Ю.В.Малинковский, д.ф.-м.н., профессор

1.1. Линейные преобразования.

Пусть P — некоторое поле (чаще всего это будет поле действительных чисел $\mathbb{R}$ или поле комплексных чисел $\mathbb{C}$).

Определение 1.1.1. Множество V называется **линейным** или **векторным пространством над полем** P, если

а) любым элементам $\mathbf{x}, \mathbf{y} \in V$ поставлен в соответствие элемент $\mathbf{z} = \mathbf{x}+\mathbf{y} \in V$, называемый **суммой** $\mathbf{x}$ и $\mathbf{y}$;

б) любому элементу $\mathbf{x} \in V$ и любому числу λ из поля P поставлен в соответствие элемент $\lambda\mathbf{x}$, называемый произведением числа λ на элемент $\mathbf{x}$

так, что эти операции удовлетворяют следующим аксиомам:

I. 1. $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ (коммутативность).

2. $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$ (ассоциативность).

3. существует элемент $\mathbf{0}$ такой, что $\mathbf{x}+\mathbf{0} = \mathbf{x}$ для любого элемента $\mathbf{x} \in V$ (существование нулевого элемента).

4. для любого элемента $\mathbf{x} \in V$ существует элемент $-\mathbf{x} \in V,$ называемый **противоположным к элементу** $\mathbf{x}$ такой, что $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$ (существование противоположного элемента).

II. 1. $1 \cdot \mathbf{x} = \mathbf{x}$ для любого элемента $\mathbf{x} \in V$ (здесь 1 — единица поля P).

2. $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ для любых $\alpha,\ \beta \in P$ и любого $\mathbf{x} \in V.$

III. 1. $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}$ для любых $\alpha,\ \beta \in P$ и любого $\mathbf{x} \in V.$

2. $\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}$ для любого $\alpha \in P$ и любых элементов $\mathbf{x},\ \mathbf{y} \in V.$

Элементы линейного пространства называются **векторами**.

Определение 1.1.2. Линейное пространство V называется n**-мерным**, если в нем существует n линейно независимых векторов, а любые $n+1$ векторов линейно зависимы.

Известно, что если в линейном пространстве V существует n линейно независимых векторов таких, что любой вектор из V является их линейной комбинацией, то пространство V n-мерно.

Определение 1.1.3. Совокупность n линейно независимых векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ n-мерного пространства V называется **базисом в** V.

Теорема 1.1.1. *Любой вектор* $\mathbf{x}$ *из* n*-мерного линейного пространства* V *можно представить, и притом единственным образом, в виде линейной комбинации векторов базиса.*

Определение 1.1.4. Если $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ — базис n-мерного пространства V и

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n, \text{ где } x_1, x_2, \ldots, x_n \in P,$$

то числа $x_1, x_2, \ldots, x_n$ называются **координатами вектора** $\mathbf{x}$ **в базисе** $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

Известно, что при сложении векторов соответствующие координаты складываются, а при умножении вектора на число из поля P, его координаты умножаются на это число.

Определение 1.1.5. Линейные пространства V и V' над одним и тем же полем P называются **изоморфными**, если между векторами $\mathbf{x} \in V$ и $\mathbf{x}' \in V'$ можно установить

$\mathbf{y}$ соответствует вектор $\mathbf{y}'$, то

1) вектору $\mathbf{x}+\mathbf{y}$ соответствует вектор $\mathbf{x}'+\mathbf{y}'$;

2) вектору $\lambda\mathbf{x}$ соответствует вектор $\lambda\mathbf{x}'$ для любого $\lambda \in P$.

Известно, что *два линейных пространства изоморфны друг другу тогда и только тогда, когда их размерности одинаковы.* Поэтому, не искажая алгебраических свойств, можно считать вектора в n-мерном пространстве векторами-столбцами с n координатами, которые складываются при сложении векторов и умножаются на число из поля P при умножении вектора на это число.

Пусть в n-мерном пространстве V заданы два базиса

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \qquad \text{и} \qquad \mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n.$$

Разложим каждый вектор второго базиса по векторам первого базиса:

$$\mathbf{f}_1 = c_{11}\mathbf{e}_1 + c_{12}\mathbf{e}_2 + \ldots + c_{1n}\mathbf{e}_n$$

$$\mathbf{f}_2 = c_{21}\mathbf{e}_1 + c_{22}\mathbf{e}_2 + \ldots + c_{2n}\mathbf{e}_n$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\mathbf{f}_n = c_{n1}\mathbf{e}_1 + c_{n2}\mathbf{e}_2 + \ldots + c_{nn}\mathbf{e}_n$$

Матрица

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}$$

называется **матрицей перехода от базиса $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ к базису $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$**.

Известно, что матрица перехода от одного базиса к другому является невырожденной.

Пусть V — линейное пространство над полем P. Отображение линейного пространства в себя называется **преобразованием этого пространства**. Если A — преобразование пространства V, то через $A\mathbf{x}$ обозначается вектор из V, в который отображается вектор $\mathbf{x}$ из V под действием преобразования A. Вектор $A\mathbf{x}$ называется **образом вектора $\mathbf{x}$**.

Определение 1.1.6. Преобразование A линейного пространства V называется **линейным**, если

$$1)\ \ A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} \quad \forall\ \mathbf{x}, \mathbf{y} \in V;$$

$$2)\ \ A(\lambda\mathbf{x}) = \lambda A\mathbf{x} \quad \forall \mathbf{x} \in V \quad \text{и} \quad \forall \lambda \in P.$$

Каждому линейному преобразованию n-мерного линейного пространства V над полем P ставится в соответствие определенная квадратная матрица порядка n с элементами из поля P, называемая матрицей этого линейного преобразования. Для ее построения выберем в пространстве V фиксированный базис $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ и найдем образы базисных векторов при линейном преобразовании A. Получим вектора $A\mathbf{e}_1, A\mathbf{e}_2, \dots, A\mathbf{e}_n \in V$. Разложим каждый из них по данному базису:

$$A\mathbf{e}_1 = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \ldots + a_{1n}\mathbf{e}_n$$

$$A\mathbf{e}_2 = a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + \ldots + a_{2n}\mathbf{e}_n$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$A\mathbf{e}_n = a_{n1}\mathbf{e}_1 + a_{n2}\mathbf{e}_2 + \ldots + a_{nn}\mathbf{e}_n$$

Определение 1.1.7. Матрица A, в которой i-й столбец ($i=\overline{1,n}$) составляют координаты вектора $A\mathbf{e}_i$, называется **матрицей линейного преобразования** A **в базисе** $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$:

$$A=(A\mathbf{e}_1,A\mathbf{e}_2,\ldots,A\mathbf{e}_n)=\begin{pmatrix}a_{11}&a_{21}&\ldots&a_{n1}\\a_{12}&a_{22}&\ldots&a_{n2}\\\ldots&\ldots&\ldots&\ldots\\a_{1n}&a_{2n}&\ldots&a_{nn}\end{pmatrix}.$$

Известно, что между всеми линейными преобразованиями пространства V и всеми квадратными матрицами порядка n с элементами из поля P существует взаимно однозначное соответствие, зависящее от выбора базиса.

Пусть $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ и $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$ — два базиса в n-мерном линейном пространстве V, C — матрица перехода от базиса $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ к базису $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$. Тогда матрица B линейного преобразования в базисе $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$ выражается через матрицу A этого преобразования в базисе $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ с помощью равенства

$$B=C^{-1}AC,$$

где C^{-1} — матрица, обратная к C, служащая матрицей перехода от базиса $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$ к базису $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$.

Определение 1.1.8. Квадратные матрицы A и B одного и того же порядка называются **подобными**, если существует такая невырожденная матрица C, что

$$B=C^{-1}AC,$$

Таким образом, *матрицы линейного преобразования A в различных базисах являются подобными.*

Определение 1.1.9. Вектор $\mathbf{x}\neq\mathbf{0}$, для которого

$$A\mathbf{x}=\lambda\mathbf{x}, \tag{1.1.1}$$

называется **собственным вектором** линейного преобразования A (или соответствующей матрицы этого линейного преобразования). Соответствующее число λ называется **собственным значением (характеристическим числом)**, соответствующим собственному вектору $\mathbf{x}$.

Теорема 1.1.2. *В комплексном пространстве V всякое линейное преобразование A имеет хотя бы один собственный вектор.*

Д о к а з а т е л ь с т в о. Пусть фиксирован некоторый базис, A — матрица линейного преобразования в этом базисе. Равенство (1.1.1) эквивалентно векторному равенству

$$(A-\lambda E)\mathbf{x}=\mathbf{0}, \tag{1.1.2}$$

где E — единичная матрица. Для существования ненулевого решения этой системы линейных однородных уравнений необходимо и достаточно, чтобы ее определитель был равен нулю:

$$|A-\lambda E|=\begin{vmatrix}a_{11}-\lambda&a_{21}&\ldots&a_{n1}\\a_{12}&a_{22}-\lambda&\ldots&a_{n2}\\\ldots&\ldots&\ldots&\ldots\\a_{1n}&a_{2n}&\ldots&a_{nn}-\lambda\end{vmatrix}=0.$$

теореме алгебры он имеет хотя бы один комплексный корень. Подставим найденный корень в систему (1.1.2). Получим систему однородных линейных уравнений, определитель которой равен нулю. Следовательно, она имеет ненулевое решение $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, где $x_1, x_2, \ldots, x_n$ — координаты $\mathbf{x}$ в выбранном базисе. Для найденных λ и $\mathbf{x}$ выполнено (1.1.2), т.е. эквивалентное ему равенство (1.1.1). Согласно определению 1.1.9 это означает, что $\mathbf{x}$ — собственный вектор линейного преобразования A.

Определение 1.1.10. Многочлен $|A - \lambda E|$ n-й степени называется **характеристическим многочленом** матрицы A, а уравнение (1.1.2) —**характеристическим уравнением**.

Теорема 1.1.3. *Подобные матрицы обладают одинаковыми характеристическими многочленами, и, следовательно, одинаковыми собственными значениями.*

Д о к а з а т е л ь с т в о. Пусть матрицы A и B подобны, т.е.

$$B = C^{-1}AC, \quad \text{где} \quad |C| \neq 0.$$

Тогда характеристический многочлен матрицы B равен

$$|B - \lambda E| = |C^{-1}AC - \lambda E| = |C^{-1}AC - \lambda E C^{-1}C| = |C^{-1}AC - C^{-1}\lambda EC| =$$

$$= |C^{-1}(A - \lambda E)C| = |C|^{-1}|A - \lambda E||C| = |A - \lambda E|,$$

т.е. совпадает с характеристическим многочленом матрицы A.

Вспоминая, что матрицы линейного преобразования в разных базисах подобны, делаем вывод: *матрицы линейного преобразования в разных базисах имеют один и тот же набор собственных значений.* Во первых, это обосновывает корректность определения собственных векторов и собственных значений, во вторых, многочлен $|A - \lambda E|$ можно назвать не только характеристическим многочленом матрицы A, но и характеристическим многочленом линейного преобразования A в данном базисе.

Определение 1.1.11. Набор собственных значений, каждое из которых берется с кратностью, которую оно имеет как корень характеристического уравнения, называется **спектром линейного преобразования** A (или **матрицы** A).

Часто необходимо знать, может ли линейное преобразование A иметь в некотором базисе диагональную матрицу.

Теорема 1.1.4. *Линейное преобразование A задается в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ диагональной матрицей тогда и только тогда, когда все вектора этого базиса являются собственными векторами преобразования A.*

Д о к а з а т е л ь с т в о. Действительно, равенство

$$A\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

эквивалентно тому, что в i-й строке матрицы преобразования A в этом базисе все элементы вне главной диагонали равны 0, а на диагонали стоит λ_i:

$$A = (A\mathbf{e}_1, A\mathbf{e}_2, \ldots, A\mathbf{e}_n) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

ным собственным значениям, независимы.

Говорят, что линейное преобразование A действительного линейного пространства (т.е. V рассматривается над полем $\mathbb{R}$ действительных чисел) **имеет простой спектр**, если все характеристические числа действительны и различны. В этом случае собственные вектора образуют базис и по теореме 1.1.4 получаем, что *каждое линейное преобразование с простым спектром может быть задано диагональной матрицей (в базисе из собственных векторов).* Для матриц это утверждение принимает следующую форму.

Любая матрица, собственные значения которой действительны и различны, подобна диагональной матрице или, как говорят, приводится к диагональному виду.

Теперь рассмотрим вопрос о приведении матрицы к наиболее простому виду в общем случае. Для этого нам необходимо вспомнить понятие жордановой нормальной формы матрицы.

1.2. Канонический вид многочленных матриц.

Определение 1.2.1. Квадратная матрица порядка n, элементы которой — многочлены от переменной λ с коэффициентами из поля P, называется **многочленной матрицей** или **λ-матрицей**.

Примерами многочленных матриц могут служить 1) характеристическая матрица $A-\lambda E$; 2) матрица с элементами из поля P.

Пусть дана λ-матрица

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \dots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \dots & a_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ a_{n1}(\lambda) & a_{n2}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

Элементарными преобразованиями этой матрицы называются преобразования следующих четырех типов:

1) умножение любой строки матрицы $A(\lambda)$ на любое число $\alpha \in P, \quad \alpha \neq 0$;

2) умножение любого столбца матрицы $A(\lambda)$ на любое число $\alpha \in P, \quad \alpha \neq 0$;

3) прибавление к любой i-й строке матрицы $A(\lambda)$ любой ее j-й строки, $j \neq i$, умноженной на любой многочлен $\varphi(\lambda)$ над полем P;

4) прибавление к любому i-му столбцу матрицы $A(\lambda)$ любого ее j-го столбца, $j \neq i$, умноженного на любой многочлен $\varphi(\lambda)$ над полем P.

Утверждение 1.2.1. *Для любого элементарного преобразования λ-матрицы существует обратное преобразование, являющееся элементарным.*

Доказательство. Действительно, обратным для преобразования 1) является элементарное преобразование умножения той же строки на $\alpha^{-1} \in P, \quad \alpha^{-1} \neq 0$; для 3) — элементарное преобразование прибавления к i-й строке j-й строки, умноженной на многочлен $-\varphi(\lambda)$. Аналогичные преобразования, но применяемые к столбцам, дают обратные элементарные преобразования к 2) и 4).

Утверждение 1.2.2. *В матрице $A(\lambda)$ при помощи конечного числа элементарных преобразований можно переставить любые две строки и любые два столбца.*

Доказательство. Перестановка i-й и j-й строк может быть осуществлена с помощью четырех элементарных преобразований, применяемых только к этим строкам по

$$\begin{pmatrix} i \\ j \end{pmatrix} \longrightarrow \begin{pmatrix} i+j \\ j \end{pmatrix} \longrightarrow \begin{pmatrix} i+j \\ -i \end{pmatrix} \longrightarrow \begin{pmatrix} j \\ -i \end{pmatrix} \longrightarrow \begin{pmatrix} j \\ i \end{pmatrix}.$$

Определение 1.2.2. Говорят, что λ-матрицы $A(\lambda)$ и $B(\lambda)$ **эквивалентны** и пишут $A(\lambda) \sim B(\lambda)$, если с помощью конечного числа элементарных преобразований можно перейти от матрицы $A(\lambda)$ к матрице $B(\lambda)$.

Утверждение 1.2.3. *Отношение $\sim$ между λ-матрицами является отношением эквивалентности.*

Доказательство. 1. Данное отношение в множестве λ-матриц *рефлексивно*: $A(\lambda) \sim A(\lambda)$. Чтобы в этом убедиться, достаточно, например, первую строку матрицы умножить на $1 \in P$.

2. Это отношение *симметрично*: если $A(\lambda) \sim B(\lambda)$, то $B(\lambda) \sim A(\lambda)$. Действительно, в силу утверждения 1.2.1 для каждого из элементарных преобразований существует обратное, причем элементарное преобразование. Поэтому, если с помощью конечной цепочки элементарных преобразований можно перейти от $A(\lambda)$ к $B(\lambda)$, то то с помощью конечной цепочки обратных элементарных преобразований, выполненных в обратном порядке, мы можем перейти от $B(\lambda)$ к $A(\lambda)$.

3. Наконец, рассматриваемое отношение *транзитивно*: если $A(\lambda) \sim B(\lambda)$, а $B(\lambda) \sim C(\lambda)$, то $A(\lambda) \sim C(\lambda)$. Для построения конечной цепочки элементарных преобразований, преобразующих $A(\lambda)$ в $C(\lambda)$, достаточно последовательно провести конечное число элементарных преобразований, переводящих $A(\lambda)$ в $B(\lambda)$, а затем конечное число элементарных преобразований, переводящих $B(\lambda)$ в $C(\lambda)$.

Следствие 1.2.1. *Все квадратные λ-матрицы n-го порядка распадаются на непересекающиеся классы эквивалентных между собой матриц.*

Наша цель — отыскание среди всех λ-матриц, эквивалентных данной λ-матрице $A(\lambda)$, матрицы по возможности наиболее простого вида. Для достижения этой цели введем следующее

Определение 1.2.3. **Канонической λ-матрицей** называется многочленная матрица, обладающая следующими тремя свойствами:

а) эта матрица диагональная, т.е. имеет вид

$$\begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e_n(\lambda) \end{pmatrix};$$

б) каждый многочлен $e_i(\lambda), \quad i = \overline{2,n}$, нацело делится на многочлен $e_{i-1}(\lambda)$ (считается, что если многочлен нулевой степени равен нулю, то он делится на многочлен, равный нулю);

в) старший коэффициент каждого многочлена $e_i(\lambda), \quad i = \overline{1,n}$, равен единице, если только этот многочлен не равен тождественно нулю.

Если среди $e_i(\lambda)$ встречаются нулевые, то в силу б) они занимают последние места на главной диагонали. Если же среди $e_i(\lambda)$ встречаются отличные от нуля многочлены нулевой степени, то в силу в) они равны 1, а в силу б) они занимают на главной диагонали первые места.

числе единичная и нулевая.

Из курса линейной алгебры и геометрии известна следующая теорема существования и единственности для данной многочленной матрицы канонической λ-матрицы.

Теорема 1.2.1. *Любая λ-матрица эквивалентна некоторой и притом единственной канонической λ-матрице (т.е. приводится с помощью конечного числа элементарных преобразований к каноническому виду).*

Доказательство этой теоремы выносится в приложение 1.

Пусть $A(\lambda)$ — любая λ-матрица порядка n и пусть для всех $1 \le k \le n$

$d_k(\lambda)$ — наибольший общий делитель (НОД) всех миноров k-го порядка, взятый со старшим

коэффициентом 1.

Договоримся считать, что $d_k(\lambda) = 0$, если все миноры k-го порядка равны 0. Следовательно, матрица $A(\lambda)$ однозначно определяет набор многочленов

$$d_1(\lambda), \;\; d_2(\lambda), \;\; \ldots, \;\; d_n(\lambda).$$

При этом, очевидно, что

$$d_1(\lambda) = \text{НОД всех элементов } A(\lambda) \text{ со старшим коэффициентом } 1,$$

$$d_n(\lambda) = \frac{|A(\lambda)|}{\text{старший коэффициент многочлена } |A(\lambda)|}.$$

Лемма 1.2.1. *Многочлены $d_k(\lambda)$, $1 \le k \le n$, не меняются при выролнении элементарных преобразований над матрицей $A(\lambda)$.*

Доказательство леммы 1.2.1 выносится в приложение 1.

Следствие 1.2.1. *Всем λ-матрицам, эквивалентным матрице $A(\lambda)$, соответствует один и тот же набор многочленов $d_1(\lambda)$, $d_2(\lambda)$, $\ldots$, $d_n(\lambda)$.*

В приложении доказывается, что

$$d_k(\lambda) = e_1(\lambda)e_2(\lambda)\ldots e_k(\lambda), \qquad k = \overline{1, n},$$

и что если матрица $A(\lambda)$ имеет ранг r, то

$$e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)} \quad \text{для } \; k \le r \qquad (\text{полагаем} \quad d_0(\lambda) = 1);$$

$$e_{r+1}(\lambda) = \ldots = e_n(\lambda) = 0.$$

Многочлены $e_k(\lambda)$ называются **инвариантными множителями** λ-матрицы $A(\lambda)$. Они, как и $d_k(\lambda)$, не меняются при элементарных преобразованиях.

Таким образом, получается следующий *алгоритм вычисления инвариантных множителей*:

1-й шаг. Вычисляем

$$d_1(\lambda) = \text{НОД всех элементов } A(\lambda) \text{ со старшим коэффициентом } 1,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

. . . .

$$d_n(\lambda)=\frac{|A(\lambda)|}{\text{старший коэффициент многочлена } |A(\lambda)|}.$$

2-й шаг. Полагаем

$$e_1(\lambda)=d_1(\lambda),\quad e_2(\lambda)=\frac{d_2(\lambda)}{d_1(\lambda)},\quad \ldots,\quad e_r(\lambda)=\frac{d_r(\lambda)}{d_{r-1}(\lambda)};\quad e_{r+1}(\lambda)=\ldots=e_n(\lambda)=0,$$

где r — ранг матрицы $A(\lambda)$.

Пример 1.2.1. Необходимо привести к каноническому виду λ-матрицу

$$A(\lambda)=\begin{pmatrix}\lambda^3-\lambda & 2\lambda^2\\ \lambda^2+5\lambda & 3\lambda\end{pmatrix}.$$

1-й способ.

Последовательно применяя элементарные преобразования, приведем данную матрицу к каноническому виду:

$$A(\lambda)\sim\begin{pmatrix}\lambda^3-\lambda & \frac{2}{3}\lambda^2\\ \lambda^2+5\lambda & \lambda\end{pmatrix}\sim\begin{pmatrix}\frac{1}{3}\lambda^3-\frac{10}{3}\lambda^2-\lambda & 0\\ \lambda^2+5\lambda & \lambda\end{pmatrix}\sim\begin{pmatrix}\frac{1}{3}\lambda^3-\frac{10}{3}\lambda^2-\lambda & 0\\ 0 & \lambda\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}\lambda^3-10\lambda^2-3\lambda & 0\\ 0 & \lambda\end{pmatrix}\sim\begin{pmatrix}\lambda & 0\\ 0 & \lambda^3-10\lambda^2-3\lambda\end{pmatrix}.$$

2-й способ.

Сначала найдем $d_k(\lambda)$:

$$d_1(\lambda)=\text{НОД}\,(\lambda^3-\lambda,\ 2\lambda^2,\ \lambda^2+5\lambda,3\lambda)=\lambda;$$

$$d_2(\lambda)=|A(\lambda)|=\lambda^4-10\lambda^3-3\lambda^2;$$

в силу того, что старший коэффициент этого многочлена уже равен 1. Искомые инвариантные множители равны

$$e_1(\lambda)=d_1(\lambda)=\lambda;\quad e_2(\lambda)=\frac{d_2(\lambda)}{d_1(\lambda)}=\lambda^3-10\lambda^2-3\lambda.$$

Поэтому

$$A(\lambda)\sim\begin{pmatrix}e_1(\lambda) & 0\\ 0 & e_2(\lambda)\end{pmatrix}\sim\begin{pmatrix}\lambda & 0\\ 0 & \lambda^3-10\lambda^2-3\lambda\end{pmatrix}.$$

Определение 1.2.4. **Элементарными матрицами** называются многочленные матрицы одного из следующих двух типов:

$$(1)\quad\begin{pmatrix}1 &&&&&&&&\\ &\cdot&&&&&&&\\ &&\cdot&&&&&&\\ &&&\cdot&&&&&\\ &&&&1&&&&\\ &&&&&\alpha&&&\\ &&&&&&1&&\\ &&&&&&&\cdot&\\ &&&&&&&&\cdot\\ &&&&&&&&&1\end{pmatrix}$$

$i=\overline{1,n}$);

$$
(2)\quad \begin{pmatrix}
1 & & & & & & \cdot & & & \\
 & \cdot & & & & & \cdot & & & \\
 & & \cdot & & & & \cdot & & & \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \varphi(\lambda) & \cdot & \cdot & \cdot \\
 & & & & \cdot & & \cdot & & & \\
 & & & & & \cdot & \cdot & & & \\
 & & & & & & 1 & & & \\
 & & & & & & \cdot & \cdot & & \\
 & & & & & & \cdot & & \cdot & \\
 & & & & & & \cdot & & & 1
\end{pmatrix}
$$

(здесь произвольный многочлен $\varphi(\lambda)$ над полем P стоит на пересечении i-й строки и j-го столбца $i,j=\overline{1,n}$).

Заметим, что определитель матрицы (1) — число $\alpha \neq 0$, а определитель матрицы (2) равен 1.

Лемма 1.2.2. *Выполнение в λ-матрице $A(\lambda)$ любого элементарного преобразования равносильно умножению этой матрицы слева и справа на некоторую элементарную матрицу.*

Д о к а з а т е л ь с т в о. Легко проверяется, что

1) умножение $A(\lambda)$ слева на матрицу (1) эквивалентно умножению i-й строки $A(\lambda)$ на число α;

2) умножение $A(\lambda)$ справа на матрицу (1) эквивалентно умножению i-го столбца $A(\lambda)$ на число α;

3) умножение $A(\lambda)$ слева на матрицу (2) эквивалентно прибавлению к i-й строке матрицы $A(\lambda)$ j-й строки, умноженной на многочлен $\varphi(\lambda)$;

4) умножение $A(\lambda)$ справа на матрицу (2) эквивалентно прибавлению к j-му столбцу матрицы $A(\lambda)$ i-го стролбца, умноженного на многочлен $\varphi(\lambda)$.

1.3. Матричные многочлены. Связь подобия числовых матриц с эквивалентностью их характеристических матриц

На понятие λ-матрицы можно посмотреть с несколько иной стороны.

Определение 1.3.1. **Матричным λ-многочленом порядка n над полем** P называется многочлен от переменной λ, коэффициентами которого являются квадратные матрицы порядка n над полем P:

$$A_0\lambda^k + A_1\lambda^{k-1} + \ldots + A_{k-1}\lambda + A_k.$$

Пользуясь определением сложения матриц и умножения матрицы на число и производя указанные действия, получим некоторую λ-матрицу порядка n. Таким образом, *любой матричный многочлен n-го порядка можно записать в виде λ-матрицы n-го порядка.* Обратно, *любая λ-матрица n-го порядка может быть записана в виде матричного λ-многочлена n-го порядка.* Например,

$$\begin{pmatrix} 3\lambda^2-5 & \lambda+1 \\ \lambda^4+2\lambda & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\lambda^4 + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\lambda^3 + \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}\lambda^2 + \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\lambda + \begin{pmatrix} -5 & 1 \\ 0 & -3 \end{pmatrix}.$$

Пусть дана λ-матрица

$$A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_{m-1}\lambda + A_m,$$

$A(\lambda)$. Очевидно, m совпадает с наивысшей степенью многочленов — элементов матрицы $A(\lambda)$. Назовем матричный многочлен $A(\lambda)$ **регулярным**, если определитель матрицы A_0, $|A_0| \neq 0$.

Взгляд на λ-матрицы как на матричные многочлены позволяет развивать теорию делимости, аналогичную теории делимости для многочленов, но усложняемую некоммутативностью умножения матриц и наличием делителей нуля.

Рассмотрим основные действия над матричными многочленами. Пусть $A(\lambda)$, $B(\lambda)$ — матричные многочлены одного и того же порядка. Обозначим через m наибольшую из степеней этих многочленов. Эти многочлены можно записать в виде

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_{m-1}\lambda + A_m,$$

$$B(\lambda) = B_0\lambda^m + B_1\lambda^{m-1} + \ldots + B_{m-1}\lambda + A_m.$$

Тогда

$$A(\lambda) \pm B(\lambda) = (A_0 \pm B_0)\lambda^m + (A_1 \pm B_1)\lambda^{m-1} + \ldots + (A_{m-1} \pm B_{m-1})\lambda + (A_m \pm B_m),$$

т.е. *сумма (разность) двух матричных многочленов одного и того же порядка является многочленом, степень которого не превосходит наибольшей из степеней данных многочленов.*

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ степеней m и p одного и того же порядка n:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_{m-1}\lambda + A_m \qquad (A_0 \neq 0),$$

$$B(\lambda) = B_0\lambda^p + B_1\lambda^{p-1} + \ldots + B_{p-1}\lambda + B_p \qquad (B_0 \neq 0).$$

Тогда

$$A(\lambda)B(\lambda) = A_0B_0\lambda^{m+p} + (A_0B_1 + A_1B_0)\lambda^{m+p-1} + \ldots + A_mB_p.$$

Если бы мы умножили $B(\lambda)$ на $A(\lambda)$, то получили бы, вообще говоря, другой многочлен.

Умножение матричных многочленов обладает еще одним специфическим свойством. В отличие от скалярных многочленов произведение может иметь степень, меньшую $m+p$. Действительно, A_0B_0 может равняться 0 при $A_0 \neq 0$, $B_0 \neq 0$. Если хотя бы одна из матриц невырожденная, то из $A_0 \neq 0$, $B_0 \neq 0$ следует, что $A_0B_0 \neq 0$. Итак, *произведение двух матричных многочленов равно многочлену, степень которого не превосходит суммы степеней сомножителей. Если хотя бы один из сомножителей — регулярный многочлен, то степень произведения равна сумме степеней сомножителей.*

Рассмотрим теперь операцию деления матричных многочленов. Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ одного и того же порядка, причем $B(\lambda)$ — регулярный многочлен:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_{m-1}\lambda + A_m \qquad (A_0 \neq 0),$$

$$B(\lambda) = B_0\lambda^p + B_1\lambda^{p-1} + \ldots + B_{p-1}\lambda + B_p \qquad (|B_0| \neq 0).$$

Определение 1.3.2. Говорят, что матричные многочлены $Q(\lambda)$ и $R(\lambda)$ являются соответственно **правым частным** и **правым остатком** при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \text{ и степень } R(\lambda) < \text{степень } B(\lambda).$$

Матричные многочлены $\hat{Q}(\lambda)$ и $\hat{R}(\lambda)$ называются соответственно **левым частным** и **левым остатком** при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = B(\lambda)\hat{Q}(\lambda) + \hat{R}(\lambda) \text{ и степень } \hat{R}(\lambda) < \text{степень } B(\lambda).$$

при левом — слева. Вообще говоря, $Q(\lambda) \neq \hat{Q}(\lambda), \quad R(\lambda) \neq \hat{R}(\lambda)$.

Лемма 1.3.1. *Как правое, так и левое деление матричных многочленов одного и того же порядка всегда выполнимо и однозначно, если делитель — регулярный многочлен.*

Доказательство. Приведем доказательство для правого деления; доказательство для левого деления аналогичное. Если $m < p$, то можно положить $Q(\lambda) = 0, \; R(\lambda) = A(\lambda)$. В случае $m \geq p$ для нахождения частного $Q(\lambda)$ и остатка $R(\lambda)$ применим обычную схему деления многочлена на многочлен. "Разделим" старший член делимого $A_0\lambda^m$ на старший член делителя $B_0\lambda^p$. Получим старший член искомого частного $A_0B_0^{-1}\lambda^{m-p}$. Умножим этот член справа на делитель $B(\lambda)$ и полученное произведение вычтем из $A(\lambda)$. Найдем "первый остаток" $A^{(1)}(\lambda)$:

$$A(\lambda) = A_0B_0^{-1}\lambda^{m-p}B(\lambda) + A^{(1)}(\lambda). \tag{1.3.1}$$

Степень $m^{(1)}$ многочлена $A^{(1)}(\lambda)$ меньше m:

$$A^{(1)}(\lambda) = A_0^{(1)}\lambda^{m^{(1)}} + \ldots \qquad (A_0^{(1)} \neq 0, \;\; m^{(1)} < m) \tag{1.3.2}$$

(если $A^{(1)}(\lambda) = 0$, то полагаем $Q(\lambda) = A_0B_0^{-1}\lambda^{m-p}, \quad R(\lambda) = 0$).

Если $m^{(1)} \geq p$, то, повторяя этот процесс, получаем

$$A^{(1)}(\lambda) = A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p}B(\lambda) + A^{(2)}(\lambda), \tag{1.3.3}$$

$$A^{(2)}(\lambda) = A_0^{(2)}\lambda^{m^{(2)}} + \ldots \qquad (m^{(2)} < m^{(1)})$$

и т.д.

Так как степени многочленов $A(\lambda), \;\; A^{(1)}(\lambda), \;\; A^{(2)}(\lambda), \;\; \ldots$ убывают, то на некотором этапе придем к остатку $R(\lambda)$, степень которого меньше p. Из (1.3.1)-(1.3.3) следует, что

$$A(\lambda) = A_0B_0^{-1}\lambda^{m-p}B(\lambda) + A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p}B(\lambda) + \ldots + R(\lambda),$$

т.е.

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda),$$

где $Q(\lambda) = A_0B_0^{-1}\lambda^{m-p} + A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p} + \ldots$. Возможность правого деления доказана. Докажем его однозначность. Пусть одновременно

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda), \tag{1.3.4}$$

$$A(\lambda) = Q^*(\lambda)B(\lambda) + R^*(\lambda), \tag{1.3.5}$$

где степени $R(\lambda)$ и $R^*(\lambda)$ меньше степени $B(\lambda)$, т.е. p. Вычитая (1.3.5) из (1.3.4), получим

$$\Big[Q(\lambda) - Q^*(\lambda)\Big]B(\lambda) = R^*(\lambda) - R(\lambda). \tag{1.3.6}$$

Если бы $Q(\lambda) - Q^*(\lambda)$ не являлось бы тождественным нулем, то, поскольку $|B_0| \neq 0$, то степень левой части (1.3.6) равнялась бы сумме степеней $Q(\lambda) - Q^*(\lambda)$ и $B(\lambda)$, и потому была бы не меньше p. С другой стороны, степень $R^*(\lambda) - R(\lambda)$ строго меньше p. Противоречие. Значит, $Q(\lambda) = Q^*(\lambda)$. Но тогда из (1.3.6) следует, что $R^*(\lambda) = R(\lambda)$.

Основная теорема о подобии матриц.

Матрицы A и B с элементами из поля P подобны тогда и только тогда, когда их характеристические матрицы $A - \lambda E$ и $B - \lambda E$ эквивалентны.

невырожденная матрица, трансформирующая A в B, т.е. такая, что

$$B = R^{-1}AR.$$

может быть найдена следующим образом. Если $A - \lambda E \sim B - \lambda E$, то $A - \lambda E$ конечным числом элементарных преобразований переводится в $B - \lambda E$. Берем те из этих преобразований, которые относятся к столбцам, и произведение соответствующих элементарных матриц, взятых в том же порядке, обозначим через $V(\lambda)$. Делим $V(\lambda)$ справа на $B - \lambda E$. Остаток от этого деления и будет матрицей R.

Указанное деление можно и не выполнять, а воспользоваться следующей леммой.

Лемма 1.3.2. *Пусть*

$$V(\lambda) = V_0\lambda^s + V_1\lambda^{s-1} + \ldots + V_{s-1}\lambda + V_s, \quad V_0 \neq 0.$$

Если

$$V(\lambda) = Q(\lambda)(\lambda E - B) + R,$$

то

$$R = V_0B^s + V_1B^{s-1} + \ldots + V_{s-1}B + V_s.$$

Доказательство выносится в приложение 2.

Пример 1.3.1. Даны матрицы

$$A = \begin{pmatrix} -2 & 1 \\ 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} -10 & -4 \\ 26 & 11 \end{pmatrix}.$$

Для матрицы A

$$A - \lambda E = \begin{pmatrix} -2-\lambda & 1 \\ 0 & 3-\lambda \end{pmatrix}, \quad d_1(\lambda) = \text{НОД}(-2-\lambda,\ 1,\ 0,\ 3-\lambda) = 1,$$

$$d_2(\lambda) = \begin{vmatrix} -2-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = \lambda^2 - \lambda - 6, \quad e_1(\lambda) = d_1(\lambda) = 1, \quad e_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \lambda^2 - \lambda - 6.$$

Поэтому канонический вид матрицы $A - \lambda E$ есть

$$\begin{pmatrix} 1 & 0 \\ 0 & \lambda^2 - \lambda - 6 \end{pmatrix}.$$

Для матрицы B имеем

$$B - \lambda E = \begin{pmatrix} -10-\lambda & -4 \\ 26 & 11-\lambda \end{pmatrix}, \quad d_1(\lambda) = 1, \ \ d_2(\lambda) = \lambda^2 - \lambda - 6,$$

т.е. $B - \lambda E$ имеет такой же канонический вид, что и $A - \lambda E$. По основной теореме о подобии матриц A и B подобны. Для нахождения преобразующей матрицы R найдем какую-нибудь цепочку элементарных преобразований, переводящих $A - \lambda E$ в $B - \lambda E$:

$$A - \lambda E = \begin{pmatrix} -2-\lambda & 1 \\ 0 & 3-\lambda \end{pmatrix} \sim \begin{pmatrix} -2-\lambda & 1 \\ -16-8\lambda & 11-\lambda \end{pmatrix} \sim \begin{pmatrix} 8+4\lambda & -4 \\ -16-8\lambda & 11-\lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 40+4\lambda & -4 \\ -104 & 11-\lambda \end{pmatrix} \sim \begin{pmatrix} -10-\lambda & -4 \\ 26 & 11-\lambda \end{pmatrix} = B - \lambda E.$$

К столбцам относятся только два последних преобразования:

на элементарную матрицу

$$\begin{pmatrix} 1 & 0 \\ -8 & 1 \end{pmatrix};$$

2) первый столбец умножается на $-\frac{1}{4}$, чему соответствует умножение на элементарную матрицу

$$\begin{pmatrix} -\frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix}.$$

Поэтому

$$V(\lambda) = \begin{pmatrix} 1 & 0 \\ -8 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{4} & 0 \\ 2 & 1 \end{pmatrix}.$$

Остаток от деления $V(\lambda)$ на $B - \lambda E$ есть $V(\lambda)$, т.е.

$$R = \begin{pmatrix} -\frac{1}{4} & 0 \\ 2 & 1 \end{pmatrix}.$$

Следует отметить, что матрица, трансформирующая A в B, определяется неоднозначно. Для рассмотренного примера в качестве трансформирующей матрицы можно взять также матрицу

$$\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

1.4. Жорданова нормальная форма матрицы

Рассмотрим квадратные матрицы порядка n с элементами из поля P.

Определение 1.4.1. **Жордановой клеткой порядка** k, относяшейся к числу λ_0, называется матрица порядка k, $\quad k = \overline{1, n}$, вида

$$\begin{pmatrix} \lambda_0 & 1 & & & & \\ & \lambda_0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ & & & & & \lambda_0 \end{pmatrix}. \tag{1.4.1}$$

Например, матрицы

$$(\lambda_0), \qquad \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}, \qquad \begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$$

являются жордановыми клетками порядков $1, 2, 3$ соответственно.

Определение 1.4.2. **Жордановой матрицей порядка** n называется матрица порядка n вида

$$J = \begin{pmatrix} \boxed{J_1} & & & & \\ & \boxed{J_2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \boxed{J_s} \end{pmatrix}, \tag{1.4.2}$$

Заметим, что порядки клеток могут как совпадать, так и не совпадать, числа, к которым относятся клетки, также могут совпадать и не совпадать. Ясно, что $1 \le s \le n$.

Диагональные матрицы являются частным случаем жордановых матриц. У них все жордановы клетки первого порядка.

Найдем канонический вид для характеристической матрицы

$$\begin{pmatrix} \lambda_0-\lambda & 1 & & & \\ & \lambda_0-\lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0-\lambda \end{pmatrix}. \tag{1.4.3}$$

одной жордановой клетки (1.4.1) порядка k. Определитель этой матрицы равен $(\lambda_0-\lambda)^k = (-1)^k(\lambda-\lambda_0)^k$. Так как старший коэффициент у $d_k(\lambda)$ равен 1, то $d_k(\lambda) = (\lambda-\lambda_0)^k$. Среди миноров $(k-1)$-го порядка имеется равный 1 (получающийся вычеркиванием первого столбца и последней строки). Поэтому $d_{k-1}(\lambda) = 1$, но тогда $d_1(\lambda) = \ldots = d_{k-2}(\lambda) = 1$. Следовательно, каноническим видом для матрицы (1.4.3) служит следующая λ-матрица k-го порядка:

$$\begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 1 & \\ & & & & & (\lambda-\lambda_0)^k \end{pmatrix}. \tag{1.4.4}$$

Теорема 1.4.1. *Две жордановы матрицы подобны тогда и только тогда, когда они состоят из одних и тех же жордановых клеток, т.е. отличаются, быть может, лишь расположением этих клеток вдоль главной диагонали.*

Доказательство выносится в приложение 3.

Следствие 1.4.1. *Жорданова матрица, подобная диагональной матрице, сама диагональна. Две диагональные матрицы подобны тогда и только тогда, когда они получаются друг из друга перестановкой чисел, стоящих на главной диагонали.*

Основная теорема о приведении матрицы к жордановой нормальной форме.

Матрица A с элементами из поля P тогда и только тогда приводится в поле P к жордановой нормальной форме (т.е. подобна жордановой матрице), когда все собственные значенеия матрицы A лежат в поле P.

Приводимость матрицы A к жордановой нормальной форме **в поле** P означает, что все элементы трансформирующей матрицы содержатся в поле P. Доказательство основной теоремы выносится в приложение 3.

В силу замкнутости поля комплексных чисел отсюда немедленно вытекает

Следствие 1.4.2. *Всякая матрица с комплексными элементами приводится в поле комплексных чисел $\mathbb{C}$ к жордановой нормальной форме.*

Теорема 1.4.2 (необходимое и достаточное условие приводимости матрицы к диагональному виду). *Матрица A порядка n с элементами из поля P приводится к диагональному виду тогда и только тогда, когда все корни последнего инвариантного множителя $e_n(\lambda)$ ее характеристической матрицы лежат в поле , причем среди этих корней нет кратных.*

В приложении 3 обосновывается следующий

Алгоритм приведения матрицы к жордановой нормальной форме.

1. Находим собственные значения матрицы A и проверяем, что они принадлежат полю P. Если это не так, то делаем вывод о том, что матрица A не может быть приведена в поле P к жордановой нормальной форме. Если собственные значения матрицы A принадлежат полю P, то переходим к пункту 2.

2. Обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_t$ — все различные собственные значения матрицы A. Находим инвариантные множители характеристической матрицы $A - \lambda E$:

$$1, 1, \ldots, 1, e_{n-q+1}(\lambda), \ldots, e_{n-1}(\lambda), e_n(\lambda),$$

где $e_{n-q+1}(\lambda) \neq 1, \ldots, e_n(\lambda) \neq 1$. В приложении доказывается, что среди этих многочленов нет равных нулю и все они разлагаются на линейные множители. Сумма степеней всех многочленов равна n. Разложим эти многочлены на линейные множители:

$$e_{n-j+1}(\lambda) = \prod_{i=1}^{t} (\lambda - \lambda_i)^{k_{ij}} = (\lambda - \lambda_1)^{k_{1j}} (\lambda - \lambda_2)^{k_{2j}} \ldots (\lambda - \lambda_t)^{k_{tj}}, \qquad j = \overline{1, q}.$$

Элементарными делителями многочленов $e_{n-j+1}(\lambda), \quad j = \overline{1, q},$ называются отличные от 1 степени различных линейных двучленов, входящих в последнее разложение. **Элементарными делителями матрицы** A называются элементарные делители всех многочленов $e_{n-q+1}(\lambda), \ldots, e_n(\lambda)$.

Выписываются элементарные делители матрицы A в виде таблицы

$$\begin{array}{c}
(\lambda - \lambda_1)^{k_{11}},\ (\lambda - \lambda_1)^{k_{12}},\ \ldots,\ (\lambda - \lambda_1)^{k_{1q_1}}, \\
(\lambda - \lambda_1)^{k_{21}},\ (\lambda - \lambda_1)^{k_{22}},\ \ldots,\ (\lambda - \lambda_1)^{k_{2q_2}}, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
(\lambda - \lambda_1)^{k_{t1}},\ (\lambda - \lambda_1)^{k_{t2}},\ \ldots,\ (\lambda - \lambda_1)^{k_{tq_t}}.
\end{array}$$

Составляется жорданова матрица J порядка n, состоящая из жордановых клеток, по следующему правилу:

каждому элементарному делителю $(\lambda - \lambda_i)^{k_{ij}}$ сопоставляется жорданова клетка порядка k_{ij}, относящаяся к числу λ_i. Полученная матрица J и есть искомая матрица, т.е. матрица A подобна матрице J.

Пример 1.4.1. Найдем жорданову нормальную форму матрицы

$$A = \begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}.$$

$$|A-\lambda E| = \begin{vmatrix} 3-\lambda & -4 & 0 & 2 \\ 4 & -5-\lambda & -2 & 4 \\ 0 & 0 & 3-\lambda & -2 \\ 0 & 0 & 2 & -1-\lambda \end{vmatrix} = (\lambda+1)^2(\lambda-1)^2 = 0,$$

откуда $\lambda_1 = -1$, $\lambda_2 = 1$, $d_4(\lambda) = (\lambda+1)^2(\lambda-1)^2$. Так как имеется минор 3-го порядка

$$\begin{vmatrix} 4 & -5-\lambda & -2 \\ 0 & 0 & 3-\lambda \\ 0 & 0 & 2 \end{vmatrix} = 4\begin{vmatrix} 0 & 3-\lambda \\ 0 & 2 \end{vmatrix} = 0,$$

то $d_3(\lambda) = 1$. Аналогично, $d_2(\lambda) = d_1(\lambda) = 1$. Поэтому

$$e_4(\lambda) = \frac{d_4(\lambda)}{d_3(\lambda)} = (\lambda+1)^2(\lambda-1)^2, \quad e_3(\lambda) = \frac{d_3(\lambda)}{d_2(\lambda)} = 1, \quad e_2(\lambda) = e_1(\lambda) = 1.$$

Элементарные делители $e_4(\lambda)$, являющиеся элементарными делителями матрицы A, равны $(\lambda+1)^2$ и $(\lambda-1)^2$. Поэтому искомая жорданова нормальная форма матрицы A имеет вид

$$J = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

2.1. Обобщенная теорема Безу. Теорема Гамильтона-Кэли. Присоединенная матрица.

Матричный многочлен n-го порядка можно записать двояко:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \tag{2.1.1}$$

и

$$A(\lambda) = \lambda^m A_0 + \lambda^{m-1}A_1 + \ldots + A_m. \tag{2.1.1$'$}$$

Обе записи при скалярном аргументе λ дают один и тот же результат. Однако если мы захотим вместо скалярного аргумента λ подставить квадратную матрицу n-го порядка Λ, то результаты подстановок в (2.1.1) и (2.1.1$'$) будут, вообще говоря, различны, так как степени матрицы Λ могут быть не перестановочными с матричными коэффициентами $A_0, A_1, \ldots, A_m$. Положим

$$A(\Lambda) = A_0\Lambda^m + A_1\Lambda^{m-1} + \ldots + A_m,$$
$$\hat{A}(\Lambda) = \Lambda^m A_0 + \Lambda^{m-1}A_1 + \ldots + A_m$$

и будем называть $A(\Lambda)$ **правым**, а $\hat{A}(\Lambda)$ **левым значением матричного многочлена** $A(\lambda)$ при подстановке матрицы Λ вместо λ.

Рассмотрим два матричных многочлена

$$A(\lambda) = \sum_{i=0}^{m} A_{m-i}\lambda^i, \quad B(\lambda) = \sum_{k=0}^{p} B_{p-k}\lambda^k$$

и их произведение

$$P(\lambda) = \sum_{i=0}^{m}\sum_{k=0}^{p} A_{m-i}\lambda^i B_{p-k}\lambda^k = \sum_{i=0}^{m}\sum_{k=0}^{p} A_{m-i}B_{p-k}\lambda^{i+k} = \sum_{j=0}^{m+p}\Big(\sum_{i+k=j} A_{m-i}B_{p-k}\Big)\lambda^j \tag{2.1.2}$$

и

$$P(\lambda) = \sum_{i=0}^{m}\sum_{k=0}^{p} \lambda^i A_{m-i}\lambda^k B_{p-k} = \sum_{i=0}^{m}\sum_{k=0}^{p} \lambda^{i+k}A_{m-i}B_{p-k} = \sum_{j=0}^{m+p}\lambda^j \sum_{i+k=j} A_{m-i}B_{p-k} \tag{2.1.2$'$}$$

Преобразования в (2.1.2) справедливы при замене λ на матрицу Λ, если Λ перестановочна со всеми матричными коэффициентами B_{p-k}. Аналогично, в (2.1.2$'$) λ можно заменить на Λ, если Λ перестановочна со всеми коэффициентами A_{m-i}. В первом случае

$$P(\Lambda) = A(\Lambda)B(\Lambda),$$

во втором

$$\hat{P}(\Lambda) = \hat{A}(\Lambda)\hat{B}(\Lambda)$$

Итак,

1) правое значение произведения двух матричных многочленов равно произведению правых значений сомножителей, если матрица-аргумент Λ перестановочна со всеми коэффициентами правого сомножителя;

чений сомножителей, если матрица-аргумент Λ перестановочна со всеми коэффициентами левого сомножителя.

Рассмотрим произвольный матричный многочлен n-го порядка

$$F(\lambda) = F_0\lambda^m + F_1\lambda^{m-1} + \ldots + F_m \quad (F_0 \neq 0).$$

Разделим его на бином $\lambda E - A$ справа и слева:

$$F(\lambda) = Q(\lambda)(\lambda E - A) + R, \quad F(\lambda) = (\lambda E - A)\hat{Q}(\lambda) + \hat{R}.$$

Так как A перестановочна со всеми коэффициентами бинома $(\lambda E - A)$ $(EA = AE = A,\ AA = A)$, то правое и левое значения $F(A)$ и $\hat{F}(A)$ равны

$$F(A) = Q(A)(A - A) + R = R, \quad \hat{F}(A) = (A - A)\hat{Q}(A) + \hat{R} = \hat{R}.$$

Нами доказана

Обобщенная теорема Безу. *При правом делении матричного многочлена $F(\lambda)$ на бином $\lambda E - A$ остаток от деления равен $F(A)$. При левом делении матричного многочлена $F(\lambda)$ на бином $\lambda E - A$ остаток от деления равен $\hat{F}(A)$.*

Следствие 2.1.1. *Многочлен $F(\lambda)$ делится без остатка справа на бином $\lambda E - A$ тогда и только тогда, когда $F(A) = 0$. Многочлен $F(\lambda)$ делится без остатка слева на бином $\lambda E - A$ тогда и только тогда, когда $\hat{F}(A) = 0$.*

Пример 2.1.1. Пусть A — квадратная матрица и $f(\lambda)$ — многочлен. Тогда $F(\lambda) = f(\lambda)E - f(A)$ делится (справа и слева) на $\lambda E - A$, поскольку $F(A) = \hat{F}(A) = 0$.

Рассмотрим характеристическую матрицу $\lambda E - A$ матрицы n-го порядка A. Определитель характеристической матрицы

$$\Delta(\lambda) = |\lambda E - A| = |\lambda\delta_{ik} - a_{ik}|_{i,k=1}^{n}$$

является скалярным многочленом от λ и называется **характеристическим многочленом** матрицы A (нам удобно в этой и следующих главах в определениях заменить $A - \lambda E$ на $\lambda E - A$). Здесь

$$\delta_{ik} = \begin{cases} 1, & \text{если } i = k, \\ 0, & \text{если } i \neq k \end{cases} \text{ — символ Кронекера.}$$

Определение 2.1.1. Матрица $B(\lambda) = (b_{ik}(\lambda))_{i,k=1}^{n}$, где $b_{ik}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ki} - a_{ki}$ в определителе $\Delta(\lambda)$, называется **присоединенной матрицей** для матрицы A.

Пример 2.1.2. Для матрицы

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

характеристическая матрица есть

$$\lambda E - A = \begin{pmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{pmatrix},$$

присоединенная матрица равна

$$B(\lambda) = \begin{pmatrix} \lambda - a_{22} & a_{12} \\ a_{21} & \lambda - a_{11} \end{pmatrix}.$$

Если элементы присоединенной матрицы разделить на $\Delta(\lambda)$, то получится матрица, обратная к $\lambda E - A$. Итак, $\frac{B(\lambda)}{\Delta(\lambda)}$ — матрица, обратная к $\lambda E - A$, т.е.

$$(\lambda E - A)\frac{B(\lambda)}{\Delta(\lambda)} = \frac{B(\lambda)}{\Delta(\lambda)}(\lambda E - A) = E,$$

откуда

$$(\lambda E - A)B(\lambda) = B(\lambda)(\lambda E - A) = \Delta(\lambda)E. \tag{2.1.3}$$

Все три части этого равенства — матричные многочлены от переменной λ. Из (2.1.3) видно, что $\Delta(\lambda)E$ делится на $\lambda E - A$ слева и справа без остатка. По следствию 2.1.1 из обобщенной теоремы Безу это возможно тогда и только тогда, когда $\Delta(A)E = \Delta(A) = 0$. Нами доказана

Теорема Гамильтона-Кэли. *Всякая квадратная матрица A удовлетворяет своему характеристическому уравнению, т.е.*

$$\Delta(A) = 0.$$

Пример 2.1.3. Для матрицы

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \qquad \text{имеем} \qquad \Delta(\lambda) = \begin{vmatrix} \lambda - 1 & 0 \\ -1 & \lambda - 2 \end{vmatrix} = \lambda^2 - 3\lambda + 2,$$

поэтому действительно

$$\Delta(A) = A^2 - 3A + 2E = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} -3 & 0 \\ -3 & -6 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Пусть $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все собственные значения матрицы A, т.е. все корни характеристического многочлена $\Delta(\lambda)$ (каждое из чисел λ_i повторяется в этом ряду столько раз, какова его кратность как корня многочлена $\Delta(\lambda)$). Тогда

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n).$$

Пусть дан произвольный скалярный многочлен $g(\mu)$. Найдем собственные значения матрицы $g(A)$. Для этого разложим $g(\mu)$ на линейные множители:

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2)\ldots(\mu - \mu_l).$$

Подставим в обе части этого равенства вместо μ матрицу A:

$$g(A) = a_0(A - \mu_1 E)(A - \mu_2 E)\ldots(A - \mu_l E).$$

Тогда определитель матрицы $g(A)$ будет равен

$$|g(A)| = a_0^n |A - \mu_1 E||A - \mu_2 E|\ldots|A - \mu_l E| =$$

$$= a_0^n (-1)^n |\mu_1 E - A|(-1)^n|\mu_2 E - A|\ldots(-1)^n|\mu_l E - A| =$$

$$= a_0^n(-1)^{nl}\Delta(\mu_1)\Delta(\mu_2)\ldots\Delta(\mu_l) = a_0^n(-1)^{nl}\prod_{i=1}^{l}\Delta(\mu_i) =$$

$$= (-1)^{nl}a_0^n\prod_{i=1}^{l}\prod_{k=1}^{n}(\mu_i-\lambda_k) = a_0^n\prod_{i=1}^{l}\prod_{k=1}^{n}(\lambda_k-\mu_i) =$$

$$= \prod_{k=1}^{n}a_0\prod_{i=1}^{l}(\lambda_k-\mu_i) = \prod_{k=1}^{n}g(\lambda_k) = g(\lambda_1)g(\lambda_2)\ldots g(\lambda_n).$$

Заменив в этом равенстве многочлен $g(\mu)$ на многочлен $\lambda - g(\mu)$ от переменной μ, где λ — некоторый параметр, найдем

$$|\lambda E - g(A)| = [\lambda - g(\lambda_1)][\lambda - g(\lambda_2)]\ldots[\lambda - g(\lambda_n)].$$

Отсюда вытекает следующая

Теорема 2.1.1. *Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все собственные значения (с учетом кратностей) матрицы A, а $g(\mu)$ — некоторый скалярный многочлен, то $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$ — все собственные значения матрицы $g(A)$.*

В частности, если A имеет собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$, то матрица A^k имеет собственные значения $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ $(k = 0, 1, \ldots)$.

Выведем формулу, выражающую присоединенную матрицу $B(\lambda)$ через характеристический многочлен $\Delta(\lambda)$. Пусть

$$\Delta(\lambda) = \lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \ldots - p_n.$$

Разность $\Delta(\lambda) - \Delta(\mu)$ делится без остатка на $\lambda - \mu$. Поэтому

$$\delta(\lambda,\mu) = \frac{\Delta(\lambda)-\Delta(\mu)}{\lambda-\mu} = \lambda^{n-1} + (\mu - p_1)\lambda^{n-2} + (\mu^2 - p_1\mu - p_2)\lambda^{n-3} + \ldots$$

является многочленом относительно λ и относительно μ.

Тождество

$$\Delta(\lambda) - \Delta(\mu) = \delta(\lambda,\mu)(\lambda-\mu)$$

не нарушится, если в него вместо λ и μ подставить перестановочные между собой матрицы λE и A. Тогда, поскольку по теореме Гамильтона-Келли $\Delta(A) = 0$, то

$$\Delta(\lambda)E = \delta(\lambda E, A)(\lambda E - A).$$

Но из (2.1.3) следует, что $\Delta(\lambda)E = B(\lambda)(\lambda E - A)$. В силу однозначности операции деления из последних двух равенств вытекает искомая формула

$$B(\lambda) = \delta(\lambda E, A). \tag{2.1.4}$$

Пусть λ_0 — собственное значение матрицы A, т.е. $\Delta(\lambda_0) = 0$. Подставляя $\lambda = \lambda_0$ в (2.1.3), получим

$$(\lambda_0 E - A)B(\lambda_0) = 0. \tag{2.1.5}$$

Допустим, что $B(\lambda_0) \neq 0$, и обозначим через $\mathbf{x}$ любой ненулевой столбец этой матрицы. Тогда из (2.1.5)

$$(\lambda_0 E - A)\mathbf{x} = 0$$

значению λ_0. Итак, нами доказана

Теорема 2.1.2. *Если коэффициенты характеристического многочлена известны, то присоединенная матрица может быть найдена по формуле*

$$B(\lambda) = \delta(\lambda E, A), \quad \text{где } \delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu}.$$

Если λ_0 — собственное значение матрицы A, то ненулевые столбцы матрицы $B(\lambda_0)$ являются собственными векторами матрицы A, соответствующими собственному значению λ_0.

Пример 2.1.4. Пусть

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Характеристический многочлен матрицы A равен

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - 1)^2(\lambda - 2) = \lambda^3 - 4\lambda^2 + 5\lambda - 2.$$

Поэтому

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \frac{\lambda^3 - 4\lambda^2 + 5\lambda - \mu^3 + 4\mu^2 - 5\mu}{\lambda - \mu} =$$
$$= \frac{\lambda^3 - \mu^3}{\lambda - \mu} - 4\frac{\lambda^2 - \mu^2}{\lambda - \mu} + 5 = \lambda^2 + \lambda\mu + \mu^2 - 4(\lambda + \mu) + 5 =$$
$$= \lambda^2 + (\mu - 4)\lambda + \mu^2 - 4\mu + 5.$$

Следовательно,

$$B(\lambda) = \delta(\lambda E, A) = \lambda^2 E + \lambda(A - 4E) + A^2 - 4A + 5E,$$

т.е.

$$B(\lambda) = \begin{pmatrix} \lambda^2 - 2\lambda & -\lambda + 2 & \lambda - 2 \\ -1 & \lambda^2 - 3\lambda + 2 & \lambda - 2 \\ -\lambda + 1 & \lambda - 1 & \lambda^2 - 3\lambda + 2 \end{pmatrix}.$$

Собственному значению $\lambda_1 = 1$ соответствует собственный вектор (1,1,0), что дает первый столбец при $\lambda = 1$. Собственному значению $\lambda_2 = 2$ соответствует собственный вектор (0,1,1), что дает второй столбец при $\lambda = 2$.

2.2. Минимальный многочлен матрицы. Приведенная присоединенная матрица.

Определение 2.2.1. Скалярный многочлен $f(\lambda)$ называется **аннулирующим многочленом** квадратной матрицы A, если

$$f(A) = 0.$$

Определение 2.2.2. Аннулирующий многочлен $\psi(\lambda)$ наименьшей степени со старшим коэффициентом, равным 1, называется **минимальным многочленом** матрицы A.

Для любой матрицы аннулирующий многочлен существует. Действительно, согласно теореме Гамильтона-Келли характеристический многочлен $\Delta(\lambda)$ матрицы A является аннулирующим для этой матрицы. Однако, как будет видно, он, вообще говоря, не является минимальным.

$$f(\lambda) = \psi(\lambda)q(\lambda) + r(\lambda), \quad \text{стеиень } r(\lambda) < \text{ стеиень } \psi(\lambda).$$

Подставим сюда матрицу A вместо скалярного аргумента λ:

$$f(A) = \psi(A)q(A) + r(A).$$

Так как $f(\lambda)$ и $\psi(\lambda)$ — аннулирующие многочлены, то $f(A) = 0$, $\psi(A) = 0$, следовательно, $r(A) = 0$. Но степень $r(\lambda)$ меньше степени минимального многочлена, т.е. $r(\lambda) \equiv 0$ (в противном случае существовал бы аннулирующий многочлен $r(\lambda)$, степень которого меньше степени минимального многочлена).

Итак, *любой аннулирующий многочлен матрицы делится без остатка на ее минимальный многочлен.*

Теперь нетрудно убедиться в том, что *минимальный многочлен единственный.* Действительно, если $\psi(\lambda)$ и $\psi_1(\lambda)$ — два минимальных многочлена, то они делятся друг на друга, а так как у них старший коэффициент равен 1, то они совпадают.

Существование и единственность минимального многочлена устанавливают корректность определения 2.2.2.

Так как соответствующие миноры матриц $A - \lambda E$ и $\lambda E - A$ могут отличаться только знаком, то определенные ранее $d_k(\lambda)$, $k = \overline{1, n}$, можно определить как

$$d_k(\lambda) = \text{НОД всех миноров } k\text{-го порядка характеристической матрицы } \lambda E - A,$$

$$\text{взятый со старшим коэффициентом 1.}$$

В частности, $d_{n-1}(\lambda)$ совпадает с наибольшим общим делителем всех элементов присоединенной матрицы $B(\lambda)$, взятым со старшим коэффициентом 1.

Выведем формулу, связывающую минимальный и характеристический многочлены. Вынося наибольший общий делитель элементов присоединенной матрицы за знак матрицы, получим

$$B(\lambda) = d_{n-1}(\lambda)C(\lambda), \tag{2.2.1}$$

где $C(\lambda)$ — некоторая многочленная матрица, которая называется **приведенной присоединенной матрицей** для матрицы A. Из (2.1.3) и (2.2.1) находим

$$\Delta(\lambda)E = (\lambda E - A)C(\lambda)d_{n-1}(\lambda). \tag{2.2.2}$$

Отсюда следует, что $\Delta(\lambda)$ делится без остатка на $d_{n-1}(\lambda)$, т.е.

$$\frac{\Delta(\lambda)}{d_{n-1}(\lambda)} = \psi(\lambda), \tag{2.2.3}$$

где $\psi(\lambda)$ — некоторый многочлен. Тогда из (2.2.2)

$$\psi(\lambda)E = (\lambda E - A)C(\lambda). \tag{2.2.4}$$

Так как $\psi(\lambda)E$ делится без остатка на $\lambda E - A$, то по следствию из обобщенной теоремы Безу $\psi(A)E = \psi(A) = 0$. Итак, многочлен $\psi(\lambda)$, определяемый равенством (2.2.3), является аннулирующим для A. Докажем, что он является минимальным.

Пусть $\psi^*(\lambda)$ — минимальный многочлен A. Тогда $\psi(\lambda)$ делится без остатка на $\psi^*(\lambda)$:

$$\psi(\lambda) = \psi^*(\lambda)\chi(\lambda). \tag{2.2.5}$$

Поскольку $\psi^*(A) = 0$, то по следствию из обобщенной теоремы Безу матричный многочлен $\psi^*(\lambda)E$ делится слева без остатка на $\lambda E - A$:

$$\psi^*(\lambda)E = (\lambda E - A)C^*(\lambda).$$

$$\psi(\lambda)E = (\lambda E - A)C^*(\lambda)\chi(\lambda). \qquad (2.2.6)$$

Из (2.2.4) и (2.2.6) видно, что $C(\lambda)$ и $C^*(\lambda)\chi(\lambda)$ являются левыми частными при делении $\psi(\lambda)E$ на $\lambda E - A$. В силу однозначности деления

$$C(\lambda) = C^*(\lambda)\chi(\lambda).$$

Отсюда следует, что $\chi(\lambda)$ — общий делитель всех элементов приведенной присоединенной матрицы $C(\lambda)$. Но наибольший общий делитель всех элементов $C(\lambda)$ равен 1, поскольку эта матрица была получена из $B(\lambda)$ вынесением из нее наибольшего общего делителя ее элементов $d_{n-1}(\lambda)$. Значит, $\chi(\lambda)$ есть константа. Так как старшие коэффициенты у $\psi(\lambda)$ и $\psi^*(\lambda)$ равны 1, то в (2.2.5) $\chi(\lambda) = 1$, т.е. $\psi(\lambda) = \psi^*(\lambda)$, что и требовалось доказать.

Итак, *многочлен $\psi(\lambda)$, определяемый формулой (2.2.3), является минимальным многочленом матрицы A.*

Мы установили следующую формулу для минимального многочлена:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{d_{n-1}(\lambda)}.$$

Вспомним, что $d_n(\lambda) = \Delta(\lambda)$, т.е.

$$\psi(\lambda) = \frac{d_n(\lambda)}{d_{n-1}(\lambda)} = e_n(\lambda).$$

Таким образом, *минимальный многочлен матрицы A совпадает с последним инвариантным множителем $e_n(\lambda)$ характеристической матрицы $\lambda E - A$.*

Выведем формулу, выражающую приведенную присоединенную матрицу через минимальный многочлен. Так как $\psi(\lambda) - \psi(\mu)$ делится на $\lambda - \mu$ без остатка, то

$$\Psi(\lambda, \mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu}$$

является многочленом. Тогда подставляя в соотношение

$$\psi(\lambda) - \psi(\mu) = (\lambda - \mu)\Psi(\lambda, \mu)$$

вместо λ и μ соответственно λE и A, получим

$$\psi(\lambda)E = \psi(\lambda E) = (\lambda E - A)\Psi(\lambda E, A).$$

Сравнивая с (2.2.4), заключаем, что

$$C(\lambda) = \Psi(\lambda E, A).$$

Это и есть искомая формула для определения приведенной присоединенной матрицы через минимальный многочлен.

Запишем (2.2.4) в виде

$$(\lambda E - A)C(\lambda) = \psi(\lambda)E. \qquad (2.2.7)$$

Приравнивая определители обеих частей (2.2.7), получим:

$$\Delta(\lambda)|C(\lambda)| = [\psi(\lambda)]^n.$$

Таким образом, $\Delta(\lambda)$ делится без остатка на $\psi(\lambda)$ (аннулирующий многочлен делится на минимальный), а некоторая степень $\psi(\lambda)$ делится на $\Delta(\lambda)$. Поэтому совокупность всех различных корней у многочленов $\Delta(\lambda)$ и $\psi(\lambda)$ одна и та же. Итак, *корнями минимального многочлена $\psi(\lambda)$ служат все различные между собой собственные значения матрицы A.*

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\dots(\lambda - \lambda_s)^{n_s} \quad (\lambda_i \neq \lambda_j \quad \text{при} \quad i \neq j),$$

то

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\dots(\lambda - \lambda_s)^{m_s} \quad \text{где} \quad 0 < m_k \leq n_k. \tag{2.2.8}$$

Отметим еще одно полезное свойство матрицы $C(\lambda)$. Пусть λ_0 — собственное значение матрицы A. Тогда $\psi(\lambda_0) = 0$ и согласно (2.2.7)

$$(\lambda_0 E - A)C(\lambda_0) = 0.$$

Заметим, что всегда $C(\lambda_0) \neq 0$. Действительно, если бы это было не так, то все элементы приведенной присоединенной матрицы делились бы на $\lambda - \lambda_0$, что невозможно.

Пусть $\mathbf{x}$ — ненулевой столбец $C(\lambda_0)$. Тогда

$$(\lambda_0 E - A)\mathbf{x} = 0,$$

т.е. $A\mathbf{x} = \lambda_0 \mathbf{x}$. Итак, *любой ненулевой столбец матрицы* $C(\lambda_0)$ *(а такой всегда существует) является собственным вектором матрицы* A, *соответствующим собственному значению* λ_0.

3.1. Определение функции матрицы.

Пусть $A = (a_{ik})$ — квадратная матрица порядка n, $f(\lambda)$ — функция скалярного аргумента λ. Мы хотим определить, что следует понимать под $f(A)$, т.е. расширить $f(\lambda)$ на матричные значения аргумента.

Для простейшего случая, когда $f(\lambda) = c_0\lambda^p + c_1\lambda^{p-1} + \ldots + c_p$ — многочлен, решение этой задачи нам известно, именно, следует подставить вместо λ матрицу A:

$$f(A) = c_0A^p + c_1A^{p-1} + \ldots + c_pE$$

и совершить указанные в этой формуле действия с матрицами.

Пусть

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\ldots(\lambda - \lambda_s)^{m_s} \quad - \tag{3.1.1}$$

минимальный многочлен матрицы A (здесь $\lambda_1, \lambda_2, \ldots, \lambda_s$ — все различные собственные значения матрицы A). Степень этого многочлена $m = \sum\limits_{k=1}^{s} m_k \leq n$.

Пусть многочлены $g(\lambda)$ и $h(\lambda)$ таковы, что

$$g(A) = h(A). \tag{3.1.2}$$

Тогда $d(\lambda) = g(\lambda) - h(\lambda)$ — аннулирующий многочлен (ведь $d(A) = 0$), значит он делится на $\psi(\lambda)$ без остатка:

$$g(\lambda) \equiv h(\lambda)(mod\ \psi(\lambda)). \tag{3.1.3}$$

Очевидно, что тогда

$$d(\lambda) = (\lambda - \lambda_k)^{m_k}h_k(\lambda),$$

где $h_k(\lambda)$ — некоторый многочлен. Поэтому

$$d(\lambda_k) = 0,\ d'(\lambda_k) = 0,\ \ldots,\ d^{(m_{k-1})}(\lambda_k) = 0, \qquad k = \overline{1,s},$$

т.е.

$$g(\lambda_k) = h(\lambda_k),\ g'(\lambda_k) = h'(\lambda_k),\ \ldots,\ g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k), \qquad k = \overline{1,s}. \tag{3.1.4}$$

Определение 3.1.1. **Значениями функции** $f(\lambda)$ **на спектре матрицы** A называются m чисел

$$f(\lambda_k),\ f'(\lambda_k),\ \ldots,\ f^{(m_k-1)}(\lambda_k), \qquad k = \overline{1,s}. \tag{3.1.5}$$

Совокупность этих значений символически будем обозначать через $f(\Lambda_A)$. Если для функции $f(\lambda)$ существуют все производные, входящие в (3.1.5), то будем говорить, что **функция** $f(\lambda)$ **определена на спектре матрицы** A.

Равенство (3.1.4) показывает, что если для многочленов выполнено (3.1.2), то эти многочлены имеют одни и те же значения на спектре матрицы A. В символической записи

$$g(\Lambda_A) = h(\Lambda_A).$$

Наши рассуждения обратимы: если выполнено (3.1.4), т.е. многочлены g и h совпадают на спектре матрицы A, то для каждого $k = \overline{1,s}$ $d(\lambda)$ делится на $(\lambda - \lambda_k)^{m_k}$ и, значит, $d(\lambda)$

т.е. выполнено (3.1.2). Таким образом, *значения многочлена* $g(\lambda)$ *на спектре матрицы* A *полностью определяют матрицу* $g(A)$, т.е. все многочлены $g(\lambda)$, принимающие одни и те же значения на спектре матрицы A, имеют одно и то же матричное значение $g(A)$.

Потребуем, чтобы определение $f(A)$ в общем случае подчинялось тому же принципу: *значения* $f(\lambda)$ *на спектре матрицы* A *должны полностью определять* $f(A)$*, т.е. все функции* $f(\lambda)$*, имеющие одни и те же значения на спектре матрицы* A*, должны иметь одно и то же матричное значение* $f(A)$. Кроме того, общее определение для $f(A)$ в частном случае, когда $f(\lambda)$ — многочлен, должно давать тот же результат, что и непосредственная подстановка в многочлен вместо λ матрицы A. Но тогда, очевидно, что для определения $f(A)$ в общем случае достаточно найти такой многочлен $g(\lambda)$, который принимал бы те же значения на спектре матрицы A, что и $f(\lambda)$, и положить

$$f(A) = g(A).$$

Эти рассуждения показывают естественность следующего определения функции от матрицы.

Определение 3.1.2. Если функция $f(\lambda)$ определена на спектре матрицы A, то по определению

$$f(A) = g(A),$$

где $g(\lambda)$ — любой многочлен, принимающий на спектре матрицы A те же значения, что и $f(\lambda)$:

$$f(\Lambda_A) = g(\Lambda_A).$$

Утверждение 3.1.1. *Среди всех многочленов с комплексными коэффициентами, принимающих те же значения на спектре, что и* $f(\lambda)$*, имеется единственный многочлен* $r(\lambda)$ *степени, меньшей* m.

Д о к а з а т е л ь с т в о. С у щ е с т в о в а н и е. Позже мы построим многочлен, принимающий те же значения на спектре, что и $f(\lambda)$. Значит, такие многочлены существуют. Пусть $P(\lambda)$ — любой такой многочлен. Разделим его с остатком на минимальный многочлен (3.1.1):

$$P(\lambda) = \psi(\lambda)R(\lambda) + Q(\lambda), \quad \text{степень } Q(\lambda) < m.$$

Так как для всех $k = \overline{1, s}$

$$\psi(\lambda_k) = 0,\ \psi'(\lambda_k) = 0,\ \dots,\ \psi^{(m_k-1)}(\lambda_k) = 0,$$

то

$$P(\lambda_k) = Q(\lambda_k),\ P'(\lambda_k) = Q'(\lambda_k),\ \dots,\ P^{(m_k-1)}(\lambda_k) = Q^{(m_k-1)}(\lambda_k),$$

т.е. $Q(\lambda)$ принимает на спектре A те же значения, что $P(\lambda)$, а, значит, и $f(\lambda)$. При этом степень $Q(\lambda) < m$.

Е д и н с т в е н н о с т ь. Пусть $r_1(\lambda)$ и $r_2(\lambda)$ — два многочлена степени, меньшей, чем m, принимающие те же значения на спектре, что $f(\lambda)$. Тогда $\varphi(\lambda) = r_1(\lambda) - r_2(\lambda)$ — многочлен, степени, меньшей, чем m, принимающий на спектре значение нуль:

$$\varphi(\lambda_k) = 0,\ \varphi'(\lambda_k) = 0,\ \dots,\ \varphi^{(m_k-1)}(\lambda_k) = 0, \quad k = \overline{1, s}.$$

Но тогда $\varphi(\lambda)$ делится нацело на $(\lambda - \lambda_k)^{m_k}$, $k = \overline{1, s}$, а, значит, делится нацело и на $\psi(\lambda)$. Но многочлен степени, меньшей m, может делиться нацело на многочлен степени m только тогда, когда он тождественно равен 0, т.е. $\varphi(\lambda) = 0$, откуда $r_1(\lambda) = r_2(\lambda)$.

имеется единственный многочлен $r(\lambda)$ степени, меньшей m. Этот многочлен однозначно определяется интерполяционными условиями

$$r(\lambda_k)=f(\lambda_k),\ r'(\lambda_k)=f'(\lambda_k),\ \ldots,\ r^{(m_k-1)}(\lambda_k)=f^{(m_k-1)}(\lambda_k),\quad k=\overline{1,s}, \tag{3.1.6}$$

и называется **интерполяционным многочленом Лагранжа-Сильвестра для функции $f(\lambda)$ на спектре матрицы A**. Определение 3.1.2 можно переформулировать в виде

Определение 3.1.2$'$. Пусть $f(\lambda)$ — функция, определенная на спектре матрицы A, а $r(\lambda)$ — соответствующий интерполяционный многочлен Лагранжа-Сильвестра. Тогда по определению

$$f(A)=r(A).$$

Замечание 3.1.1. Если матрица A приводится к диагональному виду, т.е. последний ее инвариантный множитель $e_n(\lambda)$, совпадающий с мирималъным многочленом $\psi(\lambda)$, не имеет кратных корней, то для того, чтобы $f(A)$ имела смысл, достаточно, чтобы функция $f(\lambda)$ была определена в характеристических точках $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_s$. Если же $\psi(\lambda)$ имеет кратные корни, то в некоторых характеристических точках $f(\lambda)$ должна иметь производные до определенного порядка.

Пример 3.1.1. Рассмотрим жорданову клетку n-го порядка

$$J=\begin{pmatrix}\lambda_0 & 1 & & & & \\ & \lambda_0 & 1 & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & 1 \\ & & & & & \lambda_0\end{pmatrix}$$

Как мы помним, для нее все инвариантные множители равны 1, кроме последнего, который равен $e_n(\lambda)=(\lambda-\lambda_0)^n$. Значит, минимальный многочлен равен $\psi(\lambda)=(\lambda-\lambda_0)^n$. Значение $f(\lambda)$ на спектре J есть

$$f(\lambda_0),\ f'(\lambda_0),\ \ldots,\ f^{(n-1)}(\lambda_0).$$

Интерполяционный многочлен Лагранжа-Сильвестра определяется интерполяционными условиями

$$r(\lambda_0)=f(\lambda_0),\ r'(\lambda_0)=f'(\lambda_0),\ \ldots,\ r^{(n-1)}(\lambda_0)=f^{(n-1)}(\lambda_0). \tag{3.1.7}$$

Если взять

$$r(\lambda)=f(\lambda_0)+\frac{f'(\lambda_0)}{1!}(\lambda-\lambda_0)+\ldots+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(\lambda-\lambda_0)^{n-1},$$

то степень $r(\lambda)<n$ и условия (3.1.7), очевидно, выполняются. Значит, это $r(\lambda)$ и есть интерполяционный полином Лагранжа-Сильвестра. Поэтому

$$f(J)=r(J)=f(\lambda_0)E+\frac{f'(\lambda_0)}{1!}(J-\lambda_0E)+\ldots+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(J-\lambda_0E)^{n-1}=$$

$$=\begin{pmatrix}f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdot & \cdot & \cdot & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ & f(\lambda_0) & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \frac{f'(\lambda_0)}{1!} \\ & & & & & f(\lambda_0)\end{pmatrix}.$$

1. *Если* $\lambda_1,\ \lambda_2,\ \dots,\ \lambda_n$ *— собственные значения матрицы* n*-го порядка* A*, то* $f(\lambda_1)$, $f(\lambda_2),\ \dots,\ f(\lambda_n)$ *— полная система собственных значений матрицы* $f(A)$*.*

Доказательство. Для многочленов это свойство доказано раньше. Если функция $f(\lambda)$ определена на спектре матрицы A и $r(\lambda)$ — ее интерполяционный многочлен, то $f(A) = r(A)$, $f(\lambda_1) = r(\lambda_1)$, $f(\lambda_2) = r(\lambda_2)$, $\dots$, $f(\lambda_n) = r(\lambda_n)$. Так как $r(\lambda)$ — многочлен, то по доказанному ранее $r(\lambda_1),\ r(\lambda_2),\ \dots,\ r(\lambda_n)$ — полная система собственных значений $r(A)$. Значит, $f(\lambda_1),\ f(\lambda_2),\ \dots,\ f(\lambda_n)$ — полная система собственных значений матрицы $f(A)$.

2. *Если матрицы* A *и* B *подобны и матрица* T *преобразует* A *в* B*:*

$$B = T^{-1}AT,$$

то матрицы $f(A)$ *и* $f(B)$ *подобны и та же матрица* T *преобразует* $f(A)$ *в* $f(B)$*:*

$$f(B) = T^{-1}f(A)T.$$

Доказательство. Пусть $f(\lambda) = a_0\lambda^p + a_1\lambda^{p-1} + \ldots + a_p$ — многочлен. Если $B = T^{-1}AT$, то $B^k = (T^{-1}AT)^k = T^{-1}ATT^{-1}AT\ldots T^{-1}AT = T^{-1}A^kT$, но тогда

$$f(B) = a_0B^p + a_1B^{p-1} + \ldots + a_pE = a_0T^{-1}A^pT + a_1T^{-1}A^{p-1}T + \ldots + a_pT^{-1}T =$$

$$= T^{-1}(a_0A^p + a_1A^{p-1} + \ldots + a_pE)T = T^{-1}f(A)T,$$

т.е. для случая, когда f — многочлен, свойство доказано.

Если f определена на спектре A, то поскольку подобные матрицы имеют одинаковый набор собственных значений, то f определена и на спектре B. Кроме того, подобные матрицы имеют одинаковые инвариантные множители, следовательно, одинаковые минимальные многочлены $\psi(\lambda) = e_n(\lambda)$. Значит, $f(\lambda)$ принимают одни и те же значения на спектрах A и B. Поэтому существует интерполяционный многочлен $r(\lambda)$ такой, что $f(A) = r(A)$, $f(B) = r(B)$. По доказанному $r(B) = T^{-1}r(A)T$, следовательно, $f(B) = T^{-1}f(A)T$.

Определение 3.2.1. Матрица вида

$$A = \begin{pmatrix} \boxed{A_1} & & & & & \\ & \boxed{A_2} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \boxed{A_u} \end{pmatrix},$$

где $A_1,\ A_2,\ \dots,\ A_u$ — квадратные матрицы, называется **квазидиагональной** и обозначается $A = \{A_1,\ A_2,\ \dots,\ A_u\}$.

3. *Если* A *— квазидиагональная матрица:*

$$A = \{A_1,\ A_2,\ \dots,\ A_u\},$$

то $f(A)$ *— квазидиагональная матрица вида*

$$f(A) = \{f(A_1),\ f(A_2),\ \dots,\ f(A_u)\}.$$

Доказательство. Докажем сначала это свойство для многочлена, т.е. докажем, что если $P(\lambda) = a_0\lambda^p + a_1\lambda^{p-1} + \ldots + a_p$ — многочлен, то $P(A) = \{P(A_1),\ P(A_2),\ \dots,\ P(A_u)\}$.

$$A^2 = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_u} \end{pmatrix} \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_u} \end{pmatrix} =$$

$$= \begin{pmatrix} \boxed{A_1^2} & & & \\ & \boxed{A_2^2} & & \\ & & \ddots & \\ & & & \boxed{A_u^2} \end{pmatrix} = \{A_1^2,\ A_2^2,\ \ldots,\ A_u^2\}.$$

По индукции для любого натурального l

$$A^l = \{A_1^l,\ A_2^l,\ \ldots,\ A_u^l\}.$$

Поэтому, если $E_1,\ E_2,\ \ldots,\ E_s$ — единичные матрицы тех же порядков, что и у матриц $A_1,\ A_2,\ \ldots,\ A_u$ соответственно, то

$$P(A) = a_0 A^p + a_1 A^{p-1} + \ldots + a_p E =$$

$$= a_0 \begin{pmatrix} \boxed{A_1^p} & & & \\ & \boxed{A_2^p} & & \\ & & \ddots & \\ & & & \boxed{A_u^p} \end{pmatrix} + a_1 \begin{pmatrix} \boxed{A_1^{p-1}} & & & \\ & \boxed{A_2^{p-1}} & & \\ & & \ddots & \\ & & & \boxed{A_u^{p-1}} \end{pmatrix} + \ldots +$$

$$+ a_p \begin{pmatrix} \boxed{E_1} & & & \\ & \boxed{E_2} & & \\ & & \ddots & \\ & & & \boxed{E_u} \end{pmatrix} = \begin{pmatrix} \boxed{P(A_1)} & & & \\ & \boxed{P(A_2)} & & \\ & & \ddots & \\ & & & \boxed{P(A_u)} \end{pmatrix} =$$

$$= \{P(A_1),\ P(A_2),\ \ldots,\ P(A_u)\}.$$

Итак, для многочленов свойство 3 доказано.

Обозначим через $r(\lambda)$ интерполяционный многочлен Лагранжа-Сильвестра для функции $f(\lambda)$ на спектре матрицы A. По доказанному

$$f(A) = r(A) = \{r(A_1),\ r(A_2),\ \ldots,\ r(A_u)\}. \tag{3.2.1}$$

Если $\psi(\lambda)$ — минимальный многочлен для A, то $\psi(A) = \{\psi(A_1),\ \psi(A_2),\ \ldots,\ \psi(A_u)\} = 0$, откуда $\psi(A_1) = 0, \psi(A_2) = 0, \ldots, \psi(A_u) = 0$, т.е. ψ — аннулирующий многочлен для всех матриц $A_1,\ A_2,\ \ldots,\ A_u$. Множество всех собственных значений каждой матрицы A_l входит в множество собственных значений A (с учетом кратностей). Так как многочлен ψ — аннулирующий для A_l, то из равенства

$$f(\Lambda_A) = r(\Lambda_A),$$

т.е.

$$f(\lambda_k) = r(\lambda_k),\ f'(\lambda_k) = r'(\lambda_k),\ \ldots,\ f^{(m_k-1)}(\lambda_k) = r^{(m_k-1)}(\lambda_k), \quad k = \overline{1, s},$$

$$f(\Lambda_{A_1}) = r(\Lambda_{A_1}),\ f(\Lambda_{A_2}) = r(\Lambda_{A_2}),\ \dots,\ f(\Lambda_{A_u}) = r(\Lambda_{A_u})$$

(ведь m_k для A_l определяется по минимальному многочлену для A_l, который делит ψ, и, следовательно, m_k для A_l не превосходит m_k для A). Поэтому

$$f(A_1) = r(A_1),\ f(A_2) = r(A_2),\ \dots,\ f(A_u) = r(A_u),$$

и равенство (3.2.1) запишется в виде

$$f(A) = \{f(A_1),\ f(A_2),\ \dots,\ f(A_u)\}.$$

Пример 3.2.1. Диагональная матрица является частным случаем квазидиагональной матрицы с блоками вдоль главной диагонали, являющимися квадратными матрицами 1-го порядка, т.е. числами. Поэтому, если матрица приводится к диагональному виду (матрица простой структуры), т.е.

$$A = T\{\lambda_1,\ \lambda_2,\ \dots,\ \lambda_u\}T^{-1},$$

то

$$f(A) = T\{f(\lambda_1),\ f(\lambda_2),\ \dots,\ f(\lambda_u)\}T^{-1}.$$

При этом $f(A)$ имеет смысл, если функция $f(\lambda)$ определена в точках $\lambda_1,\ \lambda_2,\ \dots,\ \lambda_u$.

Пример 3.2.2. Если матрица приводится к жордановой нормальной форме

$$A = TJT^{-1},$$

где

$$J = \begin{pmatrix} \begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_1 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} \lambda_u & 1 & & & \\ & \lambda_u & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_u \end{pmatrix} \end{pmatrix},$$

причем вдоль диагонали идут жордановы клетки порядков $\nu_1,\ \dots,\ \nu_u$ соответственно, то, используя свойство 3 и результат примера 3.1.1, получим

$$f(A) = Tf(J)T^{-1},$$

$$f(J)=\begin{pmatrix} \begin{pmatrix} f(\lambda_1) & \frac{f'(\lambda_1)}{1!} & \cdots & \frac{f^{(\nu_1-1)}(\lambda_1)}{(n-1)!} \\ & f(\lambda_1) & \cdots & \cdot \\ & & \cdots & \cdot \\ & & \ddots & \cdot \\ & & & \frac{f'(\lambda_1)}{1!} \\ & & & f(\lambda_1) \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} f(\lambda_u) & \frac{f'(\lambda_u)}{1!} & \cdots & \frac{f^{(\nu_u-1)}(\lambda_u)}{(n-1)!} \\ & f(\lambda_u) & \cdots & \cdot \\ & & \cdots & \cdot \\ & & \ddots & \cdot \\ & & & \frac{f'(\lambda_u)}{1!} \\ & & & f(\lambda_u) \end{pmatrix} \end{pmatrix}.$$

3.3. Интерполяционный многочлен Лагранжа-Сильвестра.

1. Пусть характеристический многочлен $|\lambda E - A|$ не имеет кратных корней и пусть $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все его корни. Тогда все показатели m_k в (2.2.7) равны единице и минимальный многочлен матрицы A есть

$$\psi(\lambda) = |\lambda E - A| = (\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n).$$

Интерполяционные условия имеют вид

$$r(\lambda_k) = f(\lambda_k), \quad k=\overline{1,n}. \tag{3.3.1}$$

В этом случае $r(\lambda)$ — обычный интерполяционный многочлен Лагранжа для функции $f(\lambda)$ в точках $\lambda_1, \lambda_2, \ldots, \lambda_n$:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda-\lambda_1)\ldots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\ldots(\lambda-\lambda_n)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

Очевидно, он удовлетворяет интерполяционным условиям (3.3.1) и его степень меньше n. Согласно определению $3.1.2'$

$$f(A) = r(A) = \sum_{k=1}^{n} \frac{(A-\lambda_1 E)\ldots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\ldots(A-\lambda_n E)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

2.Допустим теперь, что характеристический многочлен имеет кратные корни, но минимальный многочлен, являющийся делителем характеристического, имеет только простые корни:

$$\psi(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m).$$

В этом случае, как и в предыдущем, все показатели m_k в (2.2.7) равны единице, и интерполяционные условия (3.3.1) принимают вид

$$r(\lambda_k) = f(\lambda_k), \quad k=\overline{1,m}.$$

$$r(\lambda) = \sum_{k=1}^{m} \frac{(\lambda-\lambda_1)\dots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\dots(\lambda-\lambda_m)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_m)} f(\lambda_k),$$

и, следовательно, функция от матрицы найдется как

$$f(A) = r(A) = \sum_{k=1}^{m} \frac{(A-\lambda_1 E)\dots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\dots(A-\lambda_m E)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_m)} f(\lambda_k).$$

3. Рассмотрим общий случай:

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\dots(\lambda-\lambda_s)^{m_s}, \quad m_1+m_2+\ldots+m_s = m.$$

Представим функцию $\frac{r(\lambda)}{\psi(\lambda)}$, являющуюся правильной дробью (ведь степень интерполяционного многочлена Лагранжа-Сильвестра $r(\lambda)$ меньше степени m минимального многочлена $\psi(\lambda)$), в виде суммы простых дробей:

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \Big[\frac{\alpha_{k1}}{(\lambda-\lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda-\lambda_k)^{m_k-1}} + \ldots + \frac{\alpha_{k,m_k}}{(\lambda-\lambda_k)}\Big], \qquad (3.3.2)$$

где $\alpha_{kj} \quad (j=\overline{1,m_k};\ \ k=\overline{1,s})$ — некоторые числа. Введем в рассмотрение многочлены

$$\psi^k(\lambda) = \frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}.$$

Для определения числителей простых дробей α_{kj} умножим обе части (3.3.2) на $(\lambda-\lambda_k)^{m_k}$ и воспользуемся введенным обозначением:

$$\frac{r(\lambda)}{\psi^k(\lambda)} = \alpha_{k1} + \alpha_{k2}(\lambda-\lambda_k) + \ldots + \alpha_{k,m_k}(\lambda-\lambda_k)^{m_k-1} + (\lambda-\lambda_k)^{m_k}\rho(\lambda), \quad k=\overline{1,s},$$

где $\rho(\lambda)$ — рациональная функция, не обращающаяся в ∞ при $\lambda=\lambda_k$. Отсюда

$$\alpha_{kj} = \frac{1}{(j-1)!}\Big[\frac{r(\lambda)}{\psi^k(\lambda)}\Big]^{(j-1)}_{\lambda=\lambda_k}, \quad j=\overline{1,m_k}, \quad k=\overline{1,s}. \qquad (3.3.3)$$

Формулы (3.3.3) показывают, что числители α_{kj} в правой части (3.3.2) выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, а эти значения нам известны: они равны соответствующим значениям функции $f(\lambda)$ и ее производных. Поэтому

$$\alpha_{kj} = \frac{1}{(j-1)!}\Big[\frac{f(\lambda)}{\psi^k(\lambda)}\Big]^{(j-1)}_{\lambda=\lambda_k}, \quad j=\overline{1,m_k}, \quad k=\overline{1,s}. \qquad (3.3.4)$$

После того как все α_{kj} найдены, $r(\lambda)$ можно найти из следующей формулы, которая получается умножением обеих частей равенства (3.3.2) на $\psi(\lambda)$:

$$r(\lambda) = \sum_{k=1}^{s} \Big[\alpha_{k1} + \alpha_{k2}(\lambda-\lambda_k) + \ldots + \alpha_{k,m_k}(\lambda-\lambda_k)^{m_k-1}\Big]\psi^k(\lambda). \qquad (3.3.5)$$

В этой формуле выражение в квадратных скобках, стоящее в качестве множителя перед $\psi^k(\lambda)$, в силу (3.3.4) равно сумме первых m_k членов разложения Тейлора по степеням $\lambda-\lambda_k$ для функции $\frac{f(\lambda)}{\psi^k(\lambda)}$.

$f(\lambda)$ или какой-либо ее производной, представим $r(\lambda)$ в виде

$$r(\lambda)=\sum_{k=1}^{s}\Big[f(\lambda_k)\varphi_{k1}(\lambda)+f'(\lambda_k)\varphi_{k2}(\lambda)+\ldots+f^{(m_k-1)}(\lambda_k)\varphi_{k,m_k}(\lambda)\Big]. \tag{3.3.6}$$

Здесь $\varphi_{kj}(\lambda)\quad (j=\overline{1,m_k},\quad k=\overline{1,s})$ — легко вычисляемые многочлены от λ степени, меньшей m. Эти многочлены вполне определяются заданием $\psi(\lambda)$ и не зависят от выбора функции $f(\lambda)$. Число этих многочленов равно числу значений функции на спектре матрицы A, т.е. равно m, где m — степень минимального многочлена.

Пусть все значения f на спектре матрицы A равны нулю, за исключением одного $f^{(j-1)}(\lambda_k)=$ $=1$, тогда для такой функции (3.3.6) примет вид

$$r(\lambda)=\varphi_{kj}(\lambda).$$

Таким образом, *функция $\varphi_{kj}(\lambda)$ — интерполяционный многочлен Лагранжа-Сильвестра для функции, у которой все значения на спектре матрицы A равны нулю, за исключением одного $f^{(j-1)}(\lambda_k)$, равного 1.*

Из формулы (3.3.6) следует **основная формула для $f(A)$**:

$$f(A)=\sum_{k=1}^{s}\Big[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ldots+f^{(m_k-1)}(\lambda_k)Z_{k,m_k}\Big]. \tag{3.3.7}$$

где

$$Z_{kj}=\varphi_{kj}(A),\quad j=\overline{1,m_k},\quad k=\overline{1,s}.$$

Матрицы Z_{kj} вполне определяются заданием матрицы A и не зависят от выбора функции $f(\lambda)$. В правой части (3.3.7) функция $f(\lambda)$ представлена только своими значениями на спектре матрицы A.

Матрицы $Z_{kj}\quad (j=\overline{1,m_k},\quad k=\overline{1,s})$ называются **составляющими матрицами** или **компонентами матрицы A**.

Занумеровав все элементы квадратной матрицы порядка n в некотором фиксированном (например, лексикографическом) порядке, мы превращаем множество квадратных матриц n-го порядка над полем $\mathbb{C}$ в линейное n^2-мерное пространство над этим полем с естественными операциями сложения векторов и умножения вектора на комплексное число. Тогда составляющие матрицы можно рассматривать как вектора этого пространства.

Теорема 3.3.1. *Компоненты матрицы A линейно независимы.*

Д о к а з а т е л ь с т в о. Действительно, пусть линейная комбинация m векторов Z_{kj} равна нулю:

$$\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}Z_{kj}=\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}\varphi_{kj}(A)=0. \tag{3.3.8}$$

Определим интерполяционный многочлен $r(\lambda)$ с помощью m условий

$$r^{(j-1)}(\lambda_k)=c_{kj}\quad (j=\overline{1,m_k},\quad k=\overline{1,s}). \tag{3.3.9}$$

Тогда согласно (3.3.6)

$$r(\lambda)=\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}\varphi_{kj}(\lambda). \tag{3.3.10}$$

Сравнивая (3.3.8) с (3.3.10), получим

$$r(A)=0.$$

многочлена. Так как интерполяционный многочлен оказался аннулирующим для A, то он делится на минимальный многочлен $\psi(\lambda)$ степени m. Многочлен степени, меньшей m, делится на многочлен степени m только тогда, когда он равен нулю. Следовательно, $r(\lambda) = 0$, но тогда из (3.3.9)

$$c_{kj} = 0, \quad j = \overline{1, m_k}, \quad k = \overline{1, s}. \qquad (3.3.11)$$

Итак, из равенства нулю линейной комбинации (3.3.8) следует равенство нулю коэффициентов этой линейной комбинации (3.3.11). Это означает, что вектора Z_{kj} линейно независимы.

Заметим, что из линейной независимости составляющих матриц Z_{kj} следует, что ни одна из этих компонент не равна нулю. Отметим также, что любые две из компонент перестановоцны между собой и с матрицей A, поскольку все они — многочлены от A.

3.4. Представление функций от матриц рядами.

Пусть дана квадратная матрица $A = (a_{ij})$ порядка n с минимальным многочленом

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}, \quad m_1 + m_2 + \dots + m_s = m,$$

и пусть функции $f(\lambda)$, $f_1(\lambda)$, $f_2(\lambda)$, $\dots$, $f_p(\lambda)$, $\dots$ определены на спектре матрицы A.

Определение 3.4.1. Говорят, что **последовательность функций $\{f_p(\lambda)\}$ стремится при $p \to \infty$ к некоторому пределу на спектре матрицы A**, если существуют пределы

$$\lim_{p\to\infty} f_p(\lambda_k), \ \lim_{p\to\infty} f'_p(\lambda_k), \ \dots, \ \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k), \quad k = \overline{1, s}.$$

Говорят, что **последовательность функций $\{f_p(\lambda)\}$ стремится при $p \to \infty$ к функции $f(\lambda)$ на спектре матрицы A**, и пишут

$$\lim_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A),$$

если

$$\lim_{p\to\infty} f_p(\lambda_k) = f(\lambda_k), \ \lim_{p\to\infty} f'_p(\lambda_k) = f'(\lambda_k), \ \dots, \ \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k), \quad k = \overline{1, s}.$$

Если рассматривать квадратные матрицы n-го порядка как вектора в n^2-мерном линейном пространстве $\mathbb{E}_{n^2}$, то сходимость последовательности матриц естественно определить как покоординатную сходимость (позже мы выясним, что такая сходимость эквивалентна сходимости по любой матричной норме). Таким образом, мы будем исходить из следующего определения.

Определение 3.4.2. Будем говорить, что **последовательность матриц $\{B_p = (b^p_{ij})\}$ имеет предел при $p \to \infty$**, если существуют пределы

$$\lim_{p\to\infty} b^p_{ij}, \quad i, j = \overline{1, n}.$$

Будем говорить, что **последовательность матриц $\{B_p = (b^p_{ij})\}$ сходится к матрице $B = (b_{ij})$ при $p \to \infty$**, если

$$\lim_{p\to\infty} b^p_{ij} = b_{ij}, \quad i, j = \overline{1, n}.$$

$$f(A) = \sum_{k=1}^{s} \Big[f(\lambda_k) Z_{k1} + f'(\lambda_k) Z_{k2} + \ldots + f^{(m_k-1)}(\lambda_k) Z_{k,m_k} \Big], \tag{3.4.1}$$

выражает $f(A)$ через значения $f(\lambda)$ на спектре матрицы A. В силу линейной независимости матриц Z_{kj} из этой формулы следует, что все функции $f(A)$, определенные на спектре A, при заданной фиксированной матрице A образуют m-мерное линейное подпространство в $\mathbb{E}_{n^2}$ с базисом Z_{kj} $(j = \overline{1, m_k}, \quad k = \overline{1, s})$. В этом базисе вектор $f(A)$ имеет своими координатами m значений функции $f(\lambda)$ на спектре матрицы A.

Эти соображения делают совершенно прозрачной следующую теорему.

Теорема 3.4.1. *Для того, чтобы последовательность матриц* $\{f_p(A)\}$ *при* $p \to \infty$ *сходилась к некоторому пределу, необходимо и достаточно, чтобы последовательность функций* $\{f_p(\lambda)\}$ *при* $p \to \infty$ *на спектре матрицы* A *стремилась к пределу. При этом*

$$\lim_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A)$$

тогда и только тогда, когда

$$\lim_{p\to\infty} f_p(A) = f(A).$$

Д о к а з а т е л ь с т в о. *Достаточность.* Пусть значения $f_p(\lambda)$ на спектре матрицы A при $p \to \infty$ сходятся к предельным значениям. Тогда из основной формулы (3.4.1), записанной для функций последовательности f_p:

$$f_p(A) = \sum_{k=1}^{s} \Big[f_p(\lambda_k) Z_{k1} + f_p'(\lambda_k) Z_{k2} + \ldots + f_p^{(m_k-1)}(\lambda_k) Z_{k,m_k} \Big], \tag{3.4.2}$$

следует существование предела $\lim\limits_{p\to\infty} f_p(A)$. Если к тому же $\lim\limits_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A)$, то, переходя к пределу в правой части (3.4.2), мы получим правую часть (3.4.1). Это означает, что $\lim\limits_{p\to\infty} f_p(A) = f(A)$.

Необходимость. Обратно, пусть существует $\lim\limits_{p\to\infty} f_p(A)$. Ранее мы доказали, что если рассматривать матрицы как вектора в n^2-мерном линейном пространстве, то m векторов

$$Z_{kj}, \quad j = \overline{1, m_k}, \quad k = \overline{1, s},$$

линейно независимы. Зафиксируем некоторую нумерацию элементов в матрицах (например, сначала занумеруем элементы первой строки, затем второй и т.д., наконец, n-й). Поскольку $m < n^2$, то матрица, составленная из векторов Z_{11}, Z_{12}, ..., Z_{1,m_1}, Z_{21}, Z_{22}, ..., Z_{2,m_2}, ..., Z_{21}, Z_{s1}, $Z_{s2}, \ldots$, Z_{s,m_s} как столбцов:

$$\begin{pmatrix} Z_{11}^{(1)} & \ldots & Z_{1,m_1}^{(1)} & Z_{21}^{(1)} & \ldots & Z_{2,m_2}^{(1)} & \ldots & Z_{s1}^{(1)} & \ldots & Z_{s,m_s}^{(1)} \\ Z_{11}^{(2)} & \ldots & Z_{1,m_1}^{(2)} & Z_{21}^{(2)} & \ldots & Z_{2,m_2}^{(2)} & \ldots & Z_{s1}^{(2)} & \ldots & Z_{s,m_s}^{(2)} \\ \cdot & \ldots & \cdot & \cdot & \ldots & \cdot & \ldots & \cdot & \ldots & \cdot \\ Z_{11}^{(n^2)} & \ldots & Z_{1,m_1}^{(n^2)} & Z_{21}^{(n^2)} & \ldots & Z_{2,m_2}^{(n^2)} & \ldots & Z_{s1}^{(n^2)} & \ldots & Z_{s,m_s}^{(n^2)} \end{pmatrix}$$

имеет ранг m. Поэтому существует минор m-го порядка, стоящий, например, в строках i_1, i_2, ..., i_m, который отличен от нуля. Рассматривая матрицы $f_p(A)$ и Z_{kj} как вектора в n^2-мерном линейном пространстве, а равенство (3.4.2) как равенство векторов, приравняем координаты этих векторов с номерами i_1, i_2, ..., i_m. Получим

$$f_p^{(i_l)}(A) = \sum_{k=1}^{s} \Big[f_p(\lambda_k) Z_{k1}^{(i_l)} + f_p'(\lambda_k) Z_{k2}^{(i_l)} + \ldots + f_p^{(m_k-1)}(\lambda_k) Z_{k,m_k}^{(i_l)} \Big], \quad l = \overline{1, m}.$$

Относительно m значений $f(\lambda)$ на спектре матрицы A (т.е. $f_p^{(j)}(\lambda_k)$, $j = \overline{0, m_k - 1}$, $k =$ $= \overline{1,s}$) это есть система m линейных уравнений с m неизвестными. Так как ее определитель совпадает с рассмотренным выше минором, то он отличен от нуля. Следовательно, система имеет единственное решение, и значения $f(\lambda)$ на спектре матрицы A однозначно и линейно выразятся через m значений $f_p^{(i_l)}(A)$, $l = \overline{1,m}$, т.е. через какие-то m элементов матрицы $f_p(A)$:

$$f_p^{(j)}(\lambda_k) = \sum_{l=1}^{m} \beta_{i_l} f_p^{(i_l)}(A), \quad j = \overline{0, m_k - 1}, \ \ k = \overline{1,s}.$$

Отсюда следует существование предела $\lim\limits_{p\to\infty} f_p(\Lambda_A)$ и тот факт, что из равенства $\lim\limits_{p\to\infty} f_p(A) =$ $= f(A)$ следует равенство $\lim\limits_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A)$.

Эта формула подчеркивает естественность и общность данного нами определения функции от матрицы, $f(A)$ всегда получается предельным переходом из $g_p(A)$ при $p \to \infty$, если только последовательность многочленов $g_p(\lambda)$ сходится к $f(\lambda)$ на спектре матрицы A.

Определение 3.4.3. Говорят, что **ряд** $\sum\limits_{p=0}^{\infty} u_p(\lambda)$ **сходится на спектре матрицы** A **к функции** $f(\lambda)$, и пишут

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A), \tag{3.4.3}$$

если все фигурирующие здесь функции определены на спектре матрицы A и имеют место равенства

$$f(\lambda_k) = \sum_{p=0}^{\infty} u_p(\lambda_k), \ \ f'(\lambda_k) = \sum_{p=0}^{\infty} u_p'(\lambda_k), \ \ \ldots, \ \ f^{(m_k-1)}(\lambda_k) = \sum_{p=0}^{\infty} u_p^{(m_k-1)}(\lambda_k), \ \ k = \overline{1,s}.$$

Другими словами, если положить

$$S_p(\lambda) = \sum_{q=0}^{p} u_q(\lambda), \quad p = 0, 1, 2, \ldots,$$

то (3.4.3) равносильно равенству

$$f(\Lambda_A) = \lim_{p\to\infty} S_p(\Lambda_A).$$

С использованием этого факта теорема 3.4.1 может быть переформулирована следующим образом.

Теорема 3.4.2. *Для того, чтобы ряд* $\sum\limits_{p=0}^{\infty} u_p(A)$ *сходился к некоторой матрице, необходимо и достаточно, чтобы ряд* $\sum\limits_{p=0}^{\infty} u_p(\lambda)$ *сходился на спектре матрицы* A. *При этом*

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A)$$

тогда и только тогда, когда

$$f(A) = \sum_{p=0}^{\infty} u_p(A).$$

Пусть дан степенной ряд с кругом сходимости $|\lambda - \lambda_0| < R$ и суммой $f(\lambda)$:

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p \quad (|\lambda - \lambda_0| < R). \tag{3.4.4}$$

сходимости, то ряд (3.4.4) сходится на спектре любой матрицы, собственные значения которой попадают внутрь круга сходимости.

Таким образом, имеет место

Теорема 3.4.3. *Если функция $f(\lambda)$ разлагается в степенной ряд в круге $|\lambda - \lambda_0| < R$,*

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p,$$

то

$$f(A) = \sum_{p=0}^{\infty} \alpha_p (A - \lambda_0 E)^p$$

для любой матрицы A, собственные значения которой лежат внутри круга сходимости $|\lambda - \lambda_0| < R$.

Так как во всей комплексной плоскости

$$e^{\lambda} = \sum_{p=0}^{\infty} \frac{\lambda^p}{p!}; \quad \cos\lambda = \sum_{p=0}^{\infty} (-1)^p \frac{\lambda^{2p}}{(2p)!}; \quad \sin\lambda = \sum_{p=0}^{\infty} (-1)^p \frac{\lambda^{2p+1}}{(2p+1)!},$$

то из теоремы 3.4.3 вытекают следующие разложения:

$$e^{A} = \sum_{p=0}^{\infty} \frac{A^p}{p!}; \quad \cos A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p}}{(2p)!}; \quad \sin A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p+1}}{(2p+1)!},$$

справедливые для любой матрицы A. Кроме того, поскольку

$$(1 - \lambda)^{-1} = \frac{1}{1 - \lambda} = \sum_{p=0}^{\infty} \lambda^p, \quad |\lambda| < 1,$$

то

$$(E - A)^{-1} = \sum_{p=0}^{\infty} A^p, \quad |\lambda_k| < 1, \quad k = \overline{1, s}.$$

Можно доказать следующее свойство функции от матрицы.

4. *Пусть $P(u_1, u_2, \ldots, u_l)$ — многочлен от l переменных $u_1, u_2, \ldots, u_l$; $f_1(\lambda),\ f_2(\lambda),\ \ldots,\ f_l(\lambda)$ — функции от λ, определенные на спектре матрицы A и $p(\lambda) \equiv P(f_1(\lambda),\ f_2(\lambda),\ \ldots,\ f_l(\lambda))$. Тогда из*

$$p(\Lambda_A) = 0 \tag{3.4.5}$$

следует

$$P(f_1(A),\ f_2(A),\ \ldots,\ f_l(A)) = 0.$$

Доказательство. Пусть $r_1(\lambda),\ r_2(\lambda),\ \ldots,\ r_l(\lambda)$ — интерполяционные многочлены Лагранжа-Сильвестра для $f_1(\lambda),\ f_2(\lambda),\ \ldots,\ f_l(\lambda)$. Положим

$$h(\lambda) = P(r_1(\lambda),\ r_2(\lambda),\ \ldots,\ r_l(\lambda)).$$

Очевидно, $h(\lambda)$ — многочлен, значения которого на спектре матрицы A совпадают со значениями $p(\lambda)$, т.е. $p(\lambda_A) = h(\lambda_A)$. Однако из (3.4.5) следует, что $h(\lambda_A) = 0$. Отсюда вытекает, что

$$P(f_1(A),\ f_2(A),\ \ldots,\ f_l(A)) = P(r_1(A),\ r_2(A),\ \ldots,\ r_l(A)) = h(A).$$

Согласно свойству 4 из тождества $\sin^2\lambda + \cos^2\lambda = 1$ следует, что для любой матрицы A

$$\sin^2 A + \cos^2 A = E$$

(надо взять $P(u_1, u_2) = u_1^2 + u_2^2 - 1, \quad f_1(\lambda) = \sin\lambda, \quad f_2(\lambda) = \cos\lambda$).

Беря в свойстве 4 $P(u_1, u_2) = u_1 u_2 - 1, \quad f_1(\lambda) = e^{\lambda}, \quad f_2(\lambda) = e^{-\lambda}$, получим, что для любой матрицы A

$$e^A e^{-A} = E,$$

следовательно,

$$e^{-A} = (e^A)^{-1}.$$

Беря в свойстве 4 $P(u_1, u_2, u_3) = u_1 - u_2 - iu_3, \quad f_1(\lambda) = e^{i\lambda}, \quad f_2(\lambda) = \cos\lambda, \quad f_3(\lambda) = \sin\lambda$, для любой матрицы A будем иметь

$$e^{iA} = \cos A + i\ \sin A.$$

Пусть $f(\lambda) = \frac{1}{\lambda}$ и $A = (a_{ik})$ — невырожденная матрица. Так как собственные значения невырожденной матрицы отличны от нуля, то $f(\lambda)$ определена на спектре матрицы A. Поэтому в равенстве $\lambda f(\lambda) = 1$ можно заменить λ на A:

$$Af(A) = E,$$

т.е.

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

Обозначая через $r(\lambda)$ интерполяционный многочлен для функции $\frac{1}{\lambda}$, мы можем представить обратную матрицу A^{-1} в виде многочлена от данной матрицы A:

$$A^{-1} = r(A).$$

Рассмотрим рациональную функцию $\rho(\lambda) = \frac{g(\lambda)}{h(\lambda)}$, где $g(\lambda)$ и $h(\lambda)$ — взаимно простые многочлены относительно λ. Эта функция определена на спектре матрицы A тогда и только тогда, когда собственные значения матрицы A не являются корнями многочлена $h(\lambda)$, т.е. когда $|h(\lambda)| \neq 0$. При выполнении этого условия мы можем в тождестве $\rho(\lambda)h(\lambda) = g(\lambda)$ заменить λ на A:

$$\rho(A)h(A) = g(A).$$

Отсюда

$$\rho(A) = g(A)[h(A)]^{-1} = [h(A)]^{-1}g(A).$$

В заключение, отметим, что у нас имеется два принципиально разных способа вычисления функции от матрицы:

1. С помощью интерполяционного многочлена Лагранжа-Сильвестра.
2. С помощью приведения матрицы A к жордановой нормальной форме. Если

$$A = TJT^{-1},$$

то

$$f(A) = Tf(J)T^{-1},$$

причем $f(J)$ легко выписывается.

4.1. Нормированные пространства. Векторные нормы.

Пусть $\mathbb{E}_n$ — конечномерное векторное пространство над полем действительных чисел $\mathbb{R}$ или над полем комплексных чисел $\mathbb{C}$, $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$.

Определение 4.1.1. Отображение $|| \quad || : \mathbb{E}_n \to \mathbb{R}^+$, ставящее в соответствие вектору $\mathbf{x} \in \mathbb{E}_n$ число $||\mathbf{x}|| \in \mathbb{R}^+$, называется **нормой**, если оно удовлетворяет следующим свойствам:
1. $||\mathbf{x}|| = 0$ тогда и только тогда, когда $\mathbf{x} = 0$;
2. $||\lambda\mathbf{x}|| = |\lambda|||\mathbf{x}||$, $\lambda \in \mathbb{R}$ (или $\lambda \in \mathbb{C}$);
3. $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$, $\mathbf{x}, \mathbf{y} \in \mathbb{E}_n$.

Векторное пространство с введенной на нем нормой называется **нормированным**.

Свойства нормы — свойства длины вектора. Каждое конечномерное векторное пространство можно превратить в нормированное с помощью введения нескольких норм. На данном векторном пространстве существует бесконечное множество норм, основных и вспомогательных. Имеются три основные нормы (здесь мы фиксируем в $\mathbb{E}_n$ базис, в нем каждый вектор $\mathbf{x}$ имеет координаты $\mathbf{x} = (x_1, x_2, \ldots, x_n)$):

$$||\mathbf{x}||_I = \sum_{i=1}^{n} |x_i| \text{—}\textbf{октаэдрическая норма};$$

$$||\mathbf{x}||_{II} = \sqrt{\sum_{i=1}^{n} |x_i|^2} \text{—}\textbf{сферическая норма};$$

$$||\mathbf{x}||_\infty = max\{|x_1|, |x_2|, \ldots, |x_n|\} \text{—}\textbf{кубическая норма}.$$

Если линейное пространство рассматривается над полем $\mathbb{R}$, то сферическую норму называют также **евклидовой**, а если над полем $\mathbb{C}$, то ее называют **эрмитовой**.

Докажем, что это нормы.

1. Очевидно, $||\mathbf{x}||_I = \sum\limits_{i=1}^{n} |x_i| = 0$ тогда и только тогда, когда $x_i = 0$ при каждом i, т.е. $\mathbf{x} = \mathbf{0}$.

Далее, $||\lambda\mathbf{x}||_I = \sum\limits_{i=1}^{n} |\lambda x_i| = |\lambda| \sum\limits_{i=1}^{n} |x_i| = |\lambda|||\mathbf{x}||_I$.

Наконец, $||\mathbf{x}+\mathbf{y}||_I = \sum\limits_{i=1}^{n} |x_i + y_i| \leq \sum\limits_{i=1}^{n} |x_i| + \sum\limits_{i=1}^{n} |y_i| = ||\mathbf{x}||_1 + ||\mathbf{y}||_I$.

Таким образом, $|| \quad ||_I$ — норма.

2. Очевидно, что $||\mathbf{x}||_{II} = \sqrt{\sum\limits_{i=1}^{n} |x_i|^2} = 0$ тогда и только тогда, когда $x_i = 0$ при каждом i, т.е. $\mathbf{x} = \mathbf{0}$.

Далее, $||\lambda\mathbf{x}||_{II} = \sqrt{\sum\limits_{i=1}^{n} |\lambda x_i|^2} = |\lambda|\sqrt{\sum\limits_{i=1}^{n} |x_i|^2} = |\lambda|||\mathbf{x}||_{II}$.

Наконец, $||\mathbf{x}+\mathbf{y}||_{II} = \sqrt{\sum\limits_{i=1}^{n} |x_i + y_i|^2} \leq \sqrt{\sum\limits_{i=1}^{n} |x_i|^2} + \sqrt{\sum\limits_{i=1}^{n} |y_i|^2} = ||\mathbf{x}||_{II} + ||\mathbf{y}||_{II}$. Последнее неравенство вытекает из неравенства Коши-Буняковского

$$\big(\sum_{i=1}^{n} x_i \bar{y_i}\big)^2 \leq \big(\sum_{i=1}^{n} |x_i|^2\big)\big(\sum_{i=1}^{n} |y_i|^2\big), \tag{4.1.1}$$

длин этих векторов:

$$(\mathbf{x},\mathbf{y}) \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||.$$

Действительно, извлечем квадратные корни из обеих частей неравенства (4.1.1), затем обе части полученного неравенства удвоим и прибавим к ним выражение

$$\sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} |y_i|^2,$$

в результате чего получим

$$\sum_{i=1}^{n} |x_i + y_i|^2 \leq \Big(\sqrt{\sum_{i=1}^{n} |x_i|^2} + \sqrt{\sum_{i=1}^{n} |y_i|^2}\Big)^2.$$

Остается извлечь квадратный корень из обеих частей последнего неравенства.

3. Очевидно, $||\mathbf{x}||_\infty = \max|x_i| = 0$ тогда и только тогда, когда $x_i = 0$ при каждом i, т.е. $\mathbf{x} = \mathbf{0}$.

Далее, $||\lambda\mathbf{x}||_\infty = \max|\lambda x_i| = |\lambda| \max|x_i| = |\lambda| ||\mathbf{x}||_\infty$.

Наконец, $||\mathbf{x}+\mathbf{y}||_\infty = \max|x_i + y_i| \leq \max|x_i| + \max|y_i| = ||\mathbf{x}||_\infty + ||\mathbf{y}||_\infty$.

Таким образом, $||\ \ ||_\infty$ — норма.

Определение 4.1.2. Нормы $||\cdot||_1$ и $||\cdot||_2$ называются **эквивалентными**, если существуют такие действительные числа $\beta \geq \alpha > 0$, что для всех $x \in E_n$

$$\alpha||\mathbf{x}||_1 \leq ||\mathbf{x}||_2 \leq \beta||\mathbf{x}||_1. \tag{4.1.2}$$

Покажем, что определенное отношение действительно является отношением эквивалентности.

1. *Рефлексивность:* $||\mathbf{x}|| \sim ||\mathbf{x}||$. Поскольку $||\mathbf{x}|| = ||\mathbf{x}|| = ||\mathbf{x}||$, то (4.1.2) выполняется с $\alpha = \beta = 1$.

2. *Симметричность:* Если $||\mathbf{x}||_1 \sim ||\mathbf{x}||_2$, то $||\mathbf{x}||_2 \sim ||\mathbf{x}||_1$. Пусть $||\mathbf{x}||_1 \sim ||\mathbf{x}||_2$, т.е. выполняется (4.1.2). Тогда $\frac{1}{\beta}||\mathbf{x}||_2 \leq ||\mathbf{x}||_1 \leq \frac{1}{\alpha}||\mathbf{x}||_2$, т.е. $\alpha_1||\mathbf{x}||_2 \leq ||\mathbf{x}||_1 \leq \beta_1||\mathbf{x}||_2$, где $\alpha_1 = \frac{1}{\beta}$, $\beta_1 = \frac{1}{\alpha}$, причем, в силу того, что $\beta \geq \alpha > 0$, будет выполнено $\beta_1 = \frac{1}{\alpha} \geq \frac{1}{\beta} = \alpha_1 > 0$. Таким образом, $||\mathbf{x}||_2 \sim ||\mathbf{x}||_1$.

3. *Транзитивность:* Если $||\mathbf{x}||_1 \sim ||\mathbf{x}||_2$, а $||\mathbf{x}||_2 \sim ||\mathbf{x}||_3$, то $||\mathbf{x}||_1 \sim ||\mathbf{x}||_3$. Пусть $||\mathbf{x}||_1 \sim ||\mathbf{x}||_2$, $||\mathbf{x}||_2 \sim ||\mathbf{x}||_3$. Согласно определению 4.1.2 это означает, что найдутся постоянные $\beta_1 \geq \alpha_1 > 0$, $\beta_2 \geq \alpha_2 > 0$ такие, что

$$\alpha_1||\mathbf{x}||_2 \leq ||\mathbf{x}||_1 \leq \beta_1||\mathbf{x}||_2,\ \alpha_2||\mathbf{x}||_3 \leq ||\mathbf{x}||_2 \leq \beta_2||\mathbf{x}||_3.$$

Из последних двух неравенств следует, что

$$\alpha_1\alpha_2||\mathbf{x}||_3 \leq ||\mathbf{x}||_1 \leq \beta_1\beta_2||\mathbf{x}||_3,$$

причем $\beta_1\beta_2 \geq \alpha_1\alpha_2 > 0$. Следовательно, $||\mathbf{x}||_1 \sim ||\mathbf{x}||_3$.

Утверждение 4.1.1. *Три основные нормы эквивалентны.*

$$||\mathbf{x}||_\infty = \max_i |x_i| \le ||\mathbf{x}||_1 = \sum_{i=1}^n |x_i| \le n \max_i |x_i| = n||\mathbf{x}||_\infty,$$

то $||\mathbf{x}||_1 \sim ||\mathbf{x}||_\infty$. Поскольку

$$||\mathbf{x}||_\infty = \max_i |x|_i \le ||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^n x_i^2} \le \sqrt{n \max_i x_i^2} \le \sqrt{n} \max_i |x_i| = \sqrt{n}||\mathbf{x}||_\infty,$$

то $||\mathbf{x}||_2 \sim ||\mathbf{x}||_\infty$. Утверждение доказано.

Задача 4.1.1. Докажите, что

$$||\mathbf{x}||_H = \big(\sum_{i=1}^n |x_i|^p\big)^{\frac{1}{p}},$$

где $p > 1$ — натуральное число, является нормой. Она называется **нормой Гельдера**.

4.2. Элементы топологии. Эквивалентность норм в конечномерном пространстве.

Напомним некоторые понятия из функционального анализа.

Определение 4.2.1. Говорят, что на множестве X **определена структура топологического пространства**, если задана система σ его подмножеств, обладающая свойствами:

1) само множество X и пустое множество ϕ принадлежат σ ;

2) объединение любого числа множеств системы σ принадлежит σ ;

3) пересечение любого конечного числа множеств системы σ принадлежит σ .

Система σ , удовлетворяющая условиям 1)-3), называется **топологией** на множестве X , а составляющие ее множества — **открытыми в этой топологии**.

Пара, состоящая из множества X и топологии σ называется **топологическим пространством**.

Топологическое пространство обычно обозначают через $T = (X, \sigma)$. Если понятно какая на X выбрана топология, то часто топологическое пространство обозначают той же буквой X , что и само множество.

Определение 4.2.2. **Открытой окрестностью точки** x , принадлежащей топологическому пространству $T = (X, \sigma)$, называется любое открытое множество, содержащее точку x .

Определение 4.2.3. Отображение f одного топологического пространства $T = (X, \sigma)$ в другое $S = (Y, \sigma_1)$ называется **непрерывным в точке** x , если для каждой открытой окрестности $\sigma_{f(x)}$ точки $f(x)$ найдется такая открытая окрестность σ_x точки x , что $f(\sigma_x) \subset \sigma_{f(x)}$.

Лемма 4.2.1. *Отображение $f : T \to S$ непрерывно тогда и только тогда, когда полный прообраз любого открытого множества открыт.*

Доказательство. Пусть f — непрерывное отображение и G — открытое множество в S . Если полный прообраз G не пуст, то он — открытое множество. Действительно, если $a \in f^{-1}(G)$, то, поскольку f непрерывно в точке a , существует такая открытая окрестность σ_a точки a , что $f(\sigma_a) \subset G$, т.е. σ_a принадлежит $f^{-1}(G)$ — полному прообразу множества G . Поскольку $f^{-1}(G) = \bigcup_{a \in f^{-1}(G)} \sigma_a$, то $f^{-1}(G)$ — открытое множество, как объединение открытых. Если полный прообраз G — пустое множество,то оно открыто, ибо входит в топологию.

открытое множество в T, образ которого содержится в $\sigma_{f(a)}$, т.е. f — непрерывно.

В курсе функционального анализа доказывается следующее обобщение теоремы Вейерштрасса:

Теорема 4.2.1. *Если $f: T \to \mathbb{R}$ — непрерывное отображение топологического пространства T в числовую прямую $\mathbb{R}$, M — компактное множество в T, то f ограничено и достигает своих точных нижней и верхней граней на M.*

Как мы знаем, топология на X может быть задана с помощью определяющей системы окркстностей.

Вернемся к рассмотрению пространства $\mathbb{E}_n$. Топологию на этом пространстве индуцирует норма. Назовем ε**-окрестностью** $\mathbf{a} \in \mathbb{E}_n$ множество $O_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbb{E}_n : \ ||\mathbf{x} - \mathbf{a}|| < \varepsilon\}$. Эти ε-окрестности порождают топологию в $\mathbb{E}_n$. Любое множество в $\mathbb{E}_n$, которое представимо в виде объединения таких ε-окрестностей при всевозможных $\mathbf{a}$ и всевозможных ε, назовем открытым. Тем самым $\mathbb{E}_n$ превращается в линейное топологическое пространство (на самом деле, в его частный случай — линейное нормированное пространство). Пространство $\mathbb{E}_n$ с некоторой заданной нормой можно превратить также в метрическое пространство, если ввести расстояние

$$\rho(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||. \tag{4.2.1}$$

Задача 4.2.1. Проверьте, что равенство (4.2.1) определяет метрику.

Из функционального анализа известно, что непрерывность отображения $f: X \to Y$ метрического пространства X в метрическое пространство Y эквивалентна тому, что если $x_n \to x, \quad x_n, x \in X$, то $f(x_n) \to f(x)$.

Теорема 4.2.2. *Отображение $|| \cdot || : \mathbb{E}_n \to \mathbb{R}^+$ является непрерывным.*

Доказательство. В силу свойств расстояния

$$|\rho(\mathbf{x}_n, \mathbf{0}) - \rho(\mathbf{x}, \mathbf{0})| \leq \rho(\mathbf{x}_n, \mathbf{x})$$

(разность двух сторон треугольника не превосходит третьей стороны). В силу этого неравенства при $\mathbf{x}_n \to \mathbf{x}$ будем иметь:

$$\rho(\mathbf{x}_n, \mathbf{0}) \to \rho(\mathbf{x}, \mathbf{0}),$$

что эквивалентно тому, что

$$||\mathbf{x}_n|| \to ||\mathbf{x}||.$$

В силу замечания перед теоремой 4.2.2 это эквивалентно непрерывности нормы.

Следующий результат обобщает утверждение 4.1.1.

Теорема 4.2.3 (Люстерник). *Все нормы на евклидовом пространстве эквивалентны между собой.*

Доказательство. В силу того, что эквивалентность норм транзитивна, достаточно доказать, что любая норма на $\mathbb{E}_n$ эквивалентна евклидовой норме $|| \cdot ||_{II}$. Рассмотрим любую норму $|| \cdot ||$ на единичной сфере $||\mathbf{x}||_{II} = 1$. Единичная сфера — замкнутое ограниченное множество, т.е. компактное. По теореме 4.2.2 $|| \cdot ||$ — непрерывное отображение, следовательно, по теореме 4.2.1 $|| \cdot ||$ ограничена на единичной сфере и достигает на ней своих наибольшего и наименьшего значений. Для любого ненулевого вектора $\mathbf{x}$ вектор $\frac{\mathbf{x}}{||\mathbf{x}||_{II}}$ лежит на единичной сфере, поскольку $||\frac{\mathbf{x}}{||\mathbf{x}||_{II}}||_{II} = 1$. Следовательно, найдутся числа α и β такие, что положительны и одно из них обязательно достигается. Таким образом, для любого ненулевого вектора

$$\alpha||\mathbf{x}||_{II} \leq ||\mathbf{x}|| \leq \beta||\mathbf{x}||_{II}.$$

Для вектора $\mathbf{0}$ это неравенство выполняется, так как все три его части обращаются в ноль. Значит, любая норма эквивалентна норме $||\cdot||_{II}$. Теорема доказана.

Так как все нормы в $\mathbb{E}_n$ эквивалентны, то из сходимости $\mathbf{x}_n \to \mathbf{x}, \quad \mathbf{x}_n, \mathbf{x} \in \mathbb{E}_n$, в одной норме (метрике, топологии) следует сходимость в любой другой норме (метрике, топологии). Действительно, если $\mathbf{x}_n \to \mathbf{x}$ в норме $||\cdot||_1$, т.е. $||\mathbf{x_n} - \mathbf{x}||_1 \to 0$, то поскольку $||\mathbf{x}_n - \mathbf{x}||_2 \leq \beta||\mathbf{x}_n - \mathbf{x}||_1$, то и $||\mathbf{x}_n - \mathbf{x}||_2 \to 0$, т.е. $\mathbf{x}_n \to \mathbf{x}$ в норме $||\cdot||_2$. Так как сходимость в евклидовой метрике эквивалентна покоординатной сходимости, то и любая сходимость эквивалентна покоординатной сходимости. Вспомним снова, что непрерывность отображения $f: X \to Y$ метрического пространства X в метрическое пространство Y эквивалентна тому, что если $x_n \to x, \quad x_n, x \in X$, то $f(x_n) \to f(x)$. Из сказанного выше вытекает, что непрерывная функция $f:\ \mathbb{E}_n \to \mathbb{R}$ в любой норме (метрике, топологии) останется непрерывной и в любой другой норме (метрике, топологии). Итак, сходимость и непрерывность в $\mathbb{E}_n$ инвариантны относительно выбора топологии в $\mathbb{E}_n$.

4.3. Скалярное произведение.

Пусть $\mathbb{E}_n$ — n-мерное линейное пространство над полем P. Наиболее важными для приложений являются случаи, когда поле P совпадает с полем действительных чисел $\mathbb{R}$ и когда — с полем комплексных чисел $\mathbb{C}$.

Напомним, что **действительным евклидовым пространством** $\mathbb{R}^n$ называется действительное линейное пространство, в котором определено **скалярное произведение**, т.е. отображение $(.,.):\ \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, называемое **скалярным произведением**, если оно удовлетворяет следующим аксиомам:

1. $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$.
2. $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y}), \quad \lambda \in \mathbb{R}$.
3. $(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y})$.
4. $(\mathbf{x}, \mathbf{x}) \geq \mathbf{0}\ \ \forall\, \mathbf{x}$ и $(\mathbf{x}, \mathbf{x}) = \mathbf{0}$ тогда и только тогда, когда $\mathbf{x} = \mathbf{0}$.

Если $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$ — координаты векторов $\mathbf{x}$ и $\mathbf{y}$ соответственно, то

$$(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Все пространства $\mathbb{R}^n$ изоморфны арифметическому евклидову пространству, в котором вектора будем представлять как столбцы с естественным определением операций над ними. В матричных обозначениях

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots, x_n)^T,$$

где T — операция транспонирования матрицы. В матричных обозначениях скалярное произведение запишется как

$$(\mathbf{x}, \mathbf{y}) = y^Tx = x^Ty.$$

Длиной или нормой вектора в евклидовом вещественном пространстве называется число

$$||\mathbf{x}|| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2},$$

Приведем определение скалярного произведения в комплексном конечномерном линейном пространстве. Пусть $\mathbb{C}^n$ — **комплексное евклидово пространство**, т.е. комплексное n-мерное линейное пространство со скалярным произведением.

Определение 4.3.1. Отображение $(.,.) : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ называется **скалярным произведением**, если оно удовлетворяет следующим аксиомам:

1. $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$ (здесь черта означает переход к комплексно сопряженному числу).
2. $(\lambda \mathbf{x}, \mathbf{y}) = \lambda (\mathbf{x}, \mathbf{y}), \quad \lambda \in \mathbb{C}$.
3. $(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y})$.
4. $(\mathbf{x}, \mathbf{x})$ — действительное неотрицательное число $\forall\, \mathbf{x}$, причем $(\mathbf{x}, \mathbf{x}) = \mathbf{0}$ тогда и только тогда, когда $\mathbf{x} = \mathbf{0}$.

Из аксиом 1 и 2 следует, что

$$(\mathbf{x}, \lambda \mathbf{y}) = \overline{(\lambda \mathbf{y}, \mathbf{x})} = \overline{\lambda (\mathbf{y}, \mathbf{x})} = \overline{\lambda}\ \overline{(\mathbf{y}, \mathbf{x})} = \overline{\lambda} (\mathbf{x}, \mathbf{y}).$$

Кроме того,

$$(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = \overline{(\mathbf{y}_1 + \mathbf{y}_2, \mathbf{x})} = \overline{(\mathbf{y}_1, \mathbf{x})} + \overline{(\mathbf{y}_2, \mathbf{x})} = (\mathbf{x}, \mathbf{y}_1) + (\mathbf{x}, \mathbf{y}_2).$$

Длиной или **нормой** вектора в евклидовом комплексном пространстве называется число

$$||\mathbf{x}|| = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

Векторы $\mathbf{x}$ и $\mathbf{y}$ называются **ортогональными**, если $(\mathbf{x}, \mathbf{y}) = 0$. **Ортогональным базисом** в комплексном евклидовом пространстве называется совокупность n попарно ортогональных не равных нулю векторов $\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}$. Так же, как и в случае вещественного евклидова пространства можно доказать, что эти векторы линейно независимы, т.е. образуют базис. Существование ортогонального базиса в комплексном евклидовом пространстве доказывается точно так же с помощью процесса ортогонализации, как и в вещественном евклидовом пространстве.

Выразим скалярное произведение двух векторов $\mathbf{x}$ и $\mathbf{y}$ через их координаты в ортогональном нормированном базисе (т.е. все вектора базиса единичной длины). Пусть $\mathbf{x} = x_1 \mathbf{e_1} + x_2 \mathbf{e_2} + \dots + x_n \mathbf{e_n}$ и $\mathbf{y} = y_1 \mathbf{e_1} + y_2 \mathbf{e_2} + \dots + y_n \mathbf{e_n}$ — разложения векторов по этому базису. Пользуясь аксиомами и свойствами скалярного произведения, учитывая тот факт, что $(\mathbf{e_i}, \mathbf{e_j}) = 0$ при $i \neq j$ и что $(\mathbf{e_i}, \mathbf{e_i}) = 1$, получим:

$$(\mathbf{x}, \mathbf{y}) = (x_1 \mathbf{e_1} + x_2 \mathbf{e_2} + \dots + x_n \mathbf{e_n}, y_1 \mathbf{e_1} + y_2 \mathbf{e_2} + \dots + y_n \mathbf{e_n}) = x_1 \overline{y}_1 + x_2 \overline{y}_2 + \dots + x_n \overline{y}_n.$$

В матричных обозначениях

$$(\mathbf{x}, \mathbf{y}) = \mathbf{y}^* \mathbf{x} = x_1 \overline{y}_1 + x_2 \overline{y}_2 + \dots + x_n \overline{y}_n.$$

Здесь $*$ — операция взятия сопряженной матрицы. При этом комплексная матрица A^* называется сопряженной к матрице A, если она получается из матрицы A с помощью ее транспонирования и замены в полученной матрице всех элементов комплексно сопряженными к ним элементами, т.е.

$$A^* = \overline{A}^T = \overline{A^T}.$$

Так как мы договорились рассматривать вектора как вектора-столбцы, то, в частности, $\mathbf{y}^T = (y_1, y_2, \dots, y_n)$, $\quad \mathbf{y}^* = (\overline{y_1}, \overline{y_2}, \dots, \overline{y_n})$ и

$$y^* x = (\overline{y_1}, \overline{y_2}, \dots, \overline{y_n}) \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i \overline{y_i}.$$

$$||\mathbf{x}|| = \sqrt{(\mathbf{x},\mathbf{x})} = \sqrt{\sum_{i=1}^{n} x_i \overline{x_i}} = \sqrt{\sum_{i=1}^{n} |x_i|^2},$$

т.е. совпадает с введенной ранее сферической или эрмитовой нормой $||\mathbf{x}||_{II}$.

В дальнейшем будем обозначать через $\mathbb{E}^n$ евклидово пространство над полем P, где P — либо поле комплексных чисел $\mathbb{C}$, т.е. $P = \mathbb{C}$, либо поле действительных чисел, т.е. $P = \mathbb{R}$.

4.4. Норма матрицы.

Норма матрицы может быть введена различными способами.

Способ 1. Рассматривая прямоугольные $m \times n$-матрицы как вектора в $m \times n$-мерном пространстве, можно определить норму матрицы как норму соответствующего вектора. При таком подходе для матрицы $A = (a_{ij})$, $i = \overline{1,m}$, $j = \overline{1,n}$, вводятся три основные нормы:

$$1)\ \ ||A||_I = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}| \text{ — октаэдрическая;}$$

$$2)\ \ ||A||_{II} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \text{ — сферическая;}$$

$$3)\ \ ||A||_\infty = \sqrt{mn}\max_{i,j} |a_{ij}| \quad - \quad mn - \text{норма.}$$

Остальные нормы называют вспомогательными.

Способ 2. Рассмотрим произвольную $m \times n$-матрицу A с элементами из поля P, где $P = \mathbb{C}$ или $P = \mathbb{R}$. Матрице A соответствует линейный оператор

$$\mathbf{y} = A\mathbf{x} \tag{4.4.1}$$

из пространства $\mathbb{E}^n$ в пространство $\mathbb{E}^m$. Если в пространствах $\mathbb{E}^n$ и $\mathbb{E}^m$ выбраны некоторые базисы $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ и $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$ соответственно, то линейный оператор (4.4.1) будет иметь $m \times n$-матрицу

$$A = (A\mathbf{e}_1, A\mathbf{e}_2, \dots, A\mathbf{e}_n).$$

На (4.4.1) можно смотреть как на матричное равенство, где $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ и $\mathbf{y} = (y_1, y_2, \dots, y_m)^T$ — вектора-столбцы координат векторов $\mathbf{x}$ и $\mathbf{y}$ соответственно в выбранных базисах. Более подробно

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n.$$

Таким образом, если в пространствах $\mathbb{E}^n$ и $\mathbb{E}^m$ фиксированы некоторые базисы, то между $m \times n$-матрицами и линейными операторами $A : \ \mathbb{E}^n \to \mathbb{E}^m$ устанавливается взаимно однозначное соответствие. Поэтому нормой матрицы можно считать норму соответствующего линейного оператора. Вспоминая определение нормы линейного оператора, приходим к следующему определению.

$$||A|| = \sup_{\mathbf{x}\in E^n, \mathbf{x}\neq\mathbf{0}} \frac{||A\mathbf{x}||_{E^m}}{||\mathbf{x}||_{E^n}} = \sup_{||\mathbf{x}||_{E^n}=1} ||A\mathbf{x}||_{E^m}, \tag{4.4.1}$$

где $||\cdot||_{E^n}$, $||\cdot||_{E^m}$ — нормы векторов в пространствах $\mathbb{E}^n$ и $\mathbb{E}^m$ соответственно.

Как видно из определения норма $m\times n$-матрицы A определяется как самой матрицей, так и теми векторными нормами, которые выбраны в пространствах $\mathbb{E}^n$ и $\mathbb{E}^m$. При изменении этих норм изменяется и норма матрицы.

С целью сокращения записи в дальнейшем не будем указывать к какому из пространств $\mathbb{E}^n$ или $\mathbb{E}^m$ относится векторная норма. Тогда равенство (4.4.1) перепишется как

$$||A|| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{||A\mathbf{x}||}{||\mathbf{x}||} = \sup_{||\mathbf{x}||=1} ||A\mathbf{x}||. \tag{4.4.2}$$

Рассмотрим свойства матричной нормы.

Свойство 1. $||A|| \geq 0$, *причем* $||A|| = 0$ *тогда и только тогда, когда* $A = 0$.

Д о к а з а т е л ь с т в о. Из (4.4.2) видно, что $||A|| \geq 0$.

Пусть $A = 0$, тогда $\forall \mathbf{x} \in E^n$ выполнится $A\mathbf{x} = \mathbf{0}$, значит, $||A\mathbf{x}|| = 0$, откуда с учетом (4.4.2) $||A|| = 0$.

Обратно, пусть $||A|| = 0$, тогда из (4.4.2) $||A\mathbf{x}|| = 0 \quad \forall \mathbf{x} \neq \mathbf{0}$. Но, очевидно, $||A\mathbf{0}|| = 0$. Итак, $||A\mathbf{x}|| = 0 \ \ \forall \mathbf{x}$, откуда $A = 0$ (почему?).

Свойство 2. $||\lambda A|| = |\lambda|\, ||A|| \quad \forall\, \lambda \in P$. (напомним, что $P = \mathbb{C}$ или $P = \mathbb{R}$).

Д о к а з а т е л ь с т в о. В силу (4.4.2)

$$||\lambda A|| = \sup_{||\mathbf{x}||=1} ||\lambda A\mathbf{x}|| = \sup_{||\mathbf{x}||=1} (|\lambda|\, ||A\mathbf{x}||) = |\lambda| \sup_{||\mathbf{x}||=1} ||A\mathbf{x}|| = |\lambda|\, ||A||.$$

Свойство 3. $||A + B|| \leq ||A|| + ||B||$.

Д о к а з а т е л ь с т в о. При определении нормы матрицы (4.4.2)

$$||A + B|| = \sup_{||\mathbf{x}||=1} ||(A + B)\mathbf{x}|| \leq \sup_{||\mathbf{x}||=1} (||A\mathbf{x}|| + ||B\mathbf{x}||) \leq \sup_{||\mathbf{x}||=1} ||A\mathbf{x}|| + \sup_{||\mathbf{x}||=1} ||B\mathbf{x}|| = ||A|| + ||B||.$$

Свойство 4. $||A\mathbf{x}|| \leq ||A|| \cdot ||\mathbf{x}||$.

Д о к а з а т е л ь с т в о. По свойствам точной верхней грани из определения 4.4.1 следует, что для $\mathbf{x} \neq \mathbf{0}$

$$\frac{||A\mathbf{x}||}{||\mathbf{x}||} \leq ||A||,$$

т.е. для $\mathbf{x} \neq \mathbf{0}$ действительно выполняется $||A\mathbf{x}|| \leq ||A|| \cdot ||\mathbf{x}||$. Если же $\mathbf{x} = \mathbf{0}$, то $A\mathbf{x} = \mathbf{0}$, т.е. $||A\mathbf{x}|| = 0$ и $||\mathbf{x}|| = 0$, т.е. доказываемое неравенство также выполняется.

Свойство 5. Если $p \times n$*-матрица* $B: \ E^n \to E^p$, *а* $m \times p$*-матрица* $A: \ E^p \to E^m$, *то*

$$||AB|| \leq ||A|| \cdot ||B||.$$

Д о к а з а т е л ь с т в о. В самом деле, пользуясь свойством 4, получим

$$||AB|| = \sup_{||\mathbf{x}||=1} ||AB\mathbf{x}|| \leq \sup_{||\mathbf{x}||=1} ||A||||B\mathbf{x}|| = ||A|| \sup_{||\mathbf{x}||=1} ||B\mathbf{x}|| = ||A|| \cdot ||B||.$$

$M(m,n)$ — множество всех $m \times n$-матриц над полем P ($P = \mathbb{C}$ или $P = \mathbb{R}$).

Определение 4.4.2. Отображение $||\cdot|| : M(m,n) \to \mathbb{R}^+$ называется **матричной нормой**, если оно подчиняется следующим аксиомам:

1. $||A|| = 0$ тогда и только тогда, когда $A = 0$.
2. $||\lambda A|| = |\lambda| ||A|| \quad \forall \ \lambda \in P$.
3. $||A + B|| \leq ||A|| + ||B||$.
4. $||AB|| \leq ||A|| \cdot ||B||$.

В силу свойств нормы матрицы, введенной с помощью "операторного" определения 4.4.1, такая норма удовлетворяет акстоматическому определению 4.4.2.

Определение 4.4.3. Матричная норма $||\cdot||_{\text{м}}$ и векторная норма $||\cdot||_{\text{в}}$ называются **согласованными между собой**, если

$$||A\mathbf{x}||_{\text{в}} \leq ||A||_{\text{м}} \cdot ||\mathbf{x}||_{\text{в}}.$$

Очевидно, что матричная норма "операторного" определения 4.4.1 согласована с векторной нормой, фигурирующей в этом определении.

Для любой векторной нормы всегда существует матричная норма, согласованная с ней (например, матричная норма из определения 4.4.1). Обратно, для любой матричной нормы существует согласованная с ней векторная норма. Достаточно взять ту же норму у вектора, которая бралась у матрицы (ведь вектор — это матрица размера $n \times 1$). Тогда эти нормы согласованы в силу аксиомы 4 определения 4.4.2. Итак, справедлива следующая

Теорема 4.4.1 (Люстерник). *Для любой векторной нормы существует согласованная с ней матричная норма, а для любой матричной нормы существует согласованная с ней векторная норма.*

Отметим, что основные матричные нормы $||\cdot||_I$, $||\cdot||_{II}$, $||\cdot||_\infty$ не совпадают с матричными нормами, согласованными с векторными нормами $||\cdot||_I$, $||\cdot||_{II}$, $||\cdot||_\infty$.

Определение 4.4.4. Матричные нормы $||\cdot||_1$ и $||\cdot||_2$ называются **эквивалентными**, если существуют постоянные $\beta \geq \alpha > 0$ такие, что для любой матрицы A

$$\alpha ||A||_1 \leq ||A||_2 \leq \beta ||A||_1.$$

То, что это отношение эквивалентности, доказывается точно так же, как и для векторных норм.

Каким бы способом ни была введена матричная норма, она может в конечном счете рассматриваться как норма вектора в E^{mn}, поскольку при любом подходе будут выполняться аксиомы нормы вектора. Поэтому из теоремы 4.2.3 вытекает

Теорема 4.4.2 (Люстерник). *Любые две матричные нормы эквивалентны.*

Когда рассматривалась сходимость последовательности матриц, она определялась как покоординатная сходимость. Так как такая сходимость эквивалентна сходимости по сферической норме, а по теореме 4.4.2 сходимость по сферической норме эквивалентна сходимости по любой матричной норме, то в определении сходимости последовательности матриц может быть взята любая матричная норма.

Рассмотрим специфические свойства некоторых матричных норм.

A, то число
$$R = \max(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_s|)$$
называется **спектральным радиусом** матрицы A.

Свойство 1. Матричная норма квадратной матрицы не меньше ее спектрального радиуса.

Доказательство. В силу теоремы 1.1.2 для любой квадратной матрицы существует собственный вектор. Пусть $\mathbf{x}$ — собственный вектор A, соответствующий собственному значению λ:
$$A\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \neq \mathbf{0}.$$
По теореме 4.4.1 для любой матричной нормы существует согласованная с ней векторная норма. Выберем ее. Тогда
$$||A\mathbf{x}||_{\text{в}} = ||\lambda\mathbf{x}||_{\text{в}} = |\lambda| ||\mathbf{x}||_{\text{в}} \leq ||A||_{\text{м}} ||\mathbf{x}||_{\text{в}}.$$
Так как $\mathbf{x} \neq \mathbf{0}$, то $||\mathbf{x}||_{\text{в}} > 0$. Поэтому $||A||_{\text{м}} \geq |\lambda|$. Так как это неравенство выполняется для всех собственных значений, то $||A||_{\text{м}} \geq R$.

Свойство 2. Сферическая норма квадратной матрицы подсчитывается по формуле
$$||A||_{II} = \sqrt{tr(A^*A)} = \sqrt{tr(AA^*)},$$
где tr — след матрицы, равный сумме ее диагональных элементов.

Это свойство проверяется непосредственным подсчетом.

В частности, если U — комплексная унитарная матрица (т.е. $UU^* = E$, где E — единичная матрица) или действительная ортогональная матрица (т.е. $UU^T = E$), то
$$||U_{II}|| = \sqrt{n}.$$

Определение 4.4.6. Матричная норма квадратной матрицы, построенная по векторной норме с помощью "операторного" определения 4.4.1, называется **подчиненной этой векторной норме**. Матричная норма, подчиненная сферической векторной норме, называется **спектральной**.

Свойство 3. Спектральная норма ортогональной матрицы равна 1.

Доказательство. Ортогональной матрице соответствует ортогональное линейное преобразование, которое сохраняет длины (нормы) векторов, т.е. $||U\mathbf{x}||_{II} = ||\mathbf{x}||_{II}$. Следовательно, спектральная норма будет равна
$$||U||_s = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||U\mathbf{x}||_{II}}{||\mathbf{x}||_{II}} = 1.$$

Свойство 4. Спектральная норма матрицы не меняется при умножении ее слева или справа на ортогональную матрицу.

Доказательство. Докажем для умножения слева; для умножения справа доказательство аналогично. Если U — ортогональная матрица, то U^{-1} также ортогональна. По свойству 3 $||U_s|| = ||U^{-1}||_s = 1$. Поэтому
$$||UA||_s \leq ||U||_s ||A||_s = ||A||_s = ||U^{-1}UA||_s \leq ||U^{-1}||_s ||UA||_s = ||UA||_s,$$
откуда $||UA||_s = ||A||_s$.

называются корни квадратные из собственных значений матрицы A^*A.

Для любой прямоугольной матрицы матрица A^*A является квадратной. Квадратная матрица B называется **эрмитовой**, если $B = B^*$. Для любой квадратной матрицы

$$(A\mathbf{x}, \mathbf{y}) = \mathbf{y}^*A\mathbf{x} = (A^*\mathbf{y})^*\mathbf{x} = (\mathbf{x}, A^*\mathbf{y}).$$

Значит, для эрмитовой матрицы $(B\mathbf{x}, \mathbf{y}) = (\mathbf{x}, B\mathbf{y})$. В частности, при $\mathbf{x} = \mathbf{y}$

$$(B\mathbf{x}, \mathbf{x}) = (\mathbf{x}, B\mathbf{x}) = \overline{(B\mathbf{x}, \mathbf{x})},$$

т.е. для эрмитовой матрицы скалярное произведение $(B\mathbf{x}, \mathbf{x})$ — действительное число. На самом деле,

$$(B\mathbf{x}, \mathbf{x}) = \mathbf{x}^*B\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\overline{x_i}x_j$$

есть комплексная квадратичная форма от $x_1, x_2, \ldots, x_n$.

Эрмитова матрица B называется **неотрицательно определенной**, если для всех векторов $\mathbf{x}$ имеем $(B\mathbf{x}, \mathbf{x}) \geq 0$. Все собственные значения неотрицательно определенной эрмитовой матрицы неотрицательны, поскольку

$$(B\mathbf{x}, \mathbf{x}) = (\lambda\mathbf{x}, \mathbf{x}) = \lambda(\mathbf{x}, \mathbf{x}) \geq 0.$$

Из курса алгебры известно, что вещественную симметрическую или комплексную эрмитову матрицу (являющуюся матрицей квадратичной формы) можно с помощью ортогональной матрицы или с помощью унитарной матрицы привести к диагональному виду, причем на диагонали стоят собственные значения (квадратичную форму можно привести с помощью ортогонального или унитарного преобразования к сумме квадратов, где коэффициенты перед квадратами совпадают с собственными значениями). Если к тому же матрица неотрицательно определена, то на диагонали будут стоять неотрицательные числа (коэффициенты перед квадратами неотрицательны).

В определении 4.4.7 квадратная матрица A^*A — эрмитова, так как

$$(A^*A)^* = A^*(A^*)^* = A^*A,$$

и неотрицательно определена, поскольку

$$(A^*A\mathbf{x}, \mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = (A\mathbf{x})^*A\mathbf{x} = ||A\mathbf{x}||^2 \geq 0.$$

Значит, собственные значения A^*A неотрицательны и сингулярные числа в определении 4.4.7 действительны и неотрицательны.

Свойство 5. Спектральная норма матрицы равна ее максимальному сингулярному числу.

Д о к а з а т е л ь с т в о. Пусть U — ортогональная в вещественном случае и унитарная в комплексном случае матрица такая, что A^*A приводится к диагональному виду с помощью U:

$$U^*A^*AU = \begin{pmatrix} \lambda_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_n \end{pmatrix},$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения A^*A, которые неотрицательны.

$\mathbf{x} = \mathbf{0}$, то $\xi = \mathbf{0}$. Итак, $\xi \neq \mathbf{0}$ тогда только тогда, когда $\mathbf{x} \neq \mathbf{0}$. Поэтому

$$||A||_s^2 = \Big(\sup_{||\mathbf{x}||_{II} \neq 0} \frac{||A\mathbf{x}||_{II}}{||\mathbf{x}||_{II}} \Big)^2 = \sup_{||\mathbf{x}||_{II} \neq 0} \frac{||A\mathbf{x}||_{II}^2}{||\mathbf{x}||_{II}^2} = \sup_{||\mathbf{x}||_{II} \neq 0} \frac{(A\mathbf{x})^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \sup_{||\mathbf{x}||_{II} \neq 0} \frac{\mathbf{x}^* A^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} =$$

$$= \sup_{||\xi||_{II} \neq 0} \frac{\xi^* U^* A^* A U \xi}{\xi^* \xi} = \sup_{||\xi||_{II} \neq 0} \frac{\lambda_1 \xi_1^2 + \ldots + \lambda_n \xi_n^2}{\xi_1^2 + \ldots + \xi_n^2}.$$

Пусть $i \in \{1, 2, \ldots, n\}$ таково, что $\lambda_i = \max(\lambda_1, \lambda_2, \ldots, \lambda_n)$, очевидно,

$$\frac{\lambda_1 \xi_1^2 + \ldots + \lambda_n \xi_n^2}{\xi_1^2 + \ldots + \xi_n^2} \leq \lambda_i.$$

В то же время для ненулевого вектора с координатой $\xi_i = 1$ и остальными координатами, равными 0

$$\frac{\lambda_1 \xi_1^2 + \ldots + \lambda_n \xi_n^2}{\xi_1^2 + \ldots + \xi_n^2} = \lambda_i.$$

Следовательно, $||A||_s^2 = \lambda_i$, т.е. $||A||_s = \sqrt{\lambda_i}$, значит, спектральная норма матрицы A равна ее максимальному сингулярному числу.

5.1. Треугольные матрицы.

Матрица A называется **правой** или **верхней треугольной**, если

$$a_{ij} = 0 \text{ для } i > j,$$

и **левой** или **нижней треугольной**, если

$$a_{ij} = 0 \text{ для } i < j.$$

Треугольные матрицы обладают рядом замечательных свойств, в силу которых они очень широко используются в построении самых различных методов решения задач алгебраического характера.

1. Сумма и произведение одноименных треугольных матриц есть треугольная матрица того же наименования.

2. Определитель треугольной матрицы равен произведению диагональных элементов.

3. Собственные значения треугольной матрицы совпадают с ее диагональными элементами.

4. Обратная матрица для треугольной матрицы, в которой все диагональные элементы отличны от нуля, существует и является треугольной.

Можно ли разложить квадратную матрицу в произведение двух треугольных? Допустим, что

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{pmatrix}. \tag{5.1.1}$$

Приравнивая соответствующие элементы этого равенства, получим

$$a_{11} = b_{11}c_{11}, \quad a_{12} = b_{11}c_{12}, \quad a_{13} = b_{11}c_{13},$$

$$a_{21} = b_{21}c_{11}, \quad a_{22} = b_{21}c_{12} + b_{22}c_{22}, \quad a_{23} = b_{21}c_{13} + b_{22}c_{23},$$

$$a_{31} = b_{31}c_{31}, \quad a_{32} = b_{31}c_{12} + b_{32}c_{22}, \quad a_{33} = b_{31}c_{13} + b_{32}c_{23} + b_{33}c_{33},$$

откуда, предполагая, что $b_{11}c_{11}b_{22}c_{22} \neq 0$, получим

$$c_{11} = \frac{a_{11}}{b_{11}}, \quad c_{12} = \frac{a_{12}}{b_{11}}, \quad c_{13} = \frac{a_{13}}{b_{11}},$$

$$b_{21} = \frac{a_{21}}{c_{11}}, \quad b_{31} = \frac{a_{31}}{c_{11}}, \quad c_{22} = \frac{a_{22} - b_{21}c_{12}}{b_{22}},$$

$$b_{33}c_{33} = a_{33} - b_{31}c_{13} - b_{32}c_{23}, \quad c_{23} = \frac{a_{23} - b_{21}c_{13}}{b_{22}}, \quad b_{32} = \frac{a_{32} - b_{32}c_{12}}{c_{22}}.$$

Отсюда видно, что задавая произвольным образом $b_{11} \neq 0$, $b_{22} \neq 0$, мы однозначно выразим все элементы матриц в правой части (5.1.1) через элементы матрицы A, кроме b_{33} и c_{33}, которые задаем произвольно, лишь бы $b_{33}c_{33} = a_{33} - b_{31}c_{13} - b_{32}c_{23}$. Итак, если $b_{11}c_{11}b_{22}c_{22} \neq 0$, то справедливо разложение (5.1.1). Какие условия надо наложить, чтобы $b_{11}c_{11}b_{22}c_{22} \neq 0$? Так как $a_{11} \neq 0$ тогда и только тогда, когда $b_{11}c_{11} \neq 0$, то остается посмотреть, когда $b_{22}c_{22} \neq 0$. Обозначим

$$A_1 = a_{11}, \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

видеть, что $A_2 = B_2C_2$. Но тогда $|A_2| = |B_2||C_2| = b_{11}b_{22}c_{11}c_{22}$. Поэтому, если $|A_1| \neq = 0$, $|A_2| \neq 0$, то $b_{11}b_{22}c_{11}c_{22} \neq 0$. Итак, справедлива

Теорема 5.1.1. *Если*

$$a_{11} \neq 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0,$$

то справедливо разложение (5.1.1).

Эта теорема легко обобщается на матрицы любого порядка.

Теорема 5.1.2. *Какова бы ни была квадратная матрица $A = (a_{ij})$ порядка n с отличными от нуля главными минорами*

$$a_{11} \neq 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \ \dots, \ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} \\ a_{21} & a_{22} & \dots & a_{2,n-1} \\ \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} \end{vmatrix} \neq 0,$$

ее всегда можно разложить в произведение

$$A = B\dot{C}, \tag{5.1.2}$$

где B и C — соответственно левая и правая треугольные матрицы.

Доказательство. Заметим, что если разложение (5.1.2) существует, то оно будет заведомо не единственным. В самом деле, запишем (5.1.2) в следующем виде:

$$A - BDD^{-1}C,$$

где D — диагональная матрица. Тогда матрицы BD и $D^{-1}C$ будут снова треугольными прежнего строения. Покажем теперь, что матричное уравнение (5.1.2) имеет решение.

Пусть b_{ij} и c_{ij} — элементы матриц B и C. Приравнивая между собой элементы матрицы A и произведения BC, получаем

$$a_{ij} = \sum_{p=1}^{\min(i,j)} b_{ip}c_{pj}.$$

Отсюда следует:

$$b_{11}c_{11} = a_{11}; \quad c_{1j} = \frac{a_{1j}}{b_{11}}, \quad b_{i1} = \frac{a_{j1}}{c_{11}}, \quad j = 2, 3, \dots, n; \quad b_{ii}c_{ii} = a_{ii} - \sum_{p=1}^{i-1} b_{ip}c_{pi}, \quad i = 2, 3, \dots, n;$$

$$c_{ij} = \frac{a_{ij} - \sum_{p=1}^{i-1} b_{ip}c_{pj}}{b_{ii}}, \quad b_{ji} = \frac{a_{ji} - \sum_{p=1}^{i-1} b_{jp}c_{pi}}{c_{ii}}, \quad i = 2, 3, \dots, n, \quad j = i+1, i+2, \dots, n. \tag{5.1.3}$$

Итак, уравнение (5.1.2) будет наверняка иметь решение, если все числа b_{ii} и c_{ii} ($i = = 1, 2, \dots, n-1$) не равны нулю.

Предположим, что $b_{kk}c_{kk} = 0$ при некотором k. Из (5.1.3) вытекает, что в этом случае возможно осуществить лишь разложение

$$A_k = B_kC_k,$$

где A_k, B_k и C_k — матрицы угловых миноров k-го порядка матриц A, B и C. Далее находим

$$|A_k| = |B_k|\dot{|}C_k| = \prod_{p=1}^{k} b_{pp}c_{pp} = 0.$$

$-1)$ не равны нулю и разложение (5.1.2) имеет место.

5.2. Почти треугольные матрицы.

Матрица A называется **правой почти треугольной**, если для ее элементов a_{ij} выполняются соотношения

$$a_{ij}=0 \text{ для } i>j+1, \tag{5.2.1}$$

и **левой почти треугольной**, если для ее элементов a_{ij} выполняются соотношения

$$a_{ij}=0 \text{ для } j>i+1. \tag{5.2.2}$$

Заметим, что сумма одноименных почти треугольных матриц будет матрицей почти треугольной, а произведение — нет. Как и любая квадратная матрица с отличными от нуля главными минорами, почти треугольная матрица может быть разложена в произведение двух треугольных матриц. Однако теперь это разложение становится более простым.

Теорема 5.2.1. *Если матрица A — правая почти треугольная, то в разложении (5.1.2) матрица B будет двухдиагональной, если матрица A — левая почти треугольная, то в разложении (5.1.2) матрица C будет двухдиагональной.*

Доказательство. Рассмотрим, например, случай правой почти треугольной матрицы A и докажем, что

$$b_{ji}=0 \text{ для } j>i+1. \tag{5.2.3}$$

При $i=1$ это соотношение непосредственно вытекает из первой строки (5.1.3) и определения (5.2.1). Пусть (5.2.3) выполнено для всех $i=1,2,\ldots,k$, $k>1$. Тогда получаем

$$b_{j,k+1}=\frac{a_{j,k+1}-\sum_{p=1}^{k} b_{jp}c_{p,k+1}}{c_{k+1,k+1}}=0$$

для $j>k+2$, так как для этих j числа $a_{j,k+1}=0$ по условию теоремы, а $b_{jp}=0$ по предположению. Теорема доказана.

В силу (5.2.3) формулы (5.1.3) приобретают следующий вид:

$$b_{11}c_{11}=a_{11};\quad b_{21}=\frac{a_{21}}{c_{11}},\quad c_{1j}=\frac{a_{1j}}{b_{11}},\quad j=2,3,\ldots,n;\quad b_{ii}c_{ii}=a_{ii}-b_{i,i-1}c_{i-1,i},\quad i=2,3,\ldots,n;$$

$$b_{i+1,i}=\frac{a_{i+1,i}}{c_{ii}},\quad c_{ij}=\frac{a_{ij}-b_{i,i-1}c_{i-1,j}}{b_{ii}},\quad i=2,3,\ldots,n,\quad j=i+1,i+2,\ldots,n.$$

Легко находится характеристический многочлен

$$\Delta_n(\lambda)=|\lambda E-A|$$

почти треугольной матрицы. Например, для правой почти треугольной матрицы он может быть вычислен по рекуррентной формуле

$$\Delta_n(\lambda)=(\lambda-a_{kk})\Delta_{k-1}(\lambda)-a_{k,k-1}\Delta_{k-2}(\lambda)-a_{k,k-1}a_{k-1,k-2}a_{k-2,k}\Delta_{k-3}(\lambda)-$$
$$-\ldots-a_{k,k-1}a_{k-1,k-2}\ldots a_{21}a_{1k}\Delta_0(\lambda). \tag{5.2.4}$$

$$\Delta_k(\lambda) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1k} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2k} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda - a_{kk} \end{vmatrix}, \quad k = 1, 2, \dots, n.$$

В справедливости (5.2.4) легко убедиться посредством разложения определителя $\Delta_k(\lambda)$ по элементам последнего столбца.

При этом определитель матрицы A, очевидно, равен $(-1)^n \Delta_n(0)$.

5.3. Унитарные матрицы.

Матрица A с элементами из поля комплексных чисел $\mathbb{C}$ называется **унитарной**, если она удовлетворяет уравнению

$$AA^* = E, \tag{5.3.1}$$

где A^* — матрица, сопряженная с A, т.е. $A^* = \bar{A}^T$, черта означает операцию комплексного сопряженеия, T — операция транспонирования матрицы. Если унитарная матрица является вещественной, то наше определение превращается в определение ортогональной матрицы. Таким образом, унитарность — обобщение ортогональности на случай матриц с комплексными элементами.

Соотношение (5.3.1) можно переписать в виде

$$\sum_{j=1}^{n} a_{ij}\overline{a_{kj}} = \delta_{ik}, \qquad i, k = 1, 2, \dots, n,$$

где δ_{ik} — символ Кронекера. Если обозначить i-ю строку матрицы A через A_i^T $(i = \overline{1, n})$, то через скалярное произведение последнее соотношение запишется как

$$(A_i, A_k) = A_k^* A_i = \delta_{ik}.$$

Таким образом, унитарная матрица — это такая матрица, у которой строки ортогональны и каждая строка имеет единичную длину. Из (5.3.1) следует, что $A^* = A^{-1}$, поэтому $A^*A = A^{-1}A = E$. Следовательно,

$$\sum_{j=1}^{n} a_{ji}\overline{a_{jk}} = \delta_{ik}, \qquad i, k = 1, 2, \dots, n.$$

Это означает, что столбцы матрицы A ортогональны, а каждый столбец имеет единичную длину.

Унитарные матрицы обладают следующими свойствами.

1. Если A — унитарная матрица, то ее определитель есть комплексное число, по модулю равное единице. Действительно,

$$det(AA^*) = det(A) det A^* = det(A) det(\bar{A}) = (det(A))^2 = 1,$$

откуда $|det(A)| = 1$ (здесь для определителя матрицы A использовано обозначение $det(A)$, чтобы не спутать с обозначением модуля числа).

2. Если A — унитарная матрица, то $A^* = A^{-1}$. Это утверждение вытекает непосредственно из (5.3.1) после умножения обеих его частей слева на A^{-1}.

3. Если A — унитарная матрица, то A^* тоже унитарна.

4. Произведение двух унитарных матриц есть матрица унитарная. В самом деле, если A и B — унитарные, то

$$(AB)(AB)^* = ABB^*A^* = AEA^* = E.$$

с характеристическими числами, по модулю равными единице. Это свойство примем без доказательства.

Вещественные унитарные матрицы называются **ортогональными**. Так как для вещественных матриц $A^* = A^T$, то определение ортогональной матрицы может быть записано в виде

$$AA^T = E.$$

К унитарным матрицам относятся элементарные унитарные матрицы

$$\begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cos\varphi\cdot e^{i\psi_1} & & & -\sin\varphi\cdot e^{i\psi_2} & & \\ & & & & \cdot & & & & \\ & & & & & \cdot & & & \\ & & & \sin\varphi\cdot e^{i\psi_3} & & & \cos\varphi\cdot e^{i\psi_4} & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{pmatrix}$$

где ψ_k действительны и $\psi_1 - \psi_2 = \psi_3 - \psi_4 (mod\, 2\pi)$. Они отличаются от единичной матрицы только четырьмя элементами, стоящими на пересечении i-й и j-й строк и i-го и j-го столбцов.

Покажем, что такие матрицы унитарны. Для этого достаточно доказать, что каждая ее вектор-строка имеет единичную длину, а строки ортогональны. То, что строки с номерами, отличными от i и j, имеют единичную длину, очевидно. Длина i-й строки

$$\sqrt{\sum_{k=1}^{n} |a_{ik}|^2} = \sqrt{|\cos\varphi|^2 |e^{i\psi_1}|^2 + |-\sin\varphi|^2 |e^{i\psi_2}|^2} = \sqrt{\cos^2\varphi + \sin^2\varphi} = 1,$$

длина j-й строки

$$\sqrt{\sum_{k=1}^{n} |a_{jk}|^2} = \sqrt{|\sin\varphi|^2 |e^{i\psi_3}|^2 + |\cos\varphi|^2 |e^{i\psi_4}|^2} = \sqrt{\sin^2\varphi + \cos^2\varphi} = 1.$$

Остается доказать, что строки ортогональны между собой, т.е. что при $l \neq k$ скалярное произведение l-й строки на k-ю

$$\sum_{m=1}^{n} a_{lm}\overline{a_{km}} = 0.$$

Если l и k отличны от и от i и от j, это очевидно. Пусть $l = i$, $k \neq j$, тогда

$$\sum_{m=1}^{n} a_{lm}\overline{a_{km}} = \cos\varphi\, e^{i\psi_1} \cdot 0 - \sin\varphi\, e^{i\psi_2} \cdot 0 = 0.$$

Аналогично проверяется, что скалярное произведение l-й строки на k-ю равно нулю при $l \neq = i$, $k = j$ (проверьте). Наконец, при $l = i$, $k = j$, т.е. произведение i-й строки на j-ю равно

$$\sum_{m=1}^{n} a_{im}\overline{a_{jm}} = \cos\varphi\, e^{i\psi_1} \sin\varphi\, e^{-i\psi_3} - \sin\varphi\, e^{i\psi_2} \cos\varphi\, e^{-i\psi_4} = \cos\varphi\ \sin\varphi \Big(e^{i(\psi_1 - \psi_3)} - e^{i(\psi_2 - \psi_4)} \Big).$$

$$\sum_{m=1}^{n} a_{im}\overline{a_{jm}} = \cos\varphi\ \sin\varphi\ e^{i(\psi_2-\psi_4)}\Big(e^{2\pi ni}-1\Big) = 0.$$

Таким образом, унитарность доказана.

Среди элементарных матриц являются наиболее употребительными матрицы

$$R_{ij}(\varphi,\psi) = \begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cos\varphi & & & -\sin\varphi\cdot e^{i\psi} & & \\ & & & & \cdot & & & & \\ & & & & & \cdot & & & \\ & & & \sin\varphi\cdot e^{-i\psi} & & & \cos\varphi & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & 1 \end{pmatrix},$$

которые получаются при $\psi_1=\psi_4=0,\ \ \psi_2=\psi,\ \ \psi_3=-\psi$, и при этом $\psi_1-\psi_2=-\psi=\psi_3-\psi_4$.

Очевидно, что при умножении матрицы A слева на матрицу $R_{ij}(\varphi,\psi)$ изменяются лишь i-я и j-я строки матрицы A. А именно, для элементов i-й и j-й строк матрицы $A_1 = R_{ij}(\varphi,\psi)A$ получим

$$a_{ip}^{(1)} = a_{ip}\cos\varphi - a_{jp}\sin\varphi\, e^{i\psi},$$

$$a_{jp}^{(1)} = a_{ip}\sin\varphi\, e^{-i\psi} + a_{jp}\cos\varphi, \quad p=\overline{1,n}.$$

Лемма 5.3.1. *Всегда можно подобрать параметры φ и ψ так, чтобы в произведении матрицы A слева на $R_{ij}(\varphi,\psi)$ сделать равным нулю любой элемент $a_{jp}^{(1)}$.*

Доказательство.

1. Пусть $\sqrt{|a_{ip}|^2+|a_{jp}|^2}\neq 0$. Выберем φ и ψ таким образом, что

$$\psi = \arg a_{ip} - \arg a_{jp},$$

$$\cos\varphi = \frac{|a_{ip}|}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}}, \quad \sin\varphi = -\frac{|a_{jp}|}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}}.$$

Тогда

$$a_{jp}^{(1)} = a_{ip}\sin\varphi\, e^{-i\psi} + a_{jp}\cos\varphi = \frac{-a_{ip}|a_{jp}|e^{-i\psi}+a_{jp}|a_{ip}|}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}} = \frac{-a_{ip}|a_{jp}|e^{-i\arg a_{ip}}e^{i\arg a_{jp}}+a_{jp}|a_{ip}|}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}} =$$

$$= \frac{-a_{ip}a_{jp}e^{-i\arg a_{ip}}+a_{jp}|a_{ip}|}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}} = e^{-i\arg a_{ip}}\frac{-a_{ip}a_{jp}+a_{jp}|a_{ip}|e^{i\arg a_{ip}}}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}} = e^{-i\arg a_{ip}}\frac{-a_{ip}a_{jp}+a_{jp}a_{ip}}{\sqrt{|a_{ip}|^2+|a_{jp}|^2}} = 0.$$

2. Если же $|a_{ip}|=|a_{jp}|=0$, то $a_{ip}=a_{jp}=0$, и $a_{jp}^{(1)}=0$ при любом выборе φ и ψ, напримнр, можно взять $\cos\varphi=1,\ \sin\psi=0$.

Теорема 5.3.1. *Любая комплексная матрица A преобразуется в правую треугольную матрицу посредством умножения слева на конечную цепочку элементарных унитарных матриц $R_{ij}(\varphi,\psi)$.*

выбирая их так, чтобы последовательно аннулировать все элементы первого столбца, кроме верхнего (используя $n-1$ раз лемму 5.3.1). После этих преобразований приходим к матрице

$$A^{(1)} = R_{1n}\, R_{1,n-1}\, \dots\, R_{12}\, A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \cdot & \cdot & \dots & \cdot \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix},$$

затем подбираем матрицы R_{23}, R_{24}, $\dots$, R_{2n} так, чтобы после умножения матрицы $A^{(1)}$ на них слева аннулировать все элементы второго столбца, лежащие ниже диагонали. Далее переходим к исключению поддиагональных элементов 3-го столбца и т.д. В конце процесса получаем правую треугольную матрицу $A^{(n-1)}$.

Произведение унитарных матриц R_{ij} является унитарной матрицей, поэтому из доказанной теоремы вытекает

Следствие 5.3.1. *Любая комплексная матрица есть произведение унитарной матрицы на правую треугольную матрицу.*

Д о к а з а т е л ь с т в о. Действительно, обозначим

$$R = R_{n-1,n} R_{n-2,n} R_{n-2,n-1} \dots R_{12},$$

тогда

$$A = R^* A^{(n-1)}. \tag{5.3.1}$$

Теорема 5.3.2. *Любая комплексная матрица A приводится к правой почти треугольной матрице с помощью конечной цепочки подобных преобразований с матрицами $R_{ij}(\varphi, \psi)$.*

Д о к а з а т е л ь с т в о. Умножим матрицу A слева на матрицы R_{23}, R_{24}, $\dots$, R_{2n}, выбирая их так, чтобы последовательно аннулировать все элементы первого столбца, кроме двух верхних. После выполнения этих преобразований приходим к матрице

$$\tilde{A}^{(1)} = R_{2n} R_{2,n-1} \dots R_{23} A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \tilde{a}_{21}^{(1)} & \tilde{a}_{22}^{(1)} & \dots & \tilde{a}_{2n}^{(1)} \\ 0 & \tilde{a}_{32}^{(1)} & \dots & \tilde{a}_{3n}^{(1)} \\ \cdot & \cdot & \dots & \cdot \\ 0 & \tilde{a}_{n2}^{(1)} & \dots & \tilde{a}_{nn}^{(1)} \end{pmatrix} \tag{5.3.2}$$

Легко видеть, что в матрице

$$A^{(1)} = \tilde{A}^{(1)} R_{23}^* \dots R_{2,n-1}^* R_{2n}^* \tag{5.3.3}$$

первый столбец будет таким же, как у матрицы $\tilde{A}^{(1)}$. При этом из (5.3.2),(5.3.3)

$$A^{(1)} = R_{2n} R_{2,n-1} \dots R_{23} A R_{23}^* \dots R_{2,n-1}^* R_{2n}^*,$$

т.е. $A^{(1)}$ подобна A с помощью конечной цепочки преобразований с матрицами R_{ij}.

Далее подбираем R_{34}, R_{35}, $\dots$, R_{3n} так, чтобы после подобного преобразования матрицы $A^{(1)}$ матрицей $R_{3n}\, \dots\, R_{35}\, R_{34}$ аннулировать все поддиагональные элементы второго столбца, кроме верхнего. Затем переходим к исключению элементов третьего столбца и т.д. В конце процесса получим правую почти треугольную матрицу $A^{(n-2)}$.

Следствие 5.3.2. *Любая комплексная матрица унитарно подобна правой почти треугольной.*

$$\tilde{R} = R_{n-1,n} R_{n-2,n} \dots R_{24} R_{23},$$

тогда

$$A = \tilde{R}^* A^{(n-2)} \tilde{R}.$$

Вместо элементарных унитарных можно использовать несколько иные матрицы. Рассмотрим преобразование векторного пространства, осуществляющее отражение векторов от заданной плоскости Q. Обозначим через $\mathbf{w}$ вектор-столбец единичной длины, ортогональный к плоскости Q. Разложим произвольный вектор $\mathbf{z}$ на вектор, лежащий в плоскости Q и вектор, ортогональный ей:

$$\mathbf{z} = \mathbf{x} + \mathbf{y}.$$

Здесь вектор $\mathbf{x}$ лежит в плоскости Q, т.е. $(\mathbf{x}, \mathbf{w}) = 0$, а вектор $\mathbf{y}$ ортогонален ей, т.е. коллинеарен вектору $\mathbf{w}$, иными словами, $\mathbf{y} = \alpha \mathbf{w}$, где α — некоторое комплексное число. Отражаясь от плоскости Q, вектор $\mathbf{z}$ переходит в вектор

$$\tilde{\mathbf{z}} = A\mathbf{z} = \mathbf{x} - \mathbf{y}.$$

Докажем, что *преобразование отражения A является линейным.* Действительно, если λ_1 и λ_2 — комплексные числа,

$$\mathbf{z}_1 = \mathbf{x}_1 + \mathbf{y}_1, \quad \mathbf{z}_2 = \mathbf{x}_2 + \mathbf{y}_2,$$

где $(\mathbf{x}_1, \mathbf{w}) = 0$, $\mathbf{y}_1 = \alpha_1 \mathbf{w}$, $(\mathbf{x}_2, \mathbf{w}) = 0$, $\mathbf{y}_2 = \alpha_2 \mathbf{w}$, то $\lambda_1 \mathbf{z}_1 + \lambda_2 \mathbf{z}_2 = (\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) + (\lambda_1 \mathbf{y}_1 + {}$ $+ \lambda_2 \mathbf{y}_2)$, где $(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \mathbf{w}) = \lambda_1 (\mathbf{x}_1, \mathbf{w}) + \lambda_2 (\mathbf{x}_2, \mathbf{w}) = 0$, $\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2 = (\lambda_1 \alpha_1 + \lambda_2 \alpha_2) \mathbf{w}$. Поэтому $A(\lambda_1 \mathbf{z}_1 + \lambda_2 \mathbf{z}_2) = (\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) - (\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2) = \lambda_1 (\mathbf{x}_1 - \mathbf{y}_1) + \lambda_2 (\mathbf{x}_2 - \mathbf{y}_2) = \lambda_1 \tilde{\mathbf{z}}_1 + {}$ $+ \lambda_2 \tilde{\mathbf{z}}_2 = \lambda_1 A \mathbf{z}_1 + \lambda_2 A \mathbf{z}_2$.

Итак, преобразование отражения A является линейным. Матрицу U этого преобразования будем называть **матрицей отражения**. Для нее можно указать явный вид, а именно

$$U = E - 2\mathbf{w}\mathbf{w}^*.$$

Действительно,

$$U\mathbf{z} = (E - 2\mathbf{w}\mathbf{w}^*)\mathbf{z} = \mathbf{z} - 2\mathbf{w}\mathbf{w}^*\mathbf{z} = \mathbf{z} - 2\mathbf{w}(\mathbf{z}, \mathbf{w}) = \mathbf{z} - 2\mathbf{w}(\mathbf{x}, \mathbf{w}) - 2\mathbf{w}(\mathbf{y}, \mathbf{w}) = \mathbf{x} + \alpha\mathbf{w} - 2\mathbf{w}(\alpha\mathbf{w}, \mathbf{w}) =$$

$$= \mathbf{x} + \alpha\mathbf{w} - 2\alpha\mathbf{w} = \mathbf{x} - \alpha\mathbf{w} = \mathbf{x} - \mathbf{y} = \tilde{\mathbf{z}}.$$

Матрица отражения является унитарной матрицей. Действительно,

$$UU^* = (E - 2\mathbf{w}\mathbf{w}^*)(E - 2\mathbf{w}\mathbf{w}^*) = E - 4\mathbf{w}\mathbf{w}^* + 4\mathbf{w}\mathbf{w}^*\mathbf{w}\mathbf{w}^* = E - 4\mathbf{w}\mathbf{w}^* + 4\mathbf{w}(\mathbf{w}, \mathbf{w})\mathbf{w}^* =$$

$$= E - 4\mathbf{w}\mathbf{w}^* + 4\mathbf{w}\mathbf{w}^* = E.$$

Лемма 5.3.2. *Пусть $\mathbf{s}$ и $\mathbf{l}$ — произвольные вектора-столбцы, причем $\mathbf{l}$ имеет единичную длину. Вектор нормали $\mathbf{w}$ всегда можно выбрать таким образом, чтобы построенная по нему матрица отражения U переводила вектор $\mathbf{s}$ в вектор, коллинеарный $\mathbf{l}$.*

Д о к а з а т е л ь с т в о. Возьмем

$$\mathbf{w} = \frac{1}{\rho}(\mathbf{s} - \alpha\mathbf{l}),$$

где

$$|\alpha| = ||\mathbf{s}||_{II} = \sqrt{(\mathbf{s}, \mathbf{s})}, \quad \arg\alpha = \arg(\mathbf{s}, \mathbf{l}) - \pi, \quad \rho = ||\mathbf{s} - \alpha\mathbf{l})||_{II} = \sqrt{(\mathbf{s} - \alpha\mathbf{l}, \mathbf{s} - \alpha\mathbf{l})}.$$

$$(\mathbf{s}, \alpha\mathbf{l}) = \overline{\alpha}(\mathbf{s}, \mathbf{l}) = |\overline{\alpha}|e^{i\arg\overline{\alpha}}(\mathbf{s}, \mathbf{l}) = |\alpha|e^{\pi i}e^{-i\arg(\mathbf{s},\mathbf{l})}\cdot(\mathbf{s}, \mathbf{l}) = -|\alpha|e^{-i\arg(\mathbf{s},\mathbf{l})}|(\mathbf{s}, \mathbf{l})|e^{i\arg(\mathbf{s},\mathbf{l})} = -|\alpha||(\mathbf{s}, \mathbf{l})|.$$

Поэтому

$$(\mathbf{s}-\alpha\mathbf{l}, \mathbf{s}-\alpha\mathbf{l}) = (\mathbf{s}, \mathbf{s}) - (\alpha\mathbf{l}, \mathbf{s}) - (\mathbf{s}, \alpha\mathbf{l}) + (\alpha\mathbf{l}, \alpha\mathbf{l}) = |\alpha|^2 - (\mathbf{s}, \alpha\mathbf{l}) - \overline{(\mathbf{s}, \alpha\mathbf{l})} + |\alpha|^2 = 2|\alpha|^2 + 2|\alpha||(\mathbf{s}, \mathbf{l})|.$$

и, следовательно,

$$\rho = \sqrt{2|\alpha|^2 + 2|\alpha||(\mathbf{s}, \mathbf{l})|}.$$

При этом

$$U\mathbf{s} = \mathbf{s} - 2\mathbf{w}\mathbf{w}^*\mathbf{s} = \mathbf{s} - 2\mathbf{w}(\mathbf{s}, \mathbf{w}) = \mathbf{s} - \frac{2}{\rho}(\mathbf{s}, \mathbf{s}-\alpha\mathbf{l})\mathbf{w} = \mathbf{s} - \frac{2}{\rho}[(\mathbf{s}, \mathbf{s}) - (\mathbf{s}, \alpha\mathbf{l})]\mathbf{w} =$$

$$= \mathbf{s} - \frac{2}{\rho}[|\alpha|^2 + |\alpha||(\mathbf{s}, \mathbf{l})|]\mathbf{w} = \mathbf{s} - \frac{1}{\rho}[2|\alpha|^2 + 2|\alpha||(\mathbf{s}, \mathbf{l})|]\mathbf{w} = \mathbf{s} - \frac{1}{\rho}\rho^2\mathbf{w} = \mathbf{s} - \rho\mathbf{w} = \alpha\mathbf{l},$$

т.е. вектор $\mathbf{s}$ с помощью матрицы отражения переводится в вектор $\alpha\mathbf{l}$, коллинеарный $\mathbf{l}$.

Лемма 5.3.2 позволяет применять вместо элементарных унитарных матриц матрицы отражения в процессах получения разложений комплексной матрицы в теоремах 5.3.1 и 5.3.2. Рассмотрим, например, процесс разложения в теореме 5.3.1. Пусть A — произвольная комплексная матрица. Умножим ее слева на матрицу отражения U_1, выбирая в качестве $\mathbf{s}$ вектор $(a_{11}, a_{21}, \dots, a_{n1})^T$, а в качестве $\mathbf{l}$ вектор $\mathbf{l} = (1, 0, 0, \dots, 0)^T$. Очевидно, что после этого преобразования приходим к матрице

$$A^{(1)} = U_1 A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \cdot & \cdot & \dots & \cdot \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}.$$

На втором шаге строим матрицу отражения U_2 по векторам $\mathbf{s} = (0, a_{22}^{(1)}, \dots, a_{n2}^{(1)})^T$, $\mathbf{l} = (0, 1, 0, \dots, 0)^T$. При этом в матрице

$$A^{(2)} = U_2 A^{(1)} = U_2 U_1 A$$

будут равны нулю поддиагональные элементы в первых двух столбцах и т.д. В конце процесса получим правую треугольную матрицу.

Отметим, что *с точки зрения уменьшения вычислительной работы выгоднее использовать матрицы отражения, а не элементарные унитарные матрицы* $R_{ij}(\varphi, \psi)$.

Вещественные элементарные матрицы называются **элементарными матрицами вращения** или **матрицами простого поворота**. Они имеют следующий вид:

$$T_{ij}(\varphi) = \begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cos\varphi & & & -\sin\varphi & & \\ & & & & \cdot & & & & \\ & & & & & \cdot & & & \\ & & & \sin\varphi & & & \cos\varphi & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{pmatrix}.$$

Очевидно, матрицы $T_{ij}(\varphi)$ и $R_{ij}(\varphi, \psi)$ связаны соотношением

$$T_{ij}(\varphi) = R_{ij}(\varphi, 0).$$

Матрица A называется **эрмитовой**, если она удовлетворяет равенству

$$A = A^*.$$

Эти матрицы характеризуются такими соотношениями между элементами:

$$a_{ij} = \overline{a_{ji}} \tag{5.4.1}$$

Матрицы отражения являются эрмитовыми, так как

$$U^* = (E - 2\mathbf{w}\mathbf{w}^*)^* = E - 2\Big(\mathbf{w}^*\Big)^* \mathbf{w}^* = E - 2\mathbf{w}\mathbf{w}^* = U.$$

Эрмитовы матрицы обладают следующими свойствами.

1. *Если A — эрмитова, то ее определитель — действительное число.*

Действительно,

$$|A| = |A^*| = |\bar{A}^T| = |\bar{A}| = \overline{|A|}.$$

2. *Если A эрмитова, то A^{-1} также эрмитова.*

В самом деле,

$$(A^{-1})^* = (A^*)^{-1} = A^{-1}.$$

3. *Произведение двух перестановочных эрмитовых матриц является эрмитовой матрицей.*

Действительно, если A и B эрмитовы и $AB = BA$, то

$$AB = BA = B^*A^* = (AB)^*.$$

4. *Если матрица A эрмитова, то любая унитарно ей подобная матрица также эрмитова.*

В самом деле, пусть $B = U^*AU$, где U — унитарная матрица. Тогда

$$B^* = U^*A(U^*)^* = U^*AU = B.$$

5. *Эрмитова матрица порядка n имеет n взаимно ортогональных собственных векторов с действительными собственными значениями.*

Примем свойство 5 без доказательства.

Для любой квадратной матрицы A и любых векторов-столбцов $\mathbf{x}$ и $\mathbf{y}$

$$(A\mathbf{x}, \mathbf{y}) = \mathbf{y}^*A\mathbf{x} = (A^*\mathbf{y})^*\mathbf{x} = (\mathbf{x}, A^*\mathbf{y}).$$

В случае, когда матрица A эрмитова, это дает

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y}).$$

Полагая $\mathbf{y} = \mathbf{x}$, получим

$$(A\mathbf{x}, \mathbf{x}) = (\mathbf{x}, A\mathbf{x}) = \overline{(A\mathbf{x}, \mathbf{x})}.$$

Таким образом, *для эрмитовой матрицы скалярное произведение $(A\mathbf{x}, \mathbf{x})$ есть действительное число.*

Среди эрмитовых матриц особое место занимают положительно определенные матрицы.

Определение 5.4.1. Эрмитова матрица A называется **положительно определенной**, если скалярное произведение $(A\mathbf{x}, \mathbf{x}) > 0$ для любого вектора $\mathbf{x} \neq \mathbf{0}$.

Ввиду того, что

$$(A\mathbf{x}, \mathbf{x}) = \mathbf{x}^*A\mathbf{x} = \sum_{i,j=1}^{n} a_{ij}\overline{x_i}x_j$$

това матрица с матрицей A является положительно определенной тогда и только тогда, когда положительно определенной является квадратичная форма с матрицей A, т.е. когда для любых комплексных чисел $x_1, x_2, \ldots, x_n$, не равных нулю одновременно,

$$\sum_{i,j=1}^{n} a_{ij}\overline{x_i}x_j > 0.$$

Собственные значения положительно определенной матрицы строго положительны.

Действительно, пусть $\mathbf{u}$ — собственный вектор матрицы A, соответствующий собственному значению λ, т.е. $A\mathbf{u} = \lambda\mathbf{u}$, причем $\mathbf{u} \neq \mathbf{0}$. Так как A положительно определена и $\mathbf{u} \neq \mathbf{0}$, то

$$(A\mathbf{u}, \mathbf{u}) = (\lambda\mathbf{u}, \mathbf{u}) = \lambda(\mathbf{u}, \mathbf{u}) > 0.$$

Поскольку для $\mathbf{u} \neq \mathbf{0}$ выполняется $(\mathbf{u}, \mathbf{u}) = ||\mathbf{u}||_{II}^2 > 0$, то из последнего неравенства следует, что $\lambda > 0$.

Матрицы вида

$$A = B^*B,$$

где B — прямоугольная матрица с линейно независимыми столбцами, являются положительно определенными.

В самом деле,

$$(B^*B\mathbf{x}, \mathbf{x}) = (B^*(B\mathbf{x}), \mathbf{x}) = (B\mathbf{x}, B\mathbf{x}) > 0$$

для всех $\mathbf{x} \neq \mathbf{0}$.

Вещественные эрмитовы матрицы называются **симметрическими**. Так как для вещественных матриц $A = \bar{A}$, то определение симметрической матрицы перепишется в виде

$$A^T = A.$$

Заметим, что треугольная эрмитова матрица является диагональной матрицей. Поэтому теорема 5.3.1 и следствие 5.3.1 для эрмитового случая примут вид

Теорема 5.4.1. *Любая эрмитова матрица A преобразуется в диагональную матрицу посредством умножения слева на конечную цепочку элементарных унитарных матриц $R_{ij}(\varphi, \psi)$.*

Следствие 5.4.1. *Любая эрмитова матрица есть произведение унитарной матрицы на диагональную матрицу.*

Легко также понять, что почти треугольная эрмитова матрица является трехдиагональной матрицей. Для этого случая теорема 5.4.2 и следствие 5.4.2 примут форму

Теорема 5.4.2. *Любая эрмитова матрица A приводится к эрмитовой трехдиагональной матрице с помощью конечной цепочки подобных преобразований с матрицами $R_{ij}(\varphi, \psi)$.*

Следствие 5.3.4. *Любая эрмитова матрица унитарно подобна трехдиагональной эрмитовой матрице.*

В методах вычислений важную роль играет следующая *трица U такая, что матрица*

$$B = U^* A U$$

является правой треугольной матрицей.

Д о к а з а т е л ь с т в о. Из теоремы 1.1.2 следует, что каждая матрица имеет хотя бы один собственный вектор. Пусть $\mathbf{u}_1$ — собственный вектор матрицы A, соответствующий собственному значению λ_1. Построим унитарную матрицу U_1, первый столбец которой пропорционален $\mathbf{u}_1$. Тогда матрица

$$B_1 = U_1^* A U_1$$

будет иметь следующий вид:

$$B_1 = \begin{pmatrix} \lambda_1 & b_{12}^{(1)} & \dots & b_{1n}^{(1)} \\ 0 & b_{22}^{(1)} & \dots & b_{2n}^{(1)} \\ \cdot & \cdot & \dots & \cdot \\ 0 & b_{n2}^{(1)} & \dots & b_{nn}^{(1)} \end{pmatrix}.$$

Так как $\mathbf{u}_1$ — собственный вектор матрицы A, то первый столбец матрицы AU_1 будет совпадать с вектором $\lambda_1 \mathbf{u}_1$. По построению матрицы U_1 скалярное произведение вектора $\mathbf{u}_1$ на первую вектор-строку матрицы U_1^* равно единице, а на остальные — нулю, что и подтверждает вид матрицы B_1.

Применяя те же рассуждения к матрице

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & b_{22}^{(1)} & \dots & b_{2n}^{(1)} \\ \cdot & \cdot & \dots & \cdot \\ 0 & b_{n2}^{(1)} & \dots & b_{nn}^{(1)} \end{pmatrix},$$

получим

$$B_2 = U_2^* B_1 U_2 = \begin{pmatrix} \lambda_1 & b_{12}^{(2)} & b_{13}^{(2)} & \dots & b_{1n}^{(2)} \\ 0 & \lambda_2 & b_{23}^{(2)} & \dots & b_{2n}^{(2)} \\ 0 & 0 & b_{33}^{(2)} & \dots & b_{3n}^{(2)} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & b_{n3}^{(2)} & \dots & b_{nn}^{(2)} \end{pmatrix}.$$

Продолжим процесс далее. Тогда, принимая во внимание, что произведение унитарных матриц есть матрица унитарная, придем к требуемому результату.

6.1. Критерии регулярности. Условия Адамара.

Пусть $A = (a_{ij})$ — произвольная квадратная матрица порядка n с комплексными элементами. Допустим, что эта матрица вырождена, т.е. $|A| = 0$. Тогда столбцы матрицы линейно зависимы, т.е. существуют числа $x_1, x_2, \ldots, x_n$, не равные нулю одновременно такие, что линейная комбинация этих векторов равна нулю:

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{n1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{n2} \end{pmatrix} x_2 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{nn} \end{pmatrix} x_n = \mathbf{0},$$

т.е.

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = \overline{1,n}.$$

Пусть $k \in \{1, 2, \ldots, n\}$ таково, что $|x_k| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. Так как не все $x_i = 0$, то $|x_k| > 0$. Для выбранного k, в частности,

$$\sum_{j=1}^{n} a_{kj} x_j = 0.$$

Но тогда

$$|a_{kk}||x_k| = \Big| - \sum_{j \neq k} a_{kj} x_j \Big| \leq \sum_{j \neq k} |a_{kj}||x_j| \leq |x_k| \sum_{j \neq k} |a_{kj}|.$$

Сокращая на $|x_k|$, получаем

$$|a_{kk}| \leq \sum_{j \neq k} |a_{kj}|. \tag{6.1.1}$$

Поэтому, если выполняются **условия Адамара**

$$H_i \equiv |a_{ii}| - \sum_{j \neq i} |a_{ij}| > 0 \quad i = \overline{1,n}, \tag{6.1.2}$$

то неравенство (6.1.1) невозможно и, следовательно, матрица A является **регулярной** (невырожденной), т.е. $|A| \neq 0$.

Таким образом, справедлива

Теорема 6.1.1 (теорема Адамара). *Если для квадратной матрицы n-го порядка выполняются n неравенств (6.1.2), то матрица A является невырожденной.*

Условие $H_i > 0$ означает, что модуль диагонального элемента a_{ii} строго больше суммы модулей всех остальных элементов i-й строки. Такой элемент называется **доминирующим** для своей строки. Условие Адамара требует, чтобы все диагональные элементы матрицы A были доминирующими для своих строк.

Теорема 6.1.2. *Если выполняются условия Адамара (6.1.2), то для модуля определителя A справедлива следующая оценка снизу:*

$$mod|A| \geq H_1 H_2 \ldots H_n > 0. \tag{6.1.3}$$

$$f_{ij}=\frac{a_{ij}}{H_i},\qquad i,j=\overline{1,n},$$

для которой

$$|f_{ii}|-\sum_{j\neq i}|f_{ij}|=\frac{1}{H_i}(a_{ij}-\sum_{j\neq i}|a_{ij}|)=1,\quad i=\overline{1,n}. \tag{6.1.4}$$

Пусть λ_0 — собственное значение этой матрицы, которому соответствует собственный вектор $\mathbf{x}=(x_1,x_2,\dots,x_n)^T$. Пусть k таково, что $|x_k|=\max\{|x_1|,|x_2|,\dots,|x_n|\}$, причем $|x_k|>0$. Тогда

$$F\mathbf{x}=\lambda_0\mathbf{x}.$$

Приравнивая k-ю координату, получим

$$\lambda_0 x_k=\sum_j f_{kj}x_j.$$

Из этого равенства с учетом соотношений (6.1.4) получаем

$$|\lambda_0||x_k|=\Big|f_{kk}x_k+\sum_{j\neq k}f_{kj}x_j\Big|\geq|f_{kk}x_k|-\Big|\sum_{j\neq k}f_{kj}x_j\Big|\geq|f_{kk}x_k|-\sum_{j\neq k}|f_{kj}||x_j|\geq$$

$$\geq|x_k|\Big(|f_{kk}|-\sum_{j\neq k}|f_{kj}|\Big)=|x_k|.$$

Сокращая на $|x_k|$, найдем

$$|\lambda_0|\geq 1.$$

Но определитель $|F|$ равен произведению собственных значений матрицы F. Каждое из них по модулю больше или равно единицы. Поэтому и

$$mod|F|\geq 1. \tag{6.1.5}$$

С другой стороны,

$$|F|=\frac{|A|}{H_1H_2\dots H_n}. \tag{6.1.6}$$

Подставляя (6.1.6) в неравенство (6.1.5), получим искомое неравенство (6.1.3).

Отметим, что для класса матриц, удовлетворяющих условиям Адамара с заданными значениями $H_1,H_2,\dots,H_n$, оценка (6.1.3) не улучшаема, поскольку при $A=(H_i\delta_{ik})$ неравенство (6.1.3) превращается в равенство.

Замечание 6.1.1. Так как $|A|=|A^T|$, то заменяя матрицу A транспонированной матрицей A^T, получаем достаточные условия невырожденности матрицы A в виде условий Адамара для столбцов

$$G_i\equiv|a_{ii}|-\sum_{j\neq i}|a_{ji}|>0,\quad i=\overline{1,n}. \tag{6.1.7}$$

При выполнении этих условий вместо (6.1.3) получится оценка

$$mod|A|\geq G_1G_2\dots G_n. \tag{6.1.8}$$

Пусть C — произвольная невырожденная квадратная матрица n-го порядка. Тогда матрицы A и AC одновременно являются невырожденными. Поэтому в условиях (6.1.2), (6.1.7), а также в оценках (6.1.3), (6.1.8) можно матрицу A заменить на матрицу AC. Варьируя матрицу C, будем получать различные (не эквивалентные между собой) достаточные условия помощью подбора надлежащей матрицы C можно осуществить произвольную перестановку столбцов. Тогда вместо условий (6.1.2) получим условия

$$H_i' \equiv |a_{i\mu_i}| - \sum_{j\neq\mu_i} |a_{ij}| > 0, \quad i = \overline{1,n},$$

где $(\mu_1, \mu_2, \ldots, \mu_n)$ — фиксированная, но произвольная перестановка индексов $1, 2, \ldots, n$. Итак, справедлива

Теорема 6.1.3. *Матрица $A = (a_{ij})$ является невырожденной, если в каждой ее строке имеется доминирующий (не обязательно диагональный) элемент и эти n доминирующих элементов расположены в различных столбцах.*

Аналогичный результат имеет место для столбцов.

6.2. Локализация собственных значений. Круги Гершгорина.

Пусть $A = (a_{ij})$ — произвольная квадратная матрица n-го порядка и λ — некоторое ее собственное значение. Тогда $|A - \lambda E| = 0$, т.е. $A - \lambda E$ — вырожденная матрица. Следовательно для нее не могут выполняться все условия Адамара, т.е. должно иметь место хотя бы одно из неравенств

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|, \quad i = \overline{1,n}. \tag{6.2.1}$$

Каждое из неравенств (6.2.1) определяет некоторый круг в комплексной λ-плоскости радиуса $\sum\limits_{j\neq i} |a_{ij}|$ с центром в точке a_{ii}, называемый **кругом Гершгорина**. Поэтому справедлива

Теорема 6.2.1 (теорема Гершгорина). *Каждое собственное значение λ матрицы $A = (a_{ij})$ всегда расположено в одном из кругов (6.2.1).*

Таким образом, объединение всех кругов Гершгорина (6.2.1) дает некоторую **область локализации** собственных значений матрицы A, т.е. область, в которой заведомо лежат все собственные значения матрицы A. Каждый критерий регулярности приводит к своей области локализации собственных значений. Так, например, исходя из условий Адамара для столбцов, мы получим область локализации в виде объединения n кругов

$$|\lambda - a_{ii}| \leq \sum_{j\neq i} |a_{ji}|, \quad i = \overline{1,n}.$$

Пример 6.2.1. Рассмотрим симметрическую матрицу четвертого порядка

$$A = \begin{pmatrix} 0 & 4 & 1 & -1 \\ 4 & 0 & -1 & 1 \\ 1 & -1 & -1 & 15 \\ -1 & 1 & 15 & -1 \end{pmatrix}.$$

Поскольку матрица A — симметрическая, то у нее все собственные значения действительные. Поэтому вместо областей локализации в комплексной λ-плоскости можно рассматривать отрезки, высекаемые этими областями на вещественной λ-оси. "Круги" Гершгорина имеют вид

$$|\lambda| \leq 6 \quad \text{для 1-й и 2-й строк;}$$

$$|\lambda + 1| \leq 17 \quad \text{для 3-й и 4-й строк.}$$

собственные значения лежат на этом отрезке.

Пусть $P_i = \sum\limits_{j \neq i} |a_{ij}|$ — сумма модулей элементов i-й строки матрицы A без диагонального элемента, $Q_j = \sum\limits_{j \neq i} |a_{ji}|$ — сумма модулей элементов j-го столбца матрицы A без диагонального элемента.

Имеет место следующий результат, который мы примем без доказательства.

Лемма 6.2.1 (леммма Островского). *Если при некотором $\alpha \in [0, 1]$ элементы квадратной матрицы A над полем комплексных чисел $\mathbb{C}$ удовлетворяют неравенствам*

$$|a_{ii}||a_{jj}| > P_i^{\alpha} Q_j^{1-\alpha}, \quad i, j = \overline{1, n},$$

то матрица A невырождена.

Очевидным следствием леммы является

Теорема 6.2.2 (теорема Островского). *Если A — квадратная матрица над полем $\mathbb{C}$, то все ее собственные значения находятся в объединении овалов Островского*

$$|\lambda - a_{ii}||\lambda - a_{jj}| \leq P_i^{\alpha} Q_j^{1-\alpha}, \quad i, j = \overline{1, n}.$$

Задача 6.2.1. Определите область локализации собственных значений матрицы примера 6.2.1 с помощью достаточного условия Островского. Сравните области локализации Гершгорина и Островского. Найдите их пересечение, являющееся, очевидно, также областью локализации собственных значений. Для определенности возьмите в теореме Островского $\alpha = \frac{1}{2}$.

Теорема 1.2.1. *Любая λ-матрица эквивалентна некоторой и притом единственной канонической λ-матрице (т.е. приводится с помощью конечного числа элементарных преобразований к каноническому виду).*

Доказательство. Доказательство проведем с помощью индукции по n, где n — порядок λ-матрицы.

1. При $n=1$

$$A(\lambda) = (a(\lambda)).$$

Если $a(\lambda)=0$, то $A(\lambda)$ уже каноническая. Если же $a(\lambda)\neq 0$, то разделив многочлен $a(\lambda)$ на его старший коэффициент (это элементарное преобразование), получим каноническую матрицу.

2. Предположим, что утверждение теоремы выполнено для λ-матриц порядка $n-1$ и докажем, что тогда оно выполняется для λ-матриц порядка n. Рассмотрим произвольную λ-матрицу $A(\lambda)$ порядка n. Если она нулевая, то она уже каноническая и все доказано. Поэтому можно считать, что у матрицы $A(\lambda)$ имеются ненулевые элементы.

Переставляя строки и столбцы $A(\lambda)$, можно перевести один из ее ненулевых элементов в верхний левый угол. Следовательно, среди λ-матриц, эквивалентных $A(\lambda)$, есть такие, в левом верхнем углу которых стоит ненулевой многочлен. Рассмотрим все такие матрицы. Среди всех λ-матриц, эквивалентных $A(\lambda)$ и имеющих ненулевой элемент в верхнем левом углу, можно найти такую, что многочлен, стоящий в ее верхнем левом углу, имеет наименьшую возможную степень. Деля первую строку этой матрицы на старший козффициент этого многочлена, мы получим следующую λ-матрицу, эквивалентную λ-матрице $A(\lambda)$:

$$A(\lambda) = \begin{pmatrix} e_1(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ b_{21}(\lambda) & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \cdot & \cdot & \dots & \cdot \\ b_{n1}(\lambda) & b_{n2}(\lambda) & \dots & b_{nn}(\lambda) \end{pmatrix}.$$

В этой матрице $e_1(\lambda)\neq 0$, старший коэффициент $e_1(\lambda)$ равен 1 и *никакой комбинацией элементарных преобразований нельзя перейти от полученной матрицы к такой матрице, в левом верхнем углу которой стоял бы ненулевой многочлен меньшей степени.*

Докажем, что все элементы первой строки и первого столбца полученной матрицы делятся на $e_1(\lambda)$. Пусть, например, для $2\le j\le n$

$$b_{1j}(\lambda) = e_1(\lambda)q(\lambda) + r(\lambda), \quad \text{где степень } r(\lambda) < \text{ степень } e_1(\lambda).$$

Вычитая из j-го столбца первый столбец, умноженный на $q(\lambda)$, а затем переставляя первый и j-й столбцы, придем к матрице, эквивалентной $A(\lambda)$, в левом верхнем углу которой стоит $r(\lambda)$, т.е. многочлен меньшей степени, чем $e_1(\lambda)$, что возможно только в случае, когда $e_1(\lambda) = 0$. Это и требовалось доказать. Для первого столбца доказательство аналогично.

Вычитая из j-го столбца первый столбец, умноженный на $q(\lambda)$, мы заменим элемент $b_{1j}(\lambda)$ нулем. Делая такие преобразования для $j=\overline{2,n}$, мы заменим нулями все элементы $b_{1j}(\lambda)$. Аналогичным образом заменяются нулями все $b_{i1}(\lambda),\ \ i=\overline{2,n}$. Поэтому мы придем к следующей матрице, эквивалентной $A(\lambda)$, в левом верхнем углу которой стоит многочлен $e_1(\lambda)$, а все остальные элементы первой строки и первого столбца равны нулю:

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \cdot & \cdot & \dots & \cdot \\ 0 & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix}. \tag{П1.1}$$

По предположению индукции, матрица $n-1$ го порядка, стоящая в правом нижнем углу последней матрицы, элементарными преобразованиями приводится к каноническому виду:

$$\begin{pmatrix} c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \cdot & \dots & \cdot \\ c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix} \sim \begin{pmatrix} e_2(\lambda) & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & e_n(\lambda) \end{pmatrix}.$$

этом первая строка и первый столбец не изменятся — мы получим, что

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & e_n(\lambda) \end{pmatrix}. \tag{П1.2}$$

Для доказательства того, что эта матрица каноническая, остается доказать, что $e_2(\lambda)$ делится на $e_1(\lambda)$. Предположим противное: пусть

$$e_2(\lambda) = e_1(\lambda)q(\lambda) + r(\lambda),$$

где $r(\lambda) \neq 0$ и степень $r(\lambda) <$ степень $e_1(\lambda)$. Прибавляя ко второму столбцу матрицы (П1.2) ее первый столбец, умноженный на $q(\lambda)$, а затем вычитая из второй строки первую, мы заменим элемент $e_2(\lambda)$ элементом $r(\lambda)$. Переставляя первые две строки и первые два столбца, мы переместим $r(\lambda)$ в верхний левый угол матрицы. Приходим к противоречию с выбором $e_1(\lambda)$. Существование канонической λ-матрицы, эквивалентной $A(\lambda)$, доказано.

Для доказательства ее единственности сначала докажем лемму 1.2.1.

Лемма 1.2.1. *Многочлены* $d_k(\lambda),\ \ 1 \le k \le n,$ *не меняются при выролнении элементарных преобразований над матрицей* $A(\lambda)$.

Доказательство.

1. Рассмотрим *элементарное преобразование 1-го типа (2-го типа).* Если i-я строка умножается на $\alpha \in P,\ \ \alpha \neq 0$, то миноры k-го порядка, через которые i-я строка проходит, будут умножаться на α, остальные миноры не изменятся. Однако наибольший общий делитель многочленов не меняется, если какие-то из многочленов умножить на $\alpha \in P$. Для умножения столбца на $\alpha \in P,\ \ \alpha \neq 0$ доказательство аналогичное.

2. Рассмотрим *элементарное преобразование 3-го типа (4-го типа).* Пусть к i-й строке $A(\lambda)$ прибавляется j-я строка, умноженная на многочлен $\varphi(\lambda)$. Обозначим через $\bar{A}(\lambda)$ получающуюся при этом матрицу, а наибольший общий делитель ее миноров k-го порядка со старшим коэффициентом 1 — через $\bar{d}_k(\lambda)$. Миноры, через которые проходит i-я строка, не изменятся. Не меняются миноры, через которые проходят как i-я, так и j-я строки (определитель не меняется при прибавлении к строке другой строки, умноженной на число). Остается случай, когда i-я строка проходит через минор k-го порядка, а j-я не проходит. Возьмем любой такой минор и обозначим его через M. Очевидно,

$$\bar{M} = M + \varphi(\lambda)M',$$

где M' — минор $A(\lambda)$, получающийся из M заменой элементов i-й строки $A(\lambda)$ соответствующими элементами j-й строки. Так как M и $\bar{M}$ — миноры k-го порядка $A(\lambda)$, то они делятся на $d_k(\lambda)$. Значит, все миноры k-го порядка матрицы $\bar{A}(\lambda)$ делятся на $d_k(\lambda)$, поэтому $\bar{d}_k(\lambda)$ делится на $d_k(\lambda)$. Но для рассматриваемого элементарного преобразования существует обратное элементарное преобразование того же типа, следовательно, $d_k(\lambda)$ делится на $\bar{d}_k(\lambda)$. Так как старшие коэффициенты многочленов $d_k(\lambda),\ \ \bar{d}_k(\lambda)$ равны 1, то они совпадают: $d_k(\lambda) = \bar{d}_k(\lambda)$. Для преобразования 4-го типа доказательство аналогичное.

Таким образом, *всем* λ*-матрицам, эквивалентным матрице* $A(\lambda)$*, соответствует один и тот же набор многочленов* $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$.

Для доказательства теоремы 1.2.1 остается доказать единственность канонической матрицы. Пусть

$$\begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & e_n(\lambda) \end{pmatrix} -$$

Минор k-го порядка, стоящий в левом верхнем углу есть

$$e_1(\lambda)e_2(\lambda)\dots e_k(\lambda). \tag{П1.3}$$

Возьмем минор k-го порядка в строках с номерами $i_1, i_2, \dots, i_k$, где $i_1 < i_2 < \dots < i_k$ и в столбцах с теми же номерами. Он равен $e_{i_1}(\lambda)e_{i_2}(\lambda)\dots e_{i_k}(\lambda)$, что делится на (П1.3). Действительно, из того, что $1 \le i_1$, следует, что $e_{i_1}(\lambda)$ делится на $e_1(\lambda)$, из того, что $2 \le i_2$, следует, что $e_{i_2}(\lambda)$ делится на $e_2(\lambda)$, и т.д. Если в матрице взят минор k-го порядка, через который хотя бы для одного i проходит i-я строка, но не проходит i-й столбец, то этот минор содержит нулевую строку и поэтому равен 0. Итак, все миноры k-го порядка делятся на (П1.3), следовательно,

$$d_k(\lambda) = e_1(\lambda)e_2(\lambda)\dots e_k(\lambda), \quad k = \overline{1,n} \tag{П1.4}$$

Остается показать, что *многочлены* $e_k(\lambda)$, $k = \overline{1,n}$, *однозначно определяются матрицей* $A(\lambda)$. Пусть ранг матрицы $A(\lambda)$ равен r, тогда $d_r(\lambda) \neq 0$, $d_{r+1}(\lambda) = 0$. В силу (П1.4) отсюда следует, что $e_{r+1}(\lambda) = 0$. В силу свойств канонической матрицы отсюда вытекает, что если ранг $A(\lambda) = r < n$, то $e_{r+1}(\lambda) = \dots = e_n(\lambda) = 0$. С другой стороны для $k \le r$ из (П1.4) следует, что

$$e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)}.$$

Доказательство завершено.

Одновременно с доказательством мы получили способ непосредственного вычисления многочленов $e_k(\lambda)$, называемых **инвариантными множителями матрицы** $A(\lambda)$.

1.3.2.

Основная теорема о подобии матриц. *Матрицы A и B с элементами из поля P подобны тогда и только тогда, когда их характеристические матрицы $A-\lambda E$ и $B-\lambda E$ эквивалентны.*

Прежде, чем доказывать эту теорему, введем понятие унимодулярных матриц и установим вспомогательные результаты.

Из предыдущих утверждений следует, что критерий эквивалентности λ-матриц можно сформулировать в следующих двух видах.

Две λ-матрицы эквивалентны тогда и только тогда, когда они приводятся к одному и тому же каноническому виду.

Две λ-матрицы эквивалентны тогда и только тогда, когда они обдадают одинаковыми инвариантными множителями.

Выведем еще один критерий, имеющий иной характер.

Определение П2.1. λ-матрица $U(\lambda)$ называется **унимодулярной**, если ее канонической матрицей является единичная матрица, т.е. если все ее инвариантные множители равны единице.

Лемма П2.1. *λ-матрица $U(\lambda)$ унимодулярна тогда и только тогда, когда ее определитель $|U(\lambda)|$ отличен от нуля и не зависит от λ.*

Д о к а з а т е л ь с т в о. Если $U(\lambda)$ унимодулярна, т.е. $U(\lambda)\sim E$, то этим матрицам соответствует один и тот же $d_n(\lambda)$. Но для единичной матрицы $d_n(\lambda)=1$, значит $d_n(\lambda)=1$ для $U(\lambda)$. Но

$$d_n(\lambda)=\frac{|U(\lambda)|}{\text{старший коэффициент } U(\lambda)}.$$

Следовательно, $|U(\lambda)|=[\text{число}\neq 0]\cdot d_n(\lambda)=[\text{число}\neq 0]$.

Обратно, если $|U(\lambda)|\neq 0$ не зависит от λ, то $d_n(\lambda)=1$ для $U(\lambda)$, но $d_n(\lambda)=e_1(\lambda)e_2(\lambda)\ldots \ldots e_n(\lambda)$, откуда все $e_i(\lambda)=1$, т.е. канонической для $U(\lambda)$ будет E.

Следствие П2.1. *Всякая невырожденная числовая матрица является унимодулярной.*

Унимодулярная матрица, тем не менее, может иметь сложный вид. Например, λ-матрица

$$\begin{pmatrix} \lambda & \lambda^3+5 \\ \lambda^2-\lambda-4 & \lambda^4-\lambda^3-4\lambda^2+5\lambda-5 \end{pmatrix}$$

унимодулярна, так как ее определитель равен 20.

Следствие П2.2. *Произведение унимодулярных матриц является унимодулярной матрицей.*

Д о к а з а т е л ь с т в о. Пусть $A(\lambda)$ и $B(\lambda)$ унимодулярны. По теореме П2.1 определители $|A(\lambda)|$ и $|B(\lambda)|$ отличны от 0 и не зависят от λ. Тогда определитель $|A(\lambda)B(\lambda)|= =|A(\lambda)||B(\lambda)|\neq 0$ и не зависит от λ. По теореме П2.1 матрица $A(\lambda)B(\lambda)$ унимодулярна.

Теорема П2.1. *λ-матрица $U(\lambda)$ унимодулярна тогда и только тогда, когда для нее существует обратная матрица, также являющаяся λ-матрицей.*

Д о к а з а т е л ь с т в о. *Необходимость*. Пусть $U(\lambda)$ унимодулярна. При подсчете обратной матрицы мы делим алгебраические дополнения на определитель. Так как определитель

Значит, $U^{-1}(\lambda)$ — λ-матрица.

Достаточность. Пусть существует $U^{-1}(\lambda)$, являющаяся λ-матрицей. Тогда $|U(\lambda)|$ — многочлен. Так как

$$|U(\lambda)||U^{-1}(\lambda)| = 1,$$

то оба многочлена $|U(\lambda)|$ и $|U^{-1}(\lambda)|$ имеют степень 0, т.е. являются числами, отличными от нуля.

Следствие П2.3. *λ-матрица, обратная к унимодулярной матрице, сама унимодулярна.*

Теорема П2.2 (критерий эквивалентности λ-матриц). *Две λ-матрицы $A(\lambda)$ и $B(\lambda)$ порядка n эквивалентны тогда и только тогда, когда существуют унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$ того же порядка n такие, что*

$$B(\lambda) = U(\lambda)A(\lambda)V(\lambda).$$

Д о к а з а т е л ь с т в о. *Необходимость.* Пусть $A(\lambda) \sim B(\lambda)$. Тогда от $A(\lambda)$ можно перейти к $B(\lambda)$ с помощью конечного числа элементарных преобразований. Заменяя каждое из них умножением слева или справа на элементарную матрицу, получим:

$$B(\lambda) = U_1(\lambda)\ldots U_k(\lambda)A(\lambda)V_1(\lambda)\ldots V_l(\lambda), \qquad (\text{П}2.1)$$

где все матрицы $V_i(\lambda)$ и $V_j(\lambda)$ элементарны, а, значит, унимодулярны. Обозначим $U(\lambda) = U_1(\lambda)\ldots U_k(\lambda)$, $V(\lambda) = V_1(\lambda)\ldots V_l(\lambda)$. Они унимодулярны как произведение унимодулярных матриц. Итак, нашлись унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$ такие, что $B(\lambda) = U(\lambda)A(\lambda)V(\lambda)$. Необходимость доказана.

Из проведенной части доказательства следует

Следствие П2.4. *λ-матрица унимодулярна тогда и только тогда, когда она представима в виде произведения элементарных матриц.*

Действительно, пусть λ-матрица представима в виде произведения элементарных матриц. Так как элементарные матрицы унимодулярны и произведение унимодулярных матриц унимодулярно, то λ-матрица унимодулярна.

Обратно, если $W(\lambda)$ — произвольная унимодулярная матрица, то $W(\lambda) \sim E$. По схеме доказательства необходимости теоремы П2.2, примененному вместо $B(\lambda)$ и $A(\lambda)$ к $W(\lambda)$ и E, из (П2.1) получим

$$W(\lambda) = U_1(\lambda)\ldots U_k(\lambda)V_1(\lambda)\ldots V_l(\lambda),$$

т.е. $W(\lambda)$ представима в виде произведения элементарных матриц.

Достаточность. Пусть для $A(\lambda)$ и $B(\lambda)$ существуют унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$ такие, что $B(\lambda) = U(\lambda)A(\lambda)V(\lambda)$. По доказанному $U(\lambda)$ и $V(\lambda)$ можно представить в виде произведений элементарных матриц

$$U(\lambda) = U_1(\lambda)\ldots U_k(\lambda), \quad V(\lambda) = V_1(\lambda)\ldots V_l(\lambda).$$

Тогда

$$B(\lambda) = U_1(\lambda)\ldots U_k(\lambda)A(\lambda)V_1(\lambda)\ldots V_l(\lambda).$$

Так как каждое умножение на элементарную матрицу эквивалентно элементарному преобразованию, то $B(\lambda)$ получается из $A(\lambda)$ с помощью $k+l$, т.е. конечного числа элементарных преобразований. Это и означает, что $A(\lambda) \sim B(\lambda)$.

Необходимость. Пусть A и B подобны, т.е. существует невырожденная матрица C такая, что

$$B = C^{-1}AC.$$

Тогда

$$C^{-1}(A - \lambda E)C = C^{-1}AC - \lambda(C^{-1}EC) = B - \lambda E.$$

Невырожденные числовые матрицы C^{-1} и C являются унимодулярными. По критерию эквивалентности λ-матриц (теореме П2.2) $A - \lambda E \sim B - \lambda E$.

Достаточность. Пусть $A - \lambda E \sim B - \lambda E$. Тогда существуют унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$ такие, что

$$U(\lambda)(A - \lambda E)V(\lambda) = B - \lambda E. \tag{П2.2}$$

Учитывая, что для унимодулярных матриц обратные матрицы существуют и являются λ-матрицами, получим:

$$U(\lambda)(A - \lambda E) = (B - \lambda E)V^{-1}(\lambda), \tag{П2.3}$$

$$(A - \lambda E)V(\lambda) = U^{-1}(\lambda)(B - \lambda E). \tag{П2.4}$$

Произведем левое деление $U(\lambda)$ на $B - \lambda E$ (степень $B - \lambda E$ равна 1 и $B - \lambda E$ — регулярный многочлен — его старший коэффициент $-E$ имеет определитель, отличный от 0):

$$U(\lambda) = (B - \lambda E)Q_1(\lambda) + R_1, \tag{П2.5}$$

и правое деление $V(\lambda)$ на $(B - \lambda E)$:

$$V(\lambda) = Q_2(\lambda)(B - \lambda E) + R_2, \tag{П2.6}$$

где степени R_1 и R_2 меньше 1, т.е. R_1 и R_2 — постоянные.

С учетом (П2.2) из (П2.5), (П2.6)

$$R_1(A - \lambda E)R_2 = (U(\lambda) - (B - \lambda E)Q_1(\lambda))(A - \lambda E)(V(\lambda) - Q_2(\lambda)(B - \lambda E)) =$$

$$= B - \lambda E - U(\lambda)(A - \lambda E)Q_2(\lambda)(B - \lambda E) - (B - \lambda E)Q_1(\lambda)(A - \lambda E)V(\lambda) +$$

$$+ (B - \lambda E)Q_1(\lambda)(A - \lambda E)Q_2(\lambda)(B - \lambda E).$$

Используя (П2.3), (П2.4), получим

$$R_1(A - \lambda E)R_2 = (B - \lambda E) - (B - \lambda E)V^{-1}(\lambda)Q_2(\lambda)(B - \lambda E) - (B - \lambda E)Q_1(\lambda)U^{-1}(\lambda)(B - \lambda E) +$$

$$+ (B - \lambda E)Q_1(\lambda)(A - \lambda E)Q_2(\lambda)(B - \lambda E) = (B - \lambda E)\{E - [V^{-1}(\lambda)Q_2(\lambda) + Q_1(\lambda)U^{-1}(\lambda) -$$

$$- Q_1(\lambda)(A - \lambda E)Q_2(\lambda)](B - \lambda E)\}.$$

Но выражение в квадратных скобках равно 0. Если предположить противное, то поскольку это выражение — λ-матрица (ведь $U^{-1}(\lambda)$, $V^{-1}(\lambda)$ — λ-матрицы), то ее степень неотрицательна. Но тогда степень выражения в фигурных скобках больше или равна 1, откуда степень $R_1(A - \lambda E)R_2 \geq 2$ — противоречие.

Следовательно,

$$R_1(A - \lambda E)R_2 = B - \lambda E,$$

откуда

$$R_1AR_2 = B, \quad R_1R_2 = E.$$

Из второго равенства следует, что числовая матрица R_2 невырождена, причем $R_2^{-1} = R_1$. Первое равенство тогда примет вид

$$B = R_2^{-1}AR_2,$$

Одновременно мы научились находить ту невырожденную матрицу R_2, которую переобозначим R, которая трансформирует матрицу A в матрицу B. Именно, если $A-\lambda E \sim B-\lambda E$, то $A-\lambda E$ конечным числом элементарных преобразований переводится в $B-\lambda E$. Берем те из этих преобразований, которые относятся к столбцам, и произведение соответствующих элементарных матриц, взятых в том же порядке, обозначим через $V(\lambda E)$. Делим $V(\lambda E)$ справа на $B-\lambda E$. Остаток от этого деления и будет матрицей R.

Указанное деление можно не выполнять, а воспользоваться следующей леммой.

Лемма 1.3.2. *Пусть*

$$V(\lambda) = V_0\lambda^s + V_1\lambda^{s-1} + \ldots + V_{s-1}\lambda + V_s, \quad V_0 \neq 0. \tag{П2.7}$$

Если

$$V(\lambda) = Q(\lambda)(\lambda E - B) + R, \tag{П2.8}$$

то

$$R = V_0B^s + V_1B^{s-1} + \ldots + V_{s-1}B + V_s. \tag{П2.9}$$

Доказательство. Вспомним, что ранее мы обозначали $R = R_2$ и $Q(\lambda) = Q_2(\lambda)$. Доказательство состоит в последовательной проверке справедливости равенства (П2.8), если многочлен $V(\lambda)$ будет заменен его записью (П2.7), вместо $R = R_2$ будет подставлено (П2.9), а в качестве $Q(\lambda) = Q_2(\lambda)$ будет взят многочлен

$$Q(\lambda) = Q_2(\lambda) = \lambda^{s-1}V_0+\lambda^{s-2}(BV_0+V_1)+\lambda^{s-3}(B^2V_0+BV_1+V_2)+\ldots+(B^{s-1}V_0+B^{s-2}V_1+\ldots+V_{s-1}).$$

Предлагается проделать это самостоятельно.

нормальной формы, и обоснование алгоритма приведения матрицы к жордановой нормальной форме.

Сначала установим следующий результат.

Лемма П3.1. *Если многочлены* $\varphi_1(\lambda), \varphi_2(\lambda), \ldots \varphi_t(\lambda)$ *попарно взаимно просты, то*

$$\begin{pmatrix} \varphi_1(\lambda) & & & \\ & \varphi_2(\lambda) & & \\ & & \ddots & \\ & & & \varphi_t(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \prod\limits_{i=1}^{t} \varphi_i(\lambda) \end{pmatrix}.$$

Д о к а з а т е л ь с т в о. Достаточно рассмотреть случай $t = 2$; общий случай получается из него очевидной индукцией по t. Так как многочлены $\varphi_1(\lambda)$ и $\varphi_2(\lambda)$ взаимно просты, то существуют многочлены $u_1(\lambda)$ и $u_2(\lambda)$ такие, что

$$\varphi_1(\lambda)u_1(\lambda) + \varphi_2(\lambda)u_2(\lambda) = 1.$$

Поэтому

$$\begin{pmatrix} \varphi_1(\lambda) & 0 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda) + \varphi_2(\lambda)u_2(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \varphi_1(\lambda) & 1 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ \varphi_2(\lambda) & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & \varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix}.$$

Рассмотрим жорданову матрицу

$$J = \begin{pmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_s} \end{pmatrix}.$$

Ее характеристическая матрица имеет вид

$$J - \lambda E = \begin{pmatrix} \boxed{J_1 - \lambda E_1} & & & \\ & \boxed{J_2 - \lambda E_2} & & \\ & & \ddots & \\ & & & \boxed{J_s - \lambda E_s} \end{pmatrix}, \tag{П3.1}$$

где E_i — единичная матрица того же порядка, что и клетка J_i.

Пусть жордановы клетки J_i относятся к следующим *различным* числам: $\lambda_1, \lambda_2, \ldots, \lambda_t$ $(t \leq s)$, к числу λ_i относится q_i жордановых клеток ($q_i \geq 1$) и пусть порядки этих клеток, расположенные в невозрастающем порядке, будут

$$k_{i1} \geq k_{i2} \geq \ldots \geq kiq_i. \tag{П3.2}$$

рез клетку $J_i - \lambda E_i$, мы не затронем других диагональных клеток. Следовательно, в (П3.1) можно при помощи элементарных преобразований заменить каждую клетку $J_i - \lambda E_i$, $i = \overline{1, s}$, соответствующей клеткой вида (1.4.4), т.е. вида

$$\begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 1 & \\ & & & & & (\lambda - \lambda_i)^k \end{pmatrix}.$$

Итак, $J - \lambda E$ *эквивалентна диагональной матрице, на диагонали которой стоят, помимо некоторого числа единиц, также следующие многочлены, соответствующие всем жордановым клеткам матрицы* J:

$$\begin{array}{c} (\lambda-\lambda_1)^{k_{11}}, \quad (\lambda-\lambda_1)^{k_{12}}, \quad \ldots, \quad (\lambda-\lambda_1)^{k_{1q_1}}, \\ (\lambda-\lambda_1)^{k_{21}}, \quad (\lambda-\lambda_2)^{k_{22}}, \quad \ldots, \quad (\lambda-\lambda_2)^{k_{2q_2}}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (\lambda-\lambda_t)^{k_{t1}}, \quad (\lambda-\lambda_t)^{k_{t2}}, \quad \ldots, \quad (\lambda-\lambda_t)^{k_{tq_t}}. \end{array} \tag{П3.3}$$

Места, на которых стоят многочлены (П3.3), не указываются, так как в каждой диагональной λ-матрице диагональные элементы можно произвольно переставлять при помощи перестановки строк и одноименных столбцов.

Пусть $q = \max\{q_1, q_2, \ldots, q_t\}$. Обозначим через $e_{n-j+1}(\lambda)$ произведение многочленов, стоящих в j-м столбце таблицы (П3.3)

$$e_{n-j+1}(\lambda) = \prod_{i=1}^{t} (\lambda - \lambda_i)^{k_{ij}}, \quad j = \overline{1, q}. \tag{П3.4}$$

Если для некоторого i окажется $q_i < j$, то в j-м столбце имеются пустые места. Тогда соответствующие пустым местам множители в (П3.4) считаем равными единице. Так как $\lambda_1, \lambda_2, \ldots, \lambda_t$ различны, то степени линейных двучленов, стоящие в j-м столбце таблицы (П3.3), попарно взаимно просты. На основании леммы П3.1 они при помощи элементарных преобразований могут быть заменены в рассматриваемой диагональной матрице их произведениями $e_{n-j+1}(\lambda)$ и некоторым числом единиц. Проделав это для $j = \overline{1, q}$, получим

$$J - \lambda E \sim \begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & 1 & & & & & \\ & & & & e_{n-q+1}(\lambda) & & & & \\ & & & & & \cdot & & & \\ & & & & & & \cdot & & \\ & & & & & & & e_{n-1}(\lambda) & \\ & & & & & & & & e_n(\lambda) \end{pmatrix}. \tag{П3.5}$$

Это — *искомый канонический вид матрицы* $J - \lambda E$. Действительно, старшие коэффициенты всех многочленов, стоящих в (П3.5) на главной диагонали, равны единице и каждый из них делится на предыдущий в силу (П3.2).

стоят из одних и тех же жордановых клеток, т.е. отличаются, быть может, лишь расположением этих клеток вдоль главной диагонали.

Доказательство. *Достаточность.* Пусть жордановы матрицы J и J' обладают одним и тем же набором жордановых клеток. В таблице (П3.3) не отражалось расположение жордановых клеток вдоль главной диагонали и (П3.3) полностью определялось набором этих клеток. Значит, J и J' соответствует одна и та же таблица многочленов (П3.3), а потому одни и те же многочлены (П3.4). Поэтому характеристические матрицы $J-\lambda E$ и $J'-\lambda E$ обладают одинаковыми инвариантными множителями, т.е. эквивалентны. Значит, J и J' подобны.

Необходимость. Пусть J и J' подобны. Тогда $J-\lambda E$ и $J'-\lambda E$ обладают одними и теми же инвариантными множителями. Пусть многочлены (П3.4) для $j=\overline{1,q}$ будут теми из инвариантных множителей, которые отличны от 1. По многочленам (П3.4) однозначно восстанавливается таблица многочленов (П3.3) (она состоит из всех тех максимальных степеней линейных множителей, на которые разлагаются многочлены (П3.4)). Наконец, по таблице (П3.3) восстанавливаются жордановы клетки исходных жордановых матриц: каждому многочлену $(\lambda-\lambda_i)^{k_{ij}}$ из таблицы (П3.3) соответствует жорданова клетка порядка k_{ij}, относящаяся к числу λ_i. Этим доказано, что матрицы J и J' состоят из одних и тех же жордановых клеток и отличаются, быть может, только их расположением.

Следствие 1.4.1. *Жорданова матрица, подобная диагональной матрице, сама диагональна. Две диагональные матрицы подобны тогда и только тогда, когда они получаются друг из друга перестановкой чисел, стоящих на главной диагонали.*

Рассмотрим теперь вопрос о приведении матрицы к жордановой нормальной форме.

Если матрица A с элементами из поля P приводится к жордановой нормальной форме, т.е. подобна жордановой матрице, *то*, как следует из теоремы 1.4.1, *жорданова нормальная форма определяется для матрицы A однозначно с точностью до расположения жордановых клеток на главной диагонали.* Условие для того, чтобы A допускала такое приведение, указывается в следующей теореме, доказательство которой дает одновременно практический способ для разыскания жордановой матрицы, подобной A, если такая жорданова матрица существует.

В дальнейшем **приводимость в поле** P означает, что все элементы трансформирующей матрицы содержатся в поле P.

Основная теорема о приведении матрицы к жордановой нормальной форме.

Матрица A с элементами из поля P тогда и только тогда приводится в поле P к жордановой нормальной форме (т.е. подобна жордановой матрице), когда все собственные значенеия матрицы A лежат в поле P.

Доказательство. *Необходимость.* Пусть A подобна жордановой матрице J. Собственные значения J — корни многочлена $|J-\lambda E|$, равного произведению ее диагональных элементов. Значит, $|J-\lambda E|$ разлагается над полем P на линейные множители. Поэтому собственные значения J лежат в поле P. Так как матрицы A и J подобны, то они обладают одними и теми же собственными значениями. Значит, собственные значения A лежат в поле P.

Достаточность. Пусть собственные значения A лежат в поле P. Если отличные от единицы инвариантные множители матрицы $A-\lambda E$ есть

$$e_{n-q+1}(\lambda),\dots,e_{n-1}(\lambda),e_n(\lambda), \tag{П3.6}$$

то

$$|A-\lambda E|=(-1)^n e_{n-q+1}(\lambda)\dots e_{n-1}(\lambda)e_n(\lambda).$$

Действительно, определители $A-\lambda E$ и ее канонической матрицы могут отличаться только постоянным множителем ($A-\lambda E=U\cdot[$ каноническая матрица$]\cdot V$, откуда $|A-\lambda E|=$

$U(\lambda)$ и $V(\lambda)$ унимодулярны), который равен $(-1)^n$ — старшему коэффициенту $|A-\lambda E|$.

Итак, среди многочленов (П3.6) нет равных нулю, сумма степеней всех этих многочленов равна n и все они разлагаются над полем P на линейные множители (так как $|A-\lambda E|$ разлагается). Пусть

$$e_{n-j+1}(\lambda)=\prod_{i=1}^{t}(\lambda-\lambda_i)^{k_{ij}} \quad (j=\overline{1,q}) \tag{П3.7}$$

есть разложения многочленов (П3.6) в произведение степеней линейных множителей. Назовем **элементарными делителями многочлена** e_{n-j+1} $(j=\overline{1,q})$ отличные от 1 степени различных линейных двучленов, входящие в ее разложение (П3.7), т.е.

$$(\lambda-\lambda_1)^{k_{1j}}(\lambda-\lambda_2)^{k_{2j}}\dots(\lambda-\lambda_t)^{k_{tj}}.$$

В силу замкнутости поля комплексных чисел отсюда немедленно вытекает

Следствие 1.4.2. *Всякая матрица с комплексными элементами приводится в поле комплексных чисел $\mathbb{C}$ к жордановой нормальной форме.*

Теорема 1.4.2 (необходимое и достаточное условие приводимости матрицы к диагональному виду). *Матрица A порядка n с элементами из поля P приводится к диагональному виду тогда и только тогда, когда все корни последнего инвариантного множителя $e_n(\lambda)$ ее характеристической матрицы лежат в поле , причем среди этих корней нет кратных.*

Доказательство. Приводимость матрицы к диагональному виду равносильна приводимости к такому жорданову виду, все жордановы клетки которого имеют порядок 1. Другими словами, все элементарные делители матрицы A должны быть многочленами первой степени. Так как все инвариантные множители матрицы $A-\lambda E$ являются делителями многочлена $e_n(\lambda)$, то последнее условие равносильно тому, что все элементарные делители многочлена $e_n(\lambda)$ имеют степень 1.

Из доказательства основной теоремы о приведении матрицы к жордановой нормальной форме вытекает следующий

Алгоритм приведения матрицы к жордановой нормальной форме.

1. Находим собственные значения матрицы A и проверяем, что они принадлежат полю P. Если это не так, то делаем вывод о том, что матрица A не может быть приведена в поле P к жордановой нормальной форме. Если собственные значения матрицы A принадлежат полю P, то переходим к пункту 2.

2. Обозначим через $\lambda_1,\lambda_2,\dots,\lambda_t$ — все различные собственные значения матрицы A. Находим инвариантные множители характеристической матрицы $A-\lambda E$:

$$1,1,\dots,1,e_{n-q+1}(\lambda),\dots,e_{n-1}(\lambda),e_n(\lambda),$$

где $e_{n-q+1}(\lambda)\neq 1,\dots,e_n(\lambda)\neq 1$. Среди этих многочленов нет равных нулю и все они разлагаются на линейные множители. Сумма степеней всех многочленов равна n. Разложим эти многочлены на линейные множители:

$$e_{n-j+1}(\lambda)=\prod_{i=1}^{t}(\lambda-\lambda_i)^{k_{ij}}=(\lambda-\lambda_1)^{k_{1j}}(\lambda-\lambda_2)^{k_{2j}}\dots(\lambda-\lambda_t)^{k_{tj}}, \qquad j=\overline{1,q}.$$

Элементарными делителями многочленов $e_{n-j+1}(\lambda)$, $j=\overline{1,q}$, называются отличные от 1 степени различных линейных двучленов, входящих в последнее разложение. **Элементарными делителями матрицы** A называются элементарные делители всех многочленов $e_{n-q+1}(\lambda),\dots,e_n(\lambda)$.

$$(\lambda-\lambda_1)^{k_{11}},\ (\lambda-\lambda_1)^{k_{12}},\ \ldots,\ (\lambda-\lambda_1)^{k_{1q_1}},$$

$$(\lambda-\lambda_1)^{k_{21}},\ (\lambda-\lambda_1)^{k_{22}},\ \ldots,\ (\lambda-\lambda_1)^{k_{2q_2}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(\lambda-\lambda_1)^{k_{t1}},\ (\lambda-\lambda_1)^{k_{t2}},\ \ldots,\ (\lambda-\lambda_1)^{k_{tq_t}}.$$

Составляется жорданова матрица J порядка n, состоящая из жордановых клеток, по следующему правилу:

каждому элементарному делителю $(\lambda-\lambda_i)^{k_{ij}}$ *сопоставляется жорданова клетка порядка* k_{ij}*, относящаяся к числу* λ_i*. Полученная матрица* J *и есть искомая матрица, т.е. матрица* A *подобна матрице* J*.*