ЛЕКЦИИ ПО ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКЕ

МГТУ им. Н.Э. Баумана
Москва

ЛЕКЦИЯ 3

Устойчивость вычислительной задачи по входным данным

X – множество допустимых входных данных
Y – множество допустимых решений

Корректность вычислительной задачи по Адамару. Вычислительная задача называется **корректной (по Адамару)**, если выполнены следующие 3-и условия.

• Решение вычислительной задачи $y \in Y$ существует при любых входных данных $x \in X$.

• Решение вычислительной задачи единственно.

• Решение устойчиво по отношению к малым возмущениям входных данных (непрерывно зависит от входных данных).

Если хотя бы одно из условий не выполнено, то вычислительная задача называется **некорректной**.

Определение. Решение y вычислительной задачи называется **устойчивым по входным данным (абсолютно устойчивым)** x, если

$$\forall \varepsilon > 0, \quad \exists \delta(\varepsilon) > 0 : \quad \forall x^*(\Delta(x^*) < \delta(\varepsilon) \quad \Longrightarrow \quad \Delta(y^*) < \varepsilon).$$

Определение. Решение y вычислительной задачи называется **неустойчивым** x, если

$$\exists \varepsilon > 0, \quad \forall \delta > 0 : \quad \exists x^*(\Delta(x^*) < \delta(\varepsilon) \quad \Longrightarrow \quad \Delta(y^*) \geqslant \varepsilon).$$

Определение. Решение называется **относительно устойчивым**, если

$$\forall \varepsilon > 0, \quad \exists \delta(\varepsilon) > 0 : \quad \forall x^*(\delta(x^*) < \delta(\varepsilon) \quad \Longrightarrow \quad \delta(y^*) < \varepsilon).$$

Устойчивость задачи вычисления определенного интеграла $I = \int\limits_a^b f(x)\,dx$

$f^*(x)$ – приближенно заданная интегрируемая функция.

$$I^* = \int\limits_a^b f^*(x)\,dx$$

$\Delta(f^*) = \sup\limits_{x\in[a,b]} |f(x) - f^*(x)|$ - абсолютная погрешность функции $f^*(x)$.

$$\Delta(I^*) = |I - I^*| = \left|\int\limits_a^b (f(x) - f^*(x))\,dx\right| \leqslant (b-a)\Delta(f^*)\,.$$

Если потребовать: $\delta(\varepsilon) = \dfrac{\varepsilon}{b-a}$, тогда

$$\forall\varepsilon > 0, \quad \exists\delta(\varepsilon) > 0: \quad \forall f^*(x)\Big(\Delta(f^*) < \delta(\varepsilon) \quad \Longrightarrow \quad \Delta(I^*) < \varepsilon\Big).$$

Следовательно, задача вычисления определенного интеграла является **устойчивой**.

Устойчивость задачи вычисления производной

$f^*(x)$ – приближенно заданная на отрезке $[a, b]$ непрерывно дифференцируемая функция.

$$u^*(x) = f^{*\prime}(x)$$

Зададим абсолютные погрешности:

$$\Delta(f^*) = \max_{x\in[a,b]} |f(x) - f^*(x)|, \quad \Delta(u^*) = \max_{x\in[a,b]} |u(x) - u^*(x)|$$

Возьмем, например, функцию $f^*(x) = f(x) + \alpha^2 \cos\left(\dfrac{x}{\alpha^5}\right)$, где $0 < \alpha \ll 1$.

Тогда $u^*(x) = u(x) - \alpha^{-3} \sin\left(\dfrac{x}{\alpha^5}\right)$ и $\Delta(f^*) = \alpha^2$, $\Delta(u^*) = \alpha^{-3}$.

Таким образом, сколь угодно малой погрешности функции f отвечает сколь угодно большая погрешность производной f'.

Задача вычисления производной приближенно заданной функции **не является устойчивой**.

Замечание. Одна и таже задача может оказаться как устойчивой, так и неустойчивой в зависимости от выбора способа вычисления абсолютных погрешностей $\Delta(x^*)$ и $\Delta(y^*)$.

Обусловленность вычислительной задачи.
Абсолютное и относительное число обусловленности.

На практике: точность входных данных ограничена.

Вопрос: как повлияют малые, но конечные погрешности входных данных на решение?

Определение. Чувствительность решения вычислительной задачи к малым погрешностям входных данных – **обусловленность вычислительной задачи**.

Определение. Задача называется **хорошо обусловленной**, если малым погрешностям входных данных отвечают малые погрешности решения и **плохо обусловленной**, если происходят сильные изменения решения.

Определение. **Число обусловленности** (количественная мера степени обусловленности вычислительной задачи) – коэффициент возможного возрастания погрешностей в решении по отношению к вызвавшим их погрешностям входных данных.

Пусть:

$$\Delta(y^*) \leqslant \nu_\Delta \Delta(x^*), \qquad \delta(y^*) \leqslant \nu_\delta \delta(x^*).$$

Определение. величина ν_Δ – **абсолютное число обусловленности**, а ν_δ – **относительное число обусловленности**.

Замечание. В неравенства вместо Δ и δ могут быть и их границы $\overline{\Delta}$ и $\overline{\delta}$.

Для плохо обусловленнной задачи $\nu \gg 1$.

Обусловленность задачи вычисления значения функции одной переменной. Используя формулы для вычисления погрешности функции одной переменной получаем:

$$\nu_\Delta \approx |f'(x)|, \qquad \nu_\delta \approx \frac{|f'(x)|\,|x|}{|f(x)|}.$$

Обусловленность задачи вычисления интеграла. Из выше приведенного примера следует, что $\nu_\Delta = b - a$.

Положим

$$\delta(f^*) = \frac{\sup\limits_{x\in[a,b]} |f^*(x) - f(x)|}{|f(x)|}, \quad \text{где} \quad f(x) \neq 0.$$

Следовательно,

$$\Delta(I^*) \leqslant \int\limits_a^b |f^*(x) - f(x)| dx \leqslant \int\limits_a^b |f(x)| dx \, \delta(f^*).$$

Получили оценку

$$\delta(I^*) \leqslant \nu_\delta \delta(f^*), \quad \text{где} \quad \nu_\delta = \frac{\int\limits_a^b |f(x)| dx}{|\int\limits_a^b f(x) dx|}.$$

Вывод. Если подынтегральная функция знакопостоянна, то $\nu_\delta = 1$ и задача хорошо обусловлена, если же функция $f(x)$ на $[a, b]$ принимает значения разных знаков, то $\nu_\delta > 1$.

ЛЕКЦИЯ 4

Численное решение задач линейной алгебры

В линейной алгебре выделяют 4-ре основные задачи:

- решение систем линейных алгебраических уравнений;
- вычисление определителей;
- нахождение обратных матриц;
- нахождение собственных значений и собственных векторов.

Рассмотрим СЛАУ

$$Ax = f,$$

где A – матрица $m \times m$,
$x = (x_1, x_2, x_3, \ldots x_m)^T$ – искомый вектор,
$f = (f_1, f_2, f_3, \ldots f_m)^T$ – заданный вектор.

Некоторые сведения из линейной алгебры.

Определение. Функцию, заданную в линейном пространстве H, которая для $\forall x \in H$ ставит в соответствие число $||x||$, называют **нормой**, если она удовлетворяет следующим **аксиомам нормы**:

- $||x|| \geqslant 0, \quad \forall x \in H$ и $||x|| = 0 \quad \Longrightarrow \quad x = 0;$
- $||\alpha x|| = |\alpha| ||x||$, где $\alpha \in \mathbb{R}$
- $||x + y|| \leqslant ||x|| + ||y||$.

Наиболее часто применяемые нормы:

$$||x||_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}, \quad ||x||_\infty = \max_{1 \leqslant i \leqslant m} |x_i|,$$

где частными случаями нормы $||x||_p$ являются нормы:

$$||x||_1 = \sum_{i=1}^{m} |x_i| \ - \text{ октаэдрическая норма;}$$

$$||x||_2 = \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2} \ - \text{ евклидова (или сферическая) норма.}$$

Абсолютная и относительная погрешности векторов

В качестве меры степени близости векторов $||x||$ и $||x^*||$ введем **абсолютную и относительную погрешности** вектора $||x^*||$

$$\Delta(x^*) = ||x - x^*||, \quad \delta(x^*) = \frac{||x - x^*||}{||x||}.$$

Пусть $\left\{x^{(n)}\right\}_{n=1}^{\infty}$ - последовательность векторов $x^{(n)} = \left(x_1^{(n)}, x_2^{(n)}, \ldots, x_m^{(n)}\right)$.

Говорят, что последовательность векторов $x^{(n)}$ **сходится в вектору** x при $n \to \infty$, если

$$\Delta\left(x^{(n)}\right) = \left|\left|x^{(n)} - x\right|\right| \to 0, \quad \text{при} \quad n \to \infty$$

Методы численного решения задач линейной алгебры

• **Прямые методы**. Решение системы x находится за конечное число арифметических операций.
В следствии погрешностей округления при решении задач на ЭВМ, прямые методы не приводят к точному решению. Сопоставление различных прямых методов производится по числу арифметических действий, необходимых для получения решения.

• **Итерационные методы (методы последовательных приближений)**.
Решение x СЛАУ находится как предел последовательных приближений $x^{(n)}$ при $n \to \infty$.
Как правило, за конечное число итераций этот предел не достигается и вычисления производятся до тех пор, пока не будет выполнена оценка

$$\left\| x^{(n)} - x \right\| < \varepsilon,$$

где $\varepsilon > 0$ – точность. Качество различных итерационных процессов сравнивают по необходимому числу итераций $n(\varepsilon)$, которое необходимо провести для получения заданной точности.

Норма матрицы

Определение. Пусть в линейном арифметическом пространстве $\mathbb{R}^m$ задана норма $||\cdot||_*$. Норму $||\cdot||_k$ в линейном пространстве $M_m(\mathbb{R})$ называют **согласованной** с нормой $||\cdot||_*$, если для $\forall A \in M_m(\mathbb{R})$ и $\forall x \in \mathbb{R}^m$ выполняется соотношение:

$$||Ax||_* \leqslant ||A||_k ||x||_*.$$

Определение. **Определение.** Число $||A|| = \sup\limits_{x \neq 0} \dfrac{||Ax||}{||x||}$ называется **нормой** матрицы A **подчиненной** данной норме $||x||$.

Для подчиненной нормы матрицы A выполняются все аксиомы нормы:

- $||A|| \geqslant 0$ и $||A|| = 0 \quad \Longrightarrow \quad A = 0$;
- $||\alpha A|| = |\alpha| ||x||$;
- $||A + B|| \leqslant ||A|| + ||B||$ для $\forall A, B$;

Дополнительно

- $||AB|| \leqslant ||A|| ||B||$ для $\forall A, B$;
- $||Ax|| \leqslant ||A|| ||x||$.

Примеры подчиненных норм матриц

Норма $||A||_1 = \max\limits_{1\leqslant j\leqslant m} \sum\limits_{i=1}^{m} |a_{ij}|$ – **максимальная столбцевая** или **октаэдрическая норма**, подчинена норме $||x||_1$.

Норма $||A||_s = \left(\max\limits_{1\leqslant j\leqslant m} \mu_j\right)^{1/2}$ – спектральная норма матрицы A, подчиненная норме $||x||_2$, где μ_j – собственные числа оператора $A^T A$.

Норма $||A||_\infty = \max\limits_{1\leqslant i\leqslant m} \sum\limits_{j=1}^{m} |a_{ij}|$ – **максимальная строчная** или **кубическая норма** подчинена норме $||x||_\infty$.

Исключение. Евклидова норма $||A||_2 = \left(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{m} a_{ij}^2\right)^{1/2}$.

Она является согласованной с $||x||_2$, но не является подчиненной.

Причем: $||A||_s \leqslant ||A||_2$.

Обусловленность СЛАУ.

Рассмотрим СЛАУ

$$Ax = f, \quad A \in M_m(\mathbb{R}).$$

Рассмотрим два типа устойчивости:

• **устойчивость по правой части**, когда возмущается только правая часть f, а матрица A остается неизменной,

• **коэффициентная устойчивость**, когда возмущается только матрица A, а правая часть f остается неизменной.

Вместо вектора f задается близкий ему вектор $\tilde{f}$ (например, из-за погрешностей округления). Рассмотрим «возмущенную систему»

$$A\tilde{x} = \tilde{f}, \quad \text{где} \quad \Delta x = \tilde{x} - x, \qquad \Delta f = \tilde{f} - f.$$

Определение. Говорят, что система $Ax = f$ **устойчива по правой части**, если при $\forall f, \tilde{f}$ справедлива оценка

$$||\Delta x|| \leqslant M_1 ||\Delta f||,$$

где $M_1 > 0$ – постоянная, не зависящая от правых частей $f, \tilde{f}$.

Пусть $\det A \neq 0$. Покажем, что система устойчива по правой части.

$$A(\Delta x) = \Delta f \quad \Rightarrow \quad \Delta x = A^{-1}(\Delta f).$$

Используя аксиомы нормы, получаем

$$||\Delta x|| \leqslant ||A^{-1}||\;||\Delta f||.$$

Следовательно $M_1 = ||A^{-1}||$.

Исключение. Чем ближе к нулю определитель матрицы A, тем больше постоянная M_1, тем сильнее погрешность правой части может исказить искомое решение.

Рассмотрим относительные погрещности $\delta\, x$ и $\delta\, f$.
Использую аксиомы нормы получаем $||f|| \leqslant ||A||\;||x||$.Тогда

$$\frac{||\Delta x||}{||x||} \leqslant \mathrm{cond}(\mathrm{A}) \frac{||\Delta f||}{||f||},$$

где $\mathrm{cond}(\mathrm{A}) = ||A^{-1}||\;||A||$.

Определение. Число $\mathrm{cond}(\mathrm{A})$, входящее в оценку, называется **числом обусловленности** матрицы A и характеризует степень зависимости относительной погрешности решения от относительной погрешности правой части. Матрицы с большим числом обусловленности называют **плохо обусловленными матрицами**.

Замечание. Число обусловленности матрицы всегда положительно и зависит от заданной нормы матрицы.

Свойства числа обусловленности матрицы.

- $\mathrm{cond}(\mathrm{A}) = \mathrm{cond}(\mathrm{A}^{-1})$.
- $\mathrm{cond}(\mathrm{AB}) \leqslant \mathrm{cond}(\mathrm{A})\mathrm{cond}(\mathrm{B})$.
- $\mathrm{cond}(\mathrm{A}) \geqslant 1$.
- $\mathrm{cond}(\mathrm{A}) \geqslant \dfrac{|\lambda_{\mathrm{max}}|}{|\lambda_{\mathrm{min}}|}$,

где λ_{max}, λ_{min} – наибольшее и наименьшее по абсолютной величине собственные значения.

Метод Гаусса.

Прямой ход метода Гаусса.

Запишем систему $Ax = f$ в развернутом виде

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = f_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = f_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mm}x_m = f_m.
\end{aligned}
$$

Идея метода: Последовательное исключении неизвестных x_1, x_2, ..., x_m из системы.

Пусть $a_{11} \neq 0$.

Тогда a_{11} называется **главным или ведущим элементом** первого шага.

Поделим первое уравнение системы на a_{11}, получим

$$x_1 + c_{12}x_2 + \ldots + c_{1m}x_m = y_1,$$

где $c_{1j} = \dfrac{a_{1j}}{a_{11}}$, $j = 2, \ldots, m$, $y_1 = \dfrac{f_1}{a_{11}}$.

Тогда

$$
\begin{aligned}
&x_1 + c_{12}x_2 + \ldots + c_{1m}x_m = y_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = f_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mm}x_m = f_m.
\end{aligned}
$$

Вычтем первое уравнение полученной системы умноженное на a_{i1} из i-го уравнения системы, $i = 2, 3, \ldots, m$:

$$
\begin{aligned}
x_1 + c_{12}x_2 + \ldots + c_{1j}x_j + \ldots + c_{1m}x_m &= y_1,\\
a_{22}^{(1)}x_2 + \ldots + a_{2j}^{(1)}x_j + \ldots + a_{2m}^{(1)}x_m &= f_2^{(1)},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
a_{m2}^{(1)}x_2 + \ldots + a_{mj}^{(1)}x_j + \ldots + a_{mm}^{(1)}x_m &= f_m^{(1)},
\end{aligned}
$$

где $a_{ij}^{(1)} = a_{ij} - c_{1j}a_{i1}$, $f_i^{(1)} = f_i - y_1 a_{i1}$, где $i, j = 2, 3, \ldots, m$

Структура матрицы полученной системы:

$$\begin{pmatrix} 1 & \times & \dots & \times \\ 0 & \times & \dots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \times & \dots & \times \end{pmatrix}$$

Если $a_{22}^{(1)} \neq 0$ (**главный элемент второго шага**), то из системы аналогично можно исключить неизвестное x_2 и перейти к системе, матрица которой имеет следующую структуру:

$$\begin{pmatrix} 1 & \times & \times & \dots & \times \\ 0 & 1 & \times & \dots & \times \\ 0 & 0 & \times & \dots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \times & \dots & \times \end{pmatrix}$$

Исключая аналогично неизвестные x_3, x_4, ..., x_m придем к окончательной системе уравнений вида:

$$\begin{aligned} x_1 + c_{12}x_2 + \ldots + c_{1m}x_m &= y_1, \\ x_2 + \ldots + c_{2m}x_m &= y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{m-1} + c_{m-1,m}x_m &= y_{m-1}, \\ x_m &= y_m, \end{aligned}$$

Обратный ход метода Гаусса заключается в нахождении неизвестных x_1, x_2, ..., x_m.

$$x_m = y_m, \quad x_{m-1} = y_{m-1} - c_{m-1,m}x_m.$$

В общем виде формулы обратного хода имеют вид:

$$x_i = y_i - \sum_{j=i+1}^{m} c_{ij}x_j, \quad i = (m-1), \ldots 1, \quad x_m = y_m. \tag{1}$$

Подсчет числа действий.

Ограничимся вычислением количества операций умножения и деления.

• Для вычисления коэффициентов c_{ij} требуется делений:

$$(m-1)+(m-2)+\ldots+2+1=\frac{m(m-1)}{2}.$$

• Для вычисления коэффициентов $a_{ij}^{(k)}$, требуется умножений:

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

• Вычисление правых частей y_k требует m делений, а вычисление коэффициентов $f_i^{(k)}$ требует умножений:

$$(m-1)+(m-2)+\ldots+2+1=\frac{m(m-1)}{2}.$$

Осуществление прямого хода требует действий:

$$\frac{m(m-1)}{2}+\frac{(m-1)m(2m-1)}{6}+m+\frac{m(m-1)}{2}=\frac{2m^3+3m^2+2}{6};$$

Для реализации обратного хода требуется умножений:

$$1 + 2 + 3 + \ldots + (m-1) = \frac{m(m-1)}{2}$$

Итого, для реализации метода Гаусса требуется действий:

$$\frac{2m^3 + 3m^2 + 2}{6} + \frac{m(m-1)}{2} = \frac{m^3 + 3m^2 - m}{3}.$$

Метод Гаусса с выбором главного элемента.

Может оказаться так, что система имеет единственное решение, даже если какой-либо из угловых миноров матрицы A равен нулю. В этом случае обычный метод Гаусса может оказаться непригодным и применяют метод Гаусса **с выбором главного элемента.**

> **Основная идея:** на очередном шаге исключают не следующее по номеру неизвестное, а неизвестное, коэффициент при котором по модулю наибольший. Т.е. в качестве ведущего элемента выбирается наибольший по модулю элемент.

Проиллюстрируем на примере СЛАУ из 2-х уравнений.

$$a_{11}x_1 + a_{12}x_2 = f_1;$$
$$a_{21}x_1 + a_{22}x_2 = f_2.$$

Метод Гаусса с выбором главного элемента по строке. Пусть $|a_{12}| > |a_{11}|$. Тогда на первом шаге исключается переменное x_2

$$a_{12}x_2 + a_{11}x_1 = f_1;$$
$$a_{22}x_2 + a_{21}x_1 = f_2,$$

и к данной системе применяется первый шаг обычного метода Гаусса.

Метод Гаусса с выбором главного элемента по столбцу. Пусть $|a_{21}| > |a_{11}|$.

Тогда

$$a_{21}x_1 + a_{22}x_2 = f_2;$$
$$a_{11}x_1 + a_{12}x_2 = f_1,$$

и к новой системе применяют первый шаг обычного метода Гаусса.

Иногда применяют метод Гаусса с выбором главного элемента по всей матрице, когда в качестве ведущего элемента выбирают наибольший по модулю элемент матрицы системы.

LU-разложение матрицы.

Метод Гаусса преобразует систему в эквивалентную систему

$$Cx = y,$$

где C – верхняя треугольная матрица с единицами на главной диагонали.

Векторы правых частей f и y связаны соотношениями.

$$\begin{aligned}
f_1 &= y_1 a_{11}, \\
f_2 &= y_1 a_{21} + y_2 a_{22}^{(1)}, \\
f_3 &= y_1 a_{31} + y_2 a_{32}^{(1)} + y_3 a_{33}^{(2)}, \qquad\qquad \text{или} \quad f = By, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
f_m &= y_1 a_{m1} + y_2 a_{m2}^{(1)} + y_3 a_{m3}^{(2)} + \ldots + y_m a_{mm}^{(m-1)}.
\end{aligned}$$

где B - нижняя треугольная матрица с элементами $b_{ii} \neq 0$.

Так как $y = B^{-1} f \quad \Longrightarrow \quad Cx = B^{-1} f \quad \Longrightarrow \quad BCx = f.$

Следовательно, получено разложение $A = BC$, где B - **нижняя треугольная** матрица с ненулевыми элементами на главной диагонали, а C - **верхняя треугольная** матрица с единицами на главной диагонали.

В этом случае, метод Гаусса можно трактовать так:

• производиться разложение матрицы $A = BC$,

• последовательно решаются две системы уравнений:

$$By = f, \quad Cx = y.$$

Теорема об LU-разложении. Пусть

$$\Delta_1 = a_{11}, \quad \Delta_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots, \Delta_m = \det A.$$

Теорема. Пусть все угловые миноры матрицы A отличны от нуля, $\Delta_i \neq 0$, $i = 1, 2, \ldots, m$. Тогда матрицу A можно представить, причем единственным образом, в виде произведения

$$A = LU, \tag{2}$$

где L - нижняя треугольная матрица с ненулевыми диагональными элементами и U - верхняя треугольная матрица с единичной диагональю.

◄ Доказательство проводим методом математической индукции.
Пусть $m = 2$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Будем искать разложение матрицы A в виде:

$$A = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix},$$

где $l_{11}, l_{21}, l_{22}, u_{12}$ -неизвестные числа.Для их нахождения приходим к системе уравнений

$$l_{11} = a_{11}, \quad l_{11}u_{12} = a_{12}, \quad l_{21} = a_{21},$$
$$l_{21}u_{12} + l_{22} = a_{22}.$$

Данная система имеет единственное решение:

$$l_{11} = a_{11} \neq 0, \quad u_{12} = a_{12}/a_{11}, \quad l_{21} = a_{21},$$
$$l_{22} = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} \neq 0.$$

Пусть утверждение теоремы справедливо для матриц порядка $(k-1)$.
Докажем, что оно справедливо для матриц порядка k.
Представим матрицу A порядка k в виде

$$A = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1,k-1} & a_{1k} \\ \dots & \dots & \dots & \dots \\ a_{k-1,1} & \dots & a_{k-1,k-1} & a_{k-1,k} \\ \hline a_{k1} & \dots & a_{k,k-1} & a_{kk} \end{array}\right) \tag{3}$$

и обозначим

$$A_{k-1} = \begin{pmatrix} a_{11} & \dots & a_{1,k-1} \\ \dots & \dots & \dots \\ a_{k-1,1} & \dots & a_{k-1,k-1} \end{pmatrix}, \quad a_{k-1} = \begin{pmatrix} a_{1k} \\ \dots \\ a_{k-1,k} \end{pmatrix},$$

$$b_{k-1} = (a_{k1}, \dots, a_{k,k-1})$$

Согласно предположению индукции существует разложение матрицы

$$A_{k-1} = L_{k-1} U_{k-1}.$$

Будем искать разложение матрицы (3) в виде

$$A = \begin{bmatrix} L_{k-1} & 0 \\ l_{k-1} & l_{kk} \end{bmatrix} \begin{bmatrix} U_{k-1} & u_{k-1} \\ 0 & 1 \end{bmatrix}, \qquad (4)$$

где $l_{k-1} = (l_{k1}, l_{k2} \ldots, l_{k,k-1})$ и $u_{k-1} = (u_{1k}, u_{2k} \ldots, u_{k-1,k})^T$ - неизвестные векторы.

Перемножая матрицы в правой части уравнения (4) и учитывая (3), приходим к системе уравнений

$$L_{k-1}u_{k-1} = a_{k-1}, \qquad (5)$$

$$l_{k-1}U_{k-1} = b_{k-1}, \qquad (6)$$

$$l_{k-1}u_{k-1} + l_{kk} = a_{kk}. \qquad (7)$$

Из предположения индукции следует существование матриц L_{k-1}^{-1} и U_{k-1}^{-1}. Следовательно из (5)-(7) получаем

$$u_{k-1} = L_{k-1}^{-1}a_{k-1}, \quad l_{k-1} = b_{k-1}U_{k-1}^{-1}, \quad l_{kk} = a_{kk} - l_{k-1}u_{k-1}.$$

Докажем, что $l_{kk} \neq 0$. Запишем

$$\det A = (\det L_{k-1}) l_{kk} (\det U_{k-1}) = (\det L_{k-1}) l_{kk}.$$

По условию теоремы $\det A \neq 0$, следовательно $l_{kk} \neq 0$.

Таким образом, LU-разложение матрицы A порядка k существует.

Докажем **единственность** такого разложения.

Предположим противное, пусть матрицу A можно разложить двумя способами:

$$A = L_1 U_1 = L_2 U_2.$$

Тогда $U_1 U_2^{-1} = L_1^{-1} L_2$. Матрица в левой части указанного равенства является верхней треугольной, а в правой - нижней треугольной. Такое равенство возможно, когда обе матрицы $U_1 U_2^{-1}$ и $L_1^{-1} L_2$ являются диагональными.
Но на диагонали матрицы $U_1 U_2^{-1}$ стоят единицы, следовательно и на диагонали $L_1^{-1} L_2$ также стоят единицы. Таким образом эти матрицы являются единичными:

$$U_1 U_2^{-1} = L_1^{-1} L_2 = E.$$

Следовательно, $U_1 = U_2$ и $L_1 = L_2$, т.е. разложение единственно. ►