Численные методы решения систем линейных алгебраических уравнений

Итерационные методы

Курц В.В.

Санкт-Петербургский Политехнический университет Петра Великого

5 ноября 2021 г.

Содержание

Итерационные методы

Итерационные методы - альтернатива прямым методам, в частности, когда размерность системы велика.

Постановка задачи: найти решение СЛАУ

$$Ax = b \tag{1}$$

с точностью ϵ, т.е. $\left\|x^{(k)} - x^*\right\| < \epsilon$.
Строится последовательность векторов $\{x^{(k)}\}$

$$x^{(0)} \in \mathbb{R}^n, x^{(k)} = \Phi_k(A, b, x^{(k-1)}, \dots, x^{(k-r)}), \tag{2}$$

предел которой - решение задачи (1)

$$\lim_{k \to \infty} x^{(k)} = x^*. \tag{3}$$

Классификационные признаки

1. $r = 1$ – одношаговый метод, $r > 1$ – многошаговый метод.
2. Φ_k - линейная функция относительно $x^{(i)}$ $\Rightarrow$ метод называется линейным, в противном случае – нелинейным.
3. Φ_k зависит от k $\Rightarrow$ метод называется нестационарным, в противном случае – стационарным.

Будем рассматривать линейные одношаговые методы.

Канонический вид линейных одношаговых итерационных методов

$$B_k \frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \tag{4}$$

где α_k, B_k - канонические параметры, в общем случае зависят от k.

Линейные одношаговые методы

1. $\alpha_k = \alpha, B_k = B$ - стационарные методы, в противном случае - нестационарные.
2. метод называется явным, если известно B_k^{-1}, например, $B_k = E$

   $$x^{(k+1)} = x^{(k)} - \alpha_k (Ax^{(k)} - b) \tag{5}$$

   или B_k - диагональная матрица

   $$B_k x^{(k+1)} = B_k x^{(k)} - \alpha_k (Ax^{(k)} - b), \tag{6}$$

   где B_k называют предобуславливателем (с его помощью можно уменьшить количество итераций).

1. Как построить последовательность $\{x^{(k)}\}$?
2. При каких условиях последовательность $\{x^{(k)}\}$ будет сходиться?
3. При каких условиях предел последовательности $\{x^{(k)}\}$ есть решение исходной задачи (1)?
4. Когда остановить итерационный процесс, чтобы получить решение с заданной точностью?

Содержание

Метод простых итераций (МПИ)

$$Ax = b \Leftrightarrow x = Cx + g \tag{7}$$

$$x^{(k+1)} = Cx^{(k)} + g, k = 0, 1, \dots \tag{8}$$

Теорема о сходимости МПИ (достаточные условия)

Если $\|C\| < 1$, то существует единственное решение (7), итерационная последовательность (8) сходится к нему при произвольном $x^{(0)} \in \mathbb{R}^n$ и справедливы оценки

$$\left\|x^* - x^{(k)}\right\| \leq \frac{\|C\|}{1 - \|C\|} \left\|x^{(k)} - x^{(k-1)}\right\| \tag{9}$$

$$\left\|x^* - x^{(k)}\right\| \leq \frac{\|C\|^k}{1 - \|C\|} \left\|x^{(1)} - x^{(0)}\right\| \tag{10}$$

Теорема о сходимости МПИ. Доказательство

$$(8) \Rightarrow x^{(k+1)} - x^{(k)} = C(x^{(k)} - x^{(k-1)}) \Rightarrow$$

$$\left\|x^{(k+1)} - x^{(k)}\right\| \le \|C\| \left\|x^{(k)} - x^{(k-1)}\right\| \tag{11}$$

$$\begin{aligned}
\left\|x^{(k+m)} - x^{(k)}\right\| &= \left\|x^{(k+m)} - x^{(k+m-1)} + x^{(k+m-1)} - \ldots - x^{(k+1)} + x^{(k+1)} - x^{(k)}\right\| \\
&\le \left\|x^{(k+m)} - x^{(k+m-1)}\right\| + \ldots + \left\|x^{(k+2)} - x^{(k+1)}\right\| + \left\|x^{(k+1)} - x^{(k)}\right\| \\
&\le \|C\|^m \left\|x^{(k)} - x^{(k-1)}\right\| + \ldots + \|C\|^2 \left\|x^{(k)} - x^{(k-1)}\right\| + \|C\| \left\|x^{(k)} - x^{(k-1)}\right\| \\
&= \frac{\|C\| \left(1 - \|C\|^m\right)}{1 - \|C\|} \left\|x^{(k)} - x^{(k-1)}\right\| \\
&\le \frac{\|C\|^k}{1 - \|C\|} \left(1 - \|C\|^m\right) \left\|x^{(1)} - x^{(0)}\right\|.
\end{aligned}$$

Теорема о сходимости МПИ. Доказательство

Зафиксируем m, а $k \to \infty \Rightarrow$ последовательность $x^{(k)}$ - фундаментальная последовательность $\Rightarrow$ имеет предел x^* ($\mathbb{R}^n$ - полное пространство)

$$x^* = Cx^* + g. \tag{12}$$

Доказать единственность x^* самостоятельно (от противного).
Справедливость (9) и (10) следует из

$$\left\| x^{(k+m)} - x^{(k)} \right\| \leq \frac{\|C\| \left(1 - \|C\|^m\right)}{1 - \|C\|} \left\| x^{(k)} - x^{(k-1)} \right\| \leq \frac{\|C\|^k \left(1 - \|C\|^m\right)}{1 - \|C\|} \left\| x^{(1)} - x^{(0)} \right\|, \tag{13}$$

если зафиксировать k, а $m \to \infty$.

Теорема о сходимости МПИ. Замечания

1. Априорная оценка (10) грубее апостериорной (9).
2. Апостериорная оценка (9) используется для остановки итерационного процесса

$$\left\| x^{(k)} - x^{(k-1)} \right\| < \frac{1 - \|C\|}{\|C\|}\epsilon. \tag{14}$$

3. Если $\|C\| \leq \frac{1}{2}$, то в качестве критерия остановки можно использовать

$$\left\| x^{(k)} - x^{(k-1)} \right\| < \epsilon. \tag{15}$$

4. Возможность привести исходную задачу к виду $x = Cx + g$, $\|C\| < 1$ связано с существованием и единственностью решения исходной задачи.

Приведение СЛАУ к виду, удобному для итераций

$$Ax = b \Leftrightarrow 0 = -\alpha(Ax - b), \alpha \neq 0 \Leftrightarrow Bx = Bx - \alpha(Ax - b), \det(B) \neq 0$$

$$x = x - \alpha B^{-1}(Ax - b) \Leftrightarrow x = \underbrace{(E - \alpha B^{-1}A)}_{C} x + \alpha B^{-1} b \tag{16}$$

Каноническая форма метода

$$B\frac{x^{(k+1)} - x^{(k)}}{\alpha} + Ax^{(k)} = b \tag{17}$$

Приведение СЛАУ к виду, удобному для итераций

1. $B = E$

$$x^{(k+1)} = x^{(k)} - \alpha(Ax^{(k)} - b) = \underbrace{(E - \alpha A)}_{C} x^{(k)} + \underbrace{\alpha b}_{g}. \tag{18}$$

Пусть A - симметричная положительно определенная матрица.
λ_{max} - максимальное собственное число A.
Условие $\|C\|_2 < 1$ будет выполнено, если $\alpha \in (0, 2/\lambda_{max})$.
Можно взять $\alpha < 2/\|A\|$.

Приведение СЛАУ к виду, удобному для итераций

2. $\alpha = 1, B = D = diag(A)$

$$x^{(k+1)} = x^{(k)} - D^{-1}(Ax^{(k)} - b) = \underbrace{(E - D^{-1}A)}_{C}x^{(k)} + \underbrace{D^{-1}b}_{g}. \tag{19}$$

Если $D = D_A$, то получим метод Якоби.

$$\begin{cases} 6x_1 - x_2 + 3x_3 = 7 \\ -x_1 + 5x_2 + 2x_3 = 6 \\ x_1 - 2x_2 + 4x_3 = 3 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{6}(x_2 - 3x_3 + 7) \\ x_2 = \frac{1}{5}(x_1 - 2x_3 + 6) \\ x_3 = \frac{1}{4}(-x_1 + 2x_2 + 3) \end{cases} \quad C = \begin{bmatrix} 0 & \frac{1}{6} & -\frac{1}{2} \\ \frac{1}{5} & 0 & -\frac{2}{5} \\ -\frac{1}{4} & \frac{1}{2} & 0 \end{bmatrix}$$

Метод Якоби (Jacobi iterations)

$$C = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & -\frac{a_{n3}}{a_{nn}} & \dots & 0 \end{bmatrix} \tag{20}$$

Если $\|C\| = \|C\|_\infty$, то достаточное условие сходимости МПИ примет вид

$$\|C\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |c_{ij}| < 1, \tag{21}$$

что эквивалентно условию диагонального преобладания

$$\sum_{j=1, j \ne i}^{n} |a_{ij}| < |a_{ii}|, i = 1, \dots, n. \tag{22}$$

Метод Якоби. Пример

$$\begin{cases} 4x_1 - x_2 + x_3 = 4 \\ -x_1 + 3x_2 - x_3 = 1 \\ x_1 - x_2 + 5x_3 = 5 \end{cases}$$

Точное решение: $x^* = (1, 1, 1)^\top$

$$x_1^{(k+1)} = \frac{1}{4}(4 + x_2^{(k)} - x_3^{(k)})$$
$$x_2^{(k+1)} = \frac{1}{3}(1 + x_1^{(k)} + x_3^{(k)})$$
$$x_3^{(k+1)} = \frac{1}{5}(5 - x_1^{(k)} + x_2^{(k)})$$

Начальное приближение: $x^{(0)} = (1, \frac{1}{3}, 1)^\top$
Первые 3 итерации:

$$\begin{aligned} x^{(1)} &= (0.833 \quad 1.000 \quad 0.867)^\top \\ x^{(2)} &= (1.033 \quad 0.900 \quad 1.033)^\top \\ x^{(3)} &= (0.967 \quad 1.022 \quad 0.973)^\top \\ &\ldots\ldots \\ x^* &= (1 \qquad 1 \qquad 1)^\top \end{aligned}$$

Необходимые и достаточные условия сходимости МПИ

Лемма

Если все собственные числа $|\lambda_C| < 1$, то $\sum\limits_{k=0}^{\infty} C^k = (E - C)^{-1}$.

Доказательство

$S_k = E + C + \ldots + C^k$.

$$S_k(E - C) = E - C + C - C^2 + \ldots + C^k - C^{k+1} = E - C^{k+1}. \tag{23}$$

Все собственные числа $|\lambda_C| < 1 \Leftrightarrow \lim\limits_{k\to\infty} C^{k+1} = 0$. $\Rightarrow$ левая часть (23) имеет предел $\Rightarrow S_k$ имеет предел.

Пусть $\lim\limits_{k\to\infty} S_k = S \Rightarrow S(E - C) = E \Rightarrow S = (E - C)^{-1}$.

Необходимые и достаточные условия сходимости МПИ

Теорема

Последовательность $x^{(k+1)} = Cx^{(k)} + g$ сходится $\Leftrightarrow$ все собственные числа $|\lambda_C| < 1$.

Доказательство

$\Rightarrow$

Пусть $x^* = \lim\limits_{k\to\infty} x^{(k)}$.

$$x^{(k)} - x^* = C(x^{(k-1)} - x^*) = C^2(x^{(k-2)} - x^*) = \ldots = C^k(x^{(0)} - x^*). \qquad (24)$$

$x^{(0)}$ - любой вектор. Пусть $x^{(0)} - x^* = e^{(p)}$. Тогда в правой части (24) останется p-й столбец C^k, который стремится к 0 $\Rightarrow C^k \xrightarrow[k\to\infty]{} 0 \Leftrightarrow \forall\lambda_C |\lambda_C| < 1$.

$\Leftarrow$

$x^{(k)} = Cx^{(k-1)} + g = C(Cx^{(k-2)} + g) + g = C^2x^{(k-2)} + (E + C)g =$
$C^kx^{(0)} + (E + C + \ldots + C^k)g \xrightarrow[k\to\infty]{} (E - C)^{-1}g.$

Метод итераций Зейделя (Gauss-Seidel iterations)

МПИ: $x^{(k+1)} = Cx^{(k)} + g$

$$\begin{cases} x_1^{(k+1)} = c_{11}x_1^{(k)} + c_{12}x_2^{(k)} + \ldots + c_{1n}x_n^{(k)} + g_1 \\ x_2^{(k+1)} = c_{21}x_1^{(k)} + c_{22}x_2^{(k)} + \ldots + c_{2n}x_n^{(k)} + g_2 \\ \ldots = \ldots \\ x_n^{(k+1)} = c_{n1}x_1^{(k)} + c_{n2}x_2^{(k)} + \ldots + c_{nn}x_n^{(k)} + g_n \end{cases} \tag{25}$$

Метод Зейделя сводится к тому, что при вычислении компонент $(k+1)$-го вектора компоненты k-го заменить уже вычисленными $(k+1)$-ми элементами

$$\begin{cases} x_1^{(k+1)} = c_{11}x_1^{(k)} + c_{12}x_2^{(k)} + \ldots + c_{1n}x_n^{(k)} + g_1 \\ x_2^{(k+1)} = c_{21}x_1^{(k+1)} + c_{22}x_2^{(k)} + \ldots + c_{2n}x_n^{(k)} + g_2 \\ \ldots = \ldots \\ x_n^{(k+1)} = c_{n1}x_1^{(k+1)} + c_{n2}x_2^{(k+1)} + \ldots + c_{nn}x_n^{(k)} + g_n \end{cases} \tag{26}$$

Схема Зейделя-Якоби. Пример

$$\begin{cases} 4x_1 - x_2 + x_3 = 4 \\ -x_1 + 3x_2 - x_3 = 1 \\ x_1 - x_2 + 5x_3 = 5 \end{cases}$$

Точное решение: $x^* = (1, 1, 1)^\top$

$$x_1^{(k+1)} = \frac{1}{4}(4 + x_2^{(k)} - x_3^{(k)})$$

$$x_2^{(k+1)} = \frac{1}{3}(1 + x_1^{(k+1)} + x_3^{(k)})$$

$$x_3^{(k+1)} = \frac{1}{5}(5 - x_1^{(k+1)} + x_2^{(k+1)})$$

Начальное приближение: $x^{(0)} = (1, \frac{1}{3}, 1)^\top$

Первые 3 итерации:

$$\begin{aligned} x^{(1)} &= (0.833 \quad 0.944 \quad 1.022)^\top \\ x^{(2)} &= (0.981 \quad 1.001 \quad 1.004)^\top \\ x^{(3)} &= (0.999 \quad 1.001 \quad 1.000)^\top \\ &\ldots\ldots \\ x^* &= (1 \qquad 1 \qquad 1)^\top \end{aligned}$$

Метод итераций Зейделя

$$C = \underline{C} + \overline{\overline{C}}$$

$$\underline{C} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ c_{21} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{n,n-1} & 0 \end{bmatrix}, \overline{\overline{C}} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1,n-1} & c_{1n} \\ 0 & c_{22} & \dots & c_{2,n-1} & c_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & c_{nn} \end{bmatrix} \tag{27}$$

Метод Зейделя:

$$x^{(k+1)} = \underline{C}x^{(k+1)} + \overline{\overline{C}}x^{(k)} + g. \tag{28}$$

$$(E - \underline{C})x^{(k+1)} = \overline{\overline{C}}x^{(k)} + g$$
$$\det(E - \underline{C}) = 1 \Rightarrow \exists (E - \underline{C})^{-1}$$

$$x^{(k+1)} = \underbrace{(E - \underline{C})^{-1}\overline{\overline{C}}}_{B}x^{(k)} + \underbrace{(E - \underline{C})^{-1}g}_{f}. \tag{29}$$

Необходимые и достаточные условия сходимости метода Зейделя

Метод Зейделя (29) сходится $\Leftrightarrow$ все собственные числа $|\lambda_B| < 1$, где $B = (E - \underline{C})^{-1}\overline{\overline{C}}$.

$$\begin{aligned}
\chi_B(\lambda) &= \det(B - \lambda E) \\
&= -\det(\lambda E - (E - \underline{C})^{-1}\overline{\overline{C}}) \\
&= -\underbrace{\det((E - \underline{C})^{-1})}_{1} \det(\lambda(E - \underline{C}) - \overline{\overline{C}}) \\
&= -\det(\lambda E - \lambda\underline{C} - \overline{\overline{C}}) \\
&= \det(\lambda\underline{C} + \overline{\overline{C}} - \lambda E).
\end{aligned} \tag{30}$$

$$\det(\lambda\underline{C} + \overline{\overline{C}} - \lambda E) = \begin{vmatrix} c_{11} - \lambda & c_{12} & \dots & c_{1n} \\ \lambda c_{21} & c_{22} - \lambda & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ \lambda c_{n1} & \lambda c_{n2} & \dots & c_{nn} - \lambda \end{vmatrix} \tag{31}$$

Пример

Пусть исходная СЛАУ приведена к виду удобному для итераций

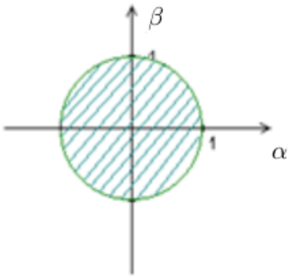
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}}_{C} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Метод простых итераций:

$\det(C - \lambda E) = (a - \lambda)^2 + \beta^2 = \lambda^2 - 2\alpha\lambda + \alpha^2 + \beta^2.$

$\lambda_{1,2} = \alpha \pm i\beta$

$|\lambda| < 1 \Leftrightarrow \sqrt{\alpha^2 + \beta^2} < 1.$



Пример (продолжение)

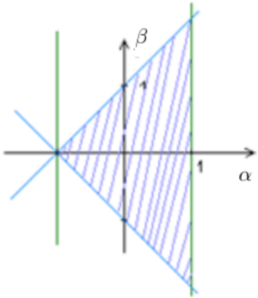
Метод итераций Зейделя:

$$\begin{vmatrix} \alpha-\lambda & -\beta \\ \lambda\beta & \alpha-\lambda \end{vmatrix} = (\alpha-\lambda)^2+\lambda\beta^2 = \lambda^2+(-2\alpha+\beta^2)\lambda+\alpha^2.$$

Теорема

Корни полинома x^2+ax+b по модулю меньше единицы $\Leftrightarrow$ $|a|< 1+b$ и $|b|< 1$.

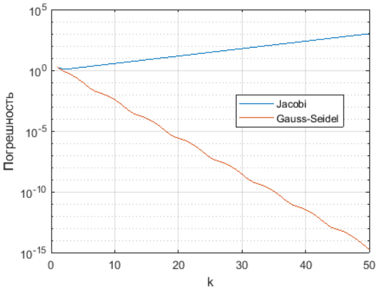
$$\begin{cases} |\beta^2-2\alpha| < 1+\alpha^2 \\ \alpha^2 < 1 \end{cases} \Rightarrow \begin{cases} |\beta| < |1+\alpha| \\ |\alpha| < 1 \end{cases}$$



Пример. Сходимость МПИ (схема Якоби) и метода Зейделя

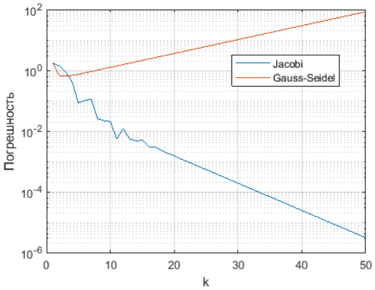
$$A_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$$

$$\max|\lambda_{C_1}| = 1.15 > 1$$
$$\max|\lambda_{B_1}| = 0.50 < 1$$



$$A_2 = \begin{bmatrix} -3 & 3 & -6 \\ -4 & 7 & -8 \\ 5 & 7 & -9 \end{bmatrix}$$

$$\max|\lambda_{C_2}| = 0.81 < 1$$
$$\max|\lambda_{B_2}| = 1.11 > 1$$



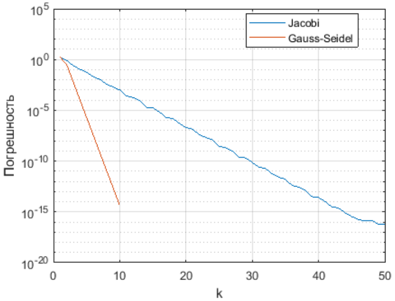
Пример. Сходимость МПИ (схема Якоби) и метода Зейделя

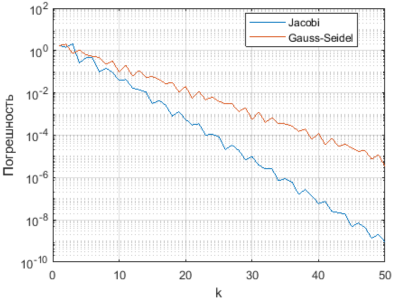
$$A_3 = \begin{bmatrix} 4 & 1 & 1 \\ 2 & -9 & 0 \\ 0 & -8 & -6 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 7 & 6 & 9 \\ 4 & 5 & -4 \\ -7 & -3 & 8 \end{bmatrix}$$

$\max|\lambda_{C_3}| = 0.44$
$\max|\lambda_{B_3}| = 0.02$

$\max|\lambda_{C_4}| = 0.64$
$\max|\lambda_{B_4}| = 0.77$





Достаточное условие сходимости метода Зейделя

$\|C\|_{1,2,\infty} < 1 \Rightarrow$ метод Зейделя (29) сходится.

Ограничимся случаем бесконечной нормы: $\|C\|_\infty = \max\limits_i \sum\limits_{j=1}^{n} |c_{ij}|$, $\|x\|_\infty = \max\limits_i |x_i|$.

Доказательство

$\|C\|_\infty < 1 \Rightarrow \exists x^*$:

$$x^* = Cx^* + g = \underline{C}x^* + \overline{\overline{C}}x^* + g. \tag{32}$$

$$x^{(k)} - x^* = \underline{C}(x^{(k)} - x^*) + \overline{\overline{C}}(x^{(k-1)} - x^*). \tag{33}$$

$$x_i^{(k)} - x_i^* = \sum_{j=1}^{i-1} c_{ij}(x_j^{(k)} - x_j^*) + \sum_{j=i}^{n} c_{ij}(x_j^{(k-1)} - x_j^*).$$

$$|x_i^{(k)} - x_i^*| \leq \sum_{j=1}^{i-1} |c_{ij}||(x_j^{(k)} - x_j^*)| + \sum_{j=i}^{n} |c_{ij}||(x_j^{(k-1)} - x_j^*)|. \tag{34}$$

Достаточное условие сходимости метода Зейделя

Введем обозначения

$$\alpha_i = \sum_{j=1}^{i-1} |c_{ij}|, \beta_i = \sum_{j=i}^{n} |c_{ij}|. \tag{35}$$

(34), (35) $\Rightarrow$

$$|x_i^{(k)} - x_i^*| \leq \alpha_i \left\| x^{(k)} - x^* \right\|_\infty + \beta_i \left\| x^{(k-1)} - x^* \right\|_\infty, \forall i. \tag{36}$$

$\exists m : |x_m^{(k)} - x_m^*| = \left\| x^{(k)} - x^* \right\|_\infty \Rightarrow$

$$\left\| x^{(k)} - x^* \right\|_\infty \leq \alpha_m \left\| x^{(k)} - x^* \right\|_\infty + \beta_m \left\| x^{(k-1)} - x^* \right\|_\infty. \tag{37}$$

$$\left\| x^{(k)} - x^* \right\|_\infty \leq \frac{\beta_m}{1 - \alpha_m} \left\| x^{(k-1)} - x^* \right\|_\infty \leq \mu \left\| x^{(k-1)} - x^* \right\|_\infty, \tag{38}$$

где $\mu = \max\limits_i \frac{\beta_i}{1-\alpha_i}$.

Достаточное условие сходимости метода Зейделя

$$\forall i \; \alpha_i + \beta_i = \sum_{j=1}^{n} |c_{ij}| \leq \max_i \sum_{j=1}^{n} |c_{ij}| = \|C\|_\infty < 1.$$

$\beta_i \leq \|C\|_\infty - \alpha_i < \|C\|_\infty - \alpha_i \|C\|_\infty \Rightarrow \frac{\beta_i}{1-\alpha_i} < \|C\|_\infty < 1 \Rightarrow \mu < \|C\|_\infty < 1.$

(38) $\Rightarrow$

$$\begin{aligned} \left\| x^{(k)} - x^* \right\|_\infty &\leq \mu \left\| x^{(k-1)} - x^* \right\|_\infty \\ &\leq \mu^2 \left\| x^{(k-2)} - x^* \right\|_\infty \\ &\leq \dots \\ &\leq \mu^k \left\| x^{(0)} - x^* \right\|_\infty \xrightarrow[k\to\infty]{} 0. \end{aligned} \tag{39}$$

$$\left\|x^{(k)} - x^*\right\|_\infty \leq \mu \left\|x^{(k-1)} - x^* \pm x^{(k)}\right\|_\infty \leq \mu \left\|x^{(k)} - x^{(k-1)}\right\|_\infty + \mu \left\|x^{(k)} - x^*\right\|_\infty$$

$$\left\| x^{(k)} - x^* \right\|_\infty \leq \frac{\mu}{1-\mu} \left\| x^{(k)} - x^{(k-1)} \right\|_\infty \leq \frac{\|C\|_\infty}{1 - \|C\|_\infty} \left\| x^{(k)} - x^{(k-1)} \right\|_\infty . \tag{40}$$

Метод Зейделя. Замечания

Для метода Зейделя в качестве критерия остановки можно использовать ту же оценку, что и в случае МПИ

$$\left\| x^{(k)} - x^{(k-1)} \right\|_\infty < \epsilon \frac{1 - \|C\|_\infty}{\|C\|_\infty}. \tag{41}$$

Однако условие

$$\left\| x^{(k)} - x^{(k-1)} \right\|_\infty < \epsilon \frac{1 - \mu}{\mu}. \tag{42}$$

является более точным.

Метод релаксаций (Successive over-relaxation method, SOR)

$x = Cx + g$

$x^{(k+1)} = Cx^{(k)} + g \pm x^{(k)} = x^{(k)} + \underbrace{\left(Cx^{(k)} + g - x^{(k)}\right)} = x^{(k)} + r^{(k)}$

Метод релаксаций

$$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)} \tag{43}$$

$\alpha_k = \alpha$ - стационарный метод релаксаций.

$x^{(k+1)} = \underline{C}x^{(k+1)} + \overline{\overline{C}}x^{(k)} + g \pm x^{(k)} = x^{(k)} + \Delta^{(k)}$

Метод релаксаций для схемы Зейделя

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \beta\Delta^{(k)} \\ &= x^{(k)} + \beta(\underline{C}x^{(k+1)} + \overline{\overline{C}}x^{(k)} + g - x^{(k)}) \\ &= (1-\beta)x^{(k)} + \beta(\underline{C}x^{(k+1)} + \overline{\overline{C}}x^{(k)} + g) \end{aligned} \tag{44}$$

Метод релаксаций для схемы Зейделя-Якоби

$Ax = b, A = \underline{A} + D_A + \overline{A}$

$$x = -D_A^{-1}(\underline{A}x + \overline{A}x - b) = \underbrace{(-D_A^{-1}\underline{A} - D_A^{-1}\overline{A})}_{C}x + \underbrace{D_A^{-1}b}_{g} \tag{45}$$

$\underline{C} = -D_A^{-1}\underline{A}, \overline{\overline{C}} = -D_A^{-1}\overline{A}$
(44) $\Rightarrow$

$$x^{(k+1)} = (1-\beta)x^{(k)} + \beta(-D_A^{-1}\underline{A}x^{(k+1)} - D_A^{-1}\overline{A}x^{(k)} + D_A^{-1}b) \tag{46}$$

$\beta = 1$ - схема Зейделя-Якоби.

Метод релаксаций

$$x^{(k+1)} = (1-\beta)x^{(k)} + \beta\underbrace{\left(-D_A^{-1}(\underline{A}x^{(k+1)} + \overline{A}x^{(k)}) + D_A^{-1}b\right)}_{\text{метод Зейделя}} \tag{47}$$

$$\begin{cases} x_1^{(k+1)} = (1-\beta)x_1^{(k)} - \beta\dfrac{a_{12}}{a_{11}}x_2^{(k)} - \ldots - \beta\dfrac{a_{1n}}{a_{11}}x_n^{(k)} + \beta\dfrac{b_1}{a_{11}} \\ x_2^{(k+1)} = -\beta\dfrac{a_{21}}{a_{22}}x_1^{(k+1)} + (1-\beta)x_2^{(k)} - \ldots - \beta\dfrac{a_{2n}}{a_{22}}x_n^{(k)} + \beta\dfrac{b_2}{a_{22}} \\ \quad\ldots = \ldots \\ x_n^{(k+1)} = -\beta\dfrac{a_{n1}}{a_{nn}}x_1^{(k+1)} - \beta\dfrac{a_{n2}}{a_{nn}}x_2^{(k+1)} - \ldots + (1-\beta)x_n^{(k)} + \beta\dfrac{b_n}{a_{nn}} \end{cases} \tag{48}$$

Пример. Метод релаксаций для схемы Зейделя-Якоби

$$\begin{cases} 4x_1 - x_2 + x_3 = 4 \\ -x_1 + 3x_2 - x_3 = 1 \\ x_1 - x_2 + 5x_3 = 5 \end{cases}$$

Точное решение: $x^* = (1, 1, 1)^\top$

Пусть $\beta = \frac{4}{3}$

$$x_1^{(k+1)} = -\frac{1}{3}x_1^{(k)} + \frac{4}{3} \cdot \frac{1}{4}(4 + x_2^{(k)} - x_3^{(k)})$$

$$x_2^{(k+1)} = -\frac{1}{3}x_2^{(k)} + \frac{4}{3} \cdot \frac{1}{3}(1 + x_1^{(k+1)} + x_3^{(k)})$$

$$x_3^{(k+1)} = -\frac{1}{3}x_3^{(k)} + \frac{4}{3} \cdot \frac{1}{5}(5 - x_1^{(k+1)} + x_2^{(k+1)})$$

Начальное приближение: $x^{(0)} = (1, \frac{1}{3}, 1)^\top$

Первые 3 итерации:

$$x^{(1)} = (0.778 \quad 1.123 \quad 1.092)^\top$$

$$x^{(2)} = (1.085 \quad 1.037 \quad 0.957)^\top$$

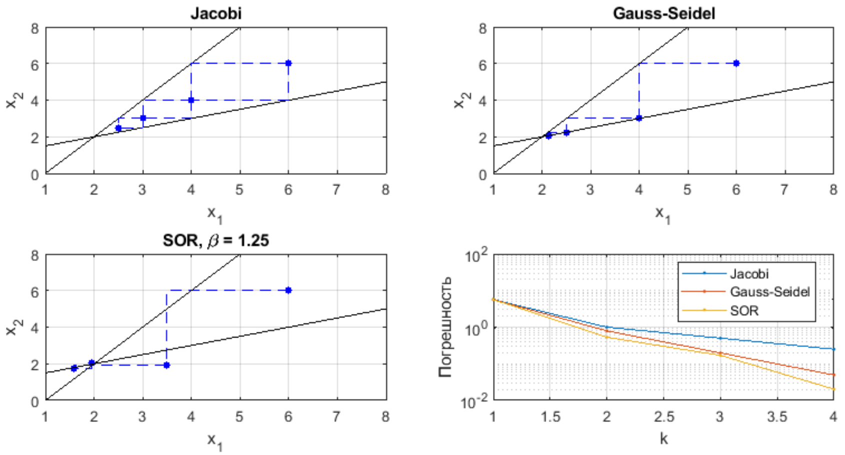
$$x^{(3)} = (0.999 \quad 0.968 \quad 1.006)^\top$$

$$\ldots\ldots$$

$$x^* = (1 \quad 1 \quad 1)^\top$$

Геометрическая интерпретация методов

$$\begin{cases} -2x_1 + x_2 = -2 \\ x_1 - 2x_2 = -2 \end{cases}$$



Метод релаксаций. Каноническая форма

$$x^{(k)} = x^{(k-1)} - \beta D_A^{-1}(\underline{A}x^{(k)} + D_A x^{(k-1)} + \overline{A}x^{(k-1)} - b)$$

$$D_A x^{(k)} - D_A x^{(k-1)} + \beta(\underline{A}x^{(k)} + (D_A + \overline{A})x^{(k-1)}) = \beta b$$

$$(D_A + \beta\underline{A})x^{(k)} - (D_A + \beta\underline{A})x^{(k-1)} + \beta\underline{A}x^{(k-1)} + \beta(D_A + \overline{A})x^{(k-1)} = \beta b$$

Каноническая форма метода релаксаций

$$(D_A + \beta\underline{A})\frac{x^{(k)} - x^{(k-1)}}{\beta} + Ax^{(k-1)} = b \qquad (49)$$

Сходимость стационарных итерационных процессов

Линейный одношаговый стационарный итерационный метод

$$B\frac{x^{(k)} - x^{(k-1)}}{\alpha} + Ax^{(k-1)} = b \tag{50}$$

Будем рассматривать СЛАУ с симметричной ($A = A^\top$) и положительно определенной матрицей A ($A > 0$).

Теорема

Пусть $A = A^\top$, $A > 0$, $\alpha > 0$ и $Q = B - \frac{\alpha}{2}A > 0$. Тогда последовательность $x^{(k)}$ сходится к точному решению $Ax = b$ при любом $x^{(0)} \in \mathbb{R}^n$.

Определим энергетическое пространство H_A, в котором $\|x\|_{H_A}^2 = (Ax, x)$.

Доказательство

$\det(A) \neq 0 \Rightarrow \exists x^* : x^* = A^{-1}b$. Пусть $e^{(k)} = x^{(k)} - x^*$. Тогда (50) примет вид

$$B\frac{e^{(k)} - e^{(k-1)}}{\alpha} + Ae^{(k-1)} = 0. \tag{51}$$

Пусть $r^{(k)} = e^{(k)} - e^{(k-1)}$. Умножим (51) скалярно на $r^{(k)}$

$$\frac{1}{\alpha}(Br^{(k)}, r^{(k)}) + (Ae^{(k-1)}, r^{(k)}) = 0. \tag{52}$$

$$\begin{aligned}
(Ae^{(k-1)}, r^{(k)}) &= \left(A\frac{e^{(k)} + e^{(k-1)}}{2} - A\frac{e^{(k)} - e^{(k-1)}}{2}, r^{(k)}\right) \\
&= \frac{1}{2}\left[(A(e^{(k)} + e^{(k-1)}), r^{(k)}) - (Ar^{(k)}, r^{(k)})\right] \\
&= \frac{1}{2}\Big[(Ae^{(k)}, e^{(k)}) + {\color{red}(Ae^{(k-1)}, e^{(k)})} - {\color{red}(Ae^{(k)}, e^{(k-1)})} - (Ae^{(k-1)}, e^{(k-1)}) \\
&\quad - (Ar^{(k)}, r^{(k)})\Big]
\end{aligned}$$

Перепишем (52)

$$
\begin{aligned}
&\frac{1}{\alpha}(Br^{(k)}, r^{(k)}) + \frac{1}{2}\left[(Ae^{(k)}, e^{(k)}) - (Ae^{(k-1)}, e^{(k-1)}) - (Ar^{(k)}, r^{(k)})\right] \\
&\quad = \frac{1}{\alpha}\left(\left(B - \frac{\alpha}{2}A\right) r^{(k)}, r^{(k)}\right) + \frac{1}{2}\left\|e^{(k)}\right\|_{H_A}^2 - \frac{1}{2}\left\|e^{(k-1)}\right\|_{H_A}^2 \\
&\quad = 0
\end{aligned}
\tag{54}
$$

$Q > 0 \Rightarrow \exists \gamma \neq 0 : (Qx, x) \geq \gamma \left\|x\right\|_2^2.$

$$
\frac{1}{\alpha}\gamma\left\|r^{(k)}\right\|_2^2 + \frac{1}{2}\left\|e^{(k)}\right\|_{H_A}^2 \leq \frac{1}{2}\left\|e^{(k-1)}\right\|_{H_A}^2 \tag{55}
$$

Последовательность $\left\|e^{(k)}\right\|_{H_A}$ является невозрастающей и ограниченной снизу нулем $\Rightarrow$ имеет предел.

$\left\|r^{(k)}\right\|_2 \xrightarrow[k\to\infty]{} 0.$

$$(51) \Rightarrow Ae^{(k-1)} = -B\frac{r^{(k)}}{\alpha} \Rightarrow e^{(k-1)} = -A^{-1}B\frac{r^{(k)}}{\alpha}$$

$$\begin{aligned} \left\|e^{(k-1)}\right\|_{H_A}^2 &= (Ae^{(k-1)}, e^{(k-1)}) \\ &= \frac{1}{\alpha^2}(Br^{(k)}, A^{-1}Br^{(k)}) \\ &\le \frac{1}{\alpha^2}\left\|A^{-1}\right\|_2 \left\|Br^{(k)}\right\|_2^2 \\ &\le \frac{1}{\alpha^2}\left\|A^{-1}\right\|_2 \|B\|_2^2 \left\|r^{(k)}\right\|_2^2 \xrightarrow[k\to\infty]{} 0 \end{aligned} \tag{56}$$

Замечание

$r^{(k)} = e^{(k)} - e^{(k-1)} = x^{(k)} - x^{(k-1)} \Rightarrow (56)$ можно использовать в качестве оценки погрешности.

Замечание

$Ax = b \Leftrightarrow A^\top Ax = A^\top b$ - ухудшается обусловленность системы.

Следствие. Метод релаксаций

Каноническая форма метода релаксаций

$$(D_A + \beta \underline{A})\frac{x^{(k)} - x^{(k-1)}}{\beta} + Ax^{(k-1)} = b \tag{57}$$

$A = A^\top, A > 0 \Rightarrow D_A > 0$

$Q = B - \frac{\alpha}{2}A = (D_A + \beta\underline{A}) - \frac{\beta}{2}(\underline{A} + D_A + \overline{A}) = (1 - \frac{\beta}{2})D_A + \frac{\beta}{2}(\underline{A} - \overline{A})$

$$\begin{aligned}(Qx, x) &= \left(1 - \frac{\beta}{2}\right)(D_A x, x) + \frac{\beta}{2}((\underline{A}x, x) - (\overline{A}x, x)) \\ &= \left(1 - \frac{\beta}{2}\right)(D_A x, x)\end{aligned} \tag{58}$$

$(Qx, x) > 0 \Leftrightarrow 0 < \beta < 2$

Следствие. Метод итераций

$$Ax = b$$
$$x^{(k)} = x^{(k-1)} - \alpha(Ax^{(k-1)} - b) = (E - \alpha A)x^{(k-1)} + \alpha b$$
$$E\frac{x^{(k)} - x^{(k-1)}}{\alpha} + Ax^{(k-1)} = b$$
$$Q = E - \frac{\alpha}{2}A$$
$$(Ax, x) \le \|A\|_2\,(x, x) \Rightarrow (Ex, x) \ge \frac{1}{\|A\|_2}(Ax, x)$$

$$\begin{aligned}(Qx, x) &= (Ex, x) - \frac{\alpha}{2}(Ax, x)\\ &\ge \frac{1}{\|A\|_2}(Ax, x) - \frac{\alpha}{2}(Ax, x)\\ &= \left(\frac{1}{\|A\|_2} - \frac{\alpha}{2}\right)(Ax, x)\end{aligned} \tag{59}$$

$$(Qx, x) > 0 \Leftrightarrow 0 < \alpha < \frac{2}{\|A\|_2} = \frac{2}{\lambda_{max}(A)}$$

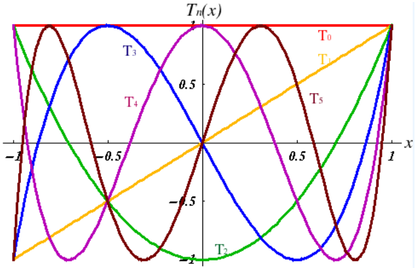
Содержание

Полиномы Чебышева

Полином Чебышева степени n

$$T_n(x) = \cos(n \arccos(x)), x \in [-1, 1]. \tag{60}$$

$T_0(x) = 1, T_1(x) = x.$



Полиномы Чебышева. Свойства

1. Верно рекуррентное соотношение

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{61}$$

Пусть $\phi = \arccos(x)$. Тогда $x = cos(\phi)$ и $T_n(x) = \cos(n\phi)$.

$$T_{n+1}(x) = \cos((n+1)\phi) = \cos(n\phi)\cos(\phi) - \sin(n\phi)\sin(\phi)$$
$$T_{n-1}(x) = \cos((n-1)\phi) = \cos(n\phi)\cos(\phi) + \sin(n\phi)\sin(\phi)$$
$$\Rightarrow T_{n+1}(x) + T_{n-1}(x) = 2\cos(n\phi)\cos(\phi) = 2xT_n(x)$$

$$T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1$$
$$T_3(x) = 2xT_2(x) - T_1(x) = 4x^3 - 3x$$
$$T_4(x) = \dots$$

Полиномы Чебышева. Свойства

2. Все корни полинома Чебышева вещественны, различны и лежат на промежутке $[-1, 1]$.

$$T_n(x) = 0 \Rightarrow \cos(n\phi) = 0 \Rightarrow n\phi_k = \frac{\pi}{2} \pm \pi k \Rightarrow \phi_k = \frac{\pi(1 \pm 2k)}{2n}, k = 0, 1, \dots$$
$$x_k = \cos(\phi_k) = \cos\left(\frac{\pi(1 \pm 2k)}{2n}\right)$$
$$x_k = \cos(\phi_k) = \cos\left(\frac{\pi(1 + 2k)}{2n}\right), k = 0, \dots, n-1.$$

3. На отрезке $[-1, 1]$ полином $T_n(x)$ имеет $n+1$ экстремум.

$$|T_n(x)| = 1 \Rightarrow n\phi_k = \pi k, k = 0, \pm 1, \pm 2, \dots$$
$$\arccos(\bar{x}_k) = \frac{\pi k}{n} \Rightarrow \bar{x}_k = \cos\left(\frac{\pi k}{n}\right), k = 0, \dots, n.$$

$$T_n(\bar{x}_k) = (-1)^k$$

Следствие: $|T_n(x)| > 1$ при $x \notin [-1, 1]$.

Полиномы Чебышева. Свойства

4. Полиномы Чебышева ортогональны с весом $p(x) = \frac{1}{\sqrt{1-x^2}}$ на $[-1, 1]$

$$\int\limits_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx = C_{nm}\delta_{nm} \tag{62}$$

$$\int\limits_{-1}^{1} \frac{\cos(n \arccos(x))\cos(m \arccos(x))}{\sqrt{1-x^2}}dx = -\int\limits_{0}^{\pi} \frac{\cos(n\phi)\cos(m\phi)}{\sin(\phi)}(-\sin(\phi))d\phi$$

$$\frac{1}{2}\int\limits_{0}^{\pi}(\cos((n+m)\phi) + \cos((n-m)\phi))d\phi = \begin{cases} 0, & n \neq m \\ \pi, & n = m = 0 \\ \frac{\pi}{2}, & n = m \neq 0 \end{cases}$$

Полиномы Чебышева. Свойства

5. Из всех многочленов степени n со старшим коэффициентом 1 нормированный многочлен Чебышева $\overline{T}_n(x) = \frac{1}{2^{n-1}} T_n(x) = x^n + \ldots$ наименее уклоняется от 0 на отрезке $[-1, 1]$.

   Если $P_n(x) = x^n + \ldots$, то $\max\limits_{[-1,1]} |P_n(x)| \geq \frac{1}{2^{n-1}}$.
   Предположим, что это не так, т.е. $\exists P_n(x) : \max\limits_{[-1,1]} |P_n(x)| = \frac{\alpha}{2^{n-1}}, 0 < \alpha < 1$.
   $Q_{n-1}(x) = \overline{T}_n(x) - P_n(x)$.
   $sign(Q_{n-1}(\overline{x}_k)) = sign\left(\frac{(-1)^k}{2^{n-1}} - P_n(\overline{x}_k)\right) = (-1)^k$

   $Q_{n-1}(x)$ меняет свой знак в $n+1$ точке $\Rightarrow Q_{n-1}(x)$ имеет по крайней мере n корней. Противоречие.

Содержание

Метод Ричардсона решения СЛАУ

Пусть $A = A^\top, A > 0$ и $\det(A) \neq 0$. Тогда итерационный процесс

$$x^{(k+1)} = x^{(k)} - \alpha(Ax^{(k)} - b) \tag{63}$$

при $0 < \alpha < \frac{2}{\|A\|}$ сходится к решению $Ax = b$.

Заменим α на α_k. Процесс станет не стационарным.
Возьмем начальное приближение $x^{(0)} \in \mathbb{R}^n$ и выполним m итераций, получим $x^{(m)}$ и

$$e^{(m)} = x^{(m)} - x^*. \tag{64}$$

Подберем $\alpha_0, \ldots, \alpha_{m-1}$ так, чтобы ошибка $e^{(m)}$ была минимальна.

Метод Ричардсона решения СЛАУ

Базис из ортонормированных собственных векторов матрицы A

$$\begin{aligned} &Aw^{(j)} = \lambda_j w^{(j)}, j = 1, \ldots, n \\ &(w^{(j)}, w^{(i)}) = \delta_{ji} \\ &\left\| w^{(j)} \right\| = 1 \end{aligned} \tag{65}$$

Любой вектор можно разложить по векторам базиса из ортонормированных собственных векторов матрицы A

$$e^{(k)} = x^{(k)} - x^* = \sum_{j=1}^{n} \beta_j^{(k)} w^{(j)}. \tag{66}$$

(63) $\Rightarrow$

$$e^{(k+1)} = e^{(k)} - \alpha_k A e^{(k)} = (E - \alpha_k A) e^{(k)} \tag{67}$$

Метод Ричардсона решения СЛАУ

$$\sum_{j=1}^{n} \beta_j^{(k+1)} w^{(j)} = (E - \alpha_k A) \sum_{j=1}^{n} \beta_j^{(k)} w^{(j)} = \sum_{j=1}^{n} \beta_j^{(k)} (1 - \alpha_k \lambda_j) w^{(j)}.$$

$$\beta_j^{(k+1)} = \beta_j^{(k)} (1 - \alpha_k \lambda_j) = \beta_j^{(k-1)} (1 - \alpha_{k-1} \lambda_j)(1 - \alpha_k \lambda_j) = \ldots = \beta_j^{(0)} \prod_{s=0}^{k} (1 - \alpha_s \lambda_j).$$

$$\left\| e^{(m)} \right\| = \left\| \sum_{j=1}^{n} \beta_j^{(m)} w^{(j)} \right\| = \left\| \sum_{j=1}^{n} \left(\beta_j^{(0)} \prod_{s=0}^{m-1} (1 - \alpha_s \lambda_j) \right) w^{(j)} \right\|$$

$$\left\| e^{(m)} \right\| \leq \sum_{j=1}^{n} |\beta_j^{(0)}| \underbrace{\left| \prod_{s=0}^{m-1} (1 - \alpha_s \lambda_j) \right|}_{\Psi_j} \leq M_m \sum_{j=1}^{n} |\beta_j^{(0)}| = M_m \left\| e^{(0)} \right\|_1 \leq M_m c \left\| e^{(0)} \right\|,$$

где $M_m = \max\limits_{j} \Psi_j$.

Подберем $\alpha_0, \ldots, \alpha_{m-1}$ так, чтобы M_m получилось минимальным.

Метод Ричардсона решения СЛАУ

$P_m(\lambda) = \prod\limits_{s=0}^{m-1} (1 - \alpha_s \lambda)$ - полином относительно λ.

$$M_m = \max_j \left| \prod_{s=0}^{m-1} (1 - \alpha_s \lambda_j) \right| = \max_j \left| P_m(\lambda_j) \right| \xrightarrow[\alpha_0, \ldots, \alpha_{m-1}]{} \min \tag{68}$$

Пусть имеется оценка для собственных чисел A: $\delta \leq \lambda_j \leq \Delta, j = 1, \ldots, n$.
Вместо задачи (68) будем решать

$$\max_{\lambda \in [\delta, \Delta]} |P_m(\lambda)| \xrightarrow[\alpha_0, \ldots, \alpha_{m-1}]{} \min \tag{69}$$

Сделаем замену переменной: $\lambda = \frac{\delta + \Delta}{2} + \frac{\Delta - \delta}{2} t$.
Тогда $P_m(\lambda) = \overline{P}_m(t) = \mu T_m(t)$, где $T_m(t)$ - полином Чебышева.

Метод Ричардсона решения СЛАУ

$\lambda = 0$ при $t = t_0 \Rightarrow t_0 = -\frac{\Delta+\delta}{\Delta-\delta} \Rightarrow |t_0| > 1 \Rightarrow |T_m(t_0)| > 1$
$P_m(0) = 1 = \mu T_m(t_0) \Rightarrow |\mu| = \frac{1}{|T_m(t_0)|} < 1.$

Если $t_k = \cos\left(\frac{\pi(2k-1)}{2m}\right), k = 1, \dots, m$ - корни полинома Чебышева $T_m(t)$, то

$$\lambda_k = \frac{\delta + \Delta}{2} + \frac{\Delta - \delta}{2} t_k \tag{70}$$

корни полинома $P_m(\lambda) \Rightarrow (1 - \alpha_s \lambda_k) = 0 \Rightarrow$

$$\alpha_s = \frac{1}{\lambda_k} = \frac{2}{(\delta + \Delta) + (\Delta - \delta) t_k}. \tag{71}$$

При вычислении α_s необходимо брать разные t_k ($s = 0, \dots, m-1, k = 1, \dots, m$).

Алгоритм метода Ричардсона

1. Задать m
2. Вычислить t_k и $\lambda_k, k = 1, \dots, m$
3. Вычислить $\alpha_s, s = 0, \dots, m-1$
4. Использовать (63), подставляя на каждой итерации α_s

$$(68) \Rightarrow M_m \leq \max_{\lambda \in [\delta, \Delta]} |P_m(\lambda)| = \max_{t \in [-1,1]} |\mu T_m(t)| = |\mu| < 1.$$

$$\left\| e^{(m)} \right\| \leq c M_m \left\| e^{(0)} \right\| \leq c|\mu| \left\| e^{(0)} \right\|$$

$$\left\| e^{(mp)} \right\| \leq c|\mu|^p \left\| e^{(0)} \right\| \xrightarrow[p \to \infty]{} 0$$

Метод Ричардсона. Замечания

1. Целесообразно выбирать $\alpha_s, s = 0, \ldots, m-1$ в порядке убывания.
2. $m = 1$ - метод с оптимальным выбором итерационного параметра.

$$\alpha_0 = \alpha = \frac{2}{\delta + \Delta}. \tag{72}$$

В идеале

$$\alpha = \frac{2}{\lambda_{min}(A) + \lambda_{max}(A)}. \tag{73}$$

Причем $\alpha < \frac{2}{\lambda_{max}(A)} = \frac{2}{\|A\|_2}$.

Градиентный метод решения СЛАУ (the gradient method, the steepest descent method)

Пусть $A = A^\top, A > 0$ и

$$Ax = b. \tag{74}$$

$\det(A) \neq 0 \Rightarrow \exists!\, x^* : Ax^* = b.$

Определим квадратичную форму $F(y)$

$$F(y) = (Ay, y) - 2(b, y). \tag{75}$$

Теорема

x^* - решение (74) $\Leftrightarrow$ x^* сообщает минимум (75).

Эта теорема позволяет заменить задачу решения СЛАУ (74) на задачу нахождения минимума функции $F(y)$ (75).

Доказательство

$\Rightarrow$

Пусть x^* - решение (74). Тогда

$F(y) = (Ay, y) - 2(b, y) = (Ay, y) - 2(Ax^*, y) + (Ax^*, x^*) - (Ax^*, x^*) =$
$(A(y - x^*), y) - [(Ax^*, y) - (Ax^*, x^*)] - (Ax^*, x^*) = (A(y - x^*), y - x^*) - (Ax^*, x^*)$

$$F(y) = \|y - x^*\|^2_{H_A} - \|x^*\|^2_{H_A}. \tag{76}$$

$\Leftarrow$

Пусть x^* сообщает минимум (75).

$u \in \mathbb{R}^n$ - фиксировано, $\alpha \in \mathbb{R}, y = x^* + \alpha u$

$F(y) = F(x^* + \alpha u) = (A(x^* + \alpha u), x^* + \alpha u) - 2(b, x^* + \alpha u) =$
$(Ax^*, x^*) + (Ax^*, \alpha u) + (A\alpha u, x^*) + (A\alpha u, \alpha u) - 2(b, x^*) - 2(b, \alpha u)$

$$F(x^* + \alpha u) = F(x^*) + 2\alpha(Ax^* - b, u) + \alpha^2(Au, u) = \phi(\alpha). \tag{77}$$

Доказательство

Необходимое и достаточное условие минимума $\phi(\alpha)$ в точке α^*

$$\begin{cases} \phi'(\alpha^*) = 0 \\ \phi''(\alpha^*) > 0 \end{cases} \tag{78}$$

(77) $\Rightarrow$

$$\begin{cases} \phi'(\alpha) = 2\alpha(Au, u) + 2(Ax^* - b, u) \\ \phi''(\alpha) = 2(Au, u) \end{cases} \tag{79}$$

$\Rightarrow (Ax^* - b, u) = 0, \forall u \in \mathbb{R}^n.$

Пусть $u = Ax^* - b \Rightarrow \|Ax^* - b\|^2 = 0 \Rightarrow x^*$ - решение (74).

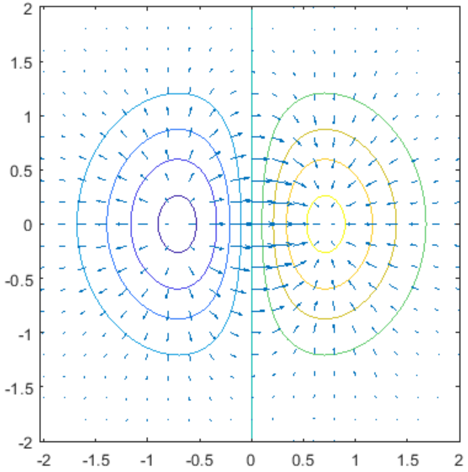
Градиент функции

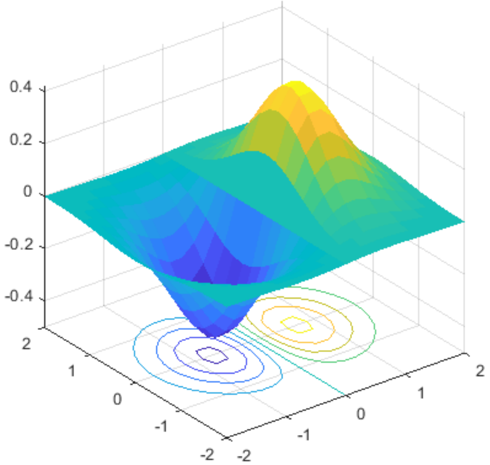
Градиент функции многих переменных $F(x) = F(x_1, x_2, \ldots, x_n)$ - это вектор

$$\nabla F(x) = gradF(x) = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n} \right)^\top . \tag{80}$$

- Градиент функции в точке показывает направление наибольшего возрастания функции.
- Модуль градиента = скорость роста в направлении наибольшего возрастания.
- Градиент функции в точке перпендикулярен линии уровня.

Градиент функции





Градиентный метод решения СЛАУ

Каждое следующее приближение нужно получать, двигаясь в направлении, противоположном градиенту

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla F(x^{(k)}), \alpha_k > 0, \tag{81}$$

причем α_k выбирать так, чтобы $F(x^{(k+1)}) \underset{\alpha}{\to} \min$.

$\nabla F(x^{(k)}) = 2(Ax^{(k)} - b) = g^{(k)}$ (проверить самостоятельно!).
$\nabla F(x^{(k)}) = -2r^{(k)}, r^{(k)} = b - Ax^{(k)}$ - невязка (residual).

$$F(x^{(k+1)}) = F(x^{(k)} - \alpha_k g^{(k)}) = \phi(\alpha_k) \overset{(77)}{=} F(x^{(k)}) - 2\alpha_k(\underbrace{Ax^{(k)} - b}_{\frac{1}{2}g^{(k)}}, g^{(k)}) + \alpha_k^2(Ag^{(k)}, g^{(k)})$$

$\phi'(\alpha_k) = 0 \Rightarrow$

$$\alpha_k = \frac{(g^{(k)}, g^{(k)})}{2(Ag^{(k)}, g^{(k)})} \tag{82}$$

Сходимость градиентного метода

Теорема

$A = A^\top > 0 \Rightarrow$ градиентный метод сходится при любом $x^{(0)}$ и верна оценка

$$\left\| x^{(k+1)} - x^* \right\|_{H_A} \leq \left(\frac{\Delta - \delta}{\Delta + \delta} \right) \left\| x^{(k)} - x^* \right\|_{H_A}, \tag{83}$$

где $0 < \delta < \lambda_j < \Delta, \forall j$.

Доказательство

Вычисление следующего приближения по градиентному методу

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)} = x^{(k)} - 2\alpha_k (Ax^{(k)} - b). \tag{84}$$

Определим $y^{(k+1)}$ как

$$y^{(k+1)} = x^{(k)} - \frac{2}{\delta + \Delta}(Ax^{(k)} - b). \tag{85}$$

$$F(x^{(k+1)}) \le F(y^{(k+1)}) \overset{(76)}{\Leftrightarrow} \left\|x^{(k+1)} - x^*\right\|_{H_A}^2 - \|x^*\|_{H_A}^2 \le \left\|y^{(k+1)} - x^*\right\|_{H_A}^2 - \|x^*\|_{H_A}^2$$

Введем векторы ошибок: $e^{(k)} = x^{(k)} - x^*$ и $r^{(k)} = y^{(k)} - x^* \Rightarrow$

$$\left\|e^{(k+1)}\right\|_{H_A}^2 \le \left\|r^{(k+1)}\right\|_{H_A}^2 . \tag{86}$$

Выразим $r^{(k+1)}$ через $e^{(k)}$

$$r^{(k+1)} = e^{(k)} - \frac{2}{\delta + \Delta}(Ae^{(k)}) = \left(E - \frac{2}{\delta + \Delta}A\right) e^{(k)}. \tag{87}$$

$A = A^\top, A > 0 \Rightarrow A$ имеет базис из собственных векторов. Выберем из него ортонормированный базис $w^{(j)}, j = 1, \ldots, n$.

$e^{(k)} = \sum\limits_{j=1}^{n} \beta_j^{(k)} w^{(j)} \Rightarrow Ae^{(k)} = \sum\limits_{j=1}^{n} \beta_j^{(k)} \lambda_j w^{(j)}$

$\Rightarrow \left\|e^{(k)}\right\|_{H_A}^2 = (Ae^{(k)}, e^{(k)}) = \sum\limits_{j=1}^{n} \left(\beta_j^{(k)}\right)^2 \lambda_j$

$$r^{(k+1)} = \left(E - \tfrac{2}{\delta+\Delta}A\right)\sum_{j=1}^{n}\beta_j^{(k)}w^{(j)} = \sum_{j=1}^{n}\beta_j^{(k)}\left(E - \tfrac{2}{\delta+\Delta}A\right)w^{(j)}$$
$$\Rightarrow \left\|r^{(k+1)}\right\|_{H_A}^2 = \sum_{j=1}^{n}\left(\beta_j^{(k)}\right)^2\left(1 - \tfrac{2\lambda_j}{\delta+\Delta}\right)^2\lambda_j$$

Выберем M так, чтобы $|1 - \frac{2\lambda_j}{\delta+\Delta}| \le M, \forall j = 1, \ldots, n \Rightarrow$

$$\left\|r^{(k+1)}\right\|_{H_A}^2 \le M^2\left\|e^{(k)}\right\|_{H_A}^2. \tag{88}$$

Докажем, что $M < 1$.
$\delta < \lambda_j < \Delta \Leftrightarrow -\frac{\Delta-\delta}{\Delta+\delta} \le 1 - \frac{2\lambda_j}{\delta+\Delta} \le \frac{\Delta-\delta}{\Delta+\delta}$ (проверить самостоятельно).
В качестве M можно взять $\frac{\Delta-\delta}{\Delta+\delta}$, которое меньше 1.

Тогда $\left\|e^{(k+1)}\right\|_{H_A}^2 \overset{(86)}{\le} \left\|r^{(k+1)}\right\|_{H_A}^2 \overset{(88)}{\le} M^2\left\|e^{(k)}\right\|_{H_A}^2 \le \ldots \le M^{2(k+1)}\left\|e^{(0)}\right\|_{H_A}^2$.
$\Rightarrow \left\|x^{(k+1)} - x^*\right\|_{H_A}^2 \le \left(\frac{\Delta-\delta}{\Delta+\delta}\right)^{2(k+1)}\left\|x^{(0)} - x^*\right\|_{H_A}^2 \xrightarrow[k\to\infty]{} 0$

О скорости сходимости

$$\left\| e^{(k+1)} \right\|_{H_A} \leq M \left\| e^{(k)} \right\|_{H_A} \tag{89}$$

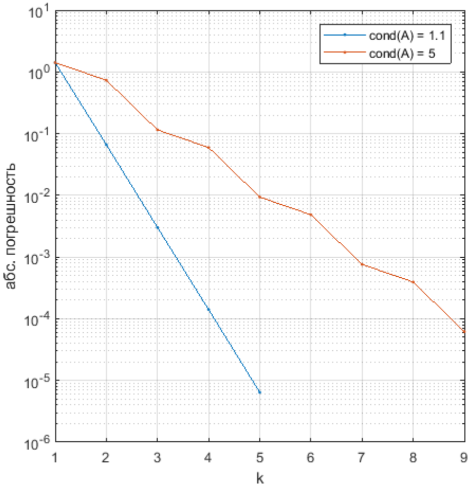
где $M = \frac{\Delta-\delta}{\Delta+\delta} = \frac{\frac{\Delta}{\delta}-1}{\frac{\Delta}{\delta}+1} = \frac{cond(A)-1}{cond(A)+1}$.

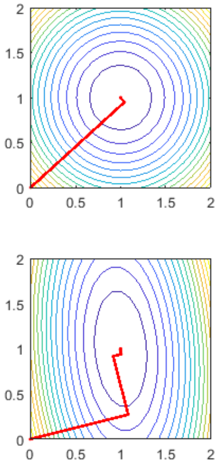
Об окончании итерационного процесса

$$\left\| x^{(k+1)} - x^* \right\|_{H_A} \leq M \left\| x^{(k)} - x^* \pm x^{(k+1)} \right\|_{H_A} \leq M \left\| x^{(k)} - x^{(k+1)} \right\|_{H_A} + M \left\| x^{(k+1)} - x^* \right\|_{H_A}$$

$$\left\| x^{(k+1)} - x^* \right\|_{H_A} \leq \frac{M}{1-M} \left\| x^{(k+1)} - x^{(k)} \right\|_{H_A} \tag{90}$$

Градиентный метод





Метод сопряженных направлений (the Conjugate Gradient Method)

Градиентный метод - движение в направлении антиградиента F. Каждое новое направление ортогонально предыдущему.
Не существует ли более разумного способа выбора нового направления?
Рассмотрим движение в произвольном направлении $p^{(k)}$

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}. \tag{91}$$

$$F(x^{(k+1)}) = F(x^{(k)} + \alpha_k p^{(k)}) \overset{(77)}{=} F(x^{(k)}) + 2\alpha_k(\underbrace{Ax^{(k)} - b}_{\frac{1}{2}g^{(k)}}, p^{(k)}) + \alpha_k^2(Ap^{(k)}, p^{(k)}) = \phi(\alpha_k)$$

$\phi'(\alpha_k) = 0 \Rightarrow$

$$\alpha_k = -\frac{(g^{(k)}, p^{(k)})}{2(Ap^{(k)}, p^{(k)})} = \frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})} \tag{92}$$

Как определить направление $p^{(k)}$?

Сопряженные направления

Векторы x и y называются A-сопряженными (A-ортогональными), если

$$(x, Ay) = 0. \tag{93}$$

Направление $p^{(k)}$ должно быть A-сопряженными всем предыдущим направлениям $p^{(j)}, j = 0, \dots, k-1$

$$(p^{(k)}, Ap^{(j)}) = 0. \tag{94}$$

Будем искать $p^{(k)}$ в виде

$$p^{(k)} = r^{(k)} - \beta_{k-1} p^{(k-1)}. \tag{95}$$

$(p^{(k)}, Ap^{(k-1)}) = 0 \Rightarrow$

$$\beta_{k-1} = \frac{(r^{(k)}, Ap^{(k-1)})}{(p^{(k-1)}, Ap^{(k-1)})}. \tag{96}$$

$(p^{(k)}, Ap^{(j)}) = 0, j = 0, \dots, k-2$ - проверить самостоятельно.

$x^{(0)}, p^{(0)} = r^{(0)}$
$k = 0, 1, \dots$
α_k по формуле (92) и $x^{(k+1)}$ по формуле (91)
$r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}$
β_k по формуле (96) и $p^{(k+1)}$ по формуле (95)

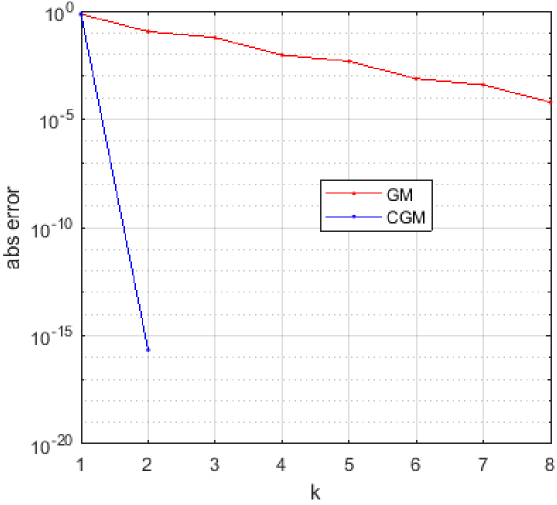
Пусть $A = A^\top > 0$. Тогда точное решение СЛАУ $Ax = b$ по методу сопряженных направлений находится не более чем за n шагов.
Метод может быть отнесен к прямым методам.

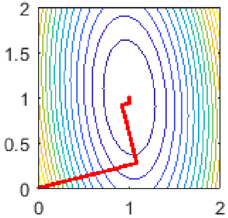
Замечание

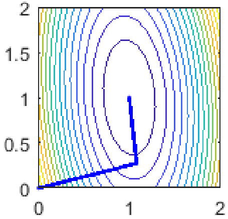
Метод сопряженных направлений применяется как итерационный

- ошибки округления
- число шагов для достижения желаемой точности может оказаться значительно меньшим n.

Градиентный метод и метод сопряженных направлений



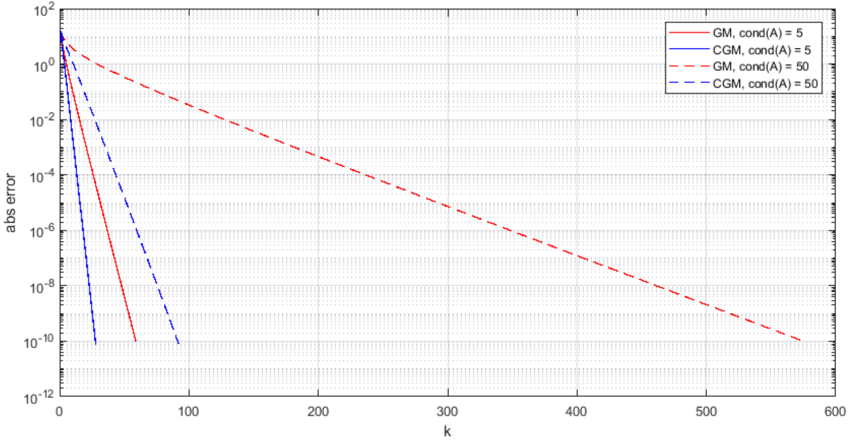




Градиентный метод и метод сопряженных направлений

Пусть $A \in \mathbb{R}^{1000\times 1000}$, $\epsilon = 10^{-10}$.

$cond(A) = 5$, $cond(A) = 50$



О критерии остановки в итерационных методах

МПИ: $x^{(k+1)} = Cx^{(k)} + g \Rightarrow e^{(k+1)} = Ce^{(k)} \Rightarrow \left\| e^{(k+1)} \right\| \leq \|C\| \left\| e^{(k)} \right\| \Rightarrow$

$$\left\| e^{(k+1)} \right\| \leq \frac{\|C\|}{1 - \|C\|} \left\| x^{(k+1)} - x^{(k)} \right\| \tag{97}$$

$\|C\|$ можно оценить:
$x^{(k+1)} - x^{(k)} = C(x^{(k)} - x^{(k-1)}) \Rightarrow$

$$\|C\| \geq \frac{\left\| x^{(k+1)} - x^{(k)} \right\|}{\left\| x^{(k)} - x^{(k-1)} \right\|} = \frac{\delta_{k+1}}{\delta_k}. \tag{98}$$

Учитывая (98) в (97), получим индикатор ϵ_{k+1} для $\left\| e^{(k+1)} \right\|$

$$\epsilon_{k+1} = \frac{\delta_{k+1}^2}{\delta_k - \delta_{k+1}}. \tag{99}$$

О критерии остановки в итерационных методах

Замечание

ϵ_{k+1} не является верхней границей для $\left\|e^{(k+1)}\right\|$.

Замечание

Если про ошибки на соседних итерациях известно, что $\left\|e^{(k+1)}\right\| \leq q \left\|e^{(k)}\right\|$ (например, градиентный метод, формула (83)), то можно провести аналогичные рассуждения.

О критерии остановки в итерационных методах

"Часто" останавливаются по невязке: $\left\|r^{(k)}\right\| \leq \epsilon$. Для абсолютной погрешности верно

$$\left\|x^* - x^{(k)}\right\| = \left\|A^{-1}b - x^{(k)}\right\| = \left\|A^{-1}r^{(k)}\right\| \leq \left\|A^{-1}\right\| \left\|r^{(k)}\right\| \leq \left\|A^{-1}\right\| \epsilon. \tag{100}$$

Пусть контролируется нормализованная невязка: $\frac{\left\|r^{(k)}\right\|}{\|b\|} \leq \epsilon$. Тогда справедлива оценка для относительной погрешности

$$\frac{\left\|x^* - x^{(k)}\right\|}{\|x^*\|} \leq \frac{\left\|A^{-1}\right\| \left\|r^{(k)}\right\|}{\|x^*\|} \leq cond(A) \frac{\left\|r^{(k)}\right\|}{\|b\|} \leq cond(A)\epsilon. \tag{101}$$

О критерии остановки в итерационных методах

Пусть $A \in \mathbb{R}^{1000\times 1000}$, $cond(A) = 100$.

Метод Зейделя.