

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ОБУЧАЮЩИХСЯ
ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Б1.В.ДВ.05.02 Теория матриц

Направление подготовки 20.03.01 Техносферная безопасность

Профиль подготовки «Безопасность жизнедеятельности в техносфере»

Квалификация выпускника бакалавр

Форма обучения заочная

СОДЕРЖАНИЕ

1	Конспект лекций.....	3
1.1	Алгоритм Гаусса и некоторые его применения . Матрица и действия над ними	3
1.2	Уравнения различных типов	27
1.3	Извлечение корня m -ной степени из матрицы	32
1.4	Скалярное уравнение. Логарифм матрицы	35
1.5	Сингулярные пучки матриц	39
1.6	Приложение теории матриц к исследованию систем линейных дифференциальных уравнений Добавление неравенства для собственных и сингулярных чисел	49
2	Методические указания по выполнению практических занятий.....	63
2.1	Арифметические действия над матрицами. Применение метода Гаусса (интерактивная форма)	63
2.2	Решение матричных уравнений	70
2.3	Извлечение корня m -ной степени	76
2.4	Решение скалярного уравнения	81
2.5	Нахождение логарифма матрицы (интерактивная форма)	60
2.6	Решение дифференциальных уравнений. Неравенства Неймана-Хорна	84

1. КОНСПЕКТ ЛЕКЦИЙ

1.1 Лекция № 1 (1 час).

Тема: «Матрица и действия над ними»

1.1.1 Вопросы лекции:

1. Матрицы и основные обозначения.
2. Сложение и умножение прямоугольных матриц.
3. Квадратные матрицы.

1.1.2 Краткое содержание вопросов:

1. Матрицы и основные обозначения.

1. Пусть дано некоторое числовое поле K .

Определение 1. Прямоугольную таблицу чисел из поля K

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad (1)$$

будем называть матрицей. Если $m = n$, то матрица называется квадратной, а число m , равное n , — ее порядком. В общем же случае матрица называется прямоугольной (с размерами $m \times n$) или $m \times n$ -матрицей. Числа, составляющие матрицу, называются ее элементами.

Обозначения. При двухиндексном обозначении элементов первый индекс всегда указывает номер строки, а второй индекс — номер столбца, на пересечении которых стоит данный элемент.

Наряду с обозначениями матрицы (1) будем употреблять и сокращенное обозначение:

$$\|a_{ik}\| (i = 1, 2, \dots, m; k = 1, 2, \dots, n).$$

Часто матрицу (1) будем обозначать также одной буквой, например матрица A .

Если A — квадратная матрица порядка n , то будем

писать: $A = \|a_{ik}\|_1^n$ Определитель квадратной матрицы $A = \|a_{ik}\|_1^n$ будем обозначать

так: $|a_{ik}|_1^n$ или $|A|$.

Введем сокращенные обозначения для определителей, составленных из элементов данной матрицы:

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \dots & a_{i_2 k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_p k_1} & a_{i_p k_2} & \dots & a_{i_p k_p} \end{vmatrix} \quad (2)$$

Определитель (3) называется минором P -го порядка матрицы A ,

если $1 \leq i_1 < i_2 < \dots < i_p \leq m$ и $1 \leq k_1 < k_2 < \dots < k_p \leq n$.

$m \times n$ -матрица $A = \|a_{ik}\|$ имеет $C_m^p C_n^p$ миноров P -го порядка

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \left(\begin{matrix} 1 \leq i_1 < i_2 < \dots < i_p \leq m \\ 1 \leq k_1 < k_2 < \dots < k_p \leq n \end{matrix} ; p \leq m, n \right). \quad (2')$$

Миноры (2'), у которых $i_1 = k_1, i_2 = k_2, \dots, i_p = k_p$, называются главными. В обозначениях

(2) определитель квадратной матрицы $A = \|a_{ik}\|_1^n$ запишется так:

$$|A| = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Наибольший из порядков отличных от нуля миноров, порождаемых матрицей, называется рангом матрицы. Если r — ранг прямоугольной матрицы A с размерами $m \times n$, то, очевидно, $r \leq m, n$.

Прямоугольную матрицу, состоящую из одного столбца

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

мы будем называть столбцовой и обозначать так: $(x_1, x_2, \dots, x_n)$.

Прямоугольную матрицу, состоящую из одной строки

$$\|z_1, z_2, \dots, z_n\|,$$

мы будем называть строчной и обозначать так: $[z_1, z_2, \dots, z_n]$.

Квадратную матрицу, у которой все элементы, расположенные вне главной диагонали, равны нулю,

$$\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

мы будем называть диагональной и обозначать так: $\|d_i \delta_{ik}\|_1^n$ или

$$\{d_1, d_2, \dots, d_n\}.$$

Введем еще специальные обозначения для строк и столбцов $m \times n$ -матрицы $A = \|a_{ik}\|$.

Будем обозначать i -ю строку матрицы A через $a_{i\cdot}$, а j -й столбец — через $a_{\cdot j}$:

$$a_{i\cdot} = [a_{i1}, a_{i2}, \dots, a_{in}],$$

$$a_{\cdot j} = [a_{1j}, a_{2j}, \dots, a_{mj}],$$

$$(i = 1, \dots, m; j = 1, \dots, n).$$

(3)

Пусть m величин $y_1, y_2, \dots, y_m$ выражаются линейно и однородно через n других величин $x_1, x_2, \dots, x_n$:

$$\left. \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\dots \dots \dots \dots \dots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned} \right\} \quad (4)$$

или, в сокращенной записи

$$y_i = \sum_{k=1}^n a_{ik}x_k \quad (i = 1, 2, \dots, m) \quad (4')$$

Преобразование величин $x_1, x_2, \dots, x_n$ в величины $y_1, y_2, \dots, y_m$ при помощи формул (4) называется линейным преобразованием. Коэффициенты этого преобразования образуют $m \times n$ -матрицу (1). Задание линейного преобразования (4) однозначно определяет матрицу (1) и наоборот.

В следующем параграфе, исходя из свойств линейных преобразований (4), мы определим основные операции над прямоугольными матрицами.

2. Сложение и умножение прямоугольных матриц.

Определим основные операции над матрицами: сложение матриц, умножение матрицы на число и умножение матриц.

1. Пусть величины $y_1, y_2, \dots, y_m$ выражаются через величины $x_1, x_2, \dots, x_n$ при помощи линейного преобразования

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i = 1, 2, \dots, m), \quad (5)$$

а величины $z_1, z_2, \dots, z_m$ — через те же величины $x_1, x_2, \dots, x_n$ при помощи преобразования

$$z_i = \sum_{k=1}^n b_{ik} x_k \quad (i = 1, 2, \dots, m). \quad (6)$$

Тогда

$$y_i + z_i = \sum_{k=1}^n (a_{ik} + b_{ik}) x_k \quad (i = 1, 2, \dots, m). \quad (7)$$

В соответствии с этим мы устанавливаем

Определение 2. Суммой двух прямоугольных матриц $A = \|a_{ik}\|$ и $B = \|b_{ik}\|$ одинаковых размеров $m \times n$ называется матрица $C = \|c_{ik}\|$ тех же размеров, элементы которой равны суммам соответствующих элементов данной матрицы:

$$C = A + B,$$

если

$$c_{ik} = a_{ik} + b_{ik} \quad (i = 1, 2, \dots, m, k = 1, 2, \dots, n).$$

Операция нахождения суммы данных матриц называется сложением матриц.

Пример

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} + \begin{vmatrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix} = \begin{vmatrix} a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ b_1 + d_1 & b_2 + d_2 & b_3 + d_3 \end{vmatrix}.$$

Согласно определению 2, складывать можно только прямоугольные матрицы одинаковых размеров.

В силу этого же определения матрица коэффициентов в преобразовании (7) есть сумма матриц коэффициентов в преобразованиях (5) и (6).

Из определения сложения матриц непосредственно следует, что эта операция обладает переместительным и сочетательным свойствами:

$$1^\circ \quad A + B = B + A,$$

$$2^\circ \quad (A + B) + C = A + (B + C).$$

Здесь A, B, C — произвольные прямоугольные матрицы одинаковых размеров.

Операция сложения матриц естественным образом распространяется на случай любого числа слагаемых.

2. Умножим в преобразовании (5) величины $y_1, y_2, \dots, y_m$ на некоторое число α из K . Тогда

$$\alpha y_i = \sum_{k=1}^n (\alpha a_{ik}) x_k \quad (i = 1, 2, \dots, m).$$

В соответствии с этим имеет место

Определение 3. Произведением матрицы $A = \|a_{ik}\|$ ($i = 1, 2, \dots, m, k = 1, 2, \dots, n$) на

число α из K называется матрица $C = \|c_{ik}\|$ ($i = 1, 2, \dots, m, k = 1, 2, \dots, n$), элементы которой получаются из соответствующих элементов матрицы A умножением на число α :

$$C = \alpha A,$$

если

$$c_{ik} = \alpha a_{ik} \quad (i = 1, 2, \dots, m, k = 1, 2, \dots, n).$$

Операция нахождения произведения матрицы на число называется умножением матрицы на число.

Пример

$$\alpha \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} + \begin{vmatrix} \alpha a_1 & \alpha a_2 & \alpha a_3 \\ \alpha b_1 & \alpha b_2 & \alpha b_3 \end{vmatrix}.$$

Легко видеть, что

- 1° $\alpha(A+B) = \alpha A + \alpha B$,
- 2° $(\alpha + \beta)A = \alpha A + \beta A$,
- 3° $(\alpha\beta)A = \alpha(\beta A)$.

Здесь A, B — прямоугольные матрицы одинаковых размеров, α, β — числа из поля K . Разность $A - B$ двух прямоугольных матриц одинаковых размеров определяется равенством

$$A - B = A + (-1)B.$$

Если A — квадратная матрица порядка n , а α — число из K , то

$$|\alpha A| = \alpha^n |A|.$$

3. Пусть величины $z_1, z_2, \dots, z_m$ выражаются через величины $y_1, y_2, \dots, y_m$ при помощи преобразования

$$z_i = \sum_{k=1}^n a_{ik} y_k \quad (i = 1, 2, \dots, m) \quad (8)$$

а величины $y_1, y_2, \dots, y_m$ — через те же величины $x_1, x_2, \dots, x_q$ при помощи формул

$$y_k = \sum_{j=1}^q b_{kj} x_j \quad (k = 1, 2, \dots, m). \quad (9)$$

Тогда, подставляя эти выражения для y_k ($k = 1, 2, \dots, m$) в формулы (8), мы

выразим $z_1, z_2, \dots, z_m$ через $x_1, x_2, \dots, x_q$ при помощи «составного» преобразования:

$$z_i = \sum_{k=1}^n a_{ik} \sum_{j=1}^q b_{kj} x_j = \sum_{j=1}^q \left(\sum_{k=1}^n a_{ik} b_{kj} \right) x_j \quad (i = 1, 2, \dots, m). \quad (10)$$

В соответствии с этим имеет место

Определение 4. Произведением двух прямоугольных матриц

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nq} \end{pmatrix}$$

называется матрица

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1q} \\ c_{21} & c_{22} & \dots & c_{2q} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mq} \end{pmatrix},$$

у которой элемент c_{ij} , стоящий на пересечении i -й строки и j -го столбца, равен «произведению» i -й строки первой матрицы A на j -й столбец второй матрицы B :

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, q). \quad (11)$$

Операция нахождения произведения данных матриц называется умножением матриц.

Пример

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & e_1 & f_1 \\ c_2 & d_2 & e_2 & f_2 \\ c_3 & d_3 & e_3 & f_3 \end{pmatrix} = \begin{pmatrix} a_1 c_1 + a_2 c_2 + a_3 c_3, & a_1 d_1 + a_2 d_2 + a_3 d_3, & a_1 e_1 + a_2 e_2 + a_3 e_3, & a_1 f_1 + a_2 f_2 + a_3 f_3 \\ b_1 c_1 + b_2 c_2 + b_3 c_3, & b_1 d_1 + b_2 d_2 + b_3 d_3, & b_1 e_1 + b_2 e_2 + b_3 e_3, & b_1 f_1 + b_2 f_2 + b_3 f_3 \end{pmatrix}$$

По определению 4 матрица коэффициентов в преобразовании (10) равна произведению матрицы коэффициентов в (8) на матрицу коэффициентов (9).

Заметим, что операция умножения двух прямоугольных матриц выполнима лишь в том случае, когда число столбцов в первом сомножителе равно числу строк во втором. В частности, умножение всегда выполнимо, если оба сомножителя — квадратные матрицы одного и того же порядка. Обратим внимание читателя и на то, что даже в этом частном случае умножение матриц не обладает переместительным свойством. Так, например,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 8 & -2 \\ 18 & -4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix}.$$

Если $AB = BA$, то матрицы A и B называются перестановочными или коммутирующими между собой.

Пример. Матрицы

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} -3 & 2 \\ -2 & -4 \end{pmatrix}$$

перестановочны между собой, так как

$$AB = \begin{pmatrix} -7 & -6 \\ 6 & -4 \end{pmatrix}, \quad BA = \begin{pmatrix} -7 & -6 \\ 6 & -4 \end{pmatrix}.$$

Легко проверяется сочетательное свойство умножения матриц, а также распределительное свойство умножения относительно сложения:

$$\begin{aligned} 1^\circ & \quad (AB)C = A(BC), \\ 2^\circ & \quad (A+B)C = AC + BC, \\ 3^\circ & \quad A(B+C) = AB + AC. \end{aligned} \quad (12)$$

4. Если воспользоваться произведением прямоугольных матриц, то линейное преобразование

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 \dots + a_{2n}x_n \\ &\dots \qquad \qquad \qquad \dots \qquad \qquad \dots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 \dots a_{mn}x_n \end{aligned} \quad (13)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$
$$y = Ax, \quad (13')$$

Равенства (13) выражают собой тот факт, что столбец $\mathcal{Y}$ является линейной комбинацией столбцов матрицы A с коэффициентами $x_1, x_2, \dots, x_n$:

$$y = x_1 a_{.1} + x_2 a_{.2} + \dots + x_n a_{.n} = \sum_{k=1}^n x_k a_{.k}. \quad (13'')$$

$$C = AB \quad (14)$$
$$c_{.j} = \sum_{k=1}^n b_{kj} a_{.k} \quad (j=1, 2, \dots, q) \quad (14')$$
$$c_i = \sum_{k=1}^n a_{ik} b_k \quad (i = 1, 2, \dots, m). \quad (14'')$$

Аналогично, любая i -я строка в матрице C является линейной комбинацией строк матрицы B , а коэффициентами этой линейной зависимости являются элементы i -й строки матрицы A .

Остановимся еще на том частном случае, когда в произведении $C = AB$ второй сомножитель является квадратной и притом диагональной матрицей $B = \{a_1, a_2, \dots, a_n\}$. Тогда из формул (11) следует:

$$c_{ij} = a_{ij} d_j \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, n).$$

8

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} \begin{vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{vmatrix} = \begin{vmatrix} a_{11}d_1 & a_{12}d_2 & \dots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \dots & a_{2n}d_n \\ \dots & \dots & \dots & \dots \\ a_{m1}d_1 & a_{m2}d_2 & \dots & a_{mn}d_n \end{vmatrix}.$$

Аналогично

$$\begin{vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_m \end{vmatrix} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} = \begin{vmatrix} d_1a_{11} & d_1a_{12} & \dots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \dots & d_2a_{2n} \\ \dots & \dots & \dots & \dots \\ d_ma_{m1} & d_ma_{m2} & \dots & d_ma_{mn} \end{vmatrix}.$$

Таким образом, при умножении прямоугольной матрицы A справа (слева) на диагональную матрицу $\{d_1, d_2, \dots\}$ все столбцы (соответственно строки) матрицы A помножаются на числа $d_1, d_2, \dots$

5. Пусть квадратная матрица $C = \|c_{ij}\|$ является произведением двух прямоугольных матриц $A = \|a_{ik}\|$ и $B = \|b_{kj}\|$ соответственно размеров $m \times n$ и $n \times m$:

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} \begin{vmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nm} \end{vmatrix}, \quad (15)$$

т. е.

$$c_{ij} = \sum_{\alpha=1}^n a_{i\alpha} b_{\alpha j} \quad (i, j = 1, 2, \dots, m). \quad (15')$$

Установим важную формулу Бине-Коши,

выражающую определитель $|C|$ через миноры матриц A и B :

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \begin{vmatrix} b_{k_1 1} & \dots & b_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & b_{k_m m} \end{vmatrix} \quad (16)$$

или в специальных обозначениях:

$$C \begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \quad (16')$$

Согласно этой формуле определитель матрицы C равен сумме произведений всевозможных миноров максимального (m -го) порядка матрицы A на соответствующие миноры того же порядка матрицы B .

Вывод формулы Бине-Коши. На основании формулы

(15') определитель матрицы C можно представить в виде

$$\begin{aligned}
\left\| \begin{matrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{matrix} \right\| &= \left\| \begin{matrix} \sum_{\alpha_1=1}^n a_{1\alpha_1} b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^n a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum_{\alpha_1=1}^n a_{m\alpha_1} b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^n a_{m\alpha_m} b_{\alpha_m m} \end{matrix} \right\| = \\
&= \sum_{\alpha_1, \dots, \alpha_m=1}^n \left\| \begin{matrix} a_{1\alpha_1} b_{\alpha_1 1} & \dots & a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1} b_{\alpha_1 1} & \dots & a_{m\alpha_m} b_{\alpha_m m} \end{matrix} \right\| = \sum_{\alpha_1, \dots, \alpha_m=1}^n A \begin{pmatrix} 1 & 2 \dots m \\ \alpha_1 & \alpha_2 \dots \alpha_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m}
\end{aligned} \quad (16'')$$

Если $m > n$, то среди чисел $\alpha_1, \alpha_2, \dots, \alpha_m$ всегда найдутся равные между собой числа и, следовательно, каждое слагаемое в правой части равенства (16'') будет равно нулю.

Значит, в этом случае $|C| = 0$.

Пусть теперь $m \leq n$. Тогда в сумме, стоящей в правой части равенства (16''), будут равны нулю те слагаемые, у которых хотя бы два из индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ равны между собой. Все же остальные слагаемые этой суммы можно разбить на группы по $m!$ слагаемых в каждой, объединяя в одну группу те слагаемые, которые отличаются друг от друга только порядком индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ (индексы $\alpha_1, \alpha_2, \dots, \alpha_m$ в пределах каждой группы слагаемых имеют одну и ту же совокупность значений). Тогда в пределах одной такой группы сумма соответствующих слагаемых будет равна

$$\begin{aligned}
&\sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} = \\
&= A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} = \sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} = \\
&= A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}.
\end{aligned}$$

Поэтому из (16'') получаем (16').

Пример 1.

$$\left\| \begin{matrix} a_1 c_1 + a_2 c_2 + \dots + a_n c_n & a_1 d_1 + a_2 d_2 + \dots + a_n d_n \\ b_1 c_1 + b_2 c_2 + \dots + b_n c_n & b_1 d_1 + b_2 d_2 + \dots + b_n d_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 & d_1 \\ c_2 & d_2 \\ \dots & \dots \\ c_n & d_n \end{matrix} \right\|.$$

Поэтому формула (16) дает так называемое тождество Коши

$$\left\| \begin{matrix} a_1 c_1 + a_2 c_2 + \dots + a_n c_n & a_1 d_1 + a_2 d_2 + \dots + a_n d_n \\ b_1 c_1 + b_2 c_2 + \dots + b_n c_n & b_1 d_1 + b_2 d_2 + \dots + b_n d_n \end{matrix} \right\| = \sum_{1 \leq i < k \leq n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix} \begin{vmatrix} c_i & d_i \\ c_k & d_k \end{vmatrix}$$

Полагая в этом тождестве $a_i = c_i$ и $b_i = d_i$ ($i = 1, 2, \dots, n$), получим:

$$\left| \begin{matrix} a_1^2 + a_2^2 + \dots + a_n^2 & a_1 b_1 + a_2 b_2 + \dots + a_n b_n \\ a_1 b_1 + a_2 b_2 + \dots + a_n b_n & b_1^2 + b_2^2 + \dots + b_n^2 \end{matrix} \right| = \sum_{1 \leq i < k \leq n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix}^2.$$

В случае когда a_i и b_i ($i = 1, 2, \dots, n$) - вещественные числа, отсюда следует известное неравенство:

$$(a_1 b_1 + a_2 b_2 + \dots + a_n b_n)^2 \leq (a_1^2 + a_2^2 + \dots + a_n^2)(b_1^2 + b_2^2 + \dots + b_n^2).$$

a_i При этом знак равенства имеет место тогда и только тогда, когда все числа пропорциональны соответствующим b_i ($i = 1, 2, \dots, n$).

Пример 2.

$$\begin{vmatrix} a_1c_1 + b_1d_1 & \dots & a_1c_n + b_1d_n \\ \dots & \dots & \dots \\ a_nc_1 + b_nd_1 & \dots & a_nc_n + b_nd_n \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ \dots & \dots \\ a_n & b_n \end{vmatrix} \begin{vmatrix} c_1 \dots c_n \\ d_1 \dots d_n \end{vmatrix}.$$

Поэтому при $n > 2$

$$\begin{vmatrix} a_1c_1 + b_1d_1 & \dots & a_1c_n + b_1d_n \\ \dots & \dots & \dots \\ a_nc_1 + b_nd_1 & \dots & a_nc_n + b_nd_n \end{vmatrix} = 0.$$

Рассмотрим частный случай, когда A и B — квадратные матрицы одного и того же порядка n , и положим в (16') $m = n$. Тогда приходим к известной теореме об умножении определителей:

$$C \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} B \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix},$$

или в других обозначениях:

$$|C| = |AB| = |A| \cdot |B| \quad (17)$$

Таким образом, определитель произведения двух квадратных матриц равен произведению определителей перемножаемых матриц.

6. Формула Бине—Коши дает возможность в самом общем случае выразить миноры произведения двух прямоугольных матриц через миноры сомножителей. Пусть

$$A = \|a_{ik}\|, \quad B = \|b_{kj}\|, \quad C = \|c_{ij}\|$$

$$(i = 1, 2, \dots, m, \quad k = 1, 2, \dots, n, \quad j = 1, 2, \dots, q)$$

и

$$C = AB.$$

Рассмотрим произвольный минор матрицы C :

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \begin{pmatrix} 1 \leq i_1 < i_2 < \dots < i_p \leq m \\ 1 \leq j_1 < j_2 < \dots < j_p \leq q \end{pmatrix}; p \leq m, q.$$

Матрица, составленная из элементов этого минора, представляет собой произведение двух прямоугольных матриц

$$\begin{vmatrix} a_{i_1 1} a_{i_2 2} \dots a_{i_p p} \\ \dots \dots \dots \\ a_{i_p 1} a_{i_p 2} \dots a_{i_p p} \end{vmatrix}, \quad \begin{vmatrix} b_{1 j_1} & \dots & b_{1 j_p} \\ b_{2 j_1} & \dots & b_{2 j_p} \\ \dots & \dots & \dots \\ b_{n j_1} & \dots & b_{n j_p} \end{vmatrix}.$$

Поэтому, применяя формулу Бине—Коши, получаем:

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_p \leq n} A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \quad (18)$$

При $p = 1$ формула (18) переходит в формулу (11). При $p > 1$ формула (18) является естественным обобщением формулы (11).

Отметим еще одно следствие из формулы (18).

Ранг произведения двух прямоугольных матриц не превосходит ранга любого из сомножителей.

Если $C = AB$ и r_A, r_B, r_C — ранги матриц A, B, C , то

$$r_C \leq r_A, r_B.$$

7. Если X — решение матричного уравнения $AX = C$ (размеры матриц A, X и C соответственно $n \times m, n \times q$ и $m \times q$), то $r_X \geq r_C$. Покажем, что среди решений матричного уравнения $AX = C$ существует решение X_0 минимального ранга, для которого $r_{X_0} \geq r_C$.

Действительно, пусть $r = r_C$. Тогда среди столбцов матрицы C имеется r линейно независимых. Пусть для конкретности первые r столбцов $C_1, \dots, C_r$ линейно независимы, а остальные столбцы $C_{r+1}, \dots, C_q$ являются линейными комбинациями первых r :

$$C_{\cdot j} = \sum_{k=1}^r \alpha_{jk} C_{\cdot k} \quad (j = r+1, \dots, q).$$

Пусть X — произвольное решение уравнения $AX = C$. Тогда (см. стр. 19)

$$AX_{\cdot k} = C_{\cdot k} \quad (k = 1, \dots, r) \quad (20)$$

Определим столбцы $\tilde{X}_{\cdot r+1}, \dots, \tilde{X}_{\cdot q}$ равенствами

$$X_{\cdot j} = \sum_{k=1}^r \alpha_{jk} X_{\cdot k} \quad (j = r+1, \dots, q).$$

Умножая эти равенства слева почленно на A , в силу равенств (19) и (20) находим:

$$A\tilde{X}_{\cdot j} = C_{\cdot j} \quad (j = r+1, \dots, q). \quad (20')$$

Система из q равенств (20) и (20') эквивалентна одному матричному равенству

$$AX_0 = C,$$

где $X_0 = (X_{\cdot 1}, \dots, X_{\cdot r}, \tilde{X}_{\cdot r+1}, \dots, \tilde{X}_{\cdot q})$ — матрица ранга r .

Решение X_0 минимального ранга r_C матричного уравнения $AX = C$ всегда представимо в виде

$$X_0 = VC,$$

где V — некоторая $m \times n$ -матрица.

Действительно, из равенства $AX_0 = C$ следует, что строки матрицы C являются линейными комбинациями строк матрицы X_0 . Поскольку как среди строк матрицы C , так и среди строк матрицы X_0 имеется одно и то же число r_C линейно независимых, то и, наоборот, строки матрицы X_0 являются линейными комбинациями строк матрицы C , а отсюда уже следует равенство $X_0 = VC$.

Докажем теперь следующее предложение.

Матричное уравнение

$$AXB = C,$$

где A, B — заданные, а X — искомая прямоугольная матрица, имеет решение в том и только том случае, когда одновременно имеют решения матричные уравнения

$$AY = C, \quad ZB = C, \quad (22)$$

т. е. когда столбцы матрицы C являются линейными комбинациями столбцов матрицы A , а строки матрицы C являются линейными комбинациями строк матрицы B . В самом деле, если матрица X — решение уравнения (21), то матрицы $Y = XB$ и $Z = AX$ являются решениями уравнений (22). Обратно, пусть существуют решения Y, Z уравнений (22). Тогда первое из этих уравнений имеет решение Y_0 минимального ранга r_C , которое по доказанному представимо в виде

$$Y_0 = VC.$$

Поэтому

$$C = AY_0 = AVC = AVZB.$$

Тогда матрица $X = VZ$ будет решением уравнения (21).

3. Квадратные матрицы.

Квадратную матрицу n -го порядка, у которой на главной диагонали стоят единицы, а все остальные элементы равны нулю, будем называть единичной матрицей и обозначать через $E^{(n)}$ или просто E . Название «единичная матрица» связано со следующим свойством матрицы E : для любой прямоугольной матрицы

$$A = \|a_{ik}\| \quad (i = 1, 2, \dots, m, k = 1, 2, \dots, n)$$

имеют место равенства

$$E^{(n)}A = AE^{(n)} = A.$$

Очевидно,

$$E^{(n)} = \|\delta_{ik}\|_n$$

Пусть $A = \|a_{ik}\|_n$ — квадратная матрица. Тогда степень матрицы определяется обычным образом:

$$A^p = \underbrace{AA \dots A}_p \quad (p = 1, 2, \dots); \quad A^0 = E.$$

Из сочетательного свойства умножения матриц следует:

$$A^q = A^{p+q}.$$

Здесь p, q — произвольные целые неотрицательные числа.

Рассмотрим многочлен (целую рациональную функцию) с коэффициентами из поля K :

$$f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + \dots + \alpha_m.$$

Тогда под $f(A)$ будем понимать матрицу

$$f(A) = \alpha_0 A^m + \alpha_1 A^{m-1} + \dots + \alpha_m E.$$

Так определяется многочлен от матрицы.

Пусть многочлен $f(t)$ равен произведению многочленов $h(t)$ и $g(t)$:

$$f(t) = h(t)g(t).$$

Многочлен $f(t)$ получается из $h(t)$ и $g(t)$ путем почленного перемножения и приведения подобных членов. При этом используется правило перемножения степеней: $t^p t^q = t^{p+q}$. Так как все эти действия правомерны и при замене скалярной величины t на матрицу A , то

$$f(A) = h(A)g(A)$$

Отсюда, в частности,

$$h(A)g(A) = g(A)h(A),$$

т. е. два многочлена от одной и той же матрицы всегда перестановочны между собой.

Примеры.

Условимся p -й наддиагональю (поддиагональю) в прямоугольной

матрице $A = \|a_{ik}\|$ называть ряд элементов a_{ik} , у

которых $k-i = p$ (соответственно $i-k = p$). Обозначим через H квадратную матрицу n -го порядка, у которой элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Тогда

$$H = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad H^2 = \begin{pmatrix} 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \text{ и т. д.};$$

$$H^p = 0 \quad (p \geq n)$$

В силу этих равенств если:

$$f(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_{n-1} t^{n-1} + \dots$$

— многочлен относительно t , то

$$f(H) = a_0 E + a_1 H + a_2 H^2 + \dots = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & a_0 & a_1 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & a_2 \\ \vdots & \vdots & \vdots & \ddots & a_1 \\ 0 & 0 & 0 & \dots & a_0 \end{pmatrix}.$$

Аналогично, если F — квадратная матрица n -го порядка, у которой все элементы первой поддиагонали равны единице, а все остальные, нулю, то

$$f(F) = a_0 E + a_1 F + a_2 F^2 = \begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_2 \\ \vdots & \vdots & \ddots & a_1 \\ a_{n-1} & \dots & a_1 & a_0 \end{pmatrix}.$$

Предлагаем читателю проверить следующие свойства матриц H и F :

1° В результате умножения произвольной $m \times n$ -матрицы A слева на матрицу H (матрицу F) m -го порядка все строки матрицы подминаются (опускаются) на одно место вверх (вниз), первая (последняя) строка матрицы A исчезает, а последняя (первая) строка произведения заполняется нулями. Так, например,

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \text{ и } |A| \neq 0,$$

то

$$A^{-1} = \frac{1}{|A|} = \begin{vmatrix} b_2c_3 - b_3c_2 & a_3c_2 - a_2c_3 & a_2b_3 - a_3b_2 \\ b_3c_1 - b_1c_3 & a_1c_3 - a_3c_1 & a_3b_1 - a_1b_3 \\ b_1c_2 - b_2c_1 & a_2c_1 - a_1c_2 & a_1b_2 - a_2b_1 \end{vmatrix}.$$

Образуя составное преобразование из данного преобразования (23) и обратного (24) в одном и в другом порядке, мы в обоих случаях получаем тождественное преобразование (с единичной матрицей коэффициентов); поэтому

$$AA^{-1} = A^{-1}A = E. \quad (26)$$

В справедливости равенств (26) можно убедиться и непосредственным перемножением матриц A и A^{-1} . Действительно, в силу (25)

$$\left[AA^{-1} \right]_{ij} = \sum_{k=1}^n a_{ik} a_{kj}^{(-1)} = \frac{1}{|A|} \sum_{k=1}^n a_{ik} A_{jk} = \delta_{ij} \quad (i, j = 1, 2, \dots, n).$$

Аналогично

$$\left[A^{-1}A \right]_{ij} = \sum_{k=1}^n a_{ij}^{(-1)} a_{kj} = \frac{1}{|A|} \sum_{k=1}^n a_{kj} A_{ki} = \delta_{ij} \quad (i, j = 1, 2, \dots, n).$$

Нетрудно видеть, что матричные уравнения

$$AX = E \text{ и } XA = E \quad (|A| \neq 0) \quad (27)$$

никаких других решений, кроме решения $X = A^{-1}$ не имеют. Действительно, умножая обе части первого уравнения слева, а второго — справа на A^{-1} и используя сочетательное свойство произведения матриц, а также равенства (26), мы в обоих случаях получим:

$$X = A^{-1}.$$

Этим же способом доказывается, что каждое из матричных уравнений

$$AX = B, \quad XA = B \quad (|A| \neq 0), \quad (28)$$

где X и B — прямоугольные матрицы равных размеров, A — квадратная матрица соответствующего размера, имеет одно и только одно решение:

$$X = A^{-1}B \text{ и соответственно } X = BA^{-1} \quad (29)$$

Матрицы (29) являются как бы «левым» и «правым» частными от «деления» матрицы B на матрицу A . Из (28) и (29) следует соответственно (см. стр.

22) $r_B \leq r_X$ и $r_X \leq r_B$, т. е. $r_X = r_B$. Сопоставляя с (28), имеем:

При умножении прямоугольной матрицы слева или справа на неособенную матрицу ранг исходной матрицы не изменяется.

Заметим еще, что из (26) вытекает $|A| \cdot |A^{-1}| = 1$, т. е.

$$|A^{-1}| = \frac{1}{|A|}.$$

Для произведения двух неособенных матриц имеем:

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (30)$$

3. Все матрицы n -го порядка образуют кольцо с единичным элементом E . Поскольку в этом кольце определена операция умножения на число из поля K и

существует базис из n^2 линейно независимых матриц, через которые линейно выражаются все матрицы n -го порядка, то кольцо матриц n -го порядка является алгеброй.

Все квадратные матрицы n -го порядка образуют коммутативную группу относительно операции сложения. Все неособенные матрицы n -го порядка образуют (некоммутативную) группу относительно операции умножения.

Квадратная матрица $A = \|a_{ik}\|$ называется верхней треугольной (нижней треугольной), если равны нулю все элементы матрицы, расположенные под главной диагональю (над главной диагональю):

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

(1)
(2)

Диагональная матрица является частным случаем как верхней, так и нижней треугольной матрицы.

Так как определитель треугольной матрицы равен произведению ее диагональных элементов, то треугольная (и, в частности, диагональная) матрица является неособенной только тогда, когда все ее диагональные элементы отличны от нуля.

Легко проверить, что сумма и произведение двух диагональных (верхних треугольных, нижних треугольных) матриц есть диагональная (соответственно верхняя треугольная, нижняя треугольная) матрица и что обратная матрица для неособенной диагональной (верхней треугольной, нижней треугольной) матрицы является матрицей того же типа. Поэтому

1° Все диагональные, все верхние треугольные, все нижние треугольные матрицы n -го порядка образуют три коммутативные группы относительно операции сложения.

2° Все неособенные диагональные матрицы образуют коммутативную группу относительно умножения.

3° Все неособенные верхние (нижние) треугольные матрицы составляют группу (некоммутативную) относительно умножения

4. В заключение этого параграфа укажем на две важные операции над матрицами — транспонирование матрицы и переход к сопряженной матрице.

Если $A = \|a_{ik}\|$ ($i = 1, 2, \dots, m, k = 1, 2, \dots, n$), то транспонированная матрица A' определяется

равенством $A' = \|a'_{ik}\|$, где $a'_{ik} = a_{ki}$ ($i = 1, 2, \dots, m, k = 1, 2, \dots, n$). Сопряженная

же матрица $A^* = \|a^*_{ik}\|$, где $a^*_{ik} = \bar{a}'_{ik} = \bar{a}_{ki}$ ($i = 1, 2, \dots, m, k = 1, 2, \dots, n$). Если

матрица A имеет размеры $m \times n$, то матрицы A' и A^* имеют размеры $n \times m$.

Легко проверяются свойства:

$$1 \quad (A+B)' = A' + B', \quad (A+B)^* = A^* + B^*.$$

$$2 \quad (\alpha A)' = \alpha A', \quad (\alpha A)^* = \bar{\alpha} A^*.$$

$$3 \quad (AB)' = B' A', \quad (AB)^* = B^* A^*.$$

$$4 \quad (A^{-1})' = (A')^{-1}, \quad (A^{-1})^* = (A^*)^{-1}.$$

$$5 \quad (A')' = A, \quad (A^*)^* = A.$$

Продолжая этот алгоритм далее, мы на $(n-1)$ -м этапе приведем исходную систему (1) к треугольной рекуррентной системе

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= y_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ a_{33}^{(2)}x_3 + \dots + a_{3n}^{(2)}x_n &= y_3^{(2)} \\ &\dots\dots\dots \\ a_{nn}^{(n-1)}x_n &= y_n^{(n-1)} \end{aligned} \right\} \quad (6)$$

Это приведение выполнимо в том и только в том случае, если в процессе приведения все числа $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \dots, a_{n-1,n-1}^{(n-2)}$, оказываются отличными от нуля. Изложенный нами алгоритм Гаусса состоит из однотипных операции, которые легко выполняются на современных счетных машинах.

3. Выразим коэффициенты и правые части приведенной системы через коэффициенты и правые части исходной системы (1). При этом мы не будем предполагать, что в процессе приведения все числа $a_{11}, a_{22}^{(1)}, \dots, a_{n-1,n-1}^{(n-2)}$ оказываются отличными от нуля, а рассмотрим общий случай, когда первые p из этих чисел отличны от нуля:

$$a_{11} \neq 0, a_{22}^{(1)} \neq 0, \dots, a_{pp}^{(p-1)} \neq 0 \quad (p \leq n-1), \quad (7)$$

что дает возможность (на p -м этапе приведения) привести исходную систему уравнений к виду:

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ &\dots\dots\dots \\ a_{pp}^{(p-1)}x_p + \dots + a_{pn}^{(p-1)}x_n &= y_p^{(p-1)}, \\ a_{p+1,p+1}^{(p)}x_{p+1} + \dots + a_{p+1,n}^{(p)}x_n &= y_{p+1}^{(p)}, \\ &\dots\dots\dots \\ a_{n,p+1}^{(p)}x_{p+1} + \dots + a_{nn}^{(p)}x_n &= y_n^{(p)}. \end{aligned} \right\} \quad (8)$$

Матрицу коэффициентов в этой системе уравнений обозначим через G_p :

$$G_p = \left\| \begin{array}{cccc} a_{11} & a_{12} \dots a_{1p} & a_{1,p+1} & \dots a_{1n} \\ 0 & a_{22}^{(1)} \dots a_{2p}^{(1)} & a_{2,p+1}^{(1)} & \dots a_{2n}^{(1)} \\ \dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0 & 0 \dots a_{pp}^{(p-1)} & a_{p,p+1}^{(p-1)} & \dots a_{pn}^{(p-1)} \\ 0 & 0 \dots 0 & a_{p+1,p+1}^{(p)} & \dots a_{p+1,n}^{(p)} \\ \dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0 & 0 \dots\dots 0 & a_{n,p+1}^{(p)} & \dots a_{nn}^{(p)} \end{array} \right\|. \quad (9)$$

Переход от матрицы A к матрице G_p совершался следующим образом: к каждой строке матрицы A , начиная со 2-й и кончая n -й, последовательно прибавлялись какие-то предыдущие строки (из числа первых p), помноженные на некоторые коэффициенты.

Поэтому у матриц A и G_p одинаковы все миноры p -го порядка, содержащиеся в

первых P строках, а также все миноры $(p+1)$ -го порядка, содержащиеся в строках с номерами: $1, 2, \dots, p, i$ ($i > p$).

$$\left. \begin{aligned} A \begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} &= G_p \begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\ (1 \leq k_1 < k_2 < \dots < k_p \leq n), \\ A \begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} &= G_p \begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} \\ (1 \leq k_1 < k_2 < \dots < k_{p+1} \leq n). \end{aligned} \right\} \quad (10)$$

Из этих формул, учитывая структуру (9) матрицы G_p , найдем:

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)}, \quad (11)$$

$$A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)} a_{ik}^{(p)} \quad (i, k = p+1, \dots, n). \quad (12)$$

Деля почленно второе из этих равенств на первое, получим основные формулы

$$a_{ik}^{(p)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} \quad (i, k = p+1, \dots, n). \quad (13)$$

Если условия (7) выполнены для данного значения P , то такие же условия выполнены для любого меньшего значения P . Поэтому формулы (13) имеют место не только для данного значения P , но и для всех меньших значений P . То же можно сказать и о формуле (11).

Поэтому вместо этой формулы можно написать равенства

$$A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = a_{11}, \quad A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = a_{11} a_{22}^{(1)}, \quad A \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = a_{11} a_{22}^{(1)} a_{33}^{(2)}, \dots \quad (14)$$

Таким образом, условия (7), т. е. необходимые и достаточные условия выполнимости первых P этапов алгоритма Гаусса, могут быть записаны в виде следующих неравенств:

$$A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq 0, \quad A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \neq 0, \quad A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0. \quad (15)$$

Тогда из (14) находим:

$$\begin{aligned} a_{11} &= A \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad a_{22}^{(1)} = \frac{A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}{A \begin{pmatrix} 1 \\ 1 \end{pmatrix}}, \\ a_{33}^{(2)} &= \frac{A \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}{A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}, \quad a_{pp}^{(p-1)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}}. \end{aligned} \quad (16)$$

Для того чтобы в алгоритме исключения Гаусса можно было последовательно

исключить $x_1, x_2, \dots, x_p$ нужно, чтобы все величины (16) были отличны от нуля, т. е. чтобы

выполнялись неравенства (15). В то же время формулы для $a_{ik}^{(p)}$ имеют смысл, если выполняется только последнее из условий (15).

4. Пусть матрица коэффициентов в системе уравнений (1) имеет ранг r . Тогда надлежащей перестановкой уравнений и изменением нумерации неизвестных можно добиться выполнения неравенств

$$A \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \quad (j=1, 2, \dots, r). \quad (17)$$

Это позволяет последовательно исключить $x_1, x_2, \dots, x_r$ и получить систему уравнений

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ &\dots \\ a_{rr}^{(r-1)}x_r + \dots + a_{rn}^{(r-1)}x_n &= y_r^{(r-1)}, \\ a_{r+1,r+1}^{(r)}x_{r+1} + \dots + a_{rn}^{(r)}x_n &= y_{r+1}^{(r)}, \\ &\dots \\ a_{n,r+1}^{(r)}x_{r+1} + \dots + a_{nn}^{(r)}x_n &= y_n^{(r)}. \end{aligned} \right\} \quad (18)$$

Здесь коэффициенты определяются по формуле (13). Из этих формул, поскольку

ранг матрицы $A = \|a_{ik}\|_1^n$ равен r , следует, что

$$a_{ik}^{(r)} = 0 \quad (i, k = r+1, \dots, n)$$

и матрица G_r получающаяся из матрицы $A = \|a^{(-1)}_{ik}\|_1^n$ после применения r -этапного алгоритма исключения Гаусса, имеет вид

$$G_r = \left\| \begin{array}{cccccc} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array} \right\|. \quad (19)$$

Последние $n-r$ уравнений (18) сводятся к условиям совместности

$$y_i^{(r)} = 0 \quad (i = r+1, \dots, n). \quad (20)$$

Заметим, что столбец свободных членов при алгоритме исключения подвергается таким же преобразованиям, как и любой столбец коэффициентов. Поэтому, дополняя

матриц $A = \|a_{ik}\|_1^n$ $(n+1)$ -м столбцом из свободных членов, мы получим:

$$y_i^{(p)} = \frac{A \begin{pmatrix} 1 & \dots & p & i \\ 1 & \dots & p & n+1 \end{pmatrix}}{A \begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix}} \quad (j = 1, 2, \dots, n; p = 1, 2, \dots, r). \quad (21)$$

В частности, условия совместности (20) сводятся к известным условиям

$$A \begin{pmatrix} 1 & \dots & r & r+j \\ 1 & \dots & r & n+1 \end{pmatrix} = 0 \quad (j=1, 2, \dots, n-r). \quad (22)$$

Если $r = n$, т. е. матрица $A = \|a_{ik}\|$ неособенная, и

$$A \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \quad (j=1, 2, \dots, n),$$

то при помощи алгоритма Гаусса можно последовательно исключить $x_1, x_2, \dots, x_{n-1}$ и привести систему уравнений к виду (6).

2. Механическая интерпретация метода Гаусса.

Рассмотрим произвольную упругую статическую систему S , закрепленную по краям (например, струну, стержень, многопролетный стержень, мембрану, пластину или дискретную систему), и возьмем на ней n точек $(1), (2), \dots, (n)$.

Мы будем рассматривать перемещения

(прогибы) $y_1, y_2, \dots, y_n$ точек $(1), (2), \dots, (n)$ системы S под действием сил $F_1, F_2, \dots, F_n$, приложенных в этих же точках. Мы будем предполагать, что силы и перемещения параллельны одному и тому же направлению и потому определяются своими алгебраическими величинами (рис. 1).

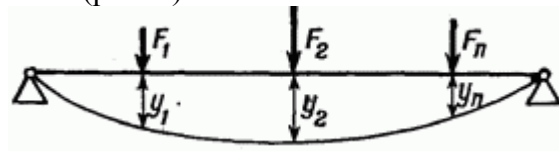


Рис. 1.

Кроме того, мы примем, что имеет место принцип линейного наложения сил:

1° При суммарном наложении двух систем сил соответствующие прогибы складываются.

2° При умножении величин всех сил на одно и то же вещественное число все прогибы умножаются на это число.

Обозначим через a_{ik} коэффициент влияния точки (k) на точку (i) , т. е. прогиб в точке (i) под действием единичной силы, приложенной в точке (k) ($i, k = 1, 2, \dots, n$) (рис.

2). Тогда при совместном действии сил $F_1, F_2, \dots, F_n$ прогибы $y_1, y_2, \dots, y_n$ определяются по формулам

$$\sum_{k=1}^n a_{ik} F_k = y_i \quad (i = 1, 2, \dots, n). \quad (23)$$

Сопоставляя (23) с исходной системой (1), мы задачу отыскания решения системы уравнений (1) можем интерпретировать так:

Даны прогибы $y_1, y_2, \dots, y_n$. Ищутся соответствующие силы $F_1, F_2, \dots, F_n$.

Обозначим через S_p статическую систему, получающуюся

из S введением p неподвижных шарнирных опор в точках $(1), (2), \dots, (p)$ ($p \leq n$).

Коэффициенты влияния для оставшихся подвижных

точек $(p+1), \dots, (n)$ системы S_p обозначим через

$$a_{ik}^{(p)} \quad (i, k = p+1, \dots, n)$$

(см. рисунок 3 для $p = 1$).

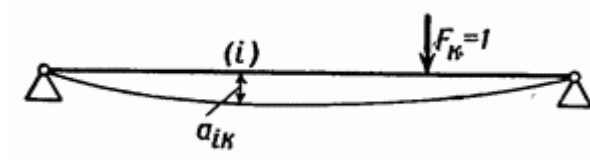


Рис. 2.

Коэффициент $a_{ik}^{(p)}$ можно рассматривать как прогиб в точке (i) системы S при действии единичной силы в точке (k) и сил реакций $R_1, R_2, \dots, R_p$ в закрепленных точках $(1), (2), \dots, (p)$. Поэтому

$$a_{ik}^{(p)} = R_1 a_{i1} + \dots + R_p a_{ip} + a_{ik}. \quad (24)$$

С другой стороны, при этих же силах прогибы системы в S точках $(1), (2), \dots, (p)$ равны нулю:

$$\left. \begin{aligned} R_1 a_{11} + \dots + R_p a_{1p} + a_{1k} &= 0, \\ R_1 a_{p1} + \dots + R_p a_{pp} + a_{pk} &= 0. \end{aligned} \right\} \quad (25)$$

Если

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0,$$

то мы можем из (25) определить $R_1, R_2, \dots, R_p$ и полученные выражения подставить в (24).

Это исключение $R_1, R_2, \dots, R_p$ можно сделать и так. К системе равенств (25) прибавим равенство (24), записанное в виде

$$R_1 a_{i1} + \dots + R_p a_{ip} + a_{ik} - a_{ik}^{(p)} = 0. \quad (24')$$

Рассматривая (25) и (24') как систему $p+1$ однородных уравнений, имеющую ненулевое решение $R_1, R_2, \dots, R_p, R_{p+1} = 1$, получаем, что определитель этой системы равен нулю:

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & a_{ik} \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & a_{pk} \\ a_{i1} & \dots & a_{ip} & a_{ik} - a_{ik}^{(p)} \end{vmatrix} = 0.$$

Отсюда

$$a_{ik}^{(p)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} \quad (i, k = p+1, \dots, n). \quad (26)$$

По этим формулам коэффициенты влияния «опорной» системы S_p выражаются через коэффициенты влияния исходной системы S .

Но формулы (26) совпадают с формулами (13) предыдущего параграфа. Поэтому для любого $(p \leq n-1)$ коэффициенты $a_{ik}^{(p)}$ $(i, k = p+1, \dots, n)$ в алгоритме Гаусса являются коэффициентами влияния опорной системы S_p .

В справедливости этого основного положения можно убедиться из чисто механических соображений, но опираясь на алгебраический вывод формул (13). Для этого рассмотрим

сначала частный случай одной опоры: $p = 1$ (рис. 3). В этом случае коэффициенты влияния системы S_1 определяются по формулам [полагаем $p = 1$ в (26)]:

$$\alpha_{ik}^{(1)} = \frac{A \begin{pmatrix} 1 & i \\ 1 & k \end{pmatrix}}{A \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \alpha_{ik} - \frac{\alpha_{ik}}{\alpha_{11}} \alpha_{1k} \quad (i, k = 1, 2, \dots, n).$$

Эти формулы совпадают с формулами (3').

Таким образом, если коэффициенты α_{ik} ($i, k = 1, 2, \dots, n$) в системе уравнений (1) являются коэффициентами влияния статической системы S , то коэффициенты $\alpha_{ik}^{(1)}$ ($i, k = 2, \dots, n$) в алгоритме Гаусса являются коэффициентами влияния системы S_1 . Применяя эти же соображения к системе S_1 и вводя в ней вторую опору в точке (2), получим, что коэффициенты $\alpha_{ik}^{(2)}$ ($i, k = 3, \dots, n$) в системе уравнений (4) являются коэффициентами влияния опорной системы S_2 и вообще для любого $(p \leq n-1)$ коэффициенты $\alpha_{ik}^{(p)}$ ($i, k = p+1, \dots, n$) в алгоритме Гаусса являются коэффициентами влияния опорной системы S_p .

Из механических соображений очевидно, что последовательное введение p опор равносильно одновременному введению этих опор.

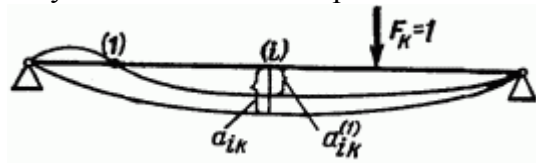


Рис. 3.

Замечание. Обращаем внимание на то, что при механической интерпретации алгоритма исключения не было необходимости предполагать, что точки, в которых рассматриваются прогибы, совпадают с точками приложения сил $F_1, F_2, \dots, F_n$. Можно считать, что $y_1, y_2, \dots, y_n$ — прогибы точек $(1), (2), \dots, (n)$, а силы $F_1, F_2, \dots, F_n$ приложены в точках $(1'), (2'), \dots, (n')$. Тогда α_{ik} — коэффициент влияния точки (k') на точку (i) . В этом случае вместо опоры в точке (j) следует от рассматривать обобщенную опору в точках $(j), (j')$, при которой прогиб в точке (j) поддерживается все время равным нулю за счет надлежащим образом выбранной вспомогательной силы R_j в точке (j') . Условие возможности введения p обобщенных опор в точках $(1), (1'); (2), (2'); \dots; (p), (p')$, т. е. возможность удовлетворить условиям $y_1 = 0, y_2 = 0, \dots, y_p = 0$ при любых $F_{p+1}, \dots, F_n$ за счет надлежащих $R_1 = F_1, \dots, R_p = F_p$, выражается неравенством

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0.$$

3. Разбиение матрицы на блоки.

Часто приходится пользоваться матрицами, разбитыми на прямоугольные части — «клетки» или «блоки». Рассмотрению таких «блочных» матриц мы посвящаем настоящий параграф.

1. Пусть дана прямоугольная матрица

$$A = \|a_{ik}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n). \quad (57)$$

При помощи горизонтальных и вертикальных линий рассечем матрицу A на прямоугольные блоки:

$$A = \left(\begin{array}{c|c|c|c} \overbrace{A_{11} \ A_{12} \ \dots \ A_{1t}}^{n_1} & & & \\ \hline \overbrace{A_{21} \ A_{22} \ \dots \ A_{2t}}^{n_2} & & & \\ \hline \dots & & & \\ \hline \overbrace{A_{s1} \ A_{s2} \ \dots \ A_{st}}^{n_s} & & & \end{array} \right) \begin{array}{l} \} m_1 \\ \} m_2 \\ \} \dots \\ \} m_s \end{array} \quad (58)$$

Про матрицу (58) будем говорить, что она разбита

на st блоков $A_{\alpha\beta}$ размером $m_\alpha \times n_\beta$ ($\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t$) или что она представлена в виде блочной матрицы. Вместо (58) будем сокращенно писать:

$$A = (A_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t). \quad (59)$$

В случае $s = t$ будем пользоваться и такой записью:

$$A = (A_{\alpha\beta})_1^s. \quad (60)$$

Действия над блочными матрицами производятся по тем же формальным правилам, как и в случае, когда вместо блоков имеем числовые элементы. Пусть, например, даны две прямоугольные матрицы одинаковых размеров и с одинаковым разбиением на блоки:

$$A = (A_{\alpha\beta}), \quad B = (B_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t). \quad (61)$$

Легко усмотреть, что

$$A + B = (A_{\alpha\beta} + B_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t). \quad (62)$$

Подробнее остановимся на умножении блочных матриц. Известно (см. гл. I, стр. 17), что при умножении двух прямоугольных матриц A и B длина строк в первом сомножителе A должна совпадать с высотой столбцов во втором сомножителе B . Для возможности блочного умножения этих матриц мы дополнительно потребуем, чтобы при разбиении на блоки все горизонтальные размеры в первом сомножителе совпадали с соответствующими вертикальными размерами во втором:

$$A = \left(\begin{array}{c|c|c|c} \overbrace{A_{11} \ A_{12} \ \dots \ A_{1t}}^{n_1} & & & \\ \hline \overbrace{A_{21} \ A_{22} \ \dots \ A_{2t}}^{n_2} & & & \\ \hline \dots & & & \\ \hline \overbrace{A_{s1} \ A_{s2} \ \dots \ A_{st}}^{n_s} & & & \end{array} \right) \begin{array}{l} \} m_1 \\ \} m_2 \\ \} \dots \\ \} m_s \end{array}, \quad B = \left(\begin{array}{c|c|c|c} \overbrace{B_{11} \ B_{12} \ \dots \ B_{1u}}^{p_1} & & & \\ \hline \overbrace{B_{21} \ B_{22} \ \dots \ B_{2u}}^{p_2} & & & \\ \hline \dots & & & \\ \hline \overbrace{B_{t1} \ B_{t2} \ \dots \ B_{tu}}^{p_t} & & & \end{array} \right) \begin{array}{l} \} n_1 \\ \} n_2 \\ \} \dots \\ \} n_t \end{array}. \quad (63)$$

Тогда легко проверить, что

$$AB = C = (C_{\alpha\beta}), \quad \text{где} \quad C_{\alpha\beta} = \sum_{\delta=1}^t A_{\alpha\delta} B_{\delta\beta} \quad \left(\begin{array}{l} \alpha = 1, 2, \dots, s \\ \beta = 1, 2, \dots, u \end{array} \right). \quad (64)$$

Отдельно отметим тот частный случай, когда одним из сомножителей является квазидиагональная матрица. Пусть A — квазидиагональная матрица, т.

е. $s = t$ и $A_{\alpha\beta} = 0$ при $\alpha \neq \beta$. В этом случае формула (64) нам дает:

$$C_{\alpha\beta} = A_{\alpha\alpha} B_{\beta\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (65)$$

При умножении блочной матрицы слева на квазидиагональную матрицу строки блочной матрицы умножаются слева на соответствующие диагональные клетки квазидиагональной матрицы.

Пусть теперь B — квазидиагональная матрица, т. е. $t = u$ и $B_{\alpha\beta} = 0$ при $\alpha \neq \beta$. Тогда из (64) получаем:

$$C_{\alpha\beta} = A_{\alpha\beta} B_{\beta\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (66)$$

При умножении блочной матрицы справа на квазидиагональную все столбцы блочной матрицы умножаются справа на соответствующие диагональные клетки квазидиагональной матрицы.

Заметим, что умножение квадратных блочных матриц одного и того же порядка всегда выполнимо, когда сомножители разбиты на одинаковые квадратные схемы блоков и в каждом из сомножителей на диагональных местах стоят квадратные матрицы.

Блочная матрица (58) называется верхней (нижней) квазитреугольной, если $s = t$ и все $A_{\alpha\beta} = 0$ при $\alpha > \beta$ (соответственно все $A_{\alpha\beta} = 0$ при $\alpha < \beta$). Частным случаем квазитреугольной матрицы является квазидиагональная матрица.

Из формулы (64) легко усмотреть, что произведение двух верхних (нижних) квазитреугольных матриц является снова верхней (нижней) квазитреугольной матрицей; при этом диагональные блоки произведения получаются путем перемножения соответствующих диагональных блоков сомножителей.

Действительно, полагая в (64) $s = t$ и

$$A_{\alpha\beta} = 0, \quad B_{\alpha\beta} = 0 \quad \text{при } \alpha < \beta,$$

найдем:

$$\left. \begin{array}{l} C_{\alpha\beta} = 0 \text{ при } \alpha < \beta \\ \text{и} \\ C_{\alpha\alpha} = A_{\alpha\alpha} B_{\alpha\alpha} \end{array} \right\} \quad (\alpha, \beta = 1, 2, \dots, s).$$

Аналогично разбирается случай нижних квазитреугольных матриц.

Отметим правило вычисления определителя квазитреугольной матрицы. Это правило можно получить, исходя из разложения Лапласа.

Если A — квазитреугольная (в частности, квазидиагональная) матрица с квадратными диагональными блоками, то определитель этой матрицы равен произведению определителей диагональных блоков:

$$|A| = |A_{11}| |A_{22}| \dots |A_{ss}| \quad (67)$$

1.3 Лекция №3 (1 час).

Тема: «Уравнения различных типов»

1.3.1 Вопросы лекции:

1. Уравнение вида $AX = XB$.

1.3.2 Краткое содержание вопросов:

Пусть дано уравнение

$$AX = XB, \quad (1)$$

где A и B — две заданные квадратные матрицы (вообще говоря, разных порядков)

$$A = \|\alpha_{ij}\|_1^m, \quad B = \|\beta_{kl}\|_1^n,$$

а X – искомая прямоугольная матрица размером $m \times n$:

$$X = \|x_{jk}\| \quad (j=1, 2, \dots, m, k=1, 2, \dots, n).$$

Выпишем элементарные делители матриц A и B (в поле комплексных чисел):

$$(A) : (\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (p_1 + p_2 + \dots + p_u = m),$$

$$(B) : (\lambda - \mu_1)^{q_1}, (\lambda - \mu_2)^{q_2}, \dots, (\lambda - \mu_v)^{q_v} \quad (q_1 + q_2 + \dots + q_v = n).$$

В соответствии с этими элементарными делителями приведем матрицы A и B к нормальной жордановой форме

$$A = U \tilde{A} U^{-1}, \quad B = V \tilde{B} V^{-1}, \quad (2)$$

где U и V – квадратные неособенные матрицы соответственно порядка m и n ,

а $\tilde{A}$ и $\tilde{B}$ – жордановы матрицы:

$$\tilde{A} = \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \right\},$$

$$\tilde{B} = \left\{ \mu_1 E^{(q_1)} + H^{(q_1)}, \mu_2 E^{(q_2)} + H^{(q_2)}, \dots, \mu_v E^{(q_v)} + H^{(q_v)} \right\}. \quad (3)$$

Подставляя в уравнение (1) вместо A и B их выражения (2), получим:

$$U \tilde{A} U^{-1} X = X V \tilde{B} V^{-1}.$$

Умножим обе части этого равенства слева на U^{-1} , а справа – на V :

$$A U^{-1} X V = U^{-1} X V \tilde{B}. \quad (4)$$

Вводя вместо искомой матрицы X новую искомую матрицу $\tilde{X}$ (тех же размеров $m \times n$)

$$\tilde{X} = U^{-1} X V, \quad (5)$$

мы уравнение (4) запишем так:

$$\tilde{A} \tilde{X} = \tilde{X} \tilde{B}. \quad (6)$$

Мы заменили матричное уравнение (1) уравнением (6) того же вида, но в котором заданные матрицы имеют нормальную жорданову форму.

В соответствии с квазидиагональным видом матриц $\tilde{A}$ и $\tilde{B}$ разобьем матрицу $\tilde{X}$ на блоки:

$$\tilde{X} = \{X_{\alpha\beta}\} \quad (\alpha=1, 2, \dots, u; \beta=1, 2, \dots, v)$$

(здесь $X_{\alpha\beta}$ – прямоугольная матрица размером $p_\alpha \times q_\beta$; $\alpha=1, 2, \dots, u$; $\beta=1, 2, \dots, v$).

Используя правило умножения блочной матрицы на квазидиагональную (см. стр. 56), произведем умножение матриц в левой и правой частях уравнения (6). Тогда это уравнение распадается на uv матричных уравнений

$$\left[\lambda_\alpha E^{(p_\alpha)} + H^{(p_\alpha)} \right] X_{\alpha\beta} = X_{\alpha\beta} \left[\mu_\beta E^{(q_\beta)} + H^{(q_\beta)} \right] \quad (\alpha=1, 2, \dots, u; \beta=1, 2, \dots, v),$$

которые перепишем еще так:

$$(\mu_\beta - \lambda_\alpha) X_{\alpha\beta} = H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta \quad (\alpha=1, 2, \dots, u; \beta=1, 2, \dots, v); \quad (7)$$

при этом мы ввели сокращенные обозначения

$$H_\alpha = H^{(p_\alpha)}, G_\beta = H^{(q_\beta)} \quad (\alpha=1, 2, \dots, u; \beta=1, 2, \dots, v). \quad (8)$$

Возьмем какое-нибудь из уравнений (7). Могут представиться два случая:

1. $\lambda_\alpha \neq \mu_\beta$. Проитерлируем $r-1$ раз равенство (7):

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{\sigma+\tau=r} (-1)^\tau H_\alpha^\sigma X_{\alpha\beta} G_\beta^\tau. \quad (9)$$

Заметим, что в силу (8)

$$H_\alpha^{p_\alpha} = G_\beta^{q_\beta} = 0. \quad (10)$$

Если в (9) взять $r \geq p_\alpha + q_\beta - 1$, то в каждом члене суммы, стоящей в правой части равенства (9), выполняется по крайней мере одно из соотношений

$$\sigma \geq p_\alpha, \quad \tau \geq q_\beta,$$

и потому в силу (10) либо $H_\alpha^\sigma = 0$, либо $G_\beta^\tau = 0$. Так как, кроме того, в рассматриваемом случае $\lambda_\alpha \neq \mu_\beta$, то из (9) находим:

$$X_{\alpha\beta} = 0. \quad (11)$$

2. $\lambda_\alpha = \mu_\beta$. В этом случае уравнение (7) принимает вид

$$H_\alpha X_{\alpha\beta} = X_{\alpha\beta} G_\beta. \quad (12)$$

В матрицах H_α и G_β элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Учитывая эту специфичную структуру матриц H_α и G_β и полагая

$$X_{\alpha\beta} = \|\xi_{ik}\| \quad (i = 1, 2, \dots, p_\alpha; k = 1, 2, \dots, q_\beta),$$

мы заменим матричное уравнение (12) следующей эквивалентной ему системой скалярных соотношений:

$$\xi_{i+1,k} = \xi_{i,k-1} \left(\xi_{i0} = \xi_{p_\alpha+1,k} = 0; i = 1, 2, \dots, p_\alpha; k = 1, 2, \dots, q_\beta \right). \quad (13)$$

Равенства (13) означают:

- 1) В матрице $X_{\alpha\beta}$ на каждой линии, параллельной главной диагонали, стоят равные между собой элементы,
- 2) $\xi_{21} = \xi_{31} = \dots = \xi_{p_\alpha 1} = \xi_{p_\alpha 2} = \dots = \xi_{p_\alpha, q_\beta-1} = 0$.

Пусть $p_\alpha = q_\beta$. В этом случае $X_{\alpha\beta}$ – квадратная матрица. Из 1), 2) следует, что в матрице $X_{\alpha\beta}$ все элементы, расположенные под главной диагональю, равны нулю, все элементы главной диагонали равны некоторому числу $c_{\alpha\beta}$, все элементы первой наддиагонали равны некоторому числу $c'_{\alpha\beta}$ и т. д., т. е.

$$X_{\alpha\beta} = \left\| \begin{array}{ccccc} c_{\alpha\beta} & c'_{\alpha\beta} & \dots & \dots & c_{\alpha\beta}^{(p_\alpha-1)} \\ 0 & c_{\alpha\beta} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & c'_{\alpha\beta} \\ 0 & \dots & \dots & 0 & c_{\alpha\beta} \end{array} \right\| = T_{p_\alpha}; \quad (14)$$

$$(p_\alpha = q_\beta)$$

здесь $c_{\alpha\beta}, c'_{\alpha\beta}, \dots, c_{\alpha\beta}^{(p_\alpha-1)}$ – произвольные параметры (уравнения (12) не накладывают никаких ограничений на значения этих параметров).

Легко видеть, что при $p_\alpha < q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} q_\beta - p_\alpha & \\ & \ddots \\ & & 0 \\ & & & T_{p_\alpha} \end{pmatrix}, \quad (15)$$

а при $p_\alpha > q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} T_{q_\beta} & \\ & \ddots \\ & & 0 \end{pmatrix}_{p_\alpha - q_\beta}. \quad (16)$$

Про матрицы (14), (15) и (16) мы будем говорить, что они имеют правильную верхнюю треугольную форму. Число произвольных параметров в $X_{\alpha\beta}$ равно наименьшему из чисел p_α и q_β . Приведенная ниже схема показывает структуру матрицы $X_{\alpha\beta}$ при $\lambda_\alpha = \mu_\beta$ (произвольные параметры здесь обозначены через a, b, c, d):

$$X_{\alpha\beta} = \begin{pmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{pmatrix}, \quad X_{\alpha\beta} = \begin{pmatrix} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{pmatrix}, \quad X_{\alpha\beta} = \begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}.$$

$(p_\alpha = q_\beta = a) \qquad (p_\alpha = 3, q_\beta = 5) \qquad (p_\alpha = 5, q_\beta = 3)$

Для того чтобы при подсчете произвольных параметров в матрице $\tilde{X}$ охватить и случай 1, обозначим через $d_{\alpha\beta}(\lambda)$ наибольший общий делитель элементарных делителей $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \mu_\beta)^{q_\beta}$, а через $\delta_{\alpha\beta}$ – степень многочлена $d_{\alpha\beta}(\lambda)$ ($\alpha = 1, 2, \dots, u; \beta = 1, 2, \dots, v$). В случае 1 $\delta_{\alpha\beta} = 0$; в случае 2 имеем: $\delta_{\alpha\beta} = \min(p_\alpha, q_\beta)$. Таким образом, в обоих случаях число произвольных параметров в $X_{\alpha\beta}$ равно $\delta_{\alpha\beta}$. Число произвольных параметров в $\tilde{X}$ определяется формулой

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta}.$$

В дальнейшем нам удобно будет общее решение уравнения (6) обозначить через $X_{\overline{A}\overline{B}}$ (до сих пор мы это решение обозначали буквой $\tilde{X}$).

Полученные в этом параграфе результаты можно сформулировать в виде следующей теоремы:

Теорема 1. Общее решение матричного уравнения

$$AX = XB,$$

где

$$A = \|a_{ik}\|_l^m = U \tilde{A} U^{-1} = U \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \right\} U^{-1},$$

$$B = \|b_{ik}\|_l^n = V \tilde{B} V^{-1} = V \left\{ \mu_1 E^{(q_1)} + H^{(q_1)}, \dots, \mu_v E^{(q_v)} + H^{(q_v)} \right\} V^{-1},$$

задается формулой

$$X = UX_{\overline{AB}}V^{-1}. \quad (17)$$

Здесь $X_{\overline{AB}}$ – общее решение уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$ – имеет следующую структуру: $X_{\overline{AB}}$ разбивается на блоки

$$X_{\overline{AB}} = \left(\overbrace{X_{\alpha\beta}}^{q_{\beta}} \right) p_{\alpha} \quad (\alpha = 1, 2, \dots, u; \beta = 1, 2, \dots, v);$$

если $\lambda_{\alpha} \neq \mu_{\beta}$, то на месте $X_{\alpha\beta}$ стоит нулевая матрица, если же $\lambda_{\alpha} = \mu_{\beta}$, то на месте $X_{\alpha\beta}$ стоит произвольная правильная верхняя треугольная матрица.

$X_{\overline{AB}}$, а следовательно, и X зависят линейно от N произвольных параметров $c_1, c_2, \dots, c_N$:

$$X = \sum_{j=1}^N c_j X_j, \quad (18)$$

где N определяется формулой

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} \quad (19)$$

[здесь $\delta_{\alpha\beta}$ обозначает степень наибольшего общего делителя $(\lambda - \lambda_{\alpha})^{p_{\alpha}}$ и $(\lambda - \mu_{\beta})^{q_{\beta}}$].

Заметим, что матрицы $X_1, X_2, \dots, X_N$, фигурирующие в формуле (18), суть решения исходного уравнения (1) (матрица X_j получается из X , если параметру c_j дать значение единицы, а остальным параметрам – нулевые значения; $j = 1, 2, \dots, N$). Эти решения линейно независимы, так как в противном случае при некоторых значениях параметров $c_1, c_2, \dots, c_N$, не равных одновременно нулю, матрица X , а следовательно, и $X_{\overline{AB}}$ равнялись бы нулю, что невозможно. Таким образом, равенство (19) показывает, что любое решение исходного уравнения представляет собой линейную комбинацию N линейно независимых решений.

Если матрицы A и B не имеют

общих характеристических чисел (характеристические многочлены $|\lambda E - A|$ и $|\lambda E - B|$ вза

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} = 0$$

имно просты), то и, следовательно, $X = 0$, т. е. в этом случае уравнение (1) имеет только тривиальное нулевое решение $X = 0$.

Замечание. Пусть элементы матриц A и B принадлежат некоторому числовому полю K .

Тогда нельзя утверждать, что элементы матриц $U, V, X_{\overline{AB}}$ фигурирующих в формуле (17), также принадлежат полю K . Элементы этих матриц можно выбрать в расширенном поле K_1 , которое получается из поля K путем приобщения к последнему корней

характеристических уравнений $|\lambda E - A| = 0$ и $|\lambda E - B| = 0$. С такого рода расширением основного поля всегда приходится иметь дело, когда пользуются приведением заданных матриц к нормальной жордановой форме.

Однако матричное уравнение (1) эквивалентно системе m линейных однородных уравнений, где неизвестными служат

элементы x_{jk} $j = 1, 2, \dots, m; k = 1, 2, \dots, n$ искомой матрицы X :

$$\sum_{j=1}^m a_{ij} x_{jk} = \sum_{h=1}^n x_{ih} b_{hk} \quad (i = 1, 2, \dots, m, k = 1, 2, \dots, n) \quad (20)$$

Нами доказано, что эта система имеет N линейно независимых решений, где N определяется формулой (19). Но известно, что базисные линейно независимые решения можно выбрать в основном поле K , которому принадлежат коэффициенты уравнений (20). Таким образом, в формуле (18) матрицы $X_1, X_2, \dots, X_N$ можно выбрать так, чтобы их элементы принадлежали полю K . Тогда, придавая в формуле (18) произвольным параметрам всевозможные значения из поля K , мы получим все матрицы X с элементами из K , удовлетворяющие уравнению (1).

1.4 Лекция №4 (1 час).

Тема: «Извлечение корня m -ной степени из матрицы»

1.4.1 Вопросы лекции:

1. Правила извлечения корней из матриц

1.4.2 Краткое содержание вопросов:

1. Правила извлечения корней из матриц

Рассмотрим уравнению

$$X^m = A, \quad (54)$$

где A – заданная, а X – искомая матрицы (обе порядка n), m – данное целое положительное число.

В данном параграфе мы рассмотрим случай, когда $|A| \neq 0$ (A – неособенная матрица). В этом случае все характеристические числа матрицы A отличны от нуля (ибо $|A|$ равен произведению этих характеристических чисел). Обозначим через

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u}$$

элементарные делители матрицы A и приведем матрицу A к жордановой форме:

$$A = U \tilde{A} U^{-1} = U \{ \lambda_1 E_1 + H_1, \dots, \lambda_u E_u + H_u \} U^{-1}.$$

Так как характеристические числа искомой матрицы X при возведении в m -ю степень дают характеристические числа матрицы A , то и у матрицы X все характеристические числа отличны от нуля. Поэтому на этих характеристических числах производная от

$$f(\lambda) = \lambda^m$$

не обращается в нуль. Но в таком случае (см. гл. VI, стр. 159) элементарные делители матрицы X не «расщепляются» при возведении матрицы X в m -ю степень. Из сказанного следует, что элементарными делителями матрицы X будут:

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u},$$

где $\xi_j^m = \lambda_j$, т. е. ξ_j является одним из корней m -й степени из λ_j ($\xi_j = \sqrt[m]{\lambda_j}$; $j = 1, 2, \dots, u$).

Определим теперь $\sqrt[m]{\lambda_j E_j + H_j}$ следующим образом. Возьмем в λ -плоскости круг с центром в точке λ_j , не захватывающий нуля. В этом круге мы имеем m отдельных ветвей функции $\sqrt[m]{\lambda}$. Эти ветви можно отличать одну от другой по значениям, которые

они принимают в центре круга, в точке λ_j . Обозначим через $\sqrt[m]{\lambda}$ ту ветвь, значение которой в точке λ_j , совпадает с характеристическим числом ξ_j искомой матрицы X , и, исходя из этой ветви, определим функцию от матрицы $\sqrt[m]{\lambda_j E_j + H_j}$ с помощью обрывающегося ряда

$$\sqrt[m]{\lambda_j E_j + H_j} = \lambda_j^{\frac{1}{m}} E_j + \frac{1}{m} \lambda_j^{\frac{1}{m}-1} H_j + \frac{1}{2!} \frac{1}{m} \left(\frac{1}{m} - 1 \right) \lambda_j^{\frac{1}{m}-2} H_j^2 + \dots$$

Так как производная от рассматриваемой функции $\sqrt[m]{\lambda}$ в точке λ_j не равна нулю, то матрица (58) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$,

где $\xi_j = \sqrt[m]{\lambda_j}$ (здесь $j = 1, 2, \dots, u$). Отсюда следует, что квазидиагональная матрица

$$\left\{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \right\}$$

имеет элементарные делители (57), т. е. те же элементарные делители, что и искомая матрица X . Поэтому существует такая неособенная матрица $T \left(|T| \neq 0 \right)$, что

$$X = T \left\{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \right\} T^{-1}$$

Для определения матрицы T заметим, что, подставляя в обе части тождества

$$\left(\sqrt[m]{\lambda} \right)^m = \lambda$$

вместо λ матрицу $\lambda_j E_j + H_j$ ($j = 1, 2, \dots, u$), получим:

$$\left(\sqrt[m]{\lambda_j E_j + H_j} \right)^m = \lambda_j E_j + H_j \quad (j = 1, 2, \dots, u)$$

Теперь из (54) и (59) следует:

$$A = T \left\{ \lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u \right\} T^{-1}$$

Сопоставляя (56) и (60), найдем:

$$T = UX_{\bar{A}},$$

где $X_{\bar{A}}$ – произвольная неособенная матрица, перестановочная с $\bar{A}$ (структура матрицы $X_{\bar{A}}$ детально описана в § 2).

Подставляя в (59) вместо T выражение $UX_{\bar{A}}$, получаем формулу, охватывающую все решения уравнения (54):

$$X = UX_{\bar{A}} \left\{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \right\} X_{\bar{A}}^{-1} U^{-1}$$

Мнозначность правой части этой формулы имеет как дискретный, так и континуальный характер: дискретный (в данном случае и конечный) характер этой многозначности

получается за счет выбора различных ветвей функции $\sqrt[m]{\lambda}$ в различных клетках квазидиагональной матрицы (при этом даже при $\lambda_j = \lambda_k$ ветви $\sqrt[m]{\lambda}$ в j -й и в k -й диагональных клетках могут быть различными); континуальный характер многозначности получается за счет произвольных параметров, содержащихся в матрице $X_{\bar{A}}$.

Все решения уравнения (54) мы будем называть корнями m -й степени из матрицы A и обозначать многозначным символом $\sqrt[m]{A}$. Обратим внимание на то, что $\sqrt[m]{A}$ в общем случае не является функцией от матрицы A (т. е. не представляется в виде многочлена от A).

Замечание. Если все элементарные делители матрицы A попарно взаимно просты, т. е. числа $\lambda_1, \lambda_2, \dots, \lambda_u$ все различны, то матрица $X_{\bar{A}}$ имеет квазидиагональный вид

$$X_{\bar{A}} = \{X_1, X_2, \dots, X_u\},$$

где матрица X_j перестановочна с $\lambda_j E_j + H_j$ и, следовательно, перестановочна с любой функцией от $\lambda_j E_j + H_j$, в частности с $\sqrt[m]{\lambda_j E_j + H_j}$ ($j = 1, 2, \dots, u$). Поэтому в рассматриваемом случае формула (62) принимает вид

$$X = U \{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \} U^{-1}.$$

Таким образом, если элементарные делители матрицы A попарно взаимно просты, то в формуле для $X = \sqrt[m]{A}$ имеется только дискретная многозначность. В этом случае любое значение $\sqrt[m]{A}$ можно представить как многочлен от A .

Пример. Пусть требуется найти все квадратные корни из матрицы

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

т. е. все решения уравнения

$$X^2 = A.$$

В данном случае матрица A уже имеет нормальную жорданову форму. Поэтому в

формуле (62) можно положить $A = \tilde{A}, U = E$. Матрица $X_{\bar{A}}$ в данном случае выглядит так (см. стр. 204):

$$X_{\bar{A}} = \begin{pmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{pmatrix},$$

где a, b, c, d, e – произвольные параметры.

Формула (62), дающая все искомые решения X , в данном случае принимает следующий вид:

$$X = \begin{pmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{pmatrix} \begin{pmatrix} \varepsilon & \frac{\varepsilon}{2} & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \eta \end{pmatrix} \begin{pmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{pmatrix}^{-1} \quad (\varepsilon^2 = \eta^2 = 1)$$

Не изменяя X , мы можем в формуле (62) помножить $X_{\bar{A}}$ на такой скаляр, чтобы $|X_{\bar{A}}| = 1$.

В данном случае это приведет к равенству $a^2 e = 1$, откуда $e = a^{-2}$.

Вычислим элементы матрицы $X_{\bar{A}}^{-1}$. Для этого выпишем линейное преобразование с матрицей коэффициентов $X_{\bar{A}}$:

$$y_1 = ax_1 + bx_2 + cx_3,$$

$$y_2 = ax_2,$$

$$y_3 = dx_2 + a^{-2}x_3.$$

Разрешим эту систему уравнений относительно x_1, x_2, x_3 . Тогда получим преобразование с обратной матрицей X_A^{-1} :

$$\begin{aligned}x_1 &= a^{-1}y_1 - (a^{-2}b - cd)y_2 - acy_3, \\x_2 &= a^{-1}y_2, \\x_3 &= -ady_2 + a^2y_3.\end{aligned}$$

Отсюда находим:

$$X_A^{-1} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{vmatrix}^{-1} = \begin{vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{vmatrix}.$$

Формула (63) дает:

$$X = \begin{vmatrix} \varepsilon & (\varepsilon - \eta)acd + \frac{\varepsilon}{2} & a^2c(\eta - \varepsilon) \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta)da^{-1} & \eta \end{vmatrix} = \begin{vmatrix} \varepsilon & (\varepsilon - \eta)vw + \frac{\varepsilon}{2} & (\eta - \varepsilon)v \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta)w & \eta \end{vmatrix}$$

$$(v = a^2c; w = a^{-1}d).$$

Решение X зависит от двух произвольных параметров v и w и от двух произвольных знаков ε и η .

1.5 Лекция №5 (1 час).

Тема: «Скалярное уравнение»

1.5.1 Вопросы лекции:

1. Скалярное уравнение вида $f(x)=0$

1.5.2 Краткое содержание вопросов:

1. Скалярное уравнение вида $f(x)=0$

Рассмотрим сначала уравнение

$$g(X) = 0,$$

где

$$g(\lambda) = (\lambda - \lambda_1)^{a_1} (\lambda - \lambda_2)^{a_2} \dots (\lambda - \lambda_h)^{a_h}$$

– заданный многочлен переменной λ , а X – искомая квадратная матрица порядка n . Так как минимальный многочлен матрицы X , т. е. первый инвариантный многочлен $i_1(\lambda)$, должен быть делителем многочлена $g(\lambda)$, то элементарные делители матрицы X должны иметь следующий вид:

$$(\lambda - \lambda_{i_1})^{p_{i_1}}, (\lambda - \lambda_{i_2})^{p_{i_2}}, \dots, (\lambda - \lambda_{i_v})^{p_{i_v}} \left(\begin{array}{l} i_1, i_2, \dots, i_v = 1, 2, \dots, h, \\ p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_v} \leq a_{i_v}, \\ p_{i_1} + p_{i_2} + \dots + p_{i_v} = n \end{array} \right)$$

(среди индексов $i_1, i_2, \dots, i_v$ могут быть и равные, n – заданный порядок искомой матрицы X).

Искомая матрица X представится в виде

$$X = T \left\{ \lambda_{i_1} E^{(p_1)} + H^{(p_1)}, \dots, \lambda_{i_v} E^{(p_v)} + H^{(p_v)} \right\} T^{-1},$$

где T – произвольная неособенная матрица порядка n . Множество решений уравнения с заданным порядком искомой матрицы распадается согласно формуле на конечное число классов подобных между собой матриц.

Пример 1. Дано уравнение

$$X^m = 0.$$

Если некоторая степень матрицы равна нулю, то матрица называется нильпотентной.

Наименьший из показателей, при которых степень матрицы равна нулю, называется индексом нильпотентности данной матрицы.

Очевидно, решениями уравнения (35) являются все нильпотентные матрицы с индексом нильпотентности $\mu \leq m$. Формула, охватывающая все решения данного порядка n , выглядит так:

$$X = T \left\{ H^{(p_1)}, H^{(p_2)}, \dots, H^{(p_v)} \right\} T^{-1} \begin{pmatrix} p_1, p_2, \dots, p_v \leq m, \\ p_1 + p_2 + \dots + p_v = n \end{pmatrix}$$

(T – произвольная неособенная матрица).

Пример 2. Дано уравнение

$$X^2 = X.$$

Матрица, удовлетворяющая этому уравнению, называется идемпотентной.

Элементарными делителями идемпотентной матрицы могут быть только λ либо $\lambda - 1$.

Поэтому идемпотентную матрицу можно определить как матрицу простой структуры (т. е. приводящуюся к диагональной форме) с характеристическими числами, равными нулю или единице. Формула, охватывающая все идемпотентные матрицы данного порядка, имеет вид

$$X = T \underbrace{\{1, 1, \dots, 1, 0, \dots, 0\}}_n T^{-1},$$

где T – произвольная неособенная матрица порядка n .

Рассмотрим теперь более общее уравнение

$$f(X) = 0,$$

где $f(\lambda)$ – регулярная функция в некоторой области G плоскости комплексного

аргумента λ . От искомого решения $X = \|x_{ik}\|_1^n$ будем требовать,

чтобы характеристические числа его принадлежали области G . Выпишем все нули

функции $f(\lambda)$, лежащие в области G , и их кратности:

$$\lambda_1, \lambda_2, \dots,$$

$$a_1, a_2, \dots$$

Как и в предыдущем случае, каждый элементарный делитель матрицы X должен иметь вид

$$(\lambda - \lambda_i)^{p_i} \quad (p_i \leq a_i)$$

и потому

$$X = T \left\{ \lambda_{i_1} E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_v} E^{(p_{i_v})} + H^{(p_{i_v})} \right\} T^{-1}$$

$$\left(\begin{array}{l} i_1, i_2, \dots, i_v = 1, 2, \dots; p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_v} \leq a_{i_v}; \\ p_{i_1} + p_{i_2} + \dots + p_{i_v} = n \end{array} \right)$$

(T – произвольная неособенная матрица).

1. 6 Лекция №6 (1 час).

Тема: «Логарифм матрицы»

1.6.1 Вопросы лекции:

1. Принципы нахождения логарифма матрицы.

1.6.2 Краткое содержание вопросов:

1. Принципы нахождения логарифма матрицы.

1. Рассмотрим матричное уравнение

$$e^X = A. \quad (90)$$

Все решения этого уравнения будем называть логарифмами (натуральными) матрицы A и обозначать через $\ln A$.

Характеристические числа λ_j матрицы A связаны с характеристическими числами ξ_j матрицы X формулой $\lambda_j = e^{\xi_j}$; поэтому, если уравнение (90) имеет решение, то все характеристические числа матрицы A отличны от нуля и матрица A является неособенной ($|A| \neq 0$). Таким образом, условие $|A| \neq 0$ является необходимым для существования решения уравнения (90). Ниже мы увидим, что это условие является и достаточным.

Итак, пусть $|A| \neq 0$. Выпишем элементарные делители матрицы A :

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (\lambda_1 \lambda_2 \dots \lambda_u \neq 0, p_1 + p_2 + \dots + p_u = n). \quad (91)$$

В соответствии с этими элементарными делителями приведем матрицу A к нормальной жордановой форме:

$$A = U \tilde{A} U^{-1} = U \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \right\} U^{-1}. \quad (92)$$

Так как производная от функции e^{ξ} отлична от нуля при всех значениях ξ , то (см. гл. VI, стр. 159) при переходе от матрицы X к матрице $A = e^X$ элементарные делители не расщепляются, т. е. матрица X имеет элементарные делители

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}, \quad (93)$$

где $e^{\xi_j} = \lambda_j$ ($j = 1, 2, \dots, u$), т. е. ξ_j есть одно из значений $\ln \lambda_j$ ($j = 1, 2, \dots, u$).

Возьмем в плоскости комплексного переменного λ круг с центром в

точке λ_j радиуса $< |\lambda_j|$ и обозначим через $f_j(\lambda) = \ln \lambda$ ту из ветвей функции $\ln \lambda$ в

рассматриваемом круге, которая в точке λ_j принимает значение,

равное характеристическому числу ξ_j матрицы X ($j = 1, 2, \dots, u$). После этого полагаем:

$$\ln(\lambda_j E^{(p_j)} + H^{(p_j)}) = f_j(\lambda_j E^{(p_j)} + H^{(p_j)}) = \ln \lambda_j E^{(p_j)} + \lambda_j^{-1} H^{(p_j)} + \dots \quad (94)$$

Так как производная от $\ln \lambda$ нигде не обращается в нуль (в конечной части плоскости λ), то матрица (94) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$. В силу этого квазидиагональная матрица

$$\left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_n E^{(p_n)} + H^{(p_n)}) \right\} \quad (95)$$

имеет те же элементарные делители, что и искомая матрица X . Поэтому существует такая матрица $T(|T| \neq 0)$, что

$$X = T \left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln(\lambda_n E^{(p_n)} + H^{(p_n)}) \right\} T^{-1} \quad (96)$$

Для определения матрицы T заметим, что

$$A = e^X = T \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_n E^{(p_n)} + H^{(p_n)} \right\} T^{-1} \quad (97)$$

Сопоставляя (97) с (92), находим:

$$T = UX_{\tilde{A}}, \quad (98)$$

где $X_{\tilde{A}}$ – произвольная матрица, перестановочная с матрицей $\tilde{A}$. Подставляя выражение для T из (98) в (96), получим общую формулу, охватывающую все логарифмы матрицы:

$$X = UX_{\tilde{A}} \left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_n E^{(p_n)} + H^{(p_n)}) \right\} X_{\tilde{A}}^{-1} U^{-1} \quad (99)$$

Замечание. Если все элементарные делители матрицы A взаимно просты, то в правой части формулы (99) можно выбросить множители $X_{\tilde{A}}$ и $X_{\tilde{A}}^{-1}$ (см. аналогичное замечание на стр. 213).

2. Выясним, когда вещественная неособенная матрица A имеет вещественный логарифм X . Пусть искомая матрица имеет несколько элементарных делителей,

отвечающих характеристическому числу вида $\rho + i\pi: (\lambda - \rho - i\pi)^{q_1}, \dots, (\lambda - \rho - i\pi)^{q_s}$.

Поскольку матрица X вещественна, то она имеет и сопряженные элементарные

делители: $(\lambda - \rho + i\pi)^{q_1}, \dots, (\lambda - \rho + i\pi)^{q_s}$. При переходе от матрицы X к

матрице A элементарные делители не расщепляются,

но характеристические числа $\rho + i\pi, \rho - i\pi$ заменяются в них

числами $e^{\rho + i\pi} = -\mu, e^{\rho - i\pi} = -\mu$, где $\mu = e^{\rho} > 0$. Поэтому в системе элементарных делителей

матрицы A каждый элементарный делитель, соответствующий

отрицательному характеристическому числу (если таковые существуют),

повторяется четное число раз. Докажем теперь, что это необходимое условие является и

достаточным, т. е. что вещественная неособенная матрица A тогда и только тогда имеет

вещественный логарифм X , когда у матрицы A либо совсем нет элементарных

делителей, соответствующих отрицательным характеристическим числам, либо каждый

такой элементарный делитель повторяется четное число раз.

Действительно, пусть это условие выполнено. Тогда в квазидиагональной матрице (95) в

соответствии с формулой (94) в тех клетках, где λ_i вещественно и положительно, возьмем

для $\ln \lambda_i$ вещественное значение; если же в какой-либо клетке имеется комплексное λ_n , то

найдется другая клетка такого же размера с $\lambda_{\bar{n}} = \overline{\lambda_n}$. В этих клетках возьмем комплексно

сопряженные значения для $\ln \lambda_k$ и $\ln \lambda_{\bar{k}}$. Каждая же клетка по условию повторяется в (98) четное число раз с сохранением размера клетки. Тогда в половине этих клеток положим $\ln \lambda_k = \ln |\lambda_k| + i\pi$, а в другой половине возьмем $\ln \lambda_k = \ln |\lambda_k| - i\pi$. Тогда в квазидиагональной матрице (98) диагональные клетки либо будут вещественными, либо будут попарно комплексно сопряженными. Но такая квазидиагональная матрица всегда подобна вещественной матрице. Поэтому существует такая неособенная

матрица $T_1 (|T_1| \neq 0)$, что матрица

$$X_1 = T_1 \left\{ \ln \left(\lambda_1 E^{(p_1)} + H^{(p_1)} \right), \dots, \ln \left(\lambda_n E^{(p_n)} + H^{(p_n)} \right) \right\} T_1^{-1}$$

вещественна. Но тогда будет вещественной и матрица

$$A_1 = e^{X_1} = T_1 \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_n E^{(p_n)} + H^{(p_n)} \right\} T_1^{-1}. \quad (100)$$

Сопоставляя формулу (100) с формулой (92), заключаем, что матрицы A и A_1 подобны между собой (поскольку они подобны одной и той же жордановой матрице). Но две подобные вещественные матрицы могут быть преобразованы друг в друга с помощью

некоторой неособенной вещественной матрицы $W (|W| \neq 0)$:

$$A = W A_1 W^{-1} = W e^{X_1} W^{-1} = e^{W X_1 W^{-1}}.$$

Тогда матрица $X = W X_1 W^{-1}$ и будет искомым вещественным логарифмом матрицы A .

1. 7 Лекция №7 (1 час).

Тема: «Сингулярные пучки матриц»

1.7.1 Вопросы лекции:

1. Регулярный пучок матриц
2. Формулы приведения
3. Каноническая форма сингулярного пучка.

1.7.2 Краткое содержание вопросов:

1. Регулярный пучок матриц

1. Рассмотрим частный случай, когда пучки $A + \lambda B$ и $A_1 + \lambda B_1$ состоят из квадратных матриц ($m = n$) и $|B| \neq 0$, $|B_1| \neq 0$. В этом случае, как было доказано в главе VI (стр. 140), два понятия «эквивалентность» и «строгая эквивалентность» пучков совпадают. Поэтому, применяя к пучкам общий критерий эквивалентности λ -матриц (стр. 148), приходим к теореме:

Теорема 1. Два пучка квадратных матриц одного и того же порядка $A + \lambda B$ и $A_1 + \lambda B_1$, у которых $|B| \neq 0$ и $|B_1| \neq 0$, являются строго эквивалентными в том и только в том случае, когда эти пучки имеют одни и те же элементарные делители в поле K .

Пучок квадратных матриц $A + \lambda B$ с $|B| \neq 0$ в главе VI называется регулярным, поскольку он представляет собой частный случай регулярного матричного многочлена относительно λ (см. гл. IV, стр. 87). В предыдущем параграфе настоящей главы мы дали более широкое определение регулярного пучка. Согласно этому определению в регулярном пучке возможно равенство $|B| = 0$ (и даже $|A| = |B| = 0$).

Для того чтобы выяснить, сохранится ли теорема 1 для регулярных пучков (при расширенном определении 1), рассмотрим следующий пример:

$$A + \lambda B = \begin{vmatrix} 2 & 1 & 3 \\ 3 & 2 & 5 \\ 3 & 2 & 6 \end{vmatrix} + \lambda \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix}, \quad A_1 + \lambda B_1 = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} + \lambda \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad (3)$$

Нетрудно видеть, что здесь каждый из пучков $A + \lambda B$ и $A_1 + \lambda B_1$ имеет только один элементарный делитель $\lambda + 1$. В то же время эти пучки не являются строго эквивалентными, так как матрицы B и B_1 имеют соответственно ранги 2 и 1, а из равенства (2), если бы оно имело место, следовало бы, что ранги матриц B и B_1 равны между собой. При этом пучки (3) являются регулярными согласно определению 1, так как $|A + \lambda B| \equiv |A_1 + \lambda B_1| \equiv \lambda + 1$.

Разобранный пример показывает, что теорема 1 неверна при расширенном определении регулярного пучка.

2. Для того чтобы сохранить теорему 1, нам придется ввести понятие о «бесконечных» элементарных делителях пучка. Будем пучок $A + \lambda B$ задавать при помощи «однородных» параметров λ, μ : $\mu A + \lambda B$.

Тогда определитель $\Delta(\lambda, \mu) \equiv |\mu A + \lambda B|$ будет однородной функцией от λ, μ .

Определяя наибольший общий делитель $D_k(\lambda, \mu)$ всех миноров k -го

порядка матрицы $\mu A + \lambda B$ ($k = 1, 2, \dots, n$), получим инвариантные многочлены по известным формулам

$$i_1(\lambda, \mu) = \frac{D_n(\lambda, \mu)}{D_{n-1}(\lambda, \mu)}, \quad i_2(\lambda, \mu) = \frac{D_{n-1}(\lambda, \mu)}{D_{n-2}(\lambda, \mu)}, \dots;$$

при этом все $D_k(\lambda, \mu)$ и $i_j(\lambda, \mu)$ - однородные относительно λ и μ многочлены.

Разлагая инвариантные многочлены на степени неприводимых в поле K однородных многочленов, получим элементарные делители $e_\alpha(\lambda, \mu)$ ($\alpha = 1, 2, \dots$) пучка $\mu A + \lambda B$ в поле K .

Совершенно очевидно, что, полагая $\mu = 1$ в $e_\alpha(\lambda, \mu)$, мы вернемся к элементарным

делителям $e_\alpha(\lambda)$ пучка $A + \lambda B$. Обратно, из каждого элементарного

делителя $e_\alpha(\lambda)$ степени q пучка $A + \lambda B$ мы получим соответствующий элементарный

делитель $e_\alpha(\lambda, \mu)$ по формуле $e_\alpha(\lambda, \mu) = \mu^q e_\alpha\left(\frac{\lambda}{\mu}\right)$. Таким способом могут быть

получены все элементарные делители пучка $\mu A + \lambda B$ за исключением элементарных делителей вида μ^q .

Элементарные делители вида μ^q существуют в том и только в том случае, когда $|B| = 0$, и носят название «бесконечных» элементарных делителей для пучка $A + \lambda B$.

Поскольку из строгой эквивалентности пучков $A + \lambda B$ и $A_1 + \lambda B_1$ следует строгая эквивалентность пучков $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ то у строго эквивалентных

пучков $A + \lambda B$ и $A_1 + \lambda B_1$ должны совпадать не только «конечные», но и «бесконечные» элементарные делители.

Пусть теперь даны два регулярных пучка $A + \lambda B$ и $A_1 + \lambda B_1$, у которых соответственно совпадают все (в том числе и бесконечные) элементарные делители. Введем однородные параметры: $\mu A + \lambda B$, $A_1 + \lambda B_1$. Преобразуем параметры:

$$\begin{aligned}\lambda &= \alpha_1 \tilde{\lambda} + \alpha_2 \tilde{\mu}, \\ \mu &= \beta_1 \tilde{\lambda} + \beta_2 \tilde{\mu} \quad (\alpha_1 \beta_2 - \alpha_2 \beta_1 \neq 0)\end{aligned}$$

В новых параметрах пучки запишутся

так: $\tilde{\mu} \tilde{A} + \tilde{\lambda} \tilde{B}$, $\tilde{\mu} \tilde{A}_1 + \tilde{\lambda} \tilde{B}_1$ где $\tilde{B} = \beta_1 A + \alpha_1 B$, $\tilde{B}_1 = \beta_1 A_1 + \alpha_1 B_1$. Из регулярности пучков $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ вытекает, что можно подобрать числа α_1 и β_1 так, чтобы $|\tilde{B}| \neq 0$ и $|\tilde{B}_1| \neq 0$.

Поэтому согласно теореме 1 пучки $\tilde{\mu} \tilde{A} + \tilde{\lambda} \tilde{B}$ и $\tilde{\mu} \tilde{A}_1 + \tilde{\lambda} \tilde{B}_1$, а следовательно, и исходные пучки $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ (или, что то же $A + \lambda B$ и $A_1 + \lambda B_1$) строго эквивалентны. Таким образом, мы пришли к следующему обобщению теоремы 1.

Теорема 2. Для того чтобы два регулярных пучка $A + \lambda B$ и $A_1 + \lambda B_1$ были строго эквивалентны, необходимо и достаточно, чтобы эти пучки имели одни и те же («конечные» и «бесконечные») элементарные делители.

В разобранный ранее примере пучки (3) имели один и тот же «конечный» элементарный делитель $\lambda + 1$, но отличались «бесконечными» элементарными делителями (первый пучок имеет один «бесконечный» элементарный делитель μ^2 , а второй - два: μ , μ). Поэтому эти пучки и не оказались строго эквивалентными.

3. Пусть теперь дан произвольный регулярный пучок $A + \lambda B$. Тогда существует такое число c , что $|A + cB| \neq 0$. Данный пучок представим в виде $A_1 + (\lambda - c)B$, где $A_1 = A + cB$ и потому $|A_1| \neq 0$. Умножим пучок слева

на $A_1^{-1}: E + (\lambda - c)A_1^{-1}B$. Преобразованием подобия приводим этот пучок к виду

$$E + (\lambda - c)\{J_0, J_1\} = \{E - cJ_0 + \lambda J_0, E - cJ_1 + \lambda J_1\}, \quad (4)$$

где $\{J_0, J_1\}$ - квазидиагональная нормальная форма матрицы $A_1^{-1}B$, J_0 - жорданова нильпотентная матрица, а $|J_1| \neq 0$.

Первый диагональный блок правой части (4) умножим на $(E - cJ_0)^{-1}$.

Получим: $E + \lambda(E - cJ_0)^{-1}J_0$. Здесь коэффициент при λ - нильпотентная матрица. Поэтому преобразованием подобия этот пучок можно привести к виду

$$\begin{aligned}E + \lambda \tilde{J}_0 &= \{N^{(u_1)}, N^{(u_2)}, \dots, N^{(u_s)}\} \\ (N^{(u)} &= E^{(u)} + \lambda H^{(u)})\end{aligned} \quad (5)$$

Второй диагональный блок в правой части (4) умножением на J_1^{-1} , а затем преобразованием подобия может быть приведен к виду $J + \lambda E$, где J - матрица, имеющая нормальную форму, а E - единичная матрица. Мы пришли к теореме

Теорема 3. Произвольный регулярный пучок $A + \lambda B$ может быть приведен к (строго эквивалентному) каноническому квазидиагональному виду

$$\left\{ N^{(u_1)}, N^{(u_2)}, \dots, N^{(u_s)}, J + \lambda E \right\}$$

$$\left(N^{(u)} = E^{(u)} + \lambda H^{(u)} \right), \quad (6)$$

где первые s диагональных блоков соответствуют бесконечным элементарным делителям $\mu^{u_1}, \mu^{u_2}, \dots, \mu^{u_s}$ пучка $A + \lambda B$, а нормальная форма последнего диагонального блока $J + \lambda E$ однозначно определяется конечными элементарными делителями данного пучка.

2. Формулы приведения

Переходим к рассмотрению сингулярного пучка матриц $A + \lambda B$ с размерами $m \times n$. Обозначим через r ранг пучка, т. е. наибольший из порядков миноров, не равных тождественно нулю. Из сингулярности пучка следует, что всегда имеет место по крайней мере одно из неравенств $r < n$ или $r < m$. Пусть $r < n$. Тогда столбцы λ -матрицы $A + \lambda B$ линейно зависимы, т. е. уравнение

$$(A + \lambda B)x = 0, \quad (7)$$

где x - искомый столбец, имеет ненулевое решение. Каждое ненулевое решение этого уравнения определяет некоторую линейную зависимость между столбцами λ -

матрицы $A + \lambda B$. Мы ограничимся только теми решениями $x(\lambda)$ уравнения (7), которые являются многочленами относительно λ , и среди этих решений возьмем решение наименьшей степени ε

$$x(\lambda) = x_0 - \lambda x_1 + \lambda^2 x_2 - \dots + (-1)^\varepsilon \lambda^\varepsilon x_\varepsilon \quad (x_\varepsilon \neq 0) \quad (8)$$

Подставляя это решение в (7) и приравнявая нулю коэффициенты при степенях λ , получим:

$$Ax_0 = 0, Bx_0 - Ax_1 = 0, Bx_1 - Ax_2 = 0, \dots, Bx_{\varepsilon-1} - Ax_\varepsilon = 0, Bx_\varepsilon = 0 \quad (9)$$

Рассматривая эту систему равенств как систему линейных однородных

уравнений относительно элементов столбцов $x_0, -x_1, +x_2, \dots, (-1)^\varepsilon x_\varepsilon$, заключаем, что матрица коэффициентов этой системы

$$M_\varepsilon = M_\varepsilon[A + \lambda B] = \left\| \begin{array}{c|ccc} & \overbrace{\varepsilon+1} & & \\ A & 0 & \dots & 0 \\ B & A & & . \\ 0 & B & \ddots & . \\ \vdots & \vdots & \ddots & A \\ 0 & 0 & \dots & B \end{array} \right\| \quad (10)$$

имеет ранг $\rho_\varepsilon < (\varepsilon+1)n$. В то же время в силу минимального свойства числа ε для рангов $\rho_1, \dots, \rho_{\varepsilon-1}$ матриц

$$M_0 = \begin{bmatrix} A \\ B \end{bmatrix}, M_1 = \begin{bmatrix} A & 0 \\ B & A \\ 0 & B \end{bmatrix}, \dots, M_{\varepsilon-1} = \begin{bmatrix} \overbrace{A & 0 & \dots & 0}^{\varepsilon} \\ B & A & & \\ 0 & B & \ddots & \\ \vdots & \vdots & \ddots & A \\ 0 & 0 & \dots & B \end{bmatrix} \quad (10')$$

имеют место равенства $\rho_0 = n, \rho_1 = 2n, \dots, \rho_{\varepsilon-1} = \varepsilon n$.

Таким образом, число ε есть наименьшее значение индекса k , при котором в соотношении $\rho_k \leq (k+1)n$ имеет место знак $<$.

Теперь мы сформулируем и докажем следующую фундаментальную теорему:

Теорема 4: Если уравнение (7) имеет решение минимальной степени ε и $\varepsilon > 0$, то данный пучок $A + \lambda B$ строго эквивалентен пучку вида

$$\begin{bmatrix} L_{\varepsilon} & 0 \\ 0 & \hat{A} + \lambda \hat{B} \end{bmatrix}, \quad (11)$$

где

$$L_{\varepsilon} = \left\{ \begin{bmatrix} \overbrace{\lambda & 1 & 0 & \dots & 0 & 0}^{\varepsilon+1} \\ 0 & \lambda & 1 & & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \\ 0 & 0 & & \dots & \lambda & 1 \end{bmatrix} \right\}_{\varepsilon_1} \quad (12)$$

а $\hat{A} + \lambda \hat{B}$ - пучок матриц, для которого уравнение, аналогичное (7), имеет решений степени $< \varepsilon$.

Доказательство теоремы разобьем на три этапа. Сначала докажем, что данный пучок $A + \lambda B$ строго эквивалентен пучку вида

$$\begin{bmatrix} L_{\varepsilon} & D + \lambda F \\ 0 & \hat{A} + \lambda \hat{B} \end{bmatrix}, \quad (13)$$

где $D, F, \hat{A}, \hat{B}$ - постоянные прямоугольные матрицы соответственных размеров. Затем установим, что уравнение $(\hat{A} + \lambda \hat{B})\hat{x} = 0$ не имеет решений $\hat{x}(\lambda)$ степени $< \varepsilon$. После этого мы покажем, что дальнейшими преобразованиями пучок (13) может быть приведен к квазидиагональному виду (11).

1. Первую часть доказательства облечем в геометрическую форму. Вместо пучка матриц $A + \lambda B$ рассмотрим пучок операторов $\mathbf{A} + \lambda \mathbf{B}$ отображающих $\mathbf{R}_n$ в $\mathbf{R}_m$, и покажем, что при надлежащем выборе базисов в этих пространствах матрица, соответствующая оператору $\mathbf{A} + \lambda \mathbf{B}$, будет иметь форму (13).

Вместо, уравнения (7) возьмем векторное уравнение

$$(\mathbf{A} + \lambda \mathbf{B})\mathbf{x} = 0 \quad (14)$$

с векторным решением

$$\mathbf{x}(\lambda) = \mathbf{x}_0 - \lambda \mathbf{x}_1 + \lambda^2 \mathbf{x}_2 - \dots + (-1)^{\varepsilon} \lambda^{\varepsilon} \mathbf{x}_{\varepsilon} \quad (15)$$

равенства (9) заменяются векторными равенствами

$$\mathbf{Ax}_0 = 0, \mathbf{Ax}_1 = \mathbf{Bx}_0, \mathbf{Ax}_2 = \mathbf{Bx}_1, \dots, \mathbf{Ax}_\varepsilon = \mathbf{Bx}_{\varepsilon-1}, \mathbf{Bx}_\varepsilon = 0 \quad (16)$$

Ниже мы докажем, что векторы

$$\mathbf{Ax}_1, \mathbf{Ax}_2, \dots, \mathbf{Ax}_\varepsilon \quad (17)$$

линейно независимы. Отсюда легко будет следовать линейная независимость векторов

$$\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_\varepsilon \quad (18)$$

Действительно, поскольку $\mathbf{Ax}_0 = 0$,

из $\alpha_0 x_0 + \alpha_1 x_1 + \dots + \alpha_\varepsilon x_\varepsilon = 0$ находим: $\alpha_1 \mathbf{Ax}_1 + \dots + \alpha_\varepsilon \mathbf{Ax}_\varepsilon = 0$, откуда в

силу линейной независимости векторов (17) $\alpha_1 = \alpha_2 = \dots = \alpha_\varepsilon = 0$. Но $\mathbf{x}_0 \neq 0$, поскольку в

противном случае $\frac{1}{\lambda} \mathbf{x}(\lambda)$ было бы решением уравнения (14) степени $\varepsilon-1$, что

невозможно. Поэтому и $\alpha_0 = 0$.

Если теперь принять векторы (17) и (18) в качестве первых базисных векторов для новых

базисов соответственно в $\mathbf{R}^n$ и $\mathbf{R}^m$, то в новых базисах операторам $\mathbf{A}$ и $\mathbf{B}$ в силу (16)

будут соответствовать матрицы

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & 1 & \dots & 0_{\varepsilon+1} & * & \dots & * \\ 0 & 0 & 1 & \dots & 0_{\varepsilon+1} & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1_{\varepsilon+1} & * & \dots & * \\ 0 & 0 & \dots & 0_{\varepsilon+1} & * & \dots & * \\ 0 & 0 & \dots & 0_{\varepsilon+1} & * & \dots & * \end{pmatrix}, \quad \tilde{\mathbf{B}} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0_{\varepsilon+1} & * & \dots & * \\ 0 & 1 & \dots & 0 & 0_{\varepsilon+1} & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0_{\varepsilon+1} & * & \dots & * \\ 0 & 0 & \dots & 0 & 0_{\varepsilon+1} & * & \dots & * \\ 0 & 0 & \dots & 0 & 0_{\varepsilon+1} & * & \dots & * \end{pmatrix},$$

тогда λ -матрица $\tilde{\mathbf{A}} + \lambda \tilde{\mathbf{B}}$ будет иметь вид (13). Все предыдущие рассуждения будут обоснованными, если мы докажем, что векторы (17) линейно независимы. Допустим

противное, и пусть $\mathbf{Ax}_h$ ($h \geq 1$) - первый в ряду (17) вектор, линейно зависящий от предыдущих векторов

$$\mathbf{Ax}_h = \alpha_1 \mathbf{Ax}_{h-1} + \alpha_2 \mathbf{Ax}_{h-2} + \dots + \alpha_{h-1} \mathbf{Ax}_1$$

В силу (16) это равенство может быть переписано так:

$$\mathbf{Bx}_{h-1} = \alpha_1 \mathbf{Bx}_{h-2} + \alpha_2 \mathbf{Bx}_{h-3} + \dots + \alpha_{h-1} \mathbf{Bx}_0$$

т. е.

$$\mathbf{Bx}_{h-1}^* = 0,$$

где

$$\mathbf{x}_{h-1}^* = \mathbf{x}_{h-1} - \alpha_1 \mathbf{x}_{h-2} - \alpha_2 \mathbf{x}_{h-3} \dots - \alpha_{h-1} \mathbf{x}_0$$

Далее, опять в силу (16)

$$\mathbf{Ax}_{h-1}^* = \mathbf{B}(\mathbf{x}_{h-2} - \alpha_1 \mathbf{x}_{h-3} \dots - \alpha_{h-1} \mathbf{x}_0) = \mathbf{Bx}_{h-2}^*,$$

где

$$\mathbf{x}_{h-2}^* = \mathbf{x}_{h-2} - \alpha_1 \mathbf{x}_{h-3} \dots - \alpha_{h-2} \mathbf{x}_0$$

Продолжая этот процесс далее и вводя еще векторы

$$\mathbf{x}_{h-3}^* = \mathbf{x}_{h-3} - \alpha_1 \mathbf{x}_{h-4} - \dots - \alpha_{h-3} \mathbf{x}_0, \dots, \mathbf{x}_1^* = \mathbf{x}_1 - \alpha_1 \mathbf{x}_0, \mathbf{x}_0^* = \mathbf{x}_0,$$

мы получим цепочку равенств

$$\mathbf{Bx}_{h-1}^* = 0, \mathbf{Ax}_{h-1}^* = \mathbf{Bx}_{h-2}^*, \dots, \mathbf{Ax}_1^* = \mathbf{Bx}_0^*, \mathbf{Ax}_0^* = 0 \quad (19)$$

Из (19) следует, что

$$x^*(\lambda) = x_0^* - \lambda x_1^* + \dots + (-1)^{k-1} x_{k-1}^* \quad (x_0^* = x_0 \neq 0)$$

есть ненулевое решение уравнения (14) степени $\leq k-1 < \varepsilon$, что невозможно. Таким образом, векторы (17) линейно независимы.

2. Докажем теперь, что уравнение $(\hat{A} + \lambda \hat{B})\hat{x} = 0$ не имеет решений степени $< \varepsilon$. Сначала обратим внимание на то, что уравнение $L_\varepsilon y = 0$, как и уравнение (7), имеет ненулевое решение наименьшей степени ε . В этом можно убедиться непосредственно, если матричное уравнение $L_\varepsilon y = 0$ заменить системой обыкновенных уравнений

$$\lambda y_1 + y_2 = 0, \lambda y_2 + y_3 = 0, \dots, \lambda y_\varepsilon + y_{\varepsilon+1} = 0,$$

$$[y = (y_1, y_2, \dots, y_{\varepsilon+1})], \text{ откуда } y_k = (-1)^{k-1} y_1 \lambda^{k-1} \quad (k=1, 2, \dots, \varepsilon+1)$$

С другой стороны, если пучок имеет «треугольный» вид (13), то соответствующие этому пучку матрицы M_k ($k=1, 2, \dots, \varepsilon$) [см. (10) и (10') на стр. 335] после надлежащей перестановки строк и столбцов также могут быть приведены к треугольному виду

$$\begin{vmatrix} M_k[L_\varepsilon] & M_k[D + \lambda F] \\ 0 & M_k[\hat{A} + \lambda \hat{B}] \end{vmatrix} \quad (20)$$

При $k = \varepsilon - 1$ все столбцы этой матрицы, а значит, и столбцы матрицы $M_{\varepsilon-1}[L_\varepsilon]$, линейно независимы. Но $M_{\varepsilon-1}[L_\varepsilon]$ - квадратная матрица порядка $\varepsilon(\varepsilon+1)$. Поэтому и в матрице $M_{\varepsilon-1}[\hat{A} + \lambda \hat{B}]$ все столбцы линейно независимы, а это, как было выяснено в начале параграфа, означает, что уравнение $(\hat{A} + \lambda \hat{B})\hat{x} = 0$ не имеет решений степени $\leq \varepsilon - 1$, что и требовалось доказать.

3. Заменим пучок (13) строго эквивалентным ему пучком

$$\left\| \begin{array}{cc|cc} E_1 & Y & L_\varepsilon & D + \lambda F \\ 0 & E_2 & 0 & \hat{A} + \lambda \hat{B} \end{array} \right\| \begin{array}{c} E_3 \\ 0 \\ E_4 \end{array} - X = \left\| \begin{array}{cc|c} L_\varepsilon & D + \lambda F + Y(\hat{A} + \lambda \hat{B}) - L_\varepsilon X \\ 0 & \hat{A} + \lambda \hat{B} \end{array} \right\|, \quad (21)$$

где E_1, E_2, E_3, E_4 - квадратные единичные матрицы соответственно порядков $\varepsilon, m - \varepsilon, \varepsilon + 1$ и $n - \varepsilon - 1$, а X, Y - произвольные постоянные прямоугольные матрицы соответствующих размеров. Наша теорема будет полностью доказана, если мы покажем, что матрицы X и Y могут быть выбраны так, чтобы имело место матричное равенство

$$L_\varepsilon X = D + \lambda F + Y(\hat{A} + \lambda \hat{B}) \quad (22)$$

Введем обозначения для элементов матриц D, F, X а также для строк матрицы Y и для столбцов матриц $\hat{A}, \hat{B}$:

$$D = \|d_{ik}\|, \quad F = \|f_{ik}\|, \quad X = \|x_{jk}\| \quad (i=1, 2, \dots, \varepsilon; \quad k=1, 2, \dots, n - \varepsilon - 1; \quad j=1, 2, \dots, \varepsilon + 1),$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_\varepsilon \end{pmatrix}, \quad \hat{A} = (a_1, a_2, \dots, a_{n-\varepsilon-1}), \quad \hat{B} = (b_1, b_2, \dots, b_{n-\varepsilon-1})$$

Тогда матричное уравнение (22) можно заменить системой скалярных уравнений, записывая, что элементы k -го столбца в левой и правой частях равенства (22)

соответственно равны друг другу ($k=1, 2, \dots, n-\varepsilon-1$):

$$\left. \begin{aligned} x_{2k} + \lambda x_{1k} &= d_{1k} + \lambda f_{1k} + y_1 a_k + \lambda y_1 b_k \\ x_{3k} + \lambda x_{2k} &= d_{2k} + \lambda f_{2k} + y_2 a_k + \lambda y_2 b_k \\ x_{2k} + \lambda x_{3k} &= d_{3k} + \lambda f_{3k} + y_3 a_k + \lambda y_3 b_k \\ \hline x_{\varepsilon+1,k} + \lambda x_{\varepsilon k} &= d_{\varepsilon k} + \lambda f_{\varepsilon k} + y_{\varepsilon} a_k + \lambda y_{\varepsilon} b_k \\ &(k=1, 2, \dots, n-\varepsilon-1) \end{aligned} \right\} \quad (23)$$

В левых частях этих равенств стоят линейные двучлены относительно λ . Свободный член каждого из первых $\varepsilon - 1$ этих двучленов равен коэффициенту при λ в следующем двучлене. Но тогда и правые части должны удовлетворять этому условию. Поэтому

$$\begin{aligned} y_1 a_k - y_2 b_k &= f_{2k} - d_{1k} \\ y_2 a_k - y_3 b_k &= f_{3k} - d_{2k} \\ \vdots & \\ y_{\varepsilon-1} a_k - y_{\varepsilon} b_k &= f_{\varepsilon k} - d_{\varepsilon-1,k} \\ (k=1, 2, \dots, n-\varepsilon-1) \end{aligned} \quad (24)$$

Если равенства (24) имеют место, то, очевидно, из (23) можно определить искомые элементы матрицы X .

Теперь осталось показать, что система уравнений (24) относительно элементов матрицы

Y всегда имеет решение при любых d_{ik} и f_{ik} ($i=1, 2, \dots, \varepsilon$; $k=1, 2, \dots, n-\varepsilon-1$).

Действительно, матрица, составленная из коэффициентов при неизвестных элементах

строк $y_1, -y_2, +y_3, -y_4, \dots$, может быть записана после транспонирования в виде

$$\left| \begin{array}{cccccc} \hat{A} & 0 & \overbrace{\quad \quad \quad}^{\varepsilon-1} & & & 0 \\ \hat{B} & \hat{A} & & & & . \\ 0 & \hat{B} & . & & & . \\ . & & . & . & . & . \\ . & & . & . & . & . \\ . & & & & & \hat{A} \\ 0 & 0 & . & . & . & \hat{B} \end{array} \right|$$

Но эта матрица является матрицей $M_{\varepsilon-2}$ для пучка прямоугольных матриц $\hat{A} + \lambda \hat{B}$ [см.

(10') на стр. 335]. Ранг же этой матрицы равен $(\varepsilon-1)(n-\varepsilon-1)$, поскольку по доказанному

уравнение $(\hat{A} + \lambda \hat{B})\hat{x} = 0$ не имеет решений степени $< \varepsilon$. Таким образом, ранг системы

уравнений (24) равен числу уравнений, а такая система при любых свободных членах является совместной (непротиворечивой).

Теорема доказана полностью.

3. Каноническая форма сингулярного пучка.

Пусть дан произвольный сингулярный пучок матриц $A + \lambda B$ размеров $m \times n$. Допустим сначала, что как между столбцами, так и между строками этого пучка нет линейной зависимости с постоянными коэффициентами.

Пусть $r < n$, где r — ранг пучка, т. е. столбцы пучка $A + \lambda B$ линейно зависимы между собой. В этом случае уравнение $(A + \lambda B)x = 0$ имеет ненулевое решение минимальной степени ε_1 . Из принятого в начале этого параграфа ограничения следует, что $\varepsilon_1 > 0$. Поэтому согласно теореме 4 данный пучок можно преобразовать к виду

$$\begin{pmatrix} L_{\varepsilon_1} & 0 \\ 0 & A_1 + \lambda B_1 \end{pmatrix},$$

где уравнение $(A_1 + \lambda B_1)x^{(1)} = 0$ не имеет решений $x^{(1)}$ степени $< \varepsilon_1$.

Если это уравнение имеет ненулевое решение минимальной степени ε_2 (при этом непременно $\varepsilon_2 \geq \varepsilon_1$), то, применяя к пучку $A_1 + \lambda B_1$ теорему 4, мы данный пучок преобразуем к виду

$$\begin{pmatrix} I_{\varepsilon_1} & 0 & 0 \\ 0 & L_{\varepsilon_2} & 0 \\ 0 & 0 & A_2 + \lambda B_2 \end{pmatrix}.$$

Продолжая этот процесс далее, мы приведем данный пучок к квазидиагональному виду

$$\left\| \begin{array}{ccccccc} L_{\varepsilon_1} & & & & & & 0 \\ & L_{\varepsilon_2} & & & & & \\ & & \ddots & & & & \\ & & & L_{\varepsilon_p} & & & \\ & & & & A_p + \lambda B_p & & \\ 0 & & & & & & \end{array} \right\|, \quad (25)$$

где $0 < \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_p$, а уравнение $(A_p + \lambda B_p)x^{(p)} = 0$ не имеет ненулевых решений, т. е. столбцы матрицы $A_p + \lambda B_p$ линейно независимы.

Если строки пучка $A_p + \lambda B_p$ линейно зависимы, то транспонированный пучок $A_p' + \lambda B_p'$ может быть приведен к виду (25), где вместо чисел $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ будут фигурировать числа $0 < \eta_1 \leq \eta_2 \leq \dots \leq \eta_q$. Но тогда данный пучок $A + \lambda B$ окажется преобразованным к квазидиагональному виду

$$\begin{array}{c}
\begin{array}{c} L_{\varepsilon_1} \\ L_{\varepsilon_2} \\ \vdots \\ L_{\varepsilon_p} \\ L'_{\eta_1} \\ L'_{\eta_2} \\ \vdots \\ L'_{\eta_q} \end{array} \quad \begin{array}{c} 0 \\ \\ \\ \\ \\ \\ A_0 + \lambda B_0 \end{array} \\
\hline
\begin{array}{c} 0 \\ \\ \\ \\ \\ \\ \end{array}
\end{array} \quad (26)$$

$$(0 < \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_p, 0 < \eta_1 \leq \eta_2 \leq \dots \leq \eta_q),$$

где у пучка $A_0 + \lambda B_0$ как столбцы, так и строки линейно независимы, т. е. $A_0 + \lambda B_0$ - регулярный пучок.

Рассмотрим теперь общий случай, когда строки и столбцы данного пучка могут быть связаны линейными зависимостями с постоянными коэффициентами. Обозначим максимальное число постоянных независимых решений уравнений

$$(A + \lambda B)x = 0 \quad \text{и} \quad (A' + \lambda B')y' = 0$$

соответственно через g и h . Вместо первого из этих уравнений, подобно тому как мы это делали при доказательстве теоремы 4, рассмотрим соответствующее векторное

уравнение $(A + \lambda B)x = 0$ (A и B - операторы, отображающие $\mathbf{R}_n$ в $\mathbf{R}_m$). Линейно

независимые постоянные решения этого уравнения обозначим через $e_1, e_2, \dots, e_g$ и примем

за первые базисные векторы в $\mathbf{R}_n$. Тогда в соответствующей

матрице $\tilde{A} + \lambda \tilde{B}$ первые g столбцов будут состоять из нулей

$$\tilde{A} + \lambda \tilde{B} = \begin{pmatrix} \overset{g}{0} \\ \tilde{A}_1 + \lambda \tilde{B}_1 \end{pmatrix}. \quad (27)$$

Совершенно так же в пучке $\tilde{A}_1 + \lambda \tilde{B}_1$ первые h строк можно сделать нулевыми. Тогда данный пучок примет вид

$$\begin{array}{c} \overset{g}{h} \\ \left[\begin{array}{cc} 0 & 0 \\ 0, & A^0 + \lambda B^0 \end{array} \right] \end{array} \quad (28)$$

где строки и столбцы пучка $A^0 + \lambda B^0$ уже не связаны линейными зависимостями с постоянными коэффициентами. К пучку $A^0 + \lambda B^0$ применимо представление типа (26). Таким образом, в самом общем случае пучок $A + \lambda B$ всегда может быть приведен к каноническому квазидиагональному виду

$$\{h \begin{bmatrix} \varepsilon \\ 0 \end{bmatrix}; L_{\varepsilon_{g+1}}, \dots, L_{\varepsilon_{kp}}; L'_{\eta_{h+1}}, \dots, L'_{\eta_g}; A_0 + \lambda B_0\}. \quad (29)$$

Выбор индексов при ε и η связан с тем, что нам удобно здесь считать $\varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_g = 0$ и $\eta_1 = \eta_2 = \dots = \eta_h = 0$.

Заменяя фигурирующий в (29) регулярный пучок $A_0 + \lambda B_0$ его канонической формой (6) (см. § 2, стр. 334), получим окончательно следующую квазидиагональную матрицу

$$\{h \begin{bmatrix} \varepsilon \\ 0 \end{bmatrix}; L_{\varepsilon_{g+1}}, \dots, L_{\varepsilon_{kp}}; L'_{\eta_{h+1}}, \dots, L'_{\eta_g}; N^{(u_1)}, \dots, N^{(u_3)}; J + \lambda E\}, \quad (30)$$

где матрица J имеет жорданову или естественную нормальную форму,

$$N^{(u)} = E^{(u)} + \lambda H^{(u)}.$$

Матрица (30) представляет собой каноническую форму пучка $A + \lambda B$ в самом общем случае.

Для того чтобы по данному пучку непосредственно определить его каноническую форму (30), не осуществляя последовательно процесс приведения, мы, следуя Кронекеру, в следующем параграфе введем понятие о минимальных индексах пучка.

1. 8 Лекция №8 (1 час).

Тема: «Приложение теории матриц к исследованию систем линейных дифференциальных уравнений»

1.8.1 Вопросы лекции:

1. Основные понятия.
2. Преобразования Ляпунова.
3. Приводимые системы.

1.8.2 Краткое содержание вопросов: (тезисно изложить основное содержание рассматриваемых вопросов)

1. Основные понятия.

Пусть дана система линейных однородных дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt} = \sum_{k=1}^n p_{ik}(t) x_k \quad (i = 1, 2, \dots, n) \quad (1)$$

где $p_{ik}(t)$ ($i, k = 1, 2, \dots, n$) - комплексные функции вещественного аргумента t , непрерывные в некотором (конечном или бесконечном) интервале изменения t .

Полагая $P(t) = \|p_{ik}(t)\|_1^n$ и $x = (x_1, x_2, \dots, x_n)$, мы систему (1) запишем так:

$$\frac{dx}{dt} = P(t) x. \quad (2)$$

Интегральной матрицей системы (4) мы будем

называть квадратную матрицу $X(t) = \|x_{ik}(t)\|_1^n$, столбцами которой являются n линейно независимых решений системы.

Так как каждый столбец матрицы X удовлетворяет уравнению (2), то и интегральная матрица X удовлетворяет уравнению

$$\frac{dX}{dt} = P(t)X. \quad (3)$$

В дальнейшем мы вместо системы (1) будем рассматривать Матричное уравнение (3). Из теоремы о существовании и единственности решения системы линейных дифференциальных уравнений следует, что

интегральная матрица $X(t)$ однозначно определяется, если задано значение этой матрицы при некотором («начальном») значении $t = t_0$, $X(t_0) = X_0$. В качестве матрицы X_0 можно взять любую неособенную квадратную матрицу n -го порядка. В частном случае, когда $X(t_0) = E$, интегральную матрицу $X(t)$ будем называть нормированной.

Продифференцируем определитель матрицы X , дифференцируя последовательно строки определителя и используя при этом дифференциальные соотношения

$$\frac{dx_{ij}}{dt} = \sum_{k=1}^n p_{ik}x_{kj} \quad (i, j = 1, 2, \dots, n).$$

Тогда получим:

$$\frac{d|X|}{dt} = (p_{11} + p_{22} + \dots + p_{nn})|X|.$$

Отсюда следует известное тождество Якоби

$$|X| = ce^{\int_{t_0}^t \text{Sp } P dt} \quad (4)$$

где c - постоянная, а

$$\text{Sp } P = p_{11} + p_{22} + \dots + p_{nn}$$

- след матрицы $P(t)$.

Так как определитель $|X|$ не может тождественно равняться нулю, то $c \neq 0$. Но тогда из тождества Якоби следует, что определитель $|X|$ при любом значении аргумента отличен от нуля

$$|X| \neq 0,$$

т. е. интегральная матрица при любом значении аргумента является неособенной.

Если $\tilde{X}(t)$ - неособенное ($|\tilde{X}| \neq 0$) частное решение уравнения (3), то общее решение этого уравнения определяется формулой

$$X = \tilde{X}C, \quad (5)$$

где C - произвольная постоянная матрица.

Действительно, умножая обе части равенства

$$\frac{d\tilde{X}}{dt} = P\tilde{X} \quad (6)$$

справа на C , убеждаемся, что и матрица $\tilde{X}C$ удовлетворяет уравнению (3). С другой стороны, если X - произвольное решение уравнения (3), то из (6) следует:

$$\frac{dX}{dt} = \frac{d}{dt}(\tilde{X} \cdot \tilde{X}^{-1}X) = \frac{d\tilde{X}}{dt}\tilde{X}^{-1}X + \tilde{X} \frac{d}{dt}(\tilde{X}^{-1}X) = PX + \tilde{X} \frac{d}{dt}(\tilde{X}^{-1}X),$$

откуда в силу (3)

$$\frac{d}{dt}(\tilde{X}^{-1}X) = 0$$

и

$$\tilde{X}^{-1}X = \text{const} = C,$$

т. е. имеет место (5).

Все интегральные матрицы X системы (1) получаются по формуле (5) при $|C| \neq 0$. Рассмотрим частный случай:

$$\frac{dX}{dt} = AX, \quad (7)$$

где A - постоянная матрица. При этом $\tilde{X} = e^{At}$ есть частное неособенное решение уравнения (7) и потому общее решение этого уравнения имеет вид

$$X = e^{At}C, \quad (8)$$

где C - произвольная постоянная матрица.

Полагая в (8) $t = t_0$ найдем: $X_0 = e^{At_0}C$. Отсюда $C = e^{-At_0}X_0$ и потому формулу (8) можно представить в виде

$$X = e^{A(t-t_0)}X_0 \quad (9)$$

Эта формула эквивалентна выведенной ранее формуле (46) главы V (стр. 125).

Рассмотрим еще так называемую систему Коши:

$$\frac{dX}{dt} = \frac{A}{t-a}X \quad (A - \text{постоянная матрица}) \quad (10)$$

Этот случай сводится к предыдущему заменой аргумента:

$$u = \ln(t-a).$$

Поэтому общее решение системы (10) выглядит так:

$$X = e^{A \ln(t-a)}C = (t-a)^A C \quad (11)$$

Функции e^{At} и $(t-a)^A$, встречающиеся в формулах (8) и (11), могут быть представлены в виде (стр. 125)

$$e^{At} = \sum_{k=1}^s \left(Z_{k1} + Z_{k2}t + \dots + Z_{km_k}t^{m_k-1} \right) e^{\lambda_k t} \quad (12)$$

$$(t-a)^A = \sum_{k=1}^s \left(Z_{k1} + Z_{k2} \ln(t-a) + \dots + Z_{km_k} [\ln(t-a)]^{m_k-1} \right) (t-a)^{\lambda_k} \quad (13)$$

Здесь

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

$$(\lambda_i \neq \lambda_k \text{ при } i \neq k; i, k = 1, 2, \dots, s)$$

- минимальный многочлен матрицы A , а Z_{ik} ($j = 1, 2, \dots, m_k; k = 1, 2, \dots, s$) - линейно независимые постоянные матрицы, являющиеся многочленами от A .

Замечание. Иногда в качестве интегральной матрицы

системы дифференциальных уравнений (1) берут матрицу W , у которой строки являются линейно независимыми решениями системы. Очевидно, матрица W будет транспонированной матрицей для X :

$$W = X'.$$

Переходя в обеих частях равенства (3) к транспонированным матрицам, мы вместо (3) получим следующее уравнение для W :

$$\frac{dW}{dt} = W P(t) \quad (3')$$

В правой части этого уравнения матрица W стоит первым множителем, а не вторым, как X в уравнении (3).

2. Преобразования Ляпунова.

Допустим теперь, что в системе (1) [и в уравнении

(3)] матрица коэффициентов $P(t) = \|p_{ik}(t)\|_1^n$ - непрерывнаяограниченная функция от t в интервале $[t_0, \infty)$.

Введем вместо неизвестных функций $x_1, x_2, \dots, x_n$ новые неизвестные функции $y_1, y_2, \dots, y_n$ при помощи преобразования

$$x_i = \sum_{k=1}^n l_{ik}(t) y_k \quad (i=1, 2, \dots, n) \quad (14)$$

На матрицу преобразования $L(t) = \|l_{ik}(t)\|_1^n$ наложим следующие ограничения:

1° $L(t)$ имеет непрерывную производную $\frac{dL}{dt}$ в интервале $[t_0, \infty)$;

2° $L(t)$ и $\frac{dL}{dt}$ ограничены в интервале $[t_0, \infty)$;

3° существует постоянная m такая, что

$$0 < m < \inf_{t \geq t_0} |L(t)|,$$

т. е. определитель $|L(t)|$ ограничен по модулю снизу положительной постоянной m .

Преобразование (14), в котором матрица коэффициентов $L(t) = \|l_{ik}(t)\|_1^n$ удовлетворяет условиям 1°-3°, мы будем называть преобразованием Ляпунова, а соответствующую матрицу $L(t)$ - матрицей Ляпунова.

Такие преобразования рассматривал А. М. Ляпунов в своем знаменитом мемуаре «Общая задача об устойчивости движения» [19].

Примеры. 1. Если $L = \text{const}$ и $|L| \neq 0$, то матрица L удовлетворяет условиям 1°-3°.

Следовательно, неособенное преобразование с постоянными коэффициентами всегда является преобразованием Ляпунова.

2. Если $D = \|d_{ik}\|_1^n$ - матрица простой структуры с чисто мнимыми характеристическими числами, то матрица

$$L(t) = e^{Dt}$$

удовлетворяет условиям 1°-3° и потому является матрицей Ляпунова.

Легко проверить, что из свойств 1°-3° матрицы $L(t)$ следует, что

существует обратная матрица $L^{-1}(t)$ и что она удовлетворяет тем же условиям 1°-3°, т. е. обратное преобразование для преобразования Ляпунова снова является преобразованием Ляпунова. Точно так же проверяется, что два последовательных

преобразования Ляпунова в результате снова дают преобразование Ляпунова. Таким образом, преобразования Ляпунова образуют группу. Преобразования Ляпунова обладают следующим важным свойством:

Если при преобразовании (14) система уравнений (1) переходит в систему

$$\frac{dy_i}{dt} = \sum_{k=1}^n q_{ik}(t)y_k, \quad (15)$$

нулевое решение которой является устойчивым, асимптотически устойчивым или неустойчивым по Ляпунову (см. гл. V, § 7), то таким же свойством обладает и нулевое решение исходной системы (1).

Другими словами, преобразования Ляпунова не изменяют характеристики нулевого решения (в отношении устойчивости). Поэтому эти преобразования могут быть использованы при исследовании устойчивости для упрощения исходной системы уравнений.

Преобразование Ляпунова устанавливает одно-однозначное соответствие между решениями систем (1) и (15), при этом линейно независимые решения остаются таковыми и после преобразования. Поэтому преобразование Ляпунова переводит интегральную матрицу X системы (1) в некоторую интегральную матрицу Y системы (15), при этом

$$X = L(t)Y \quad (16)$$

В матричной записи система (15) имеет вид

$$\frac{dY}{dt} = Q(t)Y \quad (17)$$

где $Q(t) = \|q_{ik}(t)\|$ - матрица коэффициентов системы (15).

Подставляя в (3) вместо X произведение LY и сопоставляя полученное уравнение с (17), легко найдем следующую формулу, выражающую матрицу Q через матрицы P и L :

$$Q = L^{-1}PL - L^{-1}\frac{dL}{dt} \quad (18)$$

Две системы (1) и (15) или, что то же, (3) и (17) мы будем называть эквивалентными (в смысле Ляпунова), если они переводятся друг в друга преобразованием Ляпунова.

Матрицы коэффициентов P и Q эквивалентных систем всегда связаны между собой формулой (18), в которой матрица L удовлетворяет условиям 1°-3°.

3. Приводимые системы.

Среди систем линейных дифференциальных уравнений первого порядка наиболее простыми и наиболее изученными являются системы с постоянными коэффициентами. Поэтому представляют интерес системы, которые при помощи преобразования Ляпунова могут быть приведены к системам с постоянными коэффициентами. Такие системы А. М. Ляпунов называл приводимыми.

Пусть дана приводимая система

$$\frac{dX}{dt} = PX \quad (19)$$

Тогда некоторое преобразование Ляпунова

$$X = L(t)Y \quad (20)$$

переводит ее в систему

$$\frac{dY}{dt} = AY, \quad (21)$$

где A - постоянная матрица. Поэтому система (19) имеет частное решение

$$\tilde{X} = L(t) e^{At} \quad (22)$$

Легко видеть, что, и обратно, всякая система (19), имеющая частное решение вида (22), где $L(t)$ - матрица Ляпунова, а A - постоянная матрица, является приводимой и при этом она приводится к виду (21) при помощи преобразования Ляпунова (20).

Следуя А. М. Ляпунову, покажем, что всякая система (19) с периодическими коэффициентами приводима.

Пусть в данной системе (19) $P(t)$ - непрерывная функция в интервале $(-\infty; +\infty)$ с периодом τ :

$$P(t + \tau) = P(t) \quad (23)$$

Заменяя в (19) t на $t + \tau$ и используя (23), получим:

$$\frac{dX(t + \tau)}{dt} = P(t) X(t + \tau)$$

Таким образом, $X(t + \tau)$, как и $X(t)$ является интегральной матрицей системы (19).

Поэтому $X(t + \tau) = X(t)V$, где V - некоторая постоянная неособенная матрица.

Поскольку $|V| \neq 0$, то можно определить

$$V^{\frac{t}{\tau}} = e^{\frac{t}{\tau} \ln V}$$

Эта матричная функция от t , как и $X(t)$, умножается справа на V , если к аргументу прибавить τ . Поэтому «частное»

$$L(t) = X(t)V^{-\frac{t}{\tau}} = X(t)e^{-\frac{t}{\tau} \ln V}$$

является непрерывной периодической функцией с периодом τ :

$$L(t + \tau) = L(t)$$

и с определителем $|L(t)| \neq 0$. Матрица $L(t)$ удовлетворяет условиям 1°-3° предыдущего параграфа и, следовательно, является матрицей Ляпунова.

С другой стороны, поскольку решение X системы (19) представимо в виде

$$X = L(t)e^{\frac{t}{\tau} \ln V}$$

то система (19) приводима.

В данном случае преобразование Ляпунова

$$X = L(t)Y,$$

приводящее систему (19) к виду

$$\frac{dY}{dt} = \frac{1}{\tau} \ln V \cdot Y,$$

имеет периодические коэффициенты с периодом τ .

А. М. Ляпуновым был установлен весьма важный критерий устойчивости и неустойчивости по первому линейному приближению для нелинейных систем дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^n a_{ik} x_k + (**), \quad (24)$$

где в правых частях стоят сходящиеся степенные ряды относительно $x_1, x_2, \dots, x_n$,

а $(**)$ обозначает сумму членов этих рядов второго порядка и выше

относительно $x_1, x_2, \dots, x_n$; коэффициенты a_{ik} ($i, k = 1, 2, \dots, n$) в линейных членах постоянны.

Критерий Ляпунова. Нулевое решение системы (24) будет устойчивым (и притом асимптотически), если матрица коэффициентов первого линейного

приближения $A = \|a_{ik}\|$ имеет все характеристические числа с отрицательными вещественными частями, и неустойчивым, если хотя бы одно из этих характеристических чисел имеет положительную вещественную часть.

Приведенные выше рассуждения позволяют использовать этот критерий для систем с периодическими коэффициентами в линейных членах:

$$\frac{dx_i}{dt} = \sum_{k=1}^n p_{ik}(t) x_k + (**) \quad (i = 1, 2, \dots, n) \quad (25)$$

Действительно, на основании предыдущих рассуждений можно при помощи преобразования Ляпунова систему (25) привести к виду (24), где

$$A = \|a_{ik}\| = \frac{1}{\tau} \ln V$$

а V - постоянная матрица, на которую умножается интегральная матрица соответствующей линейной системы (19) при сдвиге аргумента на τ . Не нарушая общности, можем считать $\tau > 0$. В силу свойств преобразования Ляпунова нулевое решение исходной системы и нулевое решение преобразованной одновременно являются устойчивыми, асимптотически устойчивыми или неустойчивыми.

Но характеристические числа λ_i и ν_i ($i = 1, 2, \dots, n$) матриц A и V связаны между собой формулой

$$\lambda_i = \frac{1}{\tau} \ln \nu_i \quad (i = 1, 2, \dots, n).$$

Поэтому, применяя критерий Ляпунова к приведенной системе, найдем:

Нулевое решение системы (25) будет асимптотически устойчивым, если

все характеристические числа $\nu_1, \nu_2, \dots, \nu_n$ матрицы V по модулю < 1 , и неустойчивым, если хотя бы одно из этих чисел по модулю > 1 .

А. М. Ляпунов установил свой критерий устойчивости по линейному приближению для значительно более широкого класса систем, а именно для систем вида (24), у которых система линейного приближения не обязательно система с постоянными коэффициентами, но принадлежит к классу систем, названных Ляпуновым правильными. Класс правильных линейных систем содержит в себе как часть все приводимые системы. Критерий неустойчивости для случая, когда первое линейное приближение является правильной системой, был установлен Н. Г. Четаевым.

1. 9 Лекция №9 (1 час).

Тема: «Добавление неравенства для собственных и сингулярных чисел»

1.9.1 Вопросы лекции:

1. Мажорирующие последовательности.
2. Неравенство Вейля.

1.9.2 Краткое содержание вопросов: (тезисно изложить основное содержание рассматриваемых вопросов)

1. Мажорирующие последовательности.

В этом параграфе мы остановимся на ряде вспомогательных вопросов, связанных с конечными числовыми последовательностями. Рассмотрим две убывающие последовательности чисел, по n элементов в каждой:

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n, \quad (1)$$

$$\alpha'_1 \geq \alpha'_2 \geq \dots \geq \alpha'_n. \quad (2)$$

Принято говорить, что последовательность (2) мажорируется последовательностью (1), если

$$\alpha'_1 + \alpha'_2 + \dots + \alpha'_m \leq \alpha_1 + \alpha_2 + \dots + \alpha_m \quad (1 \leq m \leq n-1) \quad (3)$$

и

$$\alpha'_1 + \alpha'_2 + \dots + \alpha'_n = \alpha_1 + \alpha_2 + \dots + \alpha_n. \quad (3')$$

При выполнении условий (3) и (3') пишут

$$\alpha' < \alpha. \quad (4)$$

Квадратную матрицу $T = \|t_{ij}\|_1^n$ мы будем в дальнейшем называть двойкой стохастической, если матрицы T и T' являются стохастическими, другими словами, если $t_{ij} \geq 0$,

$$\sum_{j=1}^n t_{ij} = 1, \quad 1 \leq i \leq n, \quad (5)$$

и

$$\sum_{i=1}^n t_{ij} = 1, \quad 1 \leq j \leq n. \quad (5')$$

Справедливо следующее утверждение (см. [35], стр. 63).

Лемма 1. Последовательность α' мажорируется последовательностью α тогда и только тогда, когда существует двойка стохастическая матрица T такая, что

$$\alpha' = T\alpha. \quad (6)$$

Достаточность условия (6) доказывается легко. В самом деле,

$$\sum_{k=1}^m \alpha'_k = \sum_{k=1}^m \sum_{j=1}^n t_{kj} \alpha_j = \sum_{j=1}^n \left(\sum_{k=1}^m t_{kj} \right) \alpha_j = \sum_{j=1}^n \omega_j \alpha_j. \quad (7)$$

Мы положили

$$\omega_j = \sum_{k=1}^m t_{kj} \quad (1 \leq j \leq n). \quad (8)$$

Легко видеть, что $0 \leq \omega_j \leq 1$ и

$$\sum_{j=1}^n \omega_j = \sum_{k=1}^m \left(\sum_{j=1}^n t_{kj} \right) = m. \quad (9)$$

Имеем на основании равенства (7)

$$\sum_{j=1}^m \alpha_j - \sum_{k=1}^m \alpha'_k = \sum_{j=1}^m \alpha_j - \sum_{j=1}^n \omega_j \alpha_j = \alpha_1 (1 - \omega_1) + \dots + \alpha_m (1 - \omega_m) - \omega_{m+1} \alpha_{m+1} - \dots - \omega_n \alpha_n. \quad (10)$$

Уменьшая слагаемые в правой части, получаем

$$\sum_{j=1}^m \alpha_j - \sum_{j=1}^m \alpha'_j \geq \alpha_m (1 - \omega_1) + \dots + \alpha_m (1 - \omega_m) - \omega_{m+1} \alpha_m - \dots - \omega_n \alpha_m = \alpha_m (m - m) = 0. \quad (10')$$

Следовательно, неравенства (3) имеют место. Так как, далее, при $m = n$ согласно

(8) $\varphi_j = 1$ ($j = 1, 2, \dots, n$), то в силу (7) справедливо и равенство (3').

Таким образом, достаточность условия (6) установлена. Доказательство необходимости этого условия требует известных усилий. Мы проведем его по индукции. В

случае $n = 1$ последовательности содержат по одному элементу, $\alpha'_1 = \alpha_1$ и матрица T , очевидно, существует. Предположим, что утверждение справедливо для случая последовательностей из $n - 1$ элементов и рассмотрим две последовательности α' и α , которые связаны соотношением $\alpha' < \alpha$ и состоят из n элементов.

Из условия $\alpha'_1 \leq \alpha_1$ и равенства (3) следует, что $\alpha_n \leq \alpha'_1 \leq \alpha_1$. Поэтому найдется такое k ($1 \leq k \leq n - 1$), при котором

$$\alpha_{k+1} \leq \alpha'_1 \leq \alpha_k. \quad (11)$$

Следовательно, при некотором τ , $0 \leq \tau \leq 1$, мы имеем

$$\alpha'_1 = \tau \alpha_k + (1 - \tau) \alpha_{k+1}. \quad (12)$$

Наряду с α' и α рассмотрим две последовательности по $n - 1$ элементов в каждой:

$$\alpha'_2, \alpha'_3, \dots, \alpha'_k, \alpha'_{k+1}, \alpha'_{k+2}, \dots, \alpha'_n \quad (13)$$

и

$$\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_k + \alpha_{k+1} - \alpha'_1, \alpha_{k+2}, \dots, \alpha_n. \quad (13')$$

Обозначим эти последовательности через $\tilde{\alpha}'$ и $\tilde{\alpha}$ соответственно.

Учитывая (11), легко заключить, что элементы последовательности $\tilde{\alpha}$ расположены в порядке убывания. Без труда проверяется также соотношение $\tilde{\alpha}' < \tilde{\alpha}$. Поэтому в силу

индуктивного предположения существует такая двоякостochasticкая матрица $\tilde{T} = \|\tilde{t}_{ij}\|_{i,j=1}^{n-1}$,

что $\tilde{\alpha}' = \tilde{T} \tilde{\alpha}$, или в развернутой записи:

$$\alpha'_{s+1} = t_{s,1} \alpha_1 + \dots + t_{s,k-1} \alpha_{k-1} + t_{s,k} (\alpha_k + \alpha_{k+1} - \alpha'_1) + t_{s,k+1} \alpha_{k+2} + \dots + t_{s,n-1} \alpha_n \quad (1 \leq s \leq n-1).$$

Подставив сюда α'_1 из равенства (12), получим при $1 \leq s \leq n-1$:

$$\alpha'_{s+1} = t_{s,1} \alpha_1 + \dots + t_{s,k} (1 - \tau) \alpha_k + t_{s,k} \tau \alpha_{k+1} + t_{s,k+1} \alpha_{k+2} + \dots + t_{s,n-1} \alpha_n.$$

Добавляя сюда равенство $\alpha'_1 = \tau \alpha_k + (1 - \tau) \alpha_{k+1}$, легко убеждаемся в том, что последовательности α' и α связаны двояко стохастической матрицей

$$T = \begin{pmatrix} 0 & 0 & 0 & \dots & \tau & 1 - \tau & \dots & 0 \\ t_{11} & t_{12} & t_{13} & \dots & t_{1k} (1 - \tau) & t_{1k} \tau & \dots & t_{1,n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ t_{n-1,1} & t_{n-1,2} & t_{n-1,3} & \dots & t_{n-1,k} (1 - \tau) & t_{n-1,k} \tau & \dots & t_{n-1,n-1} \end{pmatrix}.$$

Лемма доказана полностью.

Нам понадобится ниже также следующее предложение (см. 233):

Лемма 2. Пусть $\varphi(t)$ – непрерывная выпуклая монотонно возрастающая функция. Пусть

$$\alpha'_1 \geq \alpha'_2 \geq \dots \geq \alpha'_p, \quad (14)$$

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_p \quad (15)$$

и

$$\alpha'_1 + \alpha'_2 + \dots + \alpha'_m \leq \alpha_1 + \alpha_2 + \dots + \alpha_m \quad (1 \leq m \leq p). \quad (16)$$

Тогда

$$\varphi(\alpha'_1) + \varphi(\alpha'_2) + \dots + \varphi(\alpha'_p) \leq \varphi(\alpha_1) + \varphi(\alpha_2) + \dots + \varphi(\alpha_p). \quad (17)$$

Доказательство. Предположим сначала, что при $m = p$ в соотношении (16) имеет место равенство. Тогда последовательность α' мажорируется последовательностью α и согласно лемме 1

$$\alpha'_s = \sum_{j=1}^p t_{sj} \alpha_j, \quad 1 \leq s \leq p, \quad (18)$$

где t_{sj} - элементы двояко стохастической матрицы. В силу выпуклости $\varphi(t)$ из равенства (18) следует, что

$$\varphi(\alpha'_s) \leq \sum_{j=1}^p t_{sj} \varphi(\alpha_j). \quad (19)$$

Суммируя неравенства (19), получаем

$$\sum_{s=1}^p \varphi(\alpha'_s) \leq \sum_{j=1}^p \left(\sum_{s=1}^p t_{sj} \right) \varphi(\alpha_j) = \sum_{j=1}^p \varphi(\alpha_j). \quad (20)$$

Таким образом, в указанном случае неравенство (17) выполняется.

Рассмотрим теперь общий случай. Пусть в соотношении (16) при $m = p$ имеет место знак $<$. Положим

$$\sum_{j=1}^p \alpha_j - \sum_{j=1}^p \alpha'_j = c > 0.$$

Наряду с последовательностями (14) и (15) рассмотрим две последовательности:

$$\alpha'_1 \geq \alpha'_2 \geq \dots \geq \alpha'_p \geq \alpha'_{p+1} \quad (21)$$

и

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_p \geq \alpha_{p+1}, \quad (22)$$

где α'_{p+1} и α_{p+1} – произвольные два числа, удовлетворяющие неравенствам (21) и (22) и соотношению

$$\alpha_{p+1} = \alpha'_{p+1} - c. \quad (23)$$

Легко видеть, что при таком выборе α'_{p+1} и α_{p+1} последовательность (21) мажорируется последовательностью (22), и по доказанному имеем

$$\varphi(\alpha'_1) + \varphi(\alpha'_2) + \dots + \varphi(\alpha'_p) + \varphi(\alpha'_{p+1}) \leq \varphi(\alpha_1) + \varphi(\alpha_2) + \dots + \varphi(\alpha_p) + \varphi(\alpha_{p+1}). \quad (24)$$

Так как, далее, $\varphi(t)$ – монотонно возрастающая функция и $\alpha'_{p+1} > \alpha_{p+1}$,

то $\varphi(\alpha'_{p+1}) \geq \varphi(\alpha_{p+1})$ и из (24) снова следует неравенство (17).

Лемма доказана полностью.

Замечание. Из наших рассуждений следует, что в том случае, когда последовательность (14) мажорируется последовательностью (15) [т. е. при $m = p$ в (16) достигается равенство], то неравенство (17) справедливо для любой

непрерывной выпуклой функции $\varphi(t)$ (возрастание является излишним требованием).

2. Неравенство Вейля.

Пусть $\mathbf{A}$ – линейный оператор, действующий в n -мерном унитарном пространстве R .

Собственные числа неотрицательного эрмитова оператора $\sqrt{A^* A}$ (см. стр. 249) принято называть сингулярными числами оператора $\mathbf{A}$.

В настоящем параграфе мы установим неравенства, связывающие сингулярные числа произведения двух операторов с сингулярными числами сомножителей.

Пусть $y_1, y_2, \dots, y_m$ и $z_1, z_2, \dots, z_m$ – два набора векторов из R . Введем сокращенное обозначение для определителя порядка m , связанного с данными наборами:

$$[(y_i, z_j)] = \begin{vmatrix} (y_1, z_1) & (y_2, z_1) & \dots & (y_m, z_1) \\ (y_1, z_2) & (y_2, z_2) & \dots & (y_m, z_2) \\ \dots & \dots & \dots & \dots \\ (y_1, z_m) & (y_2, z_m) & \dots & (y_m, z_m) \end{vmatrix}. \quad (25)$$

Рассмотрим далее неотрицательный эрмитов оператор H , действующий в R . Собственные значения оператора H занумеруем в убывающем порядке:

$$h_1 \geq h_2 \geq \dots \geq h_n \geq 0. \quad (26)$$

Справедливо следующее предложение, принадлежащее А. Хорну ([191b]):

Лемма 3. Пусть

$$x_1, x_2, \dots, x_m \quad (m \leq n) \quad (27)$$

– произвольный набор векторов из R . Тогда

$$[(Hx_i, x_j)] \leq h_1 h_2 \dots h_m [(x_i, x_j)]. \quad (28)$$

Для доказательства

рассмотрим ортонормированный базис собственных векторов оператора H :

$$e_1, e_2, \dots, e_n \quad (29)$$

и разложим каждый из векторов x_i ($i = 1, 2, \dots, m$) по базису (29). Вычисляя скалярное произведение, получаем:

$$(Hx_i, x_j) = \sum_{s=1}^n h_s (x_i, e_s) (\overline{x_j, e_s}) = \sum_{s=1}^n h_s (x_i, e_s) (e_s, x_j), \quad i, j = 1, 2, \dots, m \quad (30)$$

Равенство (30) позволяет рассматривать матрицу определителя $[(Hx_i, x_j)]$ как результат умножения двух прямоугольных матриц размеров $m \times n$ и $n \times m$.

Разлагая определитель по формуле Бине-Коши (см. стр. 20), получаем в принятых обозначениях для определителей:

$$[(Hx_i, x_j)] = \sum_{1 \leq s_1 < s_2 < \dots < s_m \leq n} [(x_i, h_{s_1} e_{s_1})] [(e_{s_1}, x_j)] \quad (31)$$

Здесь

$$[(x_i, h_{s_1} e_{s_1})] = \begin{vmatrix} (x_1, h_{s_1} e_{s_1}) & \dots & (x_m, h_{s_1} e_{s_1}) \\ (x_1, h_{s_2} e_{s_2}) & \dots & (x_m, h_{s_2} e_{s_2}) \\ \dots & \dots & \dots \\ (x_1, h_{s_m} e_{s_m}) & \dots & (x_m, h_{s_m} e_{s_m}) \end{vmatrix}, \quad (31')$$

а суммирование ведется по всевозможным

наборам натуральных чисел $1 \leq s_1 < s_2 < \dots < s_m \leq n$.

Оценив правую часть (31) по неравенству Коши-Буняковского, получим:

$$\left[(Hx_i, x_j) \right]^2 \leq \left(\sum_{1 \leq s_1 < \dots < s_m \leq n} \left[(x_i, h_{s_1} e_{s_1}) \right]^2 \right) \left(\sum_{1 \leq s_1 < \dots < s_m \leq n} \left[(e_{s_1}, x_j) \right]^2 \right). \quad (32)$$

Вторая сумма в правой части неравенства (32) равна определителю Грама $\left[(x_i, x_j) \right]$. В этом легко убедиться, положив в формуле (31) $H = E$, где E – единичный оператор. Впрочем, соответствующее равенство отдельно доказано на стр. 230 (формула (26)). В первой сумме правой части (32) вынесем из каждого определителя (31')

произведение $h_{s_1} h_{s_2} \dots h_{s_m}$ и заменим его большим $h_1 h_2 \dots h_m$. В результате получим:

$$\left[(Hx_i, x_j) \right]^2 \leq h_1^2 h_2^2 \dots h_m^2 \left[(x_i, x_j) \right]^2.$$

Извлекая из обеих частей этого неравенства квадратные корни, мы устанавливаем справедливость неравенства (28). Докажем далее следующий факт.

Лемма 4. Пусть K – произвольный оператор в R и

$$\chi_1 \geq \chi_2 \geq \dots \geq \chi_n \quad (33)$$

– его сингулярные числа. Тогда для произвольного набора

векторов $x_1, x_2, \dots, x_m$ ($m \leq n$) справедливо неравенство

$$\left[(Kx_i, Kx_j) \right] \leq \chi_1^2 \chi_2^2 \dots \chi_m^2 \left[(x_i, x_j) \right]. \quad (34)$$

Неравенство (34) немедленно следует из леммы 3 при $H = K^* K$.

Установим, наконец, еще одно вспомогательное предложение.

Лемма 5. Пусть A и B – линейные операторы в R , $C = AB$, и

пусть α_s, β_s и γ_s ($s = 1, 2, \dots, n$) – сингулярные числа соответственно A, B и C ,

занумерованные в убывающем порядке. Тогда при любом $m \leq n$ справедливы неравенства

$$\gamma_1 \gamma_2 \dots \gamma_m \leq \alpha_1 \alpha_2 \dots \alpha_m \beta_1 \beta_2 \dots \beta_m. \quad (35)$$

Для доказательства рассмотрим ортонормированный базис $e_1, e_2, \dots, e_n$ собственных

векторов оператора $C^* C$. Последовательно применяя (34), получаем

$$\left[(Ce_i, Ce_j) \right] = \left[(ABe_i, ABe_j) \right] \leq \alpha_1^2 \alpha_2^2 \dots \alpha_m^2 \left[(Be_i, Be_j) \right] \leq \alpha_1^2 \alpha_2^2 \dots \alpha_m^2 \beta_1^2 \beta_2^2 \dots \beta_m^2 \left[(e_i, e_j) \right]. \quad (36)$$

С другой стороны, поскольку e_i ($1 \leq i \leq m$) – собственные векторы $C^* C$, мы имеем:

$$\left[(Ce_i, Ce_j) \right] = \left[(C^* C e_i, e_j) \right] = \gamma_1^2 \gamma_2^2 \dots \gamma_n^2 \left[(e_i, e_j) \right]. \quad (37)$$

Следовательно, (35) имеет место.

Мы в состоянии теперь доказать следующую теорему, которая является основной целью настоящего параграфа.

Теорема 1 (Нейман-Хорн [218b, 191b]). Пусть A, B и C – линейные операторы в n -мерном унитарном пространстве R . Пусть $C = AB$ и пусть α_s, β_s и γ_s ($s = 1, 2, \dots, n$) – сингулярные числа операторов A, B и C , занумерованные в порядке убывания.

Пусть $f(x)$ – непрерывная при $x \geq 0$ функция такая, что $\varphi(t) = f(e^t)$ – монотонно возрастающая выпуклая функция параметра t . Тогда при всех $m \leq n$ справедливы неравенства

$$\sum_{s=1}^m f(\gamma_s) \leq \sum_{s=1}^m f(\alpha_s \beta_s) \quad (38)$$

Доказательство. Пусть сначала операторы A и B невырождены, тогда все числа α_s, β_s и γ_s положительны. Логарифмируя неравенства (35), получаем

$$\sum_{s=1}^m \ln \gamma_s \leq \sum_{s=1}^m \ln(\alpha_s \beta_s), \quad 1 \leq m \leq n \quad (39)$$

На основании леммы 2 имеем

$$\sum_{s=1}^m \varphi(\ln \gamma_s) \leq \sum_{s=1}^m \varphi(\ln \alpha_s \beta_s), \quad 1 \leq m \leq n \quad (40)$$

Так как $\varphi(t) = f(e^t)$, то отсюда следует (38). В случае вырожденных операторов неравенства (38) устанавливаются по непрерывности.

Замечание 1. В случае $f(x) = x^\sigma$ ($\sigma \geq 0$) получаем

$$\sum_{s=1}^m \gamma_s^\sigma \leq \sum_{s=1}^m \alpha_s^\sigma \beta_s^\sigma \quad (1 \leq m \leq n) \quad (41)$$

В таком виде неравенства (38) встречаются в приложениях чаще всего.

Замечание 2. При $m = n$ неравенство (39) превращается в равенство (см. сноску на стр. 540). Поэтому при $m = n$ неравенство (40) справедливо для любой

непрерывной выпуклой функции $\varphi(t)$ (см. замечание к лемме 2).

В частности, неравенство (41) при $m = n$ справедливо и для $\sigma < 0$.

Замечание 3. Пусть

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$$

– сингулярные числа оператора A и пусть

$$\alpha_{i,1} \geq \alpha_{i,2} \geq \dots \geq \alpha_{i,n}$$

– сингулярные числа оператора A^l (l – натуральное число). Тогда при любом $\sigma \geq 0$ и любом $1 \leq m \leq n$

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \sum_{s=1}^m \alpha_s^{\sigma l} \quad (42)$$

Неравенства (42) докажем индукцией по l . При $l = 1$ соотношение (42) очевидно; пусть оно выполняется для $l-1$. Так как $A^l = A^{l-1} \cdot A$, то согласно (41)

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \sum_{s=1}^m \alpha_{i-1,s}^\sigma \alpha_s^\sigma, \quad 1 \leq m \leq n \quad (43)$$

Применяя к правой части (43) неравенство Гёльдера с

$$p = \frac{l}{l-1} \text{ и } q = l \quad (p^{-1} + q^{-1} = 1),$$

получаем

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \left(\sum_{s=1}^m \alpha_{i-1,s}^{p\sigma} \right)^{1/p} \left(\sum_{s=1}^m \alpha_s^{q\sigma} \right)^{1/q} \quad (44)$$

По предположению индукции, имеем для первой суммы в правой части (44):

$$\left(\sum_{s=1}^m \alpha_{i-1,s}^{p\sigma}\right)^{1/p} \leq \left(\sum_{s=1}^m \alpha_s^{p\sigma(l-1)}\right)^{1/p} = \left(\sum_{s=1}^m \alpha_s^{l\sigma}\right)^{1-\frac{1}{l}}.$$

Учитывая, что во второй сумме правой части (44) $q = l$, легко получаем из (44):

$$\sum_{s=1}^m \alpha_{i,s}^{\sigma} \leq \sum_{s=1}^m \alpha_s^{l\sigma},$$

что и требовалось доказать.

В частности, при $\sigma = 2$ и $m = n$ из формулы (42) следует, что

$$\mathrm{Sp}\left(A^{*l}A^l\right) \leq \mathrm{Sp}\left(A^*A\right)^l. \quad (45)$$

2. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

2.1 Практическое занятие №1(1 час).

Тема: «Арифметические действия над матрицами»

2.1.1 Задание для работы:

1. Сложение и вычитание матриц
2. Умножение матриц
3. Транспонирование и нахождение обратной матрицы

2.1.2 Краткое описание проводимого занятия:

1. Умножение матрицы на число.

Пример.

$$2 \cdot \begin{pmatrix} 3 & 4 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ -4 & 10 \end{pmatrix} \text{ или } 3 \cdot \begin{pmatrix} 2 & -3 & 2 & 1 \\ 0 & 2 & -4 & 3 \\ 4 & 4 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 6 & -9 & 6 & 3 \\ 0 & 6 & -12 & 9 \\ 12 & 12 & 6 & 6 \end{pmatrix}$$

2. Сложение (вычитание) матриц.

Пример.

$$A + B = \begin{pmatrix} 3 & 6 & -2 \\ 0 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 4 & -3 & 4 \\ 5 & 2 & -6 \end{pmatrix} = \begin{pmatrix} 3+4 & 6-3 & -2+4 \\ 0+5 & 3+2 & 1-6 \end{pmatrix} = \begin{pmatrix} 7 & 3 & 2 \\ 5 & 5 & -5 \end{pmatrix}$$

3. Умножение матриц.

Если матрица A имеет размерность $m \times k$, а матрица B имеет размерность $k \times n$, то их можно умножить. В результате получается матрица $C = A \cdot B$, размерность которой будет $m \times n$. Например, A размерности 3×4 , B размерности 4×2 , то $C = A \cdot B$ будет размерности 3×2 .

Пример.

$$C = A \cdot B = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 \cdot 5 + 4 \cdot 2 & 2 \cdot 6 + 4 \cdot 4 \\ 1 \cdot 5 + 2 \cdot 2 & 1 \cdot 6 + 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 18 & 28 \\ 9 & 14 \end{pmatrix}$$

$2 \times 2 \quad 2 \times 2 \quad 2 \times 2$

Первую строку A умножаем на первый столбец B : $2 \cdot 5 + 4 \cdot 2 = 18$.

Первую строку A умножаем на второй столбец B : $2 \cdot 6 + 4 \cdot 4 = 28$.

Вторую строку A умножаем на первый столбец B : $1 \cdot 5 + 2 \cdot 2 = 9$.

Вторую строку A умножаем на второй столбец B : $1 \cdot 6 + 2 \cdot 4 = 14$.

Пример.

$$C = A \cdot B = \begin{pmatrix} -1 & 2 & 4 \\ 3 & 4 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ -6 & 5 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} -1-12+8 & -4+10+28 \\ 3-24-6 & 12+20-21 \end{pmatrix} = \begin{pmatrix} -5 & 34 \\ -27 & 11 \end{pmatrix}$$

$2 \times 3 \quad 3 \times 2 \quad 2 \times 2$

Пример.

$$C = A \cdot B = \begin{pmatrix} -2 & 3 & 1 \\ 4 & 3 & 2 \\ -2 & 5 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 3 & -3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0+9+1 & -4-9+1 \\ 0+9+2 & 8-9+2 \\ 0+15+1 & -4-15+1 \end{pmatrix} = \begin{pmatrix} 10 & -12 \\ 11 & 1 \\ 16 & -18 \end{pmatrix}$$

$3 \times 3 \quad 3 \times 2 \quad 3 \times 2$

4. Транспонирование матриц.

Чтобы транспонировать матрицу, надо строки матрицы записать в столбцы.

Пример.

Если $A = \begin{pmatrix} 3 & 2 \\ -4 & 6 \end{pmatrix}$, то транспонированная матрица $A^T = \begin{pmatrix} 3 & -4 \\ 2 & 6 \end{pmatrix}$

Если $B = \begin{pmatrix} -3 & 0 & 5 \\ 8 & 3 & 2 \end{pmatrix}$, то $B^T = \begin{pmatrix} -3 & 8 \\ 0 & 3 \\ 5 & 2 \end{pmatrix}$

Задание 1. Найти $C = A \cdot B + \alpha \cdot D^T$

$$A = \begin{pmatrix} 2 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & -3 \\ 3 & 0 \\ 2 & 1 \end{pmatrix} \quad D = \begin{pmatrix} -1 & 2 & 6 \\ 0 & -2 & 3 \end{pmatrix} \quad \alpha = 2$$

Решение.

$$C = \begin{pmatrix} 2 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 3 & 0 \\ 2 & 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} -1 & 0 \\ 2 & -2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 2+9+0 & -6+0+0 \\ -1+6+2 & 3+0+1 \\ 0+6+6 & 0+0+3 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 4 & -4 \\ 12 & 6 \end{pmatrix} =$$

$$= \begin{pmatrix} 11 & -6 \\ 7 & 4 \\ 12 & 3 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 4 & -4 \\ 12 & 6 \end{pmatrix} = \begin{pmatrix} 9 & -6 \\ 11 & 0 \\ 24 & 9 \end{pmatrix}$$

2.1.3 Результаты и выводы:

Найти $C = A \cdot B + \alpha \cdot D^T$

Вариант 1

$$A = \begin{pmatrix} 2 & 3 & 1 \\ -4 & 0 & 3 \\ 1 & -1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 4 \\ 2 & 1 \\ -5 & 3 \end{pmatrix} \quad D = \begin{pmatrix} -2 & 0 & 5 \\ 1 & 3 & 2 \end{pmatrix} \quad \alpha = -2$$

Вариант 2

$$A = \begin{pmatrix} -2 & 4 & 1 \\ 3 & 0 & 5 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 3 & 2 \\ 4 & 1 & -2 \\ -4 & 6 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 0 & -4 \\ 2 & 2 \\ 4 & 4 \end{pmatrix} \quad \alpha = 2$$

Вариант 3

$$A = \begin{pmatrix} -4 & 2 \\ 6 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -2 & 5 & 7 & 3 \\ 0 & 2 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} -3 & 4 \\ 1 & 2 \\ 2 & 2 \\ 0 & 3 \end{pmatrix} \quad \alpha = -2$$

Вариант 4

$$A = \begin{pmatrix} 0 & 0 & 3 & 1 \\ 2 & 1 & 1 & -2 \\ -1 & 0 & 0 & 2 \\ 6 & 2 & 4 & 4 \end{pmatrix} \quad B = \begin{pmatrix} -3 & 4 \\ 2 & 2 \\ 0 & 5 \\ 2 & 6 \end{pmatrix} \quad D = \begin{pmatrix} -6 & 3 & 4 & -2 \\ 3 & 0 & 3 & 4 \end{pmatrix} \quad \alpha = 2$$

Вариант 5

$$A = \begin{pmatrix} 1 & 3 & 1 \\ -4 & 3 & 0 \\ 1 & 4 & -2 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 4 \\ 2 & -1 \\ 5 & 0 \end{pmatrix} \quad D = \begin{pmatrix} -2 & 2 & 5 \\ 1 & -1 & 2 \end{pmatrix} \quad \alpha = -1$$

Вариант 6

$$A = \begin{pmatrix} -2 & -4 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 0 & -3 & 2 \\ 4 & -4 & 2 \\ -4 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 3 & -4 \\ 2 & -2 \\ -1 & 4 \end{pmatrix} \quad \alpha = -3$$

Вариант 7

$$A = \begin{pmatrix} -3 & 3 & 1 \\ -4 & 4 & -3 \\ 1 & -1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 4 \\ 4 & 0 \\ -5 & -3 \end{pmatrix} \quad D = \begin{pmatrix} -2 & -2 & 5 \\ 0 & 3 & 2 \end{pmatrix} \quad \alpha = -3$$

Вариант 8

$$A = \begin{pmatrix} 5 & -2 \\ -6 & -1 \end{pmatrix} \quad B = \begin{pmatrix} -2 & 5 & 0 & -3 \\ 1 & 7 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 3 & 4 \\ -1 & -2 \\ 2 & 0 \\ 0 & -5 \end{pmatrix} \quad \alpha = 2$$

Вариант 9

$$A = \begin{pmatrix} -2 & 4 & 3 & 1 \\ 2 & 1 & -1 & -2 \\ -1 & 0 & 0 & 7 \\ 6 & -2 & 1 & 4 \end{pmatrix} \quad B = \begin{pmatrix} -3 & 4 \\ 0 & 2 \\ 0 & 5 \\ 2 & 0 \end{pmatrix} \quad D = \begin{pmatrix} -6 & 1 & 4 & -2 \\ -3 & 0 & 3 & 1 \end{pmatrix} \quad \alpha = -2$$

Вариант 10

$$A = \begin{pmatrix} 4 & -2 \\ -6 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -2 & 5 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} -3 & 4 \\ -1 & 2 \\ 2 & 0 \\ 0 & -3 \end{pmatrix} \quad \alpha = -2$$

2.2 Практическое занятие №2(1 час).

Тема: «Применение метода Гаусса»

2.2.1 Задание для работы:

1. Решение систем линейных уравнений методом Гаусса
2. Применение метода Гаусса для различных типов систем линейных уравнений

2.2.2 Краткое описание проводимого занятия:

Метод Крамера и матричный метод применяется только для квадратных систем (число уравнений равно числу неизвестных), причем определитель должен быть не равен нулю. Если число уравнений не равно числу неизвестных, или определитель системы равен нулю, применяется метод Гаусса. Метод Гаусса можно применять для решения любых систем.

Мы рассмотрим решение системы четвертого порядка. Если применять метод Крамера, придется находить пять определителей четвертого порядка. Если решиться обращать матрицу четвертого порядка, то придется находить 16 определителей третьего порядка.

Метод Гаусса состоит в приведении матрицы системы к треугольной или трапециевидной форме.

Пример. Решить систему третьего порядка методом Гаусса.

$$\begin{cases} x + 3y - 4z = 2 \\ 2x - y + 5z = -2 \\ 4x - 2y + z = 7 \end{cases}$$

Выписываем матрицу системы, в которую включаем и правые части:

$$\begin{aligned} A = \begin{pmatrix} 1 & 3 & -4 & 2 \\ 2 & -1 & 5 & -2 \\ 4 & -2 & 1 & 7 \end{pmatrix} &\cong \begin{pmatrix} a_{11} = 1 \text{ назовем ведущим элементом} \\ \text{и с его помощью занулим числа под} \\ \text{ним в первом столбце} \end{pmatrix} \cong \begin{pmatrix} 1 & 3 & -4 & 2 \\ 0 & -7 & 13 & -6 \\ 0 & -14 & 17 & -1 \end{pmatrix} \cong \\ &\cong \begin{pmatrix} \text{элементы первой строки умножили} \\ \text{на } -2 \text{ и прибавили ко второй строке;} \\ \text{затем элементы первой строки умножили} \\ \text{на } -4 \text{ и прибавили к третьей строке} \end{pmatrix} \begin{pmatrix} a_{22} = -7 \text{ назовем ведущим элементом} \\ \text{и с его помощью занулим числа под} \\ \text{ним во втором столбце} \end{pmatrix} \cong \\ &\cong \begin{pmatrix} 1 & 3 & -4 & 2 \\ 0 & -7 & 13 & -6 \\ 0 & 0 & -9 & 11 \end{pmatrix} \begin{pmatrix} \text{элементы второй строки умножили} \\ \text{на } -2 \text{ и прибавили ко третьей строке;} \end{pmatrix} \end{aligned}$$

Матрица приведена к треугольной форме (столбец правых частей не считается). По матрице восстановим систему:

$$\begin{cases} x + 3y - 4z = 2 \\ -7y + 13z = -6 \\ -9z = 11 \end{cases}$$

Из последнего уравнения находим $z = -\frac{11}{9}$. Подставим во второе уравнение:

$$-7y - \frac{143}{9} = -6 \quad -7y = \frac{89}{9} \quad y = -\frac{89}{63}. \text{ Найденные } y \text{ и } z \text{ подставим в первое уравнение:}$$

$$x - \frac{89}{21} - \frac{44}{9} = 2 \quad x = \frac{255}{189} \quad \text{Ответ: } x = \frac{255}{189} \quad y = -\frac{89}{63} \quad z = -\frac{11}{9}$$

Задание 5. Решить систему уравнений методом Гаусса.

$$\begin{cases} x - 2y + z + 2u = 1 \\ 3x - 5y + 4z + 6u = 9 \\ x - y + 5z - u = 7 \\ 4x - 8y + 4z + 9u = 6 \end{cases}$$

2.3 Практическое занятие №2(1 час).

Тема: «Решение матричных уравнений»

2.3.1 Задание для работы:

2.3.2 Краткое описание проводимого занятия:

Типовое матричное уравнение состоит, как правило, из нескольких матриц и неизвестной матрицы X , которую предстоит найти. То есть, решением матричного уравнения является матрица.

Пример 1

Решить матричное уравнение, выполнить проверку

$$\begin{pmatrix} 1 & -3 \\ 8 & 0 \end{pmatrix} - 2X = 3 \cdot \begin{pmatrix} -1 & 1 \\ 0 & 4 \end{pmatrix}$$

Как решить матричное уравнение?

Фактически нужно использовать алгоритм решения детского уравнения с числами.

В правой части умножаем каждый элемент матрицы на три, а матрицу левой части переносим направо со сменой знака:

$$-2X = \begin{pmatrix} -3 & 3 \\ 0 & 12 \end{pmatrix} - \begin{pmatrix} 1 & -3 \\ 8 & 0 \end{pmatrix}$$

Причёмсываем правую часть:

$$-2X = \begin{pmatrix} -3-1 & 3-(-3) \\ 0-8 & 12-0 \end{pmatrix}$$

$$-2X = \begin{pmatrix} -4 & 6 \\ -8 & 12 \end{pmatrix}$$

Выразим X , для этого обе части уравнения умножим на $-\frac{1}{2}$:

$$-\frac{1}{2} \cdot (-2X) = -\frac{1}{2} \cdot \begin{pmatrix} -4 & 6 \\ -8 & 12 \end{pmatrix}$$

$$X = -\frac{1}{2} \cdot \begin{pmatrix} -4 & 6 \\ -8 & 12 \end{pmatrix}$$

Все числа матрицы делятся на 2, поэтому уместно избавиться от дроби. А заодно и от «минуса». Делим каждый элемент матрицы на -2 :

$$X = \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix}$$

$$X = \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix}$$

Ответ:

Как выполнить проверку?

$$X = \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix}$$

Подставим найденное значение $X = \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix}$ в левую часть исходного уравнения и проведём упрощения:

$$\begin{aligned} \begin{pmatrix} 1 & -3 \\ 8 & 0 \end{pmatrix} - 2X &= \begin{pmatrix} 1 & -3 \\ 8 & 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 8 & 0 \end{pmatrix} - \begin{pmatrix} 4 & -6 \\ 8 & -12 \end{pmatrix} = \\ &= \begin{pmatrix} 1-4 & -3-(-6) \\ 8-8 & 0-(-12) \end{pmatrix} = \begin{pmatrix} -3 & 3 \\ 0 & 12 \end{pmatrix} = 3 \cdot \begin{pmatrix} -1 & 1 \\ 0 & 4 \end{pmatrix} \end{aligned}$$

Последним действием вынесли «тройку» из матрицы.

Получена правая часть исходного уравнения, значит решение найдено правильно.

Кстати, всегда ли матричное уравнение вообще имеет решение? Конечно не всегда. С

$$(1 \ 2) + X = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

ходу привожу простейшее доказательство:

Пример, который мы разобрали, элементарен, и, скажу честно, вероятность столкнуться с чем-то подобным на практике невелика. Поэтому перейдём к более содержательным

заданиям, которые с вероятностью, стремящейся к 100%, встретятся вам в реальной контрольной работе. Но прежде систематизируем общий ход решения:

Распространённый алгоритм решения матричного уравнения

Итак, на голову упал стандартный персонаж, состоящий из нескольких матриц, некоторых множителей и птицы счастья X .

На первом шаге уравнение приводится к одному из двух видов:

$AX = B$ либо $XA = B$, где A, B – известные матрицы.

Примечание: существует также третий вид: $AXB = C$, но в действительности он встречается крайне редко. Тем не менее, в конце статьи я рассмотрю данный случай.

Как привести уравнение к виду $AX = B$ или $XA = B$? Все действия вы видели в Примере №1 – это перенос матриц из части в часть, «упаковывание» множителей в матрицы, матричное сложение/вычитание.

На втором шаге необходимо выразить X или, выражаясь более академично, разрешить уравнение относительно X .

1) $AX = B$. Для того, чтобы разрешить данное уравнение относительно X , умножим обе его части на A^{-1} слева (здесь и далее предполагаем, что обратная матрица существует):

$$A^{-1}AX = A^{-1}B$$

!!! Внимание! Произведение матриц не перестановочно, поэтому критически важно, с какой стороны проводить умножение.

По свойству матричных операций: $A^{-1}A = E$, поэтому:

$$EX = A^{-1}B$$

Единичную матрицу можно убрать (см. урок Свойства операций над матрицами. Матричные выражения):

$$X = A^{-1}B$$

Чего и требовалось достичь. Матрица A^{-1} нам не известна.

2) $XA = B$. Умножаем обе части уравнения на A^{-1} справа:

$$XAA^{-1} = BA^{-1}$$

Согласно свойству матричных операций $AA^{-1} = E$, получаем:

$$XE = BA^{-1}$$

Единичную матрицу убираем:

$$X = BA^{-1}$$

Готово. Матрица A^{-1} нам опять же не известна.

Таким образом, на втором шаге решение выражается в виде $X = A^{-1}B$ либо в виде $X = BA^{-1}$. Поскольку обратной матрицы мы не знаем, то третий этап решения будет состоять в её нахождении. Это стандартная задача урока Как найти обратную матрицу?

На заключительном четвёртом шаге выполняем матричное умножение $A^{-1}B$ или BA^{-1} , и, собственно, получаем ответ.

После выполнения задания желательно провести проверку, впрочем, в большинстве случаев её требуется выполнить по условию задачи. Схема обыденна – необходимо подставить найденное значение X в исходное уравнение и убедиться в том, что «всё сойдётся».

Рассмотрим примеры решений уравнений обоих видов более подробно:

Решение матричного уравнения вида $AX = B$

...и добавить нечего =)

Пример 2

Решить матричное уравнение, выполнить проверку

$$\begin{pmatrix} 3 & 7 \\ 2 & 8 \end{pmatrix} \cdot X = \begin{pmatrix} 4 & 8 \\ 6 & 2 \end{pmatrix}$$

Решение: Уравнение уже имеет вид $AX = B$, поэтому никаких предварительных действий проводить не нужно.

Для разрешения уравнения относительно X умножим обе его части на A^{-1} слева:

$$A^{-1}AX = A^{-1}B$$

$$EX = A^{-1}B$$

$$X = A^{-1}B$$

Да-да, прямо так и пишем при оформлении решения. Хотя можно ограничиться единственной фразой: «Решение ищем в виде $X = A^{-1}B$ » – без всяких пояснений и вывода формулы $X = A^{-1}B$.

$$A = \begin{pmatrix} 3 & 7 \\ 2 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 8 \\ 6 & 2 \end{pmatrix}$$

Из условия известны матрицы A и B , однако, обратной матрицы A^{-1} мы не знаем. Придётся её найти.

Обратную матрицу найдем по формуле:

$A^{-1} = \frac{1}{|A|} \cdot A^*$, где A^* – транспонированная матрица алгебраических дополнений соответствующих элементов матрицы A .

$$|A| = \begin{vmatrix} 3 & 7 \\ 2 & 8 \end{vmatrix} = 3 \cdot 8 - 2 \cdot 7 = 24 - 14 = 10$$

$M = \begin{pmatrix} 8 & 2 \\ 7 & 3 \end{pmatrix}$ – матрица миноров соответствующих элементов матрицы A .

$A^* = \begin{pmatrix} 8 & -2 \\ -7 & 3 \end{pmatrix}$ – матрица алгебраических дополнений.

$A^{*T} = \begin{pmatrix} 8 & -7 \\ -2 & 3 \end{pmatrix}$ – транспонированная матрица алгебраических дополнений.

Таким образом, обратная матрица:

$$A^{-1} = \frac{1}{10} \cdot \begin{pmatrix} 8 & -7 \\ -2 & 3 \end{pmatrix}$$

На финише проводим матричное умножение и получаем решение:

$$X = A^{-1}B = \frac{1}{10} \cdot \begin{pmatrix} 8 & -7 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 8 \\ 6 & 2 \end{pmatrix} = \frac{1}{10} \cdot \begin{pmatrix} 8 \cdot 4 - 7 \cdot 6 & 8 \cdot 8 - 7 \cdot 2 \\ -2 \cdot 4 + 3 \cdot 6 & -2 \cdot 8 + 3 \cdot 2 \end{pmatrix} = \frac{1}{10} \cdot \begin{pmatrix} -10 & 50 \\ 10 & -10 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ 1 & -1 \end{pmatrix}$$

$$X = \begin{pmatrix} -1 & 5 \\ 1 & -1 \end{pmatrix}$$

Ответ:

Проверка: Подставим найденное значение X в левую часть исходного уравнения:

$$\begin{pmatrix} 3 & 7 \\ 2 & 8 \end{pmatrix} \cdot X = \begin{pmatrix} 3 & 7 \\ 2 & 8 \end{pmatrix} \cdot \begin{pmatrix} -1 & 5 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 \cdot (-1) + 7 \cdot 1 & 3 \cdot 5 + 7 \cdot (-1) \\ 2 \cdot (-1) + 8 \cdot 1 & 2 \cdot 5 + 8 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 6 & 2 \end{pmatrix}$$

Получена правая часть исходного уравнения. Таким образом, решение найдено правильно.

Следующая задача весьма любопытна, и некоторые из вас сделают для себя неожиданное открытие:

Пример 3

Решить матричное уравнение и сделать проверку:

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Решение: Неизвестная X расположена справа от матрицы, и уравнение, очевидно, сведётся к виду $AX = B$. Используем уже знакомые из Примера №1 действия:

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$$

Для разрешения уравнения относительно X умножим обе его части на A^{-1} слева:

$$A^{-1}AX = A^{-1}B$$

$$EX = A^{-1}B$$

$$X = A^{-1}B$$

Обратную матрицу найдем по формуле:

$A^{-1} = \frac{1}{|A|} \cdot A^T$, где A^T – транспонированная матрица алгебраических дополнений соответствующих элементов матрицы A .

$$|A| = \begin{vmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{vmatrix} = 3 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} - 2 \cdot \begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} = 3 \cdot (9 - 1) - 2 \cdot (6 - 2) + (2 - 6) = 24 - 8 - 4 = 12$$

$$M = \begin{pmatrix} 8 & 4 & -4 \\ 5 & 7 & -1 \\ -1 & 1 & 5 \end{pmatrix} \text{ – матрица миноров соответствующих элементов матрицы } A.$$

$$A_{\bullet} = \begin{pmatrix} 8 & -4 & -4 \\ -5 & 7 & 1 \\ -1 & -1 & 5 \end{pmatrix} \text{ – матрица алгебраических дополнений.}$$

$$A^T_{\bullet} = \begin{pmatrix} 8 & -5 & -1 \\ -4 & 7 & -1 \\ -4 & 1 & 5 \end{pmatrix} \text{ – транспонированная матрица алгебраических дополнений.}$$

Обратная матрица:

$$A^{-1} = \frac{1}{12} \begin{pmatrix} 8 & -5 & -1 \\ -4 & 7 & -1 \\ -4 & 1 & 5 \end{pmatrix}$$

Таким образом, решение уравнения:

$$X = A^{-1}B = \frac{1}{12} \begin{pmatrix} 8 & -5 & -1 \\ -4 & 7 & -1 \\ -4 & 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 8 \cdot 5 - 5 \cdot (-1) - 1 \cdot 4 \\ -4 \cdot 5 + 7 \cdot (-1) - 1 \cdot 4 \\ -4 \cdot 5 + 1 \cdot (-1) + 5 \cdot 4 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 41 \\ -31 \\ -1 \end{pmatrix}$$

$$X = \frac{1}{12} \begin{pmatrix} 41 \\ -31 \\ -1 \end{pmatrix}$$

Ответ:

$$X = \begin{pmatrix} \frac{41}{12} \\ -\frac{31}{12} \\ -\frac{1}{12} \end{pmatrix}.$$

Проверка: Подставим найденное значение X в левую часть исходного уравнения:

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot \frac{1}{12} \begin{pmatrix} 41 \\ -31 \\ -1 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 3 \cdot 41 + 2 \cdot (-31) + 1 \cdot (-1) \\ 2 \cdot 41 + 3 \cdot (-31) + 1 \cdot (-1) \\ 2 \cdot 41 + 1 \cdot (-31) + 3 \cdot (-1) \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} =$$

$$= \frac{1}{12} \begin{pmatrix} 60 \\ -12 \\ 48 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Получена правая часть исходного уравнения, таким образом, решение найдено верно.

Напоминаю технический приём, который мы рассмотрели на уроке Свойства операций над матрицами. Матричные выражения. После подстановки X в левую часть уравнения,

константа $\frac{1}{12}$ уютно расположилась между матрицами. В подобных случаях число необходимо вынести вперёд и разобраться с ним в самом конце – после матричного умножения.

А теперь остановимся вот на каком моменте.... Вернёмся к самому началу решения, когда

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$$

мы получили матричное уравнение в виде $\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot X = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$. Задача состояла в том,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

чтобы найти неизвестный вектор-столбец

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$$

Перепишем уравнение в виде $\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$ и в левой части умножим матрицы по обычному правилу:

$$\begin{pmatrix} 3x_1 + 2x_2 + x_3 \\ 2x_1 + 3x_2 + x_3 \\ 2x_1 + x_2 + 3x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix}$$

До боли знакомая картина => Две матрицы равны, когда равны их соответствующие элементы. Это система трёх линейных уравнений с тремя неизвестными:

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 5 \\ 2x_1 + 3x_2 + x_3 = -1 \\ 2x_1 + x_2 + 3x_3 = 4 \end{cases}$$

$$X = \frac{1}{12} \begin{pmatrix} 41 \\ -31 \\ -1 \end{pmatrix}$$

И полученный нами ответ представляет собой решение данной системы:

$$x_1 = \frac{41}{12}, \quad x_2 = -\frac{31}{12}, \quad x_3 = -\frac{1}{12}.$$

Таким образом, матричный метод решения системы – это, по сути, частный случай матричного уравнения.

Пример 4

Найти X из матричного уравнения:

$$\begin{pmatrix} 3 & 1 & -1 \\ -1 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} \cdot X + 5 \cdot \begin{pmatrix} 2 & -3 \\ 1 & 1 \\ -2 & -3 \end{pmatrix} - 2 \cdot \begin{pmatrix} 2 & -8 \\ 3 & 5 \\ -8 & -5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Проверить полученный результат.

Заметьте, что справа находится нулевая матрица а не ноль. Нулевая матрица для матриц – это аналог нуля для чисел. И её можно не записывать, после того, как вы что-нибудь перенесёте в правую часть.

Полное решение и примерный чистовой образец оформления задания в конце урока.

Решение матричного уравнения вида $XA = B$

Алгоритм решения точно такой же с некоторыми содержательными и техническими отличиями:

Пример 5

Решить матричное уравнение, выполнить проверку найденного решения.

$$X \cdot \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 24 & 14 & 11 \end{pmatrix}$$

Решение: Уравнение имеет готовый вид $XA = B$, что позволяет сразу же заняться «иксом».

Для разрешения уравнения относительно X умножим обе его части на A^{-1} справа:

$$XAA^{-1} = BA^{-1}$$

$$XE = BA^{-1}$$

$$X = BA^{-1}$$

При оформлении можно записать и короче: «Решение ищем в виде $X = BA^{-1}$ ».

$$A = \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \quad A^{-1} = \frac{1}{|A|} \cdot A^T, \quad \text{где } A^T -$$

Матрица «бэ» известна. Берём матрицу транспонированную матрицу алгебраических дополнений соответствующих элементов матрицы A .

$$|A| = \begin{vmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{vmatrix} = 9 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} - 7 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} + 6 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 9 \cdot (1-2) - 7 \cdot (1-2) + 6 \cdot (1-1) = -9 + 7 + 0 = -2$$

$$M = \begin{pmatrix} -1 & -1 & 0 \\ 1 & 3 & 2 \\ 8 & 12 & 2 \end{pmatrix} - \text{матрица миноров соответствующих элементов матрицы } A.$$

$$A_{\bullet} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 3 & -2 \\ 8 & -12 & 2 \end{pmatrix} - \text{матрица алгебраических дополнений.}$$

$$A_{\bullet}^T = \begin{pmatrix} -1 & -1 & 8 \\ 1 & 3 & -12 \\ 0 & -2 & 2 \end{pmatrix} - \text{транспонированная матрица алгебраических дополнений.}$$

Таким образом, обратная матрица:

$$A^{-1} = \frac{1}{-2} \begin{pmatrix} -1 & -1 & 8 \\ 1 & 3 & -12 \\ 0 & -2 & 2 \end{pmatrix}$$

Находим решение, при этом **не забываем про порядок умножения матриц**, обратная матрица едет во втором вагоне:

$$\begin{aligned} X &= BA^{-1} = \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 24 & 14 & 11 \end{pmatrix} \cdot \left(-\frac{1}{2}\right) \begin{pmatrix} -1 & -1 & 8 \\ 1 & 3 & -12 \\ 0 & -2 & 2 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 24 & 14 & 11 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 & 8 \\ 1 & 3 & -12 \\ 0 & -2 & 2 \end{pmatrix} = \\ &= -\frac{1}{2} \begin{pmatrix} 2 \cdot (-1) + 0 \cdot 1 - 2 \cdot 0 & 2 \cdot (-1) + 0 \cdot 3 - 2 \cdot (-2) & 2 \cdot 8 + 0 \cdot (-12) - 2 \cdot 2 \\ 18 \cdot (-1) + 12 \cdot 1 + 9 \cdot 0 & 18 \cdot (-1) + 12 \cdot 3 + 9 \cdot (-2) & 18 \cdot 8 + 12 \cdot (-12) + 9 \cdot 2 \\ 24 \cdot (-1) + 14 \cdot 1 + 11 \cdot 0 & 24 \cdot (-1) + 14 \cdot 3 + 11 \cdot (-2) & 24 \cdot 8 + 14 \cdot (-12) + 11 \cdot 2 \end{pmatrix} = \\ &= -\frac{1}{2} \begin{pmatrix} -2 & 2 & 12 \\ -6 & 0 & 18 \\ -10 & -4 & 46 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -6 \\ 3 & 0 & -9 \\ 5 & 2 & -23 \end{pmatrix} \\ X &= \begin{pmatrix} 1 & -1 & -6 \\ 3 & 0 & -9 \\ 5 & 2 & -23 \end{pmatrix} \end{aligned}$$

Ответ:

Проверка: Подставим найденное значение X в левую часть исходного уравнения:

$$\begin{aligned} XA &= X \cdot \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -6 \\ 3 & 0 & -9 \\ 5 & 2 & -23 \end{pmatrix} \cdot \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 1 \cdot 9 - 1 \cdot 1 - 6 \cdot 1 & 1 \cdot 7 - 1 \cdot 1 - 6 \cdot 1 & 1 \cdot 6 - 1 \cdot 2 - 6 \cdot 1 \\ 3 \cdot 9 + 0 \cdot 1 - 9 \cdot 1 & 3 \cdot 7 + 0 \cdot 1 - 9 \cdot 1 & 3 \cdot 6 + 0 \cdot 2 - 9 \cdot 1 \\ 5 \cdot 9 + 2 \cdot 1 - 23 \cdot 1 & 5 \cdot 7 + 2 \cdot 1 - 23 \cdot 1 & 5 \cdot 6 + 2 \cdot 2 - 23 \cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 24 & 14 & 11 \end{pmatrix} = B \end{aligned}$$

Получена правая часть исходного уравнения. Таким образом, решение найдено правильно.

Усложним задание:

Пример 6

Решить матричное уравнение, сделать проверку:

$$X \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} + 2 \begin{pmatrix} 3 & 1 & 2 \\ 4 & -1 & 3 \end{pmatrix} = 5 \begin{pmatrix} 0 & -1 & 2 \\ -1 & 2 & 0 \end{pmatrix}$$

Решение: Незнакомец расположился слева от матрицы, поэтому уравнение сводится к виду $XA = B$. Упаковываем множители, переносим свободную матрицу в правую часть и выполняем вычитание матриц:

$$X \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} + \begin{pmatrix} 6 & 2 & 4 \\ 8 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & -5 & 10 \\ -5 & 10 & 0 \end{pmatrix}$$

$$X \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 0 & -5 & 10 \\ -5 & 10 & 0 \end{pmatrix} - \begin{pmatrix} 6 & 2 & 4 \\ 8 & -2 & 6 \end{pmatrix}$$

$$X \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} -6 & -7 & 6 \\ -13 & 12 & -6 \end{pmatrix}$$

Для разрешения уравнения относительно X умножим обе его части на A^{-1} **справа**:

$$XAA^{-1} = BA^{-1}$$

$$XE = BA^{-1}$$

$$X = BA^{-1}$$

Обратную матрицу найдем по формуле:

$A^{-1} = \frac{1}{|A|} \cdot A^T$, где A^T – транспонированная матрица алгебраических дополнений соответствующих элементов матрицы A .

$$|A| = \begin{vmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 4 & 2 \\ 2 & -2 \end{vmatrix} - \begin{vmatrix} -1 & 2 \\ 1 & -2 \end{vmatrix} + 3 \cdot \begin{vmatrix} -1 & 4 \\ 1 & 2 \end{vmatrix} =$$

$$= 2 \cdot (-8 - 4) - (2 - 2) + 3 \cdot (-2 - 4) = -24 - 0 - 18 = -42$$

$$M = \begin{pmatrix} -12 & 0 & -6 \\ -8 & -7 & 3 \\ -10 & 7 & 9 \end{pmatrix}$$

– матрица миноров соответствующих элементов матрицы A .

$$A^* = \begin{pmatrix} -12 & 0 & -6 \\ 8 & -7 & -3 \\ -10 & -7 & 9 \end{pmatrix}$$

– матрица алгебраических дополнений.

$$A^T = \begin{pmatrix} -12 & 8 & -10 \\ 0 & -7 & -7 \\ -6 & -3 & 9 \end{pmatrix}$$

– транспонированная матрица алгебраических дополнений.

Обратная матрица:

$$A^{-1} = \frac{1}{-42} \begin{pmatrix} -12 & 8 & -10 \\ 0 & -7 & -7 \\ -6 & -3 & 9 \end{pmatrix} = \frac{1}{42} \begin{pmatrix} 12 & -8 & 10 \\ 0 & 7 & 7 \\ 6 & 3 & -9 \end{pmatrix}$$

Здесь целесообразно внести минус в матрицу. Находим решение:

$$\begin{aligned} X &= BA^{-1} = \begin{pmatrix} -6 & -7 & 6 \\ -13 & 12 & -6 \end{pmatrix} \cdot \frac{1}{42} \begin{pmatrix} 12 & -8 & 10 \\ 0 & 7 & 7 \\ 6 & 3 & -9 \end{pmatrix} = \frac{1}{42} \begin{pmatrix} -6 & -7 & 6 \\ -13 & 12 & -6 \end{pmatrix} \cdot \begin{pmatrix} 12 & -8 & 10 \\ 0 & 7 & 7 \\ 6 & 3 & -9 \end{pmatrix} = \\ &= \frac{1}{42} \begin{pmatrix} -6 \cdot 12 - 7 \cdot 0 + 6 \cdot 6 & -6 \cdot (-8) - 7 \cdot 7 + 6 \cdot 3 & -6 \cdot 10 - 7 \cdot 7 + 6 \cdot (-9) \\ -13 \cdot 12 + 12 \cdot 0 - 6 \cdot 6 & -13 \cdot (-8) + 12 \cdot 7 - 6 \cdot 3 & -13 \cdot 10 + 12 \cdot 7 - 6 \cdot (-9) \end{pmatrix} = \\ &= \frac{1}{42} \begin{pmatrix} -36 & 17 & -163 \\ -192 & 170 & 8 \end{pmatrix} \\ X &= \frac{1}{42} \begin{pmatrix} -36 & 17 & -163 \\ -192 & 170 & 8 \end{pmatrix} \end{aligned}$$

Ответ:

Проверка: Подставим найденное значение X в левую часть исходного уравнения:

$$\begin{aligned} X \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} + 2 \begin{pmatrix} 3 & 1 & 2 \\ 4 & -1 & 3 \end{pmatrix} &= \frac{1}{42} \begin{pmatrix} -36 & 17 & -163 \\ -192 & 170 & 8 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & 4 & 2 \\ 1 & 2 & -2 \end{pmatrix} + 2 \begin{pmatrix} 3 & 1 & 2 \\ 4 & -1 & 3 \end{pmatrix} = \\ &= \frac{1}{42} \begin{pmatrix} -36 \cdot 2 + 17 \cdot (-1) - 163 \cdot 1 & -36 \cdot 1 + 17 \cdot 4 - 163 \cdot 2 & -36 \cdot 3 + 17 \cdot 2 - 163 \cdot (-2) \\ -192 \cdot 2 + 170 \cdot (-1) + 8 \cdot 1 & -192 \cdot 1 + 170 \cdot 4 + 8 \cdot 2 & -192 \cdot 3 + 170 \cdot 2 + 8 \cdot (-2) \end{pmatrix} + \begin{pmatrix} 6 & 2 & 4 \\ 8 & -2 & 6 \end{pmatrix} = \\ &= \frac{1}{42} \begin{pmatrix} -252 & -294 & 252 \\ -546 & 504 & -252 \end{pmatrix} + \begin{pmatrix} 6 & 2 & 4 \\ 8 & -2 & 6 \end{pmatrix} = \\ &= \begin{pmatrix} -6 & -7 & 6 \\ -13 & 12 & -6 \end{pmatrix} + \begin{pmatrix} 6 & 2 & 4 \\ 8 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & -5 & 10 \\ -5 & 10 & 0 \end{pmatrix} = 5 \begin{pmatrix} 0 & -1 & 2 \\ -1 & 2 & 0 \end{pmatrix} \end{aligned}$$

Получена правая часть исходного уравнения, таким образом, решение найдено верно.

Пример 7

Решить матричное уравнение и сделать проверку:

$$2 \begin{pmatrix} 4 & 2 & -1 \end{pmatrix} - X \cdot \begin{pmatrix} -1 & -4 & -2 \\ 1 & -1 & 1 \\ 2 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

Это пример для самостоятельного решения. В заключение коротко рассмотрим ещё один тип матричного уравнения, который практически не встречается: $AXB = C$, где A, B, C – известные матрицы. То есть, наш партизан залёг между двумя матрицами.

Разрешим данное уравнение относительно X . Сначала умножим обе части на A^{-1} слева:

$$A^{-1}AXB = A^{-1}C$$

$$EXB = A^{-1}C$$

$$XB = A^{-1}C$$

Теперь умножим обе части на B^{-1} **справа**:

$$XB B^{-1} = A^{-1} C B^{-1}$$

$$XE = A^{-1} C B^{-1}$$

$$X = A^{-1} C B^{-1}$$

Готового примера у себя в коллекции я не нашёл, но сейчас всё равно что-нибудь подберу из этой оперы.... Вот:

$$\begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix} \cdot X \cdot \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 10 \\ -2 & -1 \end{pmatrix}$$

Да, работёнки здесь побольше. Раза в два. Как решить данное уравнение?

– для матрицы $A = \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix}$ находим обратную матрицу A^{-1} ;

– для матрицы $B = \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix}$ находим обратную матрицу B^{-1} ;

– перемножаем три матрицы $X = A^{-1} C B^{-1}$ ответ: $X = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$.

Решения и ответы:

Пример 4: Решение: Приведем уравнение к виду $AX = B$:

$$\begin{pmatrix} 3 & 1 & -1 \\ -1 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} \cdot X = -5 \cdot \begin{pmatrix} 2 & -3 \\ 1 & 1 \\ -2 & -3 \end{pmatrix} + 2 \cdot \begin{pmatrix} 2 & -8 \\ 3 & 5 \\ -8 & -5 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 & -1 \\ -1 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} \cdot X = \begin{pmatrix} -10 & 15 \\ -5 & -5 \\ 10 & 15 \end{pmatrix} + \begin{pmatrix} 4 & -16 \\ 6 & 10 \\ -16 & -10 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 & -1 \\ -1 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} \cdot X = \begin{pmatrix} -6 & -1 \\ 1 & 5 \\ -6 & 5 \end{pmatrix}$$

Для разрешения уравнения относительно X умножим обе его части на A^{-1} слева:

$$A^{-1} A X = A^{-1} B$$

$$E X = A^{-1} B$$

$$X = A^{-1} B$$

Обратную матрицу найдем по формуле:

$A^{-1} = \frac{1}{|A|} \cdot A^T$, где A^T – транспонированная матрица алгебраических дополнений соответствующих элементов матрицы A .

$$|A| = \begin{vmatrix} 3 & 1 & -1 \\ -1 & 1 & 2 \\ 1 & 3 & 1 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} =$$

$$= 3 \cdot (1 - 6) - (-1 - 2) - (-3 - 1) = -15 + 3 + 4 = -8$$

2.4 Практическое занятие №2(1 час).

Тема: «Извлечение корня m-ной степени»

2.4.1 Задание для работы:

2.4.2 Краткое описание проводимого занятия

1. Извлечение корня из матрицы

Матрица B такая, что $B^k = A$, называется Алгебраическим корнем степени k из матрицы A и обозначается $\sqrt[k]{A}$.

Степени и корни матрицы обладают теми же свойствами, что степени и корни чисел: $A^k \cdot A^m = A^{k+m}$, $(A^n)^k = A^{n \cdot k}$, $\sqrt[k]{A^n} = (\sqrt[k]{A})^n$ и т. д.

2. Многочлен от матрицы

Пусть $P_k(x) = a_0 + a_1x + a_2x^2 + \dots + a_kx^k$ ($a_k \neq 0$) – многочлен степени k от скалярной переменной x . Такой многочлен называют скалярным многочленом. Если A – квадратная матрица, то $P_k(A) = a_0 \cdot I + a_1 \cdot A + a_2 \cdot A^2 + \dots + a_k \cdot A^k$ называется многочленом от матрицы. Очевидно, $P_k(A)$ – квадратная матрица того же порядка, что и матрица A .

Матрица A и любой её многочлен перестановочны, т. е. если $B = P_k(A)$, то $A \cdot B = B \cdot A$. Для большинства матриц справедливо и обратное утверждение: если квадратные матрицы A и B перестановочны, то одна из них (а чаще – каждая из них) является многочленом от другой, причём степень многочлена должна быть меньше, чем порядок этих матриц. Например, диагональные матрицы перестановочны и по операции сложения, и по операции умножения, т. е. если A и B – диагональные матрицы, то $A + B = B + A$, $A \cdot B = B \cdot A$.

Введём обозначение для диагональной матрицы:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$$

Тогда

$$A^k = \text{diag}(a_{11}^k, a_{22}^k, \dots, a_{nn}^k), \quad P_k(A) = \text{diag}(P_k(a_{11}), P_k(a_{22}), \dots, P_k(a_{nn})).$$

3. Трансцендентные функции

Из теории рядов известны разложения в ряд Маклорена функций $y = e^x$, $y = \sin x$, $y = \cos x$ и т. д. Например,

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^k}{k!} + \dots = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (27)$$

Разложение (27) можно представить в виде

$$e^x = \lim_{k \rightarrow \infty} \left(1 + x + \frac{x^2}{2!} + \dots + \frac{x^k}{k!} \right) \quad (28)$$

Или в виде 2-го замечательного предела:

$$e^x = \lim_{k \rightarrow \infty} \left(1 + \frac{x}{k} \right)^k. \quad (29)$$

Из (28) и (29) переходя к матрицам, получим:

$$e^A = \lim_{k \rightarrow \infty} \left(I + A + \frac{1}{2!} A^2 + \dots + \frac{1}{k!} A^k \right) \quad \text{и} \quad e^A = \lim_{k \rightarrow \infty} \left(I + \frac{1}{k} \cdot A \right)^k \quad (30)$$

Формулами (30) можно пользоваться, если эти пределы существуют. Однако могут возникнуть вычислительные сложности при возведении матрицы A в степень.

Для диагональных матриц формулы (30) применимы всегда. К сожалению, в реальных задачах диагональные матрицы практически не встречаются. Однако, большую роль в технических приложениях играют матрицы, имеющие различные собственные значения, и симметрические матрицы, которые в результате преобразования подобия всегда могут быть приведены к диагональному виду.

Пусть в результате преобразования подобия из матрицы A получена диагональная матрица $D(A)$; $D(A) = C^{-1} \cdot A \cdot C$, откуда

$$A = C \cdot D(A) \cdot C^{-1}. \quad (31)$$

Представление (31) матрицы A обладает следующим свойством:

$$A^2 = (C \cdot D(A) \cdot C^{-1}) \cdot (C \cdot D(A) \cdot C^{-1}) = C \cdot D^2(A) \cdot C^{-1}, \quad A^3 = C \cdot D^3(A) \cdot C^{-1} \dots,$$

т. е. оно сохраняет свой вид при возведении A в любую степень.

Если $D(A) = \text{diag}(d_1, d_2, \dots, d_n)$, то в случае многочлена от матрицы

$$P_k(A) = C \cdot \text{diag}(P_k(d_1), P_k(d_2), \dots, P_k(d_n)) \cdot C^{-1}.$$

В общем случае, если значения $f(d_i) \ (i = \overline{1, n})$ существуют, то:

$$f(A) = C \cdot \text{diag}(f(d_1), f(d_2), \dots, f(d_n)) \cdot C^{-1}. \quad (32)$$

Например, $e^A = C \cdot \text{diag}(e^{d_1}, e^{d_2}, \dots, e^{d_n}) \cdot C^{-1}$ и при этом $(e^A)^m = e^{mA}$,
но $e^{(A+B)} = e^A \cdot e^B$ только, если матрицы A и B перестановочны.

Аналогично,

$$\sin A = C \cdot \text{diag}(\sin d_1, \sin d_2, \dots, \sin d_n) \cdot C^{-1},$$

$$\cos A = C \cdot \text{diag}(\cos d_1, \cos d_2, \dots, \cos d_n) \cdot C^{-1}.$$

Если матрица $\cos A$ невырожденная, то можно найти матричный тангенс $\operatorname{tg} A = \sin A \cdot (\cos A)^{-1}$ и матричный котангенс $\operatorname{ctg} A = \cos A \cdot (\sin A)^{-1}$, если матрица $\sin A$ невырожденная.

Справедливо и основное тригонометрическое тождество: $(\sin A)^2 + (\cos A)^2 = I$.

Пример 22. Найти $\sqrt{A}$, если $A = \begin{pmatrix} 6 & 3 \\ 2 & 7 \end{pmatrix}$.

Решение. Диагонализация матриц возможна с помощью преобразования подобия: $\bar{A} = C^{-1} \cdot A \cdot C$, где C – матрица из собственных векторов матрицы A , $\bar{A}$ – диагональная матрица, полученная из A , т. е. $\bar{A} = D(A)$. Найдём собственные значения матрицы A :

$$\begin{vmatrix} 6 - \lambda & 3 \\ 2 & 7 - \lambda \end{vmatrix} = 0, \lambda^2 - 13\lambda + 36 = 0, \lambda_1 = 4, \lambda_2 = 9$$

Соответствующие собственные векторы:

$$1) \quad \underline{\lambda_1 = 4}: \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{matrix} 2x_1 = -3x_2 \\ x_1 = 3, x_2 = -2 \end{matrix} \Rightarrow X^{(1)} = \begin{pmatrix} 3 \\ -2 \end{pmatrix},$$

$$2) \quad \underline{\lambda_2 = 9}: \begin{pmatrix} -3 & 3 \\ 2 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_1 = x_2 \Rightarrow X^{(2)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$C = \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix}, C^{-1} = \frac{1}{5} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$$

Следовательно,

$$D(A) = C^{-1} \cdot A \cdot C = \frac{1}{5} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 6 & 3 \\ 2 & 7 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} = \operatorname{diag}(4, 9).$$

Так как по формуле (31) $A = C \cdot D(A) \cdot C^{-1}$, То по формуле (32) получим один из четырёх возможных ответов:

$$\sqrt{A} = C \cdot \operatorname{diag}(\sqrt{4}, \sqrt{9}) \cdot C^{-1} = \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 12 & 3 \\ 2 & 13 \end{pmatrix},$$

$$\sqrt{A} = \begin{pmatrix} 12/5 & 3/5 \\ 2/5 & 13/5 \end{pmatrix}.$$

И окончательно,

Можно выполнить проверку полученного результата, воспользовавшись определением корня из матрицы A . По определению $(\sqrt{A})^2 = \sqrt{A} \cdot \sqrt{A} = A$,

$$\text{тогда } (\sqrt{A})^2 = \frac{1}{5} \begin{pmatrix} 12 & 3 \\ 2 & 13 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 12 & 3 \\ 2 & 13 \end{pmatrix} = \frac{1}{25} \begin{pmatrix} 150 & 75 \\ 50 & 175 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 2 & 7 \end{pmatrix} = A.$$

$$\sqrt{A} = \begin{pmatrix} 12/5 & 3/5 \\ 2/5 & 13/5 \end{pmatrix}.$$

Ответ:

Пример 23. Найти e^A , Если $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$.

Решение. Для нахождения собственных значений матрицы A составим характеристическое уравнение: $\begin{vmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = 0, \lambda^2 - 5\lambda + 4 = 0$, откуда $\lambda_1 = 1, \lambda_2 = 4$.
Найдём собственные векторы:

$$1) \quad \underline{\lambda_1 = 1}: \quad \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{matrix} x_1 = -2x_2 \\ x_1 = 2, x_2 = -1 \end{matrix} \Rightarrow X^{(1)} = \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

$$2) \quad \underline{\lambda_2 = 4}: \quad \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_1 = x_2 \Rightarrow X^{(2)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Теперь можем составить матрицу C и найти C^{-1} :

$$C = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \quad C^{-1} = \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

Так как $D(A) = \text{diag}(1, 4)$, то по формуле (32)

$$\begin{aligned} e^A &= C \cdot \text{diag}(e, e^4) \cdot C^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} e & 0 \\ 0 & e^4 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \\ &= \begin{pmatrix} 2e & e^4 \\ -e & e^4 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2e + e^4 & -2e + 2e^4 \\ -e + e^4 & e + 2e^4 \end{pmatrix} = e \begin{pmatrix} 2/3 & -2/3 \\ -1/3 & 1/3 \end{pmatrix} + e^4 \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix}. \\ e^A &= e \begin{pmatrix} 2/3 & -2/3 \\ -1/3 & 1/3 \end{pmatrix} + e^4 \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix}. \end{aligned}$$

Ответ:

2.5 Практическое занятие №2(1 час).

Тема: «Решение скалярного уравнения»

2.5.1 Задание для работы:

2.5.2 Краткое описание проводимого занятия

Скалярное уравнение $f(X)=0$

Рассмотрим сначала уравнение

$$g(X) = 0, \quad (33)$$

где

$$g(\lambda) = (\lambda - \lambda_1)^{a_1} (\lambda - \lambda_2)^{a_2} \dots (\lambda - \lambda_n)^{a_n}$$

– заданный многочлен переменной λ , а X – искомая квадратная матрица порядка n . Так как минимальный многочлен матрицы X , т. е. первый инвариантный многочлен $i_1(\lambda)$,

должен быть делителем многочлена $g(\lambda)$, то элементарные делители матрицы X должны иметь следующий вид:

$$(\lambda - \lambda_{i_1})^{p_{i_1}}, (\lambda - \lambda_{i_2})^{p_{i_2}}, \dots, (\lambda - \lambda_{i_v})^{p_{i_v}} \left(\begin{array}{l} i_1, i_2, \dots, i_v = 1, 2, \dots, h, \\ p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_v} \leq a_{i_v}, \\ p_{i_1} + p_{i_2} + \dots + p_{i_v} = n \end{array} \right)$$

(среди индексов $i_1, i_2, \dots, i_v$ могут быть и равные, n – заданный порядок искомой матрицы X).

Искомая матрица X представится в виде

$$X = T \left\{ \lambda_{i_1} E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_v} E^{(p_{i_v})} + H^{(p_{i_v})} \right\} T^{-1}, \quad (34)$$

где T – произвольная неособенная матрица порядка n . Множество решений уравнения (33) с заданным порядком искомой матрицы распадается согласно формуле (34) на конечное число классов подобных между собой матриц.

Пример 1. Дано уравнение

$$X^m = 0. \quad (35)$$

Если некоторая степень матрицы равна нулю, то матрица называется нильпотентной. Наименьший из показателей, при которых степень матрицы равна нулю, называется индексом нильпотентности данной матрицы.

Очевидно, решениями уравнения (35) являются все нильпотентные матрицы с индексом нильпотентности $\mu \leq m$. Формула, охватывающая все решения данного порядка n , выглядит так:

$$X = T \left\{ H^{(p_1)}, H^{(p_2)}, \dots, H^{(p_v)} \right\} T^{-1} \left(\begin{array}{l} p_1, p_2, \dots, p_v \leq m, \\ p_1 + p_2 + \dots + p_v = n \end{array} \right) \quad (36)$$

(T – произвольная неособенная матрица).

Пример 2. Дано уравнение

$$X^2 = X. \quad (37)$$

Матрица, удовлетворяющая этому уравнению, называется идемпотентной.

Элементарными делителями идемпотентной матрицы могут быть только λ либо $\lambda - 1$. Поэтому идемпотентную матрицу можно определить как матрицу простой структуры (т. е. приводящуюся к диагональной форме) с характеристическими числами, равными нулю или единице. Формула, охватывающая все идемпотентные матрицы данного порядка, имеет вид

$$X = T \left\{ \underbrace{1, 1, \dots, 1, 0, \dots, 0}_n \right\} T^{-1}, \quad (38)$$

где T – произвольная неособенная матрица порядка n .

Рассмотрим теперь более общее уравнение

$$f(X) = 0, \quad (39)$$

где $f(\lambda)$ – регулярная функция в некоторой области G плоскости комплексного

аргумента λ . От искомого решения $X = \|x_{ik}\|_n^2$ будем требовать,

чтобы характеристические числа его принадлежали области G . Выпишем все нули

функции $f(\lambda)$, лежащие в области G , и их кратности:

$$\lambda_1, \lambda_2, \dots,$$

$$a_1, a_2, \dots$$

Как и в предыдущем случае, каждый элементарный делитель матрицы X должен иметь вид

$$(\lambda - \lambda_i)^{p_i} \quad (p_i \leq a_i)$$

и потому

$$X = T \left\{ \lambda_{i_1} E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_v} E^{(p_{i_v})} + H^{(p_{i_v})} \right\} T^{-1}$$

$$\left(\begin{array}{l} i_1, i_2, \dots, i_v = 1, 2, \dots; p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_v} \leq a_{i_v}; \\ p_{i_1} + p_{i_2} + \dots + p_{i_v} = n \end{array} \right) \quad (40)$$

(T – произвольная неособенная матрица).

2.6 Практическое занятие №2(1 час).

Тема: «Нахождение логарифма матрицы»

2.6.1 Задание для работы:

2.6.2 Краткое описание проводимого занятия

. Рассмотрим матричное уравнение

$$e^X = A. \quad (90)$$

Все решения этого уравнения будем называть логарифмами (натуральными) матрицы A и обозначать через $\ln A$.

Характеристические числа λ_j матрицы A связаны с характеристическими числами ξ_j

матрицы X формулой $\lambda_j = e^{\xi_j}$; поэтому, если уравнение (90) имеет решение, то все характеристические числа матрицы A отличны от нуля и матрица A является

неособенной ($|A| \neq 0$). Таким образом, условие $|A| \neq 0$ является необходимым для существования решения уравнения (90). Ниже мы увидим, что это условие является и достаточным.

Итак, пусть $|A| \neq 0$. Выпишем элементарные делители матрицы A :

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (\lambda_1 \lambda_2 \dots \lambda_u \neq 0, p_1 + p_2 + \dots + p_u = n). \quad (91)$$

В соответствии с этими элементарными делителями приведем матрицу A к нормальной жордановой форме:

$$A = U \tilde{A} U^{-1} = U \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \right\} U^{-1}. \quad (92)$$

Так как производная от функции e^{ξ} отлична от нуля при всех значениях ξ , то (см. гл. VI, стр. 159) при переходе от матрицы X к матрице $A = e^X$ элементарные делители не расщепляются, т. е. матрица X имеет элементарные делители

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}, \quad (93)$$

где $e^{\xi_j} = \lambda_j$ ($j = 1, 2, \dots, u$), т. е. ξ_j есть одно из значений $\ln \lambda_j$ ($j = 1, 2, \dots, u$).

Возьмем в плоскости комплексного переменного λ круг с центром в точке λ_j радиуса $< |\lambda_j|$ и обозначим через $f_j(\lambda) = \ln \lambda$ ту из ветвей функции $\ln \lambda$ в

рассматриваемом круге, которая в точке λ_j принимает значение, равно характеристическому числу ξ_j матрицы X ($j = 1, 2, \dots, u$). После этого полагаем:

$$\ln(\lambda_j E^{(p_1)} + H^{(p_1)}) = f_j(\lambda_j E^{(p_1)} + H^{(p_1)}) = \ln \lambda_j E^{(p_1)} + \lambda_j^{-1} H^{(p_1)} + \dots \quad (94)$$

Так как производная от $\ln \lambda$ нигде не обращается в нуль (в конечной части плоскости λ), то матрица (94) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_1}$. В силу этого квазидиагональная матрица

$$\left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \right\} \quad (95)$$

имеет те же элементарные делители, что и искомая матрица X . Поэтому существует

такая матрица T ($|T| \neq 0$), что

$$X = T \left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \right\} T^{-1} \quad (96)$$

Для определения матрицы T заметим, что

$$A = e^X = T \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \right\} T^{-1} \quad (97)$$

Сопоставляя (97) с (92), находим:

$$T = UX_{\tilde{A}}, \quad (98)$$

где $X_{\tilde{A}}$ – произвольная матрица, перестановочная с матрицей $\tilde{A}$. Подставляя выражение для T из (98) в (96), получим общую формулу, охватывающую все логарифмы матрицы:

$$X = UX_{\tilde{A}} \left\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \right\} X_{\tilde{A}}^{-1} U^{-1} \quad (99)$$

Замечание. Если все элементарные делители матрицы A взаимно просты, то в правой части формулы (99) можно выбросить множители $X_{\tilde{A}}$ и $X_{\tilde{A}}^{-1}$ (см. аналогичное замечание на стр. 213).

2. Выясним, когда вещественная неособенная матрица A имеет вещественный логарифм X . Пусть искомая матрица имеет несколько элементарных делителей,

отвечающих характеристическому числу вида $\rho + i\pi: (\lambda - \rho - i\pi)^{q_1}, \dots, (\lambda - \rho - i\pi)^{q_s}$.

Поскольку матрица X вещественна, то она имеет и сопряженные элементарные

делители: $(\lambda - \rho + i\pi)^{q_1}, \dots, (\lambda - \rho + i\pi)^{q_s}$. При переходе от матрицы X к матрице A

элементарные делители не расщепляются, но характеристические числа $\rho + i\pi, \rho - i\pi$

заменяются в них числами $e^{\rho + i\pi} = -\mu, e^{\rho - i\pi} = -\mu$, где $\mu = e^{\rho} > 0$. Поэтому в системе

элементарных делителей матрицы A каждый элементарный делитель, соответствующий отрицательному характеристическому числу (если таковые существуют),

повторяется четное число раз. Докажем теперь, что это необходимое условие является и

достаточным, т. е. что вещественная неособенная матрица A тогда и только тогда имеет

вещественный логарифм X , когда у матрицы A либо совсем нет элементарных

делителей, соответствующих отрицательным характеристическим числам, либо каждый такой элементарный делитель повторяется четное число раз.

Действительно, пусть это условие выполнено. Тогда в квазидиагональной матрице (95) в

соответствии с формулой (94) в тех клетках, где λ_j вещественно и положительно, возьмем

для $\ln \lambda_k$ вещественное значение; если же в какой-либо клетке имеется комплексное λ_k , то найдется другая клетка такого же размера с $\lambda_{\bar{k}} = \overline{\lambda_k}$. В этих клетках возьмем комплексно сопряженные значения для $\ln \lambda_k$ и $\ln \lambda_{\bar{k}}$. Каждая же клетка по условию повторяется в (98) четное число раз с сохранением размера клетки. Тогда в половине этих клеток положим $\ln \lambda_k = \ln |\lambda_k| + i\pi$, а в другой половине возьмем $\ln \lambda_k = \ln |\lambda_k| - i\pi$. Тогда в квазидиагональной матрице (98) диагональные клетки либо будут вещественными, либо будут попарно комплексно сопряженными. Но такая квазидиагональная матрица всегда подобна вещественной матрице. Поэтому существует такая неособенная

матрица $T_1 (|T_1| \neq 0)$, что матрица

$$X_1 = T_1 \left\{ \ln \left(\lambda_1 E^{(p_1)} + H^{(p_1)} \right), \dots, \ln \left(\lambda_n E^{(p_n)} + H^{(p_n)} \right) \right\} T_1^{-1}$$

вещественна. Но тогда будет вещественной и матрица

$$A_1 = e^{X_1} = T_1 \left\{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_n E^{(p_n)} + H^{(p_n)} \right\} T_1^{-1}. \quad (100)$$

Сопоставляя формулу (100) с формулой (92), заключаем, что матрицы A и A_1 подобны между собой (поскольку они подобны одной и той же жордановой матрице). Но две подобные вещественные матрицы могут быть преобразованы друг в друга с помощью некоторой неособенной вещественной матрицы $W (|W| \neq 0)$:

$$A = W A_1 W^{-1} = W e^{X_1} W^{-1} = e^{W X_1 W^{-1}}.$$

Тогда матрица $X = W X_1 W^{-1}$ и будет искомым вещественным логарифмом матрицы A .

2.7 Практическое занятие №2(1 час).

Тема: «Неравенства Неймана-Хорна»

2.7.1 Задание для работы:

2.7.2 Краткое описание проводимого занятия

Пусть $\mathbf{A}$ – линейный оператор, действующий в $\mathbb{R}$ -мерном унитарном пространстве $\mathbf{R}$.

Собственные числа неотрицательного эрмитова оператора $\sqrt{\mathbf{A}^* \mathbf{A}}$ (см. стр. 249) принято называть сингулярными числами оператора $\mathbf{A}$.

В настоящем параграфе мы установим неравенства, связывающие сингулярные числа произведения двух операторов с сингулярными числами сомножителей.

Пусть $y_1, y_2, \dots, y_m$ и $z_1, z_2, \dots, z_m$ – два набора векторов из $\mathbf{R}$. Введем сокращенное обозначение для определителя порядка m , связанного с данными наборами:

$$[(y_i, z_j)] = \begin{vmatrix} (y_1, z_1) & (y_2, z_1) & \dots & (y_m, z_1) \\ (y_1, z_2) & (y_2, z_2) & \dots & (y_m, z_2) \\ \dots & \dots & \dots & \dots \\ (y_1, z_m) & (y_2, z_m) & \dots & (y_m, z_m) \end{vmatrix}. \quad (25)$$

Рассмотрим далее неотрицательный эрмитов оператор $\mathbf{H}$, действующий в $\mathbf{R}$. Собственные значения оператора $\mathbf{H}$ занумеруем в убывающем порядке:

$$h_1 \geq h_2 \geq \dots \geq h_n \geq 0. \quad (26)$$

Справедливо следующее предложение, принадлежащее А. Хорну ([191b]):

Лемма 3. Пусть

$$x_1, x_2, \dots, x_m \quad (m \leq n) \quad (27)$$

– произвольный набор векторов из R . Тогда

$$[(Hx_i, x_j)] \leq h_1 h_2 \dots h_m [(x_i, x_j)]. \quad (28)$$

Для доказательства

рассмотрим ортонормированный базис собственных векторов оператора H :

$$e_1, e_2, \dots, e_n \quad (29)$$

и разложим каждый из векторов x_i ($i = 1, 2, \dots, m$) по базису (29). Вычисляя скалярное произведение, получаем:

$$(Hx_i, x_j) = \sum_{s=1}^n h_s (x_i, e_s) (\overline{x_j, e_s}) = \sum_{s=1}^n h_s (x_i, e_s) (e_s, x_j), \quad i, j = 1, 2, \dots, m \quad (30)$$

Равенство (30) позволяет рассматривать матрицу определителя $[(Hx_i, x_j)]$ как результат умножения двух прямоугольных матриц размеров $m \times n$ и $n \times m$.

Разлагая определитель по формуле Бине-Коши (см. стр. 20), получаем в принятых обозначениях для определителей:

$$[(Hx_i, x_j)] = \sum_{1 \leq s_1 < s_2 < \dots < s_m \leq n} [(x_i, h_{s_1} e_{s_1})] [(e_{s_1}, x_j)] \quad (31)$$

Здесь

$$[(x_i, h_s e_s)] = \begin{vmatrix} (x_1, h_{s_1} e_{s_1}) & \dots & (x_m, h_{s_1} e_{s_1}) \\ (x_1, h_{s_2} e_{s_2}) & \dots & (x_m, h_{s_2} e_{s_2}) \\ \dots & \dots & \dots \\ (x_1, h_{s_m} e_{s_m}) & \dots & (x_m, h_{s_m} e_{s_m}) \end{vmatrix}, \quad (31')$$

а суммирование ведется по всевозможным

наборам натуральных чисел $1 \leq s_1 < s_2 < \dots < s_m \leq n$.

Оценив правую часть (31) по неравенству Коши-Буняковского, получим:

$$[(Hx_i, x_j)]^2 \leq \left(\sum_{1 \leq s_1 < \dots < s_m \leq n} [(x_i, h_{s_1} e_{s_1})]^2 \right) \left(\sum_{1 \leq s_1 < \dots < s_m \leq n} [(e_{s_1}, x_j)]^2 \right) \quad (32)$$

Вторая сумма в правой части неравенства (32) равна определителю Грама $[(x_i, x_j)]$. В этом легко убедиться, положив в формуле (31) $H = E$, где E – единичный оператор. Впрочем, соответствующее равенство отдельно доказано на стр. 230 (формула (26)). В первой сумме правой части (32) вынесем из каждого определителя (31')

произведение $h_{s_1} h_{s_2} \dots h_{s_m}$ и заменим его большим $h_1 h_2 \dots h_m$. В результате получим:

$$[(Hx_i, x_j)]^2 \leq h_1^2 h_2^2 \dots h_m^2 [(x_i, x_j)]^2.$$

Извлекая из обеих частей этого неравенства квадратные корни, мы устанавливаем справедливость неравенства (28). Докажем далее следующий факт.

Лемма 4. Пусть K – произвольный оператор в R и

$$x_1 \geq x_2 \geq \dots \geq x_n \quad (33)$$

– его сингулярные числа. Тогда для произвольного набора векторов $x_1, x_2, \dots, x_m$ ($m \leq n$) справедливо неравенство

$$\left[(Kx_i, Kx_j) \right] \leq x_1^2 x_2^2 \dots x_m^2 \left[(x_i, x_j) \right]. \quad (34)$$

Неравенство (34) немедленно следует из леммы 3 при $H = K^* K$.

Установим, наконец, еще одно вспомогательное предложение.

Лемма 5. Пусть A и B – линейные операторы в R , $C = AB$, и пусть α_s, β_s и γ_s ($s = 1, 2, \dots, n$) – сингулярные числа соответственно A, B и C , занумерованные в убывающем порядке. Тогда при любом $m \leq n$ справедливы неравенства

$$\gamma_1 \gamma_2 \dots \gamma_m \leq \alpha_1 \alpha_2 \dots \alpha_m \beta_1 \beta_2 \dots \beta_m. \quad (35)$$

Для доказательства рассмотрим ортонормированный базис $e_1, e_2, \dots, e_n$ собственных векторов оператора $C^* C$. Последовательно применяя (34), получаем

$$\left[(Ce_i, Ce_j) \right] = \left[(ABe_i, ABe_j) \right] \leq \alpha_1^2 \alpha_2^2 \dots \alpha_m^2 \left[(Be_i, Be_j) \right] \leq \alpha_1^2 \alpha_2^2 \dots \alpha_m^2 \beta_1^2 \beta_2^2 \dots \beta_m^2 \left[(e_i, e_j) \right]. \quad (36)$$

С другой стороны, поскольку e_i ($1 \leq i \leq m$) – собственные векторы $C^* C$, мы имеем:

$$\left[(Ce_i, Ce_j) \right] = \left[(C^* Ce_i, e_j) \right] = \gamma_1^2 \gamma_2^2 \dots \gamma_n^2 \left[(e_i, e_j) \right]. \quad (37)$$

Следовательно, (35) имеет место.

Мы в состоянии теперь доказать следующую теорему, которая является основной целью настоящего параграфа.

Теорема 1 (Нейман-Хорн [218b, 191b]). Пусть A, B и C – линейные операторы в n -мерном унитарном пространстве R . Пусть $C = AB$ и пусть α_s, β_s и γ_s ($s = 1, 2, \dots, n$) – сингулярные числа операторов A, B и C , занумерованные в порядке убывания.

Пусть $f(x)$ – непрерывная при $x \geq 0$ функция такая, что $\varphi(t) = f(e^t)$ – монотонно возрастающая выпуклая функция параметра t . Тогда при всех $m \leq n$ справедливы неравенства

$$\sum_{s=1}^m f(\gamma_s) \leq \sum_{s=1}^m f(\alpha_s \beta_s). \quad (38)$$

Доказательство. Пусть сначала операторы A и B невырождены, тогда все числа α_s, β_s и γ_s положительны. Логарифмируя неравенства (35), получаем

$$\sum_{s=1}^m \ln \gamma_s \leq \sum_{s=1}^m \ln (\alpha_s \beta_s), \quad 1 \leq m \leq n. \quad (39)$$

На основании леммы 2 имеем

$$\sum_{s=1}^m \varphi(\ln \gamma_s) \leq \sum_{s=1}^m \varphi(\ln \alpha_s \beta_s), \quad 1 \leq m \leq n. \quad (40)$$

Так как $\varphi(t) = f(e^t)$, то отсюда следует (38). В случае вырожденных операторов неравенства (38) устанавливаются по непрерывности.

Замечание 1. В случае $f(x) = x^\sigma$ ($\sigma \geq 0$) получаем

$$\sum_{s=1}^m \gamma_s^\sigma \leq \sum_{s=1}^m \alpha_s^\sigma \beta_s^\sigma \quad (1 \leq m \leq n). \quad (41)$$

В таком виде неравенства (38) встречаются в приложениях чаще всего.

Замечание 2. При $m = n$ неравенство (39) превращается в равенство (см. сноску на стр. 540). Поэтому при $m = n$ неравенство (40) справедливо для любой

непрерывной выпуклой функции $\varphi(t)$ (см. замечание к лемме 2).

В частности, неравенство (41) при $m = n$ справедливо и для $\sigma < 0$.

Замечание 3. Пусть

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$$

– сингулярные числа оператора A и пусть

$$\alpha_{i,1} \geq \alpha_{i,2} \geq \dots \geq \alpha_{i,n}$$

– сингулярные числа оператора A^l (l – натуральное число). Тогда при любом $\sigma \geq 0$ и любом $1 \leq m \leq n$

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \sum_{s=1}^m \alpha_s^{\sigma l} . \quad (42)$$

Неравенства (42) докажем индукцией по l . При $l = 1$ соотношение (42) очевидно; пусть оно выполняется для $l-1$. Так как $A^l = A^{l-1} \cdot A$, то согласно (41)

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \sum_{s=1}^m \alpha_{i-1,s}^\sigma \alpha_s^\sigma, \quad 1 \leq m \leq n . \quad (43)$$

Применяя к правой части (43) неравенство Гёльдера с

$$p = \frac{l}{l-1} \text{ и } q = l \quad \left(p^{-1} + q^{-1} = 1 \right),$$

получаем

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \left(\sum_{s=1}^m \alpha_{i-1,s}^{p\sigma} \right)^{1/p} \left(\sum_{s=1}^m \alpha_s^{q\sigma} \right)^{1/q} . \quad (44)$$

По предположению индукции, имеем для первой суммы в правой части (44):

$$\left(\sum_{s=1}^m \alpha_{i-1,s}^{p\sigma} \right)^{1/p} \leq \left(\sum_{s=1}^m \alpha_s^{p\sigma(l-1)} \right)^{1/p} = \left(\sum_{s=1}^m \alpha_s^{l\sigma} \right)^{1-\frac{1}{l}} .$$

Учитывая, что во второй сумме правой части (44) $q = l$, легко получаем из (44):

$$\sum_{s=1}^m \alpha_{i,s}^\sigma \leq \sum_{s=1}^m \alpha_s^{l\sigma} ,$$

что и требовалось доказать.

В частности, при $\sigma = 2$ и $m = n$ из формулы (42) следует, что

$$\text{Sp} \left(A^{*l} A^l \right) \leq \text{Sp} \left(A^* A \right)^l . \quad (45)$$