

Модуль 11. Итерационные методы линейной алгебры для решения задач большой размерности (конспект без доказательств)

Метод простой итерации. Метод минимальных невязок.

Метод с чебышевским набором параметров. Метод сопряженных градиентов.

Расчет параметров, теоремы о сходимости и построение сходящихся методов на основе оценок собственных чисел

11.1. Метод простой итерации

Рассмотрим СЛАУ (систему линейных алгебраических уравнений) вида

$$Ax = b, \quad (11.1)$$

где $x \in R^n$, $b \in R^n$, $A (n \times n)$, $\det A \neq 0$ (невырожденная матрица).

Через x^* обозначим точное решение системы, $x^* \in R^n$.

Методом простой итерации называют явный стационарный итерационный метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau} + Ax^{(s)} = b \quad (11.2)$$

где τ – число (постоянный параметр метода), $x^{(0)} \in R^n$ – начальное приближение для запуска итераций (его можно выбирать любым), $s = 0, 1, \dots$ – номер шага метода.

Запись метода в виде (11.2) называется *канонической*. Для расчетов вместо (11.2) используется формула

$$x^{(s+1)} = x^{(s)} + \tau \cdot (b - Ax^{(s)}) = x^{(s)} - \tau \cdot r^{(s)} \quad (11.3)$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$.

Основные свойства метода описывают следующие теоремы.

Теорема 1. При решении СЛАУ (11.1) **методом простой итерации** (11.2) оценка погрешности метода на шаге s имеет вид

$$\|z^{(s)}\|_2 \leq \|G(\tau)\|_2^s \|z^{(0)}\|_2 \quad (11.4)$$

достаточным условием сходимости метода является условие

$$\|G(\tau)\|_2 < 1. \quad (11.5)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $G(\tau) = E - \tau A$ – переходная матрица метода,

$\|G(\tau)\|_2$ – норма переходной матрицы, подчиненная евклидовой норме вектора,

$\|G(\tau)\|_2^s$ – норма переходной матрицы, подчиненная евклидовой норме вектора и

возведенная в степень s , $\|z^{(s)}\|_2$ и $\|z^{(0)}\|_2$ – евклидова норма векторов.

Теорема 2. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) оценка погрешности метода на шаге s имеет вид

$$\|z^{(s)}\|_2 \leq (\max\{|1 - \lambda_1 \tau|, |1 - \lambda_n \tau|\})^s \|z^{(0)}\|_2 \quad (11.6)$$

необходимым и достаточным условием сходимости метода является

$$\tau \in (0, \frac{2}{\lambda_n}) \quad (11.7)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Комментарии

1) Выражение в круглых скобках (см. (11.6)) есть норма переходной матрицы $G(\tau)$, подчиненная евклидовой норме вектора:

$$\|G(\tau)\|_2 = \max\{|1 - \lambda_1 \tau|, |1 - \lambda_n \tau|\} \quad (11.8)$$

Запись (11.8) следует из того, что собственные числа матрицы $G(\tau)$ можно выразить через собственные числа матрицы A как

$$\lambda_i(G) = 1 - \tau \cdot \lambda_i(A), i = 1, \dots, n$$

и норма симметричной матрицы $G(\tau)$, подчиненная евклидовой норме вектора, определяется модулями ее «крайних» собственных чисел, а именно

$$|\lambda_1(G)| = |1 - \lambda_1(A) \cdot \tau|, |\lambda_n(G)| = |1 - \lambda_n(A) \cdot \tau|.$$

2) условия сходимости (11.7) получены из условия $\|G(\tau)\|_2 < 1$ как решение системы следующих двух неравенств:

$$|1 - \lambda_1 \tau| < 1, |1 - \lambda_n \tau| < 1.$$

Теорема 3. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) оптимальным является

$$\tau_{opt}^* = \frac{2}{\lambda_1 + \lambda_n} \quad (11.9)$$

для которого оценка погрешности метода на шаге s имеет вид

$$\|z^{(s)}\|_2 \leq \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \|z^{(0)}\|_2 \quad (11.10)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Через μ_A обозначено число обусловленности матрицы A , определяемое на основе нормы матрицы, подчиненной евклидовой норме вектора (далее кратко – *число обусловленности, определяемое на основе евклидовой нормы*):

$$\mu_A = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_n}{\lambda_1} \quad (11.11)$$

Комментарии

Оптимальным считается такое значение τ , при котором метод (11.2) сходится и оценка погрешности метода на шаге s (см. (11.4) или (11.6)) является в некотором смысле оптимальной. В данном случае для отыскания оптимального значения τ ставится задача минимизации

$$\|G(\tau)\|_2 \xrightarrow{\tau \in \left(0, \frac{2}{\lambda_n}\right)} \min \quad (11.12)$$

Минимальное значение функционала $\|G(\tau)\|_2$ достигается при $\tau_{opt}^* = \frac{2}{\lambda_1 + \lambda_n}$.

Минимальным значением является $\|G(\tau_{opt}^*)\|_2$, и в ходе доказательства теоремы эта величина будет найдена как

$$\begin{aligned} \min_{\tau \in R} \|G(\tau)\|_2 &= \min_{\tau \in R} \max \{ |1 - \lambda_1 \tau|, |1 - \lambda_n \tau| \} = \\ &= \max \{ |1 - \lambda_1 \tau_{opt}^*|, |1 - \lambda_n \tau_{opt}^*| \} = \\ &= \max \left\{ \left| 1 - \frac{\lambda_1 \cdot 2}{\lambda_1 + \lambda_n} \right|, \left| 1 - \frac{\lambda_n \cdot 2}{\lambda_1 + \lambda_n} \right| \right\} = \frac{\mu_A - 1}{\mu_A + 1} \end{aligned} \quad (11.13)$$

откуда следует (11.10).

Построение метода на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 2, чтобы построить сходящийся метод простой итерации, нужно знать максимальное собственное число (см. (11.7)), чтобы оценить погрешность метода – знать минимальное и максимальное собственные числа (см. (11.6)).

Покажем, как построить сходящийся метод **на основе оценок собственных чисел**.

Пусть собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, а именно, положительные числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$$\lambda_i(A) \in [M_{\min}, M_{\max}], \quad i = 1, \dots, n.$$

Тогда для λ_1 и λ_n – минимального и максимального собственных чисел – верно

$$0 < M_{\min} \leq \lambda_1 \leq \lambda_n \leq M_{\max} \quad (11.14)$$

(оценку вида (11.14) называют **оценкой границ спектра**).

Теорема 4. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) в случае, когда известна оценка границ спектра $0 < M_{\min} \leq \lambda_1 \leq \lambda_n \leq M_{\max}$, достаточным условием сходимости метода является

$$\tau \in (0, \frac{2}{M_{\max}}) \quad (11.15)$$

Оценка погрешности метода на шаге s имеет вид

$$\|z^{(s)}\|_2 \leq (\max\{|1 - \tau M_{\min}|, |1 - \tau M_{\max}|\})^s \|z^{(0)}\|_2 \quad (11.16)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Доказательство

Используем Теорему 2, условие сходимости (11.7) и оценку погрешности (11.6).

Так как $0 < \frac{2}{M_{\max}} < \frac{2}{\lambda_n}$, интервал $(0, \frac{2}{M_{\max}})$ принадлежит области сходимости метода.

Далее для каждого фиксированного значения τ очевидна оценка

$$\max\{|1 - \lambda_1 \tau|, |1 - \lambda_n \tau|\} \leq \max_{\lambda \in [\lambda_1, \lambda_n]} |1 - \lambda \tau| \leq \max_{\lambda \in [M_{\min}, M_{\max}]} |1 - \lambda \tau|$$

Здесь максимальное из двух неизвестных чисел $|1 - \lambda_1 \tau|, |1 - \lambda_n \tau|$ оценивается максимальным значением функции $|1 - \lambda \tau|$ с аргументом λ , пробегающим отрезок $\lambda \in [\lambda_1, \lambda_n]$, и затем оценивается максимальным значением той же функции $|1 - \lambda \tau|$ с аргументом λ , пробегающим отрезок $\lambda \in [M_{\min}, M_{\max}]$, который включает в себя и значения λ_1 и λ_n , и отрезок $\lambda \in [\lambda_1, \lambda_n]$.

Известно, что максимальное значение непрерывной функции на отрезке достигается либо на концах отрезка, либо в точке локального максимума. Так как при фиксированном τ функция $|1 - \lambda \tau|$ с аргументом $\lambda \in R$ непрерывна и не имеет локального максимума, максимальное значение функции $|1 - \lambda \tau|$ достигается на концах отрезка. Поэтому

$$\max_{\lambda \in [M_{\min}, M_{\max}]} |1 - \lambda \tau| = \max\{|1 - \tau M_{\min}|, |1 - \tau M_{\max}|\}$$

Таким образом, получена оценка

$$\max \{ |1 - \lambda_1 \tau|, |1 - \lambda_n \tau| \} \leq \max \{ |1 - \tau \cdot M_{\min}|, |1 - \tau \cdot M_{\max}| \} \quad (11.17)$$

Подставляя (11.17) в (11.6), получим (11.16).

Выбор оптимального параметра на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 3, чтобы построить метод простой итерации с оптимальной оценкой сходимости, нужно знать минимальное и максимальное собственные числа матрицы. **Рассмотрим возможности оптимизации на основе оценок собственных чисел.**

Пусть собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, а именно, положительные числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$$\lambda_i(A) \in [M_{\min}, M_{\max}], \quad i = 1, \dots, n.$$

Тогда для λ_1 и λ_n – минимального и максимального собственных чисел – верно

$$0 < M_{\min} \leq \lambda_1 \leq \lambda_n \leq M_{\max}$$

и для числа обусловленности μ_A , определяемого на основе евклидовой нормы, верно

$$1 \leq \mu_A = \frac{\lambda_n}{\lambda_1} \leq \frac{M_{\max}}{M_{\min}} = M \quad (1.18)$$

Теорема 5. При решении СЛАУ (11.1) с **симметричной, положительно определенной** матрицей $A = A^T > 0$ **методом простой итерации** в случае, когда известна оценка границ спектра $0 < M_{\min} \leq \lambda_1 \leq \lambda_n \leq M_{\max}$, метод (11.2) с параметром

$$\tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}} \quad (11.19)$$

имеет оптимальные свойства и сходится с оценкой

$$\|z^{(s)}\|_2 \leq \left(\frac{M-1}{M+1} \right)^s \|z^{(0)}\|_2 \quad (11.20)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них, число M является верхней оценкой числа обусловленности μ_A (определенного в евклидовой норме).

Доказательство

Используем Теорему 4, условие сходимости (11.15) и оценку погрешности (11.16).

Во-первых, значение $\tau = \tilde{\tau}^*$ принадлежит области сходимости метода:

$$\tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}} \in \left(0, \frac{2}{M_{\max}}\right) \subset \left(0, \frac{2}{\lambda_n}\right).$$

Во-вторых, подстановкой значения $\tau = \tilde{\tau}^*$ можно показать, что

$$\max \{ |1 - \tilde{\tau}^* M_{\min}|, |1 - \tilde{\tau}^* M_{\max}| \} = \frac{M_{\max} - M_{\min}}{M_{\max} + M_{\min}} = \frac{M - 1}{M + 1}$$

поэтому из (11.16) следует (11.20).

Комментарий

Оптимальные свойства метода вытекают из следующих обстоятельств. Рассмотрим класс методов (11.2) с параметром (11.15):

$$\tau \in (0, \frac{2}{M_{\max}})$$

и оценкой погрешности (11.16):

$$\|z^{(s)}\|_2 \leq \left(\max \{ |1 - \tau M_{\min}|, |1 - \tau M_{\max}| \} \right)^s \|z^{(0)}\|_2$$

Оптимальным считается такое значение τ , при котором метод сходится и оценка погрешности метода на шаге s (см. (11.16)) является в некотором смысле оптимальной. В данном случае для отыскания оптимального значения τ ставится задача минимизации

$$\max \{ |1 - \tau M_{\min}|, |1 - \tau M_{\max}| \} \min_{\tau \in \left(0, \frac{2}{M_{\max}}\right)} \quad (11.21)$$

Используя графики функций $|1 - \tau M_{\min}|$ и $|1 - \tau M_{\max}|$ с аргументом $\tau \in R$, несложно показать, что минимальное значение функционала достигается при

$$\tau = \tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}}$$

При этом минимальное значение функционала

$$\max \{ |1 - \tau M_{\min}|, |1 - \tau M_{\max}| \}$$

может быть найдено подстановкой значения $\tau = \tilde{\tau}^*$:

$$\begin{aligned} \min_{\tau \in R} \max \{ |1 - \tau M_{\min}|, |1 - \tau M_{\max}| \} = \\ = \max \{ |1 - \tilde{\tau}^* \cdot M_{\min}|, |1 - \tilde{\tau}^* \cdot M_{\max}| \} = \frac{M - 1}{M + 1} \end{aligned} \quad (11.22)$$

Комментарий к параграфу 11.1

В формулировках теорем 1-5 погрешность метода на шаге s оценивается начальной погрешностью, см. (11.4), (11.6), (11.10), (11.16), 11.20), все оценки в евклидовой норме. В свою очередь, **погрешность метода на начальном шаге можно оценить по начальной невязке**, используя норму обратной матрицы или ее оценку:

$$\|z^{(0)}\|_2 \leq \|A^{-1}\|_2 \cdot \|r^{(0)}\|_2, \text{ где } r^{(0)} = Ax^{(0)} - b \text{ и ее можно вычислить.}$$

11.2. Метод минимальных невязок

Методом минимальных невязок называют явный нестационарный итерационный метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau_s} + Ax^{(s)} = b \quad (11.23)$$

где параметр τ_s (число) на каждом шаге метода выбирают так, чтобы для уже вычисленного приближения $x^{(s)} \in R^n$ следующая невязка $r^{(s+1)} \in R^n$, то есть $r^{(s+1)} = Ax^{(s+1)} - b$, была минимальной.

Если $A = A^T > 0$ и невязка измеряется в евклидовой норме, следует выбирать

$$\tau_s = \frac{(Ar^{(s)}, r^{(s)})}{(Ar^{(s)}, Ar^{(s)})} \quad (11.24)$$

Запись метода в виде (11.23), (11.24) называется *канонической*. Для расчетов вместо (11.23) используется формула

$$x^{(s+1)} = x^{(s)} + \tau_s \cdot (b - Ax^{(s)}) = x^{(s)} - \tau_s \cdot r^{(s)} \quad (11.25)$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$, τ_s – параметр шага $s + 1$.

Теорема 6. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом минимальных невязок** (11.23), (11.24) метод сходится и для погрешности метода на шаге s верна оценка

$$\|z^{(s)}\|_2 \leq \mu_A \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \|z^{(0)}\|_2 \quad (11.26)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них. Через μ_A обозначено число обусловленности матрицы A , определяемое на основе евклидовой нормы:

$$\mu_A = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_n}{\lambda_1}.$$

Оценка погрешности на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 6, для решения СЛАУ методом минимальных невязок знать собственные числа матрицы A не требуется, но для оценки погрешности метода – см. (11.26) – такая информация необходима.

Теорема 7. Пусть собственные числа *симметричной, положительно определенной* матрицы $A = A^T > 0$ неизвестны, но известны их оценки, то есть числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что $\lambda_i(A) \in [M_{\min}, M_{\max}]$, $i = 1, \dots, n$. Тогда для λ_1 и λ_n верно $0 < M_{\min} \leq \lambda_1 \leq \lambda_n \leq M_{\max}$, а для числа обусловленности μ_A (основанного на евклидовой норме) верна оценка

$$1 \leq \mu_A = \frac{\lambda_n}{\lambda_1} \leq \frac{M_{\max}}{M_{\min}} = M \quad (11.27)$$

Для оценки погрешности метода минимальных невязок можно использовать неравенство

$$\|z^{(s)}\|_2 \leq M \left(\frac{M-1}{M+1} \right)^s \|z^{(0)}\|_2 \quad (11.28)$$

Доказательство

Так как

$$1 \leq \mu_A \leq M \text{ и}$$

$$0 \leq \frac{\mu_A - 1}{\mu_A + 1} \leq \frac{M - 1}{M + 1} < 1$$

(оценивается возрастающая функция вида $\varphi(x) = \frac{x-1}{x+1}$ при $x \geq 1$)

из (11.26) получим

$$\|z^{(s)}\|_2 \leq \mu_A \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \|z^{(0)}\|_2 \leq M \left(\frac{M - 1}{M + 1} \right)^s \|z^{(0)}\|_2$$

что и требовалось доказать.

11.3. Метод с чебышевским набором параметров (k параметров)

Рассмотрим СЛАУ (11.1) с *симметричной, положительно определенной* матрицей A . Пусть $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное и λ_n – максимальное из них.

Явным нестационарным итерационным методом с чебышевским набором параметров (k параметров) называют метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau_s} + Ax^{(s)} = b \quad (11.29)$$

где k – натуральное число (параметр метода), $\tau_0, \tau_1, \dots, \tau_{k-1}$ – вещественные параметры метода (их количество равно k), $x^{(0)} \in R^n$ – начальное приближение, которое можно выбрать любым, $s = 0, 1, \dots$ – номер шага метода.

Значения параметров $\tau_0, \tau_1, \dots, \tau_{k-1}$ определяют как

$$\tau_s = \left(\frac{\lambda_1 + \lambda_n}{2} + \frac{\lambda_n - \lambda_1}{2} \cdot \cos \left(\frac{\pi}{2k} (1 + 2s) \right) \right)^{-1}, s = 0, \dots, k-1 \quad (11.30)$$

то есть (при $\lambda_1 < \lambda_n$) как величины, обратные корням полинома Чебышева (полинома степени k , наименее уклоняющегося от нуля на отрезке $[\lambda_1, \lambda_n]$, $0 < \lambda_1 < \lambda_n$ в классе полиномов степени k со свободным слагаемым, равным 1). Случай $\lambda_1 = \lambda_n$ не является типичным и заслуживает отдельного рассмотрения, так как в данном случае все собственные числа одинаковы: $0 < \lambda_1 = \lambda_2 = \dots = \lambda_{n-1} = \lambda_n$.

Первые k шагов метода проводятся с использованием параметров $\tau_0, \tau_1, \dots, \tau_{k-1}$. При вычислении $x^{(1)}$ используют τ_0 , при вычислении $x^{(2)}$ используют τ_1 , ... при вычислении $x^{(k)}$ используют τ_{k-1} .

На последующих шагах метода указанные выше k параметров используют циклически: при вычислении $x^{(k+1)}$ снова используется τ_0 , при вычислении $x^{(k+2)}$ используем τ_1 , ... при вычислении $x^{(2k)}$ используется τ_{k-1} и т.д.

Таким образом, метод (11.29), (11.30) строит последовательность

$$\begin{aligned} & x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(k-1)}, x^{(k)} .. \\ & x^{(k+1)}, x^{(k+2)}, x^{(k+3)}, \dots, x^{(2k-1)}, x^{(2k)} .. \\ & x^{(2k+1)}, x^{(2k+2)}, x^{(2k+3)}, \dots, x^{(3k-1)}, x^{(3k)} \dots \end{aligned} \quad (11.31)$$

Приближенными решениями задачи (11.1) следует считать следующие элементы последовательности:

$$x^{(0)}, x^{(k)}, x^{(2k)}, x^{(3k)}, x^{(4k)} \dots x^{(Nk)} \dots \quad (11.32)$$

Остальные элементы последовательности имеют вспомогательное значение.

Запись метода в виде (11.29), (11.30) называется *канонической*. Для расчетов вместо (11.29) используется формула

$$x^{(s+1)} = x^{(s)} + \tau_s \cdot (b - Ax^{(s)}) = x^{(s)} - \tau_s \cdot r^{(s)} \quad (11.33)$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$, τ_s – параметр шага $s + 1$.

Свойства метода описывает следующая теорема.

Теорема 8. При решении задачи (11.1) с симметричной, положительно определенной матрицей $A = A^T > 0$ методом (11.29) с параметрами (11.30) оценка погрешности метода на шаге k имеет вид

$$\|z^{(k)}\|_2 \leq \frac{2\rho^k}{1 + \rho^{2k}} \|z^{(0)}\|_2 \quad (11.34)$$

Оценка погрешности метода на шаге Nk (N – количество циклов) имеет вид

$$\|z^{(Nk)}\|_2 \leq \left(\frac{2\rho^k}{1+\rho^{2k}} \right)^N \|z^{(0)}\|_2 \quad (11.35)$$

и метод сходится в следующем смысле:

$$\forall x^{(0)} \in R^n \quad \lim_{N \rightarrow +\infty} \|z^{(Nk)}\|_2 = 0. \quad (11.36)$$

(к решению x^* сходятся элементы последовательности (11.32)).

В формулах (11.34)-(11.36) $z^{(Nk)} = x^{(Nk)} - x^*$ – погрешность метода на шаге Nk , $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq \dots \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное и λ_n – максимальное из них.

Значение ρ в формулах (11.34), (11.35) определено как

$$\rho = \frac{\sqrt{\mu_A} - 1}{\sqrt{\mu_A} + 1} \quad (11.37)$$

Через μ_A обозначено число обусловленности матрицы A , определяемое на основе евклидовой нормы:

$$\mu_A = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_n}{\lambda_1} \quad (11.38)$$

Оптимальное свойство метода состоит в следующем: среди всех методов вида (11.29) метод (11.29) с параметрами (11.30) дает **наилучшую гарантию убывания погрешности через k шагов**, справедливую **для всех симметричных положительно определенных матриц $A = A^T > 0$, собственные числа которых расположены в диапазоне $[\lambda_1, \lambda_n]$, $0 < \lambda_1 < \lambda_n$** .

Построение метода на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 8, для построения метода и оценки погрешности нужно знать минимальное и максимальное собственные числа матрицы A .

Предположим, что собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, то есть известны числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$$\lambda_i(A) \in [M_{\min}, M_{\max}], \quad i = 1, \dots, n.$$

Рассмотрим метод (11.29) с натуральным параметром k и вещественными параметрами

$$\tilde{\tau}_s = \left(\frac{M_{\min} + M_{\max}}{2} + \frac{M_{\max} - M_{\min}}{2} \cdot \cos \left(\frac{\pi}{2k} (1 + 2s) \right) \right)^{-1}, \quad s = 0, \dots, k-1 \quad (11.39)$$

Теорема 9. При решении задачи (11.1) с *симметричной, положительно определенной* матрицей A методом (11.29) с параметрами (11.39) оценка погрешности метода на шаге k имеет вид

$$\|z^{(k)}\|_2 \leq \frac{2\tilde{\rho}^k}{1 + \tilde{\rho}^{2k}} \|z^{(0)}\|_2 \quad (11.40)$$

Оценка погрешности метода на шаге Nk (N есть количество циклов) имеет вид

$$\|z^{(Nk)}\|_2 \leq \left(\frac{2\tilde{\rho}^k}{1 + \tilde{\rho}^{2k}} \right)^N \|z^{(0)}\|_2 \quad (11.41)$$

и метод сходится в следующем смысле:

$$\forall x^{(0)} \in R^n \quad \lim_{N \rightarrow +\infty} \|z^{(Nk)}\|_2 = 0. \quad (11.42)$$

(к решению x^* сходятся элементы последовательности (11.32)).

В формулах (11.40)-(11.42) значение $\tilde{\rho}$ определено как

$$\tilde{\rho} = \frac{\sqrt{M} - 1}{\sqrt{M} + 1}, \text{ где } M = \frac{M_{\max}}{M_{\min}}.$$

Оптимальное свойство метода (11.29), (11.39) состоит в следующем: среди всех методов вида (11.29) метод с параметрами (11.39) дает **наилучшую гарантию убывания погрешности за k шагов**, справедливую для **всех симметричных положительно определенных матриц** $A = A^T > 0$, **собственные числа которых расположены в диапазоне** $[M_{\min}, M_{\max}]$, $0 < M_{\min} < M_{\max}$.

Комментарий

Если собственные числа $A = A^T > 0$ расположены в диапазоне $[\lambda_1, \lambda_n]$, $0 < \lambda_1 < \lambda_n$, метод (11.29) с параметрами (1.30) обеспечит лучшую гарантию убывания погрешности, чем метод (11.29) с параметрами (11.39).

Действительно, в данном случае $1 < \mu_A \leq M$ и $0 < \frac{\sqrt{\mu_A} - 1}{\sqrt{\mu_A} + 1} \leq \frac{\sqrt{M} - 1}{\sqrt{M} + 1} < 1$

(оценивается возрастающая функция вида $\varphi(x) = \frac{\sqrt{x} - 1}{\sqrt{x} + 1}$ при $x > 1$).

Поэтому для множителей из оценок (11.34), (11.35) и (11.40), (11.41) верно

$$0 < \frac{2\rho^k}{1 + \rho^{2k}} \leq \frac{2\tilde{\rho}^k}{1 + \tilde{\rho}^{2k}} < 1, \quad 0 < \left(\frac{2\rho^k}{1 + \rho^{2k}} \right)^N \leq \left(\frac{2\tilde{\rho}^k}{1 + \tilde{\rho}^{2k}} \right)^N < 1.$$

При этом метод (11.29) с параметрами (11.30) (тот, что лучше сходится) требует знания собственных чисел, а метод с параметрами (11.39) использует их оценки.

11.4. Метод сопряженных градиентов

Сведение решения СЛАУ к решению задачи оптимизации. Сопряженные направления и их свойства. Сведение k -мерной (многомерной) задачи оптимизации к решению k одномерных оптимизационных задач.

Рассмотрим систему линейных алгебраических уравнений (СЛАУ) с симметричной положительно определенной матрицей

$$Ax = b \quad (11.43)$$

где $x \in R^n$, $b \in R^n$, $A (n \times n)$, $A = A^T > 0$.

Через x^* обозначим точное решение СЛАУ, $x^* \in R^n$.

Введем функционал $F(x) = (Ax, x) - 2(b, x)$ и рассмотрим задачу оптимизации

$$F(x) = (Ax, x) - 2(b, x) \rightarrow \min \quad (11.44)$$

Используя свойства $F(x)$, несложно показать, что единственным решением задачи (11.44) является $x^* \in R^n$, который является единственным решением задачи (11.43).

Таким образом, методы решения многомерной (n -мерной) задачи оптимизации (11.44) могут быть использованы для отыскания решения СЛАУ (11.43).

Определение 1. Пусть $A = A^T > 0$. Векторы (направления) $h', h'' \in R^n$, $h', h'' \neq 0$ называются **сопряженными** относительно A , если $(Ah', h'') = 0$.

Определение 2. Пусть $A = A^T > 0$. Ненулевые векторы (направления) $h^{(0)}, \dots, h^{(k-1)} \in R^n$ называются **взаимно сопряженными** относительно A , если $(Ah^{(i)}, h^{(j)}) = 0$ для $\forall i, j = 0, k-1, i \neq j$

Способ решения задачи оптимизации (11.44) опирается на следующие результаты.

Утверждение. Если ненулевые $h^{(0)}, \dots, h^{(k-1)} \in R^n$ взаимно сопряжены относительно $A = A^T > 0$, тогда они линейно независимы. Если ненулевые $h^{(0)}, \dots, h^{(n-1)} \in R^n$ взаимно сопряжены относительно $A = A^T > 0$, тогда они образуют базис в R^n

Теорема 10. Пусть $h^{(0)}, \dots, h^{(k-1)} \in R^n$, $k \leq n$, представляют собой ненулевые взаимно сопряженные направления относительно $A = A^T > 0$ (они линейно независимы) и пусть $x^{(0)} \in R^n$.

Рассматривается многообразие $L_k(x_0, h^{(0)}, \dots, h^{(k-1)})$ размерности k , элементами которого являются линейные комбинации сопряженных направлений и вектора $x^{(0)} \in R^n$, то есть векторы вида

$$x = x^{(0)} + \alpha_0 h^{(0)} + \dots + \alpha_{k-1} h^{(k-1)}, \text{ где } \alpha_i \in R, i = 0, \dots, k-1.$$

Решением многомерной (k -мерной) задачи оптимизации

$$F(x) = \min_{x \in L_k(x^{(0)}, h^{(0)}, \dots, h^{(k-1)})} \quad (11.45)$$

на указанном выше многообразии является вектор

$$x^{(k)} = x^{(0)} + \alpha_0^* h^{(0)} + \dots + \alpha_{k-1}^* h^{(k-1)}$$

где коэффициенты $\alpha_s^*, s = 0, \dots, k-1$, есть решения k одномерных задач оптимизации

$$\alpha_s^2(Ah^{(s)}, h^{(s)}) + 2\alpha_s(Ax^{(0)} - b, h^{(s)}) \rightarrow \min_{\alpha_s \in R}, s = 0, k-1. \quad (11.46)$$

(записаны полиномы второй степени относительно α_s), и указанные коэффициенты вычисляются как аргументы вершины параболы:

$$\alpha_s^* = -\frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})}, s = 0, \dots, k-1 \quad (11.47)$$

Следствие. Пусть $h^{(0)}, \dots, h^{(n-1)} \in R^n$ есть ненулевые взаимно сопряженные направления относительно $A = A^T > 0$ (они образуют базис в R^n) и пусть $x^{(0)} \in R^n$. Так как многообразие $L_n(x_0, h^{(0)}, \dots, h^{(n-1)})$ совпадает с пространством R^n , решением многомерной (n -мерной) задачи оптимизации

$$F(x) \rightarrow \min_{x \in R^n}$$

является

$$x^* = x^{(0)} + \alpha_0^* h^{(0)} + \dots + \alpha_{n-1}^* h^{(n-1)}$$

где коэффициенты $\alpha_s^*, s = 0, \dots, n-1$, есть решения n одномерных задач оптимизации

$$\alpha_s^2(Ah^{(s)}, h^{(s)}) + 2\alpha_s(Ax^{(0)} - b, h^{(s)}) \rightarrow \min_{\alpha_s \in R}, s = 0, n-1. \quad (11.48)$$

(записаны полиномы второй степени относительно α_s) и указанные коэффициенты вычисляются как аргументы вершины параболы:

$$\alpha_s^* = -\frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})}, s = 0, \dots, n-1 \quad (11.49)$$

Комментарий

Если **сопряженные направления известны**, решение многомерной задачи оптимизации (11.44) может быть сведено к **явному решению нескольких одномерных оптимизационных задач**.

Описание метода сопряженных градиентов

Идея пошагового построения ненулевых взаимно сопряженных направлений реализована в **методе сопряженных градиентов**. В качестве $x^{(0)} \in R^n$ выбирают элемент R^n , «удобный» как начальное приближение к решению x^* .

Первый шаг метода

Приближение $x^{(1)} \in R^n$ найдем по формуле

$$x^{(1)} = x^{(0)} + \alpha_0 h^{(0)} \quad (11.50)$$

где направление (вектор) $h^{(0)} \in R^n$ определяется начальной невязкой:

$$h^{(0)} = -r^{(0)} = Ax^{(0)} - b. \quad (11.51)$$

Чтобы найти α_0 , решаем одномерную задачу оптимизации

$$F(x^{(0)} + \alpha_0 h^{(0)}) \rightarrow \min_{\alpha_0 \in R} \quad (11.52)$$

(задача минимизации по аргументу α_0 , где $x^{(0)}, h^{(0)} \in R^n$ известны). Так как

$$F(x^{(0)} + \alpha_0 h^{(0)}) = F(x^{(0)}) + \alpha_0^2 (Ah^{(0)}, h^{(0)}) + 2\alpha_0 (Ax^{(0)} - b, h^{(0)})$$

(полином второй степени относительно α_0), решением (11.52) является

$$\alpha_0 = -\frac{(Ax^{(0)} - b, h^{(0)})}{(Ah^{(0)}, h^{(0)})} \quad (11.53)$$

(аргумент вершины параболы).

Второй шаг метода

Приближение $x^{(2)} \in R^n$ найдем по формуле

$$x^{(2)} = x^{(1)} + \alpha_1 h^{(1)} \quad (11.54)$$

где направление (вектор) $h^{(1)} \in R^n$ определяется невязкой предыдущего шага

$$r^{(1)} = Ax^{(1)} - b$$

и предыдущим направлением $h^{(0)}$:

$$h^{(1)} = -r^{(1)} + \beta_1 h^{(0)} \quad (11.55)$$

Нужно, чтобы $h^{(1)}, h^{(0)}$ были сопряженными относительно $A = A^T > 0$:

$$(Ah^{(0)}, h^{(1)}) = 0.$$

Получим условие

$$(Ah^{(0)}, h^{(1)}) = (Ah^{(0)}, h^{(0)})\beta_1 - (Ah^{(0)}, r^{(1)}) = 0,$$

откуда следует

$$\beta_1 = \frac{(Ah^{(0)}, r^{(1)})}{(Ah^{(0)}, h^{(0)})} \quad (11.56)$$

Чтобы найти α_1 , решаем одномерную задачу оптимизации

$$F(x^{(1)} + \alpha_1 h^{(1)}) \rightarrow \min_{\alpha_1 \in R} \quad (11.57)$$

(задача минимизации по аргументу α_1 , где $x^{(1)}, h^{(1)} \in R^n$ известны). Так как

$$F(x^{(1)} + \alpha_1 h^{(1)}) = F(x^{(1)}) + \alpha_1^2 (Ah^{(1)}, h^{(1)}) + 2\alpha_1 (Ax^{(1)} - b, h^{(1)})$$

(полином второй степени относительно α_1), решением (11.57) является

$$\alpha_1 = -\frac{(Ax^{(1)} - b, h^{(1)})}{(Ah^{(1)}, h^{(1)})} \quad (11.58)$$

(аргумент вершины параболы).

Третий шаг и далее

Шаг $s+1$, где $s \geq 2$ (т.е. третий шаг и далее) аналогичен шагу 2. Приближение $x^{(s+1)} \in R^n$ находим в виде

$$x^{(s+1)} = x^{(s)} + \alpha_s h^{(s)} \quad (11.59)$$

где направление (вектор) $h^{(s)} \in R^n$ определяется невязкой предыдущего шага

$$r^{(s)} = Ax^{(s)} - b$$

и предыдущим направлением $h^{(s-1)}$:

$$h^{(s)} = -r^{(s)} + \beta_s h^{(s-1)} \quad (11.60)$$

Нужно, чтобы $h^{(s)}, h^{(s-1)}$ были сопряженными относительно $A = A^T > 0$:

$$(Ah^{(s-1)}, h^{(s)}) = 0.$$

Получим условие

$$(Ah^{(s-1)}, h^{(s)}) = (Ah^{(s-1)}, h^{(s-1)})\beta_s - (Ah^{(s-1)}, r^{(s)}) = 0$$

откуда следует

$$\beta_s = \frac{(Ah^{(s-1)}, r^{(s)})}{(Ah^{(s-1)}, h^{(s-1)})} \quad (11.61)$$

Чтобы найти α_s , решаем одномерную задачу оптимизации

$$F(x^{(s)} + \alpha_s h^{(s)}) \rightarrow \min_{\alpha_s \in R} \quad (11.62)$$

(задача минимизации по аргументу α_s , где $x^{(s)}, h^{(s)} \in R^n$ известны).

Так как

$$F(x^{(s)} + \alpha_s h^{(s)}) = F(x^{(s)}) + \alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2\alpha_s (Ax^{(s)} - b, h^{(s)})$$

(полином второй степени относительно α_s), решением (11.62) является

$$\alpha_s = - \frac{(Ax^{(s)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})} \quad (11.63)$$

(аргумент вершины параболы).

Результат работы метода

На шаге $s+1$ будет построен $x^{(s+1)} \in R^n$, его связь с начальным приближением $x^{(0)} \in R^n$ описывается формулой

$$\begin{aligned} x^{(s+1)} &= x^{(s)} + \alpha_s h^{(s)} = x^{(s-1)} + \alpha_{s-1} h^{(s-1)} + \alpha_s h^{(s)} = \dots \\ &= x^{(0)} + \alpha_0 h^{(0)} + \dots + \alpha_s h^{(s)} \end{aligned} \quad (11.64)$$

Здесь сопряженные направления $h^{(i)}, i = 0, \dots, s$ вычислены по формулам (11.51), (11.55) и (11.60), коэффициенты $\alpha_i, i = 0, \dots, s$ вычислены по формулам (11.53), (11.58) и (11.63), коэффициенты $\beta_i, i = 1, \dots, s$, необходимые для расчета сопряженных направлений – по формулам (11.56) и (11.61).

Основные свойства метода

Свойство 1. Если на шаге $s+1$ получено $r^{(s+1)} = 0$, то $x^{(s+1)}$ – точное решение СЛАУ (11.43) и задачи оптимизации (11.44).

Свойство 2. Если в процессе работы метода получены невязки $r^{(0)}, \dots, r^{(s+1)} \neq 0$ (то есть за $s+1$ шагов точное решение задач (11.43) и (11.44) еще не найдено), тогда:

- невязки $r^{(0)}, \dots, r^{(s+1)} \in R^n$ взаимно ортогональны;
- векторы $h^{(0)}, \dots, h^{(s)} \in R^n$ взаимно сопряжены;
- значение коэффициента α_s , заданного формулой (11.53), (11.58) или (11.63), совпадает со значением, заданным формулой (11.49), то есть.

$$\alpha_s = - \frac{(Ax^{(s)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})} = - \frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})};$$

– приближение $x^{(s+1)} \in R^n$ обеспечивает минимальное значение функционала $F(x) = (Ax, x) - 2(b, x)$ на многообразии $L_{s+1}(x_0, h^{(0)}, \dots, h^{(s)})$.

Свойство 3. Не позднее чем на шаге n метод сопряженных градиентов строит точное решение СЛАУ (11.43) и задачи оптимизации (11.44).

Комментарии к параграфу 11.4

1) Метод сопряженных градиентов может использоваться как прямой метод или как итерационный.

Если на шаге n или ранее получено точное решение задач (11.43) и (11.44) – значит, метод использован как прямой.

Если каждое $x^{(s)}$ рассматривается как приближенное решение задач (11.43) и (11.44) – метод используется как итерационный.

2) Приближение $x^{(s+1)}$ лучше, чем предыдущее приближение $x^{(s)}$, так как обеспечивает минимальное значение функционала $F(x)$ на многообразии $L_{s+1}(x_0, h^{(0)}, \dots, h^{(s)})$ размерности $s + 1$, включающем предыдущее многообразие $L_s(x_0, h^{(0)}, \dots, h^{(s-1)})$ размерности s , и соответственно значение функционала $F(x)$ с каждым шагом убывает (не возрастает).

3) В силу накопления вычислительной погрешности взаимно сопряженные направления строятся приближенно и при большом числе шагов s векторы

$$h^{(0)}, \dots, h^{(s)} \in R^n$$

теряют свойство взаимной сопряженности.

Если есть проблемы отыскания $x^{(s+1)}$, текущее приближение $x^{(s)}$ можно принять за начальное приближение и запустить метод заново (то есть заново строить взаимно сопряженные направления).

4) В качестве численного решения СЛАУ (11.43) и задачи оптимизации (11.44) может быть выбрано такое приближение $x^{(s)}$, для которого:

- выполнен критерий отыскания точного решения: $r^{(s)} = 0$, то есть $x^{(s)} = x^*$,
- либо выполнен критерий остановки по точности: $\|x^{(s)} - x^{(s-1)}\| \leq \varepsilon$;
- либо выполнен критерий остановки по числу шагов: $s + 1 > N_{max}$;
- решение СЛАУ найдено с достаточно малой невязкой: $\|r^{(s)}\| \leq \varepsilon^*$ и др.

5) погрешность решения СЛАУ на шаге s можно оценить по текущей невязке, используя норму обратной матрицы или ее оценку:

$$\|z^{(s)}\| \leq \|A^{-1}\| \cdot \|r^{(s)}\| \quad (\text{нормы матрицы и вектора должны быть согласованы}):$$

6) оценки погрешности метода на шаге s по *начальной невязке* см. в литературе.