

ВНИМАНИЕ!
спасибо за внимание

на правах самописной шпоры

Товарищи, надеюсь, вам больше не придётся тратить своё бесценное время на сей богопротивный предмет

РАСПЕЧАТАТЬ БЕЗ ПЕРВОГО ЛИСТА, ТОГДА 1-ый БИЛЕТ БУДЕТ С ДРУГОЙ СТОРОНЫ ОТ 2-ого

Некоторые билеты лучше раписать по-лучше)

1. Связь метода Гаусса с разложением матрицы А=ВС

Рассмотрим матричное уравнение вида

$$Ax=f,$$

где

|
A
|
≠
0
,

A
(
m
×
m
)
,

x
=
(

x

1

,

x

2

,
.
.
.
,

x

m

)

,

f
=
(

f

1

,

f

2

,
.
.
.
,

f

m

)

T

.
 Матрица *A*, вообще говоря, может быть матрицей с комплексными элементами. Рассмотрим факторизацию (разложение в произведение) матрицы *A* (*m* × *m*)

$$A=B\cdot C,$$

где *B*—нижняя треугольная матрица, а *C*—верхняя треугольная матрица с единицами на главной диагонали:

$$B=\left(\begin{matrix}b_{11}&0&\cdots&0\\b_{21}&b_{22}&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\b_{m1}&b_{m2}&\cdots&b_{mm}\end{matrix}\right),\quad C=\left(\begin{matrix}1&c_{12}&\cdots&c_{1m}\\0&1&\cdots&c_{2m}\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{matrix}\right).$$

Ясно, что не любую матрицу *A* можно представить в виде (2). В дальнейшем покажем, что нахождение элементов матриц *B* и *C* возможно при определенном ограничении на матрицу *A*. Запишем выражение элемента *a*_{*ij*} матрицы *A* = *BC* как произведение *i*-й строки матрицы *B* и *j*-ого столбца матрицы *C*:

$$a_{ij}=\sum_{l=1}^mb_{il}c_{lj}.$$

Выделим *j*-ое слагаемое:

$$a_{ij}=\sum_{l=1}^{j-1}b_{il}c_{lj}+b_{ij}c_{jj}+\sum_{l=j+1}^mb_{il}c_{lj}.$$

Учитывая структуру матрицы *C* (*c*_{*ij*} = 0, *l* > *j*, *c*_{*jj*} = 1), получим

$$b_{ij}=a_{ij}-\sum_{l=1}^{j-1}b_{il}c_{lj},\quad i\geq j.$$

Аналогично, в определении элемента *a*_{*ij*} матрицы *A* выделим *i*-ое слагаемое:

$$a_{ij}=\sum_{l=1}^{i-1}b_{il}c_{lj}+b_{ii}c_{ij}+\sum_{l=i+1}^mb_{il}c_{lj}.$$

Исходя из вида матрицы *B* (*b*_{*il*} = 0, *l* > *i*), получим

$$b_{ii}c_{ij}=a_{ij}-\sum_{l=1}^{i-1}b_{il}c_{lj}.$$

Предполагая, что *b*_{*ii*} ≠ 0, поделим левую и правую части уравнения на *b*_{*ii*}:

$$c_{ij}=\frac{a_{ij}-\sum_{l=1}^{i-1}b_{il}c_{lj}}{b_{ii}},\quad i<j.$$

Уравнения (3) и (4) позволяют сформулировать следующий алгоритм нахождения элементов матриц *B* и *C*.

- Положим *b*₁₁ = *a*₁₁. Найдем элементы 1-й строки матрицы *C*:

$$c_{1j}=\frac{a_{1j}}{b_{11}},\quad j=\overline{2,m}.$$

- Рассмотрим элементы 1-ого столбца матрицы *B*:

$$b_{i1}=a_{i1},\quad i=\overline{2,m}.$$

- Положим *b*₂₂ = *a*₂₂ − *b*₂₁*c*₁₂. Далее, аналогично первому шагу, найдем элементы 2-й строки матрицы *C* по формулам:

$$c_{2j}=\frac{a_{2j}-b_{21}c_{1j}}{b_{22}},\quad j=\overline{3,m}.$$

- Вычислим элементы 2-ого столбца матрицы *B* аналогично второму шагу:

$$b_{i2}=a_{i2}-b_{i1}c_{12},\quad i=\overline{3,m}.$$

- Повторяя последовательно шаги алгоритма для столбцов матрицы *B* и строк матрицы *C*, найдем все элементы матриц *B* и *C*.

Утверждение. Пусть все угловые миноры матрицы *A* отличны от нуля. Тогда представление матрицы *A* в виде (2) существует и единственно.

Доказательство. Обозначим

|

A

1

|
=

a

11

≠
0
,

A

2

=
(

a

11

a

12

a

21

a

22

)
,
⋯
,

A

i

=
(

a

i1

⋯

a

i1

⋮

⋱

⋮

)
,

i
=
1
,
m
.
 Поскольку

|

A

i

|
≠
0
,

i
=
1
,
m
,

введем для определенности

|

A

0

|
=
1
. Ясно, что

$$A_i=B_i\cdot C_i,\quad i=\overline{1,m},$$

где *B*_{*i*} и *C*_{*i*} матрицы угловых миноров *i*-го порядка для матриц *B* и *C* соответственно.

Вычислим значение определителя матрицы *A*_{*i*}, приняв во внимание вид матриц *C*_{*i*} и *B*_{*i*} и равенство

|

C

i

|
=
1
:

$$|A_i|=|B_i||C_i|=\underbrace{b_{11}b_{22}\cdots b_{i-1,i-1}}_{|A_{i-1}|}b_{ii},$$

Следовательно,

$$b_{ii}=\frac{|A_i|}{|A_{i-1}|}\neq 0,\quad i=\overline{1,m}.$$

Таким образом, факторизация матрицы *A* в виде (2) существует и определяется единственным образом. □

Задача. Показать, что для вычисления элементов матриц *B* и *C* по формулам (3) и (4) требуется

m

3

−
m

3

умножений и делений.
(Число умножений и делений далее будем называть числом операций.)

Решение. Оценим необходимое число операций для вычисления элементов *b*_{*ij*} по формуле (3). Для вычисления фиксированного *b*_{*ij*} потребуется (*j* − 1) умножение. Зафиксировав *i* и учитывая, что *i* ≥ *j*, получим

$$\sum_{j=1}^i(j-1)=\frac{i(i-1)}{2}.$$

Далее, варьируя *i* от 1 до *m*, получим

$$\sum_{i=1}^m\frac{i(i-1)}{2}=\frac{1}{2}\left(\sum_{i=1}^mi^2-\sum_{i=1}^mi\right)=$$

$$=\frac{1}{2}\left(\frac{m(m+1)(2m+1)}{6}-\frac{m(m+1)}{2}\right)=\frac{m(m-1)(m+1)}{6}.$$

Оценим необходимое число операций для вычисления элементов *c*_{*ij*} по формуле (4). Для вычисления фиксированного *c*_{*ij*} потребуется (*i* − 1) умножение и одно деление. При фиксированном *j* получим

$$\sum_{i=1}^{j-1}\frac{j(j-1)}{2}.$$

Далее, варьируя *j* от 1 до *m*, получим аналогичную формулу:

$$\sum_{j=1}^m\frac{j(j-1)}{2}=\frac{m(m-1)(m+1)}{6}.$$

Сложив необходимое число операций для вычисления *b*_{*ij*} и *c*_{*ij*}, получим искомый результат:

$$\frac{m(m-1)(m+1)}{6}+\frac{m(m-1)(m+1)}{6}=\frac{m^3-m}{3}.$$

□

Замечание. Классическим методом решения СЛАУ вида (1) является метод Гаусса.

Число действий, необходимых для преобразований матрицы в прямом ходе метода Гаусса равно

m

3

−
m

3

.
 Заметим, что матрица *A'*, к которой приводится матрица *A* в прямом ходе метода Гаусса, в точности совпадает с матрицей *C*, полученной в результате факторизации матрицы *A* в виде (2). Таким образом факторизация матрицы *A* в виде (2) требует такое же число действий, что и сведение матрицы *A* к *A'* в прямом ходе метода Гаусса.

3. Метод квадратного коря

Определение. *Квадратная матрица A называется эрмитовой (самосопряженной), если ее элементы связаны соотношением

a

i
j

=
a

j
i

¯
.
 В этом случае будем записывать

A
=

A

∗

.*

Рассмотрим задачу

$$Ax=f,$$

где

A
∈

C

m
×
m

,

A
=

A

∗

,

|
A
|
≠
0
,

x
=
(

x

1

,

x

2

,
.
.
.
,

x

m

)

T

,

f
=
(

f

1

,

f

2

,
.
.
.
,

f

m

)

T

,
 и один из прямых методов ее решения — метод квадратного корня (метод Холецкого).

Факторизуем эрмитову матрицу *A* в виде

$$A=S^{\ast }DS,$$

где матрица *S*—верхнетреугольная матрица с положительными элементами на главной диагонали, а *D*—диагональная матрица со значениями ±1 на главной диагонали:

$$S=\left(\begin{matrix}s_{11}&s_{12}&\cdots &s_{1m}\\0&s_{22}&\cdots &s_{2m}\\ \vdots &\vdots &\ddots &\vdots \\0&0&\cdots &s_{mm}\end{matrix}\right),\quad s_{ii}>0,$$

$$D=\left(\begin{matrix}d_{11}&0&\cdots &0\\0&d_{22}&\cdots &0\\ \vdots &\vdots &\ddots &\vdots \\0&0&\cdots &d_{mm}\end{matrix}\right),\quad d_{ii}=\pm 1.$$

Покажем, что факторизация (6) возможна на примере вещественной симметрической матрицы второго порядка. Не ограничивая общности, будем полагать *a*₁₁ ≠ 0.

$$A=\left(\begin{matrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{matrix}\right)=A^T,\quad a_{12}=a_{21}.$$

Матрицы *S* и *D* будем искать в виде

$$S=\left(\begin{matrix}s_{11}&s_{12}\\0&s_{22}\end{matrix}\right),\quad s_{ii}>0,\quad i=1,2,$$

$$S^{\ast }=S^T=\left(\begin{matrix}s_{11}&0\\s_{12}&s_{22}\end{matrix}\right),$$

$$D=\left(\begin{matrix}d_{11}&0\\0&d_{22}\end{matrix}\right),\quad d_{ii}=\pm 1,\quad i=1,2.$$

Найдем матрицу *DS*:

$$DS=\left(\begin{matrix}d_{11}&0\\0&s_{22}\end{matrix}\right)\left(\begin{matrix}s_{11}&s_{12}\\0&s_{22}\end{matrix}\right)=\left(\begin{matrix}d_{11}s_{11}&d_{11}s_{12}\\0&d_{22}s_{22}\end{matrix}\right).$$

Домножим матрицу *DS* слева на *S*^{*T*}:

$$S^TDS==\left(\begin{matrix}d_{11}s_{11}^2&d_{11}s_{11}s_{12}\\d_{11}s_{11}s_{12}&d_{11}s_{12}^2+d_{22}s_{22}^2\end{matrix}\right).$$

Приравняем элементы матриц *A* и *S*^{*T*}*DS*:

$$\left\{\begin{array}{l}a_{11}=d_{11}s_{11}^2\\a_{12}=d_{11}s_{11}s_{12}\\a_{22}=d_{11}s_{12}^2+d_{22}s_{22}^2.\end{array}\right.$$

Из неравенства *s*₁₁ > 0 и из уравнения (7) следует, что

$$d_{11}=\operatorname{sgn}a_{11},\quad s_{11}=\sqrt{|a_{11}|}.$$

Рассмотрим уравнение (8). Заметим, что *s*₁₁*d*₁₁ ≠ 0, тогда

$$s_{12}=\frac{a_{12}}{s_{11}d_{11}}.$$

Наконец, рассмотрим уравнение (9). Получим соотношение *s*₂₂²*d*₂₂ = *a*₂₂ − *d*₁₁*s*₁₂², правая часть которого известна. Следовательно,

$$d_{22}=\operatorname{sgn}(a_{22}-s_{12}^2d_{11}),\quad s_{22}=\sqrt{|a_{22}-s_{12}^2d_{11}|}.$$

Таким образом, вещественную симметрическую матрицу второго порядка можно факторизовать в виде (6). Рассмотрим теперь произвольную эрмитову матрицу *A* (*m* × *m*). Запишем уравнение для элементов матрицы *DS*:

$$(DS)_{ij}=\sum_{l=1}^md_{il}s_{lj},\quad i,j=\overline{1,m}.$$

Учитывая диагональную структуру матрицы *D*, получим:

$$(DS)_{ij}=d_{ii}s_{ij}.$$

Домножим матрицу *DS* слева на *S*^{***}:

$$a_{ij}=(S^{\ast }DS)_{ij}=\sum_{l=1}^m(S^{\ast })_{il}d_{il}s_{lj},\quad i,j=\overline{1,m}.$$

Выделим *i*-ое слагаемое из последней суммы и учтем, что

(

S

∗

)

i
j

=
s
¯

j
i

:

$$a_{ij}=\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{ij}+\bar{s}_{ii}d_{ii}s_{ij}+\sum_{l=i+1}^m\bar{s}_{li}d_{li}s_{lj},\quad i,j=\overline{1,m}.$$

Третье слагаемое из равенства равно нулю в силу того, что матрица *S*^{***} является нижнетреугольной:

s
¯

i
l

=
0
,

l
>
i
. Тогда получим:

$$a_{ij}=\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{ij}+\bar{s}_{ii}d_{ii}s_{ij},\quad i,j=\overline{1,m}.$$

Так как матрица *A* эрмитова, достаточно рассматривать это равенство только в случае *i* ≤ *j*. При *i* = *j* получим:

$$a_{ii}=\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{li}+\bar{s}_{ii}d_{ii}s_{ii},\quad i=\overline{1,m}.$$

Выразим *d*_{*ii*} и *s*_{*ii*}:

$$d_{ii}=\operatorname{sgn}(a_{ii}-\sum_{l=1}^{i-1}|s_{li}|^2d_{ll}),\quad i=\overline{1,m},$$

$$s_{ii}=\sqrt{\left|a_{ii}-\sum_{l=1}^{i-1}|s_{li}|^2d_{ll}\right|},\quad i=\overline{1,m}.$$

Рассмотрим случай *i* ≠ *j* (*i* < *j*). В уравнении (10) выделим второе слагаемое:

$$\bar{s}_{ii}d_{ii}s_{ij}=a_{ij}-\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{lj},\quad i,j=\overline{1,m}.$$

В силу того, что *s*_{*ii*} — вещественные положительные числа, получим

$$s_{ii}d_{ii}s_{ij}=a_{ij}-\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{lj},\quad i,j=\overline{1,m}.$$

Так как *s*_{*ii*}*d*_{*ii*} ≠ 0, то получим выражения для коэффициентов *s*_{*ij*}:

$$s_{ij}=\frac{a_{ij}-\sum_{l=1}^{i-1}\bar{s}_{li}d_{li}s_{lj}}{s_{ii}d_{ii}},\quad i,j=\overline{1,m},\quad i<j.$$

Таким образом, для вычисления элементов матриц в разложении (6) были получены явные формулы (11)–(13).

Метод квадратного корня позволяет примерно вдвое уменьшить число операций, необходимых для решения системы (5), по сравнению с методом Гаусса—до ~

m

3

6

 умножений и делений. Кроме этого необходимо *m* операций извлечения квадратного корня. Заметим, что метод справедлив только в случае, если матрица системы линейных уравнений эрмитова.

5 Теорема самарского о сходимости итерационных методов

Рассмотрим матричное уравнение вида

$$Ax=f,$$

где

|
A
|
≠
0
,

A
(
m
×
m
)
,

x
=
(

x

1

,

x

2

,
.
.
.
,

x

m

)

T

,

f
=
(

f

1

,

f

2

,
.
.
.
,

f

m

)

T

.
 Рассмотрим также двухслойный стационарный метод решения уравнения (14):

$$B\frac{x^{n+1}-x^n}{\tau}+Ax^n=f,$$

где *n* ∈

Z

+

, начальное приближение *x*⁰ задано,

τ
—положительное вещественное число, *B*—обратимая матрица порядка (*m* × *m*).

Введем евклидову норму:

‖
x
‖
=
√
(
x
,
x
)
=
(

∑

i
=
1

m

x

i

2

)

1
2

. Эту норму также часто называют среднеквадратичной нормой.

Определение. *Линейный оператор D называется положительным (неотрицательным), если

(
D
x
,
x
)
>
0

∀
x
∈
H
,

x
≠
θ
 (соответственно

(
D
x
,
x
)
≥
0

∀
x
∈
H
). Положительностью оператора *D* обозначается как *D* > 0.*

Определение. *Скалярным произведением в смысле оператора D называется скалярное произведение, определяемое соотношением

(
x
,
y

)

D

=
(
D
x
,
y
)
.*

Определение. *Энергетической нормой, порождаемой линейным самосопряженным положительно определенным оператором *D*, называется норма, задаваемая соотношением

‖
x
‖

D

=
√
(
x
,
x

)

D

=
√
(
D
x
,
x
)
.*

Задача. Пусть *D* = *D*^{***} > 0. Доказать, что ∃ δ > 0 :

(
D
x
,
x
)
≥
δ
(
x
,
x
)
=
δ
‖
x
‖

2

.

Рассмотрим свойства положительного самосопряженного линейного оператора.

Если *D* = *D*^{***} > 0, то определены *D*^{−1} =

(

D

−
1

)

∗
 > 0,

D

1
2

=
(

D

1
2

)

∗
 > 0,

D

−

1
2

=
(

D

−

1
2

)

∗
 > 0.

Определение. *Погрешностью итерационного метода на *n*-й итерации называется вектор*

$$v^n=x^n-x.$$

Определение. *Итерационный метод сходится в норме

‖
⋅
‖
, если

‖

v

n

‖
→
0
 при

n
→
∞
.*

Выразим *x*^{*n*} из формулы (16) и подставим в уравнение (15). Получим однородное уравнение:

$$B\frac{v^{n+1}-v^n}{\tau}+Av^n=0,$$

где *n* ∈

Z

+

, *v*⁰ = *x*⁰ − *x*.

Приступим к исследованию задачи (17). Выразим (*n* + 1)-ю итерацию через *n*-ю с учетом того, что для матрицы *B* существует обратная. Домножим уравнение (17) на *B*^{−1} слева:

$$\frac{v^{n+1}-v^n}{\tau}+B^{-1}Av^n=0.$$

Выразим из уравнения погрешность на (*n* + 1)-й итерации:

$$v^{n+1}=v^n-\tau B^{-1}Av^n=(E-\tau B^{-1}A)v^n=Su^n.$$

Таким образом, мы получили матрицу *S*, которая связывает погрешность на предыдущей итерации с погрешностью на следующей:

Определение. *Матрица

S
=
E
−
τ

B

−
1

A
 называется матрицей перехода от *n*-й итерации к (*n* + 1)-й.*

Теорема 1. *Итерационный метод (15) решения системы (14) сходится при любом начальном приближении тогда и только тогда, когда все собственные значения матрицы перехода *S* по модулю меньше единицы. (Без доказательства)*

Заметим, что данная теорема практически неприменима, так как задача нахождения полного спектра матрицы *S* аналитически решается крайне редко.

Теорема 2 (теорема Самарского). Пусть *A* — самосопряженный положительно определенный оператор,

τ
 — положительное вещественное число и выполнено операторное неравенство

$$B-\frac{\tau}{2}A>0.$$

Тогда итерационный метод (15) решения системы (14)

8. Попеременно-треугольный итерационный метод. Реализация метода. Теорема о сходимости.

стр. 21, 31,32
А = R1 + R2, R1 — нижнетреуг. матрица, R2 — верхнетреуг. матрица:

$$R_1 = \begin{pmatrix} 0.5a_{11} & 0 & 0 \\ a_{21} & 0.5a_{22} & 0 \\ \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & 0.5a_{mm} \end{pmatrix}, \quad R_2 = \begin{pmatrix} 0.5a_{11} & a_{12} & a_{1m} \\ 0 & 0.5a_{22} & a_{2m} \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0.5a_{mm} \end{pmatrix}.$$

Попеременно-треугольный метод имеет вид

$$(E + \omega R_1)(E + \omega R_2)\frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad n \in \mathbb{Z}_+, \quad (38)$$

где $\tau > 0$, $\omega > 0$ — итерационные параметры, позволяющие, вообще говоря, ускорить процесс сходимости итерационного метода, матрица E — единичная. Рассматриваемый метод формально является неявным, однако можно показать, что $(n+1)$ -я итерация выражается с помощью явных формул за три шага.Введем обозначения:

$$W^{n+1} = (E + \omega R_2)\frac{x^{n+1} - x^n}{\tau},$$

$$v^{n+1} = \frac{x^{n+1} - x^n}{\tau}.$$

Определение. Вектор $r^n = f - Ax^n$ называется *невязкой* на n -й итерации.

В нашем случае невязка r^n известна. Предположим, что матрицы $E + \omega R_1$ и $E + \omega R_2$ имеют обратные. На первом шаге решим уравнение

$$(E + \omega R_1)W^{n+1} = r^n.$$

Заметим, что $(E + \omega R_1)$ — нижнетреугольная матрица. Нахождение вектора решения системы с нижнетреугольной матрицей осуществляется по явным формулам, начиная с первой компоненты вектора W^{n+1} . На втором шаге аналогично решим уравнение с верхнетреугольной матрицей $(E + \omega R_2)$:

$$(E + \omega R_2)v^{n+1} = W^{n+1}.$$

На третьем шаге найдем $(n+1)$ -ую итерацию по формуле

$$x^{n+1} = x^n + \tau v^{n+1}.$$

Таким образом, несмотря на то, что попеременно-треугольный итерационный метод является неявным, его реализация не представляет никакой трудности. Обозначим

$$B = (E + \omega R_1)(E + \omega R_2).$$

Теорема 4 (о сходимости ПТИМ). Пусть A — самосопряженный положительно определенный оператор и $\omega > \frac{\tau}{4}$. Тогда ПТИМ сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Доказательство. Раскроем скобки в выражении для B , учитывая, что $R_1 = R_2^*$:

$$B = (E + \omega R_2^*)(E + \omega R_2) = E + \omega(R_2^* + R_2) + \omega^2 R_2^* R_2 = E + \omega A + \omega^2 R_2^* R_2. \quad (39)$$

Очевидно, что

$$B = (E - \omega R_2^*)(E - \omega R_2) + 2\omega A. \quad (40)$$

Кроме того,

$$((E - \omega R_2^*)(E - \omega R_2)x, x) = ((E - \omega R_2)x, (E - \omega R_2)x) \geq 0.$$

Тогда из равенства (67) следует неравенство

$$B \geq 2\omega A. \quad (41)$$

Учитывая условие теоремы ($\omega > \frac{\tau}{4}$), получим, что $B > \frac{\tau}{2}A$ и ПТИМ сходится по теореме Самарского при любом начальном приближении x^0 .

6. Достаточные условия сходимости методов Якоби, Зейделя, простой итерации.

стр. 25-27

Метод Якоби

Метод Якоби является явным итерационным методом и задается правилом

$$x_i^{n+1} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^n - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

Метод Зейделя

$$x_i^{n+1} = -\frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^{n+1} - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

Следствие 2. Пусть $A = A^* > 0$. Тогда метод Якоби сходится в среднеквадратичной норме при любом начальном приближении, если выполнено неравенство:

$$2D > A,$$

$$ade \; A = R_1 + D + R_2, \; D = \operatorname{diag}(a_{11}, a_{22}, \dots, a_{mm}).$$

Доказательство. В методе Якоби $\tau = 1$, а $B = D$. По теореме Самарского метод сходится, если

$$B - \frac{\tau}{2}A > 0.$$

В нашем случае

$$D - \frac{1}{2}A > 0,$$

а это выполняется в силу условия $2D > A$. Следовательно, метод Якоби сходится в среднеквадратичной норме при любом начальном приближении.

Следствие 3. Пусть положительная симметричная матрица ($A = A^* > 0$) является матрицей со строгим диагональным преобладанием:

$$a_{ii} > \sum_{j=1, j \neq i}^m |a_{ij}|, \quad i = \overline{1, m}.$$

Тогда метод Якоби сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Доказательство. Рассмотрим квадратичную форму с матрицей A :

$$(Ax, x) = \sum_{i,j=1}^m a_{ij}x_ix_j \leq \sum_{i,j=1}^m |a_{ij}| \, |x_i| \, |x_j|. \quad (42)$$

Для дальнейшей оценки квадратичной формы (42) воспользуемся неравенством $ab \leq \frac{a^2 + b^2}{2}$:

$$(Ax, x) \leq \frac{1}{2} \sum_{i,j=1}^m |a_{ij}| \, |x_i|^2 + \frac{1}{2} \sum_{i,j=1}^m |a_{ij}| \, |x_j|^2.$$

Преобразуем правую часть неравенства с учетом того, что матрица A является симметричной ($|a_{ij}| = |a_{ji}|$):

$$\frac{1}{2} \sum_{i,j=1}^m |a_{ij}| \, |x_i|^2 + \frac{1}{2} \sum_{i,j=1}^m |a_{ij}| \, |x_i|^2 = \sum_{i,j=1}^m |a_{ij}| \, |x_i|^2.$$

Вынесем суммирование по индексу i и воспользуемся свойством диагонального преобладания матрицы A :

$$\sum_{i=1}^m |x_i|^2 \left(a_{ii} + \sum_{j=1, j \neq i}^m |a_{ij}| \right) < \sum_{i=1}^m 2x_i^2 a_{ii} = (2Dx, x),$$

где $D = \operatorname{diag}(a_{11}, a_{22}, \dots, a_{mm})$. Таким образом, мы получили, что $(Ax, x) < (2Dx, x)$.

Из этого неравенства следует, что $2D > A$. Следовательно, выполняется условие следствия 1, и итерационный метод Якоби сходится при любом начальном приближении.

Задача. Пусть матрица $A = A^* > 0$. Доказать, что $a_{ii} > 0$, $i = \overline{1, m}$.

Следствие 4. Пусть $A = A^* > 0$. Тогда метод Зейделя сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Доказательство. Из условия теоремы Самарского следует, что для сходимости метода Зейделя достаточно выполнения неравенства

$$B - \frac{\tau}{2}A > 0. \quad (43)$$

Представим матрицу A в виде $A = R_1 + D + R_2$. В канонической записи метода Зейделя $\tau = 1$, $B = R_1 + D$. Тогда достаточное условие (43) преобразуется к виду

$$D + R_1 - \frac{R_1 + D + R_2}{2} > 0.$$

И, следовательно,

$$D + R_1 - R_2 > 0. \quad (44)$$

Запишем это неравенсто в виде

$$(Dx, x) + (R_1x, x) - (R_2x, x) > 0, \quad x \neq \theta.$$

Так как $A = A^*$, то $R_2^* = R_1$. Тогда

$$(R_2x, x) = (x, R_2^*x) = (x, R_1x) = (R_1x, x).$$

Следовательно, неравенство (44) принимает вид

$$(Dx, x) > 0, \quad x \neq \theta. \quad (45)$$

Если матрица симметричная и положительно определенная, то все ее диагональные элементы больше нуля (см. задачу). Следовательно, матрица D также является положительно определенной, откуда следует неравенство (45).

Следствие 5. Пусть $A = A^* > 0$, $\gamma_2 = \max_{1 \leq k \leq m} \lambda_k > 0$. Если $0 < \tau < \frac{2}{\gamma_2}$, то метод простой итерации сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Доказательство. Из условия теоремы Самарского следует, что для того, чтобы метод простой итерации сохлился в среднеквадратичной норме при любом начальном приближении, достаточно выполнения неравенства

$$B - \frac{\tau}{2}A > 0. \quad (46)$$

В методе простой итерации $B = E$. Следовательно, условие (46) преобразуется к виду

$$E - \frac{\tau}{2}A > 0. \quad (47)$$

Неравенство (47) выполнено, если $\lambda^E - \frac{\tau}{2}\lambda^A > 0$, что справедливо, если

$$1 - \frac{\tau}{2}\gamma_2 > 0.$$

Из положительности параметра τ следует, что для сходимости метода простой итерации достаточно выполнения условия

$$0 < \tau < \frac{2}{\gamma_2}.$$

4 Примеры на канонический вид двухслойного итерационного метода решения СЛАУ

стр18-21

Рассмотрим матричное уравнение

$$Ax = f, \quad (48)$$

где $|A| \neq 0$, $A \, (m \times m)$, $x = (x_1, x_2, \dots, x_m)^T$, $f = (f_1, f_2, \dots, f_m)^T$. Распишем систему (48) покоординатно:

$$\sum_{j=1}^m a_{ij}x_j = f_i, \quad i = \overline{1, m}. \quad (49)$$

Выделим i -ое слагаемое в сумме:

$$\sum_{j=1}^{i-1} a_{ij}x_j + a_{ii}x_i + \sum_{j=i+1}^m a_{ij}x_j = f_i, \quad i = \overline{1, m}.$$

Предположим, что элементы главной диагонали матрицы A отличны от нуля: $a_{ii} \neq 0$, $i = \overline{1, m}$. Тогда уравнение (49) разрешимо относительно x_i :

$$x_i = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^m a_{ij}x_j}{a_{ii}}, \quad i = \overline{1, m}.$$

Все итерационные методы основаны на построении последовательности векторов $x^n = (x_1^n, \dots, x_m^n)$ такой, что $x^n \rightarrow x$ при $n \rightarrow \infty$, где x — точное решение матричного уравнения (48). Вектор x^n называется n -й итерацией метода.

Определение. Итерационный метод называется *двухслойным*, если для вычисления текущей итерации используются только элементы предыдущей итерации.

Замечание. Двухслойный итерационный метод также называют *одношаговым*.

Метод Якоби

Метод Якоби является явным итерационным методом и задается правилом

$$x_i^{n+1} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^n - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

Метод Зейделя

$$x_i^{n+1} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^{n+1} - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

В правой части уравнения используются координаты $(n+1)$ -й итерации, поэтому метод Зейделя является неявным. Однако, каждый элемент $(n+1)$ -й итерации можно найти по явным формулам от уже известных элементов, вычисляя значения x столбец за столбом.

Каноническая запись итерационных методов

Представим матрицу A в виде

$$A = R_1 + D + R_2,$$

где R_1 — нижнетреугольная матрица с нулевой главной диагональю, D — диагональная матрица, R_2 — верхнетреугольная матрица с нулевой главной диагональю:

$$R_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} \end{pmatrix},$$

$$R_2 = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1m} \\ 0 & 0 & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Перепишем матричное уравнение (48) в виде

$$(R_1 + D + R_2)x = f.$$

Оставим в левой части слагаемое с матрицей D , остальные слагаемые перенесем в правую часть уравнения:

$$Dx = f - R_1x - R_2x.$$

Предположим, что матрица D обратима ($a_{ii} \neq 0$, $i = \overline{1, m}$). Тогда получим:

$$x = D^{-1}f - D^{-1}R_1x - D^{-1}R_2x. \quad (50)$$

Запишем итерационные методы Якоби и Зейделя исходя из уравнения (50):

$$\begin{aligned} \text{Метод Якоби:} \quad x^{n+1} &= D^{-1}f - D^{-1}R_1x^n - D^{-1}R_2x^n, \quad n \in \mathbb{Z}_+, \\ \text{Метод Зейделя:} \quad x^{n+1} &= D^{-1}f - D^{-1}R_1x^{n+1} - D^{-1}R_2x^n, \quad n \in \mathbb{Z}_+. \end{aligned}$$

Наконец, перепишем эти соотношения в виде

$$\text{Метод Якоби:} \quad D(x^{n+1} - x^n) + Ax^n = f, \quad n \in \mathbb{Z}_+, \quad (51)$$

$$\text{Метод Зейделя:} \quad (D + R_1)(x^{n+1} - x^n) + Ax^n = f, \quad n \in \mathbb{Z}_+. \quad (52)$$

Определение. Канонической формой записи двухслойного итерационного метода решения системы (48) называется его запись в виде

$$B_{n+1} \frac{x^{n+1} - x^n}{\tau_{n+1}} + Ax^n = f, \quad (53)$$

где $n \in \mathbb{Z}_+$, начальное приближение x^0 задано, τ_{n+1} — положительное вещественное число, называемое итерационным параметром, B_{n+1} — некоторая обратимая матрица.

Определение. Если в методе (53) параметр τ_{n+1} и матрица B_{n+1} не зависят от номера итерации ($B_{n+1} = B$, $\tau_{n+1} = \tau$), то такой метод называется *стационарным*, в противном случае — *нестационарным*.

Определение. Если $B_{n+1} = E$, то метод (53) называется *явным*, в противном случае — *неявным*.

Метод простой итерации

Метод простой итерации (метод релаксации) определяется итерационной схемой вида

$$\frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad \tau > 0, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ — задано.} \quad (54)$$

Метод Ричардсона

Метод Ричардсона определяется итерационной схемой вида

$$\frac{x^{n+1} - x^n}{\tau_{n+1}} + Ax^n = f, \quad \tau_{n+1} > 0, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ — задано.} \quad (55)$$

2. обращение матрицы методом Гаусса-Жордана

стр13-15

Рассмотрим задачу обращения (поиска обратной матрицы) невырожденной матрицы $A \, (m \times m)$. Согласно критерию обратимости матрицы, для невырожденной матрицы всегда существует обратная. Введем обозначение: $A^{-1} = X = (x_{ij})$, $i, j = \overline{1, m}$. С учетом этого задача обращения матрицы состоит в решении системы

$$AX = E, \quad (56)$$

где $A \, (m \times m)$, $|A| \neq 0$, или, если записать поэлементно:

$$\sum_{l=1}^m a_{il}x_{lj} = \delta_{ij}, \quad i = \overline{1, m}, \quad j = \overline{1, m}. \quad (57)$$

Покажем, что существует способ обращения матрицы, требующий ровно m^3 операций.

Сведем уравнение (57) к решению m систем линейных уравнений с матрицей A . Для этого введем вектор-столбец матрицы X : $X^{(j)} = (x_{1j}, x_{2j}, \dots, x_{mj})^T$ и вектор-столбец правой части $\delta^{(j)} = (0, 0, \dots, 0, 1, 0, \dots, 0)^T$ с единицей на j -й позиции. Теперь можем записать матричное уравнение (56) в виде m систем:

$$AX^{(j)} = \delta^{(j)}, \quad j = \overline{1, m}. \quad (58)$$

Факторизуем матрицу A в виде

$$A = B \cdot C. \quad (59)$$

Для этого требуется $\frac{m^3 - m}{3}$ умножений и делений. Получаем две системы линейных уравнений:

$$\begin{cases} By^{(j)} = \delta^{(j)}, & j = \overline{1, m}, \\ Cx^{(j)} = y^{(j)}. \end{cases} \quad (60)$$

$$(61)$$

При фиксированном j решение систем (60) и (61) потребует m^2 действий. Для решения m таких систем при $j = \overline{1, m}$ потребуется m^3 действий. Значит, в целом для обращения матрицы A необходимо $m^3 + \frac{m^3 - m}{3} \sim \frac{4}{3}m^3$ операций. Покажем теперь, что это число операций можно уменьшить. Рассмотрим систему уравнений (60):

$$\begin{aligned} b_{11}y_1^{(j)} &= 0 &\Rightarrow y_1^{(j)} &= 0, \\ b_{21}y_1^{(j)} + b_{22}y_2^{(j)} &= 0 &\Rightarrow y_2^{(j)} &= 0, \\ b_{31}y_1^{(j)} + b_{32}y_2^{(j)} + b_{33}y_3^{(j)} &= 0 &\Rightarrow y_3^{(j)} &= 0, \\ &\dots && \\ b_{j-1,1}y_1^{(j)} + \dots + b_{j-1,j-1}y_{j-1}^{(j)} &= 0 &\Rightarrow y_{j-1}^{(j)} &= 0. \end{aligned}$$

Рассмотрим j -ое уравнение: $b_{jj}y_j^{(j)} = 1$. Предполагая, что $b_{jj} \neq 0$, получим:

$$y_j^{(j)} = \frac{1}{b_{jj}}. \quad (62)$$

Запишем уравнения системы при $i > j$

$$b_{ij}y_j^{(j)} + b_{i,j+1}y_{j+1}^{(j)} + \dots + b_{ii}y_i^{(j)} = 0, \quad i = \overline{(j+1), m}, \quad (63)$$

и выразим из них $y_i^{(j)}$:

$$y_i^{(j)} = -\frac{\sum_{l=j}^{i-1} b_{il}y_l^{(j)}}{b_{ii}}, \quad i = \overline{(j+1), m}. \quad (64)$$

Перейдем к подсчету числа операций, необходимых для решения систем уравнений (60) и (61). При фиксированных i и j в формуле (64) получаем $(i-j)$ умножений и одно деление в уравнении (62). Варьируя индекс i от 1 до m , при фиксированном j получаем

9 Исследование скорости сходимости ПТИМ

Рассмотрим матричное уравнение вида $Ax = f$, где $|A| \neq 0$, A ($m \times m$), $x = (x_1, x_2, \ldots, x_m)^T$, $f = (f_1, f_2, \ldots, f_m)^T$. Представим матрицу A в виде $A = R_1 + R_2$,

$$R_1 = \begin{pmatrix} 0.5a_{11} & 0 & \cdots & 0 \\ a_{21} & 0.5a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & 0.5a_{mm} \end{pmatrix} \\ , \quad R_2 = \begin{pmatrix} 0.5a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & 0.5a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0.5a_{mm} \end{pmatrix}.$$

каноническая форма (ПТИМ):

$$(E + \omega R_1)(E + \omega R_2) \frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad \omega > 0, \quad \tau > 0, \quad n \in \mathbb{Z}_+.$$

Обозначим $B = (E + \omega R_1)(E + \omega R_2)$.

Теорема 5 (о сходимости ПТИМ). *Пусть A — самосопряженный положительно определенный оператор и $\omega > \frac{\tau}{4}$. Тогда ПТИМ сходится в среднеквадратичной норме при любом начальном приближении x^0 .*

Доказательство. тк $R_1 = R_2^*$:

$$B = (E + \omega R_2^*)(E + \omega R_2) = E + \omega(R_2^* + R_2) + \omega^2 R_2^* R_2 = E + \omega A + \omega^2 R_2^* R_2. \quad (66)$$

$$B = (E - \omega R_2^*)(E - \omega R_2) + 2\omega A. \quad (67)$$

$$((E - \omega R_2^*)(E - \omega R_2)x, x) = ((E - \omega R_2)x, (E - \omega R_2)x) \geq 0.$$

Тогда из равенства (67) следует неравенство

$$B \geq 2\omega A. \quad (68)$$

Учитывая условие теоремы ($\omega > \frac{\tau}{4}$), получим, что $B > \frac{\tau}{2}A$ и ПТИМ сходится по теореме Самарского при любом начальном приближении x^0 .

Теорема 6 (о скорости сходимости ПТИМ). *Пусть A — самосопряженный положительно определенный оператор и числа $\delta > 0$, $\Delta > 0$ таковы, что выполняются неравенства*

$$A \geq \delta E, \quad R_2^* R_2 \leq \frac{\Delta}{4} A. \quad (69)$$

Положим

$$\omega = \frac{2}{\sqrt{\delta \Delta}}, \quad \tau = \frac{2}{\gamma_1 + \gamma_2}, \quad \gamma_1 = \frac{\sqrt{\delta}}{2} \left(\frac{\sqrt{\delta \Delta}}{\sqrt{\delta} + \sqrt{\Delta}} \right), \quad \gamma_2 = \frac{\sqrt{\delta \Delta}}{4}.$$

Тогда ПТИМ сходится и имеет место оценка

$$\|x^{n+1} - x\|_B \leq \rho \|x^n - x\|_B,$$

$$\text{где } \rho = \frac{1 - \sqrt{\eta}}{1 + 3\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta}.$$

Доказательство. Покажем, что из неравенств (69) следует $\eta \leq 1$.

$$R_2^* R_2 \leq \frac{\Delta}{4} A \quad \Rightarrow \quad (R_2^* R_2 x, x) = (R_2 x, R_2 x) = \|R_2 x\|^2 \leq \frac{\Delta}{4} (Ax, x). \quad (70)$$

Рассмотрим первое неравенство:

$$A \geq \delta E \quad \Rightarrow \quad (Ax, x) \geq \delta \|x\|^2.$$

$$(Ax, x) = (R_2^* x, x) + (R_2 x, x) = 2(R_2 x, x).$$

Предположим, что x — ненулевой вектор, и получим

$$\delta \|x\|^2 \leq (Ax, x) = \frac{(Ax, x)^2}{(Ax, x)} = \frac{4(R_2 x, x)^2}{(Ax, x)}.$$

Воспользуемся неравенством Коши-Буняковского и неравенством (70):

$$\delta \|x\|^2 \leq \frac{4\|R_2 x\|^2 \|x\|^2}{(Ax, x)} \leq \frac{4\Delta(Ax, x) \|x\|^2}{4(Ax, x)} = \Delta \|x\|^2.$$

Таким образом, справедливо неравенство $\delta \leq \Delta$. Чтобы воспользоваться следствием 1 из теоремы об оценке скорости сходимости, найдем из условия теоремы числа γ_1 и γ_2 такие, что

$$\gamma_1 B \leq A \leq \gamma_2 B. \quad (71)$$

Из неравенства (68) ($B \geq 2\omega A$), следует оценка $A \leq \frac{B}{2\omega}$. Тогда можно

положить в неравенстве (71) $\gamma_2 = \frac{1}{2\omega}$.

Оценим выражение (66), воспользовавшись неравенствами (69):

$$B = E + \omega A + \omega^2 R_2^* R_2 \leq \frac{1}{\delta} A + \omega A + \frac{\Delta \omega^2}{4} A = \left(\frac{1}{\delta} + \omega + \frac{\Delta \omega^2}{4} \right) A.$$

Тем самым неравенство (71) выполнено с постоянной $\gamma_1 = \left(\frac{1}{\delta} + \omega + \frac{\Delta \omega^2}{4} \right)^{-1}$. Для нахождения максимально возможной скорости сходимости будем минимизировать функцию $\rho(\omega)$

$$\rho(\omega) = \frac{1 - \xi(\omega)}{1 + \xi(\omega)}, \quad \xi(\omega) = \frac{\gamma_1(\omega)}{\gamma_2(\omega)},$$

что эквивалентно минимизации функции $f(\omega)$:

$$f(\omega) = \frac{\gamma_2(\omega)}{\gamma_1(\omega)} = \frac{1}{2} \left(1 + \frac{1}{\omega \delta} + \frac{\Delta \omega}{4} \right).$$

Для нахождения экстремальных точек найдем производную $f(\omega)$ и приравняем ее к нулю:

$$f'(\omega) = \frac{1}{2} \left(\frac{\Delta}{4} - \frac{1}{\omega^2 \delta} \right) = 0 \quad \Rightarrow \quad \omega = \omega_0 = \frac{2}{\sqrt{\delta \Delta}}.$$

Учтем, что $\omega > 0$, и проверим, что точка ω_0 доставляет минимум функции $f(\omega)$, найдя знак второй производной функции в этой точке:

$$f''(\omega) = \frac{1}{\delta \omega^3} > 0.$$

Подставим ω_0 в выражения для γ_1, γ_2, ρ :

$$\gamma_1 = \frac{1}{\frac{1}{\delta} + \frac{2}{\sqrt{\delta \Delta}} + \frac{\Delta}{4} \frac{4}{\delta \Delta}} = \frac{1}{\frac{2}{\delta} + \frac{2}{\sqrt{\delta \Delta}}} = \frac{\delta \sqrt{\Delta}}{2\sqrt{\Delta} + 2\sqrt{\delta}} = \frac{\sqrt{\delta}}{2} \left(\frac{\sqrt{\delta \Delta}}{\sqrt{\Delta} + \sqrt{\delta}} \right)$$

$$\gamma_2 = \frac{1}{2\omega_0} = \frac{\sqrt{\delta \Delta}}{4}.$$

$$\xi(\omega) = \frac{\gamma_1(\omega)}{\gamma_2(\omega)} = \frac{4}{\sqrt{\delta \Delta}} \frac{\sqrt{\delta}}{2} \left(\frac{\sqrt{\delta \Delta}}{\sqrt{\Delta} + \sqrt{\delta}} \right) = \frac{2\sqrt{\delta}}{\sqrt{\Delta} + \sqrt{\delta}}$$

$$\left. \begin{aligned} 1 - \xi &= 1 - \frac{2\sqrt{\delta}}{\sqrt{\Delta} + \sqrt{\delta}} = \frac{\sqrt{\Delta} - \sqrt{\delta}}{\sqrt{\Delta} + \sqrt{\delta}} \\ 1 + \xi &= 1 + \frac{2\sqrt{\delta}}{\sqrt{\Delta} + \sqrt{\delta}} = \frac{\sqrt{\Delta} + 3\sqrt{\delta}}{\sqrt{\Delta} + \sqrt{\delta}} \end{aligned} \right\} \\ \Rightarrow \rho = \frac{1 - \xi}{1 + \xi} = \frac{\sqrt{\Delta} - \sqrt{\delta}}{\sqrt{\Delta} + 3\sqrt{\delta}} = \frac{1 - \sqrt{\eta}}{1 + 3\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta} \quad (\Delta \neq 0).$$

Исходя из полученных соотношений и следствия 1, получаем оценку

$$\|x^{n+1} - x\|_B \leq \rho \|x^n - x\|_B. \quad \square$$

ПТИМ сходится быстрее метода простой итерации, метода Зейделя и Якоби.

Число итераций, необходимое для достижения заданной точности $\varepsilon > 0$ равно

$$n_0(\varepsilon) = \left\lceil \ln \frac{\frac{1}{\rho}}{\frac{1}{\rho} - \varepsilon} \right\rceil,$$

где $[x]$ означает целую часть числа x , а $\ln \frac{1}{\rho}$ — скорость сходимости итерационного метода.

Оценим скорость сходимости ПТИМ:

$$\frac{1}{\rho} = \frac{1 + 3\sqrt{\eta}}{1 - \sqrt{\eta}} = \frac{(1 + 3\sqrt{\eta})(1 + \sqrt{\eta})}{1 - \eta} \approx 1 + 4\sqrt{\eta},$$

$$\ln \frac{1}{\rho} \approx \ln(1 + 4\sqrt{\eta}) = O(m^{-1}), \quad n_0(\varepsilon) = O(m).$$

Оценим скорость сходимости метода простой итерации:

$$\rho = \frac{1 - \xi}{1 + \xi} = \frac{1 - \eta}{1 + \eta}, \quad \frac{1}{\rho} = \frac{1 + \eta}{1 - \eta} = \frac{(1 + \eta)^2}{1 - \eta^2} \approx 1 + 2\eta,$$

$$\ln \frac{1}{\rho} \approx \ln(1 + 2\eta) = O(m^{-2}), \quad n_0(\varepsilon) = O(m^2).$$

Таким образом, метод простой итерации сходится на порядок медленнее, чем ПТИМ. Методы Якоби и Зейделя имеют тот же порядок сходимости, что и метод простой итерации.

11 Метод обратных итераций и обратных итераций со сдвигом

Пусть матрица A — невырожденная. Рассмотрим следующую формулу записи неявного итерационного метода:

$$Ax^{n+1} = x^n, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ задано.}$$

Будем называть такой метод методом обратных итераций. Умножим обе части равенства слева на A^{-1} и получим формулу степенного метода для матрицы A^{-1} :

$$x^{n+1} = A^{-1}x^n, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ задано.} \quad (72)$$

Из свойств обратной матрицы следует, что собственные значения невырожденной матрицы A и обратной к ней матрицы A^{-1} связаны соотношением

$$\lambda_k^{A^{-1}} = \frac{1}{\lambda_k^A}, \quad k = \overline{1, m}.$$

Заметим, что если собственные значения λ_k^A упорядочены по возрастанию модулей, то соответствующие им собственные значения $\lambda_k^{A^{-1}}$ будут упорядочены по убыванию модулей. В данном методе обозначим $\lambda_k = \lambda_k^A$, и пусть $\{\lambda_k\}$ упорядочены по возрастанию модулей.

Сформулируем три условия:

А) В пространстве $\mathbb{R}^m$ существует базис $\{e_k\}$ из собственных векторов матрицы A .

В) $\left| \frac{\lambda_1}{\lambda_2} \right| < 1$.

С) $x^0 = c_1 e_1 + c_2 e_2 + \ldots + c_m e_m$, $c_1 \neq 0$.

Утверждение. *Пусть невырожденная вещественная матрица A ($m \times m$) такова, что выполнены условия А)–С). Тогда метод обратных итераций сходится по направлению к собственному вектору, отвечающему минимальному по модулю собственному значению:*

$$x^n \xrightarrow[n \rightarrow \infty]{} e_1.$$

Доказательство. Разложим n -ю итерацию по базису $\{e_k\}$ из собственных векторов матрицы A :

$$x^n = A^{-n}x^0 = \sum_{k=1}^m c_k A^{-n}e_k = \sum_{k=1}^m c_k \lambda_k^{-n} e_k = c_1 \lambda_1^{-n} e_1 + c_2 \lambda_2^{-n} e_2 + \ldots + c_m \lambda_m^{-n} e_m.$$

В силу условия С) $c_1 \neq 0$. Кроме того, поскольку матрица A невырождена, $\lambda_1^i \neq 0$. Поделив равенство на $c_1 \lambda_1^{-n}$, получим

$$\frac{x^n}{c_1 \lambda_1^{-n}} = e_1 + \frac{c_2}{c_1} \left(\frac{\lambda_1}{\lambda_2} \right)^n e_2 + \ldots + \frac{c_m}{c_1} \left(\frac{\lambda_1}{\lambda_m} \right)^n e_m.$$

Перейдя к пределу при $n \rightarrow \infty$ и учитывая условие В), получим, что x^n сходится по направлению к e_1 :

$$\lim_{n \rightarrow \infty} x^n = e_1. \quad \square$$

Сформулируем утверждения о вычислении минимального собственного значения в виде задачи.

Метод обратных итераций со сдвигом

Рассмотрим итерационный метод, задаваемый формулой

$$(A - \alpha E)x^{n+1} = x^n, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ задано,}$$

где α — такое вещественное число, что матрица $(A - \alpha E)$ невырождена. Домножим обе части равенства слева на $(A - \alpha E)^{-1}$ и получим формулу степенного метода с матрицей $(A - \alpha E)^{-1}$:

$$x^{n+1} = (A - \alpha E)^{-1}x^n. \quad (73)$$

Таким образом, метод обратных итераций со сдвигом эквивалентен степенному методу, записанному для матрицы $B = (A - \alpha E)^{-1}$. Следовательно, векторы x^n будут сходиться при $n \rightarrow \infty$ по направлению к такому собственному вектору e_r матрицы A , для которого величина

$$|\lambda_r - \alpha|^{-1} = \max_{1 \leq k \leq m} |\lambda_k - \alpha|^{-1}.$$

Это означает, что если требуется найти собственный вектор e_r , отвечающий данному собственному значению λ_r , то надо задать число α , близкое к λ_r , и вычислить векторы x^n , исходя из формулы (73). Само собственное значение λ_r находится из выражения:

$$\lambda_r = \lim_{n \rightarrow \infty} \left(\alpha + \frac{x_i^{(n)}}{x_{n+1}^{(i)}} \right), \quad i = \overline{1, m}.$$

Следовательно, метод обратных итераций со сдвигом позволяет в принципе отыскать любое собственное значение матрицы A . Этот метод очень часто используют для нахождения и уточнения собственных векторов, если собственные значения уже известны.

13. Понятие о QR-алгоритме решения полной проблемы собственных значений

Утверждение. *Произвольная матрица A ($m \times m$) может быть представлена в виде:*

$$A = QR, \quad (74)$$

где Q — ортогональная матрица, а R — матрица, имеющая верхнюю треугольную форму (ВТФ).

Доказательство. Возьмем вектор $x = (a_{11}, a_{21}, \ldots, a_{m1})^T$ — первый столбец матрицы A . Рассмотрим вектор

$$v = x + \|x\|z, \quad z = \underbrace{(1, 0, \ldots, 0)}_m^T$$

$$\text{и построим матрицу} \quad H_1 = E - 2 \frac{vv^T}{\|v\|^2}.$$

По доказанному выше

$$H_1 x = (-\|x\|, 0, 0, \ldots, 0)^T.$$

Тогда матрица $A_1 = H_1 A$ будет иметь следующий вид:

$$A_1 = H_1 A = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \times & \times & \cdots & \times \end{pmatrix}.$$

Пусть теперь $x = \left(\overset{(1)}{a_{22}}, \overset{(1)}{a_{32}}, \ldots, \overset{(1)}{a_{m2}} \right)$, где $\overset{(1)}{a_{i2}}$, $i = \overline{2, m}$ элементы второго столбца A_1 . По вектору x однозначно определяется элементарное отражение с матрицей $H \left((m-1) \times (m-1) \right)$, удовлетворяющей равенству

$$Hx = (-\|x\|, 0, \ldots, 0)^T.$$

Пусть $H_2 = \begin{pmatrix} 1 & & \\ & \theta^T & \\ & & H \end{pmatrix}$. Тогда матрица $A_2 = H_2 A_1$ имеет следующий вид:

$$A_2 = H_2 H_1 A = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \times & \cdots & \times \end{pmatrix}.$$

После $(m-1)$ -го шага получим матрицу $R = H_{m-1} H_{m-2} \ldots H_2 H_1 A$, имеющую ВТФ:

$$R = H_{m-1} H_{m-2} \ldots H_2 H_1 A = \begin{pmatrix} \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ 0 & 0 & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \times \end{pmatrix}.$$

Введем матрицу $Q = H_1 H_2 \ldots H_{m-1}$. Покажем, что матрица Q ортогональная, воспользовавшись свойством ортогональности элементарного отражения:

$$Q^{-1} = H_{m-1}^{-1} \ldots H_2^{-1} H_1^{-1} = H_{m-1}^T \ldots H_2^T H_1^T = Q^T.$$

Таким образом, справедливо разложение (74) матрицы A . В силу того, что в ходе преобразований на матрицу A не накладывались ограничения, разложение справедливо для произвольной матрицы.

Замечание. *Число операций, необходимые для вычисления QR-разложения матрицы A , зависит от вида матрицы A . Для произвольной матрицы число операций можно оценить величиной порядка m^3 , для матрицы с ВПТФ, — порядка m^2 , для тредиагональной матрицы — порядка m .*

Утверждение. *Если все собственные значения матрицы A вещественны, то последовательность матриц $\{A_k\}$ сходится к матрице, имеющей ВТФ:*

$$A_k \xrightarrow[k \rightarrow \infty]{} \begin{pmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{pmatrix}.$$

Если же матрица имеет комплексную пару собственных значений $\lambda_0 \pm i\lambda_1$, то её на главной диагонали предельной матрицы будут соответствовать клетка размера 2×2 :

$$A_k \xrightarrow[k \rightarrow \infty]{} \begin{pmatrix} \times & & & \times \\ & \times & & \\ & & \lambda_0 & \lambda_1 \\ & & -\lambda_0 & \lambda_0 \\ & & & \ddots \\ 0 & & & \times \end{pmatrix}.$$

Замечание 1. *Итерационный процесс останавливается, когда все элементы ниже главной диагонали, либо ниже побочной (в случае комплексно-сопряженных собственных значений) матрицы A_n при некотором n становятся равными нулю. Однако следует заметить, что в данном случае под нулем мы понимаем либо машинный ноль, либо число, меньшее некоторой заданной величины — необходимой точности вычисления.*

Замечание 2. *QR-алгоритм применим к произвольной матрице A .*

Замечание 3. *QR-алгоритм является очень затратным по необходимому числу операций и объему памяти, используемому для хранения промежуточных матриц.*

15. Метод ньютона решения нелинейных уравнений

Рассмотрим функцию $f(x)$, $x \in \mathbb{R}$ и уравнение

$$f(x) = 0. \quad (75)$$

Пусть x^* — вещественный корень этого уравнения, и определена его окрестность радиуса a , не содержащая других корней уравнения:

$$U_a(x^*) = \{x: |x - x^*| < a\},$$

причем заданная функция $f(x)$ определена на этой окрестности. Будем считать, что начальное приближение $x^0 \in U_a(x^*)$ задано. Пусть в $U_a(x^*)$ существует и не обращается в ноль непрерывная первая производная функции $f(x)$:

$$f'(x) \neq 0, \quad x \in U_a(x^*).$$

Разложим $f(x^*)$ по формуле Тейлора в малой окрестности точки $x \in U_a(x^*)$:

$$f(x^*) = f(x) + (x^* - x)f'(x) + \ldots$$

и отбросим в этом разложении величины, имеющие второй и выше порядок малости по $(x^* - x)$. Заменив x^* на x^{n+1} и x на x^n , получим уравнение

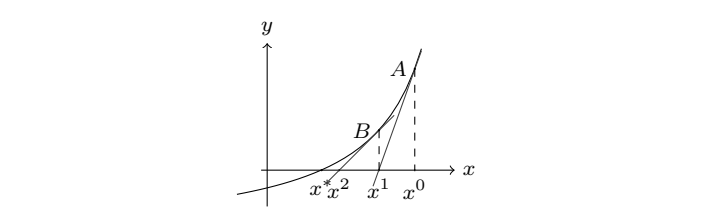
$$f(x^n) + (x^{n+1} - x^n)f'(x^n) = 0, \quad n \in \mathbb{Z}_+.$$

Учитывая, что $f'(x^n) \neq 0$, и разрешив последнее уравнение относительно x^{n+1} , имеем:

$$x^{n+1} = x^n - \frac{f(x^n)}{f'(x^n)}, \quad n \in \mathbb{Z}_+.$$

Определение. *Итерационный процесс поиска корня уравнения (75), задаваемый формулой (76), называется итерационным методом Ньютона.*

Дадим геометрическую интерпретацию метода Ньютона. Рассмотрим точку $A(x^0, f(x^0))$. Определим первую итерацию x^1 рассматриваемого процесса как абсциссу точки пересечения с осью Ox касательной к функции $f(x)$ в точке A . Аналогично получаем значение x^2 как точку пересечения с осью Ox касательной к функции $f(x)$ в точке $B(x^1, f(x^1))$. Продолжая таким образом, на n -м шаге получаем значение x^n , приближающее корень x^* уравнения (75) с заданной точностью.



17. Явная разностная схема для первой краевой задачи для уравнения теплопроводности. Аппроксимация, сходимость, устойчивость стр 80-87 (8 страниц, надо катать с телефона)

Определение. *Сеткой в заданной области называется совокупность конечного числа точек, принадлежащих данной области. Эти точки называются узлами сетки.*

В частности, равномерная сетка размера $(N-1)\times M$, $N, M\in\mathbb{N}$ в рассматриваемой области G вводится так:

$$\omega_h=\left\{x_i=ih,\; i=\overline{1,(N-1)}\right\},\quad \omega_\tau=\left\{t_j=j\tau,\; j=\overline{1,M}\right\},$$

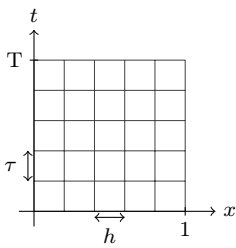
$$h=\frac{1}{N}>0,\;\tau=\frac{T}{M}>0.$$

Величину h назовем шагом по переменной x , величину τ —шагом по времени.

Тогда множество точек

$$\omega_{\tau h}=\omega_\tau\times\omega_h\subset G$$

задает равномерную сетку с шагом h по переменной x и шагом τ по времени в области G . Эта сетка изображена на рисунке.



Аналогичным образом введем равномерную сетку размера $(N+1)\times (M+1)$ на замыкании области G с теми же размерами шагов h и τ по переменной x и по переменной t соответственно. Эту сетку задаст множество точек

$$\overline{\omega}_{\tau h}=\overline{\omega}_\tau\times\overline{\omega}_h\subset\overline{G}=\{(x,t):\; x\in[0,1],\; t\in[0,T]\},$$

где

$$\overline{\omega}_h=\{x_i=ih,\; i=\overline{0,N}\},\quad \overline{\omega}_\tau=\{t_j=j\tau,\; j=\overline{0,M}\}.$$

В дальнейшем везде, где мы рассматриваем уравнение теплопроводности, будем использовать введенные сетки, если не указано иное. Рассмотрим уравнение теплопроводности с краевыми условиями первого рода:

$$\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (x,t)\in G=\{(x,t):\; x\in(0,1),\; t\in(0,T]\}$$

$$(102)$$

$$\begin{cases} u(0,t)=\mu_1(t)\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],\quad (103)$$

$$u(x,0)=u_0(x),\quad x\in[0,1]\quad (104)$$

Определение. *Сеточной функцией называется функция дискретного аргумента на заданной сетке, то есть такая функция определена только в узлах данной сетки.*

$$\frac{y_i^{n+1}-y_i^n}{\tau}=\frac{y_{i-1}^n-2y_i^n+y_{i+1}^n}{h^2}+f(x_i,t_n),\quad (x_i,t_n)\in\omega_{\tau h}.\quad (105)$$

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\overline{\omega}_\tau,\quad (106)$$

$$y_i^0=u_0(x_i),\quad x_i\in\overline{\omega}_h,\quad y_i^0=u_0(x_i),\quad x_i\in\overline{\omega}_h.\quad (107)$$

Определение. *Дискретным аналогом задачи (102)–(104), или ее разностной схемой, называется система линейных уравнений (105)–(107).*

Замечание 10. *В первой краевой задаче численные значения решения y_0^{n+1} и y_N^{n+1} равны значениям функций $\mu_1(t)$ и $\mu_2(t)$ соответственно при $t=t_{n+1}$ (хотя это и не обязательно). В случае краевых условий иного типа, аппроксимация краевых условий должна быть согласована по порядку погрешности с порядком аппроксимации уравнения. Определение аппроксимации и порядка погрешности аппроксимации будет дано ниже.*

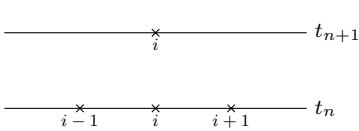
Замечание 11. *Заметим, что в уравнении (105) значения функции $f(x,t)$ не обязательно брать именно в узлах рассматриваемой сетки, можно использовать значения этой функции с некоторой «поправкой». Что именно имеется в виду под «поправкой», будет рассмотрено далее, а также будет показано, что выбор значений функции $f(x,t)$ для разностной схемы, использующей такую «поправку», позволяет получить более высокий порядок погрешности аппроксимации, а стало быть и более точное решение исходного уравнения.*

Замечание 12. *Качество и скорость решения численной задачи (105)–(107) во многом зависит от выбора числа узлов сетки $\omega_{\tau h}$: чем меньше узлов в сетке, тем меньше уравнений содержится в системе, тем проще и быстрее ее решать, но и приближение решения краевой задачи в этом случае будет более грубым.*

Замечание. *Вопросы сходимости и устойчивости разностной схемы являются ключевыми, однако обычно достаточно рассмотреть только один из этих двух вопросов: в конце курса будет доказано, что из устойчивости разностной схемы следует ее сходимость к решению исходной задачи при условии, что разностная схема аппроксимирует исходную задачу.*

Определение. *Совокупность узлов, которые участвуют в записи разностной схемы, называют шаблоном. Вернемся к изучению явной разностной схемы (105)–(107).*

В рассматриваемой разностной схеме использован четырехточечный шаблон, схематично изображенный на рисунке.



Определение. *Сеточная функция вида*

$$z_i^n=z(x_i,t_n)=y_i^n-u_i^n,\quad (x_i,t_n)\in\overline{\omega}_{\tau h}\quad (108)$$

называется погрешностью решения разностной схемы (105)–(107).

$$\psi_i^n=\frac{u_{i-1}^n-2u_i^n+u_{i+1}^n}{h^2}-\frac{u_{i-1}^{n+1}-u_i^{n+1}}{\tau}+f_i^n,\quad (109)$$

Определение. *Сеточная функция, задаваемая равенством (109) называется погрешностью аппроксимации разностной схемы (105)–(107) на решении исходной задачи.*

Задача. *Доказать, что $\psi_i^n=O(\tau+h^2)$.*

Теорема. *Пусть решение $u(x,t)$ задачи (102)–(104) обладает достаточной гладкостью (четыре раза дифференцируема по x и два раза по t). Тогда для сходимости решения разностной схемы (105)–(107) к решению исходной задачи (102)–(104) в норме $\|\cdot\|_C$ необходимо и достаточно, чтобы выполнялось условие:*

$$\gamma=\frac{\tau}{h^2}\leqslant 0.5.$$

При этом условии, выполняется оценка: $\|z^{n+1}-u^{n+1}\|_C\leqslant M_1(\tau+h^2)$, $n=0,1,\dots$ где $M_1>0$ —константа, не зависящая от τ и h .

Замечание 13. *Разностные схемы могут сходиться условно (и быть условно устойчивыми) и абсолютно. Условная сходимость определяется наличием ограничений на шаг сетки любого характера, для абсолютной сходимости требуется, чтобы такие ограничения отсутствовали. В приимеденной выше теореме условие сходимости имеет вид $\frac{\tau}{h^2}\leqslant 0.5$. Следовательно, явная разностная схема является условно сходящейся.*

Замечание 14. *Важно помнить, что сходимость и устойчивость разностной схемы доказываются в конкретной норме. В данном параграфе доказана сходимость и устойчивость решений разностной схемы (105)–(107) в норме $\|\cdot\|_C$, которая является достаточно сильной нормой, а значит, обеспечивает более точную оценку, по сравнению, например, со среднеквадратичной нормой.*

19. Симметричная разностная схема. Аппроксимация, сходимость, устойчивость стр 91-97

Рассмотрим уравнение теплопроводности с краевыми условиями первого рода:

$$\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (110)$$

$$(x,t)\in G=\{(x,t):\; x\in(0,1),\; t\in(0,T]\},\quad (111)$$

$$\begin{cases} u(0,t)=\mu_1(t)\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],\quad (112)$$

$$u(x,0)=u_0(x),\quad x\in[0,1].\quad (113)$$

Воспользуемся сетками $\omega_{\tau h}$ и $\overline{\omega}_{\tau h}$, введенными в первом параграфе данной главы на множествах G и $\overline{G}$ соответственно.

Введем вторую разностную производную для дискретной функции $y_i^n=y(x_i,t_n)$, определенной на множестве $\overline{\omega}_{\tau h}$:

$$y_{\overline{x},i}^n=\frac{y_{i-1}^n-2y_i^n+y_{i+1}^n}{h^2}.$$

Эта производная является дискретным аналогом второй производной по x функции $u(x,t)$.

Поставим в соответствие уравнению (110) его дискретный аналог в виде

$$\frac{y_i^{n+1}-y_i^n}{\tau}=\frac{y_{\overline{x},i}^{n+1}+y_{\overline{x},i}^n}{2}+f(x_i,t_{n+\frac{1}{2}}),\quad (114)$$

где $(x_i,t_{n+\frac{1}{2}})=(x_i,t_n+\frac{\tau}{2})\in\omega_{\tau h}$.

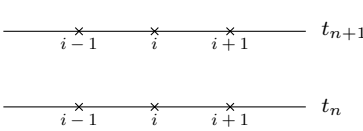
Определение. *Слой $t_{n+\frac{1}{2}}=t_n+\frac{\tau}{2}$ называется полуцелым слоем.*

Добавим краевые и начальные условия:

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\overline{\omega}_\tau,\quad (115)$$

$$y_i^0=u_0(x_i),\quad x_i\in\overline{\omega}_h.\quad (116)$$

В рассматриваемой разностной схеме использован шеститочечный шаблон вида



Заметим, что данная схема, с точки зрения нахождения численного решения, похожа на ту, которую мы рассматривали в предыдущем параграфе, в частности, матрица системы, соответствующей этой схеме, является трехдиагональной со строгим диагональным преобладанием. Это значит, что решение разностной схемы (114)–(116) такой задачи существует, единственно и находится с помощью метода прогонки.

Введем погрешность решения разностной схемы:

$$z_i^n=z(x_i,t_n)=y_i^n-u_i^n,$$

где $u_i^n=u(x_i,t_n)$, $(x_i,t_n)\in\overline{\omega}_{\tau h}$.

Выразив y_i^n из этого выражения и подставив его в уравнение (114), получим задачу относительно z_i^n :

$$\frac{z_i^{n+1}-z_i^n}{\tau}=\frac{z_{\overline{x},i}^{n+1}+z_{\overline{x},i}^n}{2}+\psi_i^n,\quad (x_i,t_n),(x_i,t_{n+1})\in\omega_{\tau h},\quad (117)$$

$$\begin{cases} z_0^{n+1}=0\\ z_N^{n+1}=0, \end{cases}\quad t_{n+1}\in\overline{\omega}_\tau,\quad (118)$$

$$z_i^0=0,\quad x_i\in\overline{\omega}_h,\quad (119)$$

где ψ_i^n —погрешность аппроксимации на решении исходной задачи (110)–(113):

$$\psi_i^n=\psi(x_i,t_n)=-\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{\overline{x},i}^{n+1}+u_{\overline{x},i}^n}{2}+f(x_i,t_{n+\frac{1}{2}}),\quad (120)$$

$$(x_i,t_n),(x_i,t_{n+1})\in\omega_{\tau h}.\quad (121)$$

Решение Разложим u_i^{n+1} и u_i^n в ряд Тейлора в окрестности точки $(x_i,t_{n+\frac{1}{2}})$:

$$u_i^{n+1}=u_i^{n+\frac{1}{2}}+u_{t,i}^{n+\frac{1}{2}}\frac{\tau}{2}+\frac{1}{2}u_{tt,i}^{n+\frac{1}{2}}\left(\frac{\tau}{2}\right)^2+O(\tau^3),$$

$$u_i^n=u_i^{n+\frac{1}{2}}-u_{t,i}^{n+\frac{1}{2}}\frac{\frac{1}{2}\tau}{2}+\frac{1}{2}u_{tt,i}^{n+\frac{1}{2}}\left(\frac{\tau}{2}\right)^2+O(\tau^3).$$

Подставим эти разложения в формулу для ψ_i^n :

$$\begin{aligned} \psi_i^n &= -\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{\overline{x},i}^{n+1}+u_{\overline{x},i}^n}{2}+f(x_i,t_n+0.5\tau)= \\ &= -u_{t,i}^{n+\frac{1}{2}}+O(\tau^2)+\frac{u_{\overline{x},i}^{n+1}+u_{\overline{x},i}^n}{2}+f_i^{n+\frac{1}{2}}. \end{aligned}$$

Теперь в представлении второй разностной производной разложим все вхождения функции $u(x_i,t_n)$ в ряд Тейлора по переменной x в окрестности точки $(x_i,t_{n+\frac{1}{2}})$. Приведа подобные слагаемые, получим

$$u_{\overline{x},i}^n=u_{\overline{x},i}^n+u_{x^4,i}^n\frac{h^2}{12}+O(h^4).$$

Применим это разложение для $u_{\overline{x},i}^{n+1}$, затем проведем еще одно разложение в ряд Тейлора в точке $(x_i,t_{n+\frac{1}{2}})$, теперь по переменной t :

$$\begin{aligned} u_{\overline{x},i}^{n+1} &= u_{\overline{x},i}^{n+1}+u_{x^4,i}^{n+1}\frac{h^2}{12}+O(h^4)= \\ &= u_{\overline{x},i}^{n+\frac{1}{2}}+u_{x^4,i}^{n+\frac{1}{2}}\frac{\tau}{2}+u_{x^4,i}^{n+1}\frac{h^2}{12}+u_{x^4,t,i}^{n+1}\frac{h^2}{12}\frac{\tau}{2}+O(\tau^2+h^4). \end{aligned}$$

То же самое проделаем для $u_{\overline{x},i}^n$:

$$\begin{aligned} u_{\overline{x},i}^{n+1} &= u_{\overline{x},i}^{n+1}+u_{x^4,i}^{n+1}\frac{h^2}{12}+O(h^4)= \\ &= u_{\overline{x},i}^{n+\frac{1}{2}}-u_{x^4,i}^{n+\frac{1}{2}}\frac{\tau}{2}+u_{x^4,i}^{n+1}\frac{h^2}{12}-u_{x^4,t,i}^{n+1}\frac{h^2}{12}\frac{\tau}{2}+O(\tau^2+h^4). \end{aligned}$$

Подставим эти разложения в формулу (120) для ψ_i^n и учтем выполнение уравнения теплопроводности (110):

$$\psi_i^n=(-u_{t,i}^{n+\frac{1}{2}}+u_{x^4,i}^{n+\frac{1}{2}}+f_i^{n+\frac{1}{2}})+u_{x^4,i}^{n+\frac{1}{2}}\frac{h^2}{12}+O(\tau^2+h^4)=O(\tau^2+h^2).$$

Теорема. *Пусть функция $u(x,t)$, являющаяся решением задачи для уравнения теплопроводности (110)–(113), имеет достаточную гладкость. Тогда симметричная разностная схема (114)–(116) сходится к решению исходной задачи со вторым порядком по τ и вторым порядком по h в $L_2(\omega_h)$ -норме пространства сеточных функций.*

Замечание. *Если в разностной задаче (114)–(116) взять нулевые краевые условия*

$$y_0^{n+1}=y_N^{n+1}=0,$$

то для y_i^n можно вывести априорную оценку, аналогичную полученной выше оценке (??):

$$\|y^{n+1}\|_{L_2(\omega_h)}\leqslant\|u_0\|_{L_2(\omega_h)}+\tau\sum_{j=0}^n\|f(t_j)\|_{L_2(\omega_h)}.$$

Эта оценка означает, что решение разностной схемы устойчиво в норме $L_2(\omega_h)$ по начальному условию u_0 и правой части уравнения.

21. Методы решения разностной задачи Дирихле стр 106-107

Рассмотрим разностную задачу Дирихле:

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}=\frac{y_{i-1,j}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}+y_{i,j+1}}{h_2^2}-f_{ij},\\ i=1,(N_1-1),\; j=1,(N_2-1),\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}\quad (122)$$

Для нахождения решения этой разностной схемы нужно решить СЛАУ с матрицей порядка $(N_1-1)\times(N_2-1)$. Заметим, что матрица системы разрежена, то есть среди элементов этой матрицы содержится большое число нулей. Очевидно, что использование классического метода Гаусса для решения такой системы не будет оптимальным. Существуют значительно более эффективные как прямые, так и итерационные методы решения системы (122) (см. [?]). Рассмотрим несколько итерационных методов, решающих поставленную задачу: методы Якоби, Зейделя и попеременно-треугольный итерационный метод.

Метод Якоби

Итерационный процесс задается схемой

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}^{(s+1)}=\frac{y_{i-1,j}^{(s)}+y_{i+1,j}^{(s)}}{h_1^2}+\frac{y_{i,j-1}^{(s)}+y_{i,j+1}^{(s)}}{h_2^2}-f_{ij},\\ i=1,(N_1-1),\; j=1,(N_2-1),\\ y_{ij}^{(s+1)}|_{\Gamma_h}=\mu_{ij}, \end{cases}$$

где $s\in\mathbb{Z}_+$, $y_{ij}^{(0)}$ —задано.

В методе Зейделя итерации определяются по правилу:

$$\begin{aligned} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}^{(s+1)}&=\frac{1}{h_1^2}\left(y_{i-1,j}^{(s+1)}+y_{i+1,j}^{(s)}\right)+\frac{1}{h_2^2}\left(y_{i,j-1}^{(s)}+y_{i,j+1}^{(s+1)}\right)-f_{ij},\\ i&=\overline{1,N_1-1},j=\overline{1,N_2-1},s=0,1,2,\dots,y_{ij}^{(s+1)}|_{\Gamma_h}=\mu_{ij}. \end{aligned}$$

Несмотря на то, что метод Зейделя формально является неявным, нетрудно предложить алгоритм вычисления решения на $s+1$ -й итерации по явным формулам. Порядок счета здесь следующий. Сначала находятся значения $y_{i,j}^{(s+1)}$, $j=\overline{1,N_2-1}$, используя $y_{i,j}^{(s)}$ и граничные значения μ_{ij} . Затем, используя $y_{i,j}^{(s+1)}$, находят $y_{i+1,j}^{(s+1)}$, $j=\overline{1,N_2-1}$ и т.д. Это означает, что вычисление значений $y_{i,j}^{(s+1)}$ на новой итерации осуществляется последовательно от левого нижнего угла (точки $x_{1,1}$) до правого верхнего угла (точки x_{N_1-1,N_2-1}). Эффективным методом решения системы разностных уравнений (122) является попеременно-треугольный итерационный метод. Запишем систему (122) в матричной форме

$$Ay=\varphi$$

с симметричной положительно определенной матрицей A и представим матрицу A в виде суммы

$$A=R_1+R_2,$$

где R_1 —нижняя треугольная, а R_2 —верхняя треугольная матрица. На главных диагоналях R_1 и R_2 стоят элементы $0.5a_{ii}$, где a_{ii} —элементы A , стоящие на главной диагонали.

Попеременно-треугольный итерационный метод имеет вид

$$(E+\omega R_1)(E+\omega R_2)\frac{y^{(s+1)}-y^{(s)}}{\tau}+Ay^{(s)}=\varphi,$$

где E —единичная матрица, ω и τ —числовые параметры. Метод сходится при $\omega>\frac{\tau}{2}>0$. В случае системы (122) матрицы R_1 и R_2 определяются соотношениями

$$(R_1y)_{ij}=\frac{y_{ij}-y_{i-1,j}}{h_1^2}+\frac{y_{ij}-y_{i,j-1}}{h_2^2},$$

$$(R_2y)_{ij}=\frac{y_{ij}-y_{i+1,j}}{h_1^2}+\frac{y_{ij}-y_{i,j+1}}{h_2^2}.$$

Алгоритм нахождения $y^{(s+1)}$ сводится к последовательному решению двух уравнений

$$(E+\omega R_1)W^{(s)}=\varphi-Ay^{(s)},$$

$$(E+\omega R_2)\frac{y^{(s+1)}-y^{(s)}}{\tau}=W^{(s)},$$

каждое из которых решается путем обращения треугольных матриц.

Замечание. *В параграфе 8 главы 1 было показано, что попеременно-треугольный итерационный метод решения систем линейных алгебраических уравнений требует для достижения заданной точности $\varepsilon>0$, числа итераций $\mathbf{n}_0(\varepsilon)$ на порядок меньше, чем методы Якоби, Зейделя, простой итерации. В силу этого он широко применяется для решения практических задач.*

23.1 Примеры численных методов решения задачи Коши. Погрешность аппроксимации общего 2-этапного метода Рунге - Кутты стр 112-117

В этой главе рассматривается задача Коши для системы обыкновенных дифференциальных уравнений

$$\begin{cases} \frac{du}{dt}=f(t,u(t)),\quad t>0,\\ u(0)=u_0, \end{cases}\quad (123)$$

где $u(t)=(u_1(t),u_2(t),\dots,u_m(t))^T$, $f(t,u(t))=(f_1(t,u(t)),\dots,f_m(t,u(t)))^T$.

В дальнейшем при построении и исследовании численных методов будем предполагать, что искомое решение задачи (123) $u(t)$ существует, единственно и обладает требуемыми свойствами гладкости.

В настоящее время наибольшее распространение получили две группы численных методов решения задачи Коши:

- Методы Рунге-Кутты;
- Многошаговые разностные методы, наиболее известными из которых являются методы Адамса.

Рассмотрим следующую задачу Коши:

$$\begin{cases} \frac{du}{dt}=f(t,u(t)),\quad t>0,\\ u(0)=u_0, \end{cases}$$

Введем сетку по времени с постоянным шагом $\tau>0$, то есть множество точек

$$\omega_\tau=\{t_n=n\tau,\; n\in\mathbb{Z}_+\},$$

и обозначим $u_n=u(t_n)$, $f_n=f(t_n,u_n)$. В дальнейшем точное решение задачи (123) будем обозначать буквой u , а приближенное решение—буквой y .

Пример 1. Метод Эйлера представляет собой разностное уравнение:

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=f(t_n,y_n),\quad t_n\in\omega_\tau\\ y_0=u_0,\quad n\in\mathbb{Z}_+. \end{cases}\quad (125)$$

Эта схема является явной, так как значение численного решения в каждой следующей точке t_{n+1} , $n\in\mathbb{Z}_+$ находится по явной формуле:

$$y_{n+1}=y_n+\tau f_n,\quad n\$$

23.2 Примеры численных методов решения задачи Коши. Погрешность аппроксимации общего 2 - этапного метода Рунге - Кутта Оценка погрешности общего двухэтапного метода Рунге-Кутта.

Рассмотрим общий вид двухэтапного метода Рунге–Кутта для уравнения (124):

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=\sigma_1K_1+\sigma_2K_2, & n\in\mathbb{Z}_+\\ y_0=u_0,\\ K_1=f(t_n,y_n), & K_2=f(t_n+a_2\tau,y_n+b_{21}\tau f(t_n,y_n)), \end{cases}\quad (134)$$

где $\sigma_1,\sigma_2,a_2,b_{21}\in\mathbb{R}$ —некоторые числа, от выбора которых зависит как погрешность аппроксимации, так и точность численного решения. Подставим значения K_1 и K_2 в первое уравнение системы (134):

$$\frac{y_{n+1}-y_n}{\tau}=\sigma_1f(t_n,y_n)+\sigma_2f(t_n+a_2\tau,y_n+b_{21}\tau f(t_n,y_n)).$$

Рассмотрим погрешность аппроксимации разностной схемы (134) на решении задачи (124):

$$\psi_n=-\frac{u_{n+1}-u_n}{\tau}+\sigma_1f(t_n,u_n)+\sigma_2f(t_n+a_2\tau,u_n+b_{21}\tau f(t_n,u_n)).\quad (135)$$

Разложим u_{n+1} в ряд Тейлора в окрестности точки t_n :

$$\frac{u_{n+1}-u_n}{\tau}=u_n'+\frac{\tau}{2}u_n''+\mathrm{O}(\tau^2).$$

Далее разложим $f(t_n+a_2\tau,u_n+b_{21}\tau f_n)$ в окрестности точки (t_n,u_n) :

$$f(t_n+a_2\tau,u_n+b_{21}\tau f(t_n,u_n))=f(t_n,u_n)+a_2\tau\frac{\partial f_n}{\partial t}+b_{21}\tau f_n\frac{\partial f_n}{\partial u}+\mathrm{O}(\tau^2).$$

Заметим, что в силу уравнения (124)

$$u''=\frac{d}{dt}\left(\frac{du}{dt}\right)=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial u}\frac{\partial u}{\partial t}=\frac{\partial f}{\partial t}+f\frac{\partial f}{\partial u}.$$

Тогда погрешность аппроксимации ψ_n принимает вид:

$$\begin{aligned} \psi_n&=-u_n'-0.5\tau\left(\frac{\partial f_n}{\partial t}+f_n\frac{\partial f_n}{\partial u}\right)+\mathrm{O}(\tau^2)+\sigma_1f(t_n,u_n)+\\ &+\sigma_2\left(f(t_n,u_n)+\tau a_2\frac{\partial f_n}{\partial t}+\tau b_{21}f_n\frac{\partial f_n}{\partial u}\right)+\mathrm{O}(\tau^2). \end{aligned}$$

Сгруппируем слагаемые следующим образом:

$$\begin{aligned} \psi_n&=-u_n'+(\sigma_1+\sigma_2)f(t_n,u_n)+\tau\left((a_2\sigma_2-0.5)\frac{\partial f_n}{\partial t}+ \right.\\ &\left. (b_{21}\sigma_2-0.5)f_n\frac{\partial f_n}{\partial u}+\mathrm{O}(\tau^2)\right). \end{aligned}$$

Чтобы получить оценку погрешности аппроксимации ψ_n со вторым порядком по τ , необходимо избавиться от слагаемых, содержащих τ в первой степени. Для этого потребуем выполнение следующих условий:

- $\sigma_1+\sigma_2=1$ (это условие называется условием аппроксимации).
- $\sigma_2a_2=\sigma_2b_{21}=0.5$.

Тогда погрешность аппроксимации этого метода имеет второй порядок малости по τ :

$$\psi_n=\mathrm{O}(\tau^2).$$

Замечание. В случае выполнения только первого условия погрешность аппроксимации имеет первый порядок по τ .

В записи общего метода Рунге–Кутта используется несколько параметров, что обеспечивает широту класса описываемых этим методом разностных схем. Однако в двухэтапном методе Рунге–Кутта не имеет смысла пользоваться двумя параметрами σ_1 и σ_2 , так наилучшая оценка погрешности метода достигается при $\sigma_1+\sigma_2=1$, поэтому, как правило, в двухэтапном методе Рунге–Кутта выбирают один параметр $\sigma=\sigma_2$, тогда $\sigma_1=1-\sigma$. Если положить $a=a_2=b_{12}$, то двухэтапный метод Рунге–Кутта запишется, как однопараметрическое по σ семейство разностных схем вида:

$$\frac{y_{n+1}-y_n}{\tau}=(1-\sigma)K_1+\sigma K_2,$$

где $K_1=f(t_n,y_n)$, $K_2=f(t_n+a\tau,y_n+a\tau f(t_n,y_n))$.

Пример. Рассмотрим примеры разностных схем, являющихся частными случаями общего двухэтапного метода Рунге–Кутта.

- При $\sigma=1$, $a=a_2=0.5$, $b=b_{21}=0.5$ мы получим схему Рунге–Кутта «предиктор–корректор» (129), которую мы уже рассматривали. Погрешность этой схемы равна $\mathrm{O}(\tau^2)$.
- Если положить $\sigma=0.5$, $a=1$, $b=1$, то мы получим симметричную разностную схему:

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=0.5(f(t_n,y_n)+f(t_n+\tau,y_n+\tau f_n)), & n\in\mathbb{Z}_+\\ y_0=u_0. \end{cases}\quad (136)$$

Эта разностная схема является очень эффективной, имеет второй порядок точности по τ и часто используется на практике.

22. Основные понятия теории разностных схем: аппроксимация, сходимость, устойчивость

стр 107-111
Пусть дана исходная дифференциальная задача. запишем ее в форме операторного уравнения.

$$L(u(x))=f(x),\quad x\in G,\quad (137)$$

где G – область изменения независимых переменных x (аргумент x может быть многомерным), $f(x)$ – заданная функция, L – линейный дифференциальный оператор. Предполагается, что начальные и граничные условия учитываются видом оператора L и правой частью $f(x)$. Будем считать, что исходная задача корректно поставлена. Это означает, что ее решение существует, оно единственно и непрерывно зависит от правой части $f(x)$.

Для построения разностной схемы, прежде всего в области G вводится разностная сетка G_h — конечное множество точек, принадлежащих G . Точки $x\in G_h$ называются узлами сетки. Параметр h (шаг сетки) характеризует плотность заполнения области G точками G_h . Будем считать h вектором, для которого определена норма $|h|$. Фактически имеют дело с последовательностью сеток, число узлов которых $N=N(h)$ увеличивается с уменьшением $|h|$. Обычно число узлов $N=N(h)$ сетки G_h неограниченно возрастает при $|h|\rightarrow 0$.

После того, как введена сетка G_h , функции непрерывного аргумента x , определенные для $x\in G$, заменяют сеточными функциями, т.е. функциями, определенными только в точках сетки G_h . Правую часть $f(x)$ уравнения (137) замениют приближенно некоторой сеточной функцией $\varphi_h(x)$, $x\in G_h$, а дифференциальный оператор L — линейным разностным оператором L_h . В результате вместо дифференциального уравнения (137) получают систему разностных уравнений

$$L_hy_h(x)=\varphi_h(x),\quad x\in G_h,\quad (138)$$

которая называется разностной схемой.

Будем предполагать в дальнейшем, что нормы $\|\cdot\|_0$ и $\|\cdot\|_h$ согласованы в том смысле, что для любого элемента $u\in B_0$ существует

$$\lim_{|h|\rightarrow 0}\|u_h\|_h=\|u\|_0.\quad (139)$$

Введем основные понятия теории разностных схем.

Пусть $u(x)$ — решение исходной задачи (137), $u_h(x)=P_hu(x)$ — проекция этого решения на пространство сеточных функций B_h , y_h — решение разностной задачи (138).

Определение. *Сеточная функция*

$$z_h(x)=y_h(x)-u_h(x),\,x\in G_h,$$

называется *погрешностью разностной схемы* (138).

Подставим выражение $y_h(x)=z_h(x)+u_h(x)$ в уравнение (138) и, используя линейность оператора L_h , получим

$$L_hz_h(x)+L_hu_h(x)=\varphi_h(x).$$

Отсюда следует

$$L_hz_h(x)=\psi_h(x),x\in G_h,\quad (140)$$

где

$$\psi_h(x)=\varphi_h(x)-L_hu_h(x).\quad (141)$$

Определение. *Сеточная функция $\psi_h(x)$, определенная по формуле (141), называется погрешностью аппроксимации (или невязкой) разностной задачи (138) на решении исходной дифференциальной задачи (137).*

Таким образом, для погрешности разностной схемы $z_h(x)$ получаем задачу (140) той же структуры, что и разностная задача (138), но с правой частью, представляющей собой погрешность аппроксимации на решении.

Спроектируем исходное уравнение (137) на пространство B_h , т.е. запишем сеточное уравнение

$$P_h(Lu)=P_hf.$$

Представим погрешность аппроксимации $\psi_h(x)$ в виде суммы

$$\psi_h(x)=\psi_h^{(1)}(x)+\psi_h^{(2)}(x),$$

где $\psi_h^{(1)}(x)=P_h(Lu(x))-L_hu_h(x)$, $\psi_h^{(2)}(x)=\varphi_h(x)-f_h(x)$, $f_h(x)=P_hf$. Функция $\psi_h^{(1)}(x)$ называется погрешностью аппроксимации дифференциального оператора L разностным оператором L_h . Функция $\psi_h^{(2)}(x)$ называется погрешностью аппроксимации правой части.

Определение. *Говорят, что разностная схема (138) аппроксимирует исходную задачу (137), если $\|\psi_h\|_h\rightarrow 0$ при $|h|\rightarrow 0$.*

Определение. *Разностная схема (138) имеет k -й порядок аппроксимации, если существуют положительные константы M_1 и k не зависящие от h и такие, что при всех достаточно малых h выполняется оценка*

$$\|\psi_h\|_h\leqslant M_1|h|^k.$$

Введем понятие корректности разностной схемы.

Определение. *Разностная схема (138) называется корректной, если при всех достаточно малых h :*

- ее решение $y_h(x)\in B_h$ существует и единственно при любых правых частях $\psi_h\in B_h$,
- существует постоянная $M_2>0$, не зависящая от h и такая, что при любых правых частях φ_h для решения задачи (138) справедлива оценка:

$$\|y_h\|_h\leqslant M_2\|\varphi_h\|_h.\quad (142)$$

Существенным отличием от определения корректности дифференциальной задачи является условие независимости константы M_2 от шагов сетки. Требование 1) в определении корректности означает существование оператора L_h^{-1} , обратного к оператору L_h , а требование 2) — равномерную по h ограниченность оператора L_h^{-1} . Свойство разностной схемы, выраженное неравенством (142), означает непрерывную и равномерную по h зависимость ее решения от правой части. Это свойство называют устойчивостью разностной схемы.

Определение. *Решение разностной задачи (138) сходится к решению исходной задачи (137), (или, более коротко, разностная схема сходится), если*

$$\lim_{h\rightarrow 0}\|z_h\|_h=\lim_{h\rightarrow 0}\|y_h-u_h\|_h=0.$$

Определение. *Говорят, что разностная схема имеет k -й порядок точности, если существует постоянная $M_3>0$, не зависящая от h и постоянная $k>0$ такие, что*

$$\|z_h\|_h\leqslant M_3|h|^k.$$

Покажем, что из аппроксимации и устойчивости разностной схемы следует ее сходимость.

Теорема 8. *(Филиппова). Предположим, что исходная задача (137) поставлена корректно. Пусть разностная схема (138) аппроксимирует исходную задачу (137) и является корректной. Тогда решение разностной задачи (138) сходится к решению исходной задачи (137), причем порядок точности разностной схемы совпадает с порядком аппроксимации.*

Доказательство. Рассмотрим задачу для погрешности (140). Так как уравнение (140) отличается от уравнения (138) только правой частью, из требования устойчивости следует, что для $z_h(x)$ справедлива оценка

$$\|z_h\|_h\leqslant M_2\|\psi_h\|_h.$$

Правая часть этого неравенства стремится к нулю при $|h|\rightarrow 0$, а так как M_2 не зависит от h и, по условию теоремы, разностная схема аппроксимирует исходную задачу. Следовательно $\|z_h\|_h\rightarrow 0$ при $|h|\rightarrow 0$, т.е. схема сходится.

Так как схема имеет k -й порядок аппроксимации, то

$$\|\psi_h\|_h\leqslant M_1|h|^k$$

и $\|z_h\|_h\leqslant M_3|h|^k$, $M_3=M_1M_2$, что означает, что схема (138) имеет k -й порядок точности. □

Доказанная теорема позволяет разделить исследование сходимости разностной схемы на два этапа: исследование погрешности аппроксимации на решении и получение оценок вида (142), т.е. исследование устойчивости. Как правило, второй этап более трудный, чем первый.

Замечание. *При доказательстве теоремы Филиппова условие согласованности норм (139) не использовалось. Это условие нужно для того, чтобы гарантировать единственность предельной функции. Теорема утверждает, что последовательность y_h сходится к точному решению $u\in B_0$ в том смысле, что $\|y_h-u_h\|_h\rightarrow 0$, при $|h|\rightarrow 0$. Но из теоремы не следует, что не может существовать такая функция $v\in B_0$ (не являющаяся решением задачи (137)), для которой также выполнено условие $\|y_h-v_h\|_h\rightarrow 0$, при $|h|\rightarrow 0$.*

Требование согласованности норм (139) устраняет эту неоднозначность. В самом деле,

$$\|u_h-v_h\|_h=\| (y_h-v_h)+(u_h-y_h)\|_h\leqslant\|y_h-v_h\|_h+\|u_h-y_h\|_h.\quad (143)$$

Так как $\|y_h-v_h\|_h\rightarrow 0$, $\|u_h-y_h\|_h\rightarrow 0$, то из (143) следует, что $\|u_h-v_h\|_h\rightarrow 0$, при $|h|\rightarrow 0$.

Но тогда из (139) получим

$$\lim_{h\rightarrow 0}\|u_h-v_h\|_h=\|u-v\|_0=0,$$

и, следовательно, $u\equiv v$.

20. Сходимость разностной задачи дирихле для ур-я Пуассона

стр 101-106

Продолжаем рассматривать задачу Дирихле

$$\begin{cases} \frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}=f(x_1,x_2), & (x_1,x_2)\in G,\\ u(x_1,x_2)|_{\Gamma}=\mu(x_1,x_2) \end{cases}\quad (144)$$

и поставим в соответствие задаче разностную схему

$$\begin{cases} y_{\overline{x}_{1,i,j}}+y_{\overline{x}_{2,2,i,j}}=f_{ij}, & x_{ij}=\left(x_1^{(i)},x_2^{(j)}\right)\in\omega_h,\\ y_{ij}|_{\Gamma_h}=\mu_{ij}, \end{cases}\quad (145)$$

Запишем разностную схему в виде:

$$\begin{cases} \frac{y_{i-1,j}-2y_{ij}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}-2y_{ij}+y_{i,j+1}}{h_2^2}=f_{ij},\\ i=\overline{1,(N_1-1)},\,j=\overline{1,(N_2-1)},\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}$$

Напомним, что f_{ij},μ_{ij} — значения непрерывных функций $f(x_1,x_2)$ и $\mu(x_1,x_2)$ в узлах сетки ω_h . Разрешим эту схему относительно центрального узла x_{ij} :

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}=\frac{y_{i-1,j}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}+y_{i,j+1}}{h_2^2}-f_{ij},\\ i=\overline{1,(N_1-1)},\,j=\overline{1,(N_2-1)},\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}\quad (146)$$

Для того чтобы эта система имела решение при любых значениях функций $f(x_1,x_2)$ и $\mu(x_1,x_2)$, необходимо и достаточно, чтобы однородная система линейных уравнений имела только тривиальное решение.

Пусть H_{N_1-1,N_2-1} —пространство сеточных функций, определенных на сетке ω_h и обращающихся в нуль на границе Γ_h . Введем норму в этом пространстве:

$$\|v\|_C=\max_{\substack{1\leqslant i\leqslant N_1-1\\ 1\leqslant j\leqslant N_2-1}}|v_{ij}|,\,v\in H_{N_1-1,N_2-1}.$$

Теорема 9. *Однородная система линейных уравнений*

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)v_{ij}=\frac{v_{i-1,j}+v_{i+1,j}}{h_1^2}+\frac{v_{i,j-1}+v_{i,j+1}}{h_2^2},\\ i=\overline{1,(N_1-1)},\,j=\overline{1,(N_2-1)},\\ v_{ij}|_{\Gamma_h}=0 \end{cases}$$

имеет единственное решение, и оно является тривиальным:

$$v_{ij}=0,\,x_{ij}\in\overline{\omega}_h.$$

Следствие. *Разностная задача*

$$\begin{cases} y_{\overline{x}_{1,i,j}}+y_{\overline{x}_{2,2,i,j}}=f_{ij}, & x_{ij}=\left(x_1^{(i)},x_2^{(j)}\right)\in\omega_h,\\ y_{ij}|_{\Gamma_h}=\mu_{ij} \end{cases}$$

имеет единственное решение при любых значениях f_{ij} и μ_{ij} , $x_{ij}\in\omega_h$.

Лемма (принцип максимума). *Пусть для сеточной функции v , определенной на сетке ω_h , выполнены неравенства*

$$v_{ij}\geqslant 0,\,x_{ij}\in\Gamma_h,$$

$$L_hv_{ij}\geqslant 0,\,x_{ij}\in\omega_h.$$

Тогда справедливо следующее неравенство:

$$v_{ij}\geqslant 0,\,x_{ij}\in\overline{\omega}_h.$$

Следствие. *Рассмотрим две разностные задачи*

$$L_hy_{ij}=\varphi_{ij},\,x_{ij}\in\omega_h,\,y_{ij}|_{\Gamma_h}-\text{заданы,}$$

$$L_hY_{ij}=\Phi_{ij},\,x_{ij}\in\omega_h,\,Y_{ij}|_{\Gamma_h}-\text{заданы.}$$

Если выполнены неравенства

$$|y_{ij}|\leqslant Y_{ij},\,x_{ij}\in\Gamma_h,$$

$$|\varphi_{ij}|\leqslant \Phi_{ij},\,x_{ij}\in\omega_h,$$

то справедливо следующее неравенство:

$$|y_{ij}|\leqslant Y_{ij},\,x_{ij}\in\overline{\omega}_h.$$

Теорема 10. *Пусть решение задачи (144) четыре раза непрерывно дифференцируемо в $\overline{G}$. Тогда решение разностной задачи (??) сходится к решению исходной задачи в сеточной норме C , и имеет место оценка*

$$\left\|y_{ij}-u\left(x_1^{(i)},x_2^{(j)}\right)\right\|_C\leqslant M\left(h_1^2+h_2^2\right),$$

где $M>0$ —константа, не зависящая от h_1 и h_2 .

18 Чисто невязная разностная схема. Аппроксимация, сходимость, устойчивость

стр 88-91

Рассмотрим уравнение теплопроводности с краевыми условиями пер-вого рода:

$$\frac{\partial u(x,t)}{\partial t}=-\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (147)$$

$$(x,t)\in G=\{(x,t):x\in(0,1),\,t\in(0,T]\}$$

$$\begin{cases} u(0,t)=\mu_1(t),\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],u(x,0)=u_0(x),\quad x\in[0,1].\quad (148)$$

Поставим в соответствие задаче (147)–(148) следующую разностную схему:

$$\begin{aligned} \frac{y_i^{n+1}-y_i^n}{\tau}&=\frac{y_{i-1}^{n+1}-2y_i^{n+1}+y_{i+1}^{n+1}}{h^2}+f(x_i,t_{n+1}),\\ (x_i,t_n),(x_i,t_{n+1})&\in\omega_{\tau h}, \end{aligned}\quad (149)$$

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\overline{\omega}_{\tau},y_i^0=u_0(x_i),\quad x_i\in\overline{\omega}_h,\quad (150)$$

где $y_i^n=y(x_i,t_n)$ —искомое численное решение в точке $(x_i,t_n)\in\overline{\omega}_{\tau h}$. Эта система имеет трехдиагональную матрицу порядка $(N-1)$:

$$A=\begin{pmatrix} 1+2\gamma & -\gamma & 0 & \ldots & 0 & 0 \\ -\gamma & 1+2\gamma & -\gamma & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+2\gamma & -\gamma \\ 0 & 0 & 0 & 0 & \ldots & -\gamma & 1+2\gamma \end{pmatrix},$$

обладающую строгим диагональным преобладанием: $a_{ii}>\sum_{\substack{j=1\\ j\neq i}}^N|a_{ij}|$, $i=\overline{1,(N-1)}$.

ψ_i^n —погрешность аппроксимации на решении:

$$\psi_i^n=\psi(x_i,t_n)=-\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}}{h^2}+f(x_i,t_{n+1}).\quad (151)$$

Задача. *Доказать, что $\psi_i^n=\mathrm{O}(\tau+h^2)$.*

Теорема. *Пусть функция $u(x,t)$ имеет достаточно гладкость (четыре раза дифференцируема по x и два раза по t). Тогда чисто невязная разностная схема сходится к решению исходной задачи в норме $\|\cdot\|_C$ с первым порядком точности по τ и вторым порядком точности по h .*

24.1 Общая формулировка m-этапного метода Рунге-Кутта. Оценка точности 2-этапного метода Рунге-Кутта

§1 Общий m-этапный метод Рунге–Кутта

Рассмотрим задачу Коши для нелинейного обыкновенного дифференциального уравнения первого порядка:

$$\begin{cases} \frac{du}{dt}=f(t,u(t)), & t>0 \\ u(0)=u_0, \end{cases} \quad (1)$$

где функции $u(t)$ и $f(t,u)$ обладают достаточной гладкостью в соответствующих областях. Считаем, решение $u(t)$ существует и единственно.

Введем равномерную сетку в области $t\geqslant 0$ с шагом $\tau>0$:

$$\omega_\tau=\{t_n=n\tau,\;\tau>0,\;n\in\mathbb{Z}_+\}.$$

Рассмотрим сеточную функцию $y_n=y(t_n)$, заданную на сетке ω_τ . Пусть значения этой функции в узлах сетки y_n приближают значения $u_n=u(t_n)$. Обозначим $f_n=f(t_n,y_n)$.
Общая идея m-этапного метода Рунге–Кутта заключается в том, что для вычисления значения приближенного решения в каждой следующей точке t_{n+1} вводятся m дополнительных этапов. Промежуточные значения на каждом шаге $n\in\mathbb{Z}_+$ вычисляются по следующим формулам:

$$\begin{aligned} K_1&=f(t_n,y_n), \\ K_2&=f(t_n+a_2\tau,y_n+b_{21}\tau K_1), \\ K_3&=f(t_n+a_3\tau,y_n+b_{31}\tau K_1+b_{32}\tau K_2), \\ &\ldots \\ K_m&=f(t_n+a_m\tau,y_n+b_{m1}\tau K_1+b_{m2}\tau K_2+\ldots+b_{mm-1}\tau K_{m-1}). \end{aligned}$$

При этом разностная схема для исходной задачи (1) имеет вид

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}= \sigma_1K_1+\sigma_2K_2+\ldots+\sigma_mK_m \\ y_0=u_0, \quad n\in\mathbb{Z}_+, \end{cases} \quad (2)$$

где $\sigma_1,\sigma_2,\ldots,\sigma_m\in\mathbb{R}$.

Будем также считать, что выполнено следующее условие аппроксимации, без которого рассмотрение метода не имеет смысла:

$$\sum_{i=1}^m\sigma_i=1.$$

Замечание. Заметим, что формулы m-этапного метода Рунге–Кутта достаточно громоздки. Это является одной из причин того, что на практике редко используются методы Рунге–Кутта для $m>4$.

Приведем примеры трех- и четырех- этапных методов Рунге–Кутта, имеющих третий и четвертый порядок точности соответственно.

Пример 1. $m=3$:

$$\frac{y_{n+1}-y_n}{\tau}=\frac{1}{6}(K_1+4K_2+K_3),$$

где

$$\begin{aligned} K_1&=f(t_n,y_n), \\ K_2&=f(t_n+0.5\tau,y_n+0.5\tau K_1), \\ K_3&=f(t_n+\tau,y_n-\tau K_1+2\tau K_2). \end{aligned}$$

Данная схема имеет третий порядок точности по τ .

Пример 2. $m=4$:

$$\frac{y_{n+1}-y_n}{\tau}=\frac{1}{6}(K_1+2K_2+2K_3+K_4),$$

где

$$\begin{aligned} K_1&=f(t_n,y_n), \\ K_2&=f(t_n+0.5\tau,y_n+0.5\tau K_1), \\ K_3&=f(t_n+0.5\tau,y_n+0.5\tau K_2), \\ K_4&=f(t_n+\tau,y_n+\tau K_3). \end{aligned}$$

Данная схема имеет четвертый порядок точности по τ .

25.1 Многошаговые разностные методы. Погрешность аппроксимации. Понятие устойчивости

стр120-129 (9 страниц, надо катать с телефона)
Рассмотрим задачу Коши для нелинейного обыкновенного дифференциального уравнения первого порядка:

$$\begin{cases} \frac{du}{dt}=f(t,u(t)), & t>0, \\ u(0)=u_0, \end{cases} \quad (3)$$

где функции $u(t)$ и $f(t,u)$ обладают достаточной гладкостью. Считаем, что решение $u(t)$ существует и единственно.

Определение. Лине́йным m-шаговым разностным методом решения задачи (3) называется разностная схема вида

$$\sum_{k=0}^ma_ky_{n-k}=\sum_{k=0}^mb_kf_{n-k}, \quad (4)$$

где $m\in\mathbb{N}$, $\tau>0$ – шаг сетки ω_τ , $a_k,b_k\in\mathbb{R}$, $k=\overline{0,m}$, причем $a_0\neq 0,b_m\neq 0$.

Замечание. Уравнение (4) следует рассматривать как рекуррентное соотношение, выражающее новое значение $y_n=y(t_n)$ через найденные ранее значения $y_{n-1},y_{n-2},\ldots,y_{n-m}$. Уравнение (4) определено для $n=m,m+1,\ldots$ и требует для начала расчета задания m начальных значений $y_0,y_1,\ldots,y_{m-1}$. Значение $y_0=u(0)$ определяется исходной задачей (3), а величины $y_1,\ldots,y_{m-1}$ можно вычислить с помощью других методов, например, с помощью рассмотренного выше метода Рунге–Кутта. В дальнейшем будем предполагать, что величины $y_0,y_1,\ldots,y_{m-1}$ уже заданы.

Если в разностной схеме (4) $b_0=0$, то рассматриваемый метод называется явным, и искомое значение y_n выражается явным образом через предыдущие: $\frac{a_0}{\tau}y_n=\sum_{k=1}^mb_kf_{n-k}-\sum_{k=1}^ma_ky_{n-k}$.

Если $b_0\neq 0$, то метод называется неявным, и для нахождения y_n приходится решать нелинейное уравнение $\frac{a_0}{\tau}y_n-b_0f(t_n,y_n)=F(y_{n-1},\ldots,y_{n-m})$, где $F(y_{n-1},\ldots,y_{n-m})=\sum_{k=1}^m(b_kf_{n-k}-\frac{a_k}{\tau}y_{n-k})$.

Обычно это уравнение решают итерационным методом Ньютона, выбирая начальное приближение равным y_{n-1} . Заметим, что коэффициенты уравнения (4) определены с точностью до множителя. Для определенности будем считать, что выполнено условие $\sum_{k=0}^mb_k=1$.

Определение. Погрешностью аппроксимации разностной схемы (4) на решении исходной задачи (3) называется сеточная функция

$$\psi_n=-\sum_{k=0}^m\frac{a_k}{\tau}u_{n-k}+\sum_{k=0}^mb_kf(t_n-k,u_{n-k}), \quad (5)$$

заданная на сетке ω_τ , где $u_n=u(t_n)$ – решение исходной задачи (3).

Для сходимости окончательно получаем следующую систему уравнений:

$$\begin{cases} \sum_{k=1}^mka_k=-1, \\ \sum_{k=0}^mk^{l-1}(ka_k+lb_k)=0, \quad l=\overline{2,p}, \end{cases} \quad (6)$$

в которой коэффициенты a_0 , b_0 вычисляются по формулам $a_0=-\sum_{k=1}^ma_k$, $b_0=1-\sum_{k=1}^mb_k$. Таким образом, мы уменьшили число уравнений в системе до p и число неизвестных до $2m$. Чтобы система не была переопределенной (в таких системах число уравнений больше числа неизвестных) необходимо выполнение условия $p\leqslant 2m$. Таким образом наибольший возможный порядок аппроксимации неявных m-шаговых разностных методов равен $2m$, явных $-(2m-1)$, так как в явных методах $b_0=0$, и число неизвестных в системе (6) меньше на единицу по сравнению с системой, записанной для неявного метода.

Замечание 1. Если убрать последние p уравнений системы (6), $p=1,(p-1)$, то получим условия, обеспечивающие порядок погрешность аппроксимации $O(\tau^{p-n})$.

Замечание 2. В практике вычислений наибольшее распространение получили методы Адамса, которые представляют собой частный случай многошаговых методов (4), когда производная $u'(t)$ в исходном уравнении аппроксимируется по двум крайним точкам t_{n-1} и t_n , то есть $a_0=1$, $a_1=-1$, $a_k=0$, $k=2,m$:

$$\frac{y_n-y_{n-1}}{\tau}=\sum_{k=0}^mb_kf_{n-k}.$$

Замечание 3. Разностные схемы вида (4), обладающие наивысшими порядками аппроксимации на решении исходного уравнения, неустойчивы и не могут быть использованы на практике. Максимальный порядок аппроксимации устойчивого неявного m-шагового метода не превосходит $(m+1)$, если m нечетно, и не превосходит $(m+2)$, если m четно. Порядок аппроксимации устойчивых явных схем не превосходит m . Подробнее понятие устойчивости m-шагового разностного метода мы рассмотрим в следующем параграфе.

Достоинства и недостатки многошаговых разностных методов по сравнению с методом Рунге–Кутта.

Достоинства:

- Формулы многошаговых методов значительно проще.
- Многошаговые методы позволяют достигать большей точности.

Недостатки:

- В многошаговых методах необходимо хранить в памяти большее число элементов — значения нескольких предыдущих шагов вместо одного.
- Многошаговые методы требуют наличия «разгонного этапа», то есть значений нескольких первых шагов, которые нельзя вычислить по многошаговым формулам. Как мы уже упоминали, эти значения обычно вычисляют с помощью метода Рунге–Кутта.

Понятие устойчивости разностного метода

Численный метод называется устойчивым, если погрешности, допущенные на каком-то этапе вычислений, не оказывают существенного влияния на результат.

$$y_{n+1}=qy_n, \quad \tilde{y}_n=y_n+\delta_n.$$

$$\tilde{y}_{n+1}=q\tilde{y}_n=qy_n+q\delta_n=y_{n+1}+q\delta_n,$$

$$\begin{cases} \frac{du(t)}{dt}+\lambda u(t)=0, \quad \lambda>0,\;t>0, \\ u(0)=u_0. \end{cases} \quad (7)$$

Пример 3. Рассмотрим, например, явную разностную схему Эйлера для решения задачи (7):

$$\frac{y_{n+1}-y_n}{\tau}+\lambda y_n=0,\;\lambda>0,$$

В этом случае схема называется условно устойчивой, а само неравенство $\tau\leqslant\frac{2}{\lambda}$ называется условием устойчивости.

Пример 4. Приведем пример абсолютно устойчивой разностной схемы. Для уравнения

$$u'(t)=f(t,u(t)), \quad (8)$$

$y_{n+1}=qy_n$, $q=(1+\tau\lambda)^{-1}$, причем $|q|<1$ при любых $\tau>0$.

Исследуем на устойчивость двухшаговый разностный метод, построенный в примере 3

$$\frac{y_{n+1}-y_n}{\tau}=\frac{3}{2}f_n-\frac{1}{2}f_{n-1}.$$

$\frac{y_{n+1}-y_n}{\tau}+\lambda(\frac{3}{2}f_n-\frac{1}{2}f_{n-1})=0$, $n=1,2,\ldots$, представляет собой разностное уравнение второго порядка с постоянными коэффициентами

$$y_{n+1}+ay_n+by_{n-1}=0, \quad (10)$$

Лемма 1. Оба корня уравнения (10) с действительными коэффициентами a , b лежат внутри или на границе единичного круга $|q|\leqslant 1$ тогда и только тогда, когда выполнены условия

$$1+a+b\geqslant 0,\;1-a+b\geqslant 0,\;b\leqslant 1. \quad (11)$$

26. Жесткие системы дифференциальных ур-ей

стр129-135 (6 страниц, надо катать с телефона)

Определение. Система линейных обыкновенных дифференциальных уравнений вида

$$\begin{cases} \frac{d\mathbf{u}(t)}{dt}=\mathbf{A}\mathbf{u}(t), \quad t>0 \\ \mathbf{u}(0)=u_0, \end{cases}$$

где $\mathbf{u}(t)=(u_1(t),u_2(t),\ldots,u_m(t))^T$, и $\mathbf{A}$ $(m\times m)$ – заданная матрица постоянных, вообще говоря, комплексных коэффициентов, называется жесткой, если:

- Действительные части всех собственных значений λ_k , $k=\overline{1,m}$ матрицы $\mathbf{A}$ отрицательные.
- Выполняется неравенство $\frac{\max\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k|}{\min\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k|}\gg 1$.

Понятие жесткости можно обобщить и на случай нелинейных систем. Рассмотрим систему нелинейных обыкновенных дифференциальных уравнений

$$\begin{cases} \frac{d\mathbf{u}(t)}{dt}=\mathbf{f}(t,\mathbf{u}(t)), \quad t>0, \\ \mathbf{u}(0)=u_0, \end{cases} \quad (12)$$

где $\mathbf{u}(t)=(u_1(t),u_2(t),\ldots,u_m(t))^T$, $\mathbf{f}(t,\mathbf{u}(t))=(f_1(t,\mathbf{u}(t)),f_2(t,\mathbf{u}(t)),\ldots,f_m(t,\mathbf{u}(t)))^T$.

Пусть $\lambda_k(t)$, $k=\overline{1,m}$ –собственные значения матрицы $\mathbf{J}(t)=\frac{\partial \mathbf{f}(t,\mathbf{v}(t))}{\partial \mathbf{u}}$.

$$\text{Введем число жесткости }S(t)=\frac{\max\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k(t)|}{\min\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k(t)|}.$$

Определение. Система (12) называется жесткой на решении $\mathbf{v}(t)$ и на данном интервале $0< t < T$ если

- $\operatorname{Re}\lambda_k(t)<0$, $k=\overline{1,m}$.
- Число жесткости $S(t)$ велико на рассматриваемом интервале $0< t < T$: $\frac{\max\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k(t)|}{\min\limits_{1\leqslant k\leqslant m}|\operatorname{Re}\lambda_k(t)|}\gg 1$.

Заметим, что первое требование означает асимптотическую устойчивость по Ляпунову решения $\mathbf{v}(t)$.

Дальнейшие определения устойчивости

$$\frac{d\mathbf{u}(t)}{dt}=\lambda\mathbf{u}(t), \quad (13)$$

Определение. Разностный метод называется $A(\alpha)$ -устойчивым, если область его устойчивости содержит угол левой полуплоскости: $|\arg(-\mu)|<\alpha$, $\mu=\tau\lambda$, $\alpha>0$.

В частности, $A\left(\frac{\pi}{2}\right)$ -устойчивость совпадает с A-устойчивостью.

Известно, что ни для какого α не существует явного $A(\alpha)$ -устойчивого линейного многошагового метода. Построены $A(\alpha)$ -устойчивые неявные методы третьего и четвертого порядка аппроксимации. К ним относятся чисто неявные многошаговые разностные схемы, у которых правая часть $\mathbf{f}(t,\mathbf{u})$ вычисляется только при новом значении $t=t_{n+m}$, а производная $\mathbf{u}'(t)$ аппроксимируется по нескольким предыдущим точкам и точке $t=t_{n+m}$.

27. Примеры разностных схем для численного интегрирования жестких систем ОДУ

стр132-135

При исследовании разностных схем для жестких систем уравнений обычно рассматривают модельное уравнение

$$\frac{du(t)}{dt} = \lambda u(t), \quad (14)$$

где λ — произвольное комплексное число. Свойства различных разностных схем изучают и сравнивают на примере этого уравнения. Для того, чтобы уравнение (14) действительно моделировало в некотором смысле исходную систему

$$\frac{du(t)}{dt} = f(t, u(t)),$$

необходимо рассматривать его при значениях λ , являющихся собственными значениями матрицы

$$J = \frac{\partial f(t, v(t))}{\partial u}.$$

Кроме обычного определения устойчивости (все корни характеристического уравнения не превосходят по модулю единицу) при рассмотрении жестких систем используют и другие, более узкие определения устойчивости. Мы рассмотрим два таких определения.

Определение. Областью устойчивости разностного метода называется множество точек комплексной плоскости $\mu = \tau\lambda$, для которых данный метод, примененный к уравнению (14), устойчив.

Рассмотрим, например, явную схему Эйлера:

$$\frac{y_{n+1} - y_n}{\tau} = f(t_n, y_n).$$

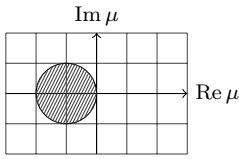
В применении к уравнению (14) эта схема примет вид

$$y_{n+1} = (1 + \mu) y_n, \quad \mu = \tau\lambda.$$

Условие устойчивости $|1 + \mu| \leq 1$ для комплексного числа $\mu = \mu_0 + i\mu_1$ означает, что

$$(\mu_0 + 1)^2 + \mu_1^2 \leq 1.$$

Таким образом, область устойчивости данного метода представляет собой круг единичного радиуса с центром в точке $(-1, 0)$.



Рассмотрим теперь неявную схему Эйлера

$$\frac{y_{n+1} - y_n}{\tau} = f(t_{n+1}, y_{n+1}).$$

В применении к уравнению (14) эта схема примет вид

$$\frac{y_{n+1} - y_n}{\tau} = \lambda y_{n+1}.$$

Перепишем это уравнение в виде

$$y_{n+1} = \frac{1}{1 - \mu} y_n.$$

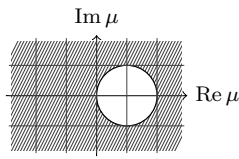
Область устойчивости метода определяется условием

$$\left| \frac{1}{1 - \mu} \right| \leq 1,$$

которое эквивалентно неравенству

$$|1 - \mu| \geq 1$$

и представляет собой внешность круга единичного радиуса с центром в точке $(1, 0)$.



Определение. Разностный метод называется *A-устойчивым*, если область его устойчивости содержит полуплоскость, задаваемую условием

$$\operatorname{Re} \mu < 0.$$

Отметим, что уравнение (14) асимптотически устойчиво при $\operatorname{Re} \lambda < 0$. Поэтому всякий *A-устойчивый* метод является абсолютно устойчивым (устойчивым при любом $\tau > 0$), если устойчиво решение исходного дифференциального уравнения. Нетрудно видеть, что неявный метод Эйлера является *A-устойчивым*, а явный метод Эйлера не является *A-устойчивым*.

Рассмотрим схему второго порядка аппроксимации:

$$\frac{y_{n+1} - y_n}{\tau} = \frac{f(t_{n+1}, y_{n+1}) + f(t_n, y_n)}{2}. \quad (15)$$

В применении к уравнению (14) эта схема примет вид

$$\frac{y_{n+1} - y_n}{\tau} = \frac{\lambda}{2} (y_{n+1} + y_n).$$

Отсюда находим

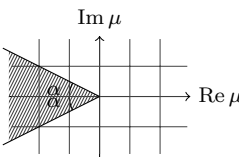
$$y_{n+1} = q y_n,$$

где $q = \frac{1 + 0.5\mu}{1 - 0.5\mu}$. Неравенство $|q| \leq 1$ выполнено при $\operatorname{Re} \mu \leq 0$. Следовательно метод (15) является *A-устойчивым*.

При решении жестких систем уравнений было бы желательно пользоваться именно *A-устойчивыми* разностными методами, так как условия их устойчивости не накладывают ограничений на шаг τ . Однако класс *A-устойчивых* методов весьма узок. Известно, что не существует явных линейных многошаговых *A-устойчивых* методов. Среди неявных линейных многошаговых методов нет *A-устойчивых* методов, имеющих порядок аппроксимации выше второго. Таким образом, схема (15) является одной из лучших *A-устойчивых* схем. В связи с тем, что класс *A-устойчивых* разностных схем весьма узок, было введено несколько определений устойчивости, являющихся менее ограничительными, чем определение *A-устойчивости*.

Определение. Разностный метод называется *A(α)-устойчивым*, если область его устойчивости содержит угол левой полуплоскости:

$$|\arg(-\mu)| < \alpha, \quad \mu = \tau\lambda, \alpha > 0.$$



В частности, $A\left(\frac{\pi}{2}\right)$ -устойчивость совпадает с *A-устойчивостью*.

Известно, что ни для какого α не существует явного $A(\alpha)$ -устойчивого линейного многошагового метода. Построены $A(\alpha)$ -устойчивые неявные методы третьего и четвертого порядка аппроксимации. К ним относятся чисто неявные многошаговые разностные схемы, у которых правая часть $f(t, u)$ вычисляется только при новом значении $t = t_{n+m}$, а производная $u'(t)$ аппроксимируется по нескольким предыдущим точкам и точке $t = t_{n+m}$. Например, схема

$$\frac{25y_{n+4} - 48y_{n+3} + 36y_{n+2} - 16y_{n+1} + 3y_n}{12\tau} = f(t_{n+4}, y_{n+4})$$

имеет четвертый порядок аппроксимации и $A(\alpha)$ -устойчива при некотором $\alpha > 0$.

25.2 Многошаговые разностные методы. Погрешность аппроксимации. Понятие устойчивости

Общий *m*-шаговый линейный разностный метод

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^m b_k f_{n-k}, \quad (16)$$

где $\tau > 0$, $y_0, y_1, \dots, y_{m-1}$ — заданы. Будем считать, что коэффициенты $a_k, b_k, k = \overline{1, m}$ не зависят от τ .

Пример. В применении к уравнению (7) метод (16) принимает вид:

$$\sum_{k=0}^m (a_k + \tau\lambda b_k) y_{n-k} = 0. \quad (17)$$

Решение этого разностного уравнения с постоянными коэффициентами будем искать в виде $y_j = q^j, j \in \mathbb{Z}_+$.

$$F_m(q, \tau) = \sum_{k=0}^m (a_k + \tau\lambda b_k) q^{m-k} = 0. \quad (18)$$

Определение. Уравнение вида (18) называется характеристическим уравнением разностной схемы (17).

$F_m(q, 0) = 0$,

$$\sum_{k=0}^m a_k q^{m-k} = 0, \quad (19)$$

Определение. Говорят, что схема (16) удовлетворяет условию (α), если все корни характеристического уравнения (19) лежат внутри или на границе единичного круга комплексной плоскости, причем на границе единичного круга нет кратных корней.

Таким образом, выполнение условия (α) соответствует устойчивости разностного метода для уравнения $u'(t) = 0$. Однако часто схему и для общего уравнения (8) называют устойчивой, если она удовлетворяет условию (α). Такая терминологическая неточность оправдана тем, что из условия (α) следует сходимость решения разностной задачи (16) к решению исходной дифференциальной задачи (8).

Теорема. Пусть разностная схема удовлетворяет условию (α) и $|f_n'| \leq L$ на отрезке $0 \leq t \leq T$. Тогда при $0 \leq t_n = n\tau \leq T$ и всех достаточно малых τ выполняется оценка

$$|y_n - u(t_n)| \leq M \left(\sum_{j=m}^n \tau |\psi_j| + \max_{0 \leq i \leq m-1} |y_i - u(t_i)| \right),$$

где $|y_i - u(t_i)|$ — погрешности в задании начальных данных, $i = 0, (m-1)$, M — константа, зависящая от L, T и не зависящая от τ, ψ_j — погрешность аппроксимации на решении исходного уравнения (8):

$$\psi_j = - \sum_{k=0}^m \frac{a_k}{\tau} u(t_{n-k}) + \sum_{k=0}^m b_k f_{n-k}.$$

(без доказательства)

Замечание 4. Методы Адамса $\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^m b_k f_{n-k}$ всегда удовлетворяют условию (α), так как для них $a_0 = -a_1 = 1$, то есть $q = q_1 = 1$, что следует из уравнения $q^n - q^{n-1} = 0$.

Замечание 5. При указанном подходе, в отличие от рассмотренных примеров, не различаются абсолютно устойчивые и условно устойчивые разностные схемы, так как параметр τ заранее считается достаточно малым.

Замечание 6. Мы уже упоминали в ?? данной главы, что наивысший достижимый порядок аппроксимации неявных *m*-шаговых методов равен $2m$, а явных — $(2m - 1)$. Однако оказывается, что методы наивысшего порядка неустойчивы в том смысле, что они не удовлетворяют условию (α). А именно, если *m* нечетно, то никакой устойчивый метод не превосходит порядка $p = m + 1$. Если *m* четно, то никакой устойчивый метод не превосходит порядка $p = m + 2$ (*p* — порядок аппроксимации). Для явных схем наивысший порядок аппроксимации устойчивых методов $p = m$.

Пример. Нетрудно привести пример схем, не удовлетворяющих условию (α). Так, явная двухшаговая схема

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2f_{n-1} + f_{n-2}}{3}$$

имеет третий порядок погрешности аппроксимации $\psi = O(\tau^3)$

24.2 Оценка точности на примере двухэтапного метода Рунге–Кутты.

Выпишем еще раз разностную схему, описывающую общий двухэтапный метод Рунге–Кутты:

$$\begin{cases} \frac{y_{n+1} - y_n}{\tau} = (1 - \sigma)f(t_n, y_n) + \sigma f(t_n + a\tau, y_n + a\tau f(t_n, y_n)), & n \in \mathbb{Z}_+ \\ y_0 = u_0. \end{cases} \quad (20)$$

Введем погрешность разностной схемы (20):

$$z_n = y_n - u_n, n \in \mathbb{Z}.$$

Подставим выражение для погрешности в разностную схему (20) и получим задачу для нахождения функции z_n :

$$\begin{cases} \frac{z_{n+1} - z_n}{\tau} = (1 - \sigma)f_n + \sigma f(t_n + a\tau, y_n + a\tau f_n) - \frac{u_{n+1} - u_n}{\tau}, & n \in \mathbb{Z}_+ \\ z_0 = 0, \end{cases} \quad (21)$$

где $f_n = f(t_n, y_n), y_n = z_n + u_n$.

Для доказательства сходимости решения разностной схемы (20) к решению исходной задачи Коши (124) достаточно показать, что

$$\lim_{n \rightarrow \infty} |z_n| = 0.$$

Покажем, что $|z_n| \leq M|\psi_n|, n \in \mathbb{Z}_+$, где константа M не зависит от шага τ, ψ_n — погрешность аппроксимации на решении исходной задачи (124):

$$\psi_n = - \frac{u_{n+1} - u_n}{\tau} + (1 - \sigma)f(t_n, u_n) + \sigma f(t_n + a\tau, u_n + a\tau f(t_n, u_n)).$$

Перепишем задачу (21) в эквивалентном виде, сформировав погрешность аппроксимации путем добавления недостающих слагаемых:

$$\begin{aligned} \frac{z_{n+1} - z_n}{\tau} &= - \frac{u_{n+1} - u_n}{\tau} + (1 - \sigma)f(t_n, u_n) + \sigma f(t_n + a\tau, u_n + a\sigma f(t_n, u_n)) + \\ &\quad + (1 - \sigma)(f(t_n, y_n) - f(t_n, u_n)) + \\ &\quad + \sigma(f(t_n + a\tau, y_n + a\tau f(t_n, y_n)) - f(t_n + a\tau, u_n + a\tau f(t_n, u_n))) = \\ &= \psi_n + \varphi_n^{(1)} + \varphi_n^{(2)}, \end{aligned} \quad (22)$$

где

$$\varphi_n^{(1)} = (1 - \sigma)(f(t_n, y_n) - f(t_n, u_n)),$$

$$\varphi_n^{(2)} = \sigma(f(t_n + a\tau, y_n + a\tau f(t_n, y_n)) - f(t_n + a\tau, u_n + a\tau f(t_n, u_n))).$$

Пусть функция $f(t, u)$ удовлетворяет условию Липшица по второму аргументу с константой $L > 0$:

$$|f(t, v) - f(t, u)| \leq L|u - v|, \quad (t, u), (t, v) \in G.$$

Замечание. Требование липшицевости функции $f(t, u)$ естественно, так как является условием того, что решение исходной задачи (124) существует и единственно.

Как правило, на практике выбирают $0 \leq \sigma \leq 1, a \geq 0$. Воспользуемся этими условиями и оценим выражения $\varphi_n^{(1)}$ и $\varphi_n^{(2)}$:

$$|\varphi_n^{(1)}| = (1 - \sigma)|f(t_n, y_n) - f(t_n, u_n)| \leq (1 - \sigma)L|y_n - u_n| = (1 - \sigma)L|z_n|,$$

$$\begin{aligned} |\varphi_n^{(2)}| &\leq \sigma L|y_n + a\tau f(t_n, y_n) - u_n - a\tau f(t_n, u_n)| \leq \\ &\leq \sigma L(|y_n - u_n| + a\tau|f(t_n, y_n) - f(t_n, u_n)|) \leq \sigma L(|z_n| + a\tau L|z_n|) = \sigma L(1 + a\tau L)|z_n|. \end{aligned}$$

Пусть $\sigma a \leq 0.5$. Оценим сумму $|\varphi_n^{(1)}| + |\varphi_n^{(2)}|$:

$$|\varphi_n^{(1)}| + |\varphi_n^{(2)}| \leq (1 - \sigma)L|z_n| + \sigma L(1 + a\tau L)|z_n| = L|z_n| + \sigma a \tau L^2|z_n| \leq L(1 + 0.5\tau L)|z_n|.$$

Приступим к получению оценки точности. Запишем равенство (22) в виде $z_{n+1} = z_n + \tau\psi_n + \tau\varphi_n^{(1)} + \tau\varphi_n^{(2)}$. Далее, очевидно, справедлива оценка:

$$|z_{n+1}| \leq |z_n| + \tau|\psi_n| + \tau(|\varphi_n^{(1)}| + |\varphi_n^{(2)}|) \leq (1 + \tau L + 0.5\tau^2 L^2)|z_n| + \tau|\psi_n|.$$

Заметим, что слагаемые в сумме $(1 + \tau L + 0.5\tau^2 L^2)$ являются первыми членами разложения функции $e^{\tau L}$ по формуле Тейлора по переменной τ в окрестности нуля. Следовательно,

$$(1 + \tau L + 0.5\tau^2 L^2) \leq e^{\tau L}.$$

Тогда

$$|z_{n+1}| \leq e^{\tau L}|z_n| + \tau|\psi_n|.$$

Введем обозначение $\rho = e^{\tau L}$. Тогда

$$|z_{n+1}| \leq \rho|z_n| + \tau|\psi_n|, \quad n \in \mathbb{Z}_+.$$

Раскроем полученное рекуррентное соотношение:

$$|z_{n+1}| \leq \rho^{n+1}|z_0| + \tau \sum_{j=0}^n \rho^{n-j} |\psi_j|.$$

Так как $z_0 = 0$, то получаем:

$$|z_{n+1}| \leq \max_{0 \leq j \leq n} |\psi_j| t_n e^{t_n L}.$$

Учтем, что $t_n \leq T$, тогда:

$$|z_{n+1}| \leq M \max_{0 \leq j \leq n} |\psi_j|,$$

где константа $M = Te^{TL} > 0$ не зависит от τ . Заметим, что

$$\lim_{\tau \rightarrow 0} |z_{n+1}| = 0,$$

так как $|\psi_j| \leq M_1(\tau^2)$ по доказанному выше. Тогда при достаточно малых τ получаем:

$$|z_{n+1}| = O(\tau^2).$$

Это означает, что рассматриваемый общий двухэтапный метод Рунге–Кутты при выполнении соответствующих условий имеет квадратичную точность по τ , совпадающую с оценкой погрешности аппроксимации на решении исходного уравнения (124).