ГЛАВА V

ЧИСЛЕННЫЕ МЕТОДЫ ЛИНЕЙНОЙ АЛГЕБРЫ

В нашем лекционном курсе мы остановимся на двух центральных проблемах численных методов линейной алгебры (ЛА). Это вопросы

I. Решение систем линейных алгебраических уравнений с невырожденной (квадратной) матрицей.

II. Нахождение собственных значений и собственных векторов для квадратных матриц — *алгебраическая проблема собственных значений.*

I. Решение систем линейных алгебраических уравнений

§1. Основные вычислительные задачи решения систем линейных алгебраических уравнений (СЛАУ)

1.1 Постановка задачи

В качестве основных вычислительных задач рассмотрим следующие три задачи:
1) Решение СЛАУ

$$Ax = f \qquad (\alpha)$$

с квадратной невырожденной матрицей $A_{n\times n}$, $\quad detA \neq 0; x = \overrightarrow{x} = (x_1, ..., x_n)^T; f = \overrightarrow{f} \in R^n$; матрица A определяет отображение $A : R^n \Rightarrow R^n$.
2) Вычисление определителя матрицы $A = \|a^i_j\|^n_n = \|a_{ij}\|_{n\times n}$

$$\Delta = detA. \qquad (\beta)$$

3) Нахождение обратной матрицы A^{-1} для невырожденной квадратной матрицы A:

$$A^{-1}A = AA^{-1} = E. \qquad (\gamma)$$

Естественно, что приведенный перечень вопросов не охватывает все, и в том числе наиболее интересные практически, проблемы, связанные с решением СЛАУ.

Мы специально ограничиваемся рассмотрением невырожденной матрицы А, $det A \neq 0$, чтобы не привлекать другого подхода к понятию решения в случае, когда оно отсутствует, либо не является единственным [*1)].

1.2 Формальное решение. Устойчивость

Формальное решение задачи (α) строится по известным формулам Крамера

$$x = A^{-1}f; \quad x_i = \frac{\Delta_i}{\Delta}.$$

Формальное решение устойчиво, т.е. непрерывно зависит от входных данных A и $\overrightarrow{f}$. Действительно, варьируя $x = A^{-1}\vec{f}$, найдем [*2)]

$$\delta x = \delta(A^{-1}f) = \delta A^{-1}f + A^{-1}\delta f = \Big| \quad \delta A^{-1} = -A^{-1}\delta A\, A^{-1}\Big| \quad \Rightarrow$$
$$\delta x = -A^{-1}\delta A\, A^{-1}f + A^{-1}\delta f = A^{-1}(\delta f - \delta A\, x). \qquad (*)$$

Таким образом $\|\delta x\| \to 0$ при $\|\delta f\|$ и $\|\delta A\| \to 0$.

1.3 Нормы

Напомним основные, используемые в R^n нормы

1) *Норма вектора* $\vec{x}$. Запишем разложение вектора $\vec{x}$ по базису $e = \{\vec{e}_i\}_n$:

$$\vec{x} = e\,X = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \cdot (x_1, x_2, \ldots, x_n)^T,$$

Базисные векторы образуют строку e, а координаты вектора $\vec{x}$ — столбец X.

а) *евклидова норма* вектора

$$\|x\| = \sqrt{(\vec{x}, \vec{x})} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2};$$

б) l_p-норма (при $p = 2$ — норма Гильберта-Шмидта)

$$\|x\|_p = \left(\frac{1}{n}\sum_{i=1}^{n} |x_i|^p\right)^{1/p},$$

(для конечномерного случая $1/n$ можно перед суммой опустить).

в) c-норма (*равномерная* или *чебышевская* норма вектора x)

$$\|x\|_c = \sup_i |x_i| = \max_i |x_i| = \|x\|_\infty.$$

[*1)] при численных расчётах грань $det A \neq 0$ и $det A = 0$ достаточно условна

[*2)] учтём, что $\delta E \equiv 0 \equiv \delta(A^{-1}A) = \delta A^{-1}A + A^{-1}\delta A$

В R^n имеют место соотношения

$$\|x\|_1 \leq \|x\|_2 \leq \|x\|_c \leq \sqrt{n}\,\|x\|_2 \leq n\,\|x\|_1,$$

т.е. в R^n все эти нормы *эквивалентны* и сходимость в любой из них влечёт сходимость в остальных нормах.
Проверим, например:

$$\|x\|_c \leq \sqrt{n}\,\|x\|_2.$$

Имеем

$$\sqrt{n}\,\|x\|_2 = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}|x_i|^2\right)^{1/2} = \left(\sum_{i=1}^{n}|x_i|^2\right)^{1/2} =$$

$$= \left(\max_i |x_i|^2 + \sum_{i\neq i_{max}}|x_i|^2\right)^{1/2} \geq \max_i |x_i| = \|x_c\| \quad \blacksquare$$

Задача. *Доказать эквивалентность введенных норм.*
2) *Норма матрицы A.* Норма матрицы A, согласованная с нормой вектора $\overrightarrow{x}$ определяется следующим образом:

$$\|A\| = \sup_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|} = \left|\begin{array}{l}\text{в силу линейности}\\ \text{преобразования } A\\ \text{и свойств нормы}\end{array}\right| = \sup_{\|x\|\neq 0}\frac{\|A\frac{x}{\|x\|}\|x\|\|}{\|x\|} = \sup_{\|x\|=1}\|Ax\|.$$

Поскольку норма вектора — непрерывная функция его координат x_i, то на замкнутом, ограниченном множестве $\|x\| = 1$ она достигает своего наибольшего значения

$$\|A\| = \max_{\|x\|=1}\|Ax\| = \max_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|}.$$

Отсюда получаем, что

$$\|Ax\| \leq \|A\|\cdot\|x\|.$$

Это условие согласования норм $\|x\|$ и $\|A\|$. Легко проверить, что введённая таким образом норма матрицы удовлетворяет неравенству

$$\|A+B\| \leq \|A\| + \|B\|; \quad \|AB\| \leq \|A\|\,\|B\|,$$

что и делает именно норму матрицы столь удобной в оценках.

Для квадратных матриц $A_{n\times n}$ наиболее употребительны следующие нормы:

$$\begin{array}{lclclcl}
\|A\|_c & = & \max\limits_i\left(\sum\limits_{j=1}^{n}|a_{ij}|\right) & & \|A\|_1 & = & \max\limits_j\left(\sum\limits_{i=1}^{n}|a_{ij}|\right)\\
\|A\|_M & = & n\cdot\max\limits_{i,j}|a_{ij}| & & \|A\|_E & = & \left(\sum\limits_i\sum\limits_j|a_{ij}|^2\right)^{1/2}\\
\|A\|_S & = & \sqrt{\max\limits_i \mu_i} & = & \max\limits_i \nu_i & &
\end{array}$$

(где μ_i — собственные значения симметричной самосопряжённой матрицы (A^*A), $\nu_i = \sqrt{\mu_i}$). Первые две нормы не имеют специальных названий, $\|A\|_M$ называется *максимальной*, $\|A\|_E$ — *сферической* или *евклидовой*, $\|A\|_S$ — *спектральной.* Эти нормы согласованы с нормами векторов в R^n норма $\|A\|_1$ согласована с нормой $\|x\|_1$, спектральная норма и сферическая — с $\|x\|_2$, максимальная норма $\|A\|_M$ — со всеми рассмотренными нормами векторов в R^n.

$$\|A\|_c \rightsquigarrow \|x\|_c; \quad \|A\|_1 \rightsquigarrow \|x\|_1; \quad \|A\|_E, \|A\|_S \rightsquigarrow \|x\|_2; \quad \|A\|_M \rightsquigarrow \|x\|_c, \|x\|_1, \|x\|_2.$$

Покажем согласованность $\|A\|_c$ и $\|x\|_c$:

$$\|A\,x\|_c = \max_i \left|\sum_j a_{ij}x_j\right| \le \max_i \sum_j |a_{ij}||x_j| \le \max_i \left(\underbrace{\max_j |x_j|}_{\|x\|_c} \sum_j |a_{ij}|\right) = \|x\|_c\,\|A\|_c \quad \blacksquare$$

Особенно часто используеться евклидова норма $\|A\|_E$, поскольку она допускает сравнительно простую и наглядную интерпретацию и, что особенно важно, широкий класс геометрических преобразований сохраняет эту норму — это ортогональные преобразования:

$$U^TU = UU^T = E; \quad U^T = U^{-1}; \quad x \to Ux -$$ вращения и отражения относительно координатных плоскостей

Задача. *Показать указанную согласованность норм.*

1.4 Обусловленность матрицы. Погрешности

Вернемся к анализу формулы (*) вариации решения x

$$\delta x = A^{-1}(\delta f - \delta A x).$$

1) Пусть матрица A известна точно ($\delta A = 0$) и погрешность решения связана лишь с погрешностью δf правой части, тогда

$$\delta x = A^{-1}\delta f \Rightarrow \|\delta x\| \le \|A^{-1}\| \cdot \|\delta f\|.$$

Из

$$f = Ax \Rightarrow \|f\| \le \|A\| \cdot \|x\|.$$

Перемножая полученные неравенства, найдем

$$\|\delta x\|\|f\| \le \|A\| \cdot \|A^{-1}\| \cdot \|\delta f\| \cdot \|x\|$$

или

$$\frac{\|\delta x\|}{\|x\|} \le \|A\| \cdot \|A^{-1}\| \frac{\|\delta f\|}{\|f\|} = Cond A\,\frac{\|\delta f\|}{\|f\|}.$$

$Cond A = \|A\|\|A^{-1}\|$ — число обусловленности матрицы A. $Cond A \ge 1$ всегда (в любой норме[*1)]). Т.о. хорошо обусловлены матрицы с малым $Cond A$, при этом относительная погрешность решения мала.

2) Пусть известно возмущение $\|\delta A\|$ матрицы A при условии, что правая часть f

[*1)] поскольку $E = A \cdot A^{-1} \Leftrightarrow \|E\| \le \|A\| \cdot \|A^{-1}\|$

задана точно. Тогда

$$\delta x = -A^{-1}\delta A\, x \Rightarrow \|\delta x\| \leq \|A^{-1}\| \cdot \|\delta A\| \cdot \|x\|$$

или

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\delta A\|}{\|A\|} = CondA \frac{\|\delta A\|}{\|A\|}.$$

В общем случае, для малых возмущений матрицы A, когда $\|\delta A\| \ll \frac{1}{\|A^{-1}\|}$, и можно гарантировать существование обратной к $(A + \delta A)$ матрицы $(A + \delta A)^{-1}$, и получить оценку ее нормы через $\|\delta A\|, \|A\|, \|A^{-1}\|$ получим

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{CondA}{\underbrace{1 - CondA \frac{\|\delta A\|}{\|A\|}}_{\|\delta A\| \cdot \|A^{-1}\| \ll 1}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta f\|}{\|f\|} \right). \tag{2}$$

Замечания:

1) Анализ погрешностей по формуле (2) можно применять к случаю нахождения погрешности округления, как соответствующих возмущений A и f. Относительно громоздкими выкладками можно получить оценку относительной погрешности через машинное эпсилон ε_M для

$$\frac{\|\delta A\|}{\|A\|} = O(n\, 2^{-t}) = O(n\varepsilon_M)$$

здесь t–разрядность ЭВМ. Аналогичная оценка имеет место и для $\|\delta f\|/\|f\|$. Тогда, опираясь на (2), получим

$$\frac{\|\delta x\|}{\|x\|} = O(CondA\, n\, \varepsilon_M). \tag{3}$$

2) Априорное нахождение $CondA$ требует построения обратной к A матрицы и нахождения её нормы — это самостоятельная и весьма трудоёмкая задача.

Перейдем непосредственно к рассмотрению алгоритмов построения решения задачи (1) $Ax = f$.

§2. Метод Гаусса последовательного исключения неизвестных. LU - разложение

Численные методы решения СЛАУ (1) делятся на две большие группы: так называемые *прямые* и *итерационные* методы решения. *Прямые методы* дают решения СЛАУ за *конечное* число шагов. Они просты с алгебраической стороны и наиболее универсальны. Их основным недостатком является ограничение на порядок $n \sim 200$ матрицы системы уравнений (1), что связано с особенностью организации памяти доступных ЭВМ.

Итерационные методы используются, в основном, для решения СЛАУ специального (разреженного, слабозаполненного) вида с числом неизвестных $10^3 \div 10^5$ и более.

2.1 Формулы метода Гаусса

Одним из основных прямых методов решения СЛАУ является метод последовательного исключения неизвестных Гаусса. Он основан на возможности приведения исходной системы к эквивалентному представлению, когда относительно x решается задача с верхне-треугольной матрицей с единичной диагональю:

$$Ax = f \Leftrightarrow Ux = y$$

где $u_{ii} = 1, u_{ij} = 0$ при $j < i$.
Получение этой системы,т.е. построение матрицы U и вектора y составляют, так называемый, *прямой ход* метода исключения Гаусса. Дальнейшее решение системы $Ux = y$ — *обратный ход* метода исключения.

а) *Прямой ход исключения.* Опишем последовательно как он выполняется.
1-ый шаг. Пусть $a_{11} \neq 0$. Тогда, деля первое уравнение на a_{11}, получим

$$x_1 + u_{12}x_2 + ... + u_{1n}x_n = y_1$$

$$u_{1k} = \frac{a_{1k}}{a_{11}}; \quad k = 2, ...n \quad y_1 = \frac{f_1}{a_{11}}$$

Комбинируя полученное уравнение с остальными уравнениями системы (1), исключая в них пееменную x_1, найдем:

$$\begin{array}{c} 0 \cdot x_1 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + ... + a_{2n}^{(1)}x_n = f_2^{(1)} \\ 0 \cdot x_1 + a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + ... + a_{3n}^{(1)}x_n = f_3^{(1)} \\ \cdots \\ 0 \cdot x_1 + a_{n2}^{(1)}x_2 + a_{n3}^{(1)}x_3 + ... + a_{nn}^{(1)}x_n = f_n^{(1)} \end{array}$$

Для оставшихся уравнений (без x_1), повторим описанную процедуру:
S-ый шаг. После $(s-1)$ шагов исключения часть переменных исключена

$$\begin{array}{rcl} x_1 + u_{12}x_2 + ... + u_{12}x_n & = & y_1 \\ x_2 + u_{23}x_3 + ... + u_{2n}x_n & = & y_2 \\ & \cdots & \\ x_{s-1} + u_{s-1,s}x_s + ... + u_{s-1,n}x_n & = & y_{s-1} \end{array}$$

и мы получаем систему:

$$\begin{cases} a_{s,s}^{(s-1)}x_s + a_{s,s+1}^{(s-1)}x_{s+1} + ... + a_{s,n}^{(s-1)} = f_s^{(s-1)} \\ .... \\ a_{n,s}^{(s-1)}x_s + ... + a_{n,n}^{(s-1)}x_n = f_n^{(s)} \end{cases} \qquad (*)$$

Положим $a_{s,s}^{(s-1)} \neq 0$. Делим s-ое уравнение на $a_{s,s}^{(s-1)}$ и находим

$$x_s + u_{s,s+1}x_{s+1} + \cdots + u_{s,n}x_n = y_s \equiv \frac{f_s^{(s-1)}}{a_{s,s}^{(s-1)}}; \quad u_{s,j} = \frac{a_{s,j}^{(s-1)}}{a_{s,s}^{(s-1)}}.$$

Используем полученное уравнение. После умножения его на $a_{i,s}^{(s-1)}, \quad i = s+1, ..., n$ и вычитания из i-го уравнения исключаем x_s в системе (*) из остальных уравнений.

Получим

$$\begin{array}{rcl} x_s \quad + \quad u_{s,s+1}x_{s+1}+\dots+u_{s,n}x_n & = & y_s \\ a^{(s)}_{s+1,s+1}x_{s+1}+\dots+a^{(s)}_{s+1,u}x_n & = & f^{(s)}_{s+1} \\ \dots & & \\ a^{(s)}_{n,s+1}x_{s+1}+\dots+a^{(s)}_{n,n}x_n & = & f^{(s)}_n, \end{array}$$

где

$$\begin{array}{rcll} a^{(s)}_{ij} & = & a^{(s-1)}_{i,j}-a^{(s-1)}_{i,s}u_{s,j} & ; \quad i,j=\overline{s+1,n} \\ f^{(s)}_i & = & f^{(s-1)}_i-a^{(s-1)}_{i,s}y_s & ; \quad i,j=\overline{s+1,n} \end{array}$$

Таким образом прямой ход в методе Гаусса

$$Ax=F \Leftrightarrow Ux=y$$

осуществляется по формулам:

$$\begin{array}{ll} u_{s,j}=\frac{a^{(s-1)}_{s,j}}{a^{(s-1)}_{s,s}}, & s=1,\dots,n; \quad j=s+1,\dots,n \\ a^{(s)}_{i,j}=a^{(s-1)}_{i,j}-a^{(s-1)}_{i,s}u_{s,j}, & i,j=s+1,\dots,n; \quad s=1,\dots,n-1 \end{array} \tag{4}$$

для матрицы и для правой части по формулам:

$$\begin{array}{ll} y_s=\frac{f^{(s-1)}_s}{a^{(s-1)}_{s,s}}; & s=1,\dots,n; \quad f^{(0)}_s=f_s \\ f^{(s)}_i=f^{(s-1)}_i-a^{(s-1)}_{i,s}y_s; & i=s+1,\dots,n; \quad s=1,\dots,n-1 \end{array} \tag{5}$$

б)*Обратный ход* метода Гаусса. Теперь решаем систему $Ux=y$ с верхне-треугольной матрицей, причём $u_{ii}=1$:

$$\begin{cases} x_n=y_n \\ x_i=y_i-\sum\limits_{j=i+1}^{n}u_jx_j, \quad i=\overline{n-1,1} \end{cases} \tag{6}$$

Замечание. Формулы (4),(5) и (6) решают задачу (1). Число наиболее продолжительных арифметических действий — умножений-делений порядка $O(\frac{n^3}{3})$ и столько же сложений-вычитаний. Таким образом $O(\frac{2n^3}{3})$ арифметических действий необходимо для осуществления метода последовательного исключения неизвестных.

2.2 LU - разложение невырожденной матрицы

При реализации метода Гаусса на каждои шаге исключения мы полагали $a^{(s-1)}_{s,s}\neq 0$. Формулы (4) и (5) можно интерпритировать так, будто имеет место представление $f=Ly$ с нижней треугольной матрицей L

$$Ux=y=L^{-1}f \Leftrightarrow LUx=f \quad \text{т.е. } A=LU.$$

Это не случайно, однако само разложение мы получили по-другому (заодно и ответим на вопрос обоснования метода Гаусса).