

Итерационные методы решения систем линейных алгебраических уравнений

Метод простых итераций

Договоримся о некоторых соглашениях, которые будут действовать до конца главы. Построим линейное пространство L_n одностробцовых матриц размера $n \times 1$ с введенными на нем естественным образом операциями сложения матриц и умножения матрицы на вещественное число. Введем каким-либо способом норму в этом пространстве. В дальнейшем одностробцовые матрицы, являющиеся элементами этого линейного пространства, будем называть, часто, *векторами*.

В дальнейшем, говоря о сходимости последовательности векторов к некоторому предельному вектору, под этим будем понимать сходимость по норме.

Пусть некоторый вектор X является приближением другого вектора \tilde{X} . Число $\|\tilde{X} - X\|$ будем называть *абсолютной погрешностью* приближенного вектора X .

Обратимся к СЛАУ, записанной в матричной форме

$$AX = B. \quad (1)$$

Эта система может быть преобразована (бесконечным числом способов) к эквивалентной ей системе вида

$$X = DX + F, \quad (2)$$

где D и F — некоторые новые матрицы размера $n \times n$ и $n \times 1$ соответственно. Построим, опираясь на (2), последовательность векторов $\{X^{(k)}\}_{k=0}^{\infty}$ следующим образом:

$$X^{(0)} \text{ — задан, } X^{(k)} = DX^{(k-1)} + F, \quad k = 1, 2, \dots \quad (3)$$

Процесс вычисления, в соответствии с определенным правилом, членов некоторой последовательности будем называть *итерационным*. Итерационный

процесс, проходящий по правилу (3), называется *методом простых итераций* (МПИ).

Заметим, что если МПИ (3) сходится к некоторому предельному вектору X^* , то X^* является решением системы (2). Сформулируем необходимое и достаточное условие сходимости МПИ (3) к решению системы (2) при любом начальном векторе $X^{(0)}$.

Теорема 1. Для того, чтобы МПИ (3) сходилась к решению системы (2) при любом выборе начального вектора $X^{(0)}$, необходимо и достаточно, чтобы все собственные числа матрицы D были по модулю меньше единицы.

Заметим, что условие теоремы 1 в реальных задачах проверять нецелесообразно в силу значительных вычислительных затрат.

Сформулируем теперь удобное для практического применения достаточное условие сходимости МПИ (3) и связь между k -м приближением системы (2) и абсолютной погрешностью этого приближения.

Предварительное замечание. Построим очевидным образом линейное пространство квадратных матриц n -го порядка. Введём в этом пространстве норму матрицы, подчинённую норме вектора из пространства L_n . Именно эта матричная норма и будет фигурировать в следующей теореме.

Теорема 2. Пусть существует такое число q , что матрица D из системы (2) удовлетворяет неравенству $\|D\| \leq q < 1$. Тогда при любом начальном векторе $X^{(0)}$ МПИ (3) сходится к единственному решению C задачи (2) и при всех натуральных k справедливы следующие оценки абсолютной погрешности вектора $X^{(k)}$:

1. $\|C - X^{(k)}\| \leq \frac{q}{1-q} \|X^{(k)} - X^{(k-1)}\|$ (апостериорная);
2. $\|C - X^{(k)}\| \leq \frac{q^k}{1-q} \|X^{(1)} - X^{(0)}\|$ (априорная).

Пусть, например, нужно получить решение системы (2) с абсолютной погрешностью, не превышающей ε . Возможны 2 варианта.

1. В соответствии с априорной оценкой теоремы 2 находим k как наименьшее целое решение неравенства $\frac{q^k}{1-q} \|X^{(1)} - X^{(0)}\| \leq \varepsilon$, после чего вычисляем $X^{(k)}$ по формулам (3).
2. В соответствии с апостериорной оценкой теоремы 2 вычисляем члены последовательности (3) до тех пор, пока при некотором значении k не будет выполнено неравенство $\|X^{(k)} - X^{(k-1)}\| \leq \frac{1-q}{q} \varepsilon$, после чего завершаем вычисления.

В обоих случаях вектор $X^{(k)}$ может быть взят как искомое приближение точного решения C системы (2) с абсолютной погрешностью, не превышающей ε .

Заметим, что априорная оценка, как правило, грубее апостериорной.

В условиях теоремы 1 сходимость (3) к решению C системы (2) гарантирована при любом начальном векторе $X^{(0)}$. Очевидно, однако, что чем ближе $X^{(0)}$ к C , тем меньше потребуется итераций для получения приближения с заданной точностью. При отсутствии дополнительной информации о решении C , в качестве $X^{(0)}$ берут, как правило, вектор F из системы (2).

МПИ можно использовать и для решения матричных уравнений. Пусть, например, в (1) X и B — неизвестная и заданная соответственно квадратные матрицы n -го порядка. Перейдём от (1) к эквивалентному матричному уравнению (2), где D и F — некоторые новые квадратные матрицы. Тогда МПИ (3) будет задавать уже последовательность квадратных матриц $\{X^k\}_{k=0}^{\infty}$, для которой очевидным образом могут быть переформулированы теоремы 1 и 2. Если в (1) B — это единичная матрица n -го порядка, то МПИ (3) при выполнении условия теоремы 1 будет сходиться к A^{-1} .

Метод Якоби

Рассмотрим один из частных случаев МПИ.

Представим матрицу A системы (1) в виде $A = G + H$, где G — диагональная матрица, а H — матрица, диагональные элементы которой — нулевые (такое представление, очевидно, единственно). Составим систему, эквивалентную (1):

$$GX = -HX + B. \quad (4)$$

Пусть среди диагональных элементов A (а значит, и G) нет нулей. Тогда матрица G обратима и система (4) эквивалентна системе

$$X = -G^{-1}HX + G^{-1}B. \quad (5)$$

Основанный на (5) итерационный процесс

$$X^{(0)} \text{ — задан, } X^{(k)} = -G^{-1}HX^{(k-1)} + G^{-1}B, \quad k = 1, 2, \dots \quad (6)$$

называется *методом Якоби*. Заметим, что метод Якоби частным случаем МПИ (3), где $D = -G^{-1}H$, $F = G^{-1}B$.

Учитывая, что

$$G = \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}^{-1} \end{pmatrix},$$

перепишем итерационный процесс (6) в координатной форме:

$$x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)} \text{ — заданы,}$$

$$\begin{cases} x_1^{(k)} = \left(b_1 - a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \dots - a_{1n}x_n^{(k-1)} \right) / a_{11}, \\ x_2^{(k)} = \left(b_2 - a_{21}x_1^{(k-1)} - a_{23}x_3^{(k-1)} - \dots - a_{2n}x_n^{(k-1)} \right) / a_{22}, \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ x_n^{(k)} = \left(b_n - a_{n1}x_1^{(k-1)} - a_{n3}x_2^{(k-1)} - \dots - a_{nn}x_n^{(k-1)} \right) / a_{nn}, \end{cases} \quad (7)$$
$$k = 1, 2, \dots$$

Сформулируем достаточное условие сходимости метода Якоби.

Теорема 3. Если матрица A в системе (1) является матрицей с диагональным преобладанием элементов, то метод Якоби (6) сходится к решению системы (5).

Доказательство. Обозначим: $S_m = (\sum_{k=1}^n |a_{mk}|) - |a_{mm}|$, где $m = \overline{1, n}$. Диагональное преобладание элементов в матрице A означает, что для любого m от 1 до n выполняется неравенство $S_m < |a_{mm}|$. Разделим обе части этого неравенства на $|a_{mm}|$ (это можно сделать в силу неотрицательности $|a_{mm}|$, следующей из этого же неравенства): $\frac{S_m}{|a_{mm}|} < 1$. Рассмотрим теперь матрицу

$$-G^{-1}H = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{pmatrix},$$

играющую в (6) роль матрицы D в (3) и вычислим её норму, подчинённую равномерной норме вектора:

$$\| -G^{-1}H \|_\infty = \max_{1 \leq m \leq n} \left\{ \left(\sum_{k=1}^n \frac{|a_{mk}|}{|a_{mm}|} \right) - 1 \right\} = \max_{1 \leq m \leq n} \left\{ \frac{S_m}{|a_{mm}|} \right\}.$$

Но из доказанного выше неравенства $\frac{S_m}{|a_{mm}|} < 1$, верного при $m = \overline{1, n}$, следует, что и

$$\max_{1 \leq m \leq n} \left\{ \frac{S_m}{|a_{mm}|} \right\} < 1,$$

T. e. $\| -G^{-1}H \|_{\infty} < 1$.

Остаётся лишь положить в теореме 2: $D = -G^{-1}H$, $q = \|-G^{-1}H\|_\infty$ (под нормой квадратной матрицы, фигурирующей в теореме, будем понимать норму, подчинённую равномерной норме вектора). Тогда условие теоремы 2 будет выполнено, а это означает, что, в соответствии с этой теоремой, метод (6) будет сходиться к решению системы (5), что и требовалось доказать.

Заметим, что в условиях теоремы 3 к методу Якоби применимы априорная и апостериорная оценки абсолютной погрешности k -го приближения, сформулированные в теореме 2.

При использовании метода Якоби (6) в качестве $X^{(0)}$ берут, как правило, вектор $G^{-1}B = \left(\frac{b_1}{a_{11}} \quad \frac{b_2}{a_{22}} \quad \dots \quad \frac{b_n}{a_{nn}} \right)^T$.

Метод Якоби можно использовать и для обращения матриц с диагональным преобладанием элементов. Для этого нужно представить обрабатываемую матрицу A в виде $G + H$ (точно таким же образом, как это делалось ранее), подставить G и H в (6), а в качестве $X^{(0)}$ взять, например, квадратную матрицу G^{-1} . Тогда метод (6) будет сходиться к A^{-1} и для него будут справедливы априорная и апостериорная оценки погрешностей, аналогичные оценкам из теоремы 2.

Метод Зейделя

Метод Зейделя заключается в таком видоизменении МПИ (3), при котором для получения x_i^k — i -й компоненты вектора $X^{(k+1)}$ используются уже найденные на этом, т. е. k -м шаге новые значения первых $i - 1$ компонент. Так, например, координатная форма записи метода Зейделя, основанная на методе Якоби (7), будет выглядеть следующим образом:

$$\begin{aligned} x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)} & \text{ — заданы,} \\ \left\{ \begin{array}{l} x_1^{(k)} = \left(b_1 - a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \dots - a_{1n}x_n^{(k-1)} \right) / a_{11}, \\ x_2^{(k)} = \left(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k-1)} - \dots - a_{2n}x_n^{(k-1)} \right) / a_{22}, \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ x_n^{(k)} = \left(b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \dots - a_{nn}x_n^{(k)} \right) / a_{nn}, \end{array} \right. \quad (8) \\ k = 1, 2, \dots \end{aligned}$$

Вплоть до окончания текущего параграфа под методом Зейделя мы будем понимать метод Зейделя, основанный на методе Якоби.

Запишем метод Зейделя (8) в матричной форме. Будем пользоваться теми же обозначениями матриц, что и в предыдущем параграфе. Представим матрицу H в виде $H = H_L + H_R$, где H_L и H_R — нижне- и верхнетреугольная матрицы соответственно, имеющие нулевые диагонали. Запишем теперь, пользуясь введёнными обозначениями, систему (8) в виде

$$X^{(k)} = -G^{-1} \left(H_R X^{(k-1)} + H_L X^{(k)} \right) + G^{-1}B.$$

Домножим обе части этого равенства на матрицу G слева и перенесём слагаемое $-H_L X^{(k)}$ в левую часть. Приходим к равенству

$$(H_L + G) X^{(k)} = -H_R X^{(k-1)} + B,$$

из которого, путём домножения на $(H_L + G)^{-1}$ слева, получаем матричную форму записи метода Зейделя(8):

$$X^{(0)} - \text{задан,} \\ X^{(k)} = -(H_L + G)^{-1} H_R X^{(k-1)} + (H_L + G)^{-1} B, \quad k = 1, 2, \dots \quad (9)$$

Заметим, что метод (9) является частным случаем МПИ (3), где

$$D = -(H_L + G)^{-1} H_R, \quad F = (H_L + G)^{-1} B.$$

Это означает, что к методу Зейделя применимы теоремы 1 и 2. Однако, для использования оценок теоремы 2, необходимо обратить треугольную матрицу $H_L + G$ и перемножить матрицы $(H_L + G)^{-1}$ и H_R , что связано с дополнительными вычислительными затратами.

Сформулируем теперь два удобных для практического применения достаточных условия сходимости метода Зейделя в виде теорем 4 и 5.

Теорема 4. Если матрица A в системе (1) является матрицей с диагональным преобладанием элементов, то метод Зейделя (8) сходится к решению системы (5), причём быстрее, чем соответствующий ему метод Якоби.

Заметим, что в методе Зейделя допустимо использование оценок погрешности метода Якоби, т. е. оценок теоремы 2, где $D = -G^{-1}H$. Разумеется, они будут заведомо грубее оценок, соответствующих $D = -(H_L + G)^{-1} H_R$.

Перед формулировкой теоремы 5 введём несколько понятий.

Вещественная квадратная матрица A называется *симметричной*, если выполняется равенство $A^T = A$.

Симметричная матрица n -го порядка A называется *положительно определённой*, если для любой ненулевой матрицы X размера $n \times 1$ единственный элемент одноэлементной матрицы $X^T A X$ положителен.

СЛАУ (1) называется *нормальной*, если матрица A — положительная симметрично определённая.

Теорема 5. Если СЛАУ (1) — нормальная, то метод Зейделя (8) сходится к решению системы (1) при любом выборе начального вектора $X^{(0)}$.

Можно доказать, что если система (1) является определённой, то эквивалентная ей система $A^T A X = A^T B$ является нормальной. Это означает, что любую определённую систему вида (1) можно, если потребуется, преобразовать к такой системе, к которой применим метод Зейделя.