

ЛИНЕЙНАЯ АЛГЕБРА

ШИШКИН
АЛЕКСАНДР АЛЕКСАНДРОВИЧ

—
ФИЗФАК МГУ



ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

Содержание

Гл. 1. Основные понятия.	3
§1. Что такое линейная алгебра?	3
§2. Числовые поля.	3
§3. Линейная зависимость столбцов и строк.	4
§4. Ранг матрицы. Теорема о базисном миноре.	6
Гл. 2. Линейные пространства.	10
§1. Определение линейного пространства	10
§2. Некоторые простейшие свойства линейных пространств.	12
§3. Линейная зависимость элементов линейного пространства.	13
§4. Базис и координаты элементов линейного пространства.	15
§5. Размерность линейного пространства.	16
§6. Преобразование базиса и координат элементов линейного пространства.	17
§7. Подпространства линейного пространства.	18
§8. Линейные оболочки.	19
§9. Ранг матрицы и размерность линейной оболочки ее столбцов. Вычисление ранга матрицы.	20
§10. Изоморфизм линейных пространств.	23
Гл. 3. Система линейных уравнений.	25
§1. Критерий совместности общей линейной системы уравнений.	25
§2. Однородные линейные системы уравнений.	25
§3. Общее решение неоднородной линейной системы уравнений.	27
Гл. 4. Евклидово пространство.	29
§1. Определение.	29
§2. Основные метрические понятия в евклидовом пространстве. Неравенство Коши-Буняковского.	30
§3. Ортогональность элементов в евклидовом пространстве. Ортонормированный базис в евклидовом пространстве.	31
§4. Разложение евклидова пространства на прямую сумму его подпространств.	32
§5. Ортогональная матрица.	36
§6. Общий вид линейного функционала.	38
§7. Изоморфизм евклидовых пространств.	38
Гл. 5. Линейные операторы в линейном конечномерном пространстве.	40
§1. Основные понятия.	40
§2. Матрица линейного оператора.	40
§3. Связь матриц оператора при переходе от одного базиса к другому.	42
§4. Действия над линейными операторами и соответствующие действия над матрицами.	42
§5. Собственные значения и собственные векторы линейных операторов.	45
§6. Симметричный (самосопряженный) линейный оператор в конечномерном евклидовом пространстве.	52

Гл. 6. Квадратичные формы.	57
§1. Общие понятия.	57
§2. Изменение квадратичной формы при линейном преобразовании переменных.	57
§3. Метод Лагранжа приведения квадратичной формы к каноническому виду.	59
§4. Приведение квадратичной формы ортогональным преобразованием к каноническому виду.	61
§5. Билинейные формы. Их связь с квадратичными формами.	63
§6. Метод Якоби приведения квадратичной формы к каноническому виду.	65
§7. Закон инерции квадратичной формы.	67
§8. Классификация квадратичных форм.	68
§9. Численные методы решения систем линейных уравнений.	70
п. 1. Итерационный метод Гаусса-Зейделя.	70
п. 2. Метод Зейделя.	70
п. 3. Метод исключения (Гаусса).	71
Литература.	73
Заключение.	74

Гл. 1. Основные понятия.

§1. Что такое линейная алгебра?

Вы уже знакомы с элементарной алгеброй чисел, с векторной алгеброй и алгеброй матриц. А теперь приступаем к изучению линейной алгебры. Что такое алгебра? Алгебра — раздел математики, исследующий операции, аналогичные сложению, вычитанию, умножению и делению, и выполняемые не только над числами, но и над другими математическими объектами, например, многочленами, векторами, матрицами, операторами и т. д., т. е. над объектами самой различной природы. Современную алгебру математики рассматривают как учение об операциях над любыми математическими объектами самой различной природы, как учение, формирующее общие понятия и методы для всей математики. В центре внимания алгебры оказываются свойства операций, а не объекты, над которыми производятся операции. Одним из разделов алгебры является линейная алгебра. Какими проблемами занимается линейная алгебра?

В школе вы складывали действительные числа по одному правилу, а вектора — по другому правилу. В алгебре матриц матрицы складываются по третьему правилу. На первый взгляд кроме общего названия в этих трех операциях нет ничего общего. Но если мы рассмотрим внимательнее операцию сложения в применении к различным типам объектов, то увидим, что эти операции обладают многими общими свойствами. Например,

для любых действительных чисел a, b : $a + b = b + a$;

для любых матриц $A, B \in A_n^m$: $A + B = B + A$;

для любых векторов \vec{a}, \vec{b} : $\vec{a} + \vec{b} = \vec{b} + \vec{a}$, т. е. выполняется переместительное свойство. Можно привести и другие примеры общих свойств этой операции. Разглядеть эти общие свойства операции нам порою мешает конкретная природа объектов рассматриваемых совокупностей. Соблазнительно и следующее соображение: знание общих свойств операции позволит избавиться от нудных и однообразных повторений при переходе от исследования одной совокупности к другой. Выход — рассмотреть совокупность объектов лишенных конкретной природы. Изучить на них свойства операций. А затем применять уже готовые результаты у объектам любой, но уже конкретной природы. Хотим — к числам, хотим — векторам, а можно и к матрицам и т. д., ибо различных алгебр (помимо трех выше перечисленных) в приложениях сколько угодно много. Вот прежде всего чем занимается линейная алгебра.

Итак, линейная алгебра является наиболее широко используемым аппаратом для всех разделов чистой и прикладной математики — от теории алгебраических чисел до квантовой механики. Логическая структура линейной алгебры проста и основывается на небольшом числе удобных в обращении понятий и аксиом. Тем не менее и сегодня линейная алгебра, вследствие абстрактного характера ее основных понятий, сохраняет за собой репутацию сложной науки. Ее изучение потребует от вас гораздо больше усилий, чем аналитическая геометрия и даже математический анализ.

§2. Числовые поля.

Главными объектами изучения далее будут матрицы, линейные пространства и многочлены от нескольких переменных. В определении каждого из них участвует некоторое множество чисел, которое обозначим символом K . Выбор K зависит от решаемой задачи и научной дисциплины. Например, с алгебраической точки зрения результаты получают наиболее законченную форму, если в качестве K выбрать мно-

жество всех комплексных чисел. Напротив - в геометрии и механике обычно рассматриваются действительные числа, а в теории чисел в качестве K естественно брать множество рациональных чисел и даже множество лишь целых чисел. Поэтому, чтобы сделать результаты применимыми к более широкому кругу задач, целесообразно заранее не фиксировать, какое именно индивидуальное множество понимается под K . Так поступим и мы, но исключим из рассмотрения комплексные числа, хотя все последующие построения справедливы и для K , содержащего комплексные числа. Это ограничение делается исключительно с целью сократить объем излагаемого материала. Таким образом, в нашем курсе будут рассматриваться различные множества действительных чисел, над которыми совершаются четыре операции, именуемые сложением, вычитанием, умножением, и делением, понимаемые всюду далее в смысле элементарной алгебры, т. е. алгебры, которую вы изучали в школе. При этом в четвертой операции делитель всегда берется отличным от нуля. Это требование будем считать выполненным всюду ниже. Пусть K_0 — множество всех действительных чисел, а поскольку далее будут использоваться только действительные числа, то договоримся их называть просто числами. Перечисленные выше четыре операции на множестве K_0 обладают следующими свойствами: 1) выполнимость операций для любых чисел из K_0 , 2) однозначность операций, 3) принадлежность результата выполнения операции к числам того же множества K_0 .

Далеко не все известные вам операции обладают совокупностью таких свойств. Например, выполнимо вычисление логарифма не на всем множестве K_0 , а только для положительных чисел. Извлечение квадратного корня выполнимо на множестве положительных чисел, но неоднозначно. Однако даже если операция однозначно выполнима для всех чисел из рассматриваемого множества, результат ее выполнения может не быть числом заданного множества. Так операция деления на множестве целых чисел однозначно выполнима для любых двух чисел из этого множества, но результат ее выполнения не обязательно будет целым числом. Итак, мы пришли к очень важному в дальнейшем понятию.

О п р е д е л е н и е. *Операция называется корректной относительно некоторого множества B чисел, если всякой паре чисел из B эта операция ставит в соответствие однозначно определенное число из того же множества B .*

Например, операция сложения корректна относительно множества N натуральных чисел, так как сумма любых двух натуральных чисел есть натуральное число. Но эта же операция не является корректной относительно множества T всех нечетных чисел, ибо сумма любых двух нечетных чисел — четное число, т. е. результат операции сложения не принадлежит множеству T . Операция умножения корректна относительно множеств N и T , а операции вычитания и деления не являются корректными относительно этих множеств.

О п р е д е л е н и е. *Числовым полем K называется такое множество K чисел, на котором корректны все все четыре операции.*

Иными словами, числовое поле K есть такое множество K чисел, на котором операции сложения, вычитания, умножения и деления однозначно выполнимы и $\forall x, y \in K$ а) $x + y \in K$, б) $x - y \in K$, в) $x \cdot y \in K$, г) $x/y \in K (y \neq 0)$.

Очевидно, что множество целых чисел не является числовым полем. А множество всех рациональных чисел и множество K являются числовыми полями.

§3. Линейная зависимость столбцов и строк.

О п р е д е л е н и е. *Произвольная система чисел a_k^p из поля K , расположенная в виде прямоугольной таблицы, содержащей m строк и n столбцов, называется*

$m \times n$ -матрицей или матрицей размеров $m \times n$ и обозначается одним из следующих символов:

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} = \|a_k^p\|_n^m = A.$$

З а м е ч а н и е. Иногда далее нам будет удобнее вместо верхнего и нижнего индексов использовать только нижние. Договоримся о правиле соответствия этих двух записей:

1-я запись	2-я запись
верхний индекс	первый индекс
нижний индекс	второй индекс

т. е. $\|a_k^p\|_n^m = \|a_{pk}\|_{m,n}$.

Как и в курсе аналитической геометрии будем пользоваться правилом сокращенного суммирования: всякий раз, когда в каком-либо выражении встречаются два одинаковых индекса, верхний и нижний, то по этому индексу будет предполагаться суммирование, т. е. запись $a^k x_k$, $k = \overline{s, n}$, (или, что то же самое, $x_k a^k$) означает

$$\sum_{k=s}^n a^k x_k.$$

Ограничимся рассмотрением матриц-столбцов, ибо все, что будет верно для столбцов, останется верным и для строк. Операции сложения столбцов и операция умножения столбца на число введены нами в гл. 4 (Алгебра матриц) курса аналитической геометрии.

О п р е д е л е н и е. Будем говорить, что один из столбцов $A_k = \|a_k^p\|_n$, $k = \overline{1, m}$, например, A_m представляет собой линейную комбинацию остальных столбцов, если найдутся такие числа x^1, x^2, \dots, x^{m-1} из поля K , что

$$A_m = A_k x^k, k = \overline{1, m-1}. \quad (1)$$

О п р е д е л е н и е. Столбцы A_1, A_2, \dots, A_m называются линейно зависимыми, если найдутся такие числа c^1, c^2, \dots, c^m из поля K , не все равные нулю, что

$$A_k c^k = \theta. \quad (2)$$

Теорема 1 Для того чтобы столбцы A_1, A_2, \dots, A_m были линейно зависимы необходимо и достаточно, чтобы один из этих столбцов представлял собой линейную комбинацию остальных.

Д о к а з а т е л ь с т в о. **Необходимость.** Пусть столбцы линейно зависимы, т. е. найдутся такие числа c^1, c^2, \dots, c^m , не все равные нулю, что имеет место (2). Ради определенности предположим, что $c^m \neq 0$. Тогда, умножив равенство (2) на $-1/c^m$ и введя обозначение $x^k = -c^k/c^m$, получим:

$$x^k A_k - A_m = \theta. \quad (3)$$

Прибавим к обеим частям равенства (3) столбец A_m . Так как $A_m + \theta = A_m$, а $A_m + (-1)A_m = \theta$, то приходим к равенству (1), т. е. столбец A_m является линейной комбинацией остальных столбцов.

Д о с т а т о ч н о с т ь. Пусть один из столбцов, например, A_m есть линейная комбинация остальных столбцов, т. е. имеем равенство (1), которое равносильно равенству (3). Поскольку из чисел $x^1, x^2, \dots, x^{m-1}, -1$ одно заведомо не равно нулю, то равенство (3) говорит о линейной зависимости рассматриваемых столбцов. Теорема доказана.

Равенство (2) можно записать в матричной форме, если через e обозначить матрицу-строку, состоящую из столбцов A_1, A_2, \dots, A_m , т. е. $e = \|A_k\|_m$, а через $C = \|c^k\|_m$ — матрицу-столбец, состоящую из чисел c^1, c^2, \dots, c^m . Тогда $A_k c^k = eC$. Следовательно, формула (2) примет вид:

$$eC = \theta \quad (4)$$

О п р е д е л е н и е. Столбцы A_1, A_2, \dots, A_m называются линейно независимыми, если равенство (4) выполняется тогда и только тогда, когда $C = \theta$, т. е. нулевой столбец.

Теорема 2 Если столбцы A_1, A_2, \dots, A_m линейно независимы и имеет место равенство $eB = eC$, то $B = C$.

Д о к а з а т е л ь с т в о. Из свойства 3^o операции сложения матриц из $eB = eC$ следует, что $eB - eC = \theta$. Из свойства 2^o умножения матриц следует: $e(B - C) = \theta$. По определению линейной независимости это равенство имеет место тогда и только тогда, когда $B - C = \theta$, т. е. $B = C$. Что и требовалось доказать.

§4. Ранг матрицы. Теорема о базисном миноре.

Пусть дана матрица $A = \|a_k^p\|_n^m$. Фиксируем в ней некоторое число r столбцов и такое же число строк. Элементы, стоящие на пересечении выделенных столбцов и строк образуют матрицу r -порядка. Ее определитель называется минором r -го порядка матрицы A . Заметим, что минор элемента определителя матрицы A — частный случай минора r -го порядка этой матрицы, а именно: 1) матрица A — квадратная, 2) $r = n - 1$.

О п р е д е л е н и е. Рангом матрицы A называется натуральное число r , если у матрицы A есть минор r -го порядка, отличный от нуля, а все миноры $r + 1$ -го порядка равны нулю, и обозначается символом $\text{rang } A$. Минор r -го порядка, неравный нулю, называется базисным минором матрицы A и обозначается символом M_b .

З а м е ч а н и е 1. Если у матрицы A хотя бы один элемент $a_k^p \neq 0$, то заведомо $\text{rang } A \geq 1$. Если все $a_k^p = 0$, то считается, что $\text{rang } A = 0$.

З а м е ч а н и е 2. Если у $n \times n$ -матрицы A отличен от нуля минор порядка n , т. е. $\det A \neq 0$, то его назовем базисным минором, а рангом — число n , хотя у матрицы A нет минора $n + 1$ -го порядка.

З а м е ч а н и е 3. У матрицы A может быть несколько базисных миноров, но все они имеют один и тот же порядок $r = \text{rang } A$.

П р и м е р. Матрица $A = \|a^p b_k\|_n^n$ имеет ранг, равный единице, так как все миноры порядка ≥ 2 равны нулю (если найдутся такие p и k , для которых $a^p b_k \neq 0$).

О п р е д е л е н и е. Столбцы (строки) матрицы A , на пересечении которых находятся элементы базисного минора, называются базисными столбцами (строками).

Теорема 3 (О базисном миноре) *Базисные столбцы матрицы A линейно независимы. Любой столбец матрицы A представляет собой линейную комбинацию ее базисных столбцов, причем такое представление единственно.*

Доказательство. 1) Докажем вначале линейную независимость базисных столбцов. Рассмотрим базисный минор матрицы A . Положим, для определенности, что базисный минор расположен в первых r строках и в первых r столбцах матрицы A , т. е. $M_b = \det \|a_k^p\|_r^r \neq 0$. Такое предположение не является ограничением, ибо, если бы это было не так, то, переставляя строки и столбцы, всегда можно расположить базисные строки и столбцы в первых r строках и столбцах, а затем перенумеровать строки и столбцы. Докажем 1) методом от противного, т. е. положим, что базисные столбцы линейно зависимы. Тогда, согласно теореме 1, один из базисных столбцов, например, r -ый представляет собой линейную комбинацию остальных. Вычитая эту линейную комбинацию из элементов r -го столбца, мы, согласно свойству 7, не изменим величину определителя M_b , но в результате r -ый столбец будет состоять из нулей. По свойству 6 этот определитель равен нулю, т. е. $M_b = 0$. Полученное противоречие доказывает первую часть теоремы.

2) Пусть e — любое парное число от 1 по n , а k — любое целое число от 1 до m . Рассмотрим определитель $(r + 1)$ -го порядка:

$$D = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_r^1 & a_e^1 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^r & a_2^r & \dots & a_r^r & a_e^r \\ a_1^k & a_2^k & \dots & a_r^k & a_e^k \end{vmatrix}$$

Могут представиться три случая:

- а) $e \leq r$, тогда $D = 0$, ибо определитель имеет два одинаковых столбца;
- б) $k \leq r$: $D = 0$, ибо определитель имеет две одинаковых строки;
- в) $e > r$, $k > r$: $D = 0$, ибо D — минор $(r + 1)$ -го порядка.

Итак, $D = 0$ при любых значениях k , e . Разложим определитель D по элементам последней строки (формула 5**):

$$\sum_{p=1}^r a_p^k A_p^k + a_e^k A_e^k = 0.$$

Заметим, что, во-первых, алгебраические дополнения элементов k -ой строки совсем не зависят от элементов этой строки, что следует из определения алгебраического дополнения, т. е. $A_1^k, A_2^k, \dots, A_r^k, A_e^k$ одни и те же для всех $k = \overline{1, m}$, и, во-вторых, $A_e^k \neq 0$, так как $A_e^k = M_b$. Поделим последнее равенство на A_e^k и перенесем все слагаемые, кроме последнего, в правую часть. Вводя обозначение $c_p = -A_p^k/A_e^k$, $p = \overline{1, r}$, получим: $a_e^k = a_p^k c^p$, где $p = \overline{1, r}$, $k = \overline{1, m}$, или в матричной форме:

$$\|a_e^k\|^m = \|a_p^k\|^m c^p,$$

где $p = \overline{1, r}$, т. е. любой e -ый столбец является линейной комбинацией базисных столбцов. Осталось показать что это представление единственно. Предположим противное: есть еще одно представление e -го столбца — $\|a_e^k\|^m = \|a_p^k\|^m b^p$, где $p = \overline{1, r}$. Вычтем из первого представления второе: $\theta = \|a_p^k\|^m (c^p - b^p)$. Но $\|a_p^k\|^m$, $p = \overline{1, r}$ — базисные столбцы и, следовательно, линейно независимы. Поэтому последнее равенство возможно тогда и только тогда, когда $c^p - b^p = 0$. т. е. $c^p = b^p$, $p = \overline{1, r}$. Полученное противоречие доказывает единственность разложения. Теорема доказана.

С л е д с т в и е 1. *Для того, чтобы определитель D n -го порядка равнялся нулю необходимо и достаточно, чтобы между его столбцами (строками) существовала линейная зависимость.*

Н е о б х о д и м о с т ь. Дано: $D = 0$. Требуется доказать: его столбцы линейно зависимы. Так как $D = 0$, то базисный минор матрицы определителя D имеет порядок $r < n$. Поэтому после выделения r базисных столбцов найдется еще по крайней мере один столбец, не попавший в число базисных. По теореме 3 такой столбец представляет собой некоторую линейную комбинацию базисных столбцов. В состав этой комбинации можно включить и все оставшиеся столбцы определителя D , поставив перед ними, например, нулевые коэффициенты. Но тогда по теореме 1 столбцы определителя D являются линейно зависимыми.

Д о с т а т о ч н о с т ь. Дано: столбцы D линейно зависимы. Требуется доказать: $D = 0$. В силу теоремы 1 один из столбцов определителя D есть линейная комбинация остальных столбцов. Если эту линейную комбинацию вычесть из этого столбца, то по свойству 7 определитель D не изменится, но у него появится нулевой столбец. Поэтому по свойству 6 $D = 0$. Что и требовалось доказать.

С л е д с т в и е 2. *Нетривиальная совместность однородной линейной системы уравнений.*

Пусть дана однородная системы уравнений:

$$a_k^p x^k = 0, \quad p = \overline{1, m}, \quad k = \overline{1, n}, \quad (5)$$

с матрицей $A = ||a_k^p||_n^m$. Система (5) всегда совместна, ибо имеет тривиальное решение: $x^k = 0, k = \overline{1, n}$.

Естественно возникает вопрос: при каких условиях система (5) имеет, кроме тривиального, еще и другие решения? Ответ на этот вопрос дает

Теорема 4 *Для того чтобы система (5) имела нетривиальные решения, необходимо и достаточно, чтобы ранг матрицы системы был меньше числа неизвестных, т. е. $r < n$.*

Н е о б х о д и м о с т ь. Дано: (5) имеет нетривиальные решения. Требуется доказать: $r < n$.

Итак, существуют числа c^1, c^2, \dots, c^n , не все равные нулю, обращающие систему (5) в верные равенства, т. е.

$$a_k^p c^k = 0, \quad p = \overline{1, m}, \quad k = \overline{1, n},$$

или в матричной форме:

$$||a_k^p||_n^m c^k = \theta, \quad k = \overline{1, n},$$

т. е. столбцы матрицы A линейно зависимы. Отсюда следует, что порядок базисного минора $r < n$, ибо базисные столбцы являются линейно независимыми.

Д о с т а т о ч н о с т ь. Дано: $r < n$. Требуется доказать: (5) имеет нетривиальные решения.

Так как $r < n$, то у матрицы A хотя бы один столбец является не базисным. По теореме 3 этот столбец представляет собой линейную комбинацию базисных столбцов, а значит и всех остальных столбцов матрицы A . Но тогда по теореме 1 столбцы матрицы A линейно зависимы, т. е. по определению линейной зависимости найдутся такие числа c^1, c^2, \dots, c^n , не все равные нулю, что справедливо равенство:

$$||a_k^p||_n^m c^k = \theta, \quad k = \overline{1, n},$$

или в координатах:

$$a_k^p c^k = 0, \quad p = \overline{1, m}, \quad k = \overline{1, n}.$$

Последние равенства говорят о том, что совокупность n чисел c^1, c^2, \dots, c^n представляет собой нетривиальное решение системы (5).

С л е д с т в и е 3. Для того, чтобы однородная линейная система n уравнений с n неизвестными имела нетривиальные решения, необходимо и достаточно, чтобы определитель матрицы этой системы был равен нулю.

С л е д с т в и е 4. Для того чтобы столбцы матрицы A были линейно зависимы, необходимо и достаточно, чтобы ее ранг был меньше числа ее столбцов, т. е. $r < n$.

Гл. 2. Линейные пространства.

Мы подошли к основному понятию нашего курса. Со времен Аристотеля известно, что в основе любой науки лежит то, что можно назвать "принципом намеренно неполного знания": абстракция и обобщение как раз и состоят в том, что определенные свойства рассматриваемых объектов систематически игнорируются. Аксиоматический метод в математике представляет собой не что иное, как применение этого принципа, и он отличается от других ситуаций, в которых работает "принцип намеренно неполного знания" лишь тем, что все свойства, которые математик не склонен игнорировать, т. е. которые он желает признать присущими изучаемым объектам, заботливо и исчерпывающе перечисляются — и в дальнейшем запрещается опираться на что-либо, кроме этих свойств (аксиом) и правил логики.

§1. Определение линейного пространства

О п р е д е л е н и е. Множество R элементов любой (но одной) природы x, y, e, \dots называется линейным пространством, если

- 1^o. Существует правило, с помощью которого любой упорядоченной паре элементов x, y из R ставится в соответствие третий элемент e из R , называемый суммой элементов x, y и обозначаемый " $x + y$ ";
- 2^o. существует еще одно правило, с помощью которого для любого элемента x из R , и любого числа c из поля K ставится в соответствие элемент ϕ из R , называемый произведением элемента x на число c и обозначаемый символом " cx " или " xc ";
- 3^o. оба этих правила удовлетворяют 8 аксиомам:

аксиомы операции сложения:

- a.1) $\forall x, y \in R \quad x + y = y + x$ (переместительное свойство),
- a.2) $\forall x, y, e \in R \quad (x + y) + e = x + (y + e)$ (сочетательное свойство),
- a.3) существует нулевой элемент θ такой, что $x + \theta = x \quad \forall x \in R$ (особая роль нулевого элемента),
- a.4) $\forall x \in R$ существует противоположный элемент x' такой, что $x + x' = \theta$;

аксиомы операции умножения:

- a.5) $\forall x \in R \quad 1 \cdot x = x$ (особая роль числового множителя 1),
- a.6) $\forall x \in R$ и $\forall c, b \in K \quad c(bx) = (cb)x$ (сочетательное относительно числовых множителей свойство);

аксиомы, связывающие операции сложения и умножения:

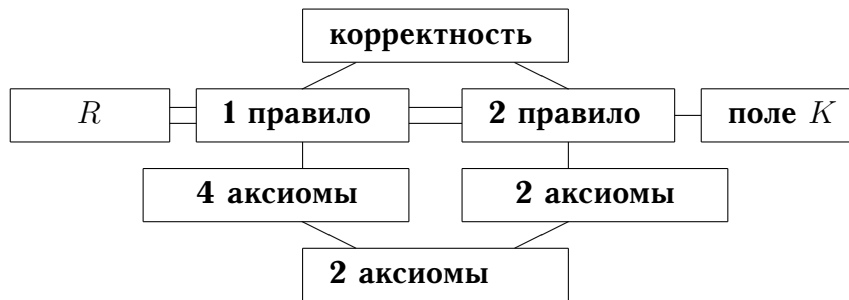
- a.7) $\forall c, b \in K$ и $\forall x \in R \quad (c + b)x = cx + bx$ (распределительное относительно суммы числовых множителей свойство),
- a.8) $\forall c \in K$ и $\forall x, y \in R \quad c(x + y) = cx + cy$ (распределительное относительно суммы элементов свойство).

З а м е ч а н и е 1. В определении множество R элементов любой, но одной природы.

З а м е ч а н и е 2. В а.6) операция умножения чисел cb и в а.7) операция сложения чисел $b + c$ понимается в смысле операций в поле K , т. е. обычных школьных операций сложения и умножения чисел.

З а м е ч а н и е 3. Аксиомы 1) - 8) не претендуют на логическую независимость. Они просто являются удобным описанием объектов, которые мы желаем изучать.

З а м е ч а н и е 4. Линейное пространство можно интерпретировать как объединение трех объектов: множества R и двух правил, причем правила 1, 2



не произвольны, а удовлетворяют условию корректности и 8 аксиомам, из которых 4 относятся к первому правилу, две — ко второму правилу и еще две аксиомы связывают эти операции друг с другом. Наконец, правило 2 зависит от поля K . Эта схема, естественно, приводит нас к мысли, что добавляя или убирая в этой схеме "кубики", т. е. вводя новые правила, или убирая часть старых, или меняя их, мы получим другие объекты алгебры. Именно так они и строятся.

З а м е ч а н и е 5. В монографиях по алгебре элементы линейного пространства часто называются точками или векторами. Так же, подчеркивая связь линейного пространства с числовым полем K , линейное пространство R называется линейным пространством R над полем K .

З а м е ч а н и е 6. Далее в этой главе, не упоминая о числовом поле K , над которым определено линейное пространство, будем полагать, что все числа принадлежат этому числовому полю K .

З а м е ч а н и е 7. При нашем введении числового поля K линейное пространство содержит или бесконечно много элементов или только один элемент. Нетрудно проверить с помощью аксиом, что последний случай — это пространство, состоящее из нулевого элемента. Это пространство будем называть тривиальным линейным пространством и обозначать так же как и сам элемент буквой θ .

З а м е ч а н и е 8. Договоримся употреблять слово "эквивалентность" и писать знак \equiv лишь тогда, когда выполняются три условия:

- а) каждый элемент эквивалентен самому себе: $\forall a, a \equiv a$ (рефлексивность)
- б) высказывание, что два элемента являются эквивалентными, не требует уточнения, какой из элементов рассматривается первым и какой вторым: $\forall a, b, a \equiv b \Rightarrow b \equiv a$ (симметричность)
- в) два элемента, эквивалентные третьему, эквивалентны между собой: $\forall a, b, c, a \equiv b \text{ и } b \equiv c \Rightarrow a \equiv c$ (транзитивность).

Примеры линейных пространств.

- 1) Множество K_0 над полем K рациональных чисел с обычными операциями сложения и умножения над числами в поле.
- 2) Множество K_0 над полем K_0 с теми же операциями, что и в 1).
- 3) Множество A_n^k всех $k \times n$ - матриц с двумя линейными операциями, введенными в Гл. 4 курса аналитической геометрии (там же доказано выполнение всех 8 аксиом определения линейного пространства). Если $n = 1$, т. е. имеем множество столбцов одной высоты k , то линейное пространство обозначается T_k , его элементы называются векторами, а элементы столбца — координатами вектора.
- 4) Множество V_3 всех свободных векторов в пространстве с двумя линейными операциями (выполнение 8 аксиом доказано в курсе аналитической геометрии).
- 5) В математическом анализе рассматривается множество всех непрерывных на сегменте $[a, b]$ функций с обычными операциями сложения и умножения на число (здесь θ есть функция, тождественно равная нулю на всем сегменте). Это линейное пространство обозначается символом $C_{[a,b]}$.
- 6) $C_{[a,b]}$ — то же, что в 5), но функции имеют k непрерывных производных на сегменте $[a, b]$.
- 7) Рассмотрим множество R векторов на плоскости, начала которых совпадают с началом координат, введенных на плоскости, а концы — в пределах первого квадранта. Будет ли такое множество векторов с обычными линейными операциями образовывать линейное пространство? Ответ: нет, так как, например, некорректна операция умножения вектора на число — при умножении вектора на -1 получаем вектор, не принадлежащий рассматриваемому множеству.
- 8) Образует ли линейное пространство множество столбцов высоты n , у каждого из которых совпадает первая и последняя координаты? Ответ: да.

Н а д о м:

- 1) Образуют ли линейное пространство множество всех векторов на плоскости за исключением векторов, параллельных некоторой фиксированной прямой, с обычными линейными операциями над векторами? Ответ: нет (некорректная операция сложения)
- 2) Рассмотрите множества векторов с обычными операциями сложения и умножения на число $\bar{x} = (a, b, c)$ таких, что а) $a = 0$, б) $a = 0$ или $b = 0$, в) $a + b = 0$, г) $a + b = 1$. В каком из этих случаев совокупность будет линейным пространством? Ответ: а) да, б) нет, в) да, г) нет.

§2. Некоторые простейшие свойства линейных пространств.

Из а.1) - а.8) можно получить следующие теоремы.

Теорема 1 В любом линейном пространстве существует единственный нулевой элемент.

Доказательство. Существование хотя бы одного нулевого элемента утверждает а.3). Предположим, что таких элементов два: θ_1 и θ_2 . Положим в а.3) $x = \theta_1$ и $\theta = \theta_2$. Тогда $\theta_1 + \theta_2 = \theta_1$. Полагая в той же аксиоме $x = \theta_2$ и $\theta = \theta_1$, получим $\theta_2 + \theta_1 = \theta_2$. Так как по а.1) $\theta_1 + \theta_2 = \theta_2 + \theta_1$, то $\theta_1 = \theta_2$. Что и требовалось доказать.

Теорема 2 В любом линейном пространстве для каждого элемента существует единственный противоположный элемент.

Доказательство. Существование хотя бы одного противоположного элемента утверждает а.4). Пусть для некоторого элемента x имеются два противоположных элемента x'_1 и x'_2 . Рассмотрим сумму $x + x'_1 + x'_2$. С одной стороны, $x + x'_1 + x'_2 = (x + x'_1) + x'_2 =$ (по а.4)) $= \theta + x'_2 =$ (по а.1)) $= x'_2 + \theta = x'_2$. С другой стороны, $x + x'_1 + x'_2 =$ (по а.1)) $= x'_1 + x + x'_2 =$ (по а.2)) $= x'_1 + (x + x'_2) =$ (по а.4)) $= x'_1 + \theta =$ (по а.3)) $= x'_1$. Отсюда следует, что $x'_1 = x'_2$. Теорема доказана.

Теорема 3 Для любого элемента x линейного пространства имеет место равенство $0 \cdot x = \theta$.

Доказательство. $x =$ (по а.5)) $= 1 \cdot x = (1 + 0)x =$ (по а.7)) $= 1 \cdot x + 0 \cdot x =$ (по а.5)) $= x + 0 \cdot x$, т. е. $x = x + 0 \cdot x$. Прибавляя к обеим частям противоположный элемент x' для x , найдем $x + x' = x + 0 \cdot x + x'$. По а.1) и а.4): $\theta = 0 \cdot x + \theta = 0 \cdot x$.

Теорема 4 Для любого элемента x линейного пространства противоположный элемент есть элемент $(-1)x$, т. е. $x' = (-1)x$.

Доказательство. $x + (-1)x =$ (по а.5)) $= 1 \cdot x + (-1)x =$ (по а.7)) $= (1 - 1)x = 0 \cdot x =$ (по т.3) $= \theta$, т. е. элемент $(-1)x$ удовлетворяет а.4), т. е. является противоположным для x .

Обозначим $(-1)x = -x$. Теперь можно ввести понятие разности элементов линейного пространства: $x - y = x + y' = x + (-y)$, т. е. разность элементов x и y есть сумма x и противоположного к y элемента.

Надо: Доказать, что 1) $a \cdot \theta = \theta$, 2) $-\theta = \theta$.

Решение. $\forall a \in K, \forall x \in R$ $ax =$ (а.3)) $= a(x + \theta) =$ (а.8)) $= ax + a\theta$, т. е. $ax = ax + a\theta$. Добавим к обеим частям последнего равенства элемент $-ax$. В силу а.1), а.4) имеем $\theta = a\theta$. В частности, при $a = -1$ имеем: $\theta = -\theta$.

§3. Линейная зависимость элементов линейного пространства.

Пусть x_1, x_2, \dots, x_n — элементы линейного пространства R , c^1, c^2, \dots, c^n — числа из K .

Определение. Элемент $y = x_k c^k$ называется линейной комбинацией элементов x_1, x_2, \dots, x_n с коэффициентами c^1, c^2, \dots, c^n .

Если $c^1 = c^2 = \dots = c^n = 0$, то в силу теоремы 3 и а.3) $y = \theta$.

Определение. Элементы x_1, x_2, \dots, x_n называются линейно зависимыми, если некоторая их линейная комбинация, не все коэффициенты которой равны нулю, дает элемент θ .

Определение. Элементы x_1, x_2, \dots, x_n называются линейно независимыми, если $x_k c^k = \theta$ тогда и только тогда, когда все $c^k = 0, k = \overline{1, n}$.

Все эти определения дословно повторяют аналогичные определения, которые были даны для векторов в векторной алгебре и для столбцов в §3 гл. 1 нашего курса. Только теперь природа элементов не конкретизируется.

Теорема 5 Для того чтобы элементы x_1, x_2, \dots, x_n были линейно зависимы, необходимо и достаточно, чтобы один из них был линейной комбинацией остальных.

Доказательство этой теоремы совпадает с доказательством теоремы 1 §3 Гл. 1.

Легко доказываются следующие теоремы.

Теорема 6 Если к элементам x_1, x_2, \dots, x_n добавить нулевой элемент θ , то элементы $x_1, x_2, \dots, x_n, \theta$ линейно зависимы.

Для доказательства достаточно в линейной комбинации $n + 1$ элементов положить $c^1 = c^2 = \dots = c^n = 0, c^{n+1} = 1$.

Теорема 7 Если элементы x_1, x_2, \dots, x_n линейно зависимы, то элементы $x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_k$ так же линейно зависимы.

Примеры.

1) В пространстве T_n : $e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$ линейно независимы, так как $y = e_k c^k = \|c^k\|^n = \theta = \|0\|^n$ только тогда, когда все $c^k = 0, k = \overline{1, n}$.

2) Элементы пространства A_n^m :

$e_1 = E_1^1 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, e_2 = E_2^1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$
 $\dots, e_{m \times n} = E_n^m = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$ линейно независимы, ибо $y = e_k c^k = \begin{pmatrix} c^1 & c^2 & \dots & c^n \\ c^{n+1} & c^{n+2} & \dots & c^{2n} \\ \dots & \dots & \dots & \dots \\ c^{n(m-1)+1} & c^{n(m-1)+2} & \dots & c^{nm} \end{pmatrix} = \theta = \|0\|_n^m$ тогда и только тогда, когда все $c^k = 0, k = \overline{1, nm}$.

3) Элементы $C_{[a,b]}$: $e_1 = \sin^2 x, e_2 = 3 \cos^2 x, e_3 = -1$ — линейно зависимы, так как $y = e_k c^k = 0$ при $c^1 = 3, c^2 = 1, c^3 = 3$.

4) Элементы $C_{[a,b]}$: $e_1 = 1, e_2 = x, e_3 = x^2, \dots, e_{n+1} = x^n$, где n — любое фиксированное натуральное число, линейно независимы, ибо линейная комбинация $y = c^1 + c^2 x + \dots + c^{n+1} x^n = 0$ — конечное уравнение которое по основной теореме алгебры при любых c^k , не всех равных нулю, имеет не более n корней относительно x на $[a, b]$. Следовательно, нельзя выбрать нетривиальную совокупность чисел c^k так, чтобы это равенство выполнялось для всех x из сегмента $[a, b]$.

Надо м: Доказать: если x, y, e — линейно независимые элементы пространства, то этим же свойством обладают элементы $x + y, y + e, e + x$.

§4. Базис и координаты элементов линейного пространства.

В векторной алгебре было введено понятие базиса. Каждый вектор в пространстве задавался тремя числами — координатами в некотором базисе. Аналогичное понятие введем в линейном пространстве. Точнее, понятие базиса в векторной алгебре — частный случай того понятия, которое мы сейчас введем.

О п р е д е л е н и е. Совокупность n линейно независимых элементов e_1, e_2, \dots, e_n называется базисом линейного пространства R и обозначается символом $(e_k)_n$, если для любого x из R существуют такие числа x^1, x^2, \dots, x^n , что $x = e_k x^k$, причем эта линейная комбинация называется разложением элемента по базису $(e_k)_n$, а числа x^k — координатами элемента x в базисе $(e_k)_n$.

Теорема 8 Разложение любого элемента линейного пространства по данному базису единственно.

Д о к а з а т е л ь с т в о. (От противного). Пусть $x = e_k a^k$ и $x = e_k b^k$. Вычтем из первого разложения элемента x второе. Тогда $\theta = e_k (a^k - b^k)$. Так как e_1, e_2, \dots, e_n — линейно независимы, то $a^k - b^k = 0$, т. е. $a^k = b^k, k = \overline{1, n}$.

П р и м е р ы.

- 1) Линейное пространство V_3 : любые три некопланарные векторы образуют базис (это доказано в курсе аналитической геометрии).
- 2) Линейное пространство T_3 : в примере 1) §3 приведены линейно независимые элементы e_1, e_2, \dots, e_n , причем для каждого $x = \|a^k\|^n = e_k a^k$, т. е. указанные элементы образуют базис.
- 3) Линейное пространство A_n^m : в примере 2) §3 указаны $n \times m$ линейно независимых элементов $E_k^p, p = \overline{1, m}, k = \overline{1, n}$, причем

$$A = \|a_k^p\|_n^m = \sum_{p=1}^m \sum_{k=1}^n E_k^p a_k^p,$$

т. е. $(e_k)_{nm}$ — базис в линейном пространстве A_n^m .

Значение базиса. После введения базиса абстрактные операции над элементами линейного пространства становятся обычными (точнее, привычными) операциями сложения и умножения над числами — координатами элементов, что утверждает следующая

Теорема 9 При сложении двух элементов линейного пространства их координаты складываются. При умножении элемента на число все его координаты умножаются на это число.

Д о к а з а т е л ь с т в о. Пусть $x = e_k a^k, y = e_k b^k$. Тогда в силу аксиом линейного пространства $x + y = e_k a^k + e_k b^k = (a.1, 2, 7) = e_k (a^k + b^k)$ и $xc = (e_k a^k)c = (a.8, 5) = e_k (a^k c)$. Так как разложение по базису в силу теоремы 8 единственно, то теорема доказана.

Разложение элемента x линейного пространства по базису $(e_k)_n$: $x = e_k a^k$ — можно записать и в другой форме. Для этого введем матрицы составленные из элементов линейного пространства (в отличии от числовых матриц будем их обозначать малыми буквами). Пусть $e = (e_1, e_2, \dots, e_n) = \|e_p\|_n$. Обозначим $X_e = \|a^k\|^n$ матрицу-столбец, состоящую из координат элемента x по базису $(e_k)_n$. Тогда разложение x по базису $(e_k)_n$ запишется так:

$$x = e_k a^k = e X_e. \quad (*)$$

Лемма 1 Пусть элементы x_1, x_2, \dots, x_m линейного пространства R разложены по базису $(e_p)_n$: $x_k = e_p a_k^p, k = \overline{1, m}$. Тогда из линейной зависимости столбцов матрицы $A = \|a_k^p\|_m^n$ следует линейная зависимость элементов x_1, x_2, \dots, x_m .

Доказательство. Введем обозначение $X_{ek} = \|a_k^p\|_m^n, k = \overline{1, m}$, тогда $x_k = e_p a_k^p = e X_{ek}$. По условию столбцы X_{ek} линейно зависимы, т. е. найдутся такие числа c^1, c^2, \dots, c^m не все равные нулю, что $X_{ek} c^k = \theta$. Рассмотрим линейную комбинацию $x_k c^k = (e X_{ek}) c^k = (\text{почему?}) = e(X_{ke} c^k) = e\theta = (e_1, e_2, \dots, e_n) \cdot \|0\|_m^n = e_1 \cdot 0 + e_2 \cdot 0 + \dots + e_n \cdot 0 = (\text{по т.3 и а.3}) = \theta$, т. е. элементы x_1, x_2, \dots, x_m линейно зависимы.

§5. Размерность линейного пространства.

О п р е д е л е н и е. *Натуральное число n называется размерностью линейного пространства R и обозначается символом $\dim R$, если в линейном пространстве R есть n линейно независимых элементов, а любые $n + 1$ элементов линейно зависимы. Само пространство R называется n -мерным и обозначается символом R_n .*

Иначе говоря, размерность линейного пространства равняется максимальному числу линейно независимых элементов этого пространства.

О п р е д е л е н и е. *Линейное пространство называется бесконечномерным, если в нем найдется любое наперед заданное число линейно независимых элементов этого пространства, и обозначается символом R_∞ .*

Например, линейное пространство $C_{[a,b]}$ бесконечномерно, ибо как показано в примере 4, §3 элементы $e_k = x^{k-1}, k = \overline{1, n}$, где n — любое натуральное число, являются линейно независимыми. Также бесконечномерным является пространство всех многочленов.

Теорема 10 *Если R_n — линейное пространство размерности n , то в нем существует базис из n элементов, причем в качестве базиса можно взять любые n линейно независимых элементов.*

Доказательство. По определению размерности линейного пространства в R_n имеются n линейно независимых элементов e_1, e_2, \dots, e_n . Для того чтобы убедиться в том, что эти элементы образуют базис, достаточно показать, что произвольный элемент x из R_n есть линейная комбинация элементов e_1, e_2, \dots, e_n . По определению размерности линейного пространства элементы x, e_1, e_2, \dots, e_n линейно зависимы, т. е. существуют числа c^0, c^1, \dots, c^n , не все равные нулю, такие, что $x c^0 + e_k c^k = \theta$. Очевидно, что $c^0 \neq 0$, ибо иначе e_1, e_2, \dots, e_n были бы линейно зависимы. Обозначив $x^k = -c^k / c^0, k = \overline{1, n}$, получим $x = e_k x^k$, т. е. $(e_k)_n$ — базис в R_n .

Теорема 11 (Обратная) *Если в линейном пространстве R есть базис из n элементов, то $\dim R = n$.*

Доказательство. Пусть $(e_p)_n$ — базис. По определению базиса e_1, e_2, \dots, e_n линейно независимы. Покажем, что любые $n + 1$ элементов линейно зависимы. Рассмотрим произвольные элементы x_1, x_2, \dots, x_{n+1} . Разложим каждый из них по базису: $x_k = e_p a_k^p, k = \overline{1, n+1}$. Рассмотрим $n \times (n+1)$ -матрицу $A = \|a_k^p\|_{n+1}^n$. Очевидно ее ранг $\leq n$, т. е. меньше числа столбцов матрицы A . Тогда по следствию 4 к теореме о базисном миноре столбцы матрицы A линейно зависимы. Следовательно, по лемме 1 элементы x_1, x_2, \dots, x_{n+1} линейно зависимы. Теорема доказана.

З а м е ч а н и е. Теперь на основании доказанных теорем можно утверждать, что

$$\dim V_3 = 3, \quad \dim T_n = n, \quad \dim A_n^m = m \cdot n.$$

§6. Преобразование базиса и координат элементов линейного пространства.

Пусть в линейном пространстве R_n даны два базиса $(e_k)_n$ и $(\phi_k)_n$. Разложим элементы входящие во второй базис, по первому базису:

$$\phi_k = e_p a_k^p, k = \overline{1, n} \quad (1)$$

Если ввести обозначения $\|a_k^p\|_n = A, \|\phi_k\|_n = \phi, \|e_k\|_n = e$, то равенство (1) можно записать в матричной форме:

$$\phi = eA. \quad (2)$$

Формула (1) (или, что то же самое, формула (2)) дает правило перехода от базиса $(e_k)_n$ к базису $(\phi_k)_n$, причем матрица A называется матрицей перехода. Заметим, что любой k -ый столбец этой матрицы составлен из координат элемента ϕ_k по базису $(e_k)_n$. При этом A — невырожденная матрица, ибо иначе по следствию 1 к теореме о базисном миноре между столбцами матрицы A была бы зависимость, а значит в силу равенства (2) и леммы 1 элементы $\phi_1, \phi_2, \dots, \phi_n$ были бы линейно зависимы, что противоречит определению базиса. Поскольку $D = \det A \neq 0$, то матрица A имеет обратную $A^{-1} = \|b_k^p\|_n$, где $b_k^p = (1/D) \sum_{k=1}^n A_p^k$, а A_p^k — алгебраическое дополнение элемента a_p^k матрицы A . Поэтому, умножив равенство (2) справа на A^{-1} , получим формулу перехода от базиса $(\phi_k)_n$ к базису $(e_k)_n$:

$$e = \phi A^{-1}, \quad (3)$$

или в координатах:

$$e_p = \phi_k b_p^k = 1/D \cdot \sum_{k=1}^n \phi_k A_p^k, \quad p = \overline{1, n}. \quad (4)$$

Итак, установлено, как преобразуется базис при линейном преобразовании.

Поставим еще одну задачу: выяснить, как при таком преобразовании изменятся координаты произвольного элемента x пространства R_n ?

Пусть элемент x в базисе $(e_k)_n$ имеет координаты b^1, b^2, \dots, b^n , т. е. определяется матрицей-столбцом $X_e = \|b^k\|_n$, а в базисе $(\phi_k)_n$ — координатами c^1, c^2, \dots, c^n , т. е. задается матрицей-столбцом $X_\phi = \|c^k\|_n$. Таким образом,

$$x = e_k b^k = e X_e, \quad x = \phi_k c^k = \phi X_\phi.$$

Сравнивая эти равенства, имеем: $e X_e = \phi X_\phi$. Но $\phi = eA$. Поэтому $\phi X_\phi = (eA) X_\phi =$ (свойство 1^o операции умножения матриц) $= e(AX_\phi)$. Следовательно,

$$e X_e = e(AX_\phi).$$

В этом равенстве слева и справа стоят разложения одного и того же элемента по базису $(e_k)_n$. В силу единственности разложения по данному базису имеем:

$$X_e = AX_\phi, \quad (5)$$

или в координатах:

$$b^k = a_p^k c^p, \quad k, p = \overline{1, n}. \quad (6)$$

Умножив равенство (5) на матрицу A^{-1} слева, получим:

$$X_\phi = A^{-1} X_e. \quad (7)$$

Равенство (7) в координатах:

$$c^k = 1/D \cdot \sum_{p=1}^n A_p^k b^p, \quad k = \overline{1, n}. \quad (8)$$

§7. Подпространства линейного пространства.

О п р е д е л е н и е. Подпространством линейного пространства R называется любое непустое множество M элементов этого пространства, на котором корректны операции сложения и умножения на число, введенные в линейном пространстве R , т. е.

а) $\forall x, y \in M \quad x + y \in M$, б) $\forall c \in K$ и $\forall x \in M \quad cx \in M$.

Сформулируем и докажем некоторые свойства подпространства линейного пространства.

Свойство 1 Всякое подпространство линейного пространства R есть линейное пространство.

Д о к а з а т е л ь с т в о. Как следует из определения подпространства нам даны множество M элементов и две линейные операции. Поэтому для доказательства этого свойства достаточно проверить выполнение аксиом линейного пространства. Выполнение а.1), 2), 5) - 8) очевидно, так как они имеют место для всех без исключения элементов пространства R . Осталось проверить а.3), а.4). Пусть x — любой элемент из M . Тогда по определению подпространства: $\forall c \quad cx \in M$. Если $c = 0$, то по теореме 3 $0 \cdot x = \theta$, и, следовательно, множество M содержит нулевой элемент θ , т. е. а.3) выполнена. Если же $c = -1$, то по теореме 4 $(-1)x$ есть элемент противоположный элементу x , и, следовательно, множество M вместе с каждым элементом x содержит и ему противоположный, т. е. а.4) выполнена.

П р и м е р ы.

- 1) Минимальное по размерности подпространство любого линейного пространства R — θ .
- 2) Максимальное по размерности подпространство любого линейного пространства — само линейное пространство.
- 3) Подпространства линейного пространства V_3 : двумерное — множество всех векторов, параллельных какой-либо плоскости; одномерное — множество всех векторов параллельных какой-либо прямой.
- 4) Пространство T_n : пусть $k \leq n$. M — множество всех элементов $\|a^p\|^n$ из T_n , для которых $a^1 = a^2 = \dots = a^k = 0$.
- 5) Пространство A_n^m : пусть $e \leq m, s \leq n$. M — множество всех матриц $\|a_k^p\|_n^m$ из A_n^m , для которых $a_k^p = 0$, $p = \overline{1, e}$, $k = \overline{1, s}$.

Свойство 2 $\dim M \leq \dim R_n$.

Д о к а з а т е л ь с т в о. Пусть x_1, x_2, \dots, x_{n+1} — какие-либо элементы из M . Но тогда они принадлежат пространству R_n , в котором любые $n + 1$ элементов линейно зависимы. Очевидно, что это верно и на M .

Свойство 3 (о пополнении базиса) . Если $(e_p)_k$ — базис в подпространстве M линейного пространства R_n , причем $k < n$, то можно так выбрать элементы в R_n $e_{k+1}, e_{k+2}, \dots, e_n$, что $(e_p)_n$ будет базисом в R_n .

Доказательство. В линейном пространстве R_n существует по крайней мере один элемент e_{k+1} такой, что элементы e_1, e_2, \dots, e_{k+1} — линейно независимы. Действительно, если предположить противное, то $\forall x \in R_n$ существуют числа c^0, c^1, \dots, c^k , не все равные нулю, такие, что $xc^0 + e_p c^p = \theta$. Если $c^0 = 0$, то e_1, e_2, \dots, e_k линейно зависимы, что противоречит предположению. Если $c^0 \neq 0$, то получаем, что любой элемент x из R_n можно разложить по базису $(e_p)_k$, но это противоречит условию $k < n$. Итак, существует элемент e_{k+1} такой, что элементы e_1, e_2, \dots, e_{k+1} линейно независимы. Если $k+1 = n$, то $(e_p)_{k+1}$ — искомый базис в R_n . Если $k+1 < n$, то можно найти такой элемент e_{k+2} , что e_1, e_2, \dots, e_{k+2} — линейно независимы. Продолжая этот процесс, через конечное число шагов построим базис $(e_p)_n$ в пространстве R_n , в котором первые k элементов являются базисом в M .

§8. Линейные оболочки.

Определение. *Линейной оболочкой заданного конечного множества элементов x_1, x_2, \dots, x_k линейного пространства R_n над полем K называется совокупность всех линейных комбинаций этих элементов и обозначается символом $L(x_1, x_2, \dots, x_k) = \{x : x = x_p c^p, c^p \in V_\infty, p = \overline{1, k}\}$, причем множество элементов x_1, x_2, \dots, x_k называется системой, порождающей линейную оболочку $L(x_1, x_2, \dots, x_k)$.*

Свойства линейной оболочки.

Свойство 1 *Линейная оболочка $L(x_1, x_2, \dots, x_k)$ элементов пространства R_n является подпространством линейного пространства R_n .*

Доказательство. Достаточно показать корректность относительно множества L операций сложения и умножения на число введенных в пространстве R_n . Пусть $x, y \in L(x_1, x_2, \dots, x_k)$, т. е. $x = x_p a^p$ и $y = x_p b^p$. Тогда
 а) $x + y = x_p a^p + x_p b^p = (a.2,7) = x_p(a^p + b^p) = x_p c^p \in L$ ($c^p = a^p + b^p$),
 б) $\forall c \in K \quad xc = (x_p a^p)c = (a.2,5) = x_p(a^p c) \in L$.

Свойство 2 *Линейная оболочка $L(x_1, x_2, \dots, x_k)$ является наименьшим подпространством содержащим элементы x_1, x_2, \dots, x_k . (Иными словами: она содержится в любом подпространстве, имеющем эти элементы.)*

Доказательство. Очевидно, что $x_1, x_2, \dots, x_k \in L(x_1, x_2, \dots, x_k)$, которая в силу свойства 1 является подпространством линейного пространства R_n . С другой стороны, любое подпространство пространства R_n , содержащее эти элементы, включают в себя все их линейные комбинации, т. е. содержит в себе $L(x_1, x_2, \dots, x_k)$.

Свойство 3 *Если какой-либо элемент из порождающей системы элементов x_1, x_2, \dots, x_k есть линейная комбинация остальных элементов этой системы, то его можно убрать из порождающей системы, не изменив линейной оболочки.*

Доказательство. Пусть x_k есть линейная комбинация остальных элементов порождающей системы. Тогда, во-первых, $x_k \in L(x_1, x_2, \dots, x_{k-1})$ и, во-вторых, любая линейная комбинация элементов x_1, x_2, \dots, x_k сводится к некоторой линейной комбинации элементов x_1, x_2, \dots, x_{k-1} , т. е. $L(x_1, x_2, \dots, x_k) \subset L(x_1, x_2, \dots, x_{k-1})$. С другой стороны, по свойству 2 $L(x_1, x_2, \dots, x_{k-1}) \subset L(x_1, x_2, \dots, x_k)$. Таким образом, из этих двух включений имеем, что $L(x_1, x_2, \dots, x_k) = L(x_1, x_2, \dots, x_{k-1})$.

Примеры.

- 1) Линейная оболочка базисных элементов e_1, e_2, \dots, e_n линейного пространства R_n совпадает со всем линейным пространством R_n , т. е. $L(e_1, e_2, \dots, e_n) = R_n$.
- 2) В пространстве V_3 линейная оболочка пары неколлинеарных векторов \vec{a}, \vec{b} состоит из всех векторов, параллельных плоскости векторов \vec{a}, \vec{b} , т. е. $L(\vec{a}, \vec{b})$ есть подпространство V_2 пространства V_3 .
- 3) В линейном пространстве $C_{[a,b]}$ линейная оболочка функций $1, x, \dots, x^k$ есть множество всех многочленов степени $\leq k$.

Свойство 4 (Размерность и базис линейной оболочки). Пусть среди порождающих элементов x_1, x_2, \dots, x_k имеются $p \leq k$ линейно независимых элементов, например, x_1, x_2, \dots, x_p , а остальные x_{p+1}, \dots, x_k представляют собой линейные комбинации данных p элементов. Тогда

- а) $L(x_1, x_2, \dots, x_k) = L(x_1, x_2, \dots, x_p)$,
- б) $\dim L(x_1, x_2, \dots, x_k) = p$,
- в) $(x_i)_p$ — базис линейной оболочки.

Доказательство. Утверждение а) следует из свойства 3. По условию элементы x_1, x_2, \dots, x_p линейно независимы и по утверждению а) любой элемент линейной оболочки $L(x_1, x_2, \dots, x_k)$ является линейной комбинацией указанных p элементов. Следовательно, эти p элементов образуют базис линейной оболочки, и ее размерность по теореме 11 равна p .

Следствие. $\dim L(x_1, x_2, \dots, x_k)$ равняется максимальному числу линейно независимых элементов в порождающей системе элементов x_1, x_2, \dots, x_k .

Рассмотрим одно из приложений понятия линейной оболочки.

§9. Ранг матрицы и размерность линейной оболочки ее столбцов. Вычисление ранга матрицы.

Рассмотрим матрицу $A = \|a_k^p\|_m^n$. Столбцы $A_k = \|a_k^p\|_m, k = \overline{1, n}$ матрицы A являются элементами линейного пространства T_m . Пусть $\text{rang } A = r$ и базисными столбцами являются первые r столбцов A_1, A_2, \dots, A_r . Образует линейную оболочку столбцов матрицы A $L(A_1, A_2, \dots, A_n)$.

Теорема 12 Размерность линейной оболочки столбцов матрицы A равна рангу матрицы A , т. е. $\dim L(A_1, A_2, \dots, A_n) = \text{rang } A$. Базисные столбцы матрицы A являются базисом линейной оболочки столбцов матрицы A .

Доказательство. По теореме о базисном миноре: 1) столбцы A_1, A_2, \dots, A_r линейно независимы, 2) любой столбец матрицы A есть линейная комбинация базисных столбцов. Но тогда из свойства 4 линейной оболочки следует утверждение теоремы.

Следствие 1. Если $\text{rang } A = r$, то любые $r + 1$ столбцов (если они есть) линейно зависимы.

Утверждение очевидно, ибо в пространстве размерности r любые $r + 1$ элементов линейно зависимы.

Следствие 2. Ранг матрицы A равен максимальному числу линейно независимых столбцов. (См. следствие к свойству 4 линейной оболочки).

Следствие 3. Максимальное число линейно независимых столбцов равняется максимальному числу линейно независимых строк. (При операции транспонирования ранг матрицы не меняется).

Рассмотрим некоторые элементарные операции над столбцами (строками) матрицы, которые, меняя матрицу, не меняют ранга. При доказательстве сохранения ранга при совершении этих операций основным инструментом будет теорема 12. Формулировки приведем только для столбцов, так как для строк они идентичны.

- 1^o Перестановка столбцов. Переставим столбцы A_1 и A_2 . Тогда $L(A_1, A_2, \dots, A_k) = L(A_2, A_1, \dots, A_k)$. Поэтому ранг не изменится.
- 2^o Умножение любого столбца на число $b \neq 0$. Пусть первый столбец умножен на b . Тогда $L(A_1b, A_2, \dots, A_k) = L(A_1, A_2, \dots, A_k)$. Действительно, если $x \in L(A_1b, A_2, \dots, A_k)$, то $x = (A_1b)c^1 + A_p c^p = A_1(bc^1) + A_p c^p \in L(A_1, A_2, \dots, A_k)$. Верно и обратное: если $x \in L(A_1, A_2, \dots, A_k)$, то $x = A_p c^p = (A_1b) \cdot c^1/b + A_p c^p$, где справа $p = \overline{2, n}$, т. е. $x \in L(A_1b, A_2, \dots, A_k)$. Поэтому ранг не изменится.
- 3^o Прибавление к одному столбцу другого, умноженного на число b . Пусть к A_1 прибавлен столбец A_2b . Тогда $L(A_1 + A_2b, A_2, \dots, A_k) = L(A_1, \dots, A_n)$. В самом деле, если $x \in L(A_1 + A_2b, A_2, \dots, A_k)$, то $x = (A_1 + A_2b)c^1 + A_2c^2 + A_p c^p = A_1c^1 + A_2(bc^1 + c^2) + A_p c^p = A_1c^1 + A_2c^2 + A_p c^p \in L(A_1, A_2, \dots, A_k)$. Аналогично доказывается, что если $x \in L(A_1, A_2, \dots, A_k)$, то $x \in L(A_1 + A_2b, A_2, \dots, A_k)$. А тогда из первого равенства следует, что ранг не меняется.
- 4^o Вычеркивание нулевого столбца. Пусть A_1 — нулевой столбец. Тогда по свойству 3 $L(A_1, A_2, \dots, A_k) = L(A_2, \dots, A_k)$, что означает, что ранг не меняется.
- 5^o Вычеркивание столбца, являющегося линейной комбинацией других столбцов. Ранг не меняется, что следует из свойства 3 линейной оболочки.

Опираясь на элементарные операции 1^o – 5^o, сформулируем метод вычисления ранга матрицы.

Если все элементы матрицы A равны нулю, то ранг матрицы A равен нулю.

Если имеются элементы матрицы, отличные от нуля, то

Шаг 1. а) Переставляя строки и столбцы, переведем один из них в левый верхний угол, т. е. на место элемента a_1^1 .

б) Далее комбинируя первую строку (первый столбец) с какой-либо другой (другим), сделаем элемент, стоящий на месте a_1^1 , равным 1 или -1.

в) Все элементы первой строки, кроме первого, сделаем равными нулю, вычитая из всех столбцов поочередно первый, умноженный на соответствующий множитель. В итоге получим:

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

Если все элементы в заштрихованной зоне равны нулю, то ранг матрицы A равен 1.

Если же в этой зоне есть элементы, отличные от нуля, то

Шаг 2. а) Переставляя строки и столбцы, поставим, один из них на место элемента a_2^2 .

б) Комбинируя вторую строку (второй столбец) с какойлибо другой (другим), сделаем элемент, стоящий на месте a_2^2 , равным 1 или -1 (при этом ни в коем случае не трогать 1 столбец и первую строку!).

в) Все элементы второй строки, лежащие правее элемента a_2^2 , сделаем равными нулю. В итоге получим:

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

Если в заштрихованной зоне все элементы равны нулю, то $\text{rang } A = 2$. Если же в этой зоне есть элементы, отличные от нуля, то все повторяем в полной аналогии с предыдущими двумя шагами, но уже в третьей строке и т. д.

В итоге на шаге r приходим к матрице:

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

ранг которой равен r .

Пример. Вычислим ранг матрицы A (знак \sim означает эквивалентность матрицы в смысле ранга).

$$A = \begin{pmatrix} 0 & \textcircled{2} & 2 \\ \Delta & -3 & -1 \\ -2 & 0 & -4 \\ 4 & 6 & 14 \end{pmatrix}$$

Вынесем из 1, 3, 4 строк множитель 2 и -2:

$$\sim \begin{pmatrix} 0 & \textcircled{1} & 1 \\ \Delta & -3 & -1 \\ 1 & 0 & 2 \\ 2 & 3 & 7 \end{pmatrix} \sim \begin{pmatrix} \textcircled{1} & 0 & 1 \\ -3 & \Delta & -1 \\ 0 & 1 & 2 \\ 3 & 2 & 7 \end{pmatrix} \sim \begin{pmatrix} \textcircled{1} & 0 & 0 \\ -3 & \Delta & 2 \\ 0 & 1 & 2 \\ 3 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} \textcircled{1} & 0 \\ -3 & \Delta \\ 0 & 1 \\ 3 & 2 \end{pmatrix}.$$

Итак, ранг матрицы равен 2. Как найти базисный минор M_b исходной матрицы? Маркируем базисные столбцы и строки в последней матрице ($\textcircled{1}$ — первая строка и первый столбец, Δ — вторая строка и второй столбец). Следим за маркированными местами, проходя в обратном порядке все преобразования по вычислению ранга матрицы. Затем вычеркиваем в исходной матрице маркированные строки и столбцы. На пересечении вычеркнутых столбцов и строк стоят элементы, образующие квадратную матрицу, определитель которой и есть искомым базисный минор исходной матрицы, т. е.

$$M_b = \begin{vmatrix} 0 & 2 \\ 1 & -3 \end{vmatrix}.$$

Докажем теперь две теоремы о ранге произведения матриц.

Теорема 13 Ранг произведения двух матриц не превосходит ранга каждого из сомножителей.

Доказательство. Пусть даны матрицы $A = \|a_k^p\|_m^n$ и $B = \|b_k^p\|_s^m$. Тогда $C = AB = \|c_k^p\|_s^n$. Докажем:

1) $\text{rang } C \leq \text{rang } A$. Так как по определению произведения матриц $c_k^p = a_e^p b_k^e$, то зафиксировав k и меняя p от 1 по n , получим:

$$C_k = A_p b_k^p,$$

где $C_k, k = \overline{1, s}$ — столбцы матрицы C , $A_p, p = \overline{1, m}$ — столбцы матрицы A . По теореме 12 $\text{rang } C = \dim L(C_1, C_2, \dots, C_s)$, а $\text{rang } A = \dim L(A_1, A_2, \dots, A_m)$. Так как каждый столбец C_k есть линейная комбинация столбцов A_1, A_2, \dots, A_m , то $L(C_1, C_2, \dots, C_s) \subset L(A_1, A_2, \dots, A_m)$.

Следовательно, $\dim L(C_1, C_2, \dots, C_s) \leq \dim L(A_1, A_2, \dots, A_m)$, а значит $\text{rang } C \leq \text{rang } A$.

2) $\text{rang } C \leq \text{rang } B$. $\text{rang } C = \text{rang } C^T = \text{rang}(AB)^T =$ (т. 4, гл. 4 курса аналитической геометрии) $= \text{rang}(B^T A^T) \leq$ (по 1) теоремы 13) $\leq \text{rang } B^T = \text{rang } B$.

Теорема 14 Если A — невырожденная $n \times n$ -матрица, а B — любая $n \times n$ -матрица, то $\text{rang } AB = \text{rang } B$.

Доказательство. Обозначим $AB = C$. По теореме 13 $\text{rang } C \leq \text{rang } B$. Так как $\det A \neq 0$, то существует обратная матрица A^{-1} . Умножив равенство $AB = C$ слева на A^{-1} , получим: $B = A^{-1}C$. Тогда по теореме 13 $\text{rang } B \leq \text{rang } C$. Из двух полученных неравенств следует, что $\text{rang } B = \text{rang } C$.

§10. Изоморфизм линейных пространств.

Оказывается, различные линейные пространства одной и той же размерности n с алгебраической точки зрения тождественны. Поэтому, если в каком-либо конкретном n -мерном линейном пространстве доказана теорема, сформулированная в терминах линейных операций, то эта теорема верна в любом другом линейном пространстве той же размерности n . Выгода такого подхода очевидна. Поэтому обоснуем его.

Определение. Соответствие Γ между элементами двух линейных пространств R и R' ($\Gamma: R \rightarrow R'$) называется взаимно однозначным, если при этом соответствии

- 1) каждому элементу из R отвечает один и только один элемент из R' ,
- 2) каждому элементу из R' отвечает один и только один элемент из R .

Определение. Два линейных пространства R и R' называются изоморфными, если между элементами этих пространств можно установить взаимно однозначное соответствие Γ так, что если элементам x, y из R отвечают соответственно элементы x', y' из R' , то элементу $x + y$ отвечает элемент $x' + y'$, а элементу bx из R ($\forall b \in K$) отвечает элемент bx' из R' (иными словами, существует $\Gamma: R \rightarrow R'$ такое, что $\Gamma(x + y) = \Gamma(x) + \Gamma(y)$ и $\Gamma(bx) = b\Gamma(x)$), в этом случае соответствие Γ называется изоморфизмом.

Какие же пространства являются изоморфными? Ответ на этот вопрос дают следующие две теоремы. Но в начале докажем одно вспомогательное утверждение.

Лемма 2 Изоморфизм $\Gamma: R \rightarrow R'$ переводит нулевой элемент θ пространства R в нулевой элемент θ' пространства R' .

Доказательство. Из определения изоморфизма следует, что $\forall b \in K \Gamma(bx) = b\Gamma(x)$. В частности, при $b = 0$ $\Gamma(0 \cdot x) = 0 \cdot \Gamma(x)$. Но по теореме 3 а) $0 \cdot x = \theta \in R$, так как $x \in R$, б) $0 \cdot \Gamma(x) = \theta' \in R'$, так как $\Gamma(x) \in R'$. Поэтому $\Gamma(\theta) = \theta'$. Что и требовалось доказать.

Теорема 15 Все линейные пространства одной и той же размерности n изоморфны между собой.

Доказательство. Достаточно доказать, что любое линейное пространство R_n изоморфно линейному пространству T_n , элементами которого являются матрицы-столбцы $\|x^p\|^n$ высоты n . Пусть $(e_k)_n$ — базис в линейном пространстве R_n . Тогда каждый элемент x из R_n имеет в этом базисе разложение: $x = e_k x^k$. Каждому элементу x из R_n поставим в соответствие n действительных чисел x^1, x^2, \dots, x^n — координат этого элемента в базисе $(e_k)_n$, т. е. вполне определенный элемент $x' = \|x^k\|^n$ пространства T_n . Установленное соответствие взаимно однозначно. Более того, если элементам $x = e_k x^k, y = e_k y^k$ из R_n отвечают соответственно столбцы $x' = \|x^k\|^n, y' = \|y^k\|^n$ из T_n , то в силу теоремы 9 элементу $x + y$ из R_n соответствует столбец $\|x^k + y^k\|^n$, т. е. элемент $x' + y'$ из T_n , а элементу bx из R_n соответствует столбец $\|bx^k\|^n$, т. е. элемент bx' из T_n . Теорема доказана.

Теорема 16 Любые два линейных пространства разных размерностей не изоморфны.

Доказательство. Рассмотрим линейные пространства R_n и R_m . Пусть $m < n$. В пространстве R_n существуют n линейно независимых элементов e_1, e_2, \dots, e_n . Если бы существовал изоморфизм, то элементы e_1, e_2, \dots, e_n перешли бы в n элементов $\phi_1, \phi_2, \dots, \phi_n$ пространства R_m , которые необходимо были бы зависимыми, ибо число элементов n больше размерности линейного пространства R_m . Постольку в силу леммы 2 при изоморфизме нулевой элемент переходит в нулевой элемент, то равной нулевому элементу линейной комбинации образов $\phi_k c^k, k = \overline{1, n}$, где не все c^k равны нулю, должна соответствовать равная нулевому элементу линейная комбинация прообразов $e_k c^k, k = \overline{1, n}$, что противоречит предположению об их линейной независимости. Теорема доказана.

Гл. 3. Система линейных уравнений.

§1. Критерий совместности общей линейной системы уравнений.

Рассмотрим линейную систему m уравнений с n неизвестными:

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

где $A = \|a_k^p\|_n^m$, $X = \|x^p\|_n$, $B = \|b^p\|_n^m$. Матрица A называется основной матрицей системы. Если в матрице A добавить еще один столбец B , то получим матрицу A^* , называемую расширенной матрицей системы (1).

Теорема 1 (Кронекера-Капелли) Для того чтобы система (1) была совместна, необходимо и достаточно, чтобы $\text{rang } A = \text{rang } A^*$.

Необходимость. Пусть система (1) совместна, т. е. существуют числа c^1, c^2, \dots, c^n , которые при подстановке в систему (1) обращают все уравнения в верные равенства. Эти равенства означают, что столбец свободных членов B есть линейная комбинация столбцов матрицы A , т. е. этот столбец можно вычеркнуть из матрицы A^* (элементарная операция 5°, §9, гл. 2), не изменив при этом ранга. Таким образом, $\text{rang } A = \text{rang } A^*$.

Достаточность. Пусть $\text{rang } A = \text{rang } A^*$. Тогда базисный минор матрицы A является базисным минором матрицы A^* . Следовательно, по теореме о базисном миноре столбец свободных членов B есть линейная комбинация базисных столбцов, а значит и всех столбцов матрицы A с числовыми коэффициентами c^1, c^2, \dots, c^n . Эти числа и являются решением системы. Теорема доказана.

§2. Однородные линейные системы уравнений.

Однородная система уравнений в матричной форме записывается так:

$$AX = \theta, \quad (2)$$

где A, X те же что и в §1, а θ — нулевой столбец высоты m . Пусть $\text{rang } A = r$. Как следует из теоремы 4 гл. 1, система (2) нетривиально совместна при $r < n$. Заметим, что всякое решение $\|c^p\|_n$ системы (2) является элементом пространства T_n . Очевидно, что не всякий элемент пространства T_n является решением системы (2). Таким образом, множество M всех решений системы (2) образует подмножество множества всех элементов линейного пространства T_n . Операции сложения и умножения на число, введенные в пространстве T_n , корректны на множестве M , т. е. а) $\forall X, Y$ — решений системы (2) $X + Y$ тоже решение (2), т. е. $X + Y \in M$; б) $\forall b \in K$ и $\forall X \in M$ $bX \in M$, так как bX — решение системы (2).

Итак, множество M всех решений системы (2) с указанными линейными операциями является подпространством линейного пространства T_n . Осталось определить его размерность.

Определение. *Фундаментальной совокупностью решений (ФСР) системы (2) называется базис в пространстве решений.*

Иными словами ФСР — максимальное число линейно независимых решений. Так как в любом линейном пространстве базисов бесконечно много, то и ФСР также бесконечно много. Если выбрана какая-либо ФСР, то любое решение системы (2) есть некоторая линейная комбинация решений из ФСР и обратно, так как ФСР —

базис в пространстве решений, то любая линейная комбинация решений из ФСР — решение системы (2). Поэтому, чтобы найти всю совокупность решений системы (2), достаточно построить ФСР. Число решений в ФСР определит согласно теореме 11 размерность пространства решений.

Будем считать, не умаляя общности рассуждений, что базисный минор M_b находится в левом верхнем углу матрицы A , т. е. $M_b = \det \|a_k^p\|_r^r$. В матрице $A = \|a_k^p\|_n^m$ каждая из последних $m - r$ строк есть линейная комбинация базисных строк, т. е. первых r строк. Поэтому в системе (2) последние $m - r$ уравнений являются следствием первых r . Значит, последние $m - r$ уравнений из системы (2) можно убрать и решать систему лишь из r первых уравнений. Кроме того, в левой части системы оставим лишь r первых неизвестных, а последние $n - r$ неизвестных перенесем в правую часть уравнений, задав их произвольным образом, т. е. $x^{r+1} = c^{r+1}, \dots, x^n = c^n$. Итак, система (2) примет вид:

$$a_k^p x^k = -a_e^p c^e, \quad p = \overline{1, r}, \quad k = \overline{1, r}, \quad e = \overline{r+1, n}. \quad (3)$$

Система (3) — система r уравнений с r неизвестными и определителем $M_b \neq 0$. Таким образом, если фиксированы значения c^{r+1}, \dots, c^n , то система (3) имеет единственное решение, определяемое формулами Крамера. Будем задавать значения c^{r+1}, \dots, c^n согласно ниже приведенной таблицы:

	1	2		$n - r$
c^{r+1}	1	0	...	0
c^{r+2}	0	1	...	0
...
c^n	0	0	...	1

 $\Rightarrow Y_1 = \begin{pmatrix} c_1^1 \\ \dots \\ c_1^r \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, Y_2 = \begin{pmatrix} c_2^1 \\ \dots \\ c_2^r \\ 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \dots, Y_{n-r} = \begin{pmatrix} c_{n-r}^1 \\ \dots \\ c_{n-r}^r \\ 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}$

т. е. каждому набору значений c^p будет соответствовать решение системы (2), причем эти решения Y_1, Y_2, \dots, Y_{n-r} линейно независимы, ибо ранг матрицы, составленной из этих столбцов, равен $n - r$. Чтобы убедиться в том, что построенное множество решений образует базис в пространстве решений, достаточно показать, что любое решение $X^* = \|c^k\|_n$ системы (2) является линейной комбинацией указанных решений. Рассмотрим линейную комбинацию

$$X = Y_k c^{k+r}, \quad k = \overline{1, n-r}, \quad (4)$$

где c^{1+r}, \dots, c^n — $(n - r)$ последних координат решения X^* . Так как у X и X^* последние $n - r$ координат совпадают, то это означает, что они удовлетворяют системе (3) с одной и той же правой частью. А так как система (3) при этом имеет единственное решение, то $X = X^*$. Таким образом, доказано, что Y_1, Y_2, \dots, Y_{n-r} — ФСР и по теореме 11 размерность пространства решений системы (2) равна $n - r$.

З а м е ч а н и е 1. Способ построения ФСР, приведенный выше, где все свободные члены полагались равными нулю, кроме одного, равного 1, дает особую ФСР, именуемую нормальной. Очевидно, значения c^{r+1}, \dots, c^n можно задавать иначе, полагая другие ФСР.

З а м е ч а н и е 2. Формула (4) при условии, что в ней c^{r+1}, \dots, c^n принимают любые значения от $-\infty$ до ∞ независимо друг от друга, определяет всю совокупность решений системы (2). Отсюда следует, что множество решений системы (2) — линейная оболочка, натянутая на порождающую систему векторов Y_1, Y_2, \dots, Y_{n-r} .

§3. Общее решение неоднородной линейной системы уравнений.

Снова рассмотрим систему (1). Будем считать ее совместной, т. е. $\text{rang } A = \text{rang } A^* = r$. Определим все множество решений системы (1). Пусть X_0 — некоторое решение системы (1), которое обычно называют частным решением системы (1). Пусть Y_1, Y_2, \dots, Y_{n-r} — ФСР соответствующей системе (1) однородной системы уравнений, т. е. системы (2). Тогда $y_k c^k, k = \overline{1, n-r}$, — общее решение системы (2), где c^1, c^2, \dots, c^{n-r} — любые действительные числа.

Теорема 2 *Общее решение совместной неоднородной системы уравнений (1), ранг матрицы которой равен r , имеет вид:*

$$X = X_0 + Y_k c^k, \quad k = \overline{1, n-r}. \quad (5)$$

Доказательство. Надо доказать: 1) при любых c^1, c^2, \dots, c^{n-r} формула (5) определяет решение системы (1) и 2) каждое решение X^* системы (1) при определенном выборе $n-r$ постоянных задается формулой (5).

1) Умножим формулу (5) на матрицу A слева: $AX = AX_0 + A(Y_k c^k) = B + \theta = B$, т. е. первое утверждение доказано.

2) Рассмотрим $X^* - X_0$. Тогда $A(X^* - X_0) = AX^* - AX_0 = B - B = \theta$, т. е. $X^* - X_0$ есть решение системы (2), а значит существуют такие числа c^1, c^2, \dots, c^{n-r} , что $X^* - X_0 = Y_k c^k$ по формуле (4). Отсюда $X^* = X_0 + Y_k c^k$. Теорема доказана.

З а м е ч а н и е. Чтобы найти какое-нибудь частное решение X_0 , нужно взять r базисных уравнений системы (1) и в них положить x^{r+1}, \dots, x^n равными чему угодно, например, равными нулю. Затем по формулам Крамера найти x^1, x^2, \dots, x^r .

О п р е д е л е н и е. Пусть M - подпространство линейного пространства R , x — любой элемент из M , а x^0 — любой фиксированный элемент из R . Тогда совокупность H всех элементов $y = x^0 + x$ называется результатом сдвига подпространства M вдоль элемента x^0 или гиперплоскостью.

З а м е ч е н и е. Если $x^0 \in M$, то гиперплоскость H является подпространством линейного пространства R той же размерности, что и M . Если же $x^0 \notin M$, то гиперплоскость H не является линейным пространством. Действительно, как было доказано ранее в линейном пространстве R нулевой элемент один (теорема 1, гл. 2). Этот же элемент является нулевым во всех подпространствах этого пространства. А в гиперплоскость H не входит нулевой элемент θ . Предположим противное, т. е. существует элемент x из M такой, что $x^0 + x = \theta$, т. е. $x = -x^0 = (-1)x^0$. Но $x^0 \notin M$ т. е. $x^0 \in R - M$. Поэтому в силу корректности операции умножения на число и элемент $(-1)x^0 = x \in R - M$, т. е. $x \notin M$, что противоречит предположению. Значит гиперплоскость H не имеет нулевого элемента и, следовательно, не является линейным пространством.

Опираясь на только что введенное понятие, можно утверждать, что совокупность решений неоднородной системы уравнений не является линейным пространством, а образует гиперплоскость.

П р и м е р. Найти общее решение следующей системы уравнений:

$$\begin{cases} x^1 - 2x^2 + x^3 + 3x^4 = 2 \\ x^2 - x^3 - x^4 = -1 \\ x^1 - x^2 + 2x^4 = 1 \end{cases} \quad (6)$$

Определим ранг матрицы A системы (6).

$A = \begin{pmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 1 & -1 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$, т. е. ранг матрицы A равен 2. Последний столбец в матрице A^* отличается от второго столбца матрицы A множителем -1 .

Значит его можно убрать из матрицы A^* , не изменив при этом ранга. Следовательно, $\text{rang } A^* = \text{rang } A$, т. е. система (6) совместна. Кроме того, $n - r = 4 - 2 = 2$. Таким образом, исходная система (6) эквивалентна, например, такой системе:

$$\begin{cases} x^2 = c^3 + c^4 - 1 \\ x^1 - x^2 = -2c^4 + 1 \end{cases} \quad (7)$$

Полагая здесь $c^3 = c^4 = 0$, получим систему уравнений:

$$\begin{cases} x^2 = -1 \\ x^1 - x^2 = 1. \end{cases}$$

Решив ее, найдем частное решение системы (6):

$$X_0 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Теперь рассмотрим соответствующую системе (7) однородную систему уравнений:

$$\begin{cases} x^2 = c^3 + c^4 \\ x^1 - x^2 = -2c^4. \end{cases}$$

Строим ФСР:

$$\left[\begin{array}{cc|c} c^3 & c^4 & \\ \hline 1 & 0 & x^1 = 1, x^2 = 1 \\ 0 & 1 & x^1 = -1, x^2 = 1 \end{array} \right] \Rightarrow Y_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Общее решение исходной системы уравнений (6):

$$X = X_0 + Y_1 a^1 + Y_2 a^2 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + a^1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + a^2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a^1 - a^2 \\ -1 + a^1 + a^2 \\ a^1 \\ a^2 \end{pmatrix}. \text{ Непосредственной проверкой убеждаемся, что это множество решений не содержит нулевого элемента } \theta = \|0\|^4, \text{ т. е. не является линейным пространством, но является гиперплоскостью.}$$

Гл. 4. Евклидово пространство.

§1. Определение.

В математическом анализе уже определено евклидово пространство как множество всевозможных упорядоченных совокупностей из n действительных чисел (x_1, x_2, \dots, x_n) , называемых точками, в котором введено расстояние между двумя точками $M_1(x_1, x_2, \dots, x_n)$ и $M_2(y_1, y_2, \dots, y_n)$ по формуле $\rho(M_1, M_2) = ((x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2)^{1/2}$. Это пример конкретного евклидова пространства ибо его элементами являются конкретные объекты — упорядоченные совокупности n действительных чисел. В нашем курсе изучаются абстрактные линейные пространства, элементами которых являются объекты любой природы (векторы, числа, столбцы, многочлены, функции, матрицы и т. д.). До сих пор мы не вводили понятия расстояния, понятий угла и длины в абстрактном линейном пространстве. Сделаем это, т. е. введем абстрактное евклидово пространство. Таким образом, евклидово пространство — это линейное пространство, в котором введены метрические соотношения, или, как говорят, введена метрика. Однако вначале вспомним метрические соотношения в пространстве V_3 , введенные нами в аналитической геометрии: длину вектора, угол между векторами и скалярное произведение двух векторов. Заметим, что знание скалярного произведения любых двух векторов позволяет найти как длину вектора: $|\vec{a}| = \sqrt{(\vec{a}, \vec{a})}$, так и угол между векторами: $\cos \phi = (\vec{a}, \vec{b}) / (|\vec{a}| |\vec{b}|)$. Следовательно, в понятии скалярного произведения заключена возможность измерения длин и углов, т. е. все, что связано с измерениями. В общем абстрактном линейном пространстве нам будет удобнее ввести сначала скалярное произведение двух элементов пространства, а затем из скалярного произведения получить определения длины вектора и угла между элементами.

В этой главе будем рассматривать только числовое поле K_0 — поле всех действительных чисел.

О п р е д е л е н и е. *Линейное пространство называется евклидовым, если в нем указано правило, ставящее в соответствие каждому двум элементам x, y число, которое обозначим (x, y) и назовем скалярным произведением элементов x, y . Это правило удовлетворяет следующим аксиомам:*

$$a.1^{\circ} \quad (x, y) = (y, x) \text{ (переместительное свойство),}$$

$$a.2^{\circ} \quad (x + y, e) = (x, e) + (y, e) \text{ (распределительное свойство),}$$

$$a.3^{\circ} \quad \forall a \in K_0 \quad (ax, y) = a(x, y),$$

$$a.4^{\circ} \quad (x, x) > 0 \text{ при } x \neq \theta \text{ и } (x, x) = 0 \text{ для } x = \theta.$$

П р и м е р ы.

1) V_3 : $(\vec{a}, \vec{b}) = |\vec{a}| |\vec{b}| \cos(\vec{a}, \vec{b})$. Аксиомы $1^{\circ} - 4^{\circ}$ были установлены в векторной алгебре (в курсе аналитической геометрии).

2) T_n : $x = \|x^k\|^n, y = \|y^k\|^n$. Введем скалярное произведение так: $(x, y) = x^1 y^1 + x^2 y^2 + \dots + x^n y^n$. Проверьте сами выполнение аксиом $1^{\circ} - 4^{\circ}$ в этом случае.

3) $C_{[a,b]}$: $x(c), y(c)$ — элементы этого пространства. Пусть $(x, y) = \int_a^b x(c)y(c)dc$. Проверьте сами выполнение аксиом $1^{\circ} - 4^{\circ}$.

Евклидово пространство обозначается символом E или, если известна его размерность n , то — E_n . Если $(e_k)_n$ — базис в евклидовом пространстве E_n и x, y —

элементы этого пространства, то $x = e_p x^p = e X_e, y = e_k y^k = e Y_e$, где использованы обозначения, введенные в §6, гл. 2. Но тогда, как следует из $a.1 - 3$,

$$(x, y) = (e_p x^p, e_k y^k) = \sum_{p,k=1}^n (e_p, e_k) x^p y^k.$$

Если ввести обозначения $a_{pk} = (e_p, e_k)$, то окончательно получим:

$$(x, y) = a_{pk} x^p y^k, \quad p, k = \overline{1, n}. \quad (1)$$

Используя определение произведения матриц, равенство (1) можно записать в матричной форме, если обозначить $A = ||a_{pk}||_{n,n}$:

$$(x, y) = X_e^T A Y_e. \quad (1')$$

Формула (1) (или, что то же самое — формула (1')) — выражение скалярного произведения в произвольном базисе. Из равенства (1) следует, что для каждого элемента x из E_n

$$(x, \theta) = (\theta, x) = 0. \quad (2)$$

§2. Основные метрические понятия в евклидовом пространстве. Неравенство Коши-Буняковского.

О п р е д е л е н и е. *Длиной элемента x пространства E называется $\sqrt{(x, x)}$ и обозначается $|x|$.*

З а м е ч а н и е. В некоторых монографиях по алгебре, например, книге Ильина В.А. и Поздняка Э.Г., таким образом введенная длина элемента называется нормой элемента из E .

О п р е д е л е н и е. *Если $|x| = 1$, то элемент x называется нормированным элементом пространства E .*

П р и м е р ы.

1) V_3 : $|\vec{a}| = \sqrt{(\vec{a}, \vec{a})}$ — обычная длина вектора.

2) T_n : $x = ||x^k||^n, |x| = (x^1 x^1 + x^2 x^2 + \dots + x^n x^n)^{1/2}$.

3) $C_{[a,b]}$: $|x| = (\int_a^b x^2(c) dc)^{1/2}$. Эту величину обозначают иногда $||x(c)||$ и называют нормой функции $x(c)$, чтобы избежать ложных ассоциаций, связанных со словами "длина функций".

О п р е д е л е н и е. *Углом между ненулевыми элементами x, y из E называется угол ϕ , определяемый условиями: 1) $\cos \phi = (x, y)/(|x| \cdot |y|)$, 2) $0 \leq \phi \leq \pi$.*

В пространстве V_3 это определение согласуется с обычным выражением угла через скалярное произведение. В общем случае надо доказать корректность этого определения, а именно показать, что для любых элементов x, y из E $(x, y)/(|x| \cdot |y|)$ по модулю не превосходит 1 или, что то же самое,

$$|(x, y)| \leq |x| \cdot |y|.$$

Последнее неравенство называется неравенством Коши-Буняковского.

Теорема 1 Для любых элементов x, y из E имеет место неравенство Коши-Буняковского.

Доказательство. Рассмотрим элемент $ax - y$, где a — любое число из K_0 . По $a.4$ для каждого a имеем: $(ax - y, ax - y) \geq 0$. А в силу $a.1 - 3$ это неравенство можно записать так: $a^2(x, x) - 2a(x, y) + (y, y) \geq 0$. Отсюда следует, что дискриминант квадратного трехчлена $D = (x, y)^2 - (x, x)(y, y) \leq 0$, т. е. $(x, y)^2 \leq (x, x) \cdot (y, y)$ или $|(x, y)| \leq \sqrt{(x, x)(y, y)} = |x||y|$. Что и требовалось доказать.

Из теоремы 1 следует корректность определения угла между элементами из E . Приведем примеры неравенства Коши-Буняковского в конкретных пространствах:

$$1) V_3 : |(\vec{a}, \vec{b})| \leq |\vec{a}| \cdot |\vec{b}|.$$

$$2) T_n : \left| \sum_{k=1}^n x^k y^k \right| \leq \left(\sum_{k=1}^n (x^k)^2 \right)^{1/2} \cdot \left(\sum_{k=1}^n (y^k)^2 \right)^{1/2}.$$

$$3) C_{[a,b]} : \left| \int_a^b x(c)y(c)dc \right| \leq \left(\int_a^b x^2(c)dc \right)^{1/2} \cdot \left(\int_a^b y^2(c)dc \right)^{1/2}.$$

На дом: Для каких элементов x, y имеет место равенство $|(x, y)| = |x| \cdot |y|$?

§3. Ортогональность элементов в евклидовом пространстве.

Ортонормированный базис в евклидовом пространстве.

Определение. Элементы x, y из E называются ортогональными, если $(x, y) = 0$.

По определению угла между элементами x, y это означает, что при $x \neq \theta, y \neq \theta$ $\cos \phi = 0$, т. е. $\phi = \pi/2$.

Лемма 1 Взаимно ортогональные ненулевые элементы x_1, x_2, \dots, x_k в E линейно независимы.

Доказательство. Предположим противное. Тогда $x_p c^p = \theta, p = \overline{1, k}$, где не все c^p равны нулю. Пусть $c^1 \neq 0$. Умножим это равенство по правилу скалярного произведения на x_1 . В силу предположения о взаимной ортогональности элементов x_1, x_2, \dots, x_k т. е. $(x_e, x_p) = 0$ при $e \neq p$, и того, что $(\theta, x_1) = 0$, мы получим: $c^1(x_1, x_1) = 0$. Отсюда следует, что $(x_1, x_1) = 0$, и, следовательно, $x_1 = \theta$, что противоречит условию леммы.

Определение. Система элементов e_1, \dots, e_n , удовлетворяющая условию $(e_p, e_k) = \delta_{pk}$, называется ортонормированной системой элементов.

Таким образом, система ортонормированных элементов — это совокупность нормированных, взаимно ортогональных элементов.

Теорема 2 В n -мерном евклидовом пространстве E_n существует ортонормированный базис $(e_k)_n$.

Доказательство. Возьмем в E_n какой-либо базис $(f_k)_n$ и с помощью некоторой процедуры, описанной ниже, построим n элементов e_1, \dots, e_n , линейно выражающихся через f_1, \dots, f_n и образующих ортонормированный базис.

Шаг 1. $e_1 = f_1/|f_1|$.

Шаг 2. e_2 будем искать в виде $af_2 + be_1$. Так как e_2 должен удовлетворять условиям: 1) $(e_1, e_2) = 0$ и 2) $(e_2, e_2) = 1$, то из 1) $\Rightarrow (e_1, e_2) = (e_1, af_2 + be_1) = a(e_1, f_2) + b(e_1, e_1) = a(e_1, f_2) + b = 0$, т. е. $b = -a(e_1, f_2)$. Поэтому $e_2 = ac_2$, где $c_2 = f_2 - (e_1, f_2)e_1$. Теперь

из 2) имеем $a = 1/|c_2|$, где $|c_2| = \sqrt{(c_2, c_2)}$. Итак, $e_2 = c_2/|c_2|$, где $c_2 = f_2 - (e_1, f_2)e_1$.
Шаг 3. e_3 будем искать в виде $af_3 + be_2 + ce_1$. Так как элемент e_3 должен удовлетворять условиям: 1) $(e_1, e_3) = 0$, 2) $(e_2, e_3) = 0$, 3) $(e_3, e_3) = 1$, то из первых двух условий имеем:

$$\begin{cases} a(f_3, e_1) + b(e_2, e_1) + c(e_1, e_1) = 0 \\ a(f_3, e_2) + b(e_2, e_2) + c(e_1, e_2) = 0 \end{cases} \Rightarrow \begin{cases} a(f_3, e_1) + c = 0 \\ a(f_3, e_2) + b = 0 \end{cases} \Rightarrow \begin{cases} c = -a(f_3, e_1) \\ b = -a(f_3, e_2) \end{cases}.$$

Таким образом, $e_3 = ac_3$, где $c_3 = f_3 - (f_3, e_2)e_2 - (f_3, e_1)e_1$. Из 3) следует, что $a = 1/|c_3|$. Итак, окончательно $e_3 = c_3/|c_3|$.

Шаг k . $e_k = c_k/|c_k|$, где $c_k = f_k - (f_k, e_{k-1})e_{k-1} - \dots - (f_k, e_1)e_1$.

Покажем, что

Шаг $k+1$. $e_{k+1} = c_{k+1}/|c_{k+1}|$, где $c_{k+1} = f_{k+1} - (f_{k+1}, e_k)e_k - \dots - (f_{k+1}, e_1)e_1$. Надо доказать: 1) $e_{k+1} \neq \theta$, 2) $(e_p, e_{k+1}) = 0$ для $p = \overline{1, k}$, 3) $|e_{k+1}| = 1$.

Утверждение 3) очевидно. Как следует из формулы для c_{k+1} , этот элемент линейно выражается через $f_1, f_2, \dots, f_k, f_{k+1}$, так как e_1, e_2, \dots, e_k — линейные комбинации элементов f_1, f_2, \dots, f_k . Но система элементов $f_1, f_2, \dots, f_k, f_{k+1}$ по предположению линейно независима, и, следовательно, их линейная комбинация $c_{k+1} \neq \theta$, так как коэффициент при f_{k+1} равен 1, т. е. отличен от 0. Значит, утверждение 1) имеет место. Умножим c_{k+1} по правилу скалярного произведения на e_p , $p = \overline{1, k}$, т. е. $(c_{k+1}, e_p) = (f_{k+1}, e_p) - \sum_{n=1}^k (f_{k+1}, e_n)(e_n, e_p) = (f_{k+1}, e_p) - \sum_{n=1}^k (f_{k+1}, e_p)\delta_{np} = (f_{k+1}, e_p) - (f_{k+1}, e_p) = 0$. Отсюда следует справедливость 2). Таким образом, по методу математической индукции представление для c_{k+1} верно для любого натурального k , т. е. для $k = n$. Значит, завершив этот процесс ортогонализации базиса $(f_p)_n$, мы приходим к ортонормированному базису $(e_p)_n$. Теорема доказана.

З а м е ч а н и е. В любом евклидовом пространстве E_n существует много ортонормированных базисов. Мы построили лишь один из них. Укажем еще один. На шаге 1 положим $e_1 = f_2/|f_2|$ и т. д.

С л е д с т в и е 1. В ортонормированном базисе $(e_p)_n$ скалярное произведение элементов $x = e_p x^p = e X_e$, $y = e_p y^p = e Y_e$ из пространства E_n имеет вид:

$$(x, y) = X_e^T Y_e \quad (3)$$

или в координатах:

$$(x, y) = \sum_{p=1}^n x^p y^p. \quad (3')$$

Д о к а з а т е л ь с т в о. В произвольном базисе скалярное произведение в силу (1) имеет вид $(x, y) = X_e^T A Y_e$, где $A = ||a_{pk}||_{n,n}$. В ортонормированном базисе $a_{pk} = \delta_{pk}$, т. е. $A = E$. Значит, из (1) следует (3).

С л е д с т в и е 2. Если в некотором базисе $(e_k)_n$ скалярное произведение элементов x, y имеет вид (3), то базис $(e_k)_n$ — ортонормированный.

Д о к а з а т е л ь с т в о. Так как в базисе $(e_k)_n$ $x = e_p x^p, y = e_p y^p$, то $(x, y) = a_{pk} x^p y^k$. Но по условию (x, y) представимо в виде (3), т. е. $a_{pk} = (e_p, e_k) = \delta_{pk}$. А это и есть условие ортонормированности базиса

§4. Разложение евклидова пространства на прямую сумму его подпространств.

О п р е д е л е н и е. Пусть M — какое-либо подпространство евклидова пространства E_n . Тогда совокупность P всех элементов y пространства E_n , ортого-

нальных к каждому элементу x подпространства M , называется ортогональным дополнением подпространства M .

Иными словами, P — множество всех элементов y из E_n , для которых $(x, y) = 0$ для любого x из M .

Теорема 3 Если M — подпространство евклидова пространства E_n размерности k , где $k \leq n$, то его ортогональное дополнение P является также подпространством евклидова пространства E_n размерности $n - k$, т. е. $\dim E = \dim M + \dim P$.

Доказательство. Покажем прежде всего, что P — подпространство пространства E_n . Так как P — множество элементов из E_n , то для этого осталось показать корректность на P введенных в E_n линейных операций. Пусть y_1, y_2 принадлежат P , т. е. для любого x из M $(x, y_1) = 0$, $(x, y_2) = 0$. Тогда: а) $(x, y_1 + y_2) = (a.1) = (y_1 + y_2, x) = (a.2) = (y_1, x) + (y_2, x) = 0 + 0 = 0$, т. е. $y_1 + y_2 \in P$. б) для любого числа b и любого x из M $(by_1, x) = b(y_1, x) = b \cdot 0 = 0$, т. е. $by_1 \in P$. Итак, P — подпространство евклидова пространства E_n . Теперь докажем утверждение о размерности P . Пусть $(e_p)_k$ — ортонормированный базис в M . Если y ортогонален к элементам e_1, e_2, \dots, e_k , то он ортогонален и к любой их линейной комбинации, т. е. к любому элементу из M . Поэтому условие $y \in P$ эквивалентно условию:

$$(y, e_i) = 0, \quad i = \overline{1, k}. \quad (4)$$

В силу свойства 3 (о пополнении базиса) подпространства линейного пространства, изложенного в §7 гл. 2, базис $(e_p)_k$ можно дополнить элементами f_{k+1}, \dots, f_n пространства E_n до базиса во всем пространстве E_n . Проведя процесс ортогонализации (как в теореме 2) элементов $e_1, e_2, \dots, e_k, f_{k+1}, \dots, f_n$, мы получим ортонормированный базис $(e_p)_n$ всего пространства E_n , т. е.

$$(e_p, e_i) = \delta_{ip}, \quad i, p = \overline{1, n}. \quad (5)$$

В этом базисе элемент y из P имеет разложение $y = e_p y^p$. И, следовательно, система (4) с учетом равенств (5) запишется так:

$$y^p \delta_{pi} = 0, \quad i = \overline{1, k}. \quad (6)$$

Матрица $\|\delta_{pi}\|_{n,k}$ однородной линейной системы уравнений (6) имеет ранг, равный k . Но тогда, как показано в §2 гл. 3, совокупность всех решений y системы (6), а значит и системы (4), образует линейное пространство размерности $n - k$, т. е. размерность P равна $n - k$. Теорема доказана.

С л е д с т в и е. Базисом в P являются $n - k$ последних элементов базиса $(e_p)_n$.

Действительно, элементы e_{k+1}, \dots, e_n линейно независимы, ибо принадлежат базису. Их $n - k$ и все они принадлежат P . Размерность P равна $n - k$. Значит по т. 10 гл. 2 указанная совокупность элементов образует базис в P .

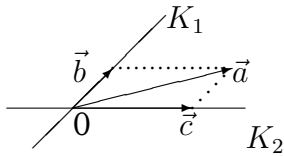
О п р е д е л е н и е. Линейное пространство E называется прямой суммой подпространств M и P , если каждый элемент x пространства E может быть представлен, и притом единственным способом, в виде суммы элемента y из M и элемента z из P .

Тот факт, что E есть прямая сумма подпространств M и P , символически записывается так: $E = M \oplus P$.

П р и м е р ы.

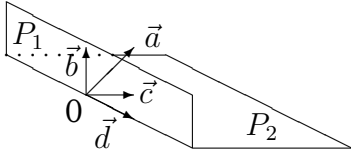
1). Пусть E_3 — линейное пространство векторов с закрепленными началами в точке O . Пусть K_1, K_2 — две различные прямые, проходящие через точку O . Тогда

плоскость P , содержащая эти две прямые, есть прямая сумма K_1 и K_2 : $P = K_1 \oplus K_2$. Действительно, для каждого вектора \vec{a} , лежащего в плоскости P , разложение на сумму векторов \vec{b} из K_1 и \vec{c} из K_2 единственно, что видно из рисунка.



2). Пусть P_1, P_2 — две различные плоскости, проходящие через точку O . Тогда любой вектор \vec{a} из E_3 представим в виде суммы вектора \vec{b} из P_1 и вектора \vec{c} из P_2 .

Но не единственным образом, так как помимо представления $\vec{a} = \vec{b} + \vec{c}$ можно предложить и такое:



$$\vec{a} = (\vec{b} + \vec{d}) + (\vec{c} - \vec{d}).$$

где \vec{d} — вектор, лежащий на пересечении плоскостей P_1 и P_2 , и тем самым вектор $\vec{b} + \vec{d}$ лежит в плоскости P_1 , а вектор $\vec{c} - \vec{d}$ в плоскости P_2 . Таким образом, пространство E_3 не является прямой суммой P_1 и P_2 , т. е. $E_3 \neq P_1 \oplus P_2$.

Теорема 4 Всякое евклидово пространство E_n можно представить в виде прямой суммы любого подпространства M и его ортогонального дополнения P , т. е. $E_n = M \oplus P$.

Доказательство. Как было показано при доказательстве теоремы 3, произвольный ортонормированный базис $(e_p)_k$ в M можно пополнить до ортонормированного базиса $(e_p)_n$ по всем пространству E_n , причем элементы e_{k+1}, \dots, e_n в нем образуют базис в P . Разложив любой элемент x из E_n по этому базису: $x = e_p x^p$, мы получим, что этот элемент однозначно представлен в силу единственности разложения по базису в виде $x = x' + x''$, где $x' = e_p x^p$, $p = \overline{1, k}$ — совершенно определенный элемент M , а $x'' = e_p x^p$, $p = \overline{k+1, n}$ — совершенно определенный элемент из P . Покажем единственность этого представления. Предположим противное, т. е. пусть есть другое представление $x = y' + y''$, где $y' \in M$, $y'' \in P$. Вычтем из первого представления второе:

$$\theta = (x' - y') + (x'' - y''). \quad (7)$$

Так как $x', y' \in M$, то $x' - y' \in M$. Так как $x'', y'' \in P$, то $x'' - y'' \in P$. Следовательно,

$$(x' - y', x'' - y'') = 0. \quad (8)$$

Умножим элемент (7) по правилу скалярного произведения на элемент $x' - y'$. Тогда $(\theta, x' - y') = (x' - y', x' - y') + (x'' - y'', x' - y')$. Здесь левая часть равна θ в силу равенства (2), второе слагаемое справа равно нулю в силу (8). Значит, $(x' - y', x' - y') = \theta$. Поэтому по а.4 $x' - y' = \theta$, т. е. $x' = y'$. Умножив элемент (7) на $x'' - y''$ по правилу скалярного произведения, аналогично только что рассмотренному случаю получаем: $x'' = y''$. Пришли к противоречию с допущением о том, что разложение по базису не единственно. Значит, это разложение единственно. Теорема доказана.

З а м е ч а н и е. Итак, доказано, что для каждого элемента x из E_n существует единственная пара элементов: y из M и z из P , где M — заданное подпространство пространства E_n , а P — ортогональное дополнение к M , такая, что $x = y + z$. При этом часто x называют наклонной к подпространству M , y — проекцией элемента

x на подпространство M , z — перпендикуляром, опущенным из элемента x на M , а длина элемента $|z|$ — расстоянием элемента x до M .

В качестве приложения и иллюстрации введенных понятий рассмотрим альтернативу Фредгольма для линейной системы n уравнений с n неизвестными:

$$AX = B, \quad (9)$$

где $A = \|a_k^p\|_n^n$, $X = \|x^p\|^n$, $B = \|b^p\|^n$. Далее нам потребуются две однородных системы, соответствующих системе (9):

$$AX = \theta \quad (10)$$

и

$$A^T X = \theta, \quad (11)$$

которую называют сопряженной однородной системой, соответствующей (9).

Теорема 5 (Альтернатива Фредгольма) *Или система (10) имеет только нулевое решение, и тогда система (9) имеет единственное решение при любом B из T_n . Или система (10) имеет ненулевое решение, и тогда система (9) совместна тогда и только тогда, когда столбец свободных членов B ортогонален пространству решений системы (11).*

Доказательство. Как было показано в курсе аналитической геометрии, линейная квадратная система уравнений имеет единственное решение тогда и только тогда, когда определитель ее матрицы не равен 0. Значит, в первом утверждении теоремы определитель матрицы системы (10), а тем самым и определитель матрицы системы (9), не равен нулю. Но тогда система (9) имеет единственное решение представимое формулами Крамера, для любого столбца B свободных членов. Докажем вторую часть теоремы. Как следует из теоремы Кронекера-Капелли (теорема 1, гл. 3) система (9) совместна тогда и только тогда, когда столбец B является линейной комбинацией столбцов A_1, A_2, \dots, A_n матрицы A этой системы. Иными словами B принадлежит линейной оболочке столбцов матрицы A $L(A_1, A_2, \dots, A_n) \equiv L(A)$ или, как следует из теоремы 4, столбец B ортогонален P — ортогональному дополнению к $L(A)$. Таким образом, любой столбец X из P ортогонален к каждому элементу из $L(A)$. Последнее имеет место, если $(X, A_k) = 0$, $k = \overline{1, n}$. Расписав эту систему равенств в координатах, получим:

$$\sum_{p=1}^n a_k^p x^p = 0, \quad k = \overline{1, n},$$

или в матричной форме: $A^T X = \theta$, т. е. X является решением системы (11), а совокупность P всех таких X — пространством решений системы (11). Итак, для совместности системы (9) необходимо и достаточно, чтобы столбец B был ортогонален к пространству P решений системы (11).

Если система (10) имеет только нулевое решение, то и система (11) имеет лишь нулевое решение, так как по свойству 1 определителя $\det A = \det A^T \neq 0$, т. е. $P = \theta$. Так как по формуле (2) $(B, \theta) = 0$, то любой столбец B ортогонален нулевому столбцу. Следовательно, в этом случае система (9) разрешима для любого B из T_n , что и доказано в первой части теоремы. Если же система (10) нетривиально совместна, то и система (11) имеет ненулевые решения. Тогда P содержит и ненулевые элементы. Поэтому здесь система (9) совместна тогда и только тогда, когда столбец B ортогонален P . Теорема доказана.

§5. Ортогональная матрица.

О п р е д е л е н и е. Квадратная матрица $Q = \|q_k^p\|_n^n$ называется ортогональной, если

$$QQ^T = E. \quad (12)$$

Сформулируем и докажем некоторые свойства ортогональной матрицы.

С в о й с т в о 1.

$$Q^T = Q^{-1}. \quad (13)$$

Д о к а з а т е л ь с т в о. По теореме 5 гл. 4 курса аналитической геометрии $\det Q^T \cdot \det Q = \det Q^T Q = \det E = 1$. Следовательно, $\det Q \neq 0$. Поэтому матрица Q имеет обратную матрицу Q^{-1} . Умножив формулу (12) на Q^{-1} слева, придем к формуле (13).

З а м е ч а н и е. Так как по определению обратной матрицы $QQ^{-1} = Q^{-1}Q = E$, то в силу равенства (13) имеем:

$$Q^T Q = E. \quad (14)$$

Каждое из равенств (13) и (14) может быть принято в качестве определения ортогональной матрицы.

С в о й с т в о 2.

$$\det Q = \pm 1. \quad (15)$$

Д о к а з а т е л ь с т в о. Как было показано при доказательстве свойства 1, $\det Q^T \cdot \det Q = 1$. Но в силу свойства 1 определителей $\det Q^T = \det Q$. Поэтому первое равенство можно записать так: $(\det Q)^2 = 1$, что и означает выполнение свойства 2.

С в о й с т в о 3. Матрица обратная к ортогональной — ортогональная матрица.

Д о к а з а т е л ь с т в о. Пусть Q — ортогональная матрица, а Q^{-1} — ей обратная. Чтобы показать, что Q^{-1} — ортогональная матрица, воспользуемся определением ортогональной матрицы, т. е. покажем выполнение равенства $(Q^{-1})^T = (Q^{-1})^{-1}$. Преобразуем левую часть равенства: $(Q^{-1})^T = (\text{свойство 1}) = (Q^T)^T = (\text{определение транспонированной матрицы}) = Q$. Теперь рассмотрим правую часть равенства: $(Q^{-1})^{-1} = Q$ (по определению обратной матрицы). Так как в результате тождественных преобразований мы пришли к одному и тому же выражению, то исходное равенство справедливо, и, следовательно, Q^{-1} — ортогональная матрица.

С в о й с т в о 4. Матрица транспонированная к ортогональной — ортогональная матрица.

Д о к а з а т е л ь с т в о. По аналогии с предыдущим: надо проверить справедливость равенства

$$(Q^T)^T = (Q^T)^{-1}.$$

Но по определению транспонированной матрицы выражение слева есть матрица Q , а справа, если заменить Q^T в силу свойства 1 на Q^{-1} и воспользоваться определением обратной матрицы, так же стоит матрица Q . Значит равенство справедливо, и Q^T — ортогональная матрица.

С в о й с т в о 5. Если Q_1 и Q_2 — ортогональные матрицы, то $Q_1 Q_2$ — ортогональная матрица.

Д о к а з а т е л ь с т в о. Покажем выполнимость равенства (13). $(Q_1 Q_2)^T =$ (по т. 4, гл. 4, курса аналитической геометрии) $= Q_2^T Q_1^T =$ (по свойству 1) $= Q_2^{-1} Q_1^{-1} =$ (по т. 9, гл. 4 курса аналитической геометрии) $= (Q_1 Q_2)^{-1}$.

Свойство 6°. $n \times n$ -матрица Q тогда и только тогда будет ортогональной, когда сумма квадратов всех элементов любого ее столбца (строки) равна 1, а сумма произведений соответствующих элементов любых двух ее столбцов (строк) равна нулю.

Необходимость. Пусть Q — ортогональная матрица, т. е. имеют место (12) и (14). Тогда из (12) следует, что

$$q_k^p (q^T)_e^k = \sum_{k=1}^n q_k^p q_k^e = \delta_{pe}. \quad (15')$$

т. е. для любых строк с индексами p, e имеет место утверждение свойства. А из (14) то же самое следует для любых двух столбцов:

$$(q^T)_k^p q_e^k = \sum_{k=1}^n q_p^k q_e^k = \delta_{pe}. \quad (15'')$$

Достаточность. Пусть выполнены (15') или (15''). Но из (15') следует, что $QQ^T = E$, т. е. согласно определению ортогональной матрицы Q — матрица ортогональная. А из (15'') следует, что $Q^T Q = E$, т. е. в силу (14) Q — ортогональная матрица.

Свойство 7°. Если $(e_k)_n$ — ортонормированный базис в E_n , то базис $(f_k)_n$, полученный преобразованием $f = eQ$, где Q — ортогональная матрица, также будет ортонормированным базисом.

Доказательство. В k -ом столбце матрицы Q стоят координаты элемента f_k в базисе $(e_p)_n$. Поэтому на основании свойства 6°, равенства (3) и ортонормированности базиса $(e_p)_n$ имеем $(f_p, f_k) = \delta_{pk}$, т. е. $(f_k)_n$ — ортонормированный базис.

Свойство 8°. Если $(f_p)_n$ и $(e_p)_n$ — ортонормированные базисы, то матрица Q , осуществляющая переход от одного базиса к другому: $f = eQ$, является ортогональной.

Доказательство. Так как $f = eQ$, то $f_k = e_p q_k^p$, $f_p = e_j q_p^j$. Тогда $(f_k, f_p) = (e_i q_k^i, e_j q_p^j) =$ (формула (3')) $= \sum_{i=1}^n q_k^i q_p^i =$ (из ортонормированности $(f_k)_n$) $= \delta_{kp}$. Но последнее равенство есть соотношение (15''). Значит в силу свойства 6° матрица Q ортогональная.

Пример. Поворот на угол α в пространстве V_2 есть ортогональное преобразование старых переменных x, y в новые x'', y'' по известным из курса аналитической геометрии формулам:

$$\begin{cases} x'' = x \cos \alpha + y \sin \alpha \\ y'' = -x \sin \alpha + y \cos \alpha \end{cases}$$

Запишем это преобразование в матричной форме: пусть

$$X'' = \begin{pmatrix} x'' \\ y'' \end{pmatrix}, X = \begin{pmatrix} x \\ y \end{pmatrix}, Q = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}.$$

Тогда $X'' = QX$ — ортогональное преобразование, а Q — ортогональная матрица. Проверим выполнение свойств. Свойство $\det Q = \cos^2 \alpha + \sin^2 \alpha = 1$, т. е. оно выполнено. Отсюда обратное преобразование задается матрицей:

$$Q^{-1} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Так как

$$Q^T = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix},$$

то $Q^T = Q^{-1}$, т. е. выполнено свойство 1°. Очевидно выполнение равенств (15), т. е. имеет место свойство 6°.

§6. Общий вид линейного функционала.

О п р е д е л е н и е. Функционалом в евклидовом пространстве E_n называется любой наперед заданный закон Φ , по которому любому элементу x из E_n сопоставляется определенное действительное число, обозначаемое символом $\Phi(x)$.

О п р е д е л е н и е. Функционал Φ , $x \in E_n$, называется линейным, если:

а) для любых x, y из E_n $\Phi(x + y) = \Phi(x) + \Phi(y)$,

б) для любого числа b из K_0 и любого x из E_n $\Phi(bx) = b\Phi(x)$.

Из условий а), б) по индукции легко получить следующее соотношение для линейного функционала: для любых x_1, x_2, \dots, x_k из E_n и любых чисел b^1, b^2, \dots, b^k из K_0 имеем:

$$\Phi(x_p b^p) = b^p \Phi(x_p). \quad (16)$$

Пример линейного функционала: Пусть x^0 — фиксированный элемент E_n , x — любой элемент E_n . Тогда скалярное произведение (x, x^0) — линейный функционал. Действительно, скалярное произведение при фиксированном x^0 — определенный закон по которому любому элементу x из E_n ставится в соответствие определенное число. Значит, это функционал. Далее из а.1° — 3° определения скалярного произведения следует, что это линейный функционал. Оказывается, приведенный пример линейного функционала играет фундаментальную роль во всей теории линейных функционалов. А именно, верна следующая

Теорема 6 *Всякий линейный функционал Φ представим единственным образом по формуле:*

$$\Phi(x) = (x, x^0), \quad (17)$$

где x^0 — некоторый фиксированный элемент из E_n .

Д о к а з а т е л ь с т в о. Пусть $(e_p)_n$ — ортонормированный базис в E_n . Так как функционал определен на всем пространстве E_n , то, в частности, определены числа $\Phi(e_1), \Phi(e_2), \dots, \Phi(e_n)$. Пусть x^0 — элемент E_n с координатами $X_e^0 = \|\Phi(e_k)\|^n$ в базисе $(e_p)_n$. Пусть x — любой элемент E_n с координатами $X_e = \|c^k\|^n$ в том же базисе $(e_p)_n$, т. е. $x = e_k c^k$. Тогда $\Phi(x) = \Phi(e_k c^k) = (\text{формула (16)}) = c^k \Phi(e_k) = (\text{так как } (e_p)_n \text{ — ортонормированный базис, то по формуле (3)}) = (x, x^0)$. Осталось доказать единственность представления функционала Φ в виде скалярного произведения. Пусть возможны два представления: $\Phi(x) = (x, x^0) = (x, y)$. Но тогда для любого x из E_n $(x, x^0) - (x, y) = (a.1^\circ, 2^\circ) = (x, x^0 - y) = 0$. В частности, при $x = x^0 - y$ имеем $(x^0 - y, x^0 - y)$, что в силу а.4° определения скалярного произведения возможно лишь при $x^0 - y = \theta$, т. е. при $x^0 = y$. Из полученного противоречия с предположением следует единственность представления. Теорема доказана.

§7. Изоморфизм евклидовых пространств.

По аналогии с изоморфизмом линейных пространств вводится понятие изоморфизма всех конечномерных евклидовых пространств.

О п р е д е л е н и е. Два евклидовых пространства E_n и E_n'' называются изоморфными, если между элементами этих пространств можно установить взаимно однозначное соответствие так, что если элементам пространства E_n x, y отвечают соответственно элементы x'', y'' пространства E_n'' , то элементу $x + y$ из E_n отвечает $x'' + y''$ из E_n'' , а элементу bx ($\forall b \in K_0$) из E_n отвечает элемент bx'' из E_n'' и, наконец, $(x, y) = (x'', y'')$.

Теорема 7 Все евклидовы пространства одной и той же размерности n изоморфны.

Д о к а з а т е л ь с т в о. Достаточно доказать, что любое n -мерное евклидово пространство E_n'' изоморфно евклидову пространству E_n упорядоченных совокупностей n действительных чисел вида $x = (x^1, x^2, \dots, x^n)$ со скалярным произведением

$$(x, y) = x^1y^1 + x^2y^2 + \dots + x^ny^n. \quad (18)$$

Согласно теореме 2 в E_n'' существует ортонормированный базис (e_k'') . Каждому элементу $x'' = e_k''x^k$ пространства E_n'' поставим в соответствие n чисел x^1, x^2, \dots, x^n — координат этого элемента в базисе (e_k) , т. е. вполне определенный элемент $x = (x^1, x^2, \dots, x^n)$ пространства E_n . Установленное соответствие взаимно однозначно. Кроме того, из теоремы 9 гл. 2 следует, что если элементам x'', y'' из E_n'' , имеющим в базисе (e_k'') координаты $X_e = \|x^k\|^n, Y_e = \|y^k\|^n$, отвечают соответственно элементы $x = (x^1, x^2, \dots, x^n)$ и $y = (y^1, y^2, \dots, y^n)$ пространства E_n , то элементу $x'' + y''$ соответствует элемент $x + y$, а элементу bx'' соответствует элемент bx . Осталось доказать, что $(x'', y'') = (x, y)$. В силу ортонормированности базиса (e_k'') и формулы (3'): $(x'', y'') = x^1y^1 + x^2y^2 + \dots + x^ny^n$. С другой стороны, в силу формулы (18), определяющей скалярное произведение в пространстве E_n , этому же выражению равно скалярное произведение (x, y) . Теорема доказана.

Гл. 5. Линейные операторы в линейном конечномерном пространстве.

§1. Основные понятия.

Пусть R_n и R_k — линейные пространства над полем K , причем $k \leq n$. Будем называть оператором \hat{A} , действующим из R_n в R_k , соответствие вида $\hat{A} : R_n \rightarrow R_k$, сопоставляющее каждому элементу x пространства R_n некоторый элемент y пространства R_k и обозначать символом $y = \hat{A}x$, при этом элемент x называется прообразом, а соответствующий ему элемент y называется образом.

О п р е д е л е н и е. Оператор \hat{A} , действующий из R_n в R_k , называется линейным, если для любых элементов x_1 и x_2 пространства R_n и любого числа b из поля K

1) $\hat{A}(x_1 + x_2) = \hat{A}(x_1) + \hat{A}(x_2)$ (свойство аддитивности оператора)

2) $\hat{A}(bx_1) = b\hat{A}(x_1)$ (свойство однородности оператора).

З а м е ч а н и е 1. Если $k = 1$, то линейный оператор \hat{A} , действующий из R_n в R_1 называется линейной формой или линейным функционалом. Этому случаю посвящен §6 гл. 4.

З а м е ч а н и е 2. Если пространство R_k совпадает с пространством R_n , то линейный оператор \hat{A} , действующий из R_n в R_n , называют также линейным преобразованием пространства R_n .

Эта глава целиком будет посвящена последнему случаю, т. е. линейному оператору, действующему из R_n в R_n . Так как в этом случае не возникает проблемы о том, куда и откуда действует оператор, то слова "действующий из R_n в R_n " за ненадобностью будем опускать.

П р и м е р ы.

1. Нуль-оператор $\hat{\theta}$: для каждого x из R $\hat{\theta}x = \theta$.
2. Единичный или тождественный оператор \hat{E} : для всех x из R $\hat{E}x = x$.
3. Оператор умножения на число из поля K или оператор подобия: для каждого x из R $\hat{A}x = bx$, где b — фиксированное число из K .
4. Оператор поворота на угол ϕ_0 в линейном пространстве V_2 : для каждого вектора $\vec{a} = \{\rho, \phi\}$ $\hat{A}\vec{a} = \vec{b} = \{\rho, \phi + \phi_0\}$.
5. Оператор дифференцирования \hat{D} , заданный в линейном пространстве непрерывно дифференцируемых на сегменте $[a, b]$ функций $C_{[a,b]}^1$, переводит элементы этого пространства в элементы пространства $C_{[a,b]}$: для каждого $x(t) \in C_{[a,b]}^1$ $\hat{D}x(t) = \dot{x}(t) \in C_{[a,b]}$. Иными словами, оператор \hat{D} действует на подпространстве $C_{[a,b]}^1$ линейного пространства $C_{[a,b]}$. Это единственный пример линейного оператора в бесконечномерном пространстве, который мы рассмотрели в нашем курсе, ибо наша цель — изучить линейные операторы, действующие в линейном конечномерном пространстве.

§2. Матрица линейного оператора.

Пусть линейный оператор \hat{A} задан в линейном пространстве R_n , а $(e_p)_n$ — некоторый базис в этом пространстве. Оказывается, достаточно выяснить результат применения оператора \hat{A} к базисным векторам, чтобы задать этот оператор во всем

пространстве R_n . Рассмотрим элементы $\hat{A}e_p$, $p = \overline{1, n}$. Разложим каждый из них по базису $(e_p)_n$:

$$\hat{A}e_p = e_k a_p^k, \quad p = \overline{1, n}, \quad k = \overline{1, n}. \quad (1)$$

Матрица $A_e = \|a_p^k\|_n^n$ называется матрицей линейного оператора \hat{A} в базисе $(e_p)_n$. Соотношение (1) запишем в матричной форме, воспользовавшись введенными в гл. 2 обозначениями $e = \|e_p\|_n$, $\hat{A}e = \|\hat{A}e_p\|_n$:

$$\hat{A}e = eA_e. \quad (1')$$

Лемма 1 Если $C = \|c_p^k\|_m^n$, то $\hat{A}(eC) = (\hat{A}e)C$.

Доказательство. $\hat{A}(eC) = (\text{определение произведения матриц}) = \hat{A} \cdot \|e_k c_p^k\|_m = \|\hat{A}(e_k c_p^k)\|_m = (\text{свойство 2) определения линейного оператора}) = \|(\hat{A}e_k) c_p^k\|_m = (\text{определение произведения матриц}) = (\hat{A}e)C$. Лемма доказана.

Теорема 1 Пусть в базисе $(e_p)_n$ линейного пространства R_n линейному оператору \hat{A} соответствует матрица A_e . Тогда, если элемент x из R_n имеет в этом базисе координаты $X_e = \|x^p\|_n^n$, то его образ $y = \hat{A}x$ будет иметь координаты, определяемые формулой:

$$Y_e = \|y^p\|_n^n = A_e X_e. \quad (2)$$

Доказательство. Если $x, y \in R$, то по базису $(e_p)_n$ они разложатся так: $x = eX_e, y = eY_e$. Тогда $y = \hat{A}x$ запишется так: $eY_e = \hat{A}(eX_e)$ или в силу леммы 1 и формулы (1') $eY_e = eA_e X_e$. Здесь слева и справа стоят разложения по базису $(e_p)_n$ равных элементов. Поэтому в силу единственности разложения по базису имеем: $Y_e = A_e X_e$.

Теорема 2 Если $(e_p)_n$ — базис в линейном пространстве R_n , то каждая $n \times n$ -матрица $A_e = \|a_p^k\|_n^n$ является матрицей некоторого линейного оператора \hat{A} в этом базисе.

Доказательство. Построим оператор \hat{A} так: каждому x с координатами X_e ставим в соответствие элемент y с координатами Y_e , заданными равенством (2).

а) Этот оператор линеен, так как элементы $x + z$ и bx имеют в силу теоремы 9, гл. 2, соответственно, координаты $X_e + Z_e, bX_e$ и по формуле (2):

$$A_e(X_e + Z_e) = A_e X_e + A_e Z_e, \quad \text{т. е. } \hat{A}(x + z) = \hat{A}x + \hat{A}z,$$

$$A_e(bX_e) = b(A_e X_e), \quad \text{т. е. } \hat{A}(bx) = b(\hat{A}x).$$

б) Вычислим матрицу этого оператора в базисе $(e_p)_n$. С этой целью вычислим координаты элемента $\hat{A}e_p$ и поставим их, как следует из определения матрицы оператора, в p -ый столбец матрицы. Тогда по формуле (2): $(\hat{A}e_p)_e = A_e \|\delta_p^i\|_n^n = \|a_i^k \delta_p^i\|_n^n = \|a_p^k\|_n^n$, где $\|\delta_p^i\|_n^n$ — координаты e_p в базисе $(e_i)_n$. Таким образом, матрица оператора в указанном базисе $\|a_p^k\|_n^n = A_e$. Теорема доказана.

Объединяя теоремы 1 и 2, можно сделать вывод: при фиксированном базисе в линейном пространстве R_n существует взаимно однозначное соответствие между линейными операторами и матрицами порядка $n \times n$.

Примеры.

1. Матрица нуль-оператора $\hat{\theta}$ в любом базисе имеет вид: $\theta_e = \|0\|_n^n$.

2. Матрица тождественного оператора \hat{E} в любом базисе имеет вид: $E_e = \|\delta_k^p\|_n^n$.

3. Матрица оператора подобия \hat{A} в любом базисе имеет вид: $A_e = \|b\delta_k^p\|_n^n$.

4. Матрица поворота на угол α в линейном пространстве V_3 для ортонормированного базиса \vec{e}_1, \vec{e}_2 : $\hat{A}\vec{e}_1 = \vec{e}_1 \cos \alpha + \vec{e}_2 \sin \alpha$, $\hat{A}\vec{e}_2 = -\vec{e}_1 \sin \alpha + \vec{e}_2 \cos \alpha$. Таким образом,

$$A_e = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} - \text{ортогональная матрица.}$$

§3. Связь матриц оператора при переходе от одного базиса к другому.

Пусть в базисах $(e_p)_n$ и $(f_p)_n$ матрицами оператора \hat{A} будут матрицы A_e и A_f соответственно. Пусть P — матрица перехода от базиса $(e_p)_n$ к базису $(f_p)_n$, т. е. $f = eP$. Имеем: $\hat{A}f = \hat{A}(eP) = (\text{лемма 1}) = (\hat{A}e)P = (\text{формула (1')}) = e(A_eP)$. С другой стороны, $\hat{A}f = fA_f = (eP)A_f = e(PA_f)$. Таким образом, $e(A_eP) = e(PA_f)$, где слева и справа стоят разложения по базису $(e_p)_n$ равных элементов. Поэтому в силу единственности разложения по базису $A_eP = PA_f$. Умножив это равенство на P^{-1} слева, получим

$$A_f = P^{-1}A_eP. \quad (3)$$

Более того

$$\det A_f = \det A_e, \quad (4)$$

что следует из формулы (3), теоремы 5 гл. 4 курса аналитической геометрии и того, что $\det P \cdot \det P^{-1} = 1$.

§4. Действия над линейными операторами и соответствующие действия над матрицами.

Пусть в линейном пространстве R_n заданы базис $(e_p)_n$ и линейные операторы \hat{A} и \hat{B} .

О п р е д е л е н и е. Операторы \hat{A} и \hat{B} называются равными, если для каждого x из R_n $\hat{A}x = \hat{B}x$.

Утверждение 1 Если операторы равны, то в любом базисе равны матрицы этих операторов.

Д о к а з а т е л ь с т в о. Так как для любого x из R_n $\hat{A}x = \hat{B}x$, то при $x = e_p$ $\hat{A}e_p = \hat{B}e_p$, т. е. при $p = \overline{1, n}$ последние равенства можно записать так: $\hat{A}e = \hat{B}e$ или по формуле (1') $eA_e = eB_e$. Поэтому $A_e = B_e$.

О п р е д е л е н и е. Суммой любых двух линейных операторов \hat{A} , \hat{B} называется оператор \hat{C} такой, что для любого x из R_n $\hat{C}x = \hat{A}x + \hat{B}x$ и обозначается символом $\hat{A} + \hat{B}$.

Утверждение 2 Если \hat{A} , \hat{B} — линейные операторы в линейном пространстве R_n , то $\hat{A} + \hat{B}$ — также линейный оператор в том же линейном пространстве.

Д о к а з а т е л ь с т в о. Пусть $\hat{C} = \hat{A} + \hat{B}$. Надо доказать, что 1) для всех x, y из R_n $\hat{C}(x+y) = \hat{C}x + \hat{C}y$, 2) для любого x из R_n и любого b из поля K $\hat{C}(bx) = b\hat{C}x$. Из определения суммы операторов следует, что для любого e из R_n $\hat{C}e = \hat{A}e + \hat{B}e$. Поэтому, если $e = x+y$, то $\hat{C}(x+y) = \hat{A}(x+y) + \hat{B}(x+y) = (\text{из линейности } \hat{A}, \hat{B}) = \hat{A}x + \hat{A}y + \hat{B}x + \hat{B}y = (\text{а.1, 2 линейного пространства}) = (\hat{A}x + \hat{B}x) + (\hat{A}y + \hat{B}y) = (\text{определение суммы операторов}) = (\hat{A} + \hat{B})x + (\hat{A} + \hat{B})y = \hat{C}x + \hat{C}y$, т. е. 1) доказано. Аналогично, $\hat{C}(bx) = (\text{определение суммы операторов}) = \hat{A}(bx) + \hat{B}(bx) = (\text{линейность } \hat{A}, \hat{B})$

$= b\hat{A}x + b\hat{B}x = (\text{а.8 линейного пространства}) = b(\hat{A}x + \hat{B}x) = (\text{определение суммы операторов}) = b(\hat{A} + \hat{B})x = b\hat{C}x.$

Утверждение 3 Матрица суммы операторов \hat{A}, \hat{B} в любом базисе $(e_p)_n$ равна сумме матриц операторов \hat{A}, \hat{B} в этом же базисе, т. е.

$$C_e = (A + B)_e = A_e + B_e.$$

Доказательство. $eC_e = (\text{формула (1')}) = \hat{C}e = \|\hat{C}e_p\|_n = \|\hat{A}e_p + \hat{B}e_p\|_n = \|\hat{A}e_p\|_n + \|\hat{B}e_p\|_n = \hat{A}e + \hat{B}e = (\text{формула (1')}) = eA_e + eB_e = e(A_e + B_e).$

Определение. Произведением оператора \hat{A} на число b из K называется оператор \hat{C} такой, что для каждого x из R_n $\hat{C}x = b(\hat{A}x)$, и обозначается символом $b\hat{A}$.

Утверждение 4 Если \hat{A} — линейный оператор, действующий в линейном пространстве R_n над полем K , а число b принадлежит K , то оператор $b\hat{A}$ — линейный.

Доказательство. Пусть $\hat{C} = b\hat{A}$. Тогда для любых x, y из R_n $\hat{C}(x + y) = b(\hat{A}(x + y)) = b(\hat{A}x + \hat{A}y) = b(\hat{A}x) + b(\hat{A}y) = \hat{C}x + \hat{C}y$, т. е. первое требование линейности оператора выполнено. Теперь рассмотрим $\hat{C}(ax)$, где $a \in K$. $\hat{C}(ax) = b(\hat{A}(ax)) = (\text{линейность } \hat{A}) = b(a\hat{A}x) = (\text{а.6 линейного пространства}) = (ba)\hat{A}x = a(b\hat{A}x) = a\hat{C}x$, т. е. и второе требование линейности оператора имеет место. Утверждение доказано.

Утверждение 5 Матрица оператора $b\hat{A}$ в любом базисе $(e_p)_n$ равна матрице оператора \hat{A} в этом же базисе, умноженной на число b , т. е.

$$(b\hat{A})_e = bA_e.$$

Доказательство. Пусть $\hat{C} = b\hat{A}$. Тогда $eC_e = (\text{формула (1')}) = \hat{C}e = \|\hat{C}e_p\|_n = \|b\hat{A}e_p\|_n = b\|\hat{A}e_p\|_n = b(\hat{A}e) = (\text{формула (1')}) = beA_e = ebA_e$. Поэтому $C_e = bA_e$.

Теорема 3 Множество всех линейных операторов, действующих в линейном пространстве R_n над полем K , с указанными операциями сложения и умножения на число b из того же поля K образуют линейное пространство.

Доказательство. Надо проверить выполнение всех восьми аксиом линейного пространства.

а.1) $\hat{A} + \hat{B} = \hat{B} + \hat{A}$. Из определения операции сложения операторов следует, что для всякого x из R_n $(\hat{A} + \hat{B})x = \hat{A}x + \hat{B}x = y_1 + y_2$ и $(\hat{B} + \hat{A})x = \hat{B}x + \hat{A}x = y_2 + y_1$. Но для любых y_1, y_2 из R_n $y_1 + y_2 = y_2 + y_1$. Поэтому а.1) выполнена.

а.2) $(\hat{A} + \hat{B}) + \hat{C} = \hat{A} + (\hat{B} + \hat{C})$ доказывается так же, как а.1).

а.3) для всех операторов \hat{A} существует оператор $\hat{\theta}$ такой, что $\hat{A} + \hat{\theta} = \hat{A}$ (доказывается по той же схеме, что и а.1)).

а.4) для всякого оператора \hat{A} существует оператор $(-1)\hat{A}$, такой что $\hat{A} + (-1)\hat{A} = \hat{\theta}$. (для всякого x из R_n $(\hat{A} + (-1)\hat{A})x = \hat{A}x + (-1)\hat{A}x = y + (-1)y = y - y = \theta$ т. е. аксиома выполнена).

а.5) $1 \cdot \hat{A} = \hat{A}$, что следует из определения операции умножения оператора на число.

а.6) для всех a, b из K $a(b\hat{A}) = (ab)\hat{A}$ (докажите самостоятельно).

а.7) для всех a, b из K и любого \hat{A} $(a + b)\hat{A} = a\hat{A} + b\hat{A}$. Для каждого x из R_n $((a + b)\hat{A})x = (a + b)(\hat{A}x) = (a + b)y = ay + by$. С другой стороны, $(a\hat{A} + b\hat{A})x = a\hat{A}x + b\hat{A}x = ay + by$. Сравнивая эти равенства, приходим к выводу, что а.7) имеет место.

а.8) для любого a из K и любых \hat{A}, \hat{B} $a(\hat{A} + \hat{B}) = a\hat{A} + a\hat{B}$. С одной стороны, для всех x из R_n $(a(\hat{A} + \hat{B}))x = a((\hat{A} + \hat{B})x) = a(\hat{A}x + \hat{B}x) = a(y_1 + y_2) = ay_1 + ay_2$. С другой стороны, $(a\hat{A} + a\hat{B})x = a\hat{A}x + a\hat{B}x = ay_1 + ay_2$. Из этих равенств следует справедливость а.8). Теорема доказана.

З а д а ч а. Найти базис и определить размерность пространства U линейных операторов, действующих в пространстве R_n над полем K .

Р е ш е н и е. Между оператором \hat{A} и его матрицей A_e в базисе $(e_p)_n$ существует взаимно однозначное соответствие, что было доказано выше (теоремы 1, 2). Поэтому, так как при этом в силу выше доказанных утверждений сумме операторов соответствует сумма их матриц, а произведению матриц на число оператор, умноженный на это число, то пространства U и A_n^n изоморфны. Следовательно, размерность пространства U равна n^2 и базис пространства U $(\hat{A}_p)_{n^2}$ — множество линейных операторов, у каждого из которых в базисе $(e_p)_n$ матрица имеет все элементы равными нулю, кроме единственного элемента, равного 1.

О п р е д е л е н и е. Произведением операторов \hat{A}, \hat{B} действующих в линейном пространстве R_n , называется оператор \hat{C} такой, что для всех x из R_n $\hat{C}x = \hat{A}(\hat{B}x)$ и обозначается так: $\hat{A}\hat{B}$.

Утверждение 6 Если \hat{A}, \hat{B} — линейные операторы, то $\hat{A}\hat{B}$ — тоже линейный оператор.

Д о к а з а т е л ь с т в о. Пусть $\hat{A}\hat{B} = \hat{C}$. Тогда для всех x, y из R_n $\hat{C}(x + y) = \hat{A}(\hat{B}(x + y)) = (\text{линейность } \hat{B}) = \hat{A}(\hat{B}x + \hat{B}y) = (\text{линейность } \hat{A}) = \hat{A}(\hat{B}x) + \hat{A}(\hat{B}y) = \hat{C}x + \hat{C}y$, т. е. первое условие линейности оператора выполнено. Так как для любых x из R_n и любого a из K $\hat{C}(ax) = \hat{A}(\hat{B}(ax)) = (\text{линейность } \hat{B}) = \hat{A}(a\hat{B}x) = (\text{линейность } \hat{A}) = a\hat{A}(\hat{B}x) = a\hat{C}x$, то выполнено и второе условие линейности оператора.

Утверждение 7 Матрица оператора $\hat{A}\hat{B}$ в любом базисе $(e_p)_n$ равна произведению матрицы оператора \hat{A} на матрицу оператора \hat{B} в том же базисе, т. е. $(AB)_e = A_e B_e$.

Д о к а з а т е л ь с т в о. Пусть $\hat{C} = \hat{A}\hat{B}$. Тогда $eC_e = (\text{формула (1')}) = \hat{C}e = (\hat{A}(\hat{B}e)) = (\text{формула (1')}) = \hat{A}(eB_e) = (\text{лемма 1}) = (\hat{A}e)B_e = (\text{формула (1')}) = (eA_e)B_e = (\text{свойство операции умножения матриц}) = e(A_e B_e)$, т. е. $eC_e = e(A_e B_e) \Rightarrow$ утверждение.

З а м е ч а н и е. $(AB)_e = A_e B_e$, $(BA)_e = B_e A_e$. Как было показано в алгебре матриц (гл. 4 курса аналитической геометрии), операция умножения матриц не обладает свойством перестановочности, т. е. вообще говоря, $A_e B_e \neq B_e A_e$. Поэтому в силу изоморфизма пространства операторов и пространства матриц A_n^n , вообще говоря $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. Например, в линейном пространстве V_2 заданы ортонормированный базис \vec{e}_1, \vec{e}_2 и два оператора: оператор \hat{A} проектирует векторы из V_2 на ось вектора \vec{e}_1 , а оператор \hat{B} поворачивает векторы из V_2 на угол $\pi/2$ против часовой стрелки. Тогда $\hat{A}\hat{B}\vec{e}_1 = \hat{A}\vec{e}_2 = \theta$, а $\hat{B}\hat{A}\vec{e}_1 = \hat{B}\vec{e}_1 = \vec{e}_2$. Но $\theta \neq \vec{e}_2$, т. е. $\hat{A}\hat{B} \neq \hat{B}\hat{A}$.

Однако есть линейные операторы, обладающие свойством перестановочности:

1) для всех \hat{A} $\hat{A}\hat{\theta} = \hat{\theta}\hat{A} = \hat{\theta}$ и 2) для всех \hat{A} $\hat{A}\hat{E} = \hat{E}\hat{A} = \hat{A}$.

Свойства операции умножения операторов.

1. Сочетательное свойство относительно числового множителя: для любого a из K и любых \hat{A}, \hat{B} из U $a(\hat{A}\hat{B}) = (a\hat{A})\hat{B}$.

2. Сочетательное свойство относительно операторного множителя: для любых $\hat{A}, \hat{B}, \hat{C}$ из U $(\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C})$.

3. Распределительное свойство относительно сложения операторов: для любых $\hat{A}, \hat{B}, \hat{C}$ из U $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$ и $\hat{C}(\hat{A} + \hat{B}) = \hat{C}\hat{A} + \hat{C}\hat{B}$.

Доказательство 1. $\hat{C} = \hat{A}\hat{B}$ — линейный оператор, как следует из утверждения 6. Поэтому для любого x из R_n $(a(\hat{A}\hat{B})x = (a\hat{C})x = (\text{линейность } \hat{C}) = \hat{C}(ax) = \hat{A}(\hat{B}(ax)))$. С другой стороны, $((a\hat{A})\hat{B})x = (\text{линейность } \hat{A}) = \hat{A}(a\hat{B}x) = (\text{линейность } \hat{B}) = \hat{A}(\hat{B}(ax))$. Из сравнения этих равенств следует свойство 1.

2. Из определения операции умножения операторов: для всех x из R_n имеем $(\hat{A}(\hat{B}\hat{C}))x = \hat{A}((\hat{B}\hat{C})x) = \hat{A}(\hat{B}(\hat{C}x))$. С другой стороны, $((\hat{A}\hat{B})\hat{C})x = (\hat{A}\hat{B})(\hat{C}x) = \hat{A}(\hat{B}(\hat{C}x))$. Отсюда следует справедливость свойства 2.

3. Докажем первое равенство, ибо второе доказывается аналогично. Для всех x из R_n $((\hat{A} + \hat{B})\hat{C})x = (\hat{A} + \hat{B})(\hat{C}x) = \hat{A}(\hat{C}x) + \hat{B}(\hat{C}x) = \hat{A}\hat{C}x + \hat{B}\hat{C}x$, что и означает выполнение свойства 3.

З а м е ч а н и е. Если перемножается несколько одинаковых операторов, то для сокращения записи введем обозначение $\hat{A} \cdot \hat{A} \cdot \dots \cdot \hat{A} = \hat{A}^n$ при условии, что таких сомножителей n . Очевидна и такая запись: $\hat{A}^{n+k} = \hat{A}^n \hat{A}^k$.

О п р е д е л е н и е. *Линейный оператор \hat{B} называется обратным к оператору \hat{A} , если $\hat{B}\hat{A} = \hat{A}\hat{B} = \hat{E}$ и обозначается символом $\hat{B} = \hat{A}^{-1}$.*

Утверждение 8 *Если оператор \hat{A} в базисе $(e_p)_n$ имеет матрицу A_e , то его обратный оператор \hat{A}^{-1} имеет в том же базисе матрицу A_e^{-1} .*

Доказательство. По определению обратного оператора и утверждению 7 $B_e A_e = A_e B_e = E$, т. е. из определения обратной матрицы имеем $B_e = A_e^{-1}$. Что и требовалось доказать.

Утверждение 9 *Линейный оператор \hat{A} имеет обратный \hat{A}^{-1} , если его матрица A_e в некотором базисе $(e_p)_n$ является невырожденной.*

Доказательство. Если в некотором базисе $(e_p)_n$ матрица A_e оператора \hat{A} невырождена, то 1) как следует из равенства (4), матрица этого оператора будет невырожденной в любом базисе пространства R_n , 2) она имеет обратную матрицу A_e^{-1} . Но тогда по теореме 2 матрица A_e^{-1} будет соответствовать линейный оператор, который и является обратным.

§5. Собственные значения и собственные векторы линейных операторов.

Пусть M — подпространство линейного пространства R_n над числовым полем K , а \hat{A} — линейный оператор, действующий в этом линейном пространстве.

О п р е д е л е н и е. *Пространство M называется инвариантным подпространством относительно оператора \hat{A} , если для любого x из M $\hat{A}x$ также принадлежит M .*

П р и м е р ы.

1) Для линейных операторов $\hat{\theta}$, \hat{E} и подобия всякое подпространство линейного пространства является инвариантным.

2) В пространстве $C_{[a,b]}^1$ линейная оболочка $L\{a \sin t + b \cos t\}$ является инвариантным подпространством относительно оператора дифференцирования $D = \partial/\partial t$.

О п р е д е л е н и е. *Число λ из поля K называется собственным значением оператора \hat{A} , если существует ненулевой элемент x из R_n такой, что*

$$\hat{A}x = \lambda x. \quad (5)$$

При этом элемент x называется собственным вектором оператора \hat{A} .

З а м е ч а н и е. Иногда равенство (5) лучше записывать так:

$$(\hat{A} - \lambda \hat{E})x = \theta. \quad (5')$$

Утверждение 10 Если x, y - собственные векторы оператора \hat{A} , соответствующие собственному значению λ , и a, b — какие угодно числа из поля K , то ненулевой вектор $ax + by$ ($a^2 + b^2 \neq 0$) тоже собственный вектор оператора \hat{A} с тем же собственным значением λ .

Д о к а з а т е л ь с т в о. $\hat{A}(ax + by) =$ (линейность \hat{A}) $= a(\hat{A}x) + b(\hat{A}y) =$ (формула (5)) $= a(\lambda x) + b(\lambda y) =$ (а.8 линейного пространства) $= \lambda(ax + by)$. Итак, $\hat{A}(ax + by) = \lambda(ax + by)$, т. е. согласно (5) $ax + by$ — собственный вектор оператора \hat{A} с собственным значением λ .

С л е д с т в и е. Каждому собственному вектору оператора \hat{A} соответствует одномерное инвариантное подпространство относительно оператора \hat{A} .

Действительно, если x — собственный вектор оператора \hat{A} , то в силу утверждения 10 $L = \{ax\}$ — множество собственных векторов и нулевой элемент θ . Поэтому при $a \neq 0$ $\hat{A}(ax) = \lambda(ax) = (\lambda a)x \in L$, а при $a = 0$ $\hat{A}\theta = \theta$, т. е. L — инвариантное относительно оператора \hat{A} подпространство.

Лемма 2 Для любого линейного оператора \hat{A} , действующего в линейном пространстве R_n , $\hat{A}\theta = \theta$.

Д о к а з а т е л ь с т в о. $\hat{A}\theta =$ (т. 3, гл. 3) $= \hat{A}(0 \cdot x) =$ (линейность \hat{A}) $= 0 \cdot \hat{A}x = 0 \cdot y = \theta$.

Теорема 4 Множество M_λ , содержащее нулевой элемент θ и все собственные векторы, соответствующие собственному значению λ оператора \hat{A} образует инвариантное относительно оператора \hat{A} подпространство.

Д о к а з а т е л ь с т в о. Если элемент y принадлежит M_λ , то

$$\hat{A}y = \lambda y, \quad (6)$$

ибо, если $y \neq \theta$, то это следует из (5), а если $y = \theta$, то из леммы 2. Далее,

а) M_λ — подпространство линейного пространства R_n , так как 1) для каждого x, y из M_λ $\hat{A}(x+y) =$ (линейность оператора \hat{A}) $= \hat{A}x + \hat{A}y =$ (формула (6)) $= \lambda x + \lambda y =$ (а.8) линейного пространства) $= \lambda(x+y)$, т. е. $x+y \in M_\lambda$. 2) для любых x из M_λ и a из K $\hat{A}(ax) =$ (линейность оператора) $= a\hat{A}x =$ (формула (6)) $= a \cdot \lambda x =$ (а.6) линейного пространства) $= \lambda(ax)$, т. е. $ax \in M_\lambda$.

б) M_λ — инвариантное относительно оператора \hat{A} подпространство, т. е. если y из M_λ такой элемент, что $\hat{A}y = \lambda y$, то $\hat{A}y \in M_\lambda$. Действительно, $\hat{A}(\hat{A}y) =$ (по формуле (6)) $= \hat{A}(\lambda y) =$ (линейность оператора) $= \lambda(\hat{A}y)$, т. е. $\hat{A}y \in M_\lambda$.

П р и м е р ы.

1) Для оператора $\hat{\theta}$ каждый элемент $x \neq \theta$ является собственным вектором с собственным значением $\lambda = 0$ т. е. $\hat{\theta}x = 0 \cdot x$.

2) Для оператора \hat{E} каждый элемент $x \neq \theta$ является собственным вектором с собственным значением $\lambda = 1$, т. е. $\hat{E}x = 1 \cdot x$.

3) Для оператора подобия с коэффициентом подобия каждый элемент $x \neq \theta$ является собственным вектором с собственным значением $\lambda = b$, т. е. $\hat{A}x = bx$.

4) Оператор поворота в V_2 на угол α ($0 < \alpha < \pi$) не имеет собственных векторов, ибо из равенства (5) следует коллинеарность векторов, а при повороте на угол α получающийся вектор не будет коллинеарен исходному вектору.

5) В линейном пространстве $C_{[a,b]}^2$ оператор $\hat{D}^2 = \partial^2/\partial t^2$ имеет собственный вектор $x = \sin t$ с собственным значением $\lambda = -1$, т. е. $D^2 \sin t = -\sin t$.

6) В пространстве V_3 оператор \hat{A} поворота вокруг некоторой оси, проходящей через нулевой элемент, на угол α , имеет собственные векторы — все векторы e этой оси с собственным значением $\lambda = 1$.

Теорема 5 *Собственные векторы оператора \hat{A} , соответствующие различным собственным значениям, линейно независимы.*

Доказательство. Воспользуемся методом математической индукции. Пусть $\lambda_1, \lambda_2, \dots, \lambda_p$ — различные собственные значения оператора \hat{A} , а e_1, e_2, \dots, e_p — соответствующие им собственные векторы. Так как $e_1 \neq \theta$, то этот вектор является линейно независимым. Следовательно, для одного вектора утверждение теоремы верно. Пусть теорема верна для k собственных векторов e_1, e_2, \dots, e_k . Присоединим у этим векторам вектор e_{k+1} и предположим, что

$$e_i a^i = \theta, \quad i = \overline{1, k+1}. \quad (7)$$

Применим к вектору (7) оператор \hat{A} , воспользуемся его линейностью и тем, что $\hat{A}e_i = \lambda_i e_i$. В итоге получаем, что

$$\lambda_i e_i a^i = \theta. \quad (8)$$

Теперь умножим вектор (7) на λ_{k+1} и вычтем из равенства (8):

$$(\lambda_i - \lambda_{k+1})e_i a^i = \theta, \quad i = \overline{1, k}.$$

Так как по условию $\lambda_i - \lambda_{k+1} \neq 0$ и e_1, e_2, \dots, e_k — линейно независимы по предположению, то последнее равенство возможно тогда и только тогда, когда $a^i = 0$, $i = \overline{1, k}$. Но тогда из равенства (7) в силу того, что $e_{k+1} \neq \theta$, следует, что $a^{k+1} = 0$. Значит, все коэффициенты линейной комбинации (7) равны нулю, т. е. векторы e_1, e_2, \dots, e_{k+1} — линейно независимы. Таким образом, на основании метода математической индукции утверждение теоремы доказано.

С л е д с т в и е. *В n -мерном линейном пространстве линейный оператор \hat{A} не может иметь более n собственных векторов с различными собственными значениями.*

Доказательство. (От противного). Пусть в R_n оператор \hat{A} имеет $n+1$ собственных векторов с различными собственными значениями. Тогда по теореме 5 указанные собственные векторы e_1, e_2, \dots, e_{n+1} линейно независимы. Следовательно, они или образуют базис в пространстве R_n или входят в базис из большего числа элементов. Значит, размерность пространства $R_n \geq n+1$, что противоречит условию теоремы.

О п р е д е л е н и е. *Уравнение*

$$\det(A_e - \lambda E) = 0 \quad (9)$$

называется характеристическим уравнением линейного оператора \hat{A} , действующего в линейном пространстве R_n , с матрицей A_e в некотором базисе $(e_p)_n$.

Утверждение 11 *Характеристическое уравнение (9) не зависит от выбора базиса в линейном пространстве.*

Доказательство. Пусть $(f_p)_n$ — базис, полученный из базиса $(e_p)_n$ при помощи невырожденного преобразования $f = eP$, т. е. $\det P \neq 0$. Тогда, поскольку $A_f = P^{-1}A_eP$ в силу (3) и $PEP^{-1} = E$, то $\det(A_f - \lambda E) = \det(P^{-1}A_eP - \lambda P^{-1}EP) =$ (свойства операции умножения матриц) $= \det P^{-1}(A_e - \lambda E)P =$ (т. 5 гл. 4 курса аналитической геометрии) $= \det P^{-1} \cdot \det(A_e - \lambda E) \cdot \det P = \det(A_e - \lambda E)$. Отсюда следует утверждение теоремы.

На вопрос: "Как найти собственные значения оператора \hat{A} "? — ответ дает следующая

Теорема 6 *Для того, чтобы число λ было собственным значением оператора \hat{A} , действующего в линейном пространстве R_n над полем K , необходимо и достаточно, чтобы это число было решением характеристического уравнения (9) оператора \hat{A} и принадлежало числовому полю K .*

Необходимость. Пусть λ — собственное значение оператора \hat{A} , а x — соответствующий собственный вектор этого оператора, т. е. $\lambda \in K$ и выполнено равенство (5'). Пусть $(e_p)_n$ — базис в R_n , а $A_e = \|\|a_k^p\|\|_n^n$ — матрица оператора \hat{A} в этом базисе. Из равенства (5') элементов пространства R_n следует равенство их координат в базисе $(e_p)_n$. По формуле (2) элемент $(\hat{A} - \lambda \hat{E})x$ имеет координаты $(A_e - \lambda E)X_e$, а элемент θ — координаты — нулевой столбец, т. е. имеем:

$$(A_e - \lambda E)X_e = \theta, \quad (10)$$

т. е. получили однородную систему уравнений относительно координат x^1, x^2, \dots, x^n вектора x в базисе $(e_p)_n$. Система (10) имеет нетривиальное решение тогда и только тогда, когда, как следует из следствия 3 к теореме 3 гл. 1 (теореме о базисном миноре), $\det(A_e - \lambda E) = 0$.

Достаточность. Пусть $\lambda \in K$ — решение характеристического уравнения (9). Но тогда система (10) имеет согласно следствию 3 к теореме 3 гл. 1 нетривиальное решение X_e . А поскольку (10) означает равенство координат элементов пространства R_n в базисе $(e_p)_n$, то равны и сами эти элементы, т. е. верно равенство (5'), причем $x \neq \theta$. Поэтому λ — собственное значение оператора \hat{A} . Теорема доказана.

Следствие 1. *Если характеристическое уравнение (9) оператора \hat{A} имеет n различных корней из поля K , то в некотором базисе $(e_p)_n$ матрица оператора \hat{A} имеет диагональный вид.*

Доказательство. Пусть $\lambda_1, \lambda_2, \dots, \lambda_n$ из K — различные собственные значения оператора \hat{A} . Каждому λ_p соответствует собственный вектор e_p . Тогда по теореме 5 векторы e_1, e_2, \dots, e_n — линейно независимы, т. е. образуют базис в пространстве R_n . Но в таком базисе $\hat{A}e_p = \lambda_p e_p$, $p = \overline{1, n}$, т. е. матрица A_e оператора \hat{A} имеет вид $\|\|\lambda_p \delta_k^p\|\|$.

Следствие 2. *Не каждый линейный оператор \hat{A} имеет собственное значение.*

Доказательство. Характеристическое уравнение (9) является уравнением степени n относительно λ . Поэтому по основной теореме алгебры уравнение (9) всегда имеет по крайней мере один корень, но не более n корней (следствие к т. 5). Однако корни этого уравнения могут не принадлежать числовому полю K , над которым действует линейное пространство R_n . В этом случае оператор \hat{A} не имеет в пространстве R_n собственных векторов.

Из теоремы 6 следует метод построения собственных значений и собственных векторов линейного оператора \hat{A} :

1) Решаем характеристическое уравнение (9). Если есть действительные корни, принадлежащие числовому полю K , то имеем собственные значения $\lambda_1, \lambda_2, \dots, \lambda_k$ ($k \leq$

n).

2) Подставляя вместо λ значение λ_p , найденное в пункте 1), в систему уравнений (10), находим координаты x^1, x^2, \dots, x^n собственных векторов x , соответствующих собственному значению λ_p , $p = \overline{1, k}$.

З а м е ч а н и е. Однородная система уравнений (10) с определителем матрицы системы, равным нулю, имеет бесконечно много решений, образующих линейное пространство. Оно является инвариантным подпространством относительно оператора \hat{A} .

П р и м е р 1. Пусть в базисе $(e_p)_3$ пространства R_3 над полем K рациональных чисел оператор \hat{A} имеет матрицу:

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

Найти все собственные векторы оператора \hat{A} .

Р е ш е н и е. Характеристическое уравнение (9):

$$\begin{vmatrix} 2 - \lambda & -1 & -1 \\ 0 & -1 - \lambda & 0 \\ 0 & 2 & 1 - \lambda \end{vmatrix} = 0.$$

Его решения: $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = -1 \in K$.

Система (10):

$$\begin{cases} (2 - \lambda)x^1 - x^2 - x^3 = 0 \\ -(1 + \lambda)x^2 = 0 \\ 2x^2 + (1 - \lambda)x^3 = 0 \end{cases},$$

которая при $\lambda_1 = 2$ имеет вид:

$$\begin{cases} -x^2 - x^3 = 0 \\ -3x^2 = 0 \\ 2x^2 - x^3 = 0 \end{cases},$$

Ее ФСР:

$$Y_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Таким образом, все множество собственных векторов, соответствующих собственному значению $\lambda = 2$, есть $X_1 = cY_1$, где $c \in K$, но $c \neq 0$.

При $\lambda_2 = 1$:

$$\begin{cases} x^1 - x^2 - x^3 = 0 \\ 2x^2 = 0 \\ 2x^2 = 0 \end{cases} \Rightarrow$$

ФСР: $Y_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Собственные векторы: $X_2 = cY_2$, где $c \in K$, $c \neq 0$.

При $\lambda_3 = -1$:

$$\begin{cases} 3x^1 - x^2 - x^3 = 0 \\ 2x^2 + 2x^3 = 0 \end{cases} \Rightarrow$$

ФСР: $Y_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Собственные векторы: $X_3 = cY_3$, где $c \in K$, $c \neq 0$.

Пример 2. Пусть в базисе $(e_p)_2$ линейного пространства R_2 над полем K рациональных чисел оператор \hat{A} имеет матрицу

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

Найти все собственные векторы \hat{A} .

Решение. Характеристическое уравнение(9): $\begin{vmatrix} -1-\lambda & 2 \\ 3 & 1-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 7 = 0 \Rightarrow \lambda_{1,2} = \pm\sqrt{7} \notin K$. Значит, нет собственных значений и собственных векторов у оператора \hat{A} .

Замечание. Если бы в примере 2 $K = K_0$, то найденные значения λ были бы собственными значениями, и по аналогии с примером 1 можно было бы найти все собственные векторы оператора \hat{A} .

Пример 3. Пусть в базисе $(e_p)_3$ линейного пространства R_3 над полем K рациональных чисел линейный оператор \hat{A} имеет матрицу

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

Найти все собственные векторы оператора \hat{A} .

Решение. Характеристическое уравнение (9):

$$\begin{vmatrix} -1-\lambda & -2 & -2 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0 \Rightarrow -(1+\lambda)(1-\lambda)^2 = 0,$$

т. е. собственные значения $\lambda_1 = -1, \lambda_{2,3} = 1 \in K$. Система (10):

$$\begin{cases} -(1+\lambda)x^1 - 2x^2 - 2x^3 = 0 \\ (1-\lambda)x^2 = 0 \\ (1-\lambda)x^3 = 0 \end{cases}$$

при $\lambda_1 = -1$ имеет вид:

$$\begin{cases} -2x^2 - 2x^3 = 0 \\ 2x^2 = 0 \\ 2x^3 = 0 \end{cases} \Rightarrow \text{ФСР: } Y_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Поэтому все собственные векторы, соответствующие собственному значению -1 , задаются линейной комбинацией $X = cY_1$, где $c \in K, c \neq 0$. При $\lambda_{2,3} = 1$ система (10) имеет вид:

$$x^1 + x^2 + x^3 = 0.$$

Ранг матрицы этой системы равен 1. Поэтому размерность пространства решений $n - r = 3 - 1 = 2$. Значит ФСР состоит из двух линейно независимых решений:

$$Y_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Поэтому множество всех собственных векторов, соответствующих собственному значению 1, задаются линейной комбинацией $X = aY_1 + bY_2$, где $a^2 + b^2 \neq 0, a, b, \in K$.

Пример 4. Решить задачу аналогичную примеру 3, но для матрицы

$$A_e = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{pmatrix}.$$

Решение. Здесь характеристическое уравнение (9) имеет вид: $(2 - \lambda)(\lambda^2 - 4\lambda + 4) = 0 \Rightarrow (2 - \lambda)^3 = 0$, т. е. оператор \hat{A} имеет единственное собственное значение $\lambda = 2$ кратности 3. В этом случае система (10) принимает вид: $x^2 = 0$, $x^3 + x^2 = 0$. Ранг матрицы этой системы равен 2. Поэтому размерность пространства решений равна $3 - 2 = 1$, т. е. ФСР состоит из одного решения $Y_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$ множество

всех собственных векторов оператора \hat{A} , соответствующих собственному значению 2, имеет вид: $X = cY_1$, $c \in K$, $c \neq 0$.

З а м е ч а н и е. 1) Если собственное значение λ — простой, т. е. не кратный, корень уравнения (9), то соответствующее ему инвариантное относительно оператора \hat{A} подпространство — одномерно. Последние играют фундаментальную роль в приложениях и именуется характеристическими направлениями.

2) Если собственное значение λ — корень кратности p характеристического уравнения (9), то соответствующее ему инвариантное подпространство может иметь размерность k от 1 до p (см. примеры 3, 4).

Пример 5. Пусть три единичных массы подвешены на пружинах с коэффициентами упругости c_1, c_2, c_3 . Вычислить основные частоты колебаний системы для случая, когда $c_1 = 8$, $c_2 = 3$, $c_3 = 11$.

Решение. Если обозначить силы, действующие на массы через K_p , а отклонения от положения равновесия через v , то по второму закону Ньютона получаем уравнения движения:

$$\begin{aligned} -y_1'' &= K_1 - K_2 = c_1 y_1 - c_2 (y_2 - y_1), \\ -y_2'' &= K_2 - K_3 = c_2 (y_2 - y_1) - c_3 (y_3 - y_2), \\ -y_3'' &= K_3 = c_3 (y_3 - y_2). \end{aligned}$$

Решение получившейся системы ищем в виде: $y_p = x_p \sin \omega t$. Тогда после подстановки предполагаемого решения в систему уравнений, получим:

$$Ax = \lambda x, \quad (11)$$

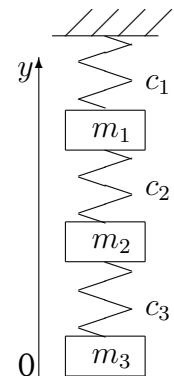
где $\lambda = \omega^2$, $x = \|x^p\|^3$,

$$A = \begin{pmatrix} c_1 + c_2 & -c_2 & 0 \\ -c_2 & c_2 + c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{pmatrix} = \begin{pmatrix} 11 & -3 & 0 \\ -3 & 14 & -11 \\ 0 & -11 & 11 \end{pmatrix}.$$

Таким образом, мы пришли к задаче на собственные значения. Характеристическое уравнение (9):

$$\det(A - \lambda E) = \begin{vmatrix} c_1 + c_2 - \lambda & -c_2 & 0 \\ -c_2 & c_2 + c_3 - \lambda & -c_3 \\ 0 & -c_3 & c_3 - \lambda \end{vmatrix} = \begin{vmatrix} 11 - \lambda & -3 & 0 \\ -3 & 14 - \lambda & -11 \\ 0 & -11 & 11 - \lambda \end{vmatrix} = 0,$$

т. е. $\lambda^3 - 36\lambda^2 + 299\lambda - 264 = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = 11, \lambda_3 = 24$. Итак, основные частоты найдены. Это $\omega_1 = 1$, $\omega_2 = \sqrt{11}$, $\omega_3 = 2\sqrt{6}$. Решая систему (11) для каждого собственного значения, найдем соответствующие собственные векторы (с точностью до



числового множителя, не равного нулю):

$$X_1 = \begin{pmatrix} 3 \\ 10 \\ 11 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 11 \\ 0 \\ -3 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 3 \\ -13 \\ 11 \end{pmatrix},$$

координаты которых — отношения амплитуд колебания масс, соответствующих данной основной частоте.

Пример 6. Протяженная изотропная среда подвержена деформации, при которой единичный куб с ребрами

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

переходит в параллелепипед с ребрами

$$f_1 = \begin{pmatrix} 1,5 \\ 0,5 \\ 0,5 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 0,5 \\ 1 \\ 0 \end{pmatrix}, \quad f_3 = \begin{pmatrix} 0,5 \\ 0 \\ 1 \end{pmatrix}.$$

Каковы главные оси деформации, т. е. направления, которые сохраняются при деформации?

Решение. Деформация описывается симметричной матрицей

$$A_e = \begin{pmatrix} 1,5 & 0,5 & 0,5 \\ 0,5 & 1 & 0 \\ 0,5 & 0 & 1 \end{pmatrix},$$

для которой $\hat{A}e_k = f_k$, $k = 1, 2, 3$. Следовательно, указанная деформация — линейный оператор \hat{A} , а искомые направления — собственные векторы этого оператора, точнее его инвариантные подпространства. Матрица A_e — матрица оператора \hat{A} в базисе $(e_p)_3$. Его характеристическое уравнение (9): $-\lambda^3 + 3,5\lambda^2 - 3,5\lambda + 1 = 0$ имеет решения: $\lambda_1 = 0,5$, $\lambda_2 = 1$, $\lambda_3 = 2$. Отвечающие им собственные векторы:

$$X_1 = c \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad X_2 = c \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad X_3 = c \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad c \in K_0, \quad c \neq 0.$$

Отметим попарную ортогональность собственных векторов, соответствующих различным собственным значениям, т. е. $X_p^T X_k = 0$, $p, k = 1, 2, 3$, $p \neq k$.

§6. Симметричный (самосопряженный) линейный оператор в конечномерном евклидовом пространстве.

Рассмотрим линейные операторы, действующие в евклидовом пространстве E_n . Поэтому здесь числовое поле $K = K_0$.

Определение. *Линейный оператор \hat{A} , действующий в пространстве E_n , называется симметричным, если для всех x, y из E_n выполняется равенство:*

$$(\hat{A}x, y) = (x, \hat{A}y). \quad (12)$$

Определение. *Матрица A называется симметричной, если $A = A^T$, т. е. $a_k^p = a_p^k$, $p, k = \overline{1, n}$.*

Утверждение 12 Матрица симметричного оператора \hat{A} в ортонормированном базисе симметрична.

Доказательство. Пусть $(e_p)_n$ — ортонормированный базис, а $A_e = \|a_k^p\|_n^n$ — матрица оператора \hat{A} в этом базисе. По формуле (1) $\hat{A}e_p = e_k a_p^k$. Поэтому $(\hat{A}e_p, e_i) = a_p^k (e_k, e_i) = (из\ ортонормированности\ базиса) = a_p^k \delta_k^i = a_p^i$. С другой стороны, $a_p^i = (\hat{A}e_p, e_i) = (симметричность\ оператора) = (e_p, \hat{A}e_i) = (e_p, e_k a_i^k) = (e_p, e_k) a_i^k = (ортogonalность\ базиса) = \delta_k^p a_i^k = a_i^p$.

Утверждение 13 Оператор \hat{A} симметричен, если в некотором ортонормированном базисе $(e_p)_n$ имеет симметричную матрицу, т. е. $A_e = A_e^T$.

Доказательство. Как следует из равенства (3) гл. 4, скалярное произведение любых двух элементов в ортонормированном базисе равняется сумме произведений соответствующих координат перемножаемых элементов. Так как элемент $\hat{A}x$ в силу (2) имеет координаты $A_e X_e$, а элемент y — координаты Y_e , то $(\hat{A}x, y) = (формула\ (3)\ гл.\ 4) = (A_e X_e)^T Y_e = (теорема\ 4\ гл.\ 4\ курса\ аналитической\ геометрии) = X_e^T A_e^T Y_e = (симметричность\ матрицы\ A_e) = X_e^T A_e Y_e$. С другой стороны, $(x, \hat{A}y) = (формула\ (3)\ гл.\ 4) = X_e^T A_e Y_e$. Из этих двух равенств следует, что $(\hat{A}x, y) = (x, \hat{A}y)$.

Утверждение 14 Симметричный оператор \hat{A} , действующий в пространстве E_n , является симметричным в любом инвариантном относительно оператора \hat{A} подпространстве M_k ($k \leq n$) пространства E_n .

Доказательство. Поскольку M_k — инвариантное относительно оператора \hat{A} подпространство, то для каждого x из M_k $\hat{A}x \in M_k$. А так как для любых x, y из E_n выполняется (12), то, в частности, для всех x, y из M_k так же имеет место (12).

Теорема 7 Все корни характеристического уравнения симметричного оператора — действительные числа.

Доказательство. (от противного). Пусть \hat{A} — симметричный оператор, действующий в E_n , а $(e_p)_n$ — ортонормированный базис в E_n . Тогда по утверждению 12 $A_e = A_e^T$. Пусть существует комплексное число $\lambda = a + ib$ такое, что $\det(A_e - \lambda E) = 0$. Тогда система (10) есть квадратная система однородных линейных уравнений с матрицей $A_e - \lambda E$, определенной на числовом поле $K = C$ — поле всех комплексных чисел. При введении числовых матриц (гл. 1) и решении систем линейных уравнений (гл. 1 и 3) мы пользовались произвольным числовым полем K . Поэтому все, что было в этих главах получено, справедливо и для числового поля C . Следовательно, система (10) имеет ненулевое решение $X_e = \|x^p\|_n$ тогда и только тогда, когда определитель матрицы этой системы равен нулю. Но у нас это условие, как отмечалось выше, выполнено. Значит, имеем решение: $X_e = B + iC = \|b^p\|_n + i\|c^p\|_n$, причем $X_e \neq \theta + i\theta$, где θ — нулевой столбец высоты n . Подставим это решение с соответствующим значением $\lambda = a + ib$ в систему (10). В итоге получим матричное тождество: $A_e(B + iC) = (a + ib)(B + iC)$ или, приравнявая отдельно действительную и отдельно мнимую части этого тождества, найдем, что

$$A_e B = aB - bC, \quad (13)$$

$$A_e C = bB - aC. \quad (14)$$

Применив операцию транспонирования к равенству (13), получим: $B^T A_e^T = aB^T - bC^T$. Так как $A_e^T = A_e$, то окончательно

$$B^T A_e = aB^T - bC^T. \quad (15)$$

Умножим равенство (15) на матрицу C справа, а равенство (14) на матрицу B^T слева и из второго равенства вычтем первое: $\theta = b(B^T B + C^T C)$, что в координатах запишется так:

$$b \sum_{k=1}^n ((b^k)^2 + (c^k)^2) = 0. \quad (16)$$

Так как $X_e \neq \theta + i\theta$, то среди чисел b^k и c^k имеются числа, отличные от 0, т. е. сумма, стоящая в равенстве (16) отлична от нуля. Поэтому из (16) следует, что $b = 0$, что противоречит предположению о том, что $\lambda = a + ib$ — комплексное число.

З а м е ч а н и е. Теорема 7, как следует из утверждения 14, справедлива в каждом инвариантном относительно симметричного оператора \hat{A} подпространстве M_k .

С л е д с т в и е. *Оператор \hat{A} симметричный в E_n всегда имеет собственные векторы.*

Теорема 8 *Если x — собственный вектор симметричного оператора \hat{A} в E_n , то совокупность всех векторов y , ортогональных вектору x , образует инвариантное относительно оператора \hat{A} подпространство M_{n-1} .*

Д о к а з а т е л ь с т в о. Так как x — собственный вектор оператора \hat{A} , то, как следует из следствия к утверждению 10, линейная оболочка $L(x) = \{ax, a \in (-\infty, +\infty)\}$ — одномерное инвариантное относительно оператора подпространство линейного пространства E_n . Но тогда совокупность M всех элементов y из E_n , ортогональных к элементу x , — ортогональное дополнение подпространства $L(x)$. Следовательно, по теореме 3 гл. 4 M — подпространство линейного пространства E_n размерности $n - 1$. Осталось показать, что оно инвариантно относительно оператора \hat{A} , т. е. если $y \in M$, то и $\hat{A}y \in M$. Действительно, $(\hat{A}y, x) = (\text{симметричность } \hat{A}) = (y, \hat{A}x) = (x - \text{собственный вектор } \hat{A}) = (y, \lambda x) = (a \cdot 3, 1 \text{ скалярного произведения}) = \lambda(x, y) = (y \in M - \text{ортогональному дополнению к } L(x)) = 0$, т. е. $\hat{A}y$ ортогонально к x , т. е. $\hat{A}y \in M$.

Теорема 9 *Для того, чтобы линейный оператор \hat{A} был симметричен, необходимо и достаточно, чтобы в E_n существовал ортонормированный базис из собственных векторов оператора \hat{A} .*

Д о к а з а т е л ь с т в о. **Н е о б х о д и м о с т ь.** Так как \hat{A} — симметричный оператор, то в силу следствия к теореме 7 существует собственный вектор этого оператора x_1 . Тогда подпространство M_{n-1} векторов, ортогональных x_1 , как следует из теоремы 8, является инвариантным. Следовательно, к этому подпространству, как утверждает замечание к теореме 7, снова применима теорема 7. Поэтому в M_{n-1} существует собственный вектор этого оператора — x_2 , причем x_2 ортогонален x_1 (ведь $x_2 \in M_{n-1}$, а M_{n-1} ортогонально x_1). Рассмотрим множество всех векторов из M_{n-1} , ортогональных к x_2 . По теореме 8 это множество есть инвариантное подпространство M_{n-2} относительно оператора \hat{A} . Следовательно, в M_{n-2} существует по теореме 7 собственный вектор оператора \hat{A} — x_3 , причем x_3 ортогонален к x_1, x_2 . Продолжая этот процесс далее получим n ортогональных друг другу векторов x_1, x_2, \dots, x_n . Как показано в лемме 1 гл. 4 ортогональные векторы линейно независимы. Поэтому n построенных векторов образуют базис в E_n . Нормируя их, т. е. полагая $e_p = x_p/|x_p|$, $p = \overline{1, n}$, получим ортонормированный базис, состоящий, как следует из утверждения 10, из собственных векторов.

Д о с т а т о ч н о с т ь. Пусть $(e_p)_n$ — ортонормированный базис в E_n из собственных векторов оператора \hat{A} , соответствующих собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_n$ (собственные значения с разными индексами могут совпадать). Как следует из формулы (1) элементы матрицы A_e оператора \hat{A} в данном базисе имеют вид: $a_p^k = \lambda_p \delta_p^k$,

т. е. матрица A_e — диагональная. Следовательно $A_e = A_e^T$. Поэтому, в силу утверждения 13, оператор \hat{A} симметричен.

Утверждение 15 *Собственные векторы симметричного оператора \hat{A} , соответствующие различным собственным значениям, ортогональны между собой.*

Доказательство. Пусть собственный вектор x_1 соответствует собственному значению λ_1 , а собственный вектор x_2 — λ_2 , ($\lambda_1 \neq \lambda_2$), т. е. $\hat{A}x_1 = \lambda_1 x_1$ и $\hat{A}x_2 = \lambda_2 x_2$. Так как \hat{A} — симметричный оператор, то $(\hat{A}x_1, x_2) = (x_1, \hat{A}x_2)$ или так: $(\lambda_1 x_1, x_2) = (x_1, \lambda_2 x_2)$ или после применения a_3^0 скалярного произведения $\lambda_1(x_1, x_2) = \lambda_2(x_1, x_2)$ или $(\lambda_1 - \lambda_2)(x_1, x_2) = 0$. Так как $\lambda_1 \neq \lambda_2$, то $(x_1, x_2) = 0$. Что и требовалось доказать.

В приложениях часто приходится сталкиваться с задачей вида: найти множество всех элементов x из E_n , удовлетворяющих равенству

$$\hat{A}x = \lambda x + b, \quad (17)$$

где \hat{A} — симметричный оператор, действующий в E_n , b — фиксированный элемент пространства E_n .

Решение. Пусть $(e_p)_n$ — базис в E_n , в котором оператор \hat{A} имеет матрицу A_e , элементы x, b — координаты X_e, B_e . Тогда в силу формулы (1*) и единственности разложения по базису, получаем что равенство (17) равносильно следующей системе линейных уравнений:

$$(A_e - \lambda E)X_e = B_e. \quad (18)$$

По альтернативе Фредгольма система (18) или имеет единственное решение для любого столбца B_e при условии, что соответствующая однородная система уравнений

$$(A_e - \lambda E)X_e = \theta \quad (19)$$

имеет только тривиальное решение, что возможно, если $\det(A_e - \lambda E) \neq 0$, т. е. λ не является собственным значением оператора \hat{A} , или совместна, если столбец B_e ортогонален пространству решений соответствующей сопряженной однородной системы уравнений, при условии, что система (19) имеет нетривиальное решение. Последнее возможно, если $\det(A_e - \lambda E) = 0$, т. е. λ — собственное значение оператора \hat{A} . Так как для симметричного оператора $A_e = A_e^T$, то сопряженная система совпадает с системой (19), решения которой в этом случае дает все множество собственных векторов оператора \hat{A} , соответствующих данному собственному значению λ . Поэтому в данном случае альтернатива Фредгольма имеет вид:

1) или λ не является собственным значением оператора \hat{A} , и тогда уравнение (17) имеет единственное решение x при любом выборе элемента b из E_n ;

2) или λ — собственное значение оператора \hat{A} , и тогда уравнение (17) имеет решение тогда и только тогда, когда элемент b ортогонален к собственным векторам оператора \hat{A} , соответствующим данному собственному значению.

Эта теорема легко распространяется на случай произвольного линейного оператора \hat{A} , если ввести еще одно понятие.

Определение. *Оператор \hat{A}^* , действующий в евклидовом пространстве E_n , называется сопряженным к линейному оператору \hat{A} , если для любых x, y из E_n выполняется соотношение:*

$$(\hat{A}x, y) = (x, \hat{A}^*y). \quad (20)$$

Утверждение 16 *Если в ортонормированном базисе $(e_p)_n$ оператор \hat{A} имеет матрицу A_e , то ему сопряженный оператор \hat{A}^* имеет матрицу $A_e^* = A_e^T$.*

Доказательство. Так как $(e_p)_n$ — ортонормированный базис, то по формуле (3) гл. 4 $(\hat{A}x, y) = (A_e, X_e)^T Y_e =$ (теорема 4 гл. 4, курса аналитической геометрии) $= X_e^T A_e^T Y_e$ и $(x, \hat{A}^*y) = X_e^T A_e^* Y_e$. Но тогда из равенства (20) следует, что $X_e^T A_e^T Y_e = X_e^T A_e^* Y_e \Rightarrow A_e^* = A_e^T$.

Утверждение 17 *Собственные значения операторов \hat{A} и \hat{A}^* совпадают.*

Доказательство. Характеристические уравнения (9) для этих операторов имеют вид:

$$\det(A_e - \lambda E) = 0 \quad \text{и} \quad \det(A_e^T - \lambda E) = 0.$$

Но $\det(A_e^T - \lambda E) = \det(A_e - \lambda E)^T =$ (1 свойство определителей n -го порядка) $= \det(A_e - \lambda E)$. Следовательно, характеристические уравнения этих операторов совпадают, что и означает совпадение собственных значений.

Собственные же векторы у этих операторов будут различны, так как определяются различными системами: для оператора \hat{A} — система (19), а для \hat{A}^* — система:

$$(A_e^T - \lambda E)X_e = \theta. \quad (21)$$

Так как для произвольного оператора, вообще говоря, $A_e \neq A_e^T$, то сопряженная однородная система уравнений, соответствующая системе (18), уже не будет совпадать с системой (19). Это система (21). Итак, в общем случае альтернатива Фредгольма может быть сформулирована следующим образом:

1) или λ не является собственным значением оператора \hat{A} , и тогда уравнение (17) имеет единственное решение x при любом выборе элемента b из E_n ;

2) или λ — собственное значение оператора \hat{A} , и тогда уравнение (17) имеет решение тогда и только тогда, когда элемент b ортогонален к собственным векторам оператора \hat{A}^* , соответствующим данному собственному значению.

О п р е д е л е н и е. *Линейный оператор \hat{Q} , действующий в евклидовом пространстве E , называется ортогональным, если для любых x, y из E выполняется равенство:*

$$(\hat{Q}x, \hat{Q}y) = (x, y). \quad (22)$$

Отсюда непосредственно следует, что если $(e_p)_n$ — ортонормированный базис в E , то $(\hat{Q}e_p)_n$ — также ортонормированный базис в E .

Утверждение 18 *В любом ортонормированном базисе $(e_p)_n$ матрица Q_e ортогонального оператора \hat{Q} является ортогональной матрицей.*

Доказательство. Если $(e_p)_n$ — ортонормированный базис, то по формуле (3) гл. 4 $(\hat{Q}x, \hat{Q}y) = (Q_e X_e)^T Q_e Y_e = X_e^T Q_e^T Q_e Y_e = X_e^T (Q_e^T Q_e) Y_e$ и $(x, y) = X_e^T Y_e$. Тогда в силу (22) $Q_e^T Q_e = E$, что означает, что Q_e — ортонормированная матрица.

Теорема 10 *Для того чтобы линейный оператор \hat{Q} был ортогонален, необходимо и достаточно, чтобы существовал оператор \hat{Q}^{-1} и было выполнено равенство*

$$\hat{Q}^* = \hat{Q}^{-1}, \quad (23)$$

где \hat{Q}^* — оператор, сопряженный к \hat{Q} , а \hat{Q}^{-1} — оператор, обратный к \hat{Q} .

Доказательство. Пусть $(e_p)_n$ — ортонормированный базис в E_n , в котором оператор \hat{Q} имеет матрицу Q_e . Тогда $Q_e^* = Q_e^T$, и равенство (23) равносильно равенству $Q_e^T = Q_e^{-1}$, которое совпадает с (13) гл. 4. Как следует из теорем 1, 2 пространство операторов изоморфно пространству $n \times n$ -матриц. Но в пространстве $n \times n$ -матриц эта теорема доказана. Следовательно, утверждение теоремы верно и в пространстве операторов.

Гл. 6. Квадратичные формы.

В этой главе будем рассматривать поле K_0 — поле всех действительных чисел.

§1. Общие понятия.

О п р е д е л е н и е. *Квадратичной формой называется функция n переменных x^1, x^2, \dots, x^n из K_0 вида*

$$\phi(x^1, x^2, \dots, x^n) = a_{pk}x^p x^k, \quad p, k = \overline{1, n} \quad (1)$$

где $a_{pk} \in K_0$, причем $a_{pk} = a_{kp}$, т. е. матрица квадратичной формы $A = \|a_{pk}\|_{nn}$ есть симметричная матрица.

П р и м е р. $\phi(x^1, x^2, x^3) = 4x^1x^2 - x^2x^2 - 6x^1x^3 + 2x^3x^3 + 2x^2x^3$. Так как у квадратичной формы $a_{pk} = a_{kp}$, то все смешанные произведения надо разбить на сумму двух равных слагаемых. Например: $4x^1x^2 = 2x^1x^2 + 2x^2x^1$. В итоге получим матрицу квадратичной формы:

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

З а м е ч а н и е. Выписывая матрицу квадратичной формы, не забывайте, что смешанные члены в ней берутся удвоенными.

Очевидно, что матрица A однозначно определяет квадратичную форму.

В ы в о д: всякой квадратичной форме соответствует единственная симметричная матрица A порядка n ; обратно: если есть симметричная матрица A порядка n , то мы можем написать квадратичную форму с матрицей A .

Квадратичную форму (1) можно записать в матричной форме, если ввести, как обычно, обозначения: $X = \|x^k\|^n$ — матрица-столбец высоты n , $X^T = (x^1, x^2, \dots, x^n)$ — матрица-строка длины n . Тогда $AX = \|a_{pk}x^k\|^n$ — матрица-столбец высоты n . Если теперь эту матрицу умножить слева на матрицу X^T , то получим $I \times I$ -матрицу с элементом $a_{pk}x^p x^k$ — квадратичную форму (1), т. е.

$$\phi(x^1, \dots, x^n) = X^T A X. \quad (2)$$

§2. Изменение квадратичной формы при линейном преобразовании переменных.

Пусть $P = \|p_{ek}\|_{n,n}$, $Q = \|q_{ek}\|_{n,n}$, $A = \|a_{ek}\|_{n,n}$, $B = \|b_{ek}\|_{n,n}$, $X = \|x^k\|^n$, $Y = \|y^k\|^n$, $Z = \|z^k\|^n$.

О п р е д е л е н и е. *Линейным преобразованием переменных y^1, \dots, y^n в переменные x^1, \dots, x^n называется преобразование вида*

$$X = PY, \quad (3)$$

причем, матрица P называется матрицей линейного преобразования (3).

З а м е ч а н и е. Здесь ради удобства изложения материала этой главы используется матричная форма записи оператора \hat{P} или, что то же самое, линейного преобразования пространства.

Если P — невырожденная матрица, то преобразование (3) называется невырожденным. В противном случае оно называется вырожденным преобразованием. Пусть P — невырожденная матрица, т. е. $\det P \neq 0$. Тогда для этой матрицы существует

обратная P^{-1} , с помощью которой можно совершить линейное преобразование от переменных x^1, \dots, x^n к переменным y^1, \dots, y^n :

$$Y = P^{-1}X. \quad (4)$$

Преобразование (4) называется обратным к преобразованию (3).

Пусть даны два последовательных преобразования $Z \rightarrow Y \rightarrow X$, т. е. преобразование (3) и преобразование

$$Y = QZ. \quad (5)$$

Их можно заменить одним преобразованием $Z \rightarrow X$, т. е. $X = BZ$, где $B = PQ$. К этому выводу мы приходим, подставив значение Y из (5) в формулу (3).

О п р е д е л е н и е. Преобразование $X = PQZ$ называется произведением преобразований (3) и (5).

Как следует из теоремы 5 гл. 4 курса аналитической геометрии, произведение преобразований невырождено тогда и только тогда, когда каждый из сомножителей является невырожденным преобразованием.

Теорема 1 Квадратичная форма X^TAX после применения линейного преобразования (3) переходит в квадратичную форму Y^TBY , где

$$B = P^TAP. \quad (6)$$

Д о к а з а т е л ь с т в о. $X^TAX =$ (формула (3)) $= (PY)^T A(PY) =$ (т. 4 гл. 4 курса аналитической геометрии и свойство 1 операции умножения матриц) $= Y^T P^T A P Y = Y^T(P^TAP)Y$. Теорема доказана.

З а м е ч а н и е. B — симметричная матрица. Действительно, $B^T =$ (формула (6)) $= (P^TAP)^T =$ (т. 4 гл. 4 курса аналитической геометрии) $= P^T A^T (P^T)^T =$ (определение транспонированной матрицы) $= P^T A^T P =$ (симметричность A) $= P^T A P =$ (формула (6)) $= B$.

Теорема 2 При любом невырожденном преобразовании переменных знак определителя матрицы квадратичной формы не меняется.

Д о к а з а т е л ь с т в о. Пусть даны квадратичная форма X^TAX и преобразование (3), причем $\det P \neq 0$. В результате этого преобразования, как следует из теоремы 1, квадратичная форма станет квадратичной формой Y^TBY , где $B = P^TAP$. Но $\det(P^TAP) =$ (т. 5 гл. 4 курса аналитической геометрии) $= \det P^T \cdot \det A \cdot \det P =$ (свойство 1 определителя) $= \det A \cdot (\det P)^2$. Откуда следует утверждение теоремы.

О п р е д е л е н и е. Рангом квадратичной формы называется ранг матрицы этой квадратичной формы.

Теорема 3 Ранг квадратичной формы сохраняется при любом невырожденном линейном преобразовании переменных.

Д о к а з а т е л ь с т в о. Пусть дана квадратичная форма (2). Совершим преобразование (3), причем $\det P \neq 0$. Тогда в силу теоремы 1 придем к квадратичной форме Y^TBY , где $B = P^TAP$. Так как $\det P = \det P^T \neq 0$, то по теореме 14 гл. 2 $\text{rang } B = \text{rang } P^T A = \text{rang } A$. Что и требовалось доказать.

Пусть квадратичная форма (2) некоторым линейным преобразованием (3) приведена к виду:

$$b_{kk}(y^k)^2. \quad (7)$$

О п р е д е л е н и е. Запись (7) называется каноническим видом квадратичной формы.

Теорема 4 Если квадратичная форма (2) невырожденным преобразованием (3) приведена к каноническому виду (7), то число отличных от нуля коэффициентов в (7) равно рангу r квадратичной формы.

Доказательство. Так как по условию теоремы $\det P \neq 0$, то в силу теоремы 3 квадратичная форма (7) имеет ранг r . С другой стороны, матрица квадратичной формы (7) имеет диагональный вид $\|b_{pk}\delta_{pk}\|_{n,n}$. Следовательно, на ее главной диагонали стоит r отличных от нуля элементов.

§3. Метод Лагранжа приведения квадратичной формы к каноническому виду.

Теорема 5 Любую квадратичную форму можно привести к каноническому виду (7) невырожденным преобразованием переменных.

Доказательство. Проведем по методу математической индукции относительно числа переменных n . Если $n = 1$, то квадратичная форма $\phi(x^1) = a_{11}x^1x^1$ — канонический вид квадратичной формы. Следовательно, в данном случае существует тождественное преобразование, т. е. для $n = 1$ утверждение теоремы имеет место. Считая теперь доказанным утверждение теоремы для $n - 1$, рассмотрим квадратичную форму $\phi(x^1, \dots, x^n) = a_{pk}x^p x^k$ и докажем для нее утверждение теоремы. Для этого придется рассмотреть 2 случая.

а) Хотя бы один из коэффициентов $a_{kk} \neq 0$, $k = \overline{1, n}$. Пусть ради определенности $a_{nn} \neq 0$. Выделим в квадратичной форме все члены, содержащие x^n :

$$a_{nn}(x^n)^2 + 2a_{kn}x^k x^n + a_{ep}x^e x^p, \quad k = \overline{1, n}; \quad p, e = \overline{1, n-1}.$$

Теперь дополним выделенные члены до полного квадрата слагаемыми, не содержащими x^n . Тогда

$$\phi(x^1, \dots, x^n) = a_{nn}(c_k x^k)^2 + \bar{\phi}(x^1, \dots, x^{n-1}).$$

где $c_k = a_{kn}/a_{nn}$, $k = \overline{1, n}$. По предположению индукции квадратичная форма $\bar{\phi}(x^1, \dots, x^{n-1})$ невырожденным преобразованием типа (3) приводится к каноническому виду:

$$\bar{\phi}(y^1, \dots, y^{n-1}) = b_{kk}y^k y^k, \quad k = \overline{1, n-1}.$$

Добавив к этому преобразованию соотношение $y^n = c_k x^k$, $k = \overline{1, n}$, получим преобразование:

$$\begin{cases} \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^{n-1} \end{pmatrix} = \bar{P} \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^{n-1} \end{pmatrix} \\ y^n = (a_{kn}/a_{nn})x^k. \end{cases} \quad (8)$$

Это преобразование невырожденное, так как определитель матрицы P этого преобразования

$$\begin{vmatrix} p_{1,1} & \dots & p_{1,n-1} & 0 \\ \dots & \dots & \dots & \dots \\ p_{n-1,1} & \dots & p_{n-1,n-1} & 0 \\ c_1 & \dots & c_{n-1} & 1 \end{vmatrix} = \det P \neq 0.$$

Следовательно, квадратичная форма $\phi(x^1, \dots, x^n)$ примет канонический вид:

$$\phi(y^1, \dots, y^n) = b_{kk}(y^k)^2 + a_{nn}(y^n)^2.$$

б) Все $a_{kk} = 0$, $k = \overline{1, n}$. Пусть $a_{12} \neq 0$. Тогда сделаем преобразование:

$$\begin{cases} x^1 = z^1 - z^2 \\ x^2 = z^1 + z^2 \\ x^k = z^k, \quad k = \overline{3, n} \end{cases}$$

Это преобразование невырожденное, так как определитель матрицы преобразования

$$\begin{vmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{vmatrix} = 2.$$

А в итоге, поскольку $2a_{12}x^1x^2 = 2a_{12}((z^1)^2 - (z^2)^2)$, в квадратичной форме появятся квадраты сразу двух переменных с отличными от нуля коэффициентами, т. е. придем к случаю а), для которого утверждение теоремы доказано.

Изложенный только что метод выделения полных квадратов называется методом Лагранжа.

Пример. Привести к каноническому виду квадратичную форму $\phi(x^1, x^2, x^3) = 2x^1x^2 + 2x^2x^3$. Матрица этой квадратичной формы

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Так как первый и третий столбцы совпадают, то $\text{rang } A = 2$. Поэтому в каноническом виде в силу теоремы 4 один коэффициент будет равен 0. Так как это случай б), то делаем преобразование:

$$\begin{cases} x^1 = z^1 - z^2 \\ x^2 = z^1 + z^2 \\ x^3 = z^3 \end{cases} \text{ или } X = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} Z = P \cdot Z, \text{ где } \det P = 2.$$

Тогда $\phi(z^1, z^2, z^3) = 2(z^1)^2 - 2(z^2)^2 + 2z^1z^3 + 2z^2z^3 = 2((z^1)^2 + z^1z^3 + 0.25(z^3)^2 - 0.25(z^3)^2) + 2z^2z^3 - 2(z^2)^2 = 2(z^1 + 0.5z^3)^2 - 2((z^2)^2 - z^2z^3 + 0.25(z^3)^2) = 2(z^1 + 0.5z^3)^2 - 2(z^2 - 0.5z^3)^2$.

Иными словами, нами сделано преобразование переменных:

$$\begin{cases} y^1 = z^1 + 0.5z^3 \\ y^2 = z^2 - 0.5z^3 \\ y^3 = z^3 \end{cases} \text{ или } Y = \begin{pmatrix} 1 & 0 & 0.5 \\ 0 & 1 & -0.5 \\ 0 & 0 & 1 \end{pmatrix} Z = P_1 \cdot Z, \text{ где } \det P_1 = 1.$$

Итак,

$$\phi(y^1, y^2, y^3) = 2(y^1)^2 - 2(y^2)^2 + 0 \cdot (y^3)^2.$$

Можно теперь указать и прямой переход от квадратичной формы $\phi(x^1, x^2, x^3)$ к квадратичной форме $\phi(y^1, y^2, y^3)$, т. е. $X = BY$, где $B = PP_1^{-1}$. Так как

$$P_1^{-1} = \begin{pmatrix} 1 & 0 & -0.5 \\ 0 & 1 & 0.5 \\ 0 & 0 & 1 \end{pmatrix}, \text{ то } B = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -0.5 \\ 0 & 1 & 0.5 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

т. е.

$$\begin{cases} x^1 = y^1 - y^2 - y^3 \\ x^2 = y^1 + y^2 \\ x^3 = y^3 \end{cases}$$

После этого преобразования получим:

$$\phi(y^1, y^2, y^3) = 2(y^1 - y^2 - y^3)(y^1 + y^2) + 2(y^1 + y^2)y^3 = 2(y^1)^2 - 2(y^2)^2.$$

§4. Приведение квадратичной формы ортогональным преобразованием к каноническому виду.

Теорема 6 Существует ортогональное преобразование $X = QY$, где Q — ортогональная матрица, приводящее квадратичную форму X^TAX к каноническому виду.

Доказательство. Рассмотрим произвольное евклидово пространство E_n с ортонормированным базисом $(e_p)_n$. Введем на этом пространстве линейный оператор \hat{A} , задав его матрицей A_e в базисе $(e_p)_n$, равной матрице A квадратичной формы $\phi(x^1, \dots, x^n) = X^TAX$. Так как $(e_p)_n$ — ортонормированный базис и $A = A^T$, то по утверждению 13 \hat{A} — симметричный оператор. По теореме 9 гл. 5 для оператора \hat{A} в пространстве E_n существует ортонормированный базис $(f_p)_n$ из собственных векторов этого оператора, в котором матрица A_f оператора \hat{A} имеет диагональный вид: $A_f = \|\lambda_p \delta_{pk}\|_{n,n}$. Обозначим через Q матрицу перехода от базиса $(f_p)_n$ к базису $(e_p)_n$, т. е. $f = eQ$. Так как оба базиса ортонормированны, то по свойству 8 ортогональной матрицы (§4, гл. 4) Q — ортогональная матрица. Как следует из равенства (3) гл. 5 $A_f = Q^{-1}A_eQ$ или, так как $A_e = A$ и $Q^{-1} = Q^T$, то $A_f = Q^T A Q$. Применим к квадратичной форме линейное преобразование $X = QY$. Тогда по теореме 1 матрица квадратичной формы X^TAX будет равна матрице $Q^T A Q$, т. е. матрице A_f — диагональной матрице.

Замечание 1. В этом случае коэффициенты канонического вида квадратичной формы $\phi(y^1, \dots, y^n)$ равны собственным значениям оператора \hat{A} . Поэтому, решив характеристическое уравнение (9), сразу можно записать канонический вид квадратичной формы.

Замечание 2. Ортогональное преобразование $X = QY$, приводящее квадратичную форму к каноническому виду, имеет матрицу Q , столбцы которой суть координаты ортонормированного базиса из собственных векторов оператора \hat{A} в базисе $(e_p)_n$.

Теорема 7 Коэффициенты канонического вида квадратичной формы не зависят от выбора ортогонального преобразования, приводящего квадратичную форму к каноническому виду.

Доказательство. Пусть ортогональное преобразование $X = QY$ приводит квадратичную форму X^TAX к каноническому виду $Y^TBY = a_p(y^p)^2$. Покажем, что $a_p = \lambda_p$, где λ_p — корень уравнения (9). Составим квадратичную форму с матрицей $A - \lambda E$: $X^T(A - \lambda E)X$. Ортогональное преобразование $X = QY$ приводит матрицу этой квадратичной формы согласно теореме 1 к виду: $Q^T(A - \lambda E)Q = B - \lambda Q^T E Q = B - \lambda E = \|(a_p - \lambda)\delta_{pk}\|_{n,n}$. Поэтому $\det Q^T(A - \lambda E)Q =$ (теорема 5 гл. 4 курса аналитической геометрии) $= \det(A - \lambda E) = \det \|(a_p - \lambda)\delta_{pk}\|_{n,n} = (a_1 - \lambda) \cdot (a_2 - \lambda) \cdot \dots \cdot (a_n - \lambda)$. Отсюда следует, что a_p — решение уравнения (9).

Пример 1.

$\phi(x^1, x^2, x^3) = 2x^1x^2 + 2x^2x^3$. $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ — матрица квадратичной формы.

Поэтому $A - \lambda E = \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix}$ и характеристическое уравнение имеет вид: $-\lambda^3 = 2\lambda = 0$. Откуда $\lambda_1 = 0$, $\lambda_2 = 2^{1/2}$, $\lambda_3 = -2^{1/2}$. Теперь решая систему уравнений

$(A - \lambda E)X_e = \theta$, т. е. систему

$$\begin{cases} -\lambda x^1 + x^2 = 0 \\ x^1 - \lambda x^2 + x^3 = 0 \\ x^2 - \lambda x^3 = 0, \end{cases}$$

при соответствующих значениях λ , найдем собственные векторы:

а) для $\lambda_1 = 0$ $\begin{cases} x^2 = 0 \\ x^1 + x^3 = 0 \end{cases} \Rightarrow x^1 = -x^3 \Rightarrow X_1 = c \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, c \neq 0;$

б) для $\lambda_2 = 2^{1/2}$ $\begin{cases} -2^{1/2}x^1 + x^2 = 0 \\ x^2 - 2^{1/2}x^3 = 0 \end{cases} \Rightarrow \begin{cases} x^1 = 2^{-1/2}x^2 \\ x^3 = 2^{-1/2}x^2 \end{cases} \Rightarrow X_2 = c \begin{pmatrix} 1 \\ 2^{1/2} \\ 1 \end{pmatrix}, c \neq 0;$

в) для $\lambda_3 = -2^{1/2}$ $\begin{cases} 2^{1/2}x^1 + x^2 = 0 \\ x^2 + 2^{1/2}x^3 = 0 \end{cases} \Rightarrow X_3 = c \begin{pmatrix} 1 \\ -2^{1/2} \\ 1 \end{pmatrix}, c \neq 0.$

$A = A^T$, т. е. соответствующий оператор \hat{A} — симметричный. Следовательно, собственные векторы, соответствующие различным собственным значениям, ортогональны. Нормируем их, т. е. требуем, чтобы $(X_k, X_k) = 1$, $k = 1, 2, 3$. Тогда: а) $2c^2 = 1$, т. е. $c = 2^{-1/2}$, б) $4c^2 = 1$, т. е. $c = 0.5$, в) $4c^2 = 1$, т. е. $c = 0.5$. Итак,

$$f_1 = \begin{pmatrix} -2^{-1/2} \\ 0 \\ 2^{-1/2} \end{pmatrix}, f_2 = \begin{pmatrix} 0.5 \\ 2^{-1/2} \\ 0.5 \end{pmatrix}, f_3 = \begin{pmatrix} 0.5 \\ -2^{-1/2} \\ 0.5 \end{pmatrix}.$$

Поэтому матрица ортогонального преобразования $X = QY$ имеет вид:

$$Q = \begin{pmatrix} -2^{-1/2} & 0.5 & 0.5 \\ 0 & 2^{-1/2} & -2^{-1/2} \\ 2^{-1/2} & 0.5 & 0.5 \end{pmatrix}.$$

Это преобразование приводит квадратичную форму к каноническому виду:

$$\phi(y^1, y^2, y^3) = \lambda_1(y^1)^2 + \lambda_2(y^2)^2 + \lambda_3(y^3)^2 = \sqrt{2}(y^2)^2 + \sqrt{2}(y^3)^2.$$

Пример 2. $\phi(x^1, x^2, x^3) = 2x^1x^2 + 2x^1x^3 + 2x^2x^3$. $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, $\det(A - \lambda E) =$

$0 \Rightarrow \lambda_{1,2} = -1, \lambda_3 = 2$. При $\lambda = -1$ система $(A - \lambda E)X_e = \theta$ принимает вид: $x^1 = -c^2 - c^3$. Поэтому базис в пространстве решений, т. е. ФСР, есть $X_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $X_2 =$

$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. Но $(x_1, x_2) \neq 0$, т. е. они не ортогональны. Поэтому применим алгоритм

ортогонализации: $e_1 = x^1/|x_1| = \begin{pmatrix} -2^{-1/2} \\ 2^{-1/2} \\ 0 \end{pmatrix}$; $e_2 = c_2/|c_2|$, где $c_2 = x_2 - (e_1, x_2)e_1 =$

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - 2^{-1/2} \cdot \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} -0.5 \\ -0.5 \\ 1 \end{pmatrix}, |c_2| = 1.5^{1/2}, e_2 = \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ \sqrt{2/3} \end{pmatrix}.$$

При $\lambda = 2$ $e_2 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$. Поэтому $X = QY$, где

$$Q = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & \sqrt{2/3} & 1/\sqrt{3} \end{pmatrix}.$$

Значит, $\phi(y^1, y^2, y^3) = -(y^1)^2 - (y^2)^2 + 2(y^3)^2$.

§5. Билинейные формы. Их связь с квадратичными формами.

Пусть в линейном пространстве R_n каждой упорядоченной паре элементов x, y по некоторому правилу B ставится в соответствие число u . Тогда $u = B(x, y)$ будем называть числовой функцией от двух аргументов x, y в линейном пространстве R_n .

О п р е д е л е н и е. Числовая функция $B(x, y)$ называется билинейной формой от x, y , если для любых элементов x, y, e из R_n выполняются условия:

1. линейность по первому аргументу при фиксированном втором:

а) $B(x + y, e) = B(x, e) + B(y, e)$,

б) для любого a из K_0 $B(ax, y) = B(x, y) \cdot a$;

2. линейность по второму аргументу при фиксированном первом:

в) $B(x, y + e) = B(x, y) + B(x, e)$,

г) для любого a из K_0 $B(x, ay) = B(x, y) \cdot a$.

Из определения билинейной формы по индукции легко получить следующее соотношение. Если x, y — линейные комбинации каких-либо элементов пространства R_n , т. е. $x = a^p \xi_p, p = \overline{1, s}; y = c^k \eta_k, k = \overline{1, m}$, то

$$B(x, y) = B(a^p \xi_p, c^k \eta_k) = a^p c^k B(\xi_p, \eta_k). \quad (9)$$

Выберем в пространстве R_n некоторый базис $(e_p)_n$. Тогда любые два элемента x, y могут быть разложены по этому базису: $x = e_p x^p, y = e_p y^p, p = \overline{1, n}$. Обозначив $B(e_p, e_k) = b_{pk}$ и применив формулу (9), получим:

$$B(x, y) = b_{pk} x^p y^k. \quad (10)$$

Представление (10) называется общим видом билинейной формы в n -мерном линейном пространстве, а матрица $B_e = \|b_{pk}\|_{n,n}$ называется матрицей билинейной формы $B(x, y)$ в базисе $(e_p)_n$.

П р и м е р ы.

1) В линейном пространстве V_3 скалярное произведение двух векторов (\vec{a}, \vec{b}) является билинейной формой, т. е. $B(\vec{a}, \vec{b}) = (\vec{a}, \vec{b})$, что следует из свойств 1-4 скалярного произведения векторов. Если в качестве базиса V_3 взять векторы $\vec{i}, \vec{j}, \vec{k}$, то $b_{pp} = 1, p = 1, 2, 3, b_{ps} = 0$ при $p \neq s$, т. е. $B_{\vec{i}, \vec{j}, \vec{k}} = \|\delta_{ps}\|_{3,3}$.

2) В пространстве $C_{[a,b]}$ билинейной формой является, например, функция $B(x, y) = \int_a^b x(c)y(c)dc$. Но так как это пространство бесконечномерно, то эта билинейная форма не имеет матрицы.

Введенную выше билинейную форму можно записать в матричной форме. Пусть в R_n задан базис $(e_p)_n$. Тогда любой элемент пространства можно разложить по

этому базису, т. е. $x = e_p x^p = e X_e$, $y = e_p y^p = e Y_e$. Поэтому равенство (10) в силу определения операции умножения матриц можно записать так:

$$B(x, y) = X_e^T B_e Y_e. \quad (11)$$

Теорема 8 Пусть B_e и B_f — матрицы билинейной формы $B(x, y)$ в базисах $(e_p)_n$ и $(f_p)_n$, соответственно. Тогда, если P — матрица невырожденного преобразования, переводящего базис $(e_p)_n$ в базис $(f_p)_n$, т. е. $f = eP$, то

$$B_f = P^T B_e P. \quad (12)$$

Доказательство. Пусть в базисе $(e_p)_n$ элементы x, y имеют координаты X_e, Y_e , а в базисе $(f_p)_n$ — X_f, Y_f соответственно. Так как $f = eP$, $\det P \neq 0$, то по формуле (6) §6 гл. 2 имеем: $X_e = P X_f$, $Y_e = P Y_f$. Тогда $B(x, y) =$ (формула (11)) $= X_e^T B_e Y_e = (P X_f)^T B_e \cdot (P Y_f) =$ (свойство 2^o операции умножения матриц) $= X_f^T (P^T B_e P) Y_f = X_f^T B_f Y_f$. Отсюда следует утверждение теоремы.

Пример. В пространстве V_3 заданы базис $\vec{e}_1 = \vec{i}, \vec{e}_2 = \vec{j}, \vec{e}_3 = \vec{k}$ и матрица $P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -3 \end{pmatrix}$ линейного преобразования, в результате которого базис $(\vec{e}_p)_3$

перейдет в базис $(\vec{f}_p)_3$ по формуле $f = eP$, т. е. $(f_1 \ f_2 \ f_3) = (\vec{e}_1 \ 2\vec{e}_2 - \vec{e}_1 \ \vec{e}_2 - 3\vec{e}_3) \Rightarrow \vec{f}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{f}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \vec{f}_3 = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}$. Найдем теперь в обоих базисах матрицу

билинейной формы, которая в базисе $(e_p)_3$ имеет вид $B(x, y) = x^1 y^1 - 2x^1 y^2 + 3x^3 y^1 - x^3 y^2$, т. е. $B_e = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 3 & -1 & 0 \end{pmatrix}$. По формуле (12) $B_f = P^T B_e P = \begin{pmatrix} 1 & -5 & -2 \\ -1 & 5 & 2 \\ -9 & 15 & 3 \end{pmatrix}$,

т. е. в этом базисе билинейная форма запишется так: $B(x, y) = x^1 y^1 - 5x^1 y^2 - x^2 y^1 + 5x^2 y^2 - 2x^1 y^3 - 9x^3 y^1 + 2x^2 y^3 + 15x^3 y^2 + 3x^3 y^3$.

Определение. Билинейная форма $B(x, y)$ называется симметричной, если для любых x, y из R_n $B(x, y) = B(y, x)$.

Теорема 9 Для того, чтобы билинейная форма $B(x, y)$ была симметричной, необходимо и достаточно, чтобы в некотором базисе $(e_p)_n$ ее матрица B_e была симметричной.

Необходимость. Пусть $B(x, y) = B(y, x)$. Тогда в базисе $(e_p)_n$ $b_{kp} = B(e_k, e_p) = B(e_p, e_k) = b_{pk}$. Что и требовалось доказать.

Достаточность. Пусть в базисе $(e_p)_n$ $B_e = B_e^T$. Тогда по формуле (10) $B(x, y) = b_{pk} x^p y^k = b_{kp} y^k x^p = B(y, x)$. Теорема доказана.

Оказывается билинейная форма связана непосредственно с ранее введенной нами квадратичной формой. А именно, если в билинейной форме положить $y = x$, то в силу (10)

$$B(x, x) = b_{pk} x^p x^k \quad (13)$$

— функция n переменных x^1, \dots, x^n , которые здесь интерпретируются как координаты элемента x пространства R_n в базисе $(e_p)_n$. В записи (13) есть подобные члены $b_{pk} x^p x^k$ и $b_{kp} x^k x^p$. Введем симметричную матрицу A : $a_{pk} = a_{kp} = (b_{pk} + b_{kp})/2$, $a_{kk} = b_{kk}$. Тогда $B(x, x) = A(x, x) = a_{pk} x^p x^k$ — квадратичная форма. Отсюда следует, что одну и ту же квадратичную форму можно получить из различных билинейных форм $B(x, y) = X^T B Y$ и $B^*(x, y) = X^T B^* Y$, если только $b_{pk} + b_{kp} = b_{pk}^* + b_{kp}^*$, $k, p = \overline{1, n}$.

Вывод: каждой билинейной форме соответствует единственная квадратичная форма; каждой квадратичной форме соответствует бесконечно много билинейных форм, но среди них есть единственная симметричная билинейная форма.

Пример. Если дана квадратичная форма $\phi(x^1, x^2, x^3) = (x^1)^2 - 2x^1x^2 + 3x^2x^3$, то соответствующая ей симметричная билинейная форма имеет вид $B(x, y) = A(x, y) = x^1y^1 - x^1y^2 - x^2y^1 + 1.5x^2y^3 + 1.5x^3y^2$.

Любая симметричная билинейная форма $A(x, y)$ может быть выражена через квадратичные формы. Действительно, $A(x + y, x + y)$ (по свойству а) билинейной формы) = $A(x, x + y) + A(y, x + y)$ (свойство б) билинейной формы) = $A(x, x) + A(x, y) + A(y, x) + A(y, y)$. Так как билинейная форма симметрична, то $A(x, y) = A(y, x)$. Следовательно,

$$A(x, y) = 0.5(A(x + y, x + y) - A(x, x) - A(y, y)).$$

Определение. Базис $(e_p)_n$ в линейном пространстве R_n называется каноническим для билинейной формы $B(x, y)$, если $B(e_p, e_k) = 0, p \neq k, p, k = \overline{1, n}$.

Теорема 9* Для того, чтобы билинейная форма $B(x, y)$ была симметричной, необходимо и достаточно, чтобы эта форма имела канонический базис.

Необходимость. Пусть форма $B(x, y)$ — симметричная билинейная форма $A(x, y)$. Пусть в базисе $(e_p)_n$ $A(x, y) = X_e^T A_e Y_e$. Соответствующая этой билинейной форме квадратичная форма $A(x, x) = X_e^T A_e X_e$ по теореме 5 невырожденным преобразованием $X = PZ$ приводится к каноническому виду, т. е. матрица $P^T A_e P$ имеет диагональный вид. Этому преобразованию соответствует преобразование базиса $(e_p)_n$ в базис $(f_p)_n$ по формуле $f = eP$. Тогда, как следует из формулы (12), $A_f = P^T A_e P$, которая по выше доказанному является диагональной. Следовательно, в этом базисе $A(x, y) = a_{kk}x^k y^k$, т. е. $a_{kp} = A(f_k, f_p) = 0$ при $k \neq p$. Но это и означает, что базис $(f_p)_n$ — канонический.

Достаточность. Пусть $(e_p)_n$ — канонический базис. Тогда матрица B_e билинейной формы в этом базисе имеет диагональный вид, так как $b_{kp} = B(e_k, e_p) = 0$ при $k \neq p$. Но диагональная матрица совпадает со своей транспонированной, т. е. $B_e = B_e^T$. А это означает, что билинейная форма $B(x, y)$ является симметричной.

§6. Метод Якоби приведения квадратичной формы к каноническому виду.

Пусть дана квадратичная форма $X^T A X$. Как отмечалось в §5, ее можно рассматривать как частный случай билинейной формы, т. е. $X^T A X = A(x, x)$, где x^1, \dots, x^n интерпретируются как координаты элемента x пространства R_n в базисе $(e_p)_n$. Поэтому проблему приведения квадратичной формы к каноническому виду можно рассматривать как проблему выбора канонического базиса. Сделаем это с помощью треугольного преобразования базисных элементов:

$$\begin{aligned} f_1 &= e_1 \\ f_2 &= c_{21}e_1 + e_2 \\ f_3 &= c_{31}e_1 + c_{32}e_2 + e_3 \\ &\dots\dots\dots \\ f_n &= c_{n1}e_1 + c_{n2}e_2 + c_{n3}e_3 + \dots + e_n \end{aligned} \tag{14}$$

Так как определитель матрицы преобразования отличен от нуля (равен 1), то f_1, \dots, f_n образуют базис в R_n . Далее, нам потребуются угловые миноры матрицы $A = A_e$ квадратичной формы $A(x, x)$ в базисе $(e_p)_n$: $\Delta_k = \det \|a_{pk}\|_{k,k}, k = \overline{1, n-1}$.

§7. Закон инерции квадратичной формы.

Как доказано в теореме 4, число ненулевых коэффициентов в каноническом виде квадратичной формы постоянно, т. е. является инвариантом квадратичной формы относительно преобразований, приводящих квадратичную форму к каноническому виду. Оказывается, что число положительных и отрицательных коэффициентов в канонической форме также постоянно, независимо от способа приведения невырожденным преобразованием к каноническому виду.

Теорема 11 (об инерции квадратичной формы) *Число положительных и отрицательных коэффициентов в каноническом виде, к которому приводится квадратичная форма невырожденным преобразованием, не зависит от выбора этого преобразования.*

Доказательство. Пусть квадратичная форма $\phi(x^1, \dots, x^n)$ ранга r невырожденным преобразованием $Y = PX$ приведена к каноническому виду:

$$a_1(y^1)^2 + \dots + a_k(y^k)^2 - a_{k+1}(y^{k+1})^2 - \dots - a_r(y^r)^2 \quad (a_p > 0, p = \overline{1, r}), \quad (20)$$

а невырожденным преобразованием $Z = CX$ — к виду:

$$b_1(z^1)^2 + \dots + b_e(z^e)^2 - b_{e+1}(z^{e+1})^2 - \dots - b_r(z^r)^2 \quad (b_p > 0, p = \overline{1, r}). \quad (21)$$

Так как $\det P \neq 0$, $\det C \neq 0$, то задав значения x^1, \dots, x^n , однозначно из $Y = PX$ и $Z = CX$ определим значения y^1, \dots, y^n и z^1, \dots, z^n . Тем самым формулы (20), (21) определяют значение квадратичной формы в точке x^1, \dots, x^n , т. е. их можно приравнять:

$$a_1(y^1)^2 + \dots + a_k(y^k)^2 - a_{k+1}(y^{k+1})^2 - \dots - a_r(y^r)^2 = b_1(z^1)^2 + \dots - b_r(z^r)^2. \quad (22)$$

Надо доказать, что $k = e$. Предположим противное, т. е. $k > e$. Как уже говорилось, в равенстве (22) переменные y^1, \dots, y^n и z^1, \dots, z^n связаны с переменными x^1, \dots, x^n соотношениями:

$$\begin{aligned} Y = PX &\Rightarrow \begin{cases} y^1 = p_{1i}x^i, \\ \dots, \\ y^k = p_{ki}x^i, \\ y^{k+1} = p_{k+1,i}x^i, \\ \dots, \\ y^n = p_{ni}x^i, \end{cases} \\ Z = CX &\Rightarrow \begin{cases} z^1 = c_{1i}x^i, \\ \dots, \\ z^e = c_{ei}x^i, \\ z^{e+1} = c_{e+1,i}x^i, \\ \dots, \\ z^n = c_{ni}x^i, \\ (i = \overline{1, n}). \end{cases} \end{aligned} \quad (23)$$

Так как $k > e$, то число уравнений в системе (23) $n - k + e < n$. Найдем из этой системы решение x^1, \dots, x^n , для которого

$$y^{k+1} = \dots = y^n = 0, \quad (24)$$

$$z^1 = \dots = z^e = 0. \quad (25)$$

В силу этих условий система (23) — однородная система уравнений, причем она нетривиально совместна, так как число уравнений меньше числа неизвестных. Поэтому существуют числа x^1, \dots, x^n , не все равные нулю, которые удовлетворяют системе (23) при условиях (24), (25). Теперь для этих значений x^1, \dots, x^n из равенств $Y = PX$ и $Z = CX$ найдем y^1, \dots, y^k и z^{e+1}, \dots, z^n и подставим в равенство (22), учитывая (24), (25):

$$a_1(y^1)^2 + \dots + a_k(y^k)^2 = -b_{e+1}(z^{e+1})^2 - \dots - b_r(z^r)^2.$$

Но это равенство возможно, поскольку $a_p > 0$, $b_p > 0$, $p = \overline{1, r}$, тогда и только тогда, когда $y^1 = \dots = y^k = 0$ и $z^{e+1} = \dots = z^r = 0$. Но тогда, учитывая (24), например, получаем, что $Y = \theta$. Поэтому и так как $\det P \neq 0$ из $Y = PX$ получаем, что $X = \theta$. Пришли к противоречию с предположением, что $X \neq \theta$, означающее, что предположение $k > e$ неверно. Аналогично доказывается, что $k < e$ быть не может. Таким образом, остается лишь один случай: $k = e$. Что и требовалось доказать.

§8. Классификация квадратичных форм.

О п р е д е л е н и е. *Квадратичная форма X^TAX называется*

- 1) *положительно (отрицательно) определенной, если для всех X $X^TAX \geq 0$ (≤ 0), причем $X^TAX = 0$ только при $X = \theta$ (такие формы называются также знакоопределенными);*
- 2) *квазиопределенной, если она может принимать значения $X^TAX \geq 0$ (≤ 0) для $X \neq \theta$.*
- 3) *неопределенной, если она может принимать как положительные, так и отрицательные значения.*

Теорема 12 *Для того, чтобы квадратичная форма была положительно (отрицательно) определенной, необходимо и достаточно, чтобы все коэффициенты в каноническом виде этой формы были строго положительны (отрицательны).*

Доказательство. **Необходимость.** Пусть $X^TAX > 0$ всюду, кроме $X = \theta$. Пусть невырожденным преобразованием $Y = PX$ эта форма приведена к каноническому виду

$$a_k(y^k)^2, \quad k = \overline{1, n}. \quad (26)$$

Требуется доказать, что $a_k > 0$, $k = \overline{1, n}$. Зафиксируем k и положим $y^1 = \dots = y^{k-1} = y^{k+1} = \dots = y^n = 0$, $y^k = 1$. Тогда из системы $Y = PX$ определим нетривиальное решение X , для которого

$$0 < X^TAX = Y^TBY = \phi(0, \dots, 0, 1, 0, \dots, 0) = a_k \Rightarrow a_k > 0.$$

Д о с т а т о ч н о с т ь. Пусть в квадратичной форме (26) $a_k > 0$, $k = \overline{1, n}$. Требуется доказать: $X^TAX > 0$ при $X \neq \theta$ и $X^TAX = 0$ при $X = \theta$. Пусть $X \neq \theta$. Тогда из системы $Y = PX$ определяем $Y \neq \theta$. Поэтому $X^TAX = Y^TBY = a_k(y^k)^2 > 0$, так как $a_k > 0$ и не все y^k равны нулю. Если же $X = \theta$, то из системы $Y = PX$ имеем $Y = P\theta = \theta$. Следовательно, $X^TAX = Y^TBY = a_1 \cdot 0 + \dots + a_n \cdot 0 = 0$.

С л е д с т в и е. *Если квадратичная форма положительно определена, то невырожденным преобразованием ее можно привести к виду: $(y^1)^2 + (y^2)^2 + \dots + (y^n)^2$.*

Теорема 13 (Критерий Сильвестра) *Для того, чтобы квадратичная форма X^TAX была положительно определена, необходимо и достаточно, чтобы угловые миноры матрицы A $\Delta_k > 0$, $k = \overline{1, n}$. Для того, чтобы эта квадратичная форма была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров чередовались, причем $\Delta_1 < 0$.*

Необходимость. Пусть квадратичная форма $X^TAX = A(x, x)$ положительно определена. Докажем, что $\Delta_k \neq 0$, $k = \overline{1, n}$. Предположим противное, т. е. для некоторого k $\Delta_k = 0$. Рассмотрим следующую квадратную однородную систему линейных уравнений:

$$a_{ep}x^p = 0, \quad e, p = \overline{1, k}. \quad (27)$$

Так как Δ_k — определитель этой системы и $\Delta_k = 0$, то система (27) имеет ненулевое решение x^1, \dots, x^k (не все x^e равны нулю). Умножим первое из уравнений системы (27) на x^1 , второе — на x^2 , ..., последнее — на x^k и сложим полученные соотношения. В итоге получим равенство: $a_{ep}x^e x^p = 0$, $e, p = \overline{1, k}$, левая часть которого представляет собой значение квадратичной формы $A(x, x)$ для ненулевого вектора x с координатами $x^1, \dots, x^k, 0, \dots, 0$. Это значение равно нулю, что противоречит знакоопределенности формы. Итак, $\Delta_k \neq 0$, $k = \overline{1, n}$. Поэтому можно применить метод Якоби приведения формы $A(x, x)$ к сумме квадратов и воспользоваться формулами (19) для канонических коэффициентов λ_k . Если $A(x, x)$ — положительно определенная форма, то все канонические коэффициенты положительны. Но тогда из равенств (19) следует, что $\Delta_p > 0$, $p = \overline{1, n}$. Если же $A(x, x)$ — отрицательно определенная форма, то все канонические коэффициенты отрицательны. Но тогда из формул (19) следует, что знаки угловых миноров чередуются, причем $\Delta_1 < 0$.

Достаточность. Пусть выполнены условия, наложенные на угловые миноры в формулировке теоремы. Тогда форму $A(x, x)$ можно привести к сумме квадратов методом Якоби, причем канонические коэффициенты могут быть найдены по формулам (19). Если все угловые миноры положительны, то из (19) следует, что все $\lambda_p > 0$, т. е. $A(x, x)$ — положительно определенная. Если же знаки Δ_p чередуются и $\Delta_1 < 0$, то из соотношений (19) следует, что $A(x, x)$ отрицательно определенная. Теорема доказана.

Задачи. Доказать:

- 1) Если A — матрица положительно определенной квадратичной формы, то A^{-1} — матрица положительно определенной квадратичной формы;
- 2) для всякой матрицы C , $\det C \neq 0$, $C^T C$ — матрица положительно определенной квадратичной формы;
- 3) если A — матрица положительно определенной квадратичной формы, то все элементы главной диагонали больше нуля;
- 4) если квадратичная форма неотрицательная и $a_{kk} = 0$, то все элементы k -го и k -ой строки матрицы этой квадратичной формы равны нулю.

Решение.

- 1) Если A — матрица положительно определенной квадратичной формы, то после преобразования (3) матрица $P^T A P$ — матрица все той же положительно определенной квадратичной формы, если $\det P \neq 0$. Пусть $P = A^{-1}$. Тогда $P^T A P = (A^{-1})^T A A^{-1} = (A^{-1})^T = (A^T)^{-1} = A^{-1}$ — матрица, как уже отмечалось выше, положительно определенной квадратичной формы.
- 2) Так как E — матрица положительно определенной формы, то при $\det C \neq 0$ $C^T E C = C^T C$ — матрица положительно определенной квадратичной формы.
- 3) Пусть, например, $a_{22} = 0$. Положим $x^1 = x^3 = \dots = x^n = 0$, $x^2 = 1$. Тогда $\phi(x^1, \dots, x^n) = a_{22}(x^2)^2 = a_{22} = 0$ при не всех x^k равных 0, т. е. ϕ не положительно определенная квадратичная форма.
- 4) Пусть, например, $a_{33} = 0$, но $a_{13} \neq 0$. Положим $x^2 = x^4 = \dots = x^n = 0$, $x^1 = a \neq 0$, $x^3 = b \neq 0$. Тогда $\phi(x^1, \dots, x^n) = a_{11}a^2 + 2a_{13}ab$. В силу 3) $a_{11} \geq 0$. Пусть

$a_{11} = 1$. Тогда $\phi = a(a + 2a_{13}b)$. Можно подобрать a, b так, что $a(a + 2a_{13}b) < 0$, что противоречит условию, т. е. $a_{13} = 0$ и тогда $\phi = a^2 \geq 0$.

§9. Численные методы решения систем линейных уравнений.

Рассмотрим систему уравнений:

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

где $A = \|a_k^p\|_n^n$, $X = \|x^p\|_n$, $B = \|b^p\|_n$. Пусть $\det A \neq 0$, т. е. система (1) имеет единственное решение.

п. 1. Итерационный метод Гаусса-Зейделя.

Если $a_p^p \neq 0$, $p = \overline{1, n}$, то системе (1) можно придать вид:

$$x^p = \frac{b^p - a_k^p x^k}{a_p^p}, \quad k = \overline{1, n}, \quad k \neq p, \quad p = \overline{1, n}. \quad (2)$$

Система (2) называется приведенной. Введя обозначения

$$m^p = \frac{b^p}{a_p^p} \quad \text{и} \quad c_k^p = \begin{cases} -a_k^p/a_p^p & \text{при } p \neq k \\ 0 & \text{при } p = k. \end{cases}$$

систему (2) запишем так:

$$X = M + CX, \quad (3)$$

где $C = \|c_k^p\|_n^n$, $M = \|m^p\|_n$. Система

$$X_k = M + CX_{k-1}, \quad k = 1, 2, 3, \dots, \quad (4)$$

называется итерационной системой уравнений для системы (3). Взяв в качестве нулевого приближения X_0 — любой числовой столбец высоты n и подставив его в рекуррентную формулу (4), получим $X_1 = M + CX_0$ и т. д. Метод итераций состоит в замене точного решения системы (1) k -ой итерацией (4) с достаточно большим номером k . Если при $k \rightarrow \infty$ $X_k \rightarrow \bar{X}$, то \bar{X} — искомое решение системы (3), обращающее эту систему уравнений в верные числовые равенства:

$$\bar{X} = M + c\bar{X}. \quad (5)$$

Будет ли схема сходящейся при $k \rightarrow \infty$? Ответ дают следующие достаточные условия сходимости метода:

$$\text{а) } |a_p^p| > \sum_{\substack{k=1 \\ k \neq p}}^n |a_k^p|, \quad p = \overline{1, n}, \quad \text{или} \quad \text{б) } |a_k^k| > \sum_{\substack{p=1 \\ p \neq k}}^n |a_k^p|, \quad k = \overline{1, n}.$$

п. 2. Метод Зейделя.

Представим матрицу C системы (3) в виде $C = B + D$, где

$$B = \begin{pmatrix} 0 & 0 & \dots & 0 \\ c_1^1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ c_n^1 & c_n^2 & \dots & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & \dots & c_1^{n-1} & c_1^n \\ 0 & 0 & \dots & c_2^{n-1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Тогда (3) есть система:

$$X = M + BX + DX, \quad (6)$$

для которой итерационная система есть система: $X_k = M + BX_k + DX_{k-1}$ или

$$(E - B)X_k = M + DX_{k-1}. \quad (7)$$

Пусть X_0 — любой числовой столбец высоты n . Тогда

$$(E - B)X_1 = T,$$

где $T = M + DX_0$ — известный нам столбец. Распишем последнюю систему:

$$\begin{cases} x^1 = t^1 \\ -c_1^2 x^1 + x^2 = t^2 \\ \dots\dots\dots \\ -c_1^n x^1 - c_2^n x^2 - \dots - c_{n-1}^n x^{n-1} + x^n = t^n \end{cases}$$

— треугольная система, которую легко решать сверху:

$$\begin{aligned} x^1 &= t^1 \\ x^2 &= t^2 + c_1^2 t^1 \\ x^3 &= t^3 + c_2^3 (t^2 + c_1^2 t^1) + c_1^3 t^1 \\ &\dots\dots\dots \end{aligned}$$

Метод сходится, если матрица A : 1) симметричная и 2) положительная, т. е. все решения уравнения $\det(A - \lambda E) = 0$ неотрицательны.

п. 3. Метод исключения (Гаусса).

Рассмотрим один из наиболее известных и широко применяемых прямых методов решения систем линейных уравнений. Применим его сначала к системе из трех уравнений с тремя неизвестными:

$$a_k^p x^k = b^p, \quad p, k = 1, 2, 3. \quad (8)$$

В такой системе по крайней мере один из коэффициентов a_3^1, a_3^2, a_3^3 должен быть отличен от нуля, иначе мы имели бы дело в этих трех уравнениях с двумя неизвестными. Если $a_3^3 = 0$, то можно переставить уравнения так, чтобы коэффициент при x^3 в третьем уравнении был отличен от нуля. Очевидно, что перестановка уравнений оставляет систему неизменной, и ее решение прежним. Теперь умножим третье уравнение вначале на a_3^1/a_3^3 и вычтем из первого уравнения, а затем — на a_3^2/a_3^3 и вычтем из второго уравнения. В итоге будем иметь:

$$\begin{cases} (a_1^1 - a_1^3 a_3^1 / a_3^3) x^1 + (a_2^1 - a_2^3 a_3^1 / a_3^3) x^2 = b^1 - b^3 a_3^1 / a_3^3 \\ (a_1^2 - a_1^3 a_3^2 / a_3^3) x^1 + (a_2^2 - a_2^3 a_3^2 / a_3^3) x^2 = b^2 - b^3 a_3^2 / a_3^3 \\ a_1^3 x^1 + a_2^3 x^2 + a_3^3 x^3 = b^3, \end{cases} \quad (9)$$

или, если переобозначить коэффициенты в первых двух уравнениях,

$$\begin{cases} \bar{a}_1^1 x^1 + \bar{a}_2^1 x^2 = \bar{b}^1 \\ \bar{a}_1^2 x^1 + \bar{a}_2^2 x^2 = \bar{b}^2 \\ a_1^3 x^1 + a_2^3 x^2 + a_3^3 x^3 = b^3. \end{cases} \quad (10)$$

Если $\bar{a}_2^2 = 0$, то переставим уравнения так, чтобы этот коэффициент стал отличным от нуля. После этого умножим второе уравнение на \bar{a}_2^1/\bar{a}_2^2 и вычтем из первого. Тогда

$$\begin{cases} \bar{a}_1^1 x^1 = \bar{b}^1 \\ \bar{a}_2^2 x^2 = \bar{b}^2 - \bar{a}_1^2 x^1 \\ a_3^3 x^3 = b^3 - a_1^3 x^1 - a_2^3 x^2, \end{cases} \quad (11)$$

где $\bar{a}_1^1 = \bar{a}_1^1 - \bar{a}_1^2 \bar{a}_2^1 / \bar{a}_2^2$, $\bar{b}^1 = \bar{b}^1 - \bar{b}^2 \bar{a}_2^1 / \bar{a}_2^2$. Если $\bar{a}_1^1 = 0$, то система вырождена. Будем считать, что $\bar{a}_1^1 \neq 0$. Тогда из системы (11), начиная с первого уравнения, легко определяются x^1, x^2, x^3 , т. е. решение исходной системы (8).

Теперь рассмотрим общий случай системы (1). Пусть $a_n^n \neq 0$. Вычтем из каждого p -го уравнения ($p = \overline{1, n-1}$) n -ое уравнение, умноженное на a_n^p/a_n^n . Введем обозначения:

$$\bar{a}_k^p = a_k^p - a_k^n a_n^p / a_n^n, \quad \bar{b}^p = b^p - b^n a_n^p / a_n^n, \quad p = \overline{1, n-1}, \quad k = \overline{1, n}.$$

Тогда система (1) запишется так:

$$\begin{cases} \bar{a}_k^p x^k = \bar{b}^p, & p, k = \overline{1, n-1} \\ a_i^n x^i = b^n, & i = \overline{1, n}. \end{cases} \quad (12)$$

Продолжая в том же духе, мы можем исключить x^{n-1} из первых $n-2$ уравнений, затем x^{n-2} из первых $n-3$ уравнений и т. д., пока не придем к треугольной системе уравнений:

$$\begin{cases} a_{11}^{(n-1)} x_1 = b_1^{(n-1)} \\ a_{21}^{(n-2)} x_1 + a_{22}^{(n-2)} x_2 = b_2^{(n-2)} \\ \dots\dots\dots \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n = b_n, \end{cases} \quad (13)$$

где ради удобства записи верхние индексы элементов матриц заменены на первые нижние. Из этой системы, начиная с первого уравнения, легко определяются неизвестные.

Литература.

Список литературы

1. В.А. Ильин, Э.Г. Поздняк, "Линейная алгебра".
2. Г.Е. Шилов, "Математический анализ. Конечномерные линейные пространства".
3. А.И. Мальцев, "Основы линейной алгебры".

Заключение.

Для получения положительной оценки на экзамене необходимо уметь решать следующие задачи:

Глава 2. Уметь находить

- 1) базис и размерность пространства,
- 2) координаты элемента x в заданном базисе,
- 3) переход от одного базиса к другому, матрицу этого перехода,
- 4) ранг матрицы.

Является ли

- 5) данное множество M подпространством линейного пространства?
- 6) данное множество элементов линейно зависимым или нет?
- 7) Образовать линейную оболочку из данной совокупности элементов. Определить ее размерность.

Глава 3. Уметь находить

- 1) ФСР,
- 2) общее решение неоднородной системы уравнений.
- 3) Является ли данная система совместной?

Глава 4. Уметь

- 1) ортогонализировать базис,
- 2) дополнять до ортогонального базиса,
- 3) доказывать, что в каждом линейном конечномерном пространстве можно ввести скалярное произведение.
- 4) Знать 4 свойства ортогональной матрицы и уметь ими пользоваться.

Глава 5. Уметь

- 1) выписывать матрицу линейного оператора. Например, для оператора проектирования в пространстве V_3 на прямую $x = y = z$.
- 2) Как изменится матрица оператора при изменении базиса?
- 3) находить собственные значения и собственные векторы линейного оператора.

Глава 6. Уметь

- 1) выписывать матрицу квадратичной формы и билинейной формы и наоборот,
- 2) приводить квадратичную к каноническому виду методами а) Якоби, б) Лагранжа, в) ортогональных преобразований.
- 3) Является ли данная квадратичная форма положительно определенной?



ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА