



МГТУ имени Н.Э. Баумана

Кафедра ИУ-1 «Системы автоматического управления»

Методы вычислений

Численные методы решения СЛАУ



Андрей Леонидович Масленников
amas@bmstu.ru

2023 г.

Система линейных алгебраических уравнений:

$$\mathbf{Ax} = \mathbf{b} \Leftrightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n \end{cases}$$

где:

n — количество уравнений

m — количество неизвестных

x_i — неизвестные

a_{ij} — коэффициенты

b_i — свободные члены

\mathbf{A} — матрица $n \times m$

\mathbf{x} — вектор $m \times 1$

\mathbf{b} — вектор $n \times 1$

Решение СЛАУ имеет вид:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Свойства СЛАУ:

- $b_i = 0$ — однородная система;
- $m = n$ — совместная система (единственное решение)
- $n > m$ — переопределенная система (множество решений)
- $n < m$ — недоопределённая система (нет решений)

Теорема Кронекера—Капелли.

Система уравнений $\mathbf{Ax} = \mathbf{b}$ разрешима (имеет хотя бы одно решение) тогда и только тогда, когда $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A}, \mathbf{b})$ где (\mathbf{A}, \mathbf{b}) — расширенная матрица системы. В этом случае количество искоемых переменных равно порядку матрицы, а решение системы будет единственным.

Прямые методы решения СЛАУ — это методы в которых вычисление решения осуществляется за конечное количество операций

- метод решения через обратную матрицу;
- метод Крамера;
- метод Гаусса;
- метод Гаусса-Жордана;
- метод через разложение Холецкого;
- метод через LU-разложение;
- и др.

Решение будет найдено всегда, если метод применим, но возможны очень большие ошибки

Итерационные методы решения СЛАУ — это методы в которых вычисление решения осуществляется в результате последовательного приближения (итерационного процесса)

- метод Якоби
(метод простой итерации)
- метод Гаусса-Зейделя;
- метод релаксации;
- метод сопряженных градиентов;
- метод бисопряженных градиентов;
- метод стабилизированных бисопряженных градиентов;
- и др.

Решение не всегда может быть найдено

Алгоритм метода

1. Вычисляем все алгебраические дополнения

$$A_{ij} = (-1)^{i+j} M_{ij}$$

2. Вычисляем обратную матрицу

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \text{adj} \mathbf{C}$$

3. Вычисляем решение

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$$

Условия применимости

- $\det(\mathbf{A}) \neq 0$.

Метод требует существенных вычислительных ресурсов, если не использовать разложение исходной матрицы \mathbf{A}

Вычисляем поэлементно решение

$$x_i = \frac{1}{\det(\mathbf{A})} \begin{vmatrix} a_{1,1} & \dots & b_1 & a_{1,i+1} \dots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & b_n & a_{n,i+1} \dots & a_{n,n} \end{vmatrix}$$

Условия применимости

- \mathbf{A} — квадратная матрица;
- $\det(\mathbf{A}) \neq 0$.

Метод требует существенных
вычислительных ресурсов

Алгоритм метода

```

for  $i = 1 : n$ 
    for  $j = i + 1 : n$ 
         $k = a_{ji} / a_{ii}$ 
         $\mathbf{a}_j = \mathbf{a}_j - k\mathbf{a}_i$ 
         $b_j = b_j - kb_i$ 
    end
end
for  $i = n : 1$ 
    
$$x_i = \frac{b_i - \sum_{k=i+1}^n a_{ik}x_k}{a_{ii}}$$

end

```

Условия применимости

- \mathbf{A} — квадратная матрица

прямой ход

обратный ход

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \xrightarrow{\text{прямой ход}} \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -3.6 \end{bmatrix} \xrightarrow{\text{обратный ход}} \mathbf{x} = \begin{bmatrix} 0.1667 \\ 0.1666 \\ 0.1667 \end{bmatrix}$$

Устойчивость метода сильно зависит от диагональных элементов матрицы \mathbf{A}

Пусть даны

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Итерация 1

 $i = 1, \quad k = i + 1 = 2:$

$$[3 \ 1 \ 2] - [1 \ 2 \ 3] \cdot \frac{3}{1} = [0 \ -5 \ -7]$$

$$[1] - [1] \cdot \frac{3}{1} = [-2]$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 2 & 3 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Итерация 2

 $i = 1, \quad k = i + 1 = 3:$

$$[2 \ 3 \ 1] - [1 \ 2 \ 3] \cdot \frac{2}{1} = [0 \ -1 \ -5]$$

$$[1] - [1] \cdot \frac{2}{1} = [-2]$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$$

Итерация 3

 $i = 2, \quad k = i + 1 = 3:$

$$[0 \ -1 \ -5] - [0 \ -5 \ -7] \cdot \frac{-1}{-5} = [0 \ 0 \ -3.6]$$

$$[(-1)] - [(-2)] \cdot \frac{-1}{-5} = [-0.6]$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -3.6 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -0.6 \end{bmatrix}$$

Обратный ход

$$x_3 = -0.6 / (-3.6) = 0.1667$$

$$x_2 = \frac{-2 - (-7 \cdot 0.1667)}{-5} = 0.1666$$

$$x_1 = \frac{1 - (2 \cdot 0.1666 + 3 \cdot 0.1667)}{1} = 0.1667$$

Решение

$$\mathbf{x} = \begin{bmatrix} 0.1667 \\ 0.1666 \\ 0.1667 \end{bmatrix}$$

Алгоритм метода

for $i = 1 : n$

for $j = i + 1 : n$

$$k = a_{ji} / a_{ii}$$

$$\mathbf{a}_j = \mathbf{a}_j - k\mathbf{a}_i$$

$$b_j = b_j - kb_i$$

end

end

for $i = n : 1$

for $j = i - 1 : 1$

$$k = a_{ji} / a_{ii}$$

$$\mathbf{a}_j = \mathbf{a}_j - k\mathbf{a}_i$$

$$b_j = b_j - kb_i$$

end

end

Условия применимости

- \mathbf{A} — квадратная матрица

прямой ход

обратный ход

$$\mathbf{C} = (\mathbf{A}, \mathbf{b}) = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 3 & 1 & 2 & 1 \\ 2 & 3 & 1 & 1 \end{bmatrix} \xrightarrow{\text{прямой ход}} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -5 & -7 & -2 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

$$\xrightarrow{\text{обратный ход}} \begin{bmatrix} 1 & 0 & 0 & 0.168 \\ 0 & -5 & 0 & -0.8333 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

$$\xrightarrow{\text{вычисление решения}} \mathbf{x} = \begin{bmatrix} 0.1680 \\ 0.1660 \\ 0.1667 \end{bmatrix}$$

После завершения прямого хода:

$$C = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -5 & -7 & -2 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

Решение

$$x_1 = 0.168 / 1 = 0.1680$$

$$x_2 = (-0.8333) / (-5) = 0.1666$$

$$x_3 = (-0.6) / (-3.6) = 0.1667$$

$$i = 3, \quad k = i - 1 = 2;$$

$$[0 \quad -5 \quad -7 \quad -2] - [0 \quad 0 \quad -3.6 \quad -0.6] \cdot \frac{-7}{-3.6} = [0 \quad -5 \quad 0 \quad -0.8333]$$

$$C = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -5 & 0 & -0.8333 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

$$i = 3, \quad k = i - 1 = 1;$$

$$[1 \quad 2 \quad 3 \quad 1] - [0 \quad 0 \quad -3.6 \quad -0.6] \cdot \frac{3}{-3.6} = [1 \quad 2 \quad 0 \quad 0.5]$$

$$C = \begin{bmatrix} 1 & 2 & 0 & 0.5 \\ 0 & -5 & 0 & -0.8333 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

$$i = 2, \quad k = i - 1 = 1;$$

$$[1 \quad 2 \quad 0 \quad 0.5] - [0 \quad -5 \quad 0 \quad -0.8333] \cdot \frac{2}{-5} = [1 \quad 0 \quad 0 \quad 0.168]$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0.168 \\ 0 & -5 & 0 & -0.8333 \\ 0 & 0 & -3.6 & -0.6 \end{bmatrix}$$

Алгоритм метода

1. Находим разложение Холецкого
2. Решаем систему $\mathbf{L}\mathbf{y} = \mathbf{b}$
3. Решаем систему $\mathbf{L}^T\mathbf{x} = \mathbf{y}$

Условия применимости

- \mathbf{A} — положительно определенная матрица

Находим разложение Холецкого

$$\mathbf{A} = \begin{bmatrix} 1.4 & 1 & 1 \\ 1 & 0.9 & 1 \\ 1 & 1 & 1.4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{L} = \begin{bmatrix} 1.18 & 0 & 0 \\ 0.85 & 0.43 & 0 \\ 0.85 & 0.66 & 0.5 \end{bmatrix}$$

Несимметричность матрицы \mathbf{A} не накладывает ограничений на применимость, но полученное решение будет не соответствовать истинному.

Решаем $\mathbf{L}\mathbf{y} = \mathbf{b}$

$$y_1 = 1 / 1.18 = 0.85$$

$$y_2 = \frac{1 - (0.85 \cdot 0.85)}{0.43} = 0.65$$

$$y_3 = \frac{1 - (0.85 \cdot 0.85 + 0.65 \cdot 0.66)}{0.5} = -0.31$$

Решаем $\mathbf{L}^T\mathbf{x} = \mathbf{y}$

$$x_3 = -0.31 / 0.5 = -0.62$$

$$x_2 = \frac{0.65 - 0.66 \cdot (-0.62)}{0.43} = 2.48$$

$$x_1 = \frac{0.85 - (0.85 \cdot (-0.62) + 0.85 \cdot 2.48)}{1.18} = -0.62$$

Итерационные методы решения СЛАУ — это методы в которых вычисление решения осуществляется в результате последовательного приближения (итерационного процесса)

- метод Якоби
(метод простой итерации)
- метод Гаусса-Зейделя;
- метод релаксации;
- метод сопряженных градиентов;
- метод бисопряженных градиентов;
- метод стабилизированных бисопряженных градиентов;
- и др.

Решение не всегда может
быть найдено

Алгоритм итерационных методов

while *criteia ok*

for $i = i : n$

$x_i = \dots$

end

$k = k + 1$

end

Критерии остановки:

- по максимальному количеству итераций

$$k \geq k_{\max}$$

- по величине относительной ошибки решения

$$\|\mathbf{x}_k - \mathbf{x}_{k-1}\| \leq \varepsilon$$

- по величине абсолютной ошибки решения

$$\|\mathbf{Ax}_k - \mathbf{b}\| \leq \varepsilon$$

Решение рассчитывается итерационно

- в векторно-матричном виде

$$\mathbf{x}_{k+1} = \mathbf{B}\mathbf{x}_k + \mathbf{g}$$

- в поэлементном виде

$$x_{i,k+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij} x_{j,k} \right)$$

где

\mathbf{L} , \mathbf{U} — нижняя и верхняя треугольные части \mathbf{A}

$$\mathbf{D} = \text{diag}(\mathbf{A})$$

$$\mathbf{B} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A} = \mathbf{D}^{-1}(\mathbf{D} - \mathbf{A})$$

$$\mathbf{B} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}) = -\mathbf{D}^{-1}(\mathbf{A} - \mathbf{D})$$

$$\mathbf{g} = \mathbf{D}^{-1}\mathbf{b}$$

Критерий сходимости

$$|\lambda_i(\mathbf{B})| < 1, \quad \forall i = 1 \dots n$$

При сходимости справедливо:

- $\|\mathbf{B}\| < 1$ при $\forall \mathbf{x}_0$
- $q = \|\mathbf{B}\|$
(q — скорость сходимости, т.е. решение будет сходиться к истинному со скоростью геометрической прогрессии)
- $\|\mathbf{x}_k - \mathbf{x}\| < q^k \|\mathbf{x}_0 - \mathbf{x}\|$

Метод требует задания начальных условия, т.е. вектора \mathbf{x}_0 , выбор которого влияет на сходимость к истинному решению

Решение рассчитывается итерационно

- в векторно-матричном виде

$$(\mathbf{L} + \mathbf{D})\mathbf{x}_{k+1} = -\mathbf{U}\mathbf{x}_k + \mathbf{B}$$

- в поэлементном виде

$$x_{i,k+1} = \sum_{j=1}^{i-1} c_{ij}x_{j,k+1} + \sum_{j=i+1}^n c_{ij}x_{j,k} + d_i$$

где

\mathbf{L} , \mathbf{U} — нижняя и верхняя треугольные части \mathbf{A}

$$\mathbf{D} = \text{diag}(\mathbf{A})$$

$$c_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}}, & j \neq i \\ 0, & j = i \end{cases}$$

$$d_i = \frac{b_i}{a_{ii}}$$

Критерий сходимости

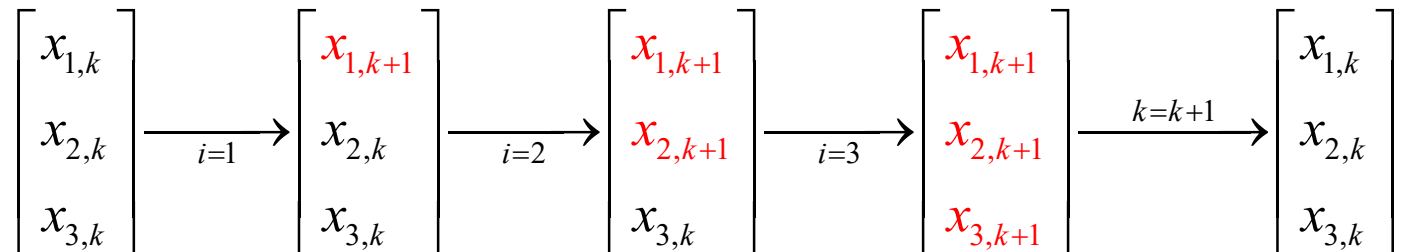
$$\|-(\mathbf{L} + \mathbf{D})^{-1}\mathbf{U}\| < 1$$

Как правило сходится для:

- симметричных и положительно определённых матриц;
- матриц с доминирующими диагональным элементом.

Может сходиться и при невыполнении этих условий

Схема вычислений



Решение рассчитывается итерационно

- в векторно-матричном виде

$$(\omega \mathbf{L} + \mathbf{D}) \mathbf{x}_{k+1} = -(\omega \mathbf{U} + (\omega - 1) \mathbf{D}) \mathbf{x}_k + \omega \mathbf{B}$$

- в поэлементном виде

$$x_{i,k+1} = (1 - \omega)x_{i,k} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} x_{j,k+1} - \sum_{j > i} a_{ij} x_{j,k} \right)$$

За счет коэффициента релаксации ω имеется возможность обеспечить более быструю сходимость

Выбор коэффициента релаксации

$$\omega = 1 + \left(\frac{\rho(\mathbf{T})}{1 + \sqrt{1 - \rho(\mathbf{T})^2}} \right)^2$$

$$\mathbf{T} = \mathbf{I} - \mathbf{D}^{-1} \mathbf{A}$$

$$\rho(\mathbf{T}) = \max |\lambda_i(\mathbf{T})|$$

где

$\rho(\cdot)$ — спектральный радиус матрицы

данный подход сформирован для трехдиагональных матрицы, а в общем случае выбор коэффициента релаксации — не тривиален

Для симметричных, положительно полуопределенных матриц метод гарантированно сходится при $0 < \omega < 2$

Методы Крыловского типа — это методы в которых, фактически, решается следующая оптимизационная задача

$$(\mathbf{Ax}, \mathbf{x}) - 2(\mathbf{b}, \mathbf{x}) \rightarrow \min$$

Метод сопряженных градиентов

\mathbf{A} – действительная, симметричная и положительно определённая матрица

Метод бисопряженных градиентов

\mathbf{A} – действительная и положительно определённая матрица

Стабилизированный метод бисопряженных градиентов

\mathbf{A} – действительная и положительно определённая матрица

Дополнительное условие выхода $\frac{\|\mathbf{r}_k\|}{\|\mathbf{b}\|} \leq \varepsilon$

увеличение применимости и повышение устойчивости (сходимости)

Метод сопряженных
градиентов

Инициализация:

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$$

$$\mathbf{z}_0 = \mathbf{r}_0$$

 k -ая итерация:

$$\alpha_k = \frac{(\mathbf{r}_{k-1}, \mathbf{r}_{k-1})}{(\mathbf{A}\mathbf{z}_{k-1}, \mathbf{z}_{k-1})}$$

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{z}_{k-1}$$

$$\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A}\mathbf{z}_{k-1}$$

$$\beta_k = \frac{(\mathbf{r}_k, \mathbf{r}_k)}{(\mathbf{r}_{k-1}, \mathbf{r}_{k-1})}$$

$$\mathbf{z}_k = \mathbf{r}_k + \beta_k \mathbf{z}_{k-1}$$

Метод бисопряженных
градиентов

Инициализация:

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0 \quad \mathbf{z}_0 = \mathbf{r}_0$$

$$\mathbf{p}_0 = \mathbf{r}_0 \quad \mathbf{s}_0 = \mathbf{r}_0$$

 k -ая итерация:

$$\alpha_k = \frac{(\mathbf{p}_{k-1}, \mathbf{r}_{k-1})}{(\mathbf{s}_{k-1}, \mathbf{A}\mathbf{z}_{k-1})}$$

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{z}_{k-1}$$

$$\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A}\mathbf{z}_{k-1}$$

$$\mathbf{p}_k = \mathbf{p}_{k-1} - \alpha_k \mathbf{A}^T \mathbf{s}_{k-1}$$

$$\beta_k = \frac{(\mathbf{p}_k, \mathbf{r}_k)}{(\mathbf{p}_{k-1}, \mathbf{r}_{k-1})}$$

$$\mathbf{z}_k = \mathbf{r}_k + \beta_k \mathbf{z}_{k-1}$$

$$\mathbf{s}_k = \mathbf{p}_k + \beta_k \mathbf{s}_{k-1}$$

Стабилизированный метод
бисопряженных градиентов

Инициализация:

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$$

$$\rho_0 = \alpha_0 = \omega_0 = 1$$

$$\tilde{\mathbf{r}}_0 = \mathbf{r}_0$$

$$\mathbf{v}_0 = \mathbf{p}_0 = 0$$

$$(\mathbf{q}, \mathbf{d}) = \sum_{i=1}^n q_i^* d_i$$

$$[\mathbf{q}, \mathbf{d}] = \sum_{i=1}^n q_i d_i$$

 k -ая итерация:

$$\rho_k = (\tilde{\mathbf{r}}, \mathbf{r}_{k-1})$$

$$\beta_k = \frac{\rho_k \alpha_{k-1}}{\rho_{k-1} \omega_{k-1}}$$

$$\mathbf{p}_k = \mathbf{r}_{k-1} + \beta_k (\mathbf{p}_{k-1} - \omega_{k-1} \mathbf{v}_{k-1})$$

$$\mathbf{v}_k = \mathbf{A}\mathbf{p}_k$$

$$\alpha_k = \frac{\rho_k}{(\tilde{\mathbf{r}}, \mathbf{v}_k)}$$

$$\mathbf{s}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{v}_k$$

$$\mathbf{t}_k = \mathbf{A}\mathbf{s}_k$$

$$\omega_k = \frac{[\mathbf{t}_k, \mathbf{s}_k]}{[\mathbf{t}_k, \mathbf{t}_k]}$$

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{p}_k + \omega_k \mathbf{s}_k$$

$$\mathbf{r}_k = \mathbf{s}_k - \omega_k \mathbf{t}_k$$