РЕШЕНИЕ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ (часть 2)

План

- Погрешность прямых методов
- Устойчивость по входным д
- Итерационные методы
  - Приведение СЛАУ к итер
  - Метод простой итерации
  - Метод Зейделя
  - Метод релаксации

Погрешность прямых методов решения СЛАУ

Источники ошибок:

- ограниченность разрядной сетки ЭВМ;
- погрешность представления исходных данных (коэффициентов системы A и (или) вектора правой части F).

Уточнение корней

Обозначим прибл. решение системы (1) через $\overline{\mathbf{x}_i}$ (i=1,2,…,n). Подставим это решение в исходную систему:

$$\mathbf{A\overline{x} = \overline{f}}.$$

Вычтем эти системы:

невязка

$$\mathbf{A(x - \overline{x}) = f - \overline{f}}.$$

Обозначим через $\mathbf{y} = \mathbf{x} - \overline{\mathbf{x}}$, $\boldsymbol{\varepsilon} = \mathbf{f} - \overline{\mathbf{f}}$

Уточнение корней

Получим новую систему:

$$\mathbf{Ay=\varepsilon}.$$

Решая ее, находим прибл. корни $\bar{y}$.

Это решение используется для уточнения $\bar{x}$:

$$\bar{\bar{\mathbf{x}}} = \mathbf{x + y}.$$

Этот процесс можно повторять.

Устойчивость по входным данным

Фактически вместо исходной системы при наличии погрешности исходных данных решается не система (1), а «возмущенная система (2):

$$A\vec{X} = \vec{f} \quad (1)$$

$$A^{*}\vec{X}^{*} = \vec{f}^{*} \quad (2)$$

Устойчивость по входным данным

Необходимо оценить, как связаны погрешность решения **ΔX=X*-X** с абсолютными погрешностями коэффициентов матрицы **ΔA=A*-A** и свободных членов **Δf=f*-f**.

Это т.н. коэффициентная устойчивость и устойчивость по правой части.

Устойчивость по входным данным

Для оценки погрешности и устойчивости системы обратимся к основным характеристикам СЛАУ, известным из курса алгебры.

Напомним некоторые определения.

Определение 1. Норма вектора

Нормой вектора X называется неотрицательное число ||X|| такое, что

$$1.\ \|X\| > 0,\ \text{при } X \neq 0;\ \|0\| = 0$$

$$2.\ \|c \cdot X\| = |c| \cdot \|X\|\ ,\ \mathrm{c} \in R;$$

$$3.\ \|X + Y\| \leq \|X\| + \|Y\|.$$

Определение 2. Норма матрицы

Нормой матрицы А называется неотрицательное число ||A|| такое, что

$$1.\ \|A\| > 0,\ \text{при } A \neq \theta;\quad \|\theta\| = 0$$

$$2.\ \|c \cdot A\| = |c| \cdot \|A\|,\ \mathrm{c} \in R;$$

$$3.\ \|A + B\| \leq \|A\| + \|B\|;$$

$$4.\ \|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

Виды норм

вектора: | **матрицы:**

$$\|X\|_I = \max_{i=\overline{1,n}} |x_i|$$

кубическая

$$\|A\|_I = \max_{i=\overline{1,n}} \sum_{j=1}^{n} |a_{ij}|$$

$$\|X\|_{II} = \sum_{i=1}^{n} |x_i|$$

октаэдрическая

$$\|A\|_{II} = \max_{j=\overline{1,n}} \sum_{i=1}^{n} |a_{ij}|$$

$$\|X\|_{III} = \sqrt{\sum_{i=1}^{n} x_i^{\,2}}$$

сферическая

$$\|A\|_{III} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2}$$

Упражнение. Найдите нормы вектора

$$X = \begin{bmatrix} -3 \\ 2 \\ 1 \\ -1 \end{bmatrix}$$

$\|X\|_I = ?$ 3

$\|X\|_{II} = ?$ 7

$\|X\|_{III} = ?$ $\sqrt{15}$

Упражнение. Найдите нормы матриц

$$A = \begin{pmatrix} 2 & -3 & -1 \\ 3 & 13 & 4 \\ 4 & 7 & 2 \end{pmatrix}$$

кубическую:

$$\|A\|_I = \max\{6,20,13\} = 20$$

октаэдрическую:

$$\|A\|_{II} = \max\{9,23,7\} = 23$$

$$A = \begin{bmatrix} -2 & -1 & 0 \\ 0 & 2 & 0 \\ 3 & -1 & -2 \end{bmatrix}$$

сферическую:

$$\|A\|_{III} = \sqrt{23}$$

Погрешность решения

Относительная погрешность решения связана с погрешностью правой части следующим образом:

$$\frac{\|\Delta X\|}{\|X\|} \le M_A \cdot \frac{\|\Delta f\|}{\|f\|}$$

Константа пропорциональности M_A- число обусловленности матрицы.

Число обусловленности матрицы

$$M_A = \|A\| \cdot \|A^{-1}\|$$

Обусловленность – это внутреннее свойство матрицы, не связанное со способом решения системы уравнений. Если это число велико, матрица называется плохо обусловленной. Решение СЛАУ считается ненадежным, если

$$M_A \geq \frac{\delta}{\varepsilon_{маш}},\ \delta \text{ допустимая отн. погрешность}$$

Примеры

Классическим примером плохо обусловленной матрицы является матрица Гильберта с элементами

$$h_{ij} = 1/(i + j + 1)$$

Уже при n=8 $M_H \geq 10^{10}$.

$$\begin{cases} 1.00x + 0.99y = 1.99, \\ 0.99x + 0.98y = 1.97. \end{cases}$$

Точное решение:
x=1, y=1.

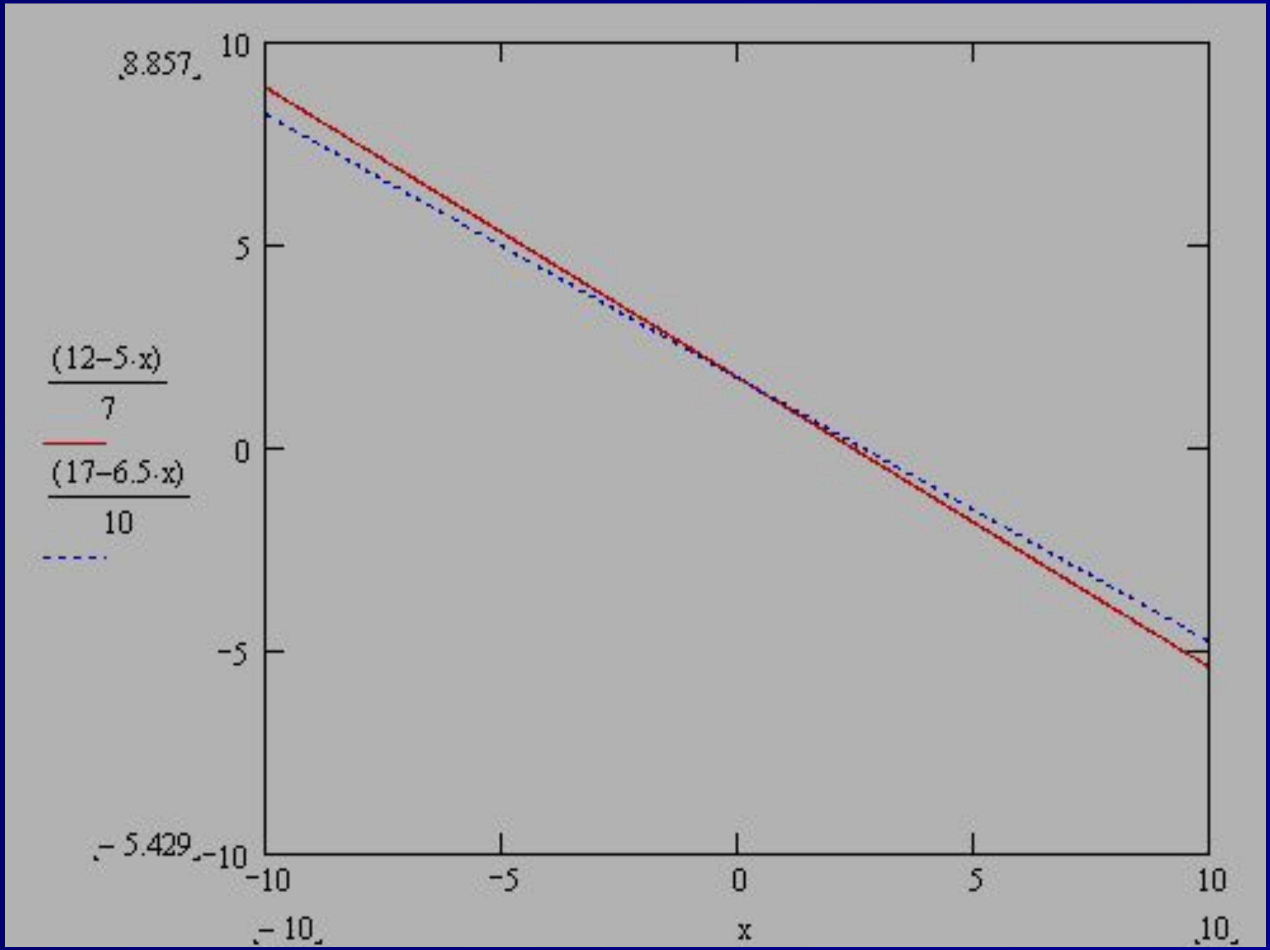
Небольшое отклонение в исходных данных резко меняет решение:

$$\begin{cases} 1.00x + 0.99y = 1.989903, \\ 0.99x + 0.98y = 1.970106. \end{cases}$$

Эта система дает решение x=3.000, y=-1.0203.

Число обусловленности матрицы А в данном случае порядка 40 000.

В двумерном случае возможна графическая трактовка понятия обусловленности: прямые, являющиеся геометрическими образами уравнений пересекаются под очень острым углом:



Графическая трактовка понятия обусловленности

Небольшое искажение данных (параллельный перенос при изменении свободных членов, поворот при возмущении матрицы коэффициентов) приводит к значительному перемещению точки пересечения (геометрического образа решения).

Какие из систем плохо обусловлены?

1) $\begin{cases} 5x - 3y = -11 \\ -2{,}5x + 1{,}501y = 5{,}5 \end{cases}$

2) $\begin{cases} -3x + 2y = 1 \\ 6{,}1x + 4y = 1 \end{cases}$

3) $\begin{cases} 2{,}1x - 4y = -5 \\ 3x + 8y = 10 \end{cases}$

4) $\begin{cases} 0{,}2x + y = 4{,}5 \\ x + 5{,}01y = 13 \end{cases}$

Ответ: 1), 4)

Число обусловленности

На практике для нахождения числа обусловленности пользуются его свойством:

$$M_A \geq \frac{|\lambda_{\max}|}{|\lambda_{\min}|}$$

где λ_{max} и λ_{min} – наибольшее и наименьшее по модулю собственные числа матрицы А.

Обусловленность матрицы

Решающую роль в обусловленности играет не близость к 0 определителя, а разброс собственных значений матрицы.

Погрешность решения

Знание числа обусловленности позволяет оценить и погрешности округления на ЭВМ. Если каждая компонента вектора правой части округляется с относительной погрешностью $O(2^{-m})$, где m – разрядность сетки ЭВМ, то решение задачи (1) не может быть найдено с точностью, большей, чем $M_A O(2^{-m})$.

Обусловленность систем

Решение плохо обусловленных систем – некорректная задача. Чтобы преодолеть некорректность используют ***метод регуляризации***, основанный на привлечении дополнительной информации о решении.

Итерационные методы

В случае произвольной матрицы А и достаточно большого числа n решение системы (1) с помощью прямых методов затруднительно. В этом случае применяются итерационные методы.

Методы решения СЛАУ

- ПРЯМЫЕ
- ИТЕРАЦИОННЫЕ
  - ПРОСТОЙ ИТЕРАЦИИ
  - ЗЕЙДЕЛЯ
  - РЕЛАКСАЦИИ
- ВЕРОЯТНОСТНЫЕ

Приведение системы к итерационному виду

Для этого систему (1) переписываем в виде, удобном для итерационного процесса:

(2)
$$\begin{cases} x_1 = \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n + \beta_1 \\ x_2 = \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n + \beta_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n + \beta_n \end{cases}$$

В матричном виде:

$$X = aX + \beta$$

Приведение системы к итерационному виду

Пример. Пусть дана система

$$\begin{cases} 5x_1 - x_2 + x_3 = 4 \\ 0,5x_1 + 2x_2 - x_3 = 1 \\ x_1 - x_2 + 4x_3 = 2 \end{cases}$$

Приведение системы к итерационному виду

Способ 1. Поделив каждое уравнение на соответствующий диагональный элемент (если они не равны нулю) и решив уравнения относительно диагональных переменных, получим:

$$\begin{cases} 5x_1 - x_2 + x_3 = 4 \\ 0{,}5x_1 + 2x_2 - x_3 = 1 \\ x_1 - x_2 + 4x_3 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = 0{,}2x_2 - 0{,}2x_3 + 0{,}8 \\ x_2 = -0{,}25x_1 + 0{,}5x_3 + 0{,}5 \\ x_3 = -0{,}25x_1 + 0{,}25x_2 + 0{,}5 \end{cases} \quad (*)$$

Приведение системы к итерационному виду

Способ 2. Вычленим единицу из каждого диагонального элемента и выразим полученные неизвестные:

$$\begin{cases} 5x_1 - x_2 + x_3 = 4 \\ 0{,}5x_1 + 2x_2 - x_3 = 1 \\ x_1 - x_2 + 4x_3 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = -4x_1 + x_2 - x_3 + 4 \\ x_2 = -0{,}5x_1 - x_2 + x_3 + 1 \\ x_3 = -x_1 + x_2 - 3x_3 + 2 \end{cases} \quad (**)$$

Эти системы эквивалентны, т.е. имеют одинаковые решения.

Все итерационные методы решения системы, приведённой к итерационному виду, состоят в том, что выбирается некоторое начальное приближение (вектор X_0), которое подставляется в правую часть итерационной системы. В левой части тогда получаем решение на первой итерации, затем подставляем его в правую часть - и т. д. Последовательность векторов строится до тех пор, пока не будет достигнута заданная точность:

$$\left\| X^{(k+1)} - X^{(k)} \right\| < \varepsilon, i = \overline{1, n}$$

Метод простой итерации

Итерационные методы различаются только по способу построения последовательности.

Формулы метода простой итерации:

$$\vec{X}^{(1)} = \alpha \vec{X}^{(0)} + \vec{\beta}$$

$$\vec{X}^{(2)} = \alpha \vec{X}^{(1)} + \vec{\beta}$$

....

$$\vec{X}^{(k+1)} = \alpha \vec{X}^{(k)} + \vec{\beta}$$

$$x_i^{(k+1)} = \sum_{j=1}^{n} \alpha_{ij} \cdot x_j^{(k)} + \beta_i, i = 1,\dots,n; k = 0,1,\dots$$

Метод простой итерации

Прерываем итерационный процесс, когда выполнится условие достижения заданной точности.

Для первой приведенной системы (*) из примера последовательность приближений решения:

$$\begin{cases} x_1 = 0,2x_2 - 0,2x_3 + 0,8 \\ x_2 = -0,25x_1 + 0,5x_3 + 0,5 \\ x_3 = -0,25x_1 + 0,25x_2 + 0,5 \end{cases}$$

$\boldsymbol{x}^{(0)}=(0,0,0)$

$\boldsymbol{x}^{(1)}=(0.8,0.5,0.5)$

$\boldsymbol{x}^{(2)}=(0.85,0.55,0.425)$

……………………….

Достаточное условие сходимости метода простой итерации

Рассмотрим, при каком условии сходится построенная последовательность векторов. Сходимость векторов рассматривается покоординатная, т. е. расстояние между векторами вычисляется по формуле:

$$\left\| \vec{x} - \vec{y} \right\| = \max_{i=1,\dots,n} \left| x_i - y_i \right|$$

Теорема. Итерационный процесс сходится к единственному решению, если какая-либо каноническая норма матрицы $\|\alpha\| < 1.$

Нормой квадратной матрицы А называется вещественное число $\|A\|$, удовлетворяющее следующим условиям:

$$1.\ \|A\| > 0 \ \text{при A} \neq 0,\ \|0\| = 0;$$

$$2.\ \|c \cdot A\| = |c| \cdot \|A\|,\ \text{c} \in R;$$

$$3.\ \|A + B\| \leq \|A\| + \|B\|;$$

$$4.\ \|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

Для канонической нормы выполняется еще два дополнительных условия:

$$5.\ |a_{ij}| \leq \|A\|,\ \text{для}\ \forall i\ j = 1,\dots,n$$

$$6.\ \det A \leq \det B \Rightarrow \|A\| \leq \|B\|$$

Доказательство

Запишем k-ое приближение итерационной последовательности

$$x^{(1)} = ax^{(0)} + \beta$$

$$x^{(2)} = ax^{(1)} + \beta$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x^{(k+1)} = ax^{(k)} + \beta$$

через все предыдущие

$$X^{(k)} = (E + \alpha + \alpha^2 + \ldots + \alpha^{k-1})\beta + \alpha^k X^{(0)}$$

$$X^{(k)} = (E + \alpha + \alpha^2 + \ldots + \alpha^{k-1})\beta + \alpha^k X^{(0)}$$

Т.к. $\|\alpha\| < 1$, то $\|\alpha\| \to 0 \text{ при } \mathrm{k} \to \infty.$

$$\lim_{k\to\infty} X^{(k)} = \lim_{k\to\infty}(E + \alpha + \alpha^2 + \ldots + \alpha^{k-1})\beta$$

Используем формулу, аналогичную сумме бесконечной убывающей геометрической прогрессии со знаменателем q<1:

$$\sum = a_0 \frac{1}{1-q}$$

$$\lim_{k\to\infty} X^{(k)} = \lim_{k\to\infty}(E+\alpha+\alpha^2+\ldots+\alpha^{k-1})\beta$$

$$\sum = a_0 \frac{1}{1-q}$$

$$\vec{X} = \vec{\beta}\cdot(E-\alpha)^{-1}$$

$$\vec{X} = \alpha\vec{X} + \vec{\beta}$$

Т.о., предельный вектор *X* является решением итерационной системы. Теорема доказана.

Следствие 1. Итерационный процесс сходится, если для системы (2) выполняется любое из неравенств:

1) $$\max_{i=\overline{1,n}} \sum_{j=1}^{n} |\alpha_{ij}| < 1$$

кубическая

2) $$\max_{j=\overline{1,n}} \sum_{i=1}^{n} |\alpha_{ij}| < 1$$

октаэдрическая

3) $$\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |\alpha_{ij}|^2} < 1$$

сферическая

Следствие 2. Для исходной системы (1) достаточным условием сходимости метода итераций является одно из:

$$\left|a_{ii}\right| > \sum_{\substack{j=1 \\ i \neq j}}^{n} \left|a_{ij}\right|$$

или

$$\left|a_{jj}\right| > \sum_{\substack{i=1 \\ i \neq j}}^{n} \left|a_{ij}\right|$$

В приведенном примере итерации сходятся для системы (*) согласно следствия 1, для системы (**) - следствия 2.

Проверка условий сходимости

$$\begin{cases} 2x_1 + 3x_2 - x_3 = -3, \\ 3x_1 - 2x_2 + 4x_3 = 13, \\ 4x_1 + x_2 - 2x_3 = 7. \end{cases}$$

$$|a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^{n} |a_{ij}|$$

Приведем систему к диагональному преобладанию путем линейной комбинации строк

У1 = У1+У2
У2 = 2*У1 – У2
У3 = У1-2*У2+У3

$$A := \begin{bmatrix} 5 & 1 & 3 & 10 \\ 1 & 8 & -6 & -19 \\ 0 & 8 & -11 & -22 \end{bmatrix}$$

Оценка погрешности метода простой итерации

Если матрица α удовлетворяет условию

$$q = \max_{i=\overline{1,n}} \sum_{j=1}^{n} \left|\alpha_{ij}\right| < 1$$

, то для оценки погрешности приближенного вектора $X^{(k)}$ имеет место формула:

$$\Delta_{X^{(k)}} = \frac{q}{1-q} \cdot \max_i \left| x_i^{(k-1)} - x_i^{(k)} \right|.$$

Метод Зейделя

В этом методе для нахождения i-ой компоненты решения на (к+1)-ой итерации используются уже найденные на (к+1) значения

$$x_1^{k+1}, x_2^{k+1}, \dots, x_{i-1}^{k+1}$$

Метод Зейделя

$$
\begin{cases}
x_1^{(k+1)} = \alpha_{11}x_1^{(k)} + \alpha_{12}x_2^{(k)} + \ldots + \alpha_{1n}x_n^{(k)} + \beta_1 \\
x_2^{(k+1)} = \alpha_{21}x_1^{(k+1)} + \alpha_{22}x_2^{(k)} + \ldots + \alpha_{2n}x_n^{(k)} + \beta_2 \\
x_3^{(k+1)} = \alpha_{31}x_1^{(k+1)} + \alpha_{32}x_2^{(k+1)} + \alpha_{33}x_3^{(k)} + \ldots + \alpha_{2n}x_n^{(k)} + \beta_2 \\
\ldots\ldots\ldots\ldots\ldots \\
x_n^{(k+1)} = \alpha_{n1}x_1^{(k+1)} + \alpha_{n2}x_2^{(k+1)} + \ldots + \alpha_{nn-1}x_{n-1}^{(k+1)} + \alpha_{nn}x_n^{(k)} + \beta_n
\end{cases}
$$

Таким образом, при вычислении очередной компоненты x_i используются вычисленные x_1, x_2, …, x_{i-1} с новой итерации, в то время как в методе простой итерации используются все значения с предыдущей.

В матричном виде метод Зейделя можно записать в следующем виде:

$$x^{(k+1)} = Lx^{(k+1)} + (R + D)x^{(k)} + \beta$$

$$L + D + R = \alpha$$

$$L = \begin{pmatrix} 0 \ldots\ldots\ldots 0 \ldots\ldots 0 \ldots\ldots\ldots 0 \\ \alpha_{21} \ldots\ldots 0 \ldots\ldots 0 \ldots\ldots\ldots 0 \\ \alpha_{31} \cdots \alpha_{32} \ldots 0 \ldots\ldots\ldots 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{n1} \cdots \alpha_{n2} \ldots\ldots\ldots\ldots 0 \end{pmatrix} \qquad R = \begin{pmatrix} 0 \ldots\ldots \alpha_{12} \cdots\cdots \alpha_{1n} \\ 0 \ldots\ldots\ldots 0 \ldots\ldots\ldots \alpha_{2n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 \ldots\ldots\ldots 0 \ldots\ldots\ldots\ldots 0 \end{pmatrix}$$

D – диагональ матрицы α

Условие сходимости

Достаточное условие сходимости практически аналогично методу простой итерации. Однако, в некоторых случаях оно менее строгое, т.е. метод Зейделя может сходиться, когда метод простой итерации расходится. Обычно метод Зейделя дает лучшую скорость сходимости, чем метод простой итерации.

Условие сходимости метода Зейделя

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

Необходимое и достаточное условие: модули всех корней уравнения (3) должны быть меньше 1.

Метод релаксации

Метод релаксации является ***обобщением метода Зейделя***. Алгоритм построения итерационной последовательности в данном случае имеет вид:

$$X^{(k+1)} = (1-\omega)\cdot X^{(k)} + \omega\cdot(LX^{(k+1)} + (R+D)X^{(k)} + \beta)$$

где ω=const>0 – параметр релаксации, управляющий скоростью сходимости итерационного процесса.

$$X^{(k+1)} = (1-\omega)\cdot X^{(k)} + \omega\cdot(LX^{(k+1)} + (R+D)X^{(k)} + \beta)$$

При **ω=1** алгоритм соответствует методу **Зейделя**,

если **ω>1** – методу **верхней релаксации**,

ω<1 – методу **нижней релаксации**.

Условие сходимости метода релаксации аналогично рассмотренным для методов простой итерации и Зейделя.

Скорость сходимости метода может быть самой высокой из итерационных методов в случае оптимального выбора параметра ω.