_i}{\partial x_j}(x_0) \right]$$

$$J(x_0)\Delta x = F(x_0 + \Delta x) - F(x_0) + O(\Delta x^2)$$

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n; \; J = \frac{\partial F}{\partial x} = \left[\frac{\partial F_i}{\partial x_j}(x_0) \right]$$

$$J(x_0)\Delta x = F(x_0 + \Delta x) - F(x_0) + O(\Delta x^2)$$

- Нельзя найти приближение к решению с заданной погрешностью

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n;\ J = \frac{\partial F}{\partial x} = \left[\frac{\partial F_i}{\partial x_j}(x_0)\right]$$

$$J(x_0)\Delta x = F(x_0 + \Delta x) - F(x_0) + O(\Delta x^2)$$

- Нельзя найти приближение к решению с заданной погрешностью
- Сложно использовать специальную структуру матрицы

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n;\ J = \frac{\partial F}{\partial x} = \left[\frac{\partial F_i}{\partial x_j}(x_0)\right]$$
$$J(x_0)\Delta x = F(x_0 + \Delta x) - F(x_0) + O(\Delta x^2)$$

- Нельзя найти приближение к решению с заданной погрешностью
- Сложно использовать специальную структуру матрицы
  Пример: если A - разреженная, матрицы L и U, Q и R не будут в общем случае разреженными

Потеря разреженности (демо)

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$Ax = b \iff x = Sx + f$$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$Ax = b \iff x = Sx + f$$
$$x^k = Sx^{k-1} + f$$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$Ax = b \iff x = Sx + f$$

$$x^k = Sx^{k-1} + f$$

$$e^k = (x^k - x) = S(x^{k-1} - x) =$$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$
\begin{aligned}
Ax = b \iff x &= Sx + f \\
x^k &= Sx^{k-1} + f \\
e^k = (x^k - x) &= S(x^{k-1} - x) = \\
&= S^k(x^0 - x) = S^k e^0
\end{aligned}
$$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$\begin{aligned} Ax = b \iff x &= Sx + f \\ x^k &= Sx^{k-1} + f \\ e^k = (x^k - x) &= S(x^{k-1} - x) = \\ &= S^k(x^0 - x) = S^k e^0 \end{aligned}$$

- Критерий сходимости:
$e^k \to 0 \iff S^k e^0 \to 0 \; \forall e^0 \iff \rho(S) < 1$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$
\begin{aligned}
Ax = b \iff x &= Sx + f \\
x^k &= Sx^{k-1} + f \\
e^k = (x^k - x) &= S(x^{k-1} - x) = \\
&= S^k(x^0 - x) = S^k e^0
\end{aligned}
$$

- Критерий сходимости:
$e^k \to 0 \iff S^k e^0 \to 0 \; \forall e^0 \iff \rho(S) < 1$
- Достаточное условие сходимости: $\|S\| = q < 1$

$$
\|e^k\| = \|S^k e^0\| \le \|S\|^k \|e^0\| = q^k \|e^0\|
$$

Общий вид простого метода

- Метод простой итерации (МПИ) основан на сведении системы к эквивалентному виду:

$$
\begin{aligned}
Ax = b \iff x &= Sx + f \\
x^k &= Sx^{k-1} + f \\
e^k = (x^k - x) &= S(x^{k-1} - x) = \\
&= S^k(x^0 - x) = S^k e^0
\end{aligned}
$$

- Критерий сходимости:
  $e^k \to 0 \iff S^k e^0 \to 0 \; \forall e^0 \iff \rho(S) < 1$
- Достаточное условие сходимости: $\|S\| = q < 1$

$$
\|e^k\| = \|S^k e^0\| \le \|S\|^k \|e^0\| = q^k \|e^0\|
$$

- Нужно задать начальное приближение x^0 и *критерий остановки*, например: $\|Ax^k - b\| \le \epsilon$

Общий вид с предобуславливателем

- Любой МПИ можно записать в виде:

$$P(x^{k+1} - x^k) = b - Ax^k,\ k = 0, 1, \dots,\ P - \text{невырожденная}$$

- $x^{k+1} = (I - P^{-1}A)x^k + P^{-1}b, \quad S = (I - P^{-1}A)$
- Требования к P:
  1. $P \approx A^{-1}$
  2. Решение системы $Py = g$ стоит $\ll O(n^3)$
- Матрицу P (иногда P^{-1}) называют *предобуславливатель* (preconditioner)

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

- Сложность $O(Kn^2)$, где K - число итераций.

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

- Сложность $O(Kn^2)$, где K - число итераций.
- Можно оценить число итераций для заданной точности $\|e^k\| \leq q^k \|e^0\| \leq \epsilon \Rightarrow k \geq \dfrac{\log(\epsilon/\|e^0\|)}{\log q}$

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

- Сложность $O(Kn^2)$, где K - число итераций.
- Можно оценить число итераций для заданной точности $\|e^k\| \leq q^k\|e^0\| \leq \epsilon \Rightarrow k \geq \dfrac{\log(\epsilon/\|e^0\|)}{\log q}$
- Нужна только процедура для вычисления произведения Sx

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

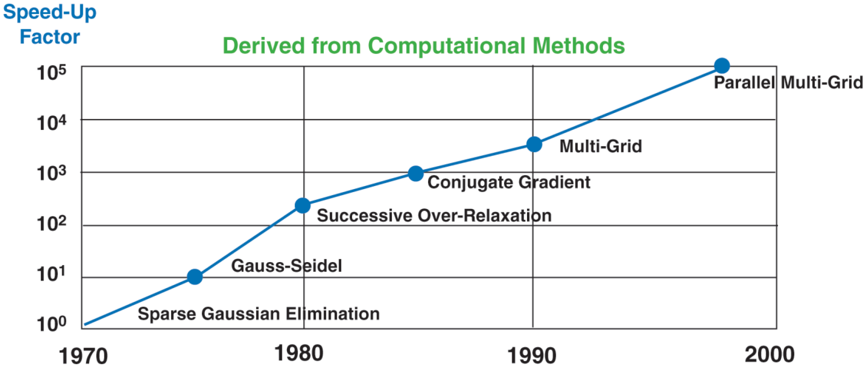
- Сложность $O(Kn^2)$, где K - число итераций.
- Можно оценить число итераций для заданной точности $\|e^k\| \leq q^k \|e^0\| \leq \epsilon \Rightarrow k \geq \dfrac{\log(\epsilon/\|e^0\|)}{\log q}$
- Нужна только процедура для вычисления произведения Sx
- Матрично-векторное умножение легко распараллеливается с высокой эффективностью.

Преимущества итерационных методов

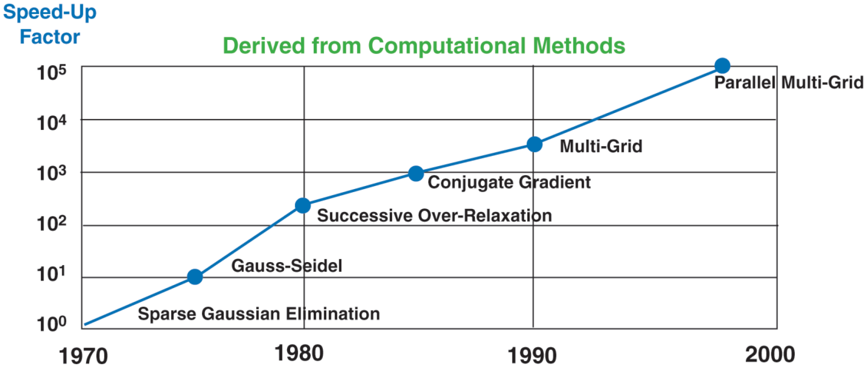
$$x^k = Sx^{k-1} + f$$

- Сложность $O(Kn^2)$, где K - число итераций.
- Можно оценить число итераций для заданной точности $\|e^k\| \leq q^k \|e^0\| \leq \epsilon \Rightarrow k \geq \dfrac{\log(\epsilon/\|e^0\|)}{\log q}$
- Нужна только процедура для вычисления произведения Sx
- Матрично-векторное умножение легко распараллеливается с высокой эффективностью.
- Для разреженных матриц сложность $O(K \times NNZ(S))$

Развитие итерационных методов



Развитие итерационных методов



Список 10 важнейших алгоритмов 20-го века, *Computing in Science and Engineering*, 2001 год:

- *итерационные методы* подпространств Крылова

Метод Ричардсона

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

Метод Ричардсона

- Рассмотрим систему $Ax = b,\ A = A^T > 0$:

$$Ax = b$$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$\begin{aligned} &Ax = b \\ &\tau(Ax - b) = 0 \end{aligned}$$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$
\begin{aligned}
&Ax = b \\
&\tau(Ax - b) = 0 \\
&x - \tau(Ax - b) = x
\end{aligned}
$$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$
\begin{aligned}
& Ax = b \\
& \tau(Ax - b) = 0 \\
& x - \tau(Ax - b) = x \\
& \underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}
$$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$
\begin{aligned}
&Ax = b \\
&\tau(Ax - b) = 0 \\
&x - \tau(Ax - b) = x \\
&\underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}
$$

τ - итерационный параметр. $S = I - \tau A$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$
\begin{aligned}
& Ax = b \\
& \tau(Ax - b) = 0 \\
& x - \tau(Ax - b) = x \\
& \underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}
$$

  τ - итерационный параметр. $S = I - \tau A$

- Критерий сходимости:

Метод Ричардсона

- Рассмотрим систему $Ax = b,\ A = A^T > 0$:

$$
\begin{aligned}
&Ax = b\\
&\tau(Ax - b) = 0\\
&x - \tau(Ax - b) = x\\
&\underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}
$$

τ - итерационный параметр. $S = I - \tau A$

- Критерий сходимости:

$$\text{sp}(I - \tau A) = 1 - \tau\ \text{sp}(A),\ \lambda_i(A) > 0$$

Метод Ричардсона

- Рассмотрим систему $Ax = b, \; A = A^T > 0$:

$$\begin{aligned} &Ax = b \\ &\tau(Ax - b) = 0 \\ &x - \tau(Ax - b) = x \\ &\underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b} \end{aligned}$$

  τ - итерационный параметр. $S = I - \tau A$

- Критерий сходимости:

$$\begin{aligned} &\text{sp}(I - \tau A) = 1 - \tau \; \text{sp}(A), \; \lambda_i(A) > 0 \\ &|1 - \tau\lambda_i| < 1, \; i = 1, \dots, n \Rightarrow \end{aligned}$$

Метод Ричардсона

- Рассмотрим систему $Ax = b,\ A = A^T > 0$:

$$
\begin{aligned}
&Ax = b\\
&\tau(Ax - b) = 0\\
&x - \tau(Ax - b) = x\\
&\underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}
$$

  τ - итерационный параметр. $S = I - \tau A$

- Критерий сходимости:

$$
\begin{aligned}
&\mathrm{sp}(I - \tau A) = 1 - \tau\ \mathrm{sp}(A),\ \lambda_i(A) > 0\\
&|1 - \tau\lambda_i| < 1,\ i = 1, \dots, n \Rightarrow\\
&(1 - \tau\lambda_i)^2 < 1 \Rightarrow\ \tau\lambda_i(\tau\lambda_i - 2) < 0 \Rightarrow \underline{0 < \tau < \frac{2}{\lambda_{max}}}
\end{aligned}
$$

Оптимальный параметр

Оптимальный параметр

- $\|e^k\|_2 = \|(I - \tau A)^k e^0\|_2 \leq \|(I - \tau A)\|_2^k \|e^0\|_2$

Оптимальный параметр

- $\|e^k\|_2 = \|(I - \tau A)^k e^0\|_2 \leq \|(I - \tau A)\|_2^k \|e^0\|_2$
- $\|(I - \tau A)\|_2 = \max\limits_{\lambda \in sp(A)} |1 - \tau\lambda|$

Оптимальный параметр

- $\|e^k\|_2 = \|(I - \tau A)^k e^0\|_2 \leq \|(I - \tau A)\|_2^k \|e^0\|_2$
- $\|(I - \tau A)\|_2 = \max\limits_{\lambda \in sp(A)} |1 - \tau\lambda|$
- Часто, известны только оценки границы спектра:

$$sp(A) \in [m, M]$$

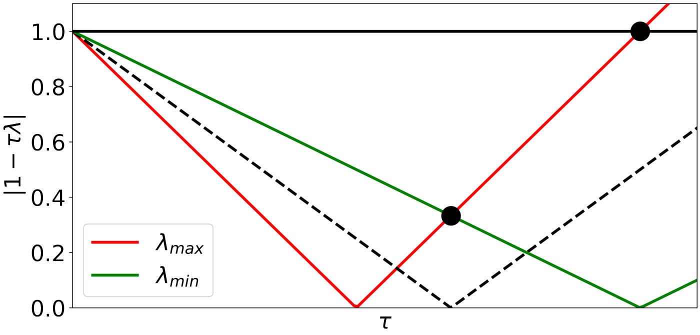
Оптимальный параметр

- $\|e^k\|_2 = \|(I - \tau A)^k e^0\|_2 \leq \|(I - \tau A)\|_2^k \|e^0\|_2$
- $\|(I - \tau A)\|_2 = \max\limits_{\lambda \in sp(A)} |1 - \tau\lambda|$
- Часто, известны только оценки границы спектра:

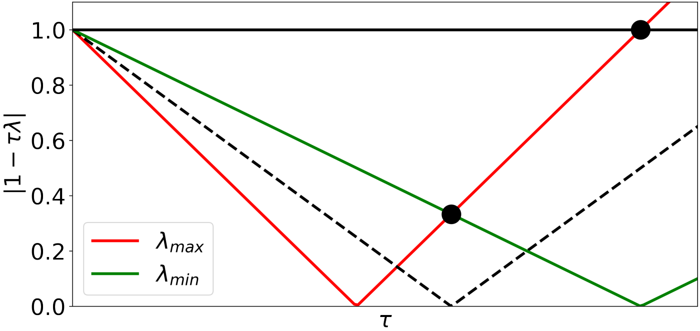
$$sp(A) \in [m, M]$$

- $\tau^* = \arg\min\limits_{\tau} \max\limits_{\lambda \in [m,M]} |1 - \tau\lambda|$

Оптимальный параметр



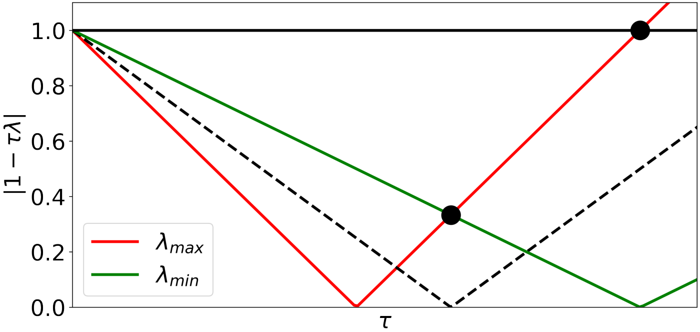
Оптимальный параметр



Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M + m} \Rightarrow$$

Оптимальный параметр

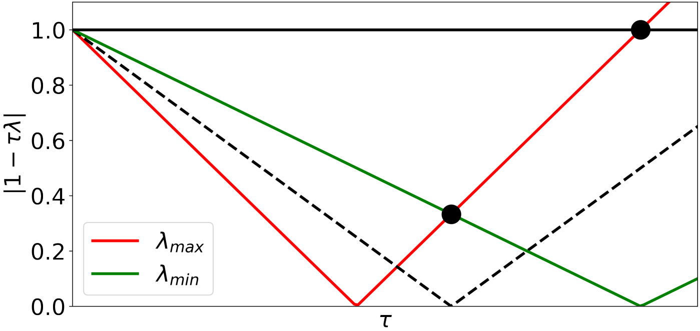


Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M+m} \Rightarrow$$

$$\|(I - \tau^* A)\|_2 = q^* = \frac{M-m}{M+m} =$$

Оптимальный параметр

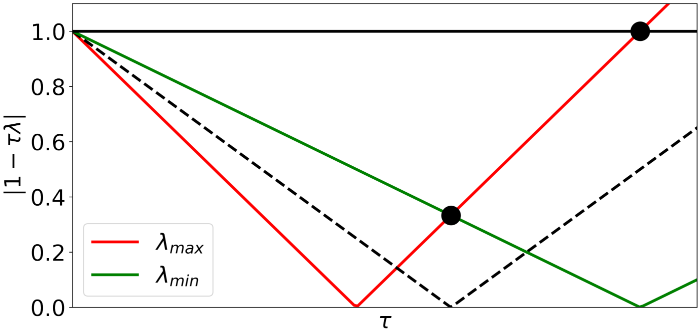


Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M+m} \Rightarrow$$

$$\|(I - \tau^* A)\|_2 = q^* = \frac{M-m}{M+m} = \frac{M/m - 1}{M/m + 1} =$$

Оптимальный параметр

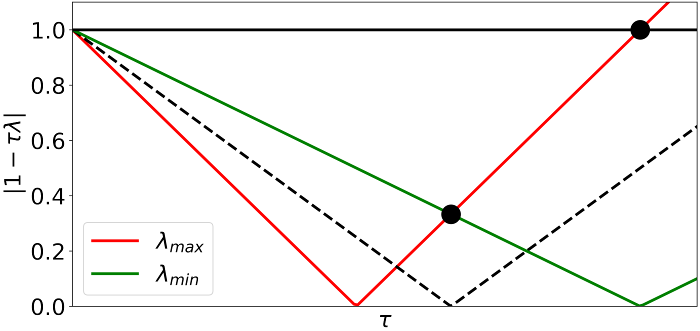


Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M+m} \Rightarrow$$

$$\|(I - \tau^* A)\|_2 = q^* = \frac{M-m}{M+m} = \frac{M/m - 1}{M/m + 1} = \frac{cond_2(A) - 1}{cond_2(A) + 1}$$

Оптимальный параметр



Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M+m} \Rightarrow$$

$$\|(I - \tau^* A)\|_2 = q^* = \frac{M-m}{M+m} = \frac{M/m - 1}{M/m + 1} = \frac{cond_2(A) - 1}{cond_2(A) + 1}$$

Скорость сходимости зависит от обусловленности!

Методы Якоби и Гаусса-Зейделя

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b$,

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$): $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
  - Диагональные матрицы – $O(n)$
  - Треугольные матрицы – $O(n^2)$

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
  - Диагональные матрицы – $O(n)$
  - Треугольные матрицы – $O(n^2)$
- $A = L + D + U$,
  L - строго нижнетреугольная часть
  D - диагональная часть
  U - строго верхнетреугольная часть.

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
  - Диагональные матрицы – $O(n)$
  - Треугольные матрицы – $O(n^2)$
- $A = L + D + U$,
  L - строго нижнетреугольная часть
  D - диагональная часть
  U - строго верхнетреугольная часть.
- **Предполагается, что D - невырожденная.**

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
  - Диагональные матрицы – $O(n)$
  - Треугольные матрицы – $O(n^2)$
- $A = L + D + U$,
  L - строго нижнетреугольная часть
  D - диагональная часть
  U - строго верхнетреугольная часть.
- **Предполагается, что D - невырожденная.**
- Метод Якоби: $Dx^{k+1} = -(L + U)x^k + b$

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
  $Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
  - Диагональные матрицы – $O(n)$
  - Треугольные матрицы – $O(n^2)$
- $A = L + D + U$,
  L - строго нижнетреугольная часть
  D - диагональная часть
  U - строго верхнетреугольная часть.
- **Предполагается, что D - невырожденная.**
- Метод Якоби: $Dx^{k+1} = -(L + U)x^k + b$
- Метод Зейделя: $(L + D)x^{k+1} = -Ux^k + b$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Теорема 7.1 Достаточное условие сходимости метода Якоби

Если матрица A имеет строчное диагональное преобладание, то метод Якоби сходится.

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Теорема 7.1 Достаточное условие сходимости метода Якоби

Если матрица A имеет строчное диагональное преобладание, то метод Якоби сходится.

- докажем, что $\|D^{-1}(L+U)\|_\infty < 1$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Теорема 7.1 Достаточное условие сходимости метода Якоби

Если матрица A имеет строчное диагональное преобладание, то метод Якоби сходится.

- докажем, что $\|D^{-1}(L+U)\|_\infty < 1$
- $|a_{ii}| > \sum\limits_{j \neq i} |a_{ij}|,\ i = 1, \ldots, n$ (определение диагонального преобладания) $\Rightarrow$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Теорема 7.1 Достаточное условие сходимости метода Якоби

Если матрица A имеет строчное диагональное преобладание, то метод Якоби сходится.

- докажем, что $\|D^{-1}(L+U)\|_\infty < 1$
- $|a_{ii}| > \sum\limits_{j \neq i} |a_{ij}|,\ i = 1, \dots, n$ (определение диагонального преобладания) $\Rightarrow$
- $\|D^{-1}(L+U)\|_\infty = \max\limits_i \sum\limits_{j \neq i} \left| \dfrac{a_{ij}}{a_{ii}} \right| = \max\limits_i \dfrac{\sum_{i \neq j} |a_{ij}|}{|a_{ii}|} < 1 \square$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Теорема 7.2 Необходимое и достаточное условие сходимости метода Якоби

Для сходимости метода Якоби необходимо и достаточно, чтобы все корни уравнения

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \lambda a_{nn} \end{vmatrix} = 0$$

по модулю не превосходили единицы ($|\lambda| \leqslant 1$).

Алгоритм метода Якоби

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, n; j \= i
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

Алгоритм метода Якоби

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, n; j \= i
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

- Метод Якоби редко используется в *чистом виде*

Алгоритм метода Якоби

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, n; j \= i
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

- Метод Якоби редко используется в *чистом виде*
- Acceleration of the Jacobi iterative method by factors exceeding 100 using scheduled relaxation, JCP, 2014

Пример программы для метода Якоби

Метод Зейделя

$$x^{k+1} = -(L + D)^{-1} U x^k + (L + D)^{-1} b$$

Метод Зейделя

$$x^{k+1} = -(L + D)^{-1} U x^k + (L + D)^{-1} b$$

Теорема 7.3 Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n \times n}$, $A = A^T > 0$, то метод Зейделя сходится.

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Теорема 7.3 Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n\times n}$, $A = A^T > 0$, то метод Зейделя сходится.

- $A = A^T > 0 \Rightarrow \quad A = L + D + L^T$

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Теорема 7.3 Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n \times n}$, $A = A^T > 0$, то метод Зейделя сходится.

- $A = A^T > 0 \Rightarrow \quad A = L + D + L^T$
- $\langle Ae_k, e_k \rangle = a_{kk} > 0 \Rightarrow D > 0$

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Теорема 7.3 Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n\times n}$, $A = A^T > 0$, то метод Зейделя сходится.

- $A = A^T > 0 \Rightarrow \quad A = L + D + L^T$
- $\langle Ae_k, e_k\rangle = a_{kk} > 0 \Rightarrow D > 0$
- $x^{k+1} = -(L+D)^{-1}L^T x^k + (L+D)^{-1}b = \boxed{(I-(L+D)^{-1}A)}x^k + (L+D)^{-1}b$

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Теорема 7.3 Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n\times n}$, $A = A^T > 0$, то метод Зейделя сходится.

- $A = A^T > 0 \Rightarrow \quad A = L + D + L^T$
- $\langle Ae_k, e_k\rangle = a_{kk} > 0 \Rightarrow D > 0$
- $x^{k+1} = -(L+D)^{-1}L^Tx^k + (L+D)^{-1}b =$ $\boxed{(I - (L+D)^{-1}A)}x^k + (L+D)^{-1}b$
- Можно ввести A-норму: $\|v\|_A^2 \equiv \langle Av, v\rangle = \langle v, v\rangle_A$, где $\langle\cdot,\cdot\rangle$ - обычное скалярное произведение

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \langle Ae^{k+1}, e^{k+1} \rangle =$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1} \right\rangle =$$
$$= \left\langle A(I - (L+D)^{-1}A)e^k, (I - (L+D)^{-1}A)e^k \right\rangle =$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1} \right\rangle =$$

$$= \left\langle A(I - (L+D)^{-1}A)e^k,\, (I - (L+D)^{-1}A)e^k \right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k \right\rangle \boxed{=}$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1} \right\rangle =$$

$$= \left\langle A(I - (L+D)^{-1}A)e^k,\ (I - (L+D)^{-1}A)e^k \right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k \right\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \langle Ae^{k+1}, e^{k+1}\rangle =$$

$$= \langle A(I-(L+D)^{-1}A)e^k,\ (I-(L+D)^{-1}A)e^k\rangle =$$

$$= \langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

$$\boxed{=} \langle Ae^k, e^k\rangle - \langle Av, e^k\rangle - \langle Ae^k, v\rangle + \langle Av, v\rangle =$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1} \right\rangle =$$

$$= \left\langle A(I-(L+D)^{-1}A)e^k,\ (I-(L+D)^{-1}A)e^k \right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k \right\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

$$\boxed{=} \left\langle Ae^k, e^k \right\rangle - \left\langle Av, e^k \right\rangle - \left\langle Ae^k, v \right\rangle + \left\langle Av, v \right\rangle =$$

$$= \|e^k\|_A^2 - \left(2\left\langle (L+D)v, v \right\rangle - \left\langle (L+D+L^T)v, v \right\rangle \right) =$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1}\right\rangle =$$

$$= \left\langle A(I - (L+D)^{-1}A)e^k,\ (I - (L+D)^{-1}A)e^k\right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k\right\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

$$\boxed{=} \left\langle Ae^k, e^k\right\rangle - \left\langle Av, e^k\right\rangle - \left\langle Ae^k, v\right\rangle + \left\langle Av, v\right\rangle =$$

$$= \|e^k\|_A^2 - \left(2\left\langle (L+D)v, v\right\rangle - \left\langle (L+D+L^T)v, v\right\rangle\right) =$$

$$= \|e^k\|_A^2 - \left\langle Dv, v\right\rangle \leq \underline{\|e^k\|_A^2 - d\|v\|_2^2\ (2)},\ d = \min_i D_{ii} = \min_i a_{ii}$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1}\right\rangle =$$

$$= \left\langle A(I-(L+D)^{-1}A)e^k,\ (I-(L+D)^{-1}A)e^k\right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k\right\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

$$\boxed{=} \left\langle Ae^k, e^k\right\rangle - \left\langle Av, e^k\right\rangle - \left\langle Ae^k, v\right\rangle + \left\langle Av, v\right\rangle =$$

$$= \|e^k\|_A^2 - \Big(2\left\langle (L+D)v, v\right\rangle - \left\langle (L+D+L^T)v, v\right\rangle\Big) =$$

$$= \|e^k\|_A^2 - \left\langle Dv, v\right\rangle \le \underline{\|e^k\|_A^2 - d\|v\|_2^2\ (2)},\ d = \min_i D_{ii} = \min_i a_{ii}$$

$$\|e^k\|_A^2 = \left\langle Ae^k, e^k\right\rangle = \left\langle (L+D)v, A^{-1}(L+D)v\right\rangle \le$$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \langle Ae^{k+1}, e^{k+1}\rangle =$$

$$= \langle A(I-(L+D)^{-1}A)e^k,\, (I-(L+D)^{-1}A)e^k\rangle =$$

$$= \langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v \ (1)}$

$$\boxed{=} \langle Ae^k, e^k\rangle - \langle Av, e^k\rangle - \langle Ae^k, v\rangle + \langle Av, v\rangle =$$

$$= \|e^k\|_A^2 - \Big(2\langle (L+D)v, v\rangle - \langle (L+D+L^T)v, v\rangle\Big) =$$

$$= \|e^k\|_A^2 - \langle Dv, v\rangle \le \underline{\|e^k\|_A^2 - d\|v\|_2^2 \ (2)},\ d = \min_i D_{ii} = \min_i a_{ii}$$

$$\|e^k\|_A^2 = \langle Ae^k, e^k\rangle = \langle (L+D)v, A^{-1}(L+D)v\rangle \le$$

$\|A^{-1}\|_2 \|(L+D)\|_2^2 \|v\|_2^2 = \lambda_{min}^{-1} \|(L+D)\|_2^2 \|v\|_2^2 \Rightarrow$ из $(2):$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \langle Ae^{k+1}, e^{k+1}\rangle =$$

$$= \langle A(I-(L+D)^{-1}A)e^k,\ (I-(L+D)^{-1}A)e^k\rangle =$$

$$= \langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v\ (1)}$

$$\boxed{=} \langle Ae^k, e^k\rangle - \langle Av, e^k\rangle - \langle Ae^k, v\rangle + \langle Av, v\rangle =$$

$$= \|e^k\|_A^2 - \Big(2\langle (L+D)v, v\rangle - \langle (L+D+L^T)v, v\rangle\Big) =$$

$$= \|e^k\|_A^2 - \langle Dv, v\rangle \le \underline{\|e^k\|_A^2 - d\|v\|_2^2\ (2)},\ d = \min_i D_{ii} = \min_i a_{ii}$$

$$\|e^k\|_A^2 = \langle Ae^k, e^k\rangle = \langle (L+D)v, A^{-1}(L+D)v\rangle \le$$

$\|A^{-1}\|_2\|(L+D)\|_2^2\|v\|_2^2 = \lambda_{min}^{-1}\|(L+D)\|_2^2\|v\|_2^2 \Rightarrow$ из (2) :

$$\underline{\|e^{k+1}\|_A^2 \le \left(1 - \lambda_{min} d\|L+D\|_2^{-2}\right)\|e^k\|_A^2}\ \square$$

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Теорема 7.4 Необходимое и достаточное условие сходимости метода Зейделя

Для сходимости метода Зейделя необходимо и достаточно, чтобы все корни уравнения

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nn} \end{vmatrix} = 0$$

по модулю не превосходили единицы $(|\lambda| \leqslant 1)$.

Алгоритм метода Зейделя

$$(L + D)x^{k+1} = -Ux^k + b$$

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, i-1
            xkp1[i] = xkp1[i] - A[i,j] * xkp1[j]
        for j = i+1, n # матрица U
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

Квадратичные функционалы и линейные системы

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n \times n}$ рассмотрим:
  $f(x) = \dfrac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \dfrac{1}{2}x^T Ax - x^T b$

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n \times n}$ рассмотрим:
  $f(x) = \dfrac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \dfrac{1}{2}x^T Ax - x^T b$
- Пусть z - решение системы $Az = b$. Если $A > 0$, то:

$$E(x) \equiv f(x) - f(z) = 0.5\langle A(x-z), x-z\rangle > 0, \forall x \neq z$$

$\Rightarrow z$ - точка минимума

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n\times n}$ рассмотрим:
$f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \frac{1}{2}x^T Ax - x^T b$
- Пусть z - решение системы $Az = b$. Если $A > 0$, то:

$$E(x) \equiv f(x) - f(z) = 0.5\langle A(x-z), x-z\rangle > 0, \forall x \neq z$$

$\Rightarrow z$ - точка минимума
- Для произвольной A можно ввести *функционал невязки*:

$$R(x) = \|b - Ax\|_2$$

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n \times n}$ рассмотрим:
  $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \frac{1}{2}x^T A x - x^T b$
- Пусть z - решение системы $Az = b$. Если $A > 0$, то:

$$E(x) \equiv f(x) - f(z) = 0.5\langle A(x-z), x-z\rangle > 0, \forall x \neq z$$

  $\Rightarrow z$ - точка минимума
- Для произвольной A можно ввести *функционал невязки*:

$$R(x) = \|b - Ax\|_2$$

- Общая идея: $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k r^k$

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n \times n}$ рассмотрим:
  $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \frac{1}{2}x^T Ax - x^T b$
- Пусть z - решение системы $Az = b$. Если $A > 0$, то:

  $$E(x) \equiv f(x) - f(z) = 0.5\langle A(x - z), x - z\rangle > 0, \forall x \neq z$$

  $\Rightarrow z$ - точка минимума
- Для произвольной A можно ввести *функционал невязки*:

  $$R(x) = \|b - Ax\|_2$$

- Общая идея: $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k r^k$
  $\tau^k = \arg\min_{\tau} F(x^k - \tau r^k),\ F(x) = E(x), R(x)$

Метод наискорейшего спуска

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\big\langle Ax, x\big\rangle - \big\langle b, x\big\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$f(x^{k+1}) = \frac{1}{2}\big\langle A(x^k - \tau r^k), x^k - \tau r^k\big\rangle - \big\langle b, x^k - \tau r^k\big\rangle =$$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$f(x^{k+1}) = \frac{1}{2}\langle A(x^k - \tau r^k), x^k - \tau r^k\rangle - \langle b, x^k - \tau r^k\rangle =$$

$$= \frac{1}{2}\langle Ax^k, x^k\rangle - \underline{\frac{1}{2}\tau\langle Ax^k, r^k\rangle - \frac{1}{2}\tau\langle Ar^k, x^k\rangle} +$$

$$+ \underline{\underline{\frac{1}{2}\tau^2\langle Ar^k, r^k\rangle}} - \langle b, x^k\rangle + \underline{\tau\langle b, r^k\rangle} =$$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$f(x^{k+1}) = \frac{1}{2}\langle A(x^k - \tau r^k), x^k - \tau r^k\rangle - \langle b, x^k - \tau r^k\rangle =$$

$$= \frac{1}{2}\langle Ax^k, x^k\rangle - \underline{\frac{1}{2}\tau\langle Ax^k, r^k\rangle - \frac{1}{2}\tau\langle Ar^k, x^k\rangle} +$$

$$+ \underline{\underline{\frac{1}{2}\tau^2\langle Ar^k, r^k\rangle}} - \langle b, x^k\rangle + \underline{\tau\langle b, r^k\rangle} =$$

$$= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle Ax^k - b, r^k\rangle + \ldots =$$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$
\begin{aligned}
f(x^{k+1}) &= \frac{1}{2}\langle A(x^k - \tau r^k), x^k - \tau r^k\rangle - \langle b, x^k - \tau r^k\rangle = \\
&= \frac{1}{2}\langle Ax^k, x^k\rangle \underline{- \frac{1}{2}\tau\langle Ax^k, r^k\rangle - \frac{1}{2}\tau\langle Ar^k, x^k\rangle} + \\
&+ \underline{\underline{\frac{1}{2}\tau^2\langle Ar^k, r^k\rangle}} - \langle b, x^k\rangle + \underline{\tau\langle b, r^k\rangle} = \\
&= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle Ax^k - b, r^k\rangle + \ldots = \\
&= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle r^k, r^k\rangle + \ldots = \Phi(\tau)
\end{aligned}
$$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$
\begin{aligned}
f(x^{k+1}) &= \frac{1}{2}\langle A(x^k - \tau r^k), x^k - \tau r^k\rangle - \langle b, x^k - \tau r^k\rangle = \\
&= \frac{1}{2}\langle Ax^k, x^k\rangle - \underline{\frac{1}{2}\tau\langle Ax^k, r^k\rangle - \frac{1}{2}\tau\langle Ar^k, x^k\rangle} + \\
&+ \underline{\underline{\frac{1}{2}\tau^2\langle Ar^k, r^k\rangle}} - \langle b, x^k\rangle + \underline{\tau\langle b, r^k\rangle} = \\
&= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle Ax^k - b, r^k\rangle + \ldots = \\
&= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle r^k, r^k\rangle + \ldots = \Phi(\tau)
\end{aligned}
$$

$$
\Phi'(\tau) = \tau\langle Ar^k, r^k\rangle - \langle r^k, r^k\rangle = 0 \Rightarrow \tau^k = \frac{\langle r^k, r^k\rangle}{\langle Ar^k, r^k\rangle}
$$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$
- Метод Ричардсона: $P = \tau^{-1} I$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$
- Метод Ричардсона: $P = \tau^{-1} I$
- Метод Якоби: $P = D$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$
- Метод Ричардсона: $P = \tau^{-1} I$
- Метод Якоби: $P = D$
- Метод Зейделя: $P = (L + D)$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$
- Метод Ричардсона: $P = \tau^{-1} I$
- Метод Якоби: $P = D$
- Метод Зейделя: $P = (L + D)$
- Можно «комбинировать» предобуславливатели, например, применить метод Ричардсона к $D^{-1}Ax = D^{-1}b$

Предобуславливание

- $P(x^{k+1} - x^k) = b - Ax^k$
- Метод Ричардсона: $P = \tau^{-1} I$
- Метод Якоби: $P = D$
- Метод Зейделя: $P = (L + D)$
- Можно «комбинировать» предобуславливатели, например, применить метод Ричардсона к $D^{-1}Ax = D^{-1}b$
- Простые методы (МПИ) используются в качестве предобуславливателей для более сложных методов

Заключение

- Сравнение прямых и итерационных методов
- Метод Ричардсона
- Методы Якоби и Зейделя
- Методы, основанные на минимизации функционала