III семестр. ММФ НГУ
Вычислительные методы линейной алгебры

КОНСПЕКТ ЛЕКЦИЙ[1]

Мацокин А.М. – проф. кафедры вычислительной математики
2003 – 2004 учебный год

Предлагаем Вашему вниманию конспект лекций семестрового курса «Вычислительные методы линейной алгебры», прочитанных профессором кафедры А.М. Мацокиным для студентов второго курса механико–математического факультета Новосибирского государственного университета.
Мы надеемся, что этот конспект будет полезен студентам ММФ НГУ для более полного усвоения курса и применения методов вычислительной математики в их дальнейшей учебной, научно–преподавательской и практической деятельности.
Ограничений на использование и распространение конспекта – нет.
Любым замечаниям автор конспекта будет только рад и принимает их по адресу
E-mail: matsokin@oapmg.sscc.ru
или
E-mail: mmf@nsu.ru

Содержание

[1] Конспект подготовлен при финансовой поддержке проекта № 274 ФЦП "Интеграция".

Лекция 1.

Традиционные задачи линейной алгебры

$$A = (a_{ij})_{i,j=1}^{n} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \text{ – матрица}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ – вектор}$$

Задачи:	**Методы** (теория определителей)**:**
решение системы уравнений $Ax = b$	метод Крамера: $x_i = \dfrac{\Delta_i}{\Delta}, \quad \Delta_i = \det\{A_i, b\}$
вычисление обратной матрицы: $AX = XA = E$	определение столбца $x^{(j)}$ матрицы X: $Ax^{(j)} = e_j, \quad x_{ij} = \dfrac{\Delta_{ji}}{\Delta}$
вычисление определителя $\Delta = \det A$	по определению $\Delta = \sum\limits_{(i_1,\ldots,i_n)} (-1)^r \cdot a_{1i_1} \cdot \ldots \cdot a_{ni_n}$
спектральная задача: $Ax = \lambda \cdot x$	собственные значения – корни полинома $P_n(\lambda) = \det(A - \lambda \cdot E)$ собственные векторы – решения систем $(A - \lambda \cdot E)x = 0$ r линейно независимых решений, где $r = \dim\{Ker(A - \lambda \cdot E)\}$

Непригодность этих методов:

количество умножений при вычислении одного определителя: $(n-1) \cdot n!$ если производительность ЭВМ 10^9 оп/сек, то	ошибки округления: $\tilde{a} = a + \varepsilon \cdot \lvert a \rvert, \quad \varepsilon \le 10^{-6}$ если $n = 6$, $\lvert a_{ij} \rvert \ge 10$, $\Delta = 1$, $a_i = a_{1i_1} \cdot \ldots \cdot a_{ni_n}$, то $a_i - \tilde{a}_i = O(1)$ $\Delta - \tilde{\Delta} = n! \cdot O(1) = O(1)$

n	время вычисления
10	$: 10^{-4}$ сек.
20	> 17 мин.
30	> 400 тыс. лет

т.е. определитель вычисляется с большой ошибкой и, следовательно, решения поставленных задач вычисляются с такой же ошибкой.

Векторные и матричные нормы

Векторные

$\forall\, x \in R^n\ (C^n)$

$\|x\| > 0, \quad x \neq 0$
$\|x\| = 0, \quad x = 0$

$\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$

$\|x + y\| \le \|x\| + \|y\|$

Примеры:

$\|x\|_\infty = \max |x_i|$ – кубическая или равномерная

$\|x\|_1 = |x_1| + \ldots + |x_n|$ – октаэдрическая

$\|x\|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$ –сферическая или евклидова

Матричные

$\forall\, A \in R^{n \times n}\ (C^{n \times n})$

аксиомы 1. – 3. – аддитивная

4. $\|AB\| \le \|A\| \cdot \|B\|$ – мультипликативная

согласованная с векторной, если $\|Ax\| \le \|A\| \cdot \|x\|$

подчиненная векторной, если $\|A\| = \sup \dfrac{\|Ax\|}{\|x\|}$

Примеры подчиненных матричных норм:

$$\|A\|_\infty = \sup \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_i \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_1 = \sup \frac{\|Ax\|_1}{\|x\|_1} = \max_j \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_2 = \sup \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\rho(A^* A)}$$

Теорема.	Любые две нормы $\|x\|_*$ и $\|x\|_{**}$ в конечномерном пространстве эквивалентны: $\exists\, \alpha, \beta:\ \forall\, x \quad \alpha \cdot \|x\|_* \le \|x\|_{**} \le \beta \cdot \|x\|_*$	Примеры: $\forall\, x \in R^n\ (C^n)$ $\|x\|_\infty \le \|x\|_1 \le n \cdot \|x\|_\infty$ $\|x\|_\infty \le \|x\|_2 \le \sqrt{n} \cdot \|x\|_\infty$ $\|x\|_2 \le \|x\|_1 \le \sqrt{n} \cdot \|x\|_2$

!!! Константы эквивалентности зависят от размерности пространства **!!!**

При решении системы линейных уравнений $Ax = b$ могут быть неточно заданы либо правая часть $\tilde{b} = b + \delta b$ либо матрица $\tilde{A} = A + \delta A$, где компоненты вектора δb и элементы матрицы δA малы по сравнению с соответствующими элементами исходных вектора и матрицы. Тогда вместо решения x мы получим его

приближение $\tilde{x} = x + \delta x$, причем компоненты вектора–ошибки δx могут быть большими.
Оценим норму ошибки через нормы возмущений правой части и матрицы системы, считая, что матричная норма подчинена векторной норме.

Число обусловленности

Определение. $\mathrm{cond}A = \|A\| \cdot \|A^{-1}\|$

1. $\mathrm{cond}A \ge 1$ т.к. $\|x\| = \|AA^{-1}x\| \le \|A\| \cdot \|A^{-1}\| \cdot \|x\|$	2. $\mathrm{cond}(AB) \le \mathrm{cond}A \cdot \mathrm{cond}B$ т.к. $\|AB\| \cdot \|(AB)^{-1}\| \le \|A\| \cdot \|B\| \cdot \|B^{-1}\| \cdot \|A^{-1}\|$

Теорема.	$Ax = b,\ \det A \ne 0$; $A(x+\delta x) = b + \delta b$ $\Rightarrow$ $\dfrac{\|\delta x\|}{\|x\|} \le \mathrm{cond}A \cdot \dfrac{\|\delta b\|}{\|b\|}$
Док-во.	$\|\delta x\| = \|A^{-1}\delta b\| \le \|A^{-1}\| \cdot \|\delta b\|$; $\|A\| \cdot \|x\| \ge \|Ax\| = \|b\|$ $\Rightarrow$ $\dfrac{\|\delta x\|}{\|x\|} \le \mathrm{cond}A \cdot \dfrac{\|\delta b\|}{\|b\|}$.

Теорема.	$Ax = b,\ \det A \ne 0,\quad (A+\delta A)(x+\delta x) = b + \delta b,\quad \|A^{-1}\| \cdot \|\delta A\| < 1$ $\Rightarrow \dfrac{\|\delta x\|}{\|x\|} \le \dfrac{\mathrm{cond}A}{1 - \mathrm{cond}A \cdot \dfrac{\|\delta A\|}{\|A\|}} \cdot \left(\dfrac{\|\delta b\|}{\|b\|} + \dfrac{\|\delta A\|}{\|A\|}\right)$
Док-во.	1. $\exists\ (A+\delta A)^{-1}$ и $\|(A+\delta A)^{-1}\| \le \dfrac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\delta A\|}$, т.к. $(A+\delta A) = A(E + A^{-1} \cdot \delta A)$ $\|(E + A^{-1}\delta A)z\| \ge \|z\| - \|A^{-1} \cdot \delta A \cdot z\| \ge \|z\| - \|A^{-1}\| \cdot \|\delta A\| \cdot \|z\| > 0$ $\exists\ (E + A^{-1}\delta A)^{-1} = E - (A^{-1}\delta A) + (A^{-1}\delta A)^2 - \ldots$ $\|(E + A^{-1}\delta A)^{-1}\| = 1 + \|A^{-1}\delta A\| + \|(A^{-1}\delta A)^2\| + \ldots$ $\dfrac{1}{1 - \|A^{-1}\delta A\|} \quad \dfrac{1}{1 - \|A^{-1}\|\,\|\delta A\|}$ 2. Т.к. $\delta x = (A+\delta A)^{-1}[b + \delta b - Ax - \delta A \cdot x] = (A+\delta A)^{-1}[\delta b - \delta A \cdot x]$, то

$$\|\delta x\| \le \|(A+\delta A)^{-1}\| \cdot (\|\delta b\| + \|\delta A\| \cdot \|x\|)$$

$$\Rightarrow \quad \frac{\|\delta x\|}{\|x\|} \le \|(A+\delta A)^{-1}\| \cdot (\frac{\|\delta b\|}{\|x\|} + \|\delta A\|) \le$$

$$\le \|(A+\delta A)^{-1}\| \cdot (\frac{\|\delta b\|}{\|b\| / \|A\|} + \|\delta A\| \cdot \frac{\|A\|}{\|A\|}) \le$$

$$\le \|(A+\delta A)^{-1}\| \cdot \|A\| \cdot (\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|}) \le$$

$$\le \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|\delta A\|} \cdot (\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|}) =$$

$$= \frac{\mathrm{cond} A}{1 - \mathrm{cond} A \cdot \frac{\|\delta A\|}{\|A\|}} \cdot (\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|})$$

Лекция 2. Прямые методы решения линейных уравнений

Метод исключения Гаусса – схема единственного деления

$$Ax=b,\quad \det A_k \equiv \det\begin{bmatrix} a_{11} & \ldots & a_{1k} \\ \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} \end{bmatrix} \neq 0,\ k=1,\ldots,n \quad \Rightarrow \quad \begin{array}{c} A=LU \\ Ly=b \\ Ux=y \end{array}$$

Схема единственного деления на примере системы третьего порядка:

Прямой ход:	Матричная формулировка:
$\begin{array}{rcrcrcl} a_{11}x_1 &+& a_{12}x_2 &+& a_{13}x_3 &=& b_1 \\ a_{21}x_1 &+& a_{22}x_2 &+& a_{23}x_3 &=& b_2 \\ a_{31}x_1 &+& a_{32}x_2 &+& a_{33}x_3 &=& b_3 \end{array}$	$Ax=b,$ $\det A_1 = a_{11} \neq 0, \quad L_1 = \begin{bmatrix} 1/a_{11} & 0 & 0 \\ -a_{21}/a_{11} & 1 & 0 \\ -a_{31}/a_{11} & 0 & 1 \end{bmatrix}$
$\begin{array}{rcrcrcl} x_1 &+& u_{12}x_2 &+& u_{13}x_3 &=& y_1 \\ && a_{22}^{(1)}x_2 &+& a_{23}^{(1)}x_3 &=& b_2^{(1)} \\ && a_{32}^{(1)}x_2 &+& a_{33}^{(1)}x_3 &=& b_3^{(1)} \end{array}$	$A^{(1)}x=b^{(1)}, \quad A^{(1)}=L_1A, \quad b^{(1)}=L_1b$ $\det A_2^{(1)} \neq 0, \quad L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/a_{22}^{(1)} & 0 \\ 0 & -a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{bmatrix}$
$\begin{array}{rcrcrcl} x_1 &+& u_{12}x_2 &+& u_{13}x_3 &=& y_1 \\ && x_2 &+& u_{23}x_3 &=& y_2 \\ &&&& a_{33}^{(2)}x_3 &=& b_3^{(2)} \end{array}$	$A^{(2)}x=b^{(2)}, \quad A^{(2)}=L_2A^{(1)}, \quad b^{(2)}=L_2b^{(1)}$ $\det A_3^{(2)} \neq 0, \quad L_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/a_{33}^{(2)} \end{bmatrix}$
$\begin{array}{rcrcrcl} x_1 &+& u_{12}x_2 &+& u_{13}x_3 &=& y_1 \\ && x_2 &+& u_{23}x_3 &=& y_2 \\ &&&& x_3 &=& y_3 \end{array}$	$Ux \equiv A^{(3)}x = b^{(3)} \equiv y, \quad A^{(3)}=L_3A^{(2)}, \quad b^{(3)}=L_3b^{(2)}$ $U=\begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}, \quad A=(L_1^{-1}\cdot L_2^{-1}\cdot L_3^{-1})\cdot U = L\cdot U$
Обратный ход: $x_3=y_3, \quad x_2=y_2-u_{23}x_3, \quad x_3=y_1-u_{12}x_2-u_{13}x_3$	Матричная формулировка: $x=U^{-1}y$

Формулы схемы единственного деления (доказать):

k–ый шаг прямого хода:

$$A^{(k-1)}x = b^{(k-1)} \;\Rightarrow\; A^{(k)}x = b^{(k)},$$

$$A^{(k)} = L_k A^{(k-1)}, \quad b^{(k)} = L_k b^{(k-1)}$$

$$L_k = \begin{bmatrix} 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1/a_{kk}^{(k-1)} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -a_{k+1k}^{(k-1)}/a_{kk}^{(k-1)} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -a_{nk}^{(k-1)}/a_{kk}^{(k-1)} & 0 & \cdots & 1 \end{bmatrix}$$

$A^{(n)} = U$ – верхняя треуг. матрица

Теорема об LU разложении

Если $\forall\, k\ \det A_k \neq 0$, то $\exists\, L\ \exists\, U:\ A = LU$, где L – нижняя, U – верхняя треугольные матрицы.

Доказательство.

Если $A = LU$, то $A_k = L_k U_k$, $\det A_k = \det L_k \cdot \det U_k = l_{11} \cdot \ldots \cdot l_{kk} \cdot u_{11} \cdot \ldots \cdot u_{kk} \neq 0$,

т.к. $\begin{bmatrix} A_k & B_{k,n-k} \\ C_{n-k,k} & A'_{n-k} \end{bmatrix} = \begin{bmatrix} L_k & O_{k,n-k} \\ L_{n-k,k} & L'_{n-k} \end{bmatrix} \cdot \begin{bmatrix} U_k & U_{k,n-k} \\ O_{n-k,k} & U'_{n-k} \end{bmatrix}$.

Предположим, что разложение $A_k = L_k U_k$ найдено ($A_1 \equiv a_{11} = L_1 U_1 \equiv l_{11} \cdot u_{11} \neq 0$). Вычислим $A_{k+1} = L_{k+1} U_{k+1}$

(т.е. последние строку матрицы L_{k+1} и столбец матрицы U_{k+1}):

т.к. $$\left[\begin{array}{ccc|c} & & & a_{1,k+1} \\ & A_k & & \vdots \\ & & & a_{k,k+1} \\ \hline a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} \end{array}\right] = \left[\begin{array}{ccc|c} & & & 0 \\ & L_k & & \vdots \\ & & & 0 \\ \hline l_{k+1,1} & \ldots & l_{k+1,k} & l_{k+1,k+1} \end{array}\right] \cdot \left[\begin{array}{ccc|c} & & & u_{1,k+1} \\ & U_k & & \vdots \\ & & & u_{k,k+1} \\ \hline 0 & \ldots & 0 & u_{k+1,k+1} \end{array}\right]$$

то $L_k \begin{bmatrix} u_{1,k+1} \\ \vdots \\ u_{k,k+1} \end{bmatrix} = \begin{bmatrix} a_{1,k+1} \\ \vdots \\ a_{k,k+1} \end{bmatrix}$, $\quad \begin{bmatrix} l_{k+1,1} & \ldots & l_{k+1,k} \end{bmatrix} U_k = \begin{bmatrix} a_{k+1,1} & \ldots & a_{k+1,k} \end{bmatrix}$ – системы с треугольными неособенными матрицами (решения $\exists!$), и

$$l_{k+1,k+1} \cdot u_{k+1,k+1} = a_{k+1,k+1} - \begin{bmatrix} l_{k+1,1} & \ldots & l_{k+1,k} \end{bmatrix} \cdot \begin{bmatrix} u_{1,k+1} \\ \vdots \\ u_{k,k+1} \end{bmatrix},$$

очевидно, что решение этого уравнения существует, но не единственно.

(так как $0 \neq \det A_{k+1} = \det L_{k+1} \cdot \det U_{k+1}$, то $\det L_{k+1} \neq 0,\ \det U_{k+1} \neq 0$.)

И, наконец, $A \equiv A_n = L_n U_n \equiv LU$.

Объем вычислений.

Так как для решения системы уравнений с треугольной матрицей порядка k достаточно выполнить $k(k+1)/2$ умножений и делений, то полагая на каждом шаге $u_{k+1,k+1} = 1$, получим, что число таких операций для вычисления последних строки и столбца матриц L_{k+1} и U_{k+1} равно $k(k+2)$, а для вычисления матриц L и U достаточно $\sum_1^{n-1} k(k+2) \approx n^3/3$ умножений или делений.

Замечание.

Если построено LU–разложение матрицы A, то ее определитель вычисляется за $2(n-1)$ умножений (перемножаются диагональные (ведущие) элементы).

Теорема (об LDU – разложении). Если $\det A_k \neq 0\ \forall k$, то разложение $A = LDU$, где $l_{kk} = u_{kk} = 1\ \forall k$, единственно.	**Док–во.** Пусть $A = L^{(1)}D^{(1)}U^{(1)} = L^{(2)}D^{(2)}U^{(2)}$, тогда $[L^{(2)}]^{-1}L^{(1)} = D^{(2)}U^{(2)}[D^{(1)}U^{(1)}]^{-1} = \mathrm{diag} = E$, (т.к. $[L^{(2)}]^{-1}L^{(1)}$ – нижняя треуг. м–ца с единицами на диагонали) $\Rightarrow L^{(1)} = L^{(2)}$ $\Rightarrow [D^{(2)}]^{-1}D^{(1)} = U^{(2)}[U^{(1)}]^{-1} = \mathrm{diag} = E$ $\Rightarrow U^{(1)} = U^{(2)}\ \&\ D^{(1)} = D^{(2)}$.

Разложение Холесского

Теорема.	Если $A = A^* > 0$ (т.е. $(Ax,x) > 0\quad \forall x \neq 0$), то $A = L \cdot D \cdot L^*$, $l_{kk} = 1$, $d_k > 0\ \forall k$.
Док–во.	Т.к. $0 < (Ax,x) = (A_k x^k, x^k)\quad \forall x = \begin{pmatrix} x^k \\ 0 \end{pmatrix} \neq 0$, то $\det A_k \neq 0\ \forall k$ $\Rightarrow A = LDU = A^* = U^* D^* L^* = LDL^*$. Т.к. $\exists\, y^{(k)} = [L^*]^{-1} e_k \neq 0\ \&\ A > 0$, то $(Ay^k, y^k) = (LDL^* y^k, y^k) = (DL^* y^k, L^* y^k) = d_k > 0\ \forall k$.

Метод квадратного корня

Теорема.	Если $A = A^* > 0$, то $A = BB^*$, где B – нижняя треугольная м–ца, и $\mathrm{cond}_2 B = \mathrm{cond}_2 B^* = \sqrt{\mathrm{cond}_2 A}$.
Док–во.	Из теоремы о разложении Холесского имеем $A = LDL^* = L(D^{1/2}D^{1/2})L^* = (LD^{1/2})(LD^{1/2})^*\ \Rightarrow\ B = LD^{1/2}$. Т.к. $\mathrm{Sp}(B^*B) = \mathrm{Sp}(BB^*) \equiv \mathrm{Sp}(A)$, то $\lVert B \rVert_2 = \lVert B^* \rVert_2 = \sqrt{\rho(A)} = \sqrt{\lVert A \rVert_2}$. Аналогично $\lVert B^{-1} \rVert_2 = \lVert (B^*)^{-1} \rVert_2 = \sqrt{\rho(A^{-1})} = \sqrt{\lVert A^{-1} \rVert_2}$. $\Rightarrow\ \mathrm{cond}_2 B = \mathrm{cond}_2 B^* = \sqrt{\mathrm{cond}_2 A}$.

Решение системы уравнений $Ax = b$ с помощью разложения $A = BB^*$ называется методом квадратного корня. Так как

$$\begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & \bar{b}_{21} & \bar{b}_{31} \\ 0 & b_{22} & \bar{b}_{32} \\ 0 & 0 & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & \bar{a}_{21} & \bar{a}_{31} \\ a_{21} & a_{22} & \bar{a}_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \Rightarrow \begin{matrix} b_{11}^2 = a_{11},\ b_{21}b_{11} = a_{21},\ b_{31}b_{11} = a_{31} \\ b_{22}^2 = a_{22} - b_{21}\bar{b}_{21},\ b_{32}b_{22} = a_{32} - b_{21}\bar{b}_{31} \\ b_{33}^2 = a_{33} - b_{31}\bar{b}_{31} - b_{32}\bar{b}_{32} \end{matrix}$$

то элементы матрицы B вычисляются по следующим формулам:

$$\left\{ b_{kk} = \sqrt{a_{kk} - \sum_{j=1}^{k-1} |b_{kj}|^2},\ \ b_{k+i,k} = (a_{k+i,k} - \sum_{j=1}^{k-1} b_{k+i,j}\bar{b}_{k,j})/b_{kk},\ i=1,\dots,n-k \right\}_{k=1}^{n}$$

Лекция 3.

Метод исключения с выбором главного элемента по столбцу

Напомним 1–ый шаг схемы единственного деления для решения $Ax = b$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \Rightarrow \quad \begin{bmatrix} 1 & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & aa_{nn}^{(1)} \end{bmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix},$$

где $a_{1j}^{(1)} = a_{1j} / a_{11}$, $b_1^{(1)} = b_1 / a_{11}$, $a_{ij}^{(1)} = a_{ij} - a_{i1} \cdot a_{1j} / a_{11}$ $(i, j = 2, \dots, n)$.

Эти операции выполнимы, если (главный элемент шага) $a_{11} \neq 0$.

Ошибки округления будут меньше, если $|a_{11}| \geq |a_{1j}|$ или $|a_{11}| \geq |a_{i1}|$.

Матрица перестановок

$P = (p_{ij})_{i,j=1}^{n}$, $\quad p_{ij} = \begin{cases} 1, & j = k_i \\ 0, & j \neq k_i \end{cases}$, где $(k_1, k_2, \dots, k_n)$ – перестановка $(1, 2, \dots, n)$.

Доказать, что $PP^* = P^*P = E$, т.е. P – ортогональная матрица.

Доказать, что $\mathrm{cond}_2(P) = 1$.

Элементарная матрица перестановок

P_{kl} – матрица перестановок k и l элементов в n–ке $(1, 2, \dots, n)$.

Доказать, что $P_{k,l} = P_{k,l}^* = P_{k,l}^{-1}$.

Доказать, что умножение на матрицу $P_{k,l}$ матрицы A слева ($P_{k,l}A$) – это перестановка k и l строк, справа ($AP_{k,l}$) – перестановка k и l столбцов матрицы A.

Выбор главного элемента по столбцу.

1–й шаг:

находим i_1: $|a_{i_1 1}| \geq \max\limits_{i=1,\dots,n} |a_{i1}|$ ($\neq 0$, если $\det A \neq 0$);

меняем местами 1 и i_1 строки: $A^{(1/2)} = P_{1,i_1} A$, $b^{(1/2)} = P_{1,i_1} b$;

обнуляем в 1–ом столбце элементы: $A^{(1)} = L_1 A^{(1/2)}$, $b^{(1)} = L_1 b^{(1/2)}$,

$$L_1 = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ -\dfrac{a_{21}^{(1/2)}}{a_{11}^{(1/2)}} & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\dfrac{a_{n1}^{(1/2)}}{a_{11}^{(1/2)}} & 0 & \ldots & 1 \end{bmatrix},$$

$$A^{(1)} = \left[\begin{array}{c|ccc} a_{11}^{(1/2)} & a_{12}^{(1/2)} & \ldots & a_{1n}^{(1/2)} \\ \hline 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{array}\right], \quad b^{(1)} = \left(\begin{array}{c} b_1^{(1/2)} \\ \hline b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{array}\right)$$

$$\det A \neq 0 \;\Rightarrow\; \det A^{(1)} \neq 0.$$

После k шагов имеем $A^{(k)}x = b^{(k)}$, где $\det A^{(k)} \neq 0$, если $\det A \neq 0$

$$A^{(k)} = \left[\begin{array}{c|ccc} U_k & & U_{k,n-k} & \\ \hline & a^{(k)}_{k+1,k+1} & \dots & a^{(k)}_{k+1,n} \\ 0 & \vdots & \ddots & \vdots \\ & a^{(k)}_{n,k+1} & \dots & a^{(k)}_{n,n} \end{array}\right], \quad b^{(k)} = \left[\begin{array}{c} y^{(k)} \\ \hline b^{(k)}_{k+1} \\ \vdots \\ b^{(k)}_n \end{array}\right]$$

$(k+1)$–й шаг:

находим i_{k+1}: $|a^{(k)}_{i_{k+1}k+1}| \geq \max\limits_{i=k+1,\dots,n} |a_{i\,k+1}|$;

меняем местами $k+1$ и i_{k+1} строки:

$$P_{k+1,i_{k+1}} = \begin{bmatrix} E_k & 0 \\ 0 & P^{(n-k)}_{1,i_{k+1}-k} \end{bmatrix}, \qquad A^{(k+1/2)} = P_{k+1,i_{k+1}} A^{(k)}, \quad b^{(k+1/2)} = P_{k+1,i_{k+1}} b^{(k)};$$

обнуляем в $(k+1)$–ом столбце элементы:

$$L_{k+1} = \begin{bmatrix} E_k & 0 \\ 0 & L^{(n-k)}_1 \end{bmatrix}, \qquad L^{(n-k)}_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -\dfrac{a^{(k+1/2)}_{k+2,k+1}}{a^{(k+1/2)}_{k+1,k+1}} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\dfrac{a^{(k+1/2)}_{n,k+1}}{a^{(k+1/2)}_{k+1,k+1}} & 0 & \dots & 1 \end{bmatrix},$$

$A^{(k+1)} = L_{k+1} A^{(k+1/2)}$, $b^{(k+1)} = L_{k+1} b^{(k+1/2)}$:

$$A^{(k+1)} = \left[\begin{array}{c|c:ccc} U_k & & & U_{k,n-k} & \\ \hline & a^{(k+1/2)}_{k+1,k+1} & a^{(k+1/2)}_{k+1,k+2} & \dots & a^{(k+1/2)}_{k+1,n} \\ \hdashline & 0 & a^{(k+1)}_{k+2,k+2} & \dots & a^{(k+1)}_{k+2,n} \\ 0 & \vdots & \vdots & \ddots & \vdots \\ & 0 & a^{(k+1)}_{n,k+2} & \dots & a^{(k+1)}_{n,n} \end{array}\right], \quad b^{(k+1)} = \left[\begin{array}{c} y^{(k)} \\ \hline b^{(k+1/2)}_{k+1} \\ \hdashline b^{(k+1)}_{k+2} \\ \vdots \\ b^{(k+1)}_n \end{array}\right].$$

$\det A^{(k)} \neq 0 \Rightarrow \det A^{(k+1)} \neq 0$.

Очевидно, что, если $\det A \neq 0$, то выполнив $n-1$ шаг, получим систему с верхней треугольной матрицей: $A^{(n-1)}x \equiv Ux = b^{(n-1)} \equiv y$.

Теорема. Если $\det A \neq 0$, то $PA = LU$, где $P = P_{n-1,i_{n-1}} \cdot \ldots \cdot P_{1,i_1}$, $L^{-1} = \tilde{L}_{n-1} \cdot \ldots \cdot \tilde{L}_1$,

$$\tilde{L}_k = P_{n-1,i_{n-1}} \cdot \ldots \cdot P_{k+1,i_{k+1}} \cdot L_k \cdot P_{k+1,i_{k+1}} \cdot \ldots \cdot P_{n-1,i_{n-1}}$$

Доказать эту теорему в качестве упражнения, проверив, что матрицы L_k и $\tilde{L}_k$ имеют одинаковую структуру.

Метод вращений решения системы уравнений

Элементарная матрица вращения

$$Q_{k,l}=\left[\begin{array}{c|ccc|c} E_{k-1} & & 0 & & 0 \\ \hline & \bar{c}_{k,l} & 0 & -\bar{s}_{k,l} & \\ 0 & 0 & E_{l-k-1} & 0 & 0 \\ & s_{k,l} & 0 & c_{k,l} & \\ \hline 0 & & 0 & & E_{n-l} \end{array}\right] \begin{array}{l} \\ -k\text{я строка} \\ \\ -l\text{я строка} \\ \\ \end{array} \qquad c_{k,l}\cdot\bar{c}_{k,l}+s_{k,l}\cdot\bar{s}_{k,l}=1, \qquad (k<l).$$

Доказать, что $Q_{k,l}$ – унитарная матрица, т.е. $Q_{k,l}(Q_{k,l})^*=(Q_{k,l})^*Q_{k,l}=E$.

Доказать, что $\det Q_{k,l}=1$.

Доказать, что при умножении на матрицу $Q_{k,l}$ матрицы A слева ($Q_{k,l}A$) изменяются только k и l строки матрицы A.

k–ый шаг метода вращений

Предположим, что после $k-1$ шага система $Ax=b$ с помощью умножения слева на ортогональную матрицу приведена к виду $A^{(k-1)}x=b^{(k-1)}$, где

$$A^{(k-1)}=\left[\begin{array}{c|ccc} R_{k-1} & & R_{k-1,n-k+1} & \\ \hline & a_{k,k}^{(k-1)} & \ldots & a_{k,n}^{(k-1)} \\ 0 & \vdots & \ldots & \vdots \\ & a_{n,k}^{(k-1)} & \ldots & a_{n,n}^{(k-1)} \end{array}\right], \quad b^{(k-1)}=\left[\begin{array}{c} y^{(k-1)} \\ \hline b_k^{(k-1)} \\ \vdots \\ b_n^{(k-1)} \end{array}\right] \quad (A^{(0)}=A,\ b^{(0)}=b).$$

Тогда k–ый шаг состоит из умножения системы $A^{(k-1)}x=b^{(k-1)}$ слева на элементарные матрицы вращений $Q_{k,k+1},\ldots,Q_{k,n}$:

$$Q_{k,k+i}=\left[\begin{array}{c|c} E_{k-1} & 0 \\ \hline 0 & Q_{1,1+i}^{(n-k+1)} \end{array}\right], \quad A^{(k-1,i)}=Q_{k,k+i}A^{(k-1,i-1)}, \quad b^{(k-1,i)}=Q_{k,k+i}b^{(k-1,i-1)},$$ где

$$c_{k,k+i}=\frac{a_{k,k}^{(k-1,i-1)}}{r_{k,k+i}}, \quad s_{k,k+i}=-\frac{a_{k+i,k}^{(k-1,i-1)}}{r_{k,k+i}},$$ если $r_{k,k+i}=\sqrt{|a_{k,k}^{(k-1,i-1)}|^2+|a_{k+i,k}^{(k-1,i-1)}|^2}\neq 0$,

$Q_{k,k+i}=E$, если $r_{k,k+i}=0$.

В результате получим $A^{(k)}=Q_kA^{(k-1)}$, $b^{(k)}=Q_kb^{(k-1)}$, где $Q_k=Q_{k,n}\ldots Q_{k,k+2}Q_{k,k+1}$.

Выполнив $n-1$ шаг, получим систему с верхней треугольной матрицей: $A^{(n-1)}x\equiv Rx=b^{(n-1)}\equiv y$ (заметим, что, если $\det A\neq 0$, то и $\det R\neq 0$).

Если определить унитарную матрицу $Q^* = Q_{n-1} \cdot \ldots \cdot Q_1$, то справедлива

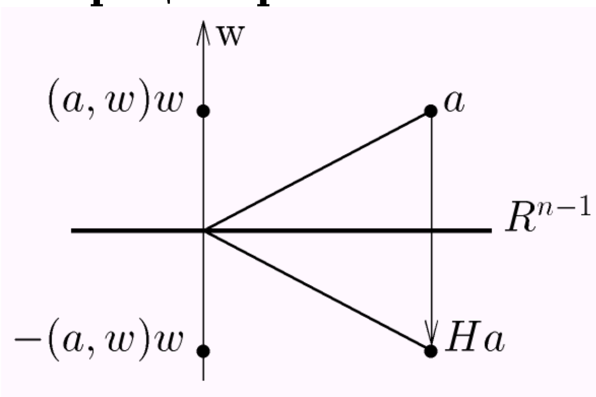
Теорема. $\forall A \ \ \exists A = Q \cdot R$.

Доказать, что $\mathrm{cond}_2 A = \mathrm{cond}_2 Q \cdot \mathrm{cond}_2 R = \mathrm{cond}_2 R$.

Лекция 4.

Метод отражений решения системы уравнений

Матрица отражения



$$Ha = a - 2(a,w)\cdot w = (E - 2\cdot w\cdot w^*)a$$

если заданы векторы a и Ha, то

$$w = \frac{a - Ha}{\|a - Ha\|_2}$$

Доказать, что $H = H^* = H^{-1}$, $\det H = -1$.

k–ый шаг метода отражений

Предположим, что после $k-1$ шага система $Ax = b$ с помощью умножения слева на ортогональную матрицу приведена к виду $A^{(k-1)}x = b^{(k-1)}$, где

$$A^{(k-1)} = \left[\begin{array}{c|ccc} R_{k-1} & & R_{k-1,n-k+1} & \\ \hline & a_{k,k}^{(k-1)} & \cdots & a_{k,n}^{(k-1)} \\ 0 & \vdots & \ddots & \vdots \\ & a_{n,k}^{(k-1)} & \cdots & a_{n,n}^{(k-1)} \end{array}\right], \quad b^{(k-1)} = \left[\begin{array}{c} y^{(k-1)} \\ \hline b_k^{(k-1)} \\ \vdots \\ b_n^{(k-1)} \end{array}\right] \quad (A^{(0)} = A,\ b^{(0)} = b).$$

Тогда k–ый шаг состоит из умножения системы $A^{(k-1)}x = b^{(k-1)}$ слева на ортогональную матрицу вращения H_k:

$$H_k = \left[\begin{array}{c|c} E_{k-1} & 0 \\ \hline 0 & E_{n-k+1} - 2w_1^{(n-k+1)}\cdot[w_1^{(n-k+1)}]^* \end{array}\right], \quad A^{(k)} = H_k A^{(k-1)}, \quad b^{(k)} = H_k b^{(k-1)},$$

где

$w_1^{(n-k+1)} = 0$	если $r_k = \sqrt{\lvert a_{k,k}^{(k-1)}\rvert^2 + \ldots + \lvert a_{n,k}^{(k-1)}\rvert^2} = 0$ или $a_{k+1,k}^{(k-1)} = \ldots = a_{n,k}^{(k-1)} = 0$,
$w_1^{(n-k+1)} = \dfrac{a_1^{(k-1)} - r_k\cdot e_1^{(k-1)}}{\lVert a_1^{(k-1)} - r_k\cdot e_1^{(k-1)}\rVert_2}$	если $r_k \neq 0$ & $a_{k,k}^{(k-1)} = 0$ (здесь $a_1^{(k-1)} = \begin{bmatrix} a_{k,k}^{(k-1)} \\ \vdots \\ a_{n,k}^{(k-1)} \end{bmatrix}$, $e_1^{(k-1)}$ – первый орт),
$w_1^{(n-k+1)} = \dfrac{a_1^{(k-1)} - \beta_k r_k\cdot e_1^{(k-1)}}{\lVert a_1^{(k-1)} - \beta_k r_k\cdot e_1^{(k-1)}\rVert_2}$	если $r_k \neq 0$ & $a_{k,k}^{(k-1)} \neq 0$ (здесь $\beta_k = -\dfrac{a_{k,k}^{(k-1)}}{\lvert a_{k,k}^{(k-1)}\rvert}$).

Выполнив $n-1$ шаг, получим систему с верхней треугольной матрицей: $A^{(n-1)}x \equiv Rx = b^{(n-1)} \equiv y$ (заметим, что, если $\det A \neq 0$, то и $\det R \neq 0$).

Решение системы с вырожденной матрицей

HR–разложение с перестановками столбцов матрицы A

1–ый шаг. Определим номер столбца j_1 матрицы $A=[a_1,\dots,a_n]$ из условия
$\|a_{j_1}\|_2=\max\|a_j\|_2$ и матрицу перестановок P_{1,j_1}.

Для матрицы $A^{(1/2)}=AP_{1,j_1}$ определим матрицу отражения H_1:

$$A^{(1)}=H_1A^{(1/2)}=H_1AP_{1,j_1}=\left[\begin{array}{c|ccc} R_1 & & \tilde{R}_{1,n-1} & \\ \hline & a^{(1)}_{2,2} & \dots & a^{(1)}_{2,n} \\ 0 & \vdots & \ddots & \vdots \\ & a^{(1)}_{n,2} & \dots & a^{(1)}_{n,n} \end{array}\right],\quad r_{11}=R_1.$$

Доказать: $|r_{11}|\ge\|a_j\|\ge|a^{(1)}_{ij}|\quad\forall\, i,j$

k–ый шаг. После $k-1$ шага имеем

$$A^{(k-1)}=\left[\begin{array}{c|ccc} R_{k-1} & & \tilde{R}_{k-1,n-k+1} & \\ \hline & a^{(k-1)}_{k,k} & \dots & a^{(k-1)}_{k,n} \\ 0 & \vdots & \ddots & \vdots \\ & a^{(k-1)}_{n,k} & \dots & a^{(k-1)}_{n,n} \end{array}\right],\quad |r_{k-1,k-1}|\ge\left\|\begin{matrix} a^{(k-1)}_{k-1,j} \\ a^{(k-1)}_{k,j} \\ \vdots \\ a^{(k-1)}_{n,j}\end{matrix}\right\|_2\ge\max_{k-1\le i,j\le n}|a^{(k-1)}_{i,j}|.$$

Определяем номер столбца j_k из условия

$$\left\|\begin{matrix} a^{(k-1)}_{k,j_k} \\ \vdots \\ a^{(k-1)}_{n,j_k}\end{matrix}\right\|_2=\max_{k\le j\le n}\left\|\begin{matrix} a^{(k-1)}_{k,j} \\ \vdots \\ a^{(k-1)}_{n,j}\end{matrix}\right\|_2$$

и для $A^{(k-1/2)}=A^{(k-1)}P_{k,j_k}$ определяем матрицу отражения H_k:

$$A^{(k)}=H_kA^{(k-1/2)}=H_kA^{(k-1)}P_{k,j_k}=\left[\begin{array}{c|ccc} R_k & & \tilde{R}_{k,n-k} & \\ \hline & a^{(k)}_{k+1,k+1} & \dots & a^{(k)}_{k+1,n} \\ 0 & \vdots & \ddots & \vdots \\ & a^{(k)}_{n,k+1} & \dots & a^{(k)}_{n,n} \end{array}\right].$$

Доказать: $|r_{k,k}|\ge\max\limits_{k\le j\le n}\left\|\begin{matrix} a^{(k)}_{k,j} \\ a^{(k)}_{k+1,j} \\ \vdots \\ a^{(k)}_{n,j}\end{matrix}\right\|_2\ge\max\limits_{k\le i,j\le n}|a^{(k)}_{i,j}|.$

Ответ: Если $t=\dim(\ker A)$, то после $n-t$ шагов имеем

$$(H_{n-t}\cdot\ldots\cdot H_1)A(P_{1,j_1}\cdot\ldots\cdot P_{n-t,j_{n-t}}) = HAP = R = \begin{bmatrix} R_{n-t} & R_{n-t,t} \\ 0 & 0 \end{bmatrix},$$

где H и P – ортогональные матрицы.

Совместность системы с вырожденной матрицей

Система $Ax=b$ называется совместной, если она имеет решение. Следовательно, система совместна $\Leftrightarrow$ $b\in \mathrm{Im}A$.

$\{x^*+y \ \ \forall\, y\in \ker A\}$– общее решение системы, где x^*– любое ее решение.

Теорема. Если система $Ax=b$ совместна ($b\in \mathrm{Im}A$), то $\forall\, B:\ \det B\neq 0$ совместна система $(BA)x=(Bb)$ и множества решений этих систем совпадают.

Система $Ax=b$ несовместна, если $b\notin \mathrm{Im}A$.

В этом случае ее обобщенным решением (относительно векторной нормы $\|\cdot\|$) называют вектор $x:\ \|Ax-b\|=\min\limits_{y}\|Ay-b\|$.

Доказать: общее решение совместной системы совпадает с множеством ее обобщенных решений.

Доказать: множество обобщенных решений $\{x:\ \|Ax-b\|_2=\min\|Ay-b\|_2\}$ совпадает с общим решением системы $A^*A\,x=A^*\,b$.

Применение HR–разложения с перестановками столбцов для решения совместной системы

Выполним эквивалентное преобразование совместной системы $Ax=b$:

$$Ry=g:\ \ R=HAP,\ \ y=P^*x,\ \ g=Hb.$$

Из–за ошибок округления эта система будет иметь вид:

$$\begin{bmatrix} R_{n-t} & R_{n-t,t} \\ 0 & \varepsilon_t \end{bmatrix}\begin{pmatrix} y^{(n-t)} \\ y^{(t)} \end{pmatrix}=\begin{pmatrix} g^{(n-t)} \\ \delta^{(t)} \end{pmatrix},$$

где матрица ε_t и вектор $\delta^{(t)}$ должны иметь малые по модулю элементы. Заменяем их на нулевые матрицу и вектор (диагональные элементы матрицы R по модулю мажорируют все левее и ниже лежащие элементы, как только очередной диагональный элемент стал “намного” меньше предыдущего, то и остальные элементы почти нулевые):

$$\begin{bmatrix} R_{n-t} & R_{n-t,t} \\ 0 & 0 \end{bmatrix}\begin{pmatrix} y^{(n-t)} \\ y^{(t)} \end{pmatrix}=\begin{pmatrix} g^{(n-t)} \\ 0 \end{pmatrix},$$

очевидно, что общее решение этой системы определяется формулой

$$\begin{pmatrix} y^{(n-t)} \\ y^{(t)} \end{pmatrix}=\begin{pmatrix} R_{n-t}^{-1}(g^{(n-t)}-R_{n-t,t}y^{(t)}) \\ y^{(t)} \end{pmatrix} \quad \forall\ y^{(t)}\in R^t,$$

а решение исходной системы $x = Py$.

Метод прогонки решения систем с трехдиагональной матрицей

LU–разложение трехдиагональной матрицы A:

$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & & 0 & a_n & b_n \end{bmatrix} = \begin{bmatrix} d_1 & 0 & & & 0 \\ a_2 & d_2 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & d_{n-1} & 0 \\ 0 & & 0 & a_n & d_n \end{bmatrix} \cdot \begin{bmatrix} 1 & u_1 & & & 0 \\ 0 & 1 & u_2 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & u_{n-1} \\ 0 & & 0 & 0 & 1 \end{bmatrix}$$

где (проверить)

$$\begin{array}{ll} d_1 = b_1 & u_1 = d_1^{-1}c_1 \\ d_2 = b_2 - a_2u_1 & u_2 = d_2^{-1}c_2 \\ \ldots\ldots & \ldots\ldots \\ d_i = b_i - a_iu_{i-1} & u_i = d_i^{-1}c_i \\ \ldots\ldots & \ldots\ldots\ldots \\ d_{n-1} = b_{n-1} - a_{n-1}u_{n-2} & u_{n-1} = d_{n-1}^{-1}c_{n-1} \\ d_n = b_n - a_nu_{n-1} & \end{array}$$

Формулы метода прогонки для системы $Ax = f$:

сначала вычисляем (рекуррентно):

$$u_1 = b_1^{-1}c_1,\ u_2 = (b_2 - a_2u_1)^{-1}c_2,\ \ldots,\ u_i = (b_i - a_iu_{i-1})^{-1}c_i,\ \ldots,\ u_{n-1} = (b_{n-1} - a_{n-1}u_{n-2})^{-1}c_{n-1}$$

и решаем систему с матрицей L (прямой ход):

$$\begin{bmatrix} d_1 & 0 & & & 0 \\ a_2 & d_2 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & d_{n-1} & 0 \\ 0 & & 0 & a_n & d_n \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix}, \qquad \begin{array}{l} y_1 = d_1^{-1}f_1 = b_1^{-1}f_1 \\ y_i = d_i^{-1}(f_i - a_iy_{i-1}) = \\ \quad = (b_i - a_iu_{i-1})^{-1}(f_i - a_iy_{i-1}) \\ \qquad i = 2, \ldots, n-1 \\ y_n = (b_n - a_nu_{n-1})^{-1}(f_n - a_ny_{n-1}) \end{array}$$

и, наконец, решаем систему $Ux = y$ (обратный ход):

$$x_n = y_n,\ \ x_{n-1} = -u_{n-1}x_n,\ \ldots,\ \ x_i = -u_ix_{i+1},\ \ldots,\ \ x_1 = -u_1x_2.$$

Теорема. Если $\forall i \ \ |b_i| > |a_i| + |c_i| \ \ (a_1 = c_n = 0)$, то $\det A_k \neq 0 \ \forall k$ (т.е. LU–разложение существует и метод прогонки применим).

Док–во. (от противного) Пусть $\exists k:\ \det A_k = 0$,

тогда $\exists x^{(k)} \neq 0:\ \ A_kx^{(k)} = 0$ и $\exists i:\ |x_i^{(k)}| = \max\limits_{1\le j\le k}|x_j^{(k)}| > 0$.

Разделим равенство $a_ix_{i-1}^{(k)} + b_ix_i^{(k)} + c_ix_{i+1}^{(k)} = 0$ на $x_i^{(k)}$ и оценим b_i:

$|b_i| \le |a_i| \cdot \dfrac{|x_{i-1}^{(k)}|}{|x_i^{(k)}|} + |c_i| \cdot \dfrac{|x_{i-1}^{(k)}|}{|x_i^{(k)}|} \le |a_i| + |c_i|$ – противоречие условию.

Лекция 5. Итерационные методы решения линейных уравнений

Мы будем рассматривать только вещественные системы линейных алгебраических уравнений, так как система уравнений $Ax = b$ над полем комплексных чисел сводится (доказать) к системе

$$\begin{bmatrix} \operatorname{Re} A & -\operatorname{Im} A \\ \operatorname{Im} A & \operatorname{Re} A \end{bmatrix} \begin{pmatrix} \operatorname{Re} x \\ \operatorname{Im} x \end{pmatrix} = \begin{pmatrix} \operatorname{Re} b \\ \operatorname{Im} b \end{pmatrix}$$

с вещественными коэффициентами.

Пример и основные определения

Пример:

пусть для матрицы системы $Ax = b$ построена обратная $A^{-1} = (LU)^{-1}$. Из–за ошибок округления мы получим не обратную матрицу, а к ней близкую: $\tilde{A}^{-1}$. Тогда $\tilde{x} = \tilde{A}^{-1} x \neq x$, а для разности $\tilde{x} - x$ имеем уравнение $A(\tilde{x} - x) = A\tilde{x} - b$, приближенное решение которого $\widetilde{(\tilde{x} - x)} = \tilde{A}^{-1}(A\tilde{x} - b) \Rightarrow \tilde{\tilde{x}} = \tilde{x} - \tilde{A}^{-1}(A\tilde{x} - b)$ или итерационное уточнение

$$x^{k+1} = x^k - \tilde{A}^{-1}(Ax^k - b), \ \ k = 0, 1, \dots .$$

Одношаговый (двухслойный) итерационный метод решения $Ax = b$**:**

$$x^{k+1} = x^k - H_k (A x^k - b)$$

$$x^0 - \text{задан}, \quad k = 0, 1, 2, \dots; \quad H_k - \text{заданные матрицы.}$$

x^k – k-тое приближение (к решению системы),

$z^k = x^k - x$ – ошибка k-той итерации	$z^{k+1} = z^k - H_k A z^k = (E - H_k A) z^k$ – процесс для ошибки, $S_k = E - H_k A$ – матрица шага для ошибки;
$r^k = A x^k - b = A z^k$ – невязка k-той итерации	$r^{k+1} = r^k - H_k A r^k = (E - A H_k) r^k$ – процесс для невязки, $T_k = E - A H_k$ – матрица шага для невязки;

Метод называется сходящимся, если $\lim\limits_{k \to \infty} \| z^k \| = 0 \quad \forall\, x^0 \in R^n$**.**

(Так как в R^n все нормы эквивалентны, то определение сходимости от нормы не зависит.)

Стационарный одношаговый итерационный метод решения $Ax = b$**:**

$$x^{k+1} = x^k - H(A x^k - b)$$

$$x^0 - \text{задан}, \quad k = 0, 1, 2, \dots; \quad H - \text{заданная матрица.}$$

Впредь мы будем предполагать, что $\det A \neq 0$ и $\det H \neq 0$.

Условия сходимости стационарного итерационного метода

$$x^{k+1} = x^k - H(Ax^k - b), \quad x^0 - \text{задан}, \quad k = 0, 1, 2, \ldots; \quad H - \text{заданная матрица}.$$

Достаточные условия:

Теорема. Если $\|S\| = \|E - HA\| < 1$, то $\|z^k\| \to 0$, т.е. $x^k \to x \ \ \forall x^0 \in R^n$.

Док-во. $\|x^k - x\| = \|z^k\| = \|Sz^{k-1}\| \le \|S\| \cdot \|z^{k-1}\| = \|S\| \cdot \|Sz^{k-2}\| \le$

$$\le \|S\|^2 \cdot \|z^{k-2}\| \le \|S\|^k \cdot \|z^0\| \to 0.$$

Теорема. Если $\|T\| = \|E - AH\| < 1$, то $\|z^k\| \to 0$, т.е. $x^k \to x \ \ \forall x^0 \in R^n$.

Док-во. $\|Ax^k - b\| = \|r^k\| = \|Tr^{k-1}\| \le \|T\| \cdot \|r^{k-1}\| = \|T\| \cdot \|Tr^{k-2}\| \le$

$$\le \|T\|^2 \cdot \|r^{k-2}\| \le \|T\|^k \cdot \|r^0\| \to 0.$$

$\Rightarrow \quad \|x^k - x\| = \|z^k\| = \|A^{-1}r^k\| \le \|A^{-1}\| \cdot \|r^k\| \to 0.$

Необходимое и достаточное условие:

Теорема. $x^k \to x \ \ \forall x^0 \in R^n \quad \Leftrightarrow \quad \rho(S) < 1.$

Док-во. **Необходимость.**

Пусть $x^k - x = z^k = S^k z^0 \to 0 \ \ \forall z^0 \in R^n$, т.е. метод сходится.

Так как $\forall \lambda \in Sp(S) \ \exists z^0 \ne 0: \ Sz^0 = \lambda \cdot z^0$, то, выбрав $x^0 = x + z^0$, получим, что $\|z^k\| = \|S^k z^0\| = \|\lambda^k z^0\| = |\lambda|^k \cdot \|z^0\| \to 0$

$\Rightarrow \ |\lambda|^k \to 0 \ \Rightarrow \ |\lambda| < 1 \ \Rightarrow \ \rho(S) < 1.$

Достаточность.

Если докажем, что $\rho(S) < 1 \Rightarrow S^k \to 0$ (нулевой матрице),

то $x^k - x = z^k = S^k z^0 \to 0 \ \ \forall z^0 \in R^n$, т.е. метод сходится.

Итак, пусть $J = \mathrm{diag}\{J_1, \ldots, J_m\}$ – жорданова форма матрицы S, т.е.

$$S = QJQ^{-1}, \qquad J_i = \begin{bmatrix} \lambda_i & & & 0 \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_i \end{bmatrix}, \qquad \lambda_i \in Sp(S), \qquad |\lambda_i| < 1.$$

Практически очевидно, что $S^k \to 0 \ \Leftrightarrow \ J_i^k \to 0 \ \forall i$.

Пусть $k > n_i$ – порядка блока J_i и $\lambda_i \ne 0$, тогда (бином Ньютона)

$$J_i^k \equiv (\lambda_i E_i + P_i)^k = \sum_{t=0}^{k} C_k^t \lambda_i^{k-t} P_i^t = \sum_{t=0}^{n_i-1} C_k^t \lambda_i^{k-t} P_i^t \text{, т.к. } P_i^t = 0 \quad \forall\, t \ge n_i \,.$$

Т.к. $C_k^t = \dfrac{k!}{t! \cdot (k-t)!} = \dfrac{(k-t+1)\ldots k}{t!} \le \dfrac{k^t}{t!} \le k^n$,

$$\lim_{k\to\infty} k^n \lambda_i^k = 0 \quad \forall\, |\lambda_i| < 1$$

$$\Rightarrow \quad \lim_{k\to\infty} C_k^t \lambda_i^{k-t} = 0 \quad \forall\, t < n_i \quad \Rightarrow \quad J_i^k \to 0,$$ что и тр.док.

Асимптотическая скорость сходимости

Сколько нужно сделать итераций, чтобы ошибка z^k итерационного процесса

$$x^{k+1} = x^k - H(Ax^k - b), \quad x^0 - \text{задан}, \quad k = 0, 1, 2, \dots; \quad H - \text{заданная матрица}.$$

уменьшилась в ε^{-1} раз: $k = k(\varepsilon) = ?: \quad \|z^k\| / \|z^0\| \le \varepsilon$.

Теорема. Если $\|S\| = \|E - HA\| < 1$, то $k(\varepsilon) = \left[\dfrac{-\ln\varepsilon}{-\ln\|S\|}\right] + 1$.

Док–во. При $k \ge k(\varepsilon)$ имеем $\|z^k\| = \|S^k z^0\| \le \|S\|^k \cdot \|z^0\| \le \varepsilon \cdot \|z^0\|$.

Средняя скорость за k **итераций:** $R_k = -\ln \sqrt[k]{\|S^k\|}$ **(Доказать:** $R_k \ge -\ln\|S\|$**)**

$\Rightarrow \quad \|S^k\| = e^{-kR_k} \le \varepsilon$, если $k \ge \dfrac{-\ln\varepsilon}{R_k}$.

Асимптотическая скорость сходимости: $R_\infty = \lim\limits_{k\to\infty} R_k = -\ln\left\{\lim\limits_{k\to\infty} \sqrt[k]{\|S^k\|}\right\}$ **.**

Теорема. Если $\rho(S) = \rho(E - HA) < 1$, то $R_\infty = -\ln\rho(S)$.

Док–во. Из док–ва теоремы о необходимом и достаточном условии сходимости $\Rightarrow$ $\|S^k\|_\infty \le ck^n[\rho(S)]^k \quad \forall\, k \ge n$

Из эквивалентности норм $\Rightarrow$ $\|S^k\| \le \beta\|S^k\|_\infty$ $\Rightarrow$

$$\overline{\lim_{k\to\infty}} \sqrt[k]{\|S^k\|} \le \overline{\lim_{k\to\infty}} \left[\sqrt[k]{\beta c}\left(\sqrt[k]{k}\right)^n \rho(S)\right] = \rho(S).$$

Т.к. $\|S^k\| \ge \rho(S^k) = [\rho(S)]^k$, то $\underline{\lim_{k\to\infty}} \sqrt[k]{\|S^k\|} \ge \rho(S)$.

$\Rightarrow$

$$\underline{\lim_{k\to\infty}} \sqrt[k]{\|S^k\|} = \overline{\lim_{k\to\infty}} \sqrt[k]{\|S^k\|} = \rho(S), \text{ т.е. } R_\infty = -\ln\rho(S).$$

Принято считать, что из двух итерационных процессов лучше тот, у которого асимптотическая скорость сходимости больше.

Но использовать асимптотическую скорость сходимости для оценки числа итераций, необходимых для уменьшения начальной ошибки в ε^{-1} раз, можно только в случае $R_k = R_\infty \quad \forall\, k$.

Лекция 6.

Один из способов построения итерационного метода решения системы линейных алгебраических уравнений $Ax = b$ состоит из представления матрицы в виде $A = B - C$, переписи системы в виде $Bx = Cx + b$ и определении очередного приближения x^{k+1} по известному приближению x^k из решения системы $Bx^{k+1} = Cx^k + b$.

Доказать: $Bx^{k+1} = Cx^k + b \;\Rightarrow\; x^{k+1} = x^k - B^{-1}(Ax^k - b)$.

Метод Якоби

Если $D = \mathrm{diag}A = \mathrm{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\}$, то итерационный процесс

$$x^{k+1} = x^k - D^{-1}(Ax^k - b)$$

называется методом Якоби для решения системы $Ax = b$.

Сходимость в случае диагонального преобладания по строкам

Теорема. Если $|a_{ii}| > \sum_{j=1,\, j\neq i}^{n} |a_{ij}| \quad \forall\, i$, то метод Якоби сходится.

Док-во. i–тая строка матрицы $S = E - D^{-1}A$: $\left[\frac{a_{i1}}{a_{ii}}, \ldots, \frac{a_{i,i-1}}{a_{ii}}, 0, \frac{a_{i,i+1}}{a_{ii}}, \ldots, \frac{a_{in}}{a_{ii}}\right]$. Из условия теоремы $\Rightarrow$ $\left|\frac{a_{i1}}{a_{ii}}\right| + \ldots + \left|\frac{a_{i,i-1}}{a_{ii}}\right| + \left|\frac{a_{i,i+1}}{a_{ii}}\right| + \ldots + \left|\frac{a_{in}}{a_{ii}}\right| < 1 \;\Rightarrow\; \|S\|_\infty < 1$, т.е. выполняется достаточное условие сходимости.

Сходимость в случае диагонального преобладания по столбцам

Теорема. Если $|a_{jj}| > \sum_{i=1,\, i\neq j}^{n} |a_{ij}| \quad \forall\, j$, то метод Якоби сходится.

Док-во. j–ый столбец матрицы $T = E - AD^{-1}$: $\left[\frac{a_{1j}}{a_{jj}}, \ldots, \frac{a_{j-1,j}}{a_{jj}}, 0, \frac{a_{j+1,j}}{a_{jj}}, \ldots, \frac{a_{nj}}{a_{jj}}\right]$. Из условия теоремы $\Rightarrow$ $\left|\frac{a_{1j}}{a_{jj}}\right| + \ldots + \left|\frac{a_{j-1,j}}{a_{jj}}\right| + \left|\frac{a_{j+1,j}}{a_{jj}}\right| + \ldots + \left|\frac{a_{nj}}{a_{jj}}\right| < 1 \;\Rightarrow$ $\rho(S) = \rho(E - D^{-1}A) = \rho(DTD^{-1}) = \rho(T) \leq \|T\|_1 < 1$, т.е. выполняется необходимое условие сходимости.

Необходимое и достаточное условие сходимости метода Якоби в случае симметричной матрицы с положительной главной диагональю

Теорема. Если $A = A^* \;\&\; D > 0$, то

$\rho(S) = \rho(E - D^{-1}A) < 1$ (т.е. метод Якоби сходится) $\Leftrightarrow$

$\Leftrightarrow\; A > 0 \;\&\; 2D - A > 0$.

Док-во. 1. собственные значения матрицы $S = E - D^{-1}A$ – вещественные:

$\lambda(S) = 1 - \lambda(D^{-1}A) = 1 - \lambda(D^{1/2}[D^{-1}A]D^{-1/2}) = 1 - \lambda(D^{-1/2}AD^{-1/2})$,

$\lambda(D^{-1/2}AD^{-1/2})$ – вещественны, т.к. $D^{-1/2}AD^{-1/2} = [D^{-1/2}AD^{-1/2}]^*$.

2. $\Rightarrow\; \rho(S) < 1 \;\Leftrightarrow\; \lambda(D^{-1}A) \in (0, 2)$

2.1. $\lambda(D^{-1}A) > 0 \;\Leftrightarrow\; A > 0$:

т.к. $(Ay, y) \equiv (\{D^{-1/2}AD^{-1/2}\}[D^{1/2}y], [D^{1/2}y]) = (\{D^{-1/2}AD^{-1/2}\}z, z)$

и $\forall\, B = B^*\;\; (Bx, x) \ge \lambda_{min}(B) \cdot (x, x)$,

то $\lambda(D^{-1}A) = \lambda(D^{-1/2}AD^{-1/2}) > 0 \;\Rightarrow\; A > 0$;

$A > 0 \;\Rightarrow\; (\{D^{-1/2}AD^{-1/2}\}z, z) > 0 \;\Rightarrow\; \lambda(D^{-1}A) = \lambda(D^{-1/2}AD^{-1/2}) > 0$;

2.2. $\lambda(D^{-1}A) < 2 \;\Leftrightarrow\; 2D - A > 0$:

$\lambda(D^{-1}A) < 2 \;\Leftrightarrow\; \lambda(2E - D^{-1}A) = \lambda(D^{-1}[2D - A]) > 0$,

$\lambda(D^{-1}[2D - A]) > 0 \;\Leftrightarrow\; [2D - A] > 0$.

Метод Зейделя (Гаусса–Зейделя, Некрасова)

Если матрицу системы $Ax = b$ представить в виде суммы $A = -L + D - R$, где $D = \mathrm{diag}A = \mathrm{diag}\{a_{11}, a_{22}, ..., a_{nn}\}$

$$L = -\begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & 0 & \\ \vdots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & 0 & \\ a_{n1} & \cdots & \cdots & a_{n,n-1} & 0 \end{bmatrix}, \quad R = -\begin{bmatrix} 0 & a_{12} & \cdots & \cdots & a_{1n} \\ & 0 & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \vdots \\ & 0 & & 0 & a_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

то итерационный процесс

$$x^{k+1} = x^k - (D - L)^{-1}(Ax^k - b)$$

называется методом Зейделя для решения системы $Ax = b$.

Доказать: $S = E - (D - L)^{-1}A = E - (A + R)^{-1}A = (A + R)^{-1}R$.

Необходимое и достаточное условие сходимости метода Зейделя в случае симметричной матрицы с положительной главной диагональю

Теорема. Если $A = A^*$ & $D > 0$, то $\rho(S) = \rho(E - (D-L)^{-1}A) < 1$ (т.е. метод Зейделя сходится) $\Leftrightarrow A > 0$.

Док-во. Необходимость:

Пусть $\rho(S) = \rho(E - (D-L)^{-1}A) < 1$, но $A \blacksquare 0$, т.е. $\exists\, \varphi:\ (A\varphi,\varphi) < 0$

($A = A^* \Rightarrow$ все $\lambda(A)$ вещественны,
$A \blacksquare 0 \Rightarrow \lambda_{\min}(A) \le 0$, иначе $(Ay,y) \ge \lambda_{\min}(A)\cdot(y,y) > 0$,
$\det A \ne 0 \Rightarrow \lambda_{\min}(A) \ne 0 \Rightarrow \lambda_{\min}(A) < 0 \Rightarrow \varphi:\ A\varphi = \lambda_{\min}(A)\varphi$).

Зададим $z^0 = \varphi$ и оценим (Az^{k+1}, z^{k+1}):

$$(Az^{k+1},z^{k+1}) = (A\{E-[A+R]^{-1}A\}z^k, \{E-[A+R]^{-1}A\}z^k) =$$
$$= (Az^k - Ay^k,\ z^k - y^k) =$$
$$= (Az^k,z^k) - (Az^k,y^k) - (Ay^k,z^k) + (Ay^k,y^k) =$$
$$= (Az^k,z^k) - ([A+R]y^k,y^k) - (y^k,[A+R]y^k) + (Ay^k,y^k) =$$
$$= (Az^k,z^k) - ([A+R]y^k,y^k) - ([A+L]y^k,y^k) + (Ay^k,y^k) =$$
$$= (Az^k,z^k) - (Dy^k,y^k) \le (Az^k,z^k) \le \ldots \le (Az^0,z^0) < 0.$$

$\Rightarrow \lim \| z^k \| \ne 0$, метод не сходится, что противоречит $\rho(S) < 1$.

Достаточность.

Докажем, что $A > 0 \Rightarrow \rho(S) = \rho([A+R]^{-1}R) < 1$, т.е. метод сходится.

$$\lambda \in Sp(S) \Rightarrow \exists\, \varphi:\ \|\varphi\|_2 = 1,\ S\varphi \equiv [A+R]^{-1}R\varphi = \lambda\varphi$$

$$\Rightarrow \begin{array}{l} (R\varphi,\varphi) = \lambda[(A\varphi,\varphi) + (R\varphi,\varphi)] \\ r + \underline{i}\cdot\mu = \lambda[a + (r + \underline{i}\cdot\mu)] \end{array} \Rightarrow |\lambda|^2 = \frac{r^2+\mu^2}{a(a+2r)+r^2+\mu^2}.$$

$\Rightarrow$ если $a(a+2r) > 0$, то $|\lambda| < 1$:

1. $a = (A\varphi,\varphi) \ge \lambda_{\min}(A)\cdot(\varphi,\varphi) = \lambda_{\min}(A) > 0$.
2. $(A\varphi,\varphi) = (D\varphi,\varphi) - (R^*\varphi,\varphi) - (R\varphi,\varphi) =$
$$= (D\varphi,\varphi) - \overline{(R\varphi,\varphi)} - (R\varphi,\varphi) = d - 2r$$
$\Rightarrow a + 2r = d \equiv (D\varphi,\varphi) > 0$

Из 1.–2. $\Rightarrow \rho(S) < 1$. Но, более того, т.к.

- $r^2 + \mu^2 = |(R\varphi,\varphi)|^2 \le \|R\|_2^2 \cdot \|\varphi\|_2^2 = \|R\|_2^2 \equiv \rho(R^*R)$,
- $d = (D\varphi,\varphi) \ge d_{\min}\cdot(\varphi,\varphi) = d_{\min} = (Ae_{\min},e_{\min}) \ge \lambda_{\min}(A)$
(здесь $d_{\min} = a_{\min,\min} = \min\{a_{ii}\} > 0$, $e_{\min}$ – орт),

то

$$|\lambda|^2 = \frac{r^2+\mu^2}{a(a+2r)+r^2+\mu^2} \le \frac{r^2+\mu^2}{\lambda_{\min}^2(A)+r^2+\mu^2} \le \max_{0<x\le\rho(R^*R)} \frac{x}{\lambda_{\min}^2(A)+x} =$$

$$= \frac{\rho(R^*R)}{\lambda_{\min}^2(A)+\rho(R^*R)},\ \text{т.е. } \rho(S) \le \sqrt{\frac{\rho(R^*R)}{\lambda_{\min}^2(A)+\rho(R^*R)}}\,.$$

Лекция 7.

Функционал ошибки

Второй из способов построения итерационного метода решения системы линейных алгебраических уравнений $Ax = b$ ($\det A \neq 0$) состоит из построения последовательности приближений $\{x^k\}_{k=0}^{\infty}$ такой, что $\|z^{k+1}\| < \|z^k\|$, т.е. строгого убывания на каждом шаге **функционала ошибки** $f(x^k) \equiv \|x^k - x\|$.

Теорема. Если $f(x^{k+1}) = \|z^{k+1}\| < f(x^k) = \|z^k\| \quad \forall\, x^k \neq x \quad (z^k \neq 0)$ и отображение $S\colon R^n \to R^n$ (оператор шага для ошибки: $z^{k+1} = S(z^k)$) непрерывно при $z \neq 0$, то $x^k \to x$, т.е. $\|z^k\| \to 0$.

Док–во. Т.к. $0 \le \|z^{k+1}\| < \|z^k\|$, то $\|z^k\| \to \alpha \ge 0$.
Предположим, что $\alpha > 0$.
Т.к. $\|z^k\| < \|z^0\|$, то $\exists\, \{z^{k_m}\}$: $z^{k_m} \to z$, $\|z\| = \alpha$.
Т.к. $z \neq 0$, то $\|S(z)\| < \|z\|$ & $S(z^{k_m}) \to S(z)$.
Тогда, выполнив предельный переход в соотношениях

$$\begin{array}{ccccc} \|z^{k_m+1}\| & = & \|S(z^{k_m})\| & < & \|z^{k_m}\| \\ \downarrow & & \downarrow & & \downarrow \\ \alpha & = & \|S(z)\| & < & \|z\| = \alpha \end{array}$$

получим противоречие: $\alpha < \alpha \;\Rightarrow\; \|z^k\| \to \alpha = 0$.

Обычно используют нормы, порождаемые симметричной положительно определенной матрицей: $\|z\|_C = \sqrt{(Cz, z)}$.

Доказать: если $C = C^* > 0$, то $(x, y)_C \equiv (C\,x, y)$ – скалярное произведение, $\|z\|_C = \sqrt{(Cz, z)}$ – норма в R^n.

Метод полной релаксации

для решения системы $Ax = b$ с матрицей $A = A^* > 0$ – очередное приближение x^{k+1} определяется по известному приближению x^k за n шагов:

$$x^{k+i/n} = x^{k+(i-1)/n} - \alpha_{k,i} \cdot e_i = \begin{pmatrix} x_1^{k+1} \\ \vdots \\ x_{i-1}^{k+1} \\ x_i^k - \alpha_{k,i} \\ x_{i+1}^k \\ \vdots \\ x_n^k \end{pmatrix}, \quad i = 1, 2, \dots, n,$$

где параметр $\alpha_{k,i}$ выбирается из условия минимума $\|z^{k+i/n}\|_A$.

Теорема. $\alpha_{k,i} = \frac{r_i^{k+(i-1)/n}}{a_{ii}} \equiv \frac{a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1} + a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k - b_i}{a_{ii}}$

и $x^k \to x$.

Док-во. Т.к. $A = A^* > 0$, то имеем

$$\| z^{k+i/n} \|_A^2 = (Az^{k+i/n}, z^{k+i/n}) =$$

$$= (A[z^{k+(i-1)/n} - \alpha_{k,i}e_i], [z^{k+(i-1)/n} - \alpha_{k,i}e_i]) =$$

$$= \| z^{k+(i-1)/n} \|_A^2 - 2\alpha_{k,i}(Az^{k+(i-1)/n}, e_i) + \alpha_{k,i}^2(Ae_i, e_i) =$$

$$= \| z^{k+(i-1)/n} \|_A^2 - 2\alpha_{k,i} \cdot r_i^{k+(i-1)/n} + \alpha_{k,i}^2 \cdot a_{i,i} =$$

$$= \| z^{k+(i-1)/n} \|_A^2 - \frac{(r_i^{k+(i-1)/n})^2 - (\alpha_{k,i} \cdot a_{i,i} - r_i^{k+(i-1)/n})^2}{a_{i,i}}.$$

Очевидно, что при $\alpha_{k,i} \cdot a_{i,i} - r_i^{k+(i-1)/n} = 0$ будет максимальное уменьшение ошибки (полная релаксация):

$$\| z^{k+i/n} \|_A^2 = \| z^{k+(i-1)/n} \|_A^2 - \frac{(r_i^{k+(i-1)/n})^2}{a_{i,i}}.$$

$$\Rightarrow \quad \| z^{k+1} \|_A^2 = \| z^k \|_A^2 - \frac{(r_1^k)^2}{a_{1,1}} - \frac{(r_2^{k+1/n})^2}{a_{2,2}} - \ldots - \frac{(r_n^{k+(n-1)/n})^2}{a_{n,n}} < \| z^k \|_A^2,$$

если хотя бы одна из компонент невязки $r_i^{k+(i-1)/n} \neq 0$

(в противном случае $x^{k+i/n} = x^{k+(i-1)/n} = x^k$, $r^k = 0$, т.е. $x^k = x$).

Итак, функционал ошибки строго убывает.

Найдем оператор шага для ошибки:

имеем (проверить!):

$$x_i^{k+1} = x_i^k - \alpha_{k,i} = x_i^k - \frac{a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1} + a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k - b_i}{a_{ii}}$$

или $x^{k+1} = x^k - D^{-1}(-Lx^{k+1} + (D-R)x^k - b) = D^{-1}(Lx^{k+1} + Rx^k - b)$

$\Rightarrow$ $x^{k+1} = x^k - (D-L)^{-1}(Ax^k - b)$ – метод Зейделя (он сходится)

$\Rightarrow$ $S = E - (D-L)^{-1}A$ – непрерывный (всюду) оператор шага

$\Rightarrow$ $x^k \to x$ по теореме о функционале ошибки.

Метод неполной релаксации

для решения системы $Ax = b$ с матрицей $A = A^* > 0$ – очередное приближение x^{k+1} определяется по известному приближению x^k за n шагов:

$$x^{k+i/n} = x^{k+(i-1)/n} - \alpha_{k,i} \cdot e_i, \quad i = 1, 2, \dots, n,$$

где параметр $\alpha_{k,i} = \omega \cdot \frac{r_i^{k+(i-1)/n}}{a_{ii}}$, т.е. ошибка уменьшается меньше, чем в методе полной релаксации ($\omega = 1$):

$$\|z^{k+i/n}\|_A^2 = \|z^{k+(i-1)/n}\|_A^2 - \frac{(r_i^{k+(i-1)/n})^2 - (\alpha_{k,i} \cdot a_{i,i} - r_i^{k+(i-1)/n})^2}{a_{i,i}} =$$

$$= \|z^{k+(i-1)/n}\|_A^2 - [1 - (\omega - 1)^2] \frac{(r_i^{k+(i-1)/n})^2}{a_{i,i}} < \|z^{k+(i-1)/n}\|_A^2 \quad \forall\, \omega \in (0, 2), \; r_i^{k+(i-1)/n} \neq 0.$$

Расчетные формулы имеют вид (проверить!):

$$x_i^{k+1} = x_i^k - \alpha_{k,i} = x_i^k - \omega \frac{a_{i,1}x_1^{k+1} + \dots + a_{i,i-1}x_{i-1}^{k+1} + a_{i,i}x_i^k + \dots + a_{i,n}x_n^k - b_i}{a_{ii}}$$

или

$$x^{k+1} = x^k - \omega D^{-1}(-Lx^{k+1} + (D - R)x^k - b)$$

$$\Rightarrow \; (D - \omega L)x^{k+1} = Dx^k - \omega Dx^k + \omega(Rx^k + b) =$$

$$= Dx^k - \omega Lx^k + \omega Lx^k - \omega Dx^k + \omega(Rx^k + b) =$$

$$= (D - \omega L)x^k - \omega(Ax^k - b)$$

Теорема. Если $A = A^* > 0$, то метод неполной релаксации сходится $\forall\, \omega \in (0, 2)$.

Док-во практически совпадает с доказательством сходимости метода полной релаксации.

Оценка сходимости методов релаксации

Итак, ошибка $z^{k+1} = Sz^k \equiv (E - \omega(D - \omega L)^{-1}A)z^k$ монотонно убывает в норме $\|z\|_A = \sqrt{(Az, z)}$. Оценим $\|S\|_A^2 = \max\limits_{z \neq 0} \frac{(ASz, Sz)}{(Az, z)}$.

Т. к.

$$\omega(D - \omega L)^{-1} = \omega[D + \omega(R_1 - 0.5D)]^{-1} = \frac{2\omega}{2 - \omega}(D + \frac{2\omega}{2 - \omega}R_1)^{-1}, \text{ где } R_1 = \frac{1}{2}D - L > 0$$

, то $S = E - \tau(D + \tau R_1)^{-1}A \equiv E - \tau B^{-1}A, \quad \tau = \frac{2\omega}{2 - \omega} \in (0, \infty)$, если $\omega \in (0, 2)$.

$$\Rightarrow \|S\|_A^2 = \max_{z \neq 0}\left[1 - 2\tau\frac{(B^{-1}Az, Az)}{(Az, z)} + \tau^2\frac{(AB^{-1}Az, B^{-1}Az)}{(Az, z)}\right].$$

Т.к. $\tau(Ay, y) = \tau([R_1 + R_1^*]y, y) = 2(\tau R_1 y, y) = 2([B - D]y, y)$, то

$$\tau^2(A[B^{-1}Az], [B^{-1}Az]) = 2\tau(B[B^{-1}Az], [B^{-1}Az]) - 2\tau(D[B^{-1}Az], [B^{-1}Az])$$

$$= 2\tau(B^{-1}Az, Az) - 2\tau(DB^{-1}Az, B^{-1}Az)$$

$$\Rightarrow \|S\|_A^2 = \max_{z\neq 0}\left[1-2\tau\frac{(DB^{-1}Az, B^{-1}Az)}{(Az,z)}\right] = 1-2\tau\cdot\min_{z\neq 0}\frac{(DB^{-1}Az, B^{-1}Az)}{(Az,z)} =$$

$$= 1-2\tau\cdot\min_{z\neq 0}\frac{([A^{1/2}(B^{-1})^* DB^{-1}A^{1/2}]A^{1/2}z, A^{1/2}z)}{(A^{1/2}z, A^{1/2}z)} = 1-2\tau\gamma,$$

где $\gamma = \lambda_{\min}(A^{1/2}(B^{-1})^* DB^{-1}A^{1/2})$.

Пусть $A^{1/2}(B^{-1})^{*}DB^{-1}A^{1/2}y=\gamma\cdot y \quad\Rightarrow\quad Av=\gamma\cdot BD^{-1}B^{*}v,\quad v=A^{-1/2}y$.

Т.к. $BD^{-1}B^{*}=(D+\tau R_1)D^{-1}(D+\tau R_1^{*})=D+\tau A+\tau^2 R_1 D^{-1}R_1^{*}$,

то $\gamma=\dfrac{(Av,v)}{(Dv,v)+\tau(Av,v)+\tau^2(R_1D^{-1}R_1^{*}v,v)}$.

Теорема.

$$\gamma\ge\frac{1}{1/\delta+\tau+\tau^2\Delta}=\frac{\delta}{1+\tau\delta+\tau^2\delta\Delta},$$

$$\|S\|_A^2\le g(\tau)=\frac{1-\tau\delta+\tau^2\delta\Delta}{1+\tau\delta+\tau^2\delta\Delta}<1\quad\forall\,\tau>0,$$

где постоянные $\delta>0$ и $\Delta>0$ таковы, что

$$\delta\cdot(Dv,v)\le(Av,v),\ \ (R_1D^{-1}R_1^{*}v,v)\le\Delta\cdot(Av,v)\ \ \forall\, v$$

$(\delta D\le A,\ \ \delta\le\lambda_{min}(D^{-1}A),\ R_1D^{-1}R_1^{*}\le\Delta\cdot A,\ \ \Delta\ge\lambda_{max}(A^{-1}R_1D^{-1}R_1^{*}))$

Док-во очевидно.

Доказать: $\min\limits_{\tau>0} g(\tau)=g(\tau_*)=\dfrac{1-\sqrt{\delta/(4\Delta)}}{1+\sqrt{\delta/(4\Delta)}},\quad \tau_*=\dfrac{1}{\sqrt{\delta\Delta}}$.

Доказать: $\omega_*=\dfrac{2}{1+2\sqrt{\delta\Delta}}$.

Пример

$$A=\begin{bmatrix}2&-1&&&0\\-1&2&-1&&\\&\ddots&\ddots&\ddots&\\&&-1&2&-1\\0&&&-1&2\end{bmatrix}=\begin{bmatrix}1&0&&&0\\-1&1&0&&\\&\ddots&\ddots&\ddots&\\&&-1&1&0\\0&&&-1&1\end{bmatrix}+\begin{bmatrix}1&-1&&&0\\0&1&-1&&\\&\ddots&\ddots&\ddots&\\&&0&1&-1\\0&&&0&1\end{bmatrix}=R_1+R_1^{*},$$

$\delta=\lambda_{min}(D^{-1}A)=0.5\ 4\ \sin^2(\pi/(2(n+1)))\quad \pi/[2(n+1)^2]=1$,

т.к. $A=2\cdot R_1D^{-1}R_1^{*}+\mathrm{diag}\{1,0,\ldots,0\}$, то $0.5\cdot A\ge R_1D^{-1}R_1^{*}$ и $\Delta=0.5$,

тогда (проверить):

верхняя релаксация	$\tau_*=\dfrac{1}{\sqrt{\delta\Delta}}\quad 2(n+1)/\pi,\quad \omega_*=\dfrac{2}{1+\pi/(n+1)}>1$, $g(\tau_*)\quad 1-\dfrac{\pi}{n+1},\ \|S\|_A\quad \sqrt{1-\dfrac{\pi}{n+1}}\quad 1-\dfrac{\pi}{2(n+1)},\ k(\varepsilon)=\dfrac{2(n+1)}{\pi}\ \ln\dfrac{1}{\varepsilon}$
полная релаксация	$g(2)-1\quad \dfrac{2\pi^2}{(n+1)^2},\ \|S\|_A\quad \sqrt{g(2)}-1\quad \dfrac{\pi^2}{(n+1)^2}$, $k(\varepsilon)=\dfrac{(n+1)^2}{\pi^2}\ \ln\dfrac{1}{\varepsilon}$

т.е. метод верхней релаксации в $(n+1)/(2\pi)$ раз дешевле.

Лекция 8.

Градиент, метод наискорейшего спуска

Как выбирать вектор У при построении итерационного метода $x^{k+1} = x^k + \alpha \cdot y$ из условия минимизации ошибки: $\|z^{k+1}\|^2 = \min_{\alpha} \|z^k + \alpha \cdot y\|^2$?

Если $f(z) \equiv \|z\|^2 = (Az, z),\ A = A^* > 0$, то

$$f(z^{k+1}) = f(z^k) + \frac{df(z^k)}{d\alpha}\alpha + O(\alpha^2) = f(z^k) + \left[\frac{\partial f(z^k)}{\partial z_1} y_1 + \ldots + \frac{\partial f(z^k)}{\partial z_n} y_n\right]\alpha + O(\alpha^2) =$$

$$= f(z^k) + (\nabla f, y)\cdot\alpha + O(\alpha^2) = \|z^k\|^2 + 2\alpha\cdot(Az^k, y) + \alpha^2 \cdot \|y\|^2 \approx$$

$$\approx \|z^k\|^2 + 2\alpha\cdot(Az^k, y) \geq \|z^k\|^2 - 2|\alpha|\cdot(Az^k, Az^k).$$

Следовательно, $y = -\nabla f = -2Az^k = -2r^k,\quad \alpha > 0$.

Теорема. Метод наискорейшего спуска

$$\boxed{\begin{aligned} x^{k+1} &= x^k - \tau_k(Ax^k - b) \\ \tau_k &= \frac{(r^k, r^k)}{(Ar^k, r^k)}, \quad k = 0, 1, \ldots \end{aligned}}$$

сходится, если $A = A^* > 0$.

Док-во.

$$\|z^{k+1}\|_A^2 = (A[z^k - \tau_k Az^k], z^k - \tau_k Az^k) = \|z^k\|_A^2 - 2\tau_k(r^k, r^k) + \tau_k^2 (Ar^k, r^k)$$

минимум правой части достигается при $\tau_k = (r^k, r^k)/(Ar^k, r^k)$:

$$\|z^{k+1}\|_A^2 = \|z^k\|_A^2 - \frac{|(r^k, r^k)|^2}{(Ar^k, r^k)} < \|z^k\|_A^2,$$ если $r^k = Az^k \neq 0$.

Очевидно, что оператор S: $z^{k+1} = S(z^k) = z^k - \tau_k(z^k)\cdot Az^k$ непрерывен всюду, кроме, быть может, 0. $\Rightarrow\quad z^k \to 0$.

Метод минимальных невязок

В итерационном процессе $x^{k+1} = x^k - \tau_k(Ax^k - b)$ параметр τ_k будем выбирать из условия минимизации невязки: $(r^{k+1}, r^{k+1}) = \min_{\tau} (r^k - \tau Ar^k,\ r^k - \tau Ar^k)$.

Теорема. Метод минимальных невязок

$$\boxed{\begin{aligned} x^{k+1} &= x^k - \tau_k(Ax^k - b) \\ \tau_k &= \frac{(Ar^k, r^k)}{(Ar^k, Ar^k)}, \quad k = 0, 1, \ldots \end{aligned}}$$

сходится, если $A > 0$.

Док-во.

$$\|z^{k+1}\|_{A^*A}^2 = (A^*A[z^k - \tau_k Az^k], z^k - \tau_k Az^k) = (r^k - \tau_k Ar^k, r^k - \tau_k Ar^k) =$$

$$= \|z^k\|_{A^*A}^2 - 2\tau_k(Ar^k, r^k) + \tau_k^2(Ar^k, Ar^k)$$

минимум правой части достигается при $\tau_k = (Ar^k, r^k)/(Ar^k, Ar^k)$:

$\|z^{k+1}\|_A^2 = \|z^k\|_A^2 - \dfrac{|(Ar^k, r^k)|^2}{(Ar^k, Ar^k)} < \|z^k\|_A^2$, если $r^k \neq 0$ & $(Ar^k, r^k) \neq 0$.

Очевидно, что оператор S: $z^{k+1} = S(z^k) = z^k - \tau_k(z^k) \cdot Az^k$ непрерывен всюду, кроме , быть может, 0. $\Rightarrow$ $z^k \to 0$.

Метод простой итерации

В методах наискорейшего спуска и минимальных невязок для определения параметра τ_k на каждом шаге нужно вычислять два скалярных произведения (с умножением невязки на матрицу системы). Использование постоянного параметра $\tau_k \equiv \tau$ существенно уменьшает объем вычислений на каждом шаге.

Теорема. Если $A = A^* > 0$, то метод простой итерации

$$x^{k+1} = x^k - \tau \cdot (Ax^k - b), \quad k = 0, 1, \ldots$$

сходится при $\forall\ \tau \in (0, \frac{2}{\rho(A)})$,

$$\| z^k \|_2 \le [\rho_\tau]^k \cdot \| z^0 \|_2$$

$$\rho_\tau = \max\{|1 - \tau \cdot \lambda_{min}(A)|, |1 - \tau \cdot \lambda_{max}(A)|\}$$

При $\tau_{опт} = \frac{2}{\lambda_{min} + \lambda_{max}}$

$$\rho_{опт} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \le \rho_\tau \quad \forall\ \tau > 0.$$

Док-во.

$$\| z^k \|_2 = \| (E - \tau \cdot A) z^{k-1} \|_2 \le \| E - \tau \cdot A \|_2 \cdot \| z^{k-1} \|_2 \le$$

$$\le (\| E - \tau \cdot A \|_2)^k \cdot \| z^0 \|_2 = \rho^k (E - \tau \cdot A) \cdot \| z^0 \|_2$$

$$\rho_\tau \equiv \rho(E - \tau \cdot A) = \max_{\lambda \in Sp(A)} |1 - \tau \cdot \lambda| = \max\{|1 - \tau \cdot \lambda_{min}(A)|, |1 - \tau \cdot \lambda_{max}(A)|\}$$

т.к. функция $g_\tau = |1 - \tau \cdot \lambda|$ выпукла вниз.

$$|1 - \tau \cdot \lambda| < 1 \quad \Leftrightarrow \quad -1 < \tau \cdot \lambda - 1 < 1 \quad \forall\ \lambda \in Sp(A) \quad \Leftrightarrow \quad 0 < \tau < 2/\lambda_{max}$$

$\Rightarrow \quad \rho_\tau = \|S\|_2 < 1 \quad \forall\ \tau \in (0, \frac{2}{\rho(A)})$, метод сходится.

Оптимальный параметр выбираем из условия

$$\rho_{опт} \equiv \| E - \tau_{опт} \cdot A \|_2 = \min_{\tau > 0} \| E - \tau \cdot A \|_2 \equiv \min_{\tau > 0} \rho_\tau$$

легко проверить, что

$$\rho_\tau = \begin{cases} 1 - \tau \cdot \lambda_{min} & 0 < \tau \le \tau_{опт} \\ \tau \cdot \lambda_{max} - 1, & \tau_{опт} \le \tau < \frac{2}{\lambda_{max}} \end{cases} \quad \Bigg| \Rightarrow \quad \tau_{опт} = \frac{2}{\lambda_{min} + \lambda_{max}},$$

$$\rho_{опт} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}$$

Оценки сходимости МНС и ММН

Теорема. Если $A = A^* > 0$, то для ошибки $\|z^k\|$ метода наискорейшего спуска:

$$x^{k+1} = x^k - \tau_k (Ax^k - b), \quad \tau_k = \frac{(r^k, r^k)}{(Ar^k, r^k)}, \quad k = 0, 1, \dots,$$

справедливы оценки:

$$\|z^k\|_A \le \left[\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right]^k \|z^0\|_A,$$

$$\|z^k\|_2 \le \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \left[\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right]^k \|z^0\|_2$$

Док-во. Так как

$$\|z^{k+1}\|_A = \inf_\tau \|z^k - \tau A z^k\|_A \le \|z^k - \tau_{опт} A z^k\|_A \le \|E - \tau_{опт} A\|_A \cdot \|z^k\|_A$$

и $\|E - \tau_{опт} A\|_A^2 = \sup\limits_{z \ne 0} \dfrac{(A[E - \tau_{опт} A]z, [E - \tau_{опт} A]z)}{(Az, z)} =$

$$= \sup_{z \ne 0} \frac{([E - \tau_{опт} A]A^{0.5}z, [E - \tau_{опт} A]A^{0.5}z)}{(A^{0.5}z, A^{0.5}z)} = \|E - \tau_{опт} A\|_2^2 = \rho_{опт}^2,$$

то $\|z^k\|_A \le [\rho_{опт}]^k \|z^0\|_A$.

Т.к. $\lambda_{min}(z,z) \le (Az,z) \le \lambda_{max}(z,z)$ $\Rightarrow$ $\sqrt{\lambda_{min}}\|z\|_2 \le \|z\|_A \le \sqrt{\lambda_{max}}\|z\|_2$

то $\|z^k\|_A \le \sqrt{\lambda_{max}/\lambda_{min}} [\rho_{опт}]^k \|z^0\|_A$.

Теорема. Если $A = A^* > 0$, то для метода минимальных невязок:

$$x^{k+1} = x^k - \tau_k (Ax^k - b), \quad \tau_k = \frac{(Ar^k, r^k)}{(Ar^k, Ar^k)}, \quad k = 0, 1, \dots,$$

справедливы оценки:

$$\|r^k\|_2 \le \left[\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right]^k \|r^0\|_2, \qquad \|z^k\|_2 \le \frac{\lambda_{max}}{\lambda_{min}} \left[\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right]^k \|z^0\|_2$$

Док-во. Так как

$$\|r^{k+1}\|_2 = \inf_\tau \|r^k - \tau A r^k\|_2 \le \|r^k - \tau_{опт} A r^k\|_2 \le \|E - \tau_{опт} A\|_2 \cdot \|r^k\|_2$$

и $\|E - \tau_{опт} A\|_2 = \rho_{опт}$, то $\|r^k\|_2 \le [\rho_{опт}]^k \|r^0\|_2$.

Т.к. $\|r^k\|_2 = \|Az^k\|_2$ и $|\lambda_{min}|^2 (z,z) \le (Az, Az) \le |\lambda_{max}|^2 (z,z)$

$$\Rightarrow \quad \lambda_{min} \| z^k \|_2 \le \| r^k \|_2 \le \lambda_{max} \| z^k \|_2 \quad \Rightarrow \quad \| z^k \|_2 \le \frac{\lambda_{max}}{\lambda_{min}} [\rho_{опт}]^k \| z^0 \|_2 \, .$$

Лекция 9. Метод Ричардсона с чебышевскими параметрами

Задача оптимизации параметров

Параметры τ_k в итерационном методе $x^{k+1} = x^k - \tau_{k+1}(Ax^k - b)$ можно выбирать из условия минимизации спектрального радиуса $\rho(S_{m,\tau}(A))$ матрицы (оператора) ошибки за m шагов: $z^m = S_{m,\tau}(A)z^0 \equiv (E-\tau_m A)\dots(E-\tau_1 A)z^0$.

Если все параметры взять одинаковыми, то мы получим метод простой итерации и он сходится при известных условиях, т.е. предлагаемый способ построения итерационного метода может привести только к лучшему методу.

Мы будем предполагать, что $Sp(A) \subseteq [\alpha,\beta]$, $\alpha > 0$, т.е. все собственные значения матрицы системы линейных уравнений $Ax = b$ положительны.

Т.к. $\min\limits_{\tau_1,\dots,\tau_m} \rho(S_{m,\tau}(A)) = \min\limits_{\tau_1,\dots,\tau_m}\{\max\limits_{\lambda\in Sp(A)} |S_{m,\tau}(\lambda)|\} \le \min\limits_{\tau_1,\dots,\tau_m}\{\max\limits_{\lambda\in[\alpha,\beta]} |S_{m,\tau}(\lambda)|\}$, а последнюю минимаксную задачу решать проще (почему?), мы будем искать $\tau_1^*, \tau_2^*, \dots, \tau_m^*$:

$$\max_{\lambda\in[\alpha,\beta]} |S_{m,\tau^*}(\lambda)| = \min_{\tau_1,\dots,\tau_m}\{\max_{\lambda\in[\alpha,\beta]} |S_{m,\tau}(\lambda)|\} = \rho_m,$$

т.е. решать задачу о поиске полинома $S_{m,\tau^*}(\lambda)$ степени m, наименее уклоняющегося от нуля на отрезке $[\alpha,\beta]$ при условии $S_{m,\tau^*}(0) = 1$.

Тогда, т.к. $S_{m,\tau^*}(\lambda) \equiv (1-\tau_m^*\lambda)\dots(1-\tau_1^*\lambda) = \dfrac{(-1)^m}{\tau_1^*\dots\tau_m^*}(\lambda-\mu_1)\dots(\lambda-\mu_m)$, где μ_i – корни полинома $S_{m,\tau^*}(\lambda)$, $\tau_k^* = \dfrac{1}{\mu_k}$ и $\rho(S_{m,\tau^*}(A)) \le \max\limits_{\lambda\in[\alpha,\beta]} |S_{m,\tau^*}(\lambda)| = \rho_m$.

Если $A = A^*$, то $\|S_{m,\tau^*}(A)\|_2 = \rho(S_{m,\tau^*}(A)) \le \rho_m$ и, следовательно,

$$\|z^{mk}\|_2 \le [\rho_m]^k \cdot \|z^0\|_2$$

– оценка сходимости метода.

Полином Чебышева $T_m(x) \equiv \cos(m\cdot\arccos x),\quad x\in[-1,1]$ и решение задачи оптимизации параметров

Очевидно, что $T_0(x) = 1$, $T_1(x) = x$ – полиномы.

Т.к. $\cos((k+1)\varphi) + \cos((k-1)\varphi) = 2\cos(\varphi)\cdot\cos(k\varphi)$, то при $\varphi = \arccos x$ имеем $T_{k+1}(x) = 2\cdot T_1(x)\cdot T_k(x) - T_{k-1}(x)$ – полином при любом $k+1>1$.

Точки экстремумов $T_m(x) \equiv \cos(m \cdot \arccos x)$**:** $\hat{x}_k = \cos\frac{k\pi}{m}, \quad k = 0, 1, \ldots, m$**:**

$$-1 = \hat{x}_m < \hat{x}_{m-1} < \ldots < \hat{x}_1 < \hat{x}_0 = 1, \qquad T_m(\hat{x}_k) = -(-1)^k.$$

Корни полинома $T_m(x) \equiv \cos(m \cdot \arccos x)$**:** $x_k = \cos\frac{(2k-1)\pi}{2m}, \quad k = 1, \ldots, m$**:**

$$-1 \le \hat{x}_k < x_k < \hat{x}_{k-1} \le 1, \qquad T_m(x) \ne 0 \quad \forall \; x \notin [-1, 1].$$

Линейное преобразование $\lambda \in [\alpha,\beta] \;\to\; x \in [-1,1]$**:** $x(\lambda) = \dfrac{2\lambda - (\beta+\alpha)}{\beta-\alpha}$.

Рассмотрим **полином**: $S_m(\lambda) = \dfrac{1}{T_m\big(x(0)\big)} \cdot T_m\big(x(\lambda)\big)$.

Очевидно, что $S_m(0) = 1$, $\mu_k = \dfrac{(\beta-\alpha)\cdot x_k + (\beta+\alpha)}{2} \in [\alpha,\beta]$ – корни полинома.

Покажем, что этот полином наименее уклоняется от нуля на интервале $[\alpha,\beta]$ среди всех полиномов $P_m(\lambda)$, $P_m(0)=1$, т.е. $P_m(\lambda) = 1 - \lambda\cdot Q_{m-1}(\lambda)$.

Теорема. Если $\alpha > 0$, то $\|S_m(\lambda)\|_{C[\alpha,\beta]} = \inf\limits_{Q_{m-1}(\lambda)} \|1-\lambda Q_{m-1}(\lambda)\|_{C[\alpha,\beta]}$.

Док-во. Пусть $\|S_m(\lambda)\|_{C[\alpha,\beta]} > \inf\limits_{Q_{m-1}(\lambda)} \|1-\lambda Q_{m-1}(\lambda)\|_{C[\alpha,\beta]}$,

тогда $\exists\, Q_{m-1}(\lambda)$: $\|S_m(\lambda)\|_{C[\alpha,\beta]} > \|1-\lambda Q_{m-1}(\lambda)\|_{C[\alpha,\beta]}$.

Так как

- полином $S_m(\lambda) = 1 - \lambda P_m(\lambda)$ имеет экстремумы (одинаковые по модулю) в точках $\hat\mu_k \in [\alpha,\beta]$:
  $\dfrac{2\hat\mu_k - (\beta+\alpha)}{\beta-\alpha} = \hat x_k$, $k = 0, 1, \ldots, m$:
  $|S_m(\hat\mu_k)| = \|S_m(\lambda)\|_{C[\alpha,\beta]}$, и $\{S_m(\hat\mu_k)\}_{k=0}^{m}$ знакопеременна:

то разность $R_m(\lambda) \equiv S_m(\lambda) - [1-\lambda Q_{m-1}(\lambda)] = \lambda Q_{m-1}(\lambda) - \lambda P_{m-1}(\lambda)$

- полином степени m,
- последовательность $\{R_m(\hat\mu_k)\}_{k=0}^{m}$ знакопеременна $\Rightarrow$ в интервале $[\alpha,\beta]$ имеется m попарно различных корней полинома $R_m(\lambda)$ (т.к. внутри интервала $[\hat\mu_k, \hat\mu_{k-1}]$ имеется хотя бы один корень),
- $\lambda = 0 \notin [\alpha,\beta]$ – корень ($(m+1)$–ый) полинома $R_m(\lambda)$ (именно здесь мы использовали условие $\alpha > 0$).

$\Rightarrow$ $R_m(\lambda) \equiv 0$, т.е. $\|S_m(\lambda)\|_{C[\alpha,\beta]} = \|1-\lambda Q_{m-1}(\lambda)\|_{C[\alpha,\beta]}$ – противоречие.

Следовательно, $\|S_m(\lambda)\|_{C[\alpha,\beta]} = \inf\limits_{Q_{m-1}(\lambda)} \|1-\lambda Q_{m-1}(\lambda)\|_{C[\alpha,\beta]}$.

Осталось вычислить $\rho_m = \|S_m(\lambda)\|_{C[\alpha,\beta]}$.

Теорема. Если $\alpha > 0$, то $\rho_m = \|S_m(\lambda)\|_{C[\alpha,\beta]} = \dfrac{2\gamma^m}{1+\gamma^{2m}} < 1$, где

$$\gamma = \frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}}.$$

Док-во. Очевидно, $\rho_m = \|S_m(\lambda)\|_{C[\alpha,\beta]} = |T_m(x(0))|^{-1}$, $x(0) = -\frac{\beta+\alpha}{\beta-\alpha} < -1$.

Для вычисления $T_m(x(0))$ воспользуемся формулой

$$T_m(x) = \frac{(x + \sqrt{x^2-1})^m + (x - \sqrt{x^2-1})^m}{2} \text{ при } |x| > 1.$$

Заметим, что

$$x(0)+\sqrt{x^2(0)-1}=\frac{-(\sqrt{\beta})^2+2\sqrt{\beta}\sqrt{\alpha}-(\sqrt{\alpha})^2}{(\sqrt{\beta}+\sqrt{\alpha})\cdot(\sqrt{\beta}-\sqrt{\alpha})}=-\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}=-\gamma,$$

$$x(0)-\sqrt{x^2(0)-1}=\frac{1}{x(0)+\sqrt{x^2(0)-1}}=-\frac{1}{\gamma}.$$

Тогда

$$T_m(x(0))=\frac{(-\gamma)^m+(-\gamma^{-1})^m}{2}=(-1)^m\frac{1+\gamma^{2m}}{2\gamma^m}\quad\Rightarrow\quad \rho_m=\frac{2\gamma^m}{1+\gamma^{2m}}<1.$$

Док-во формулы $T_k(x)=0.5[(x+\sqrt{x^2-1})^k+(x-\sqrt{x^2-1})^k]$ при $|x|>1$.

Действительно, $T_0(x)=1$ и $T_1(x)=x$.

Осталось проверить, что $T_{k+1}(x)=2\cdot T_1(x)\cdot T_k(x)-T_{k-1}(x)$ или

$$(x+\sqrt{x^2-1})^{k+1}+(x-\sqrt{x^2-1})^{k+1}=$$
$$=2x[(x+\sqrt{x^2-1})^k+(x-\sqrt{x^2-1})^k]-[(x+\sqrt{x^2-1})^{k-1}+(x-\sqrt{x^2-1})^{k-1}]$$

.

Пусть $y=\sqrt{x^2-1}$, тогда

$$(x+y)^{k+1}+(x-y)^{k+1}=x\cdot[(x+y)^k+(x-y)^k]+y\cdot[(x+y)^k-(x-y)^k]=$$
$$=x\cdot[(x+y)^k+(x-y)^k]+$$
$$+y\cdot[x\cdot\{(x+y)^{k-1}-(x-y)^{k-1}\}+y\cdot\{(x+y)^{k-1}+(x-y)^{k-1}\}]=$$
$$=x\cdot[(x+y)^k+(x-y)^k]+$$
$$+y\cdot x\cdot\{(x+y)^{k-1}-(x-y)^{k-1}\}+(x^2-1)\cdot\{(x+y)^{k-1}+(x-y)^{k-1}\}=$$
$$=2x\cdot[(x+y)^k+(x-y)^k]-\{(x+y)^{k-1}+(x-y)^{k-1}\}+$$
$$+(-x^2-xy+yx+x^2)\cdot(x+y)^{k-1}+(-x^2+xy-yx+x^2)\cdot(x-y)^{k-1}=$$
$$=2x\cdot[(x+y)^k+(x-y)^k]-\{(x+y)^{k-1}+(x-y)^{k-1}\},$$ что и тр. док.

Итак, $S_{m,\tau^*}(\lambda)=S_m(\lambda)$ – решение задачи оптимизации параметров за m шагов.

Циклический метод Ричардсона: формулы и сходимость

Теорема. Если $A=A^*>0$ и известны оценки ее спектра: $SpA\in[\alpha,\beta]$, $\alpha>0$, то циклический метод Ричардсона (с длиной цикла m) решения системы $Ax=b$:

$$\begin{cases} x^{k+1}=x^k-\tau_{k+1}(Ax^k-b), \quad k=0,1,\ldots, \\ \tau_1,\ldots,\tau_m,\ \tau_{m+1}=\tau_1,\ldots,\tau_{2m}=\tau_m,\ldots\quad(\tau_{m+j}=\tau_j), \end{cases}$$

с чебышевскими параметрами

$$\tau_k = 2\left[(\beta+\alpha)+(\beta-\alpha)\cdot\cos\frac{(2k-1)\pi}{2m}\right]^{-1}$$

сходится и $\|z^{mk}\|_2 \le [\rho_m]^k \|z^0\|_2 = \left[\dfrac{2\gamma^m}{1+\gamma^{2m}}\right]^k \|z^0\|_2$, $\gamma = \dfrac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}$

.

Об устойчивости метода Ричардсона

Из–за ошибок округления реализация формул $x^{k+1} = x^k - \tau_{k+1}(Ax^k - b)$ неустойчива, т.к. норма оператора шага $\| E - \tau_{k+1} A \|_2$ для ошибки может быть значительно больше 1 (в методе простой итерации эта норма меньше 1).

Рассмотрим модельный пример, в котором ошибка округления возникает только на шаге с $\tau = \tau_m$ ($m = 10^4 \gg 1$):

$$\tilde{z}^m = (E - \tau_m A)[(E - \tau_{m-1} A)\ldots(E - \tau_1 A) z^0 + \varepsilon z^0] = z^m + \varepsilon (E - \tau_m A) z^0, \quad A = \begin{bmatrix} 1 & 0 \\ 0 & m \end{bmatrix}.$$

Проверить: $\rho_m \approx 2e^{-200} \quad \Rightarrow \quad \| z^m \|_2 \le 2e^{-200} \| z^0 \|_2 \approx 0.$

Проверить: $\tau_m \approx 1,\ E - \tau_m A \approx \begin{bmatrix} 0 & 0 \\ 0 & 1-m \end{bmatrix}$

$\Rightarrow \quad \| \tilde{z}^m \|_2 \approx \varepsilon m | z_2^0 | = 10^4 \varepsilon | z_2^0 | \approx | z_2^0 |$, если $\varepsilon = 10^{-4}$, т.е. фактически ошибка не уменьшилась.

Изменим упорядочение параметров: $\tau_2, \tau_3, \ldots, \tau_{m-1}, \tau_m, \tau_1$:

$$\hat{z}^m = (E - \tau_1 A)(E - \tau_m A)[(E - \tau_{m-1} A)\ldots(E - \tau_2 A) z^0 + \varepsilon z^0] =$$
$$= z^m + \varepsilon (E - \tau_1 A)(E - \tau_m A) z^0.$$

Проверить: $\tau_1 = \dfrac{1}{m} + O(m^{-2}),\ (E - \tau_1 A)(E - \tau_m A) \approx \begin{bmatrix} 0 & 0 \\ 0 & O(1) \end{bmatrix}$

$\Rightarrow \quad \| \hat{z}^m \|_2 \approx \varepsilon \cdot O(| z_2^0 |)$,

т.е. реализация с точностью до ошибок округления.

Из этого примера следует, что переупорядочение параметров существенно влияет на устойчивость вычислений в методе Ричардсона. Задача об оптимальном упорядочении параметров ставится следующим образом.

Пусть	$p = [p(1), \ldots, p(m)]$ – перестановка m–ки $(1, \ldots, m)$, $\nu_j(p) = \rho[(E - \tau_{p(j)} A)\ldots(E - \tau_{p(m)} A)],\ \nu(p) = \sum_{j=1}^{m-1} \tau_{p(j)} \nu_{j+1}(p) + \tau_{p(m)}$.
Найти	$p_{опт}:\ \nu(p_{опт}) = \inf_p \nu(p)$.

Решением этой задачи для $m = 2^t$ является следующая процедура:

$$\tau_1, \tau_2, \ldots, \tau_{m-1}, \tau_m$$
$$\Downarrow$$

$$r_1^1 = (\tau_m, \tau_1),\ r_2^1 = (\tau_{m-1}, \tau_2), \ldots, r_{m/2}^1 = (\tau_{m/2+1}, \tau_{m/2-1})$$

$$\Downarrow$$

$$r_1^2 = (r_{m/2}^1, r_1^1),\ r_2^2 = (r_{m/2-1}^1, r_2^1), \ldots, r_{m/4}^2 = (r_{m/4+1}^1, r_{m/4-1}^1)$$

$$\Downarrow$$

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

$$p_{\text{опт}} = r_1^t$$

Трехчленные формулы реализации метода Ричардсона с чебышевскими параметрами

Для решения системы $Ax=b$ попытаемся построить приближения $\{x^k\}$:

$$z^k = S_k(A)z^0 \quad \forall\ k \ge 0.$$

Т.к. $S_0(A)=E$ и $S_1(A)=E-\dfrac{2}{\beta+\alpha}A$ (проверить!), то

$$x^1 = x^0 - \frac{2}{\beta+\alpha}(Ax^0-b) \equiv x^0 - \frac{2}{\beta+\alpha}r^0.$$

Если ввести обозначение $t_k = T_k(\,x(0)\,) = T_k(-[\beta+\alpha]/[\beta-\alpha])$, то

$$t_{k+1}S_{k+1}(A) = 2t_kT_1\left(\frac{2}{\beta-\alpha}A - \frac{\beta+\alpha}{\beta-\alpha}E\right)S_k(A) - t_{k-1}S_{k-1}(A) =$$

$$= 2t_k\frac{2}{\beta-\alpha}AS_k(A) + 2t_kt_1S_k(A) - t_{k-1}S_{k-1}(A)$$

$$\Rightarrow \quad t_{k+1}z^{k+1} = 2t_k\frac{2}{\beta-\alpha}r^k + 2t_kt_1z^k - t_{k-1}z^{k-1} =$$

$$= 2t_k\frac{2}{\beta-\alpha}r^k + (2t_kt_1 - t_{k-1})z^k + t_{k-1}(z^k - z^{k-1}) =$$

$$= 2t_k\frac{2}{\beta-\alpha}r^k + t_{k+1}z^k + t_{k-1}(z^k - z^{k-1})$$

$$\Rightarrow \quad t_{k+1}x^{k+1} = 2t_k\frac{2}{\beta-\alpha}r^k + t_{k+1}x^k + t_{k-1}(x^k - x^{k-1}) \qquad \Rightarrow \quad \{x^k\}_{k=0}^{\infty} \text{ построена:}$$

$$x^{k+1} = 2\frac{t_k}{t_{k+1}}\frac{2}{\beta-\alpha}r^k + x^k + \frac{t_{k-1}}{t_k}\frac{t_k}{t_{k+1}}(x^k - x^{k-1}), \quad k=1,2,\dots .$$

Преобразуем эту формулу:

$$2\frac{t_k}{t_{k+1}}\frac{2}{\beta-\alpha} = -\frac{2t_kt_1}{t_{k+1}}\frac{2}{\beta+\alpha} = -\frac{t_{k+1}+t_{k-1}}{t_{k+1}}\frac{2}{\beta+\alpha} = -(1+\frac{t_{k-1}}{t_k}\frac{t_k}{t_{k+1}})\frac{2}{\beta+\alpha}.$$

Введем обозначение $\omega_k = t_{k-1}/t_k$, т.к. $t_{k+1} = 2t_1t_k - t_{k-1}$, то

$$\omega_1 = -\frac{\beta-\alpha}{\beta+\alpha}, \quad \omega_{k+1} = \frac{1}{2t_1-\omega_k} = \frac{1}{2(\omega_1)^{-1}-\omega_k}, \quad k=1,2,\dots .$$

Тогда

$$x^{k+1} = x^k + \omega_k\omega_{k+1}(x^k - x^{k-1}) - \frac{2}{\beta+\alpha}(1+\omega_k\omega_{k+1})\cdot(Ax^k-b), \quad k=1,2,\dots,$$

– двухшаговый (трехслойный) итерационный процесс.

Лекция 10.

Многошаговые методы. Вариационная оптимизация

Для определения параметров метода Ричардсона (простой итерации при $m=1$) для решения системы $Ax=b$ необходимо предварительное вычисление (точное или приближенное) границ спектра матрицы A, чего не требуется в методах наискорейшего спуска и минимальных невязок. Попытаемся выбрать параметры метода $x^k = x^{k-1} - \tau_k(Ax^{k-1} - b)$ из условия

$$\|z^{t\cdot m}\| \equiv \|(E-\tau_m^{(t)}A)\dots(E-\tau_1^{(t)}A)z^{(t-1)m}\| = \min_{\gamma}\|(E-\gamma_m A)\dots(E-\gamma_1 A)z^{(t-1)m}\|.$$

Решим эту задачу при $t=1$ (т.к. при других t решение задачи будет таким же с точностью до обозначений), определив $\|z\|^2 \equiv \|z\|_D^2 = (Dz,z)$, где $D=D^*>0$.

Т.к.

$$z^m \equiv (E-\tau_m A)\dots(E-\tau_1 A)z^0 = z^0 - q_1(\tau)Az^0 - \dots - q_m(\tau)A^m z^0 =$$
$$= z^0 - \alpha_1(\tau)g_1 - \dots - \alpha_m(\tau)g_m,$$

где $L\{g_1, \dots, g_m\} = L\{Az^0, \dots, A^m z^0\} \equiv L_m$,

то

$$\|z^m\|_D = \min_{\gamma}\|(E-\gamma_m A)\dots(E-\gamma_1 A)z^0\|_D = \min_{\alpha}\|z^0 - \alpha_1 g_1 - \dots - \alpha_m g_m\|_D.$$

Параметры $\alpha_1, \dots, \alpha_m$ удовлетворяют системе уравнений

$$\frac{1}{2}\frac{\partial(Dz^m,z^m)}{\partial\alpha_i} = (D\frac{\partial z^m}{\partial\alpha_i}, z^m) \equiv (Dg_i,\ z^0 - \alpha_1 g_1 - \dots - \alpha_m g_m) = 0,\quad i=1, \dots, m,$$

или

$$\begin{bmatrix} (Dg_1,g_1) & (Dg_1,g_2) & \dots & (Dg_1,g_m) \\ (Dg_2,g_1) & (Dg_2,g_2) & \dots & (Dg_2,g_m) \\ \vdots & \vdots & \dots & \vdots \\ (Dg_m,g_1) & (Dg_m,g_2) & \dots & (Dg_m,g_m) \end{bmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} (Dz^0,g_1) \\ (Dz^0,g_2) \\ \vdots \\ (Dz^0,g_m) \end{pmatrix}.$$

Матрица этой системы – матрица Грамма базиса $\{g_i\}_{i=1}^m$ в L_m.

Для того, чтобы был известен вектор правой части, достаточно выбрать $D=A^*HA$ с любой матрицей $H=H^*>0$.

Если базис $\{g_i\}_{i=1}^m$ является D–ортогональным, т.е. $(Dg_i,g_j)=0,\ i\neq j$, то

$$\alpha_k = \frac{(Dz^0,g_k)}{(Dg_k,g_k)},\qquad x^m = x^0 - \alpha_1 g_1 - \alpha_2 g_2 - \dots - \alpha_m g_m,$$

а вычисление $x^{2m}, x^{3m}, \ldots$ осуществляется аналогично.

Метод сопряженных градиентов

Пусть матрица системы $Ax = b$ симметрична и положительно определена. Построим ($D = A$) A–ортогональный базис $\{g_i\}_{i=1}^m$ в $L_m = L\{Az^0, \ldots, A^m z^0\}$.

1. $g_1 = Az^0 \equiv r^0$ – базис в L_1

$$\Rightarrow \ \alpha_1 = \frac{(r^0, g_1)}{(Ag_1, g_1)}, \quad x^1 = x^0 - \alpha_1 g_1, \ \ r^1 = r^0 - \alpha_1 A g_1.$$

Заметим, что $\left.\begin{matrix} r^1 = Az^0 - \alpha_1 A^2 z^0 \\ (r^1, g_1) = 0 \end{matrix}\right\} \Rightarrow \begin{cases} r^1 \in L_2 \\ r^1 \perp L_1 \end{cases}$

Предположим, что выполнили k шагов: $L_k = \{g_1, \ldots, g_k\}$ и $\begin{cases} r^k \in L_{k+1} \\ r^k \perp L_k \end{cases}$.

Определим $g_{k+1} = r^k - \gamma_1 g_1 - \ldots - \gamma_k g_k$: $(Ag_{k+1}, g_i) = 0, \ i = 1, \ldots, k$,

т.е. $\gamma_i = (Ar^k, g_i)/(Ag_i, g_i)$, т.к. $(Ag_j, g_i) = 0, \ i \ne j$.

Заметим, что $Ag_i \in L_{i+1}$ и $r^k \perp L_{i+1} \subset L_k$

$$\Rightarrow \ (Ar^k, g_i) = (r^k, Ag_i) = 0 \ \forall \ i \le k-1 \Rightarrow$$

$k+1$–шаг.

$$g_{k+1} = r^k - \gamma_k g_k$$

$$\gamma_k = \frac{(r^k, Ag_k)}{(Ag_k, g_k)}$$

$\{g_1, \ldots, g_{k+1}\}$ – базис в L_{k+1}

$$\Rightarrow \ \alpha_{k+1} = \frac{(r^0, g_{k+1})}{(Ag_{k+1}, g_{k+1})} = \frac{(r^1 + \alpha_1 A g_1, g_{k+1})}{(Ag_{k+1}, g_{k+1})} = \ldots$$

$$= \frac{(r^k, g_{k+1})}{(Ag_{k+1}, g_{k+1})},$$

$$x^{k+1} = x^k - \alpha_{k+1} g_{k+1}, \ \ r^{k+1} = r^k - \alpha_{k+1} A g_{k+1}.$$

$r^{k+1} \perp L_{k+1}$,

т.к. $r^k \perp L_k$ & $Ag_k \perp L_k$ & $(r^{k+1}, g_{k+1}) = 0$.

Т.к. $r^k \in L_{k+1} \ \Rightarrow \ r^k = a_1 Az^0 + \ldots + a_{k+1} A^{k+1} z^0$

и $g_{k+1} \in L_{k+1} \ \Rightarrow \ g_{k+1} = b_1 Az^0 + \ldots + b_{k+1} A^{k+1} z^0$, то

$r^{k+1} = c_1 Az^0 + \ldots + c_{k+1} A^{k+1} z^0 + c_{k+2} A^{k+2} z^0 \in L_{k+2}$

$\Rightarrow$ предположения мат. индукции выполнены,

мы построили метод сопряженных градиентов.

Теорема.

Если $A = A^* > 0$, то метод сопряженных градиентов продолжается до получения решения системы $Ax = b$ за $m \le n$ итераций (пока $r^k \neq 0$) и

$$\| z^k \|_A \le \frac{2\gamma^k}{1+\gamma^{2k}} \| z^0 \|_A, \qquad \gamma = \frac{\sqrt{\lambda_{max}} - \sqrt{\lambda_{min}}}{\sqrt{\lambda_{max}} + \sqrt{\lambda_{min}}}.$$

Доказать теорему в качестве упражнения.

Переобуславливатель

Для решения системы $Ax = b$ рассмотрим итерационный метод

$$x^{k+1} = x^k - \tau B^{-1}(Ax^k - b),$$

где матрица B (переобуславливатель) эквивалентна по спектру матрице A с постоянными $\gamma_1 \ge \gamma_0 > 0$:

$$\gamma_0(Bv, v) \le (Av, v) \le \gamma_1(Bv, v) \quad \forall\ v \in R^n.$$

Теорема. $A = A^* > 0\ \ \&\ \ B = B^* > 0\ \Rightarrow\ \gamma_0 \le \lambda(B^{-1}A) \le \gamma_1.$

Док-во. $B^{-1}Av = \lambda v\ \Rightarrow\ [B^{-1/2}AB^{-1/2}](B^{1/2}v) = \lambda(B^{1/2}v)$.

$[B^{-1/2}AB^{-1/2}] = [B^{-1/2}AB^{-1/2}]^*\ \Rightarrow\ \exists\ w = B^{1/2}v$ – вещественный

$$\Rightarrow\ \gamma_0 \le \lambda = \frac{(B^{-1/2}AB^{-1/2}w, w)}{(w, w)} = \frac{(Av, v)}{(Bv, v)} \le \gamma_1.$$

Следствие. $A = A^* > 0\ \ \&\ \ B = B^* > 0$

$$\Rightarrow \begin{cases} \rho(S_\tau) = \rho(E - \tau B^{-1}A) < 1 \quad \forall\ \tau \in (0, 2/\gamma_1) \\ \|S_\tau\|_B = \|B^{1/2}S_\tau B^{-1/2}\|_2 = \rho(B^{1/2}S_\tau B^{-1/2}) = \rho(S_\tau) < 1 \\ \|S_\tau\|_A = \|A^{1/2}S_\tau A^{-1/2}\|_2 = \rho(A^{1/2}S_\tau A^{-1/2}) = \rho(S_\tau) < 1 \end{cases}$$

т.к. $B^{1/2}S_\tau B^{-1/2} = [B^{1/2}S_\tau B^{-1/2}]^*$ и $A^{1/2}S_\tau A^{-1/2} = [A^{1/2}S_\tau A^{-1/2}]^*$.

Теорема. $A = A^* > 0\ \ \&\ \ B > 0.5\tau A$ ($\tau > 0$), тогда $\rho(S_\tau) < 1$.

Док-во.

$$\begin{aligned} \|z^{k+1}\|_A &= (Az^{k+1}, z^{k+1}) = (AS_\tau z^k, S_\tau z^k) = \\ &= (Az^k, z^k) - 2\tau(AB^{-1}Az^k, z^k) + \tau^2(AB^{-1}Az^k, B^{-1}Az^k) = \\ &= (Az^k, z^k) - 2\tau(Bw^k, w^k) + \tau^2(Aw^k, w^k) = \\ &= (Az^k, z^k) - 2\tau([B - 0.5\tau A]w^k, w^k) < \|z^k\|_A \end{aligned}$$

$\Rightarrow$ функционал ошибки строго убывает и, т.к. оператор S_τ непрерывен, то итерационный процесс сходится $\Rightarrow\ \rho(S_\tau) < 1$.

Положительно определенные матрицы

$$A > 0 \text{ в } C^n (\text{или } R^n) \quad \Leftrightarrow \quad (Ax, x) > 0 \quad \forall x \in C^n (R^n),\ x \neq 0$$

Теорема 1. $(Ax, x) = \mathrm{Re}(Ax, x) \ \forall\, x \in C^n \quad \Leftrightarrow \quad A = A^*$.

Теорема 2. $A = A^* \Rightarrow$

$$\begin{cases} A > 0 \text{ в } C^n (R^n) \quad \Leftrightarrow \quad \forall\ \lambda(A) > 0, \\ \lambda_{min}(x, x) \le (Ax, x) \le \lambda_{max}(x, x) \quad \forall\ x \in C^n (R^n). \end{cases}$$

Теорема 3. $A = A^* \Rightarrow \left\{ A > 0 \text{ в } C^n (R^n) \quad \Leftrightarrow \quad \forall \det(A_k) > 0 \right\}$

(т.к. $A = LDL^*$ – разложение Холесского) – это критерий Сильвестра положительной определенности или положительности всех собственных значений симметричной (самосопряженной) матрицы.

Теорема 4. $A > 0 \text{ в } R^n \quad \Leftrightarrow \quad A + A^* > 0 \text{ в } R^n$.

Теорема 5. $A = -A^*$ – веществ. кососимметричная матрица $\Rightarrow A = 0 \text{ в } R^n$.

Теорема 6. $A > 0 \text{ в } R^n \quad \Rightarrow \quad \mathrm{Re}\,\lambda(A) > 0$.

Доказать эти утверждения в качестве упражнений.
Построить пример вещественной несимметричной, но положительно определенной в R^n матрицы.

Лекция 11. Проблема собственных значений

Для матрицы $A=\{a_{ij}\}_{i,j=1}^{n}$ нужно найти числа λ и ненулевые векторы x такие, что $Ax=\lambda x$: λ – собственное значение, x – собственный вектор.

Корректность задачи на собственные значения

Известно, что все собственные значения матрицы A являются корнями характеристического полинома

$$P_n(\lambda)=\det(A-\lambda E)=(-1)^n\lambda^n+p_{n-1}\lambda^{n-1}+\ldots+p_1\lambda+p_0,$$

а коэффициенты $p_0, \ldots, p_{n-1}$ – непрерывные функции элементов матрицы A.

Пусть δA – матрица с “малыми” по величине элементами, $P_{\delta,n}(\lambda)$ – характеристический полином матрицы $A+\delta A$. Следствием непрерывности $\det(A+\delta A)$ как функции элементов матрицы $A+\delta A$ является

Лемма 1. $\lim\limits_{\delta A\to 0} P_{\delta,n}(\lambda)=P_n(\lambda) \quad \forall\, \lambda\in C$.

Лемма 2. В любом круге на комплексной плоскости с центром в точке λ_c и радиуса $\sqrt[n]{|P_n(\lambda_c)|}$ лежит хотя бы один корень полинома $P_n(\lambda)$.

Док-во. Разложим $P_n(\lambda)$ в ряд Тейлора в точке λ_c:

$$P_n(\lambda)=P_n(\lambda_c)+\frac{P_n'(\lambda_c)}{1!}(\lambda-\lambda_c)+\ldots+\frac{P_n^{(n)}(\lambda_c)}{n!}(\lambda-\lambda_c)^n\equiv Q(z),$$

где $z=\lambda-\lambda_c$.

Пусть $z_1, \ldots, z_n$ – корни полинома $Q(z)$, среди которых корень с минимальной абсолютной величиной имеет номер $\min$.

Так как $|P_n(\lambda_c)|=|Q(0)|=|z_1\cdot\ldots\cdot z_n|\ge|z_{\min}|^n=|\lambda_{\min}-\lambda_c|^n$, то $\lambda_{\min}$ (корень полинома $P_n(\lambda)$) лежит в круге радиуса $\sqrt[n]{|P_n(\lambda_c)|}$.

Лемма 3. Если $\lambda_1, \ldots, \lambda_n$ – корни полинома $P_n(\lambda)$, то $\exists$ нумерация корней $\lambda_{\delta,1}, \ldots, \lambda_{\delta,n}$ полинома $P_{\delta,n}(\lambda)$: $\lambda_{\delta,k}\to\lambda_k \;\; \forall\, k$ при $\delta A\to 0$.

Док-во методом матиндукции по степени полинома.

$n=1 \;\Rightarrow\; \lambda_{\delta,1} = p_{\delta,0} \to p_0 = \lambda_1$.

Пусть лемма верна при $n<k$.

$n=k$: из леммы 2 $\Rightarrow\; \exists\, \lambda_{\delta,1}: |\lambda_{\delta,1} - \lambda_1| \le \sqrt[n]{|P_{\delta,n}(\lambda_1)|} \to 0$.

Т.к. $P_n(\lambda) = (\lambda - \lambda_1)R_{n-1}(\lambda)$, $P_{\delta,n}(\lambda) = (\lambda - \lambda_{\delta,1})R_{\delta,n-1}(\lambda)$

и $R_{\delta,n-1}(\lambda) \to R_{n-1}(\lambda)$, то $\lambda_{\delta,2} \to \lambda_2, \ldots, \lambda_{\delta,n} \to \lambda_n$.

Степенной метод вычисления максимального собственного значения матрицы $A = A^* \geq 0$

Идея метода: для заданного вектора x^0 рассмотрим его k–ю итерацию $A^k x^0$,

если $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} < \lambda_n = \rho(A)$ – собственные значения,

$q^{(1)}, q^{(2)}, \ldots, q^{(n)}$ – соответствующие им собственные векторы, то

$$A^k x^0 = \rho^k [\alpha_n q^{(n)} + (\frac{\lambda_1}{\rho})^k \alpha_1 q^{(1)} + \ldots + (\frac{\lambda_{n-1}}{\rho})^k \alpha_{n-1} q^{(n-1)}] \approx \rho^k \alpha_n q^{(n)},$$

$$\frac{\|A^{k+1}x^0\|}{\|A^k x^0\|} \approx \rho, \qquad \frac{1}{\|A^k x^0\|} A^{k+1} x^0 \approx \rho \cdot q^{(n)},$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n$ – коэффициенты (неизвестные!) разложения вектора x^0 по базису $q^{(1)}, q^{(2)}, \ldots, q^{(n)}$.

Итерационный процесс

$$x^0 \neq 0, \qquad x^{k+1} = A \frac{x^k}{\|x^k\|}, \qquad k = 0, 1, \ldots,$$

называется степенным методом вычисления максимального собственного значения матрицы $A = A^* \geq 0$**:**

$$\|x^k\| \to \rho(A), \qquad x^k \to x: \quad Ax = \rho \cdot x,$$

если проекция начального вектора x^0 на линейную оболочку собственных векторов, соответствующих $\rho(A)$, не равна 0.

Док-во. Пусть $0 \leq \lambda_1 \leq \ldots \leq \lambda_r < \lambda_{r+1} = \ldots = \lambda_n = \rho$ – собственные значения,

$q^{(1)}, \ldots, q^{(r)}, q^{(r+1)}, \ldots, q^{(n)}$ – собственные векторы матрицы A, и

$$x^0 = \alpha_1 q^{(1)} + \ldots + \alpha_r q^{(r)} + \alpha_{r+1} q^{(r+1)} + \ldots + \alpha_n q^{(n)} =$$

$$= \alpha_1 q^{(1)} + \ldots + \alpha_r q^{(r)} + y, \qquad y \neq 0.$$

Тогда $A^k x^0 = \rho^k [x + (\lambda_1/\rho)^k \alpha_1 q^{(1)} + \ldots + (\lambda_r/\rho)^k \alpha_r q^{(r)}]$ и,

т.к. $x^k = \dfrac{Ax^{k-1}}{\|x^{k-1}\|} = \dfrac{A^2 x^{k-2}}{\|Ax^{k-2}\|} = \ldots = \dfrac{A^k x^0}{\|A^{k-1}x^0\|}$, $\quad 0 \leq \dfrac{\lambda_1}{\rho} \leq \ldots \leq \dfrac{\lambda_r}{\rho} < 1$,

то $\|x^k\| = \rho \dfrac{\|y + (\lambda_1/\rho)^k \alpha_1 q^{(1)} + \ldots + (\lambda_r/\rho)^k \alpha_r q^{(r)}\|}{\|y + (\lambda_1/\rho)^{k-1} \alpha_1 q^{(1)} + \ldots + (\lambda_r/\rho)^{k-1} \alpha_r q^{(r)}\|} \to \rho$,

$$x^k = \rho \frac{y + (\lambda_1/\rho)^k \alpha_1 q^{(1)} + \ldots + (\lambda_r/\rho)^k \alpha_r q^{(r)}}{\| y + (\lambda_1/\rho)^{k-1} \alpha_1 q^{(1)} + \ldots + (\lambda_r/\rho)^{k-1} \alpha_r q^{(r)} \|} \to x = \rho \frac{y}{\| y \|}.$$

Замечание. Сходимость степенного метода не зависит от выбора в нем векторной нормы, т.к. все нормы в R^n эквивалентны.

Степенной метод вычисления минимального собственного значения матрицы $A = A^* \geq 0$

Задача вычисления минимального собственного значения матрицы $A = A^* \geq 0$ легко сводится к задаче вычисления максимального собственного значения матрицы $\beta \cdot E - A \geq 0$, где $\beta \geq \rho(A)$, так как $\rho(\beta \cdot E - A) = \beta - \lambda_{min}(A)$.

Оценку для $\rho(A)$ легко найти: $\beta = \|A\|_\infty \geq \rho(A)$. Тогда **итерационный процесс**

$$x^0 \neq 0, \qquad x^{k+1} = (\|A\|_\infty \, E - A)\frac{x^k}{\|x^k\|}, \qquad k = 0, 1, \ldots,$$

называется степенным методом вычисления минимального собственного значения матрицы $A = A^* \geq 0$ **:** $(\|A\|_\infty - \|x^k\|) \to \lambda_{min}(A)$**,**

если проекция начального вектора x^0 на линейную оболочку собственных векторов, соответствующих $\lambda_{min}(A)$, не равна 0.

Справедливость этого утверждения является следствием сходимости степенного метода вычисления спектрального радиуса матрицы $B = \|A\|_\infty \, E - A$.

Применение ортогонализации и степенного метода для вычисления очередного собственного значения

Предположим, что собственное значение $\lambda_n = \rho(A)$ и соответствующий ему собственный вектор (какой–то!) $q^{(n)}$ матрицы $A = A^* \geq 0$ мы приближенно (например степенным методом) вычислили: $\tilde{\lambda}_n \approx \lambda_n$, $\tilde{q}^{(n)} \approx q^{(n)}$.

Построим симметричную положительно определенную матрицу $\tilde{A}_{n-1} = \tilde{P}_n A \tilde{P}_n$, где матрица $\tilde{P}_n = E - \tilde{q}^{(n)}[\tilde{q}^{(n)}]^T$ – ортогональный проектор на подпространство $(L\{\tilde{q}^{(n)}\})^\perp$, ортогональное вектору $\tilde{q}^{(n)}$.

Докажите, что спектр матрицы A_{n-1} (т.е. $\tilde{\lambda}_n = \lambda_n$, $\tilde{q}^{(n)} = q^{(n)}$) состоит из собственных значений $\lambda_1 \leq \ldots \leq \lambda_{n-1}$ матрицы A и нуля (вектор $q^{(n)}$ принадлежит ее ядру).

Отсюда следует, что, если $\tilde{q}^{(n)} \to q^{(n)}$ (а степенной метод такую сходимость гарантирует), то $\rho(\tilde{A}_{n-1}) \to \rho(A_{n-1}) = \lambda_{n-1}(A)$.

Следовательно, применяя степенной метод для матрицы $\tilde{A}_{n-1}$, мы получим приближение к $\lambda_{n-1}(A)$ и $q^{(n-1)}$ – очередным собственным значению и вектору матрицы A.
Эту процедуру можно продолжать до тех пор, пока мы не получим все собственные значения.

Лекция 12. Метод деления пополам (бисекций)

Для самосопряженной матрицы $A = \{a_{ij}\}_{i,j=1}^{n}$ имеет место **закон инерции:**
если матрицу A конгруэнтным преобразованием привести к диагональному виду: $D = T^{*}AT$, где $\det T \neq 0$, то от матрицы T (способа преобразования) не зависит

- $\sigma_{-}(A)$ – количество отрицательных элементов,
- $\sigma_{0}(A)$ – количество нулевых элементов,
- $\sigma_{+}(A)$ – количество положительных элементов на диагонали D.

Нам известно (из теоремы и алгоритма LDU–разложения), что если все $\det A_k \neq 0$, то $A = LDL^{*}$, $D = \mathrm{diag}\{d_1, ..., d_n\}$, $\det A_k = d_1 \cdot ... \cdot d_k$.
Следовательно, в этом случае за конечное число действий мы можем определить $\sigma(A) = \{\sigma_{-}(A), \sigma_{0}(A), \sigma_{+}(A)\}$, $\sigma_{0}(A) = 0$.
Матрица $A = A^{*}$ преобразованием подобия ортогональной матрицей Q (конгруэнтным преобразованием) из собственных векторов приводится к диагональному виду $\Lambda = \mathrm{diag}\{\lambda_1, ..., \lambda_n\} = Q^{*}AQ$. Следовательно,

- $\sigma_{-}(A)$ = количеству отрицательных,
- $\sigma_{0}(A)$ = количеству нулевых,
- $\sigma_{+}(A)$ = количеству положительных собственных значений матрицы A,

и, используя LDL^{*}–разложение, мы можем эти числа определить.
Подытожим эти рассуждения в виде следующей леммы.

Лемма 1. Если матрица $A = A^{*}$ и $\det A_k \neq 0 \ \forall \ k$, то количество ее отрицательных собственных значений

$$\sigma_{-}(A) = \text{ЧПЗ}\{1, \det A_1, \det A_2, ..., \det A_n\}$$

– число перемен знака.

Док–во леммы оставляется в виде упражнения.

Идея метода бисекций вычисления $\lambda_j \in Sp(A)$

$\lambda_j \in [a_0, b_0] = [-\|A\|_\infty, \|A\|_\infty]$, т.к. $\rho(A) \le \|A\|_\infty$, т.е. все собственные значения $\lambda_1 \le \lambda_2 \le ... \le \lambda_n$ матрицы $A = A^*$ лежат в этом интервале.

Определим в какой половине интервала $[a_0, b_0]$ лежит λ_j. Для этого вычислим $\sigma_-(A - c_0 E)$ – количество собственных значений меньших $c_0 = (a_0 + b_0)/2$. Если $\sigma_-(A - c_0 E) \ge j$, то $\lambda_j \in [a_0, c_0] \equiv [a_1, b_1]$, иначе $\lambda_j \in [c_0, b_0] \equiv [a_1, b_1]$.

Через k таких шагов получим: $\lambda_j \in [a_k, b_k]$, $b_k - a_k = \|A\|_\infty / 2^{k-1} \to 0$, т.е. мы можем получить оценку искомого собственного числа с любой точностью.

Приведение самосопряженной матрицы к трехдиагональному виду ортогональным преобразованием подобия с помощью матриц вращения

Как и раньше, через $Q_{i,i+k}$ будем обозначать элементарную матрицу вращения, отличающуюся от единичной матрицы двумя диагональными элементами: $(Q_{i,i+k})_{i,i} = \bar{c}_{i,i+k}$, $(Q_{i,i+k})_{i+k,i+k} = c_{i,i+k}$, и двумя внедиагональными элементами: $(Q_{i,i+k})_{i,i+k} = -\bar{s}_{i,i+k}$, $(Q_{i,i+k})_{i+k,i} = s_{i,i+k}$, $|c_{i,i+k}|^2 + |s_{i,i+k}|^2 = 1$.

Выполним и

1–й шаг. Исключение элементов **1–го столбца** матрицы A, начиная с **3–его**, с помощью последовательного умножения на унитарные матрицы $Q_{2,3}, \ldots, Q_{2,n}$:

$$A_1 = (Q_{2,n} \cdot \ldots \cdot Q_{2,3}) A (Q_{2,n} \cdot \ldots \cdot Q_{2,3})^* \equiv Q_1 A Q_1^*.$$

2–й шаг. Исключение элементов **2–го столбца** матрицы A_1, начиная с **4–ого**, с помощью последовательного умножения на унитарные матрицы $Q_{3,4}, \ldots, Q_{3,n}$:

$$A_2 = (Q_{3,n} \cdot \ldots \cdot Q_{3,4}) A_1 (Q_{3,n} \cdot \ldots \cdot Q_{3,4})^* \equiv Q_2 A_1 Q_2^*.$$

… ……………………

k–й шаг. Исключение элементов **k–го столбца** матрицы A_{k-1}, начиная с (**k+2)–ого**, с помощью последовательного умножения на матрицы $Q_{k+1,k+2}, \ldots, Q_{k+1,n}$:

$$A_k = (Q_{k+1,n} \cdot \ldots \cdot Q_{k+1,k+2}) A_{k-1} (Q_{k+1,n} \cdot \ldots \cdot Q_{k+1,k+2})^* \equiv Q_k A_{k-1} Q_k^*.$$

… ……………………

(n–2)–й шаг. Исключение последнего элемента **(n-2)–го столбца** матрицы A_{n-3} с помощью умножения на матрицу $Q_{n-1,n}$:

$$A_{n-2} = (Q_{n-1,n}) A_{n-3} (Q_{n-1,n})^* \equiv Q_{n-2} A_{n-3} Q_{n-2}^*.$$

$$T = A_{n-2} = (Q_{n-2} \cdot \ldots \cdot Q_1) A (Q_{n-2} \cdot \ldots \cdot Q_1)^* = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \bar{\beta}_1 & \alpha_2 & \beta_2 & & & \\ & O & O & O & & \\ & & & \bar{\beta}_{n-2} & \alpha_{n-1} & \beta_{n-1} \\ & & & & \bar{\beta}_{n-1} & \alpha_n \end{bmatrix},$$

$Sp(A) = Sp(T)$.

Если $\beta_k = 0$, то $T = \begin{bmatrix} T_k & 0 \\ 0 & \hat{T}_{n-k} \end{bmatrix} \Rightarrow Sp(T) = Sp(T_k) \cup Sp(\hat{T}_{n-k})$,

т.е. поиск собственных значений самосопряженной матрицы сводится к задаче на собственные значения якобиевых трехдиагональных матриц.

Лемма 2. Самосопряженная матрица подобна трехдиагональной вещественной матрице.

Док-во. Только что мы привели самосопряженную матрицу A к трехдиагональному виду ортогональным преобразованием подобия:

$$T = (Q_{n-2} \cdot \ldots \cdot Q_1) A (Q_{n-2} \cdot \ldots \cdot Q_1)^* = \mathrm{tridiag}\{\bar{\beta}_{i-1}, \alpha_i, \beta_i\}.$$

Определим матрицу $D = \mathrm{diag}\{d_1, \ldots, d_n\}$: (предполагая $\forall\ \beta_i \neq 0$)

$$d_1 = 1,\ d_2 = \beta_1 / |\beta_1|,\ \ldots,\ d_n = \beta_1 / |\beta_1| \cdot \ldots \cdot \beta_{n-1} / |\beta_{n-1}|.$$

Тогда $D^{-1} = D^*$, $B = DTD^{-1} = \mathrm{tridiag}\{|\beta_{i-1}|, \alpha_i, |\beta_i|\}$.

Якобиевы матрицы

Вещественная матрица

$$B = \begin{bmatrix} a_1 & b_1 & & & \\ c_2 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-1} & a_{n-1} & b_{n-1} \\ & & & c_n & a_n \end{bmatrix},\ b_1 \cdot c_2 > 0,\ b_2 \cdot c_3 > 0, \ldots,\ b_{n-1} \cdot c_n > 0,$$

называется **якобиевой** (у нас $c_i = b_{i-1}$).

Лемма 3. Пусть $B = \mathrm{tridiag}\{b_{i-1}, a_i, b_i\}$ – якобиева матрица, тогда

1. $\det B_0 \equiv 1,\ \det B_1 = a_1,$

   $\det B_{i+1} = a_{i+1} \cdot \det B_i - b_i^2 \cdot \det B_{i-1},$

   $i = 1, \ldots, n-1.$

2. если $\det B_i = 0\ (i < n)$, то $\det B_{i-1} \cdot \det B_{i+1} < 0$,

   если $\det B_n = 0$, то $\det B_{n-1} \neq 0$.

Док-во оставляется читателю в качестве упражнения.

Лемма 4. Собственные значения якобиевой матрицы B попарно различные (простые).

Док-во. Т.к. размерность ядра симметричной матрицы $B_\lambda = B - \lambda E$ совпадает с кратностью $\lambda \in Sp(B)$, а из леммы 3 следует, что у вырожденной якобиевой матрицы B_λ минор $[\det B_\lambda]_{n-1} \neq 0$, то $\mathrm{rang} B_\lambda = n-1, \dim \mathrm{Ker} B_\lambda = 1$ и λ простое собственное значение матрицы B.

Теорема. Пусть $B = \mathrm{tridiag}\{b_{i-1}, a_i, b_i\}$ – якобиева матрица, тогда

$$\sigma_-(B) = ЧПЗ\{1,\ \det B_1,\ \det B_2,\ \ldots,\ \det B_n\},$$

если $\det B_k = 0$ приписать знак $\det B_{k-1}$.

Док-во. 1. Если $\det B_k \neq 0\ \forall\ k$, то это лемма 1.

2. Пусть $\exists\ k:\ \det B_k = 0$. Пусть $\mathrm{sign}(\det B_k) = \mathrm{sign}(\det B_{k-1})$.

Определим $\varepsilon_0 = \min\limits_{\lambda \in Sp(B_i),\ \lambda \neq 0,\ i=1,\ldots,n} |\lambda| > 0$ и рассмотрим якобиевы матрицы $B_{\pm\varepsilon} = B \pm \varepsilon E,\ \varepsilon \in (0, \varepsilon_0)$.

Т.к. $\lambda([B_{\pm\varepsilon}]_i) \equiv \lambda(B_i) \pm \varepsilon \neq 0\ \ \forall\ i = 1, \ldots, n$, то

а) $\det[B_{\pm\varepsilon}]_i \neq 0\ \ \forall\ i = 1, \ldots, n,$ (т.к. определитель матрицы равен произведению ее собственных значений),

б) $\mathrm{sign}(\det[B_\varepsilon]_i) = \mathrm{sign}(\det[B_{-\varepsilon}]_i) = \mathrm{sign}(\det B_i)\ \ \forall\ \det B_i \neq 0,$

в) $\mathrm{sign}(\det[B_\varepsilon]_k) \cdot \mathrm{sign}(\det[B_{-\varepsilon}]_k) < 0\ \ \forall\ \det B_k = 0,$

(т.к. из леммы 4 следует, что $\lambda(B_k) = 0$ простое и отрицательных собственных значений у матрицы $[B_{-\varepsilon}]_k$ на одно больше, чем у матрицы $[B_\varepsilon]_k$),

г) $\sigma_-(B_\varepsilon) = \sigma_-(B),\ \ \sigma_-(B_{-\varepsilon}) = \sigma_-(B) + \sigma_0(B)$.

Из леммы 1, а) и г) следует, что

$$\sigma_-(B_\varepsilon) = ЧПЗ\{1,\ \det[B_\varepsilon]_1,\ \det[B_\varepsilon]_2,\ \ldots,\ \det[B_\varepsilon]_n\} = \sigma_-(B),$$

$$\sigma_-(B_{-\varepsilon}) = ЧПЗ\{1,\ \det[B_{-\varepsilon}]_1,\ \det[B_{-\varepsilon}]_2,\ \ldots,\ \det[B_{-\varepsilon}]_n\} = \sigma_-(B) + \sigma_0(B),$$

$$ЧПЗ\{1,\ \det B_1,\ \det B_2,\ \ldots,\ \det B_n\} = ?$$

Подсчитаем эти числа:

Из б) следует, что если $\det B_j \neq 0$ и $\det B_{j+1} \neq 0$, то перемена знака происходит (или нет) одновременно в этих последовательностях.

Случай $\det B_k = 0,\ k \neq n$.

Из леммы 3 имеем $\det B_{k-1} \cdot \det B_{k+1} < 0$, отсюда и из б) следует $\det[B_{\pm\varepsilon}]_{k-1} \cdot \det[B_{\pm\varepsilon}]_{k+1} < 0$ и на участках

$$\det[B_{\pm\varepsilon}]_{k-1},\ \ \det[B_{\pm\varepsilon}]_k,\ \ \det[B_{\pm\varepsilon}]_{k+1}$$

$$\det B_{k-1},\qquad \det B_k,\qquad \det B_{k+1}$$

по одной перемене знака.

Случай $\det B_n = 0,\ \sigma_0(B) = 1$. Отсюда, из в) и г) следует, что

$$\det[B_\varepsilon]_{n-1} \cdot \det[B_\varepsilon]_n > 0,\qquad \det[B_{-\varepsilon}]_{n-1} \cdot \det[B_{-\varepsilon}]_n < 0,$$

$$\mathrm{sign}(\det B_{n-1}) \cdot \mathrm{sign}(\det B_n) > 0.$$

Следовательно, (если $\det B_k = 0$ приписать знак $\det B_{k-1}$) последовательности миноров матриц B_ε и B имеют одинаковые знаки. Теорема доказана.

О вычислении ЧПЗ

Для вычисления $\sigma_-(B) = \text{ЧПЗ}\{1, \det B_1, \det B_2, ..., \det B_n\}$ якобиевой матрицы $B = \text{tridiag}\{b_{i-1}, a_i, b_i\}$ достаточно знать знак каждого $\det B_k$. Если

$$d_0 = 1,\ d_1 = \det B_1,$$

$$d_{i+1} = a_{i+1} \cdot d_i - b_i^2 \cdot d_{i-1},\ i = 1, ..., k-1,$$

$$d_{k-1} := d_{k-1} / |t_k|,\ \ d_k := d_k / |t_k|,$$

$$d_{i+1} = a_{i+1} \cdot d_i - b_i^2 \cdot d_{i-1},\ i = k, ..., n-1,$$

(обычно выбирают $t_k = \max\{|d_{k-1}|, |d_k|\}$), то $\text{sign}\{d_i\} = \text{sign}\{\det B_i\}\ \forall\ i$ и $\sigma_-(B) = \text{ЧПЗ}\{1, d_1, ..., d_n\}$. Нормировку можно применять неоднократно, что позволит избежать быстрого роста (переполнения) чисел $\{d_i\}$.

О вычислении собственного вектора

Лемма 5. Последняя компонента собственного вектора x якобиевой матрицы $B = \text{tridiag}\{b_{i-1}, a_i, b_i\}$ не равна нулю.

Док-во. Пусть $Bx = \lambda x,\ x \neq 0$. Предположим, что $x_n = 0$. Тогда

$$x_{n-1} = -(a_n - \lambda) \cdot x_n / b_{n-1} = 0$$

$$x_{n-i} = -[(a_{n-i+1} - \lambda) \cdot x_{n-i+1} + b_{n-i+1} \cdot x_{n-i+2}] / b_{n-i} = 0,$$

$$i = 2, ..., n-1,$$

$\Rightarrow\ x = 0$ – противоречие, значит $x_n \neq 0$.

Собственный вектор x якобиевой матрицы $B = \text{tridiag}\{b_{i-1}, a_i, b_i\}$ мы можем, положив $x_n = 1$, вычислить по формулам

$$x_{n-1} = -(a_n - \lambda) \cdot x_n / b_{n-1}$$

$$x_{n-2} = -[(a_{n-1} - \lambda) \cdot x_{n-1} + b_{n-1} \cdot x_n] / b_{n-2}$$

....

$$x_1 = -[(a_2 - \lambda) \cdot x_2 + b_2 \cdot x_3] / b_1$$

или решив систему

$$\begin{bmatrix} a_1-\lambda & b_1 & & & \\ b_1 & a_2-\lambda & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-3} & a_{n-2}-\lambda & b_{n-2} \\ & & & b_{n-2} & a_{n-1}-\lambda \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -b_{n-1} \end{pmatrix}$$

с неособенной матрицей.

Лекция 13. Метод вращений (Якоби)

Для самосопряженной матрицы $A = \{a_{ij}\}_{i,j=1}^{n}$ существует унитарная матрица Q (столбцы которой – собственные векторы матрицы A):

$$Q^* A Q = \Lambda \equiv \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$$

$\Rightarrow \ \Phi(Q^* A Q) = \min\limits_{T^*T=E} \Phi(T^* A T)$, где $\Phi(A) = \sum\limits_{i \ne j} |a_{ij}|^2$.

Идея:

построить $\{A_k = Q_k^* A_{k-1} Q_k : \ Q_k^* Q_k = E, \ A_0 = A\}$: $\Phi(A_{k-1}) > \Phi(A_k) \to 0$, тогда на диагональные элементы A_k будут приближать собственные значения, а столбцы $(Q_0 \ldots Q_{k+1})$ – собственные векторы матрицы A.

Определим $S(A) = \sum_{i,j=1}^{n} |a_{ij}|^2$.

Лемма 1. Для любых квадратной матрицы A и унитарной матрицы T имеем

$$S(TA) = S(AT) = S(A).$$

Док-во. Если $A = \left[a_1 \ \ldots \ a_n\right]$, то

$$S(TA) = S([Ta_1 \ \ldots \ Ta_n]) = (Ta_1, Ta_1) + \ldots + (Ta_n, Ta_n) =$$
$$= (a_1, a_1) + \ldots + (a_n, a_n) = S(A).$$

В качестве матриц Q_k будем выбирать элементарные матрицы вращения.

Лемма 2. Пусть $A = A^*$, $\tilde{A} = Q_{ij}^* A Q_{ij} = \{\tilde{a}_{kl}\}$,

где Q_{ij} – элементарная матрица вращения, тогда

$$\Phi(\tilde{A}) = \Phi(A) + \left[|a_{ii}|^2 + |a_{jj}|^2 - |\tilde{a}_{ii}|^2 - |\tilde{a}_{jj}|^2 \right].$$

Док-во. Заметим, что изменились только строки и столбцы с номерами i, j.

Тогда, используя лемму 1, получим

$$S(A) \equiv \Phi(A) + \sum_{\substack{k=1 \\ k \ne i, k \ne j}}^{n} |a_{kk}|^2 + |a_{ii}|^2 + |a_{jj}|^2 \equiv$$

$$\equiv \Phi(\tilde{A}) + \sum_{\substack{k=1 \\ k \ne i, k \ne j}}^{n} |a_{kk}|^2 + |\tilde{a}_{ii}|^2 + |\tilde{a}_{jj}|^2 \equiv S(\tilde{A})$$

откуда следует утверждение леммы.

Выбор вращения

Для простоты будем полагать, что матрица A вещественная. Выразим разность $\Phi(A)-\Phi(\tilde{A}) = |\tilde{a}_{ii}|^2 + |\tilde{a}_{jj}|^2 - |a_{ii}|^2 - |a_{jj}|^2$ через элементы матрицы A.

Лемма 3. Пусть $A = A^*$, $\tilde{A} = Q_{ij}^* A Q_{ij} = \{\tilde{a}_{kl}\}$, где Q_{ij} – элементарная матрица вращения (α – угол вращения), тогда

$$\Phi(A)-\Phi(\tilde{A}) = 2|a_{ij}|^2 - \frac{1}{2}\left[(a_{ii}-a_{jj})\sin 2\alpha + 2a_{ij}\cos 2\alpha\right]^2 =$$

$$= 2|a_{ij}|^2 - 2|\tilde{a}_{ij}|^2,$$

Док-во. Требуемые равенства выводятся из соотношения

$$\begin{bmatrix} \tilde{a}_{ii} & \tilde{a}_{ij} \\ \tilde{a}_{ij} & \tilde{a}_{jj} \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Лемма 4. Пусть $A = A^*$, $\tilde{A} = Q_{ij}^* A Q_{ij} = \{\tilde{a}_{kl}\}$, где Q_{ij} – элементарная матрица вращения такая, что

$$|a_{ij}| = \max_{k\neq l} |a_{kl}|, \quad (a_{ii}-a_{jj})\sin 2\alpha + 2a_{ij}\cos 2\alpha = 0,$$

то $\Phi(\tilde{A}) \le \left[1-2/(n(n-1))\right]\cdot\Phi(A)$.

Док-во. Требуемое неравенство следует из равенства $\Phi(\tilde{A}) = \Phi(A) - 2|a_{ij}|^2$ и оценки $\Phi(A) \le n(n-1)|a_{ij}|^2$.

Следующая лемма обеспечивает существование для леммы 4 матрицы Q_{ij}.

Лемма 5. Решением уравнения $a\cdot\sin 2\alpha + 2b\cdot\cos 2\alpha = 0$ при $b \neq 0$ является угол α такой, что

$$\cos\alpha = \sqrt{0.5(1-a/r)},\quad r = \sqrt{|a|^2 + 4|b|^2},$$

$$\sin\alpha = \frac{2b}{r-a}\cos\alpha.$$

Док-во осуществляется непосредственной проверкой.

Из последних двух лемм следует справедливость теоремы сходимости метода.

Теорема Последовательность матриц $\{A_k\}_{k=0}^{\infty}$ **метода**

вращений:

1. $A_0 = A$, $A_k = Q_k^* A_{k-1} Q_k$, где $Q_k = Q_{i(k),j(k)}$ – матрица вращения, определяемая по формулам лемм 4 и 5, для решения полной проблемы на собственные значения $A = A^*$, сходится к диагональному виду, т.е. $\Phi(A_k) \to 0$, причем

$$\Phi(A_k) \le \left[1 - 2/(n(n-1))\right]^k \cdot \Phi(A).$$

Из теоремы 1 $\Rightarrow\ \forall\, \varepsilon > 0\ \exists\, k$:

$$\tilde A \equiv A_k = (Q_1 ... Q_k)^* A (Q_1 ... Q_k) \equiv \tilde Q^* A \tilde Q, \qquad \Phi(\tilde A) \le \varepsilon^2.$$

Пусть

$$\tilde\Lambda = \mathrm{diag}\,\tilde A = \mathrm{diag}\{\tilde\lambda_1, ..., \tilde\lambda_n\},$$

$$\Lambda = Q^* A Q = \mathrm{diag}\{\lambda_1, ..., \lambda_n\}.$$

Сходимость собственных значений

Лемма 6. $\tilde P(\lambda) \equiv \det(\tilde\Lambda - \lambda E) \ \to\ P(\lambda) \equiv \det(A - \lambda E)$ при $\varepsilon \to 0$.

Док-во. Т.к.

$$\det(\tilde\Lambda - \lambda E) = \det(\tilde Q \tilde\Lambda \tilde Q^* - \lambda E), \qquad \tilde Q \tilde\Lambda \tilde Q^* = A - \tilde Q(\tilde A - \tilde\Lambda)\tilde Q^*,$$

$$S(\tilde Q(\tilde A - \tilde\Lambda)\tilde Q^*) = S(\tilde A - \tilde\Lambda) = \Phi(\tilde A) \le \varepsilon^2 \to 0$$

то $\tilde Q \tilde\Lambda \tilde Q^* \to A \ \Rightarrow\ \det(\tilde\Lambda - \lambda E) \to \det(A - \lambda E)$.

Теорема 2 (оценка приближения собственных значений).

а) $\forall\, \lambda_i \ \ \exists\, \tilde\lambda_{j(i)}: \ \ |\lambda_i - \tilde\lambda_{j(i)}| \le \sqrt{n}\cdot\varepsilon$,

б) $\forall\, \tilde\lambda_j \ \ \exists\, \lambda_{i(j)}: \ \ |\tilde\lambda_j - \tilde\lambda_{i(j)}| \le \sqrt{n}\cdot\varepsilon$.

Док-во. Т.к. $Q\Lambda Q^* = A = \tilde Q \tilde\Lambda \tilde Q^* + \tilde Q(\tilde A - \tilde\Lambda)\tilde Q^*$, то

$$\Lambda(Q^*\tilde Q) - (Q^*\tilde Q)\tilde\Lambda = (Q^*\tilde Q)(\tilde A - \tilde\Lambda) \equiv E, \qquad |\varepsilon_{ij}|^2 \le S(E) \le \varepsilon^2.$$

$\Rightarrow\ \lambda_i \cdot r_{ij} - r_{ij}\cdot\tilde\lambda_j = \varepsilon_{ij}$, где $\{r_{ij}\} = R = Q^*\tilde Q$ – ортогональная м-ца.

а) $\forall\, i\ \exists\, j(i): |r_{ij(i)}|^2 = \max\limits_k |r_{ik}|^2 \ge 1/n$, т.к. $|r_{i1}|^2 + ... + |r_{in}|^2 = 1$.

$$\Rightarrow\ |\lambda_i - \tilde\lambda_{j(i)}| = |\varepsilon_{ij} / r_{ij(i)}| \le \sqrt{n}\cdot\varepsilon.$$

б) доказывается аналогично.

Сходимость собственных векторов

Будем предполагать, что $\lambda_1 < \lambda_2 < ... < \lambda_n$ и $\tilde\lambda_1 \le \tilde\lambda_2 \le ... \le \tilde\lambda_n$ (этого всегда можно добиться, переставив столбцы матриц Q и $\tilde Q$).

Лемма 7. Если $\lambda_1 < \lambda_2 < ... < \lambda_n$, $\tilde\lambda_1 \le \tilde\lambda_2 \le ... \le \tilde\lambda_n$, $\sqrt{n}\cdot\varepsilon < 0.5\cdot a$,

$$a = \min_{i \ne j} |\lambda_i - \lambda_j|,$$

то $|\lambda_i - \tilde\lambda_i| \le \sqrt{n}\cdot\varepsilon, \qquad |\lambda_i - \tilde\lambda_j| > 0.5\cdot a \ \ \forall\, i \ne j$.

Док-во оставляется в качестве упражнения.

Т.к. собственные векторы $Q=[q_1 \dots q_n]$ матрицы A определяются с точностью до их направления, будем считать, что $(q_i,\tilde{q}_i)\ge 0$ ($[\tilde{q}_1 \dots \tilde{q}_n]=\tilde{Q}$ – приближения к собственным векторам матрицы A), т.е. диагональные элементы матрицы $R=Q^*\tilde{Q}$ неотрицательны.

Теорема 3 (оценка приближения собственных векторов).

В условиях леммы 7 $S(Q-\tilde{Q})\le \dfrac{8}{a^2}\varepsilon^2$.

Док-во. Т.к. $S(Q-\tilde{Q})=S(E-Q^*\tilde{Q})\equiv S(E-R)$ и из доказательства теоремы 2 ($\mathsf{E}=\Lambda R-R\tilde{\Lambda}=R(\tilde{\Lambda}-\tilde{\Lambda})$) и леммы 7 следует, что

$$|r_{ij}|=\frac{|\varepsilon_{ij}|}{|\lambda_i-\tilde{\lambda}_j|}<\frac{|\varepsilon_{ij}|}{0.5\cdot a}\quad \forall\, i\ne j,$$

то

$$\Phi(E-R)<\frac{4}{a^2}S(\mathsf{E})\le\frac{4}{a^2}\varepsilon^2.$$

Осталось оценить $\sum_{i=1}^{n}(1-r_{ii})^2=\sum_{i=1}^{n}\left(1-\sqrt{1-\sum_{\substack{j=1\\ j\ne i}}^{n}|r_{ij}|^2}\right)^2$

(здесь мы воспользовались условием $r_{ii}\ge 0$).

Т.к. $(1-x)^2\le 1-x\ \ \forall\, x\in[0,1]$, то

$$\sum_{i=1}^{n}\left(1-\sqrt{1-\sum_{\substack{j=1\\ j\ne i}}^{n}|r_{ij}|^2}\right)^2\le\sum_{i=1}^{n}\left(1-\sqrt{1-\sum_{\substack{j=1\\ j\ne i}}^{n}|r_{ij}|^2}\right)=$$

$$=\sum_{i=1}^{n}\frac{\sum_{\substack{j=1\\ j\ne i}}^{n}|r_{ij}|^2}{1+\sqrt{1-\sum_{\substack{j=1\\ j\ne i}}^{n}|r_{ij}|^2}}\le\Phi(R)=\Phi(E-R)\le\frac{4}{a^2}\varepsilon^2.$$

Подводя итог, имеем

$$S(Q-\tilde{Q})=S(E-R)=\Phi(E-R)+\sum_{i=1}^{n}(1-r_{ii})^2\le\frac{8}{a^2}\varepsilon^2.$$

Литература

1. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. - М.: Л.: Физматгиз, 1963.
2. Коновалов А.Н. Введение в вычислительные методы линейной алгебры.- Новосибирск: ВО "Наука", Сибирская издательская фирма, 1993.
3. Воеводин В.В. Вычислительнные основы линейной алгебры. - М.: Наука, 1977.
4. Годунов С.К. Решение систем линейных уравнений. - Новосибирск: Наука, Сиб. отд-ние, 1980.
5. Бахвалов Н.С. Численные методы. - М.: Наука, 1975.
6. Самарский А.А., Гулин А.В. Численные методы. - М.: Наука, 1989.