Матрицы, замены координат и линейные отображения

Замечание 1. В предыдущей лекции содержался небольшой обман: условие $1v = v$ не следует из остальных. Например, абелева группа $\mathbb{R}^2$ с операцией умножения на число $\lambda(x, y) = (\lambda x, 0)$ прекрасно подходит. Так что это условие надо накладывать дополнительно.

Замечание 2. Если использовать и верхние, и нижние индексы, то стандартная конвенция в физике состоит в том, что числа (коэффициенты, координаты и пр.) имеют верхние индексы, а векторы — нижние индексы, причём по повторяющимся сверху и снизу индексам идёт суммирование. Так, физики вместо записи $\sum_{i=1}^n \lambda_i v_i$ пишут $\lambda^i v_i$, считая, что в каждый момент все знают, сколько именно векторов (чему равно n) и что λ^i — это не возведение в степень ...

Пусть в векторном пространстве V есть два базиса: $e_1, \ldots, e_n$ и $e'_1, \ldots, e'_n$. Тогда базис $\{e'_i\}$ («новый») можно выразить через базис $\{e_j\}$ («старый»):

$$\begin{aligned} e'_1 &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n, \\ e'_2 &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e'_n &= a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n. \end{aligned} \tag{1}$$

Иначе говоря, $e'_i = \sum_{j=1}^n a_{ji}e_j$.

Как преобразуются координаты в этом случае? Если вектор v имеет координаты $x'_1, \ldots, x'_n$ в новом базисе, т.е. $v = \sum_{i=1}^n x'_i e'_i$, то

$$v = \sum_{i=1}^n x'_i \Big(\sum_{j=1}^n a_{ji} e_j \Big) = \sum_{j=1}^n \Big(\sum_{i=1}^n a_{ji} x'_i \Big) e_j,$$

или же

$$\begin{aligned} x_1 &= a_{11}x'_1 + a_{12}x'_2 + \ldots + a_{1n}x'_n, \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + \ldots + a_{2n}x'_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= a_{n1}x'_1 + a_{n2}x'_2 + \ldots + a_{nn}x'_n. \end{aligned} \tag{2}$$

Видно, что если известно выражение нового базиса через старый, то легко выразить координаты в старом базисе через координаты в новом: закон определяется в каждом из случаев таблицей

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n,1} & \ldots & \ldots & a_{n,n} \end{pmatrix},$$

которую в одном из случаев, впрочем, надо отразить относительно диагонали a_{11} — a_{nn} («главной диагонали»). Эта таблица называется *матрицей перехода от базиса* $\{e_n\}$ *к базису* $\{e'_n\}$.

Замечание 3. Операция отражения относительно главной диагонали называется *транспонированием*, а отражённая матрица называется транспонированной и обозначается ${}^tA$. Таким образом,

$${}^tA = \begin{pmatrix} a_{1,1} & a_{2,1} & \ldots & a_{n,1} \\ a_{1,2} & a_{2,2} & \ldots & a_{n,2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1,n} & \ldots & \ldots & a_{n,n} \end{pmatrix}.$$

На самом деле, немного более резонно называть матрицу ${}^tA$ матрицей перехода от базиса $\{e_i\}$ к базису $\{e'_i\}$, а матрицу A называть матрицей преобразования координат. Однако мы следуем традиции обозначений в этом случае.

Предположим, что есть три базиса $\{e_i\}$, $\{e'_j\}$, $\{e''_k\}$ и матрицы преобразования координат при переходе от второго к первому (A), от третьего ко второму (B) и от третьего к первому (C). Как связаны эти матрицы?

$$x_i = \sum_{k=1}^{n} a_{ik} x'_k = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{n} b_{kj} x''_j.$$

С другой стороны, $x_i = \sum_{j=1}^{n} c_{ij} x''_j$. Приравнивая коэффициенты, получаем $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Определение 1. Эта формула задаёт на множестве матриц бинарную операцию, называемую *умножением матриц*. Обозначение:

$$\begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,n} \\ c_{2,1} & c_{2,2} & \dots & c_{2,n} \\ \dots & \dots & \dots & \dots \\ c_{n,1} & \dots & \dots & c_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \dots & \dots & \dots & \dots \\ b_{n,1} & \dots & \dots & b_{n,n} \end{pmatrix}.$$

Легко видеть, что по такому закону можно перемножить не только две матрицы $n \times n$, но и матрицу $n \times m$ и матрицу $m \times k$ (и получится матрица $n \times k$). Это правило коротко формулируют «строка на столбец», т. е. чтобы вычислить ij-ый элемент произведения двух матриц, надо приставить i-ую строку первой матрицы к j-ому столбцу второй, покоординатно перемножить и сложить результаты. В терминах умножения матриц можно равенства (1) и (2) переписать в виде

$$(e'_1, e'_2, \dots, e'_n) = (e_1, e_2, \dots, e_n) \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & \dots & a_{n,n} \end{pmatrix},$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Замечание 4. Проще говоря, для нахождения матрицы перехода к новому базису нужно записать координаты новых базисных векторов в старом базисе по столбцам в матрицу.

Матрица

$$E_n = \mathrm{id}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 1 \end{pmatrix},$$

которая получится, если выражать базис через тот же самый базис, называется *единичной матрицей*. Это название оправдано, поскольку для каждой $n \times n$-матрицы A верны равенства $AE_n = E_n A = A$. При этом раз каждый базис можно разложить по другому базису, то для каждой матрицы A перехода к другому базису существует *обратная матрица* $B = A^{-1}$, для которой $AB = BA = E_n$. Уравнение $AB = E_n$ (на матрицу B) разрешимо не для любой матрицы A. Например, если все элементы матрицы A равны нулю, то решения нет. Это уравнение — система линейных уравнений на элементы матрицы B.

Рассмотрим задачу описания решений системы m линейных уравнений с n неизвестными:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{3}$$

или в «матричном» виде $Ax = b$, где

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & \dots & \dots & a_{m,n} \end{pmatrix}, \ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Для решения линейных уравнений используют *метод Гаусса*: перенумерацией неизвестных и уравнений можно добиться, чтобы в первом уравнении коэффициент при x_1 не равен нулю, после этого вычтем первое уравнение из остальных с нужными коэффициентами так, чтобы в результате x_1 не входило в остальные уравнения, и будем продолжать в том же духе. В результате матрица системы станет верхнетреугольной, т. е. в каждой строке либо все коэффициенты будут нулевыми, либо начальный отрезок нулевых коэффициентов будет больше, чем в предыдущей строке.

Можно переформулировать задачу: мы хотим узнать, является ли вектор b линейной комбинацией столбцов матрицы A (т. е. лежит ли он в их линейной оболочке). Получаем следующее (очевидное) утверждение:

Теорема 1. (Теорема Кронекера–Капелли) Система (3) имеет решение (совместна), если и только если размерность линейной оболочки столбцов матрицы A совпадает с размерностью линейной оболочки системы векторов, состоящей из столбцов матрицы A и вектора b.

Сколько может быть решений у системы линейных уравнений? Если столбцы матрицы линейно независимы, то решение, когда оно существует, единственно. В самом деле, разность двух решений — решение *однородной системы*, т. е. системы, для которой вектор b нулевой. А у однородной системы в этом случае единственное решение по определению линейной независимости. Пусть столбцы линейно зависимы. Рассмотрим случай, когда уравнений столько, сколько и неизвестных. Ясно, что если решение есть, то решений бесконечно много, поскольку можно прибавить любое решение однородной системы. Но возможен и случай, когда решения нет, поскольку столбцы линейно зависимы и их линейная оболочка не совпадает со всем пространством. Этим доказана

Теорема 2. (Альтернатива Фредгольма) Рассмотрим систему из n линейных уравнений с n неизвестными. Тогда имеет место в точности одна из двух возможностей:

— при любом векторе b частях система имеет единственное решение,
— для некоторых векторов b система имеет бесконечно много решений, а для некоторых — несовместна.

Наконец, можно воспринимать линейные системы как описание ядер *линейных отображений.*

Определение 2. Пусть V, W — векторные пространства над полем K. Отображение $A\colon V \to W$ называется *линейным отображением из пространства V в пространство W*, если выполнены свойства:

- $\forall v_1, v_2 \in V\ A(v_1 + v_2) = A(v_1) + A(v_2)$;
- $\forall v \in V, \lambda \in K\ A(\lambda v) = \lambda A(v)$.

Пример 1. Линейное отображение из пространства V в одномерное пространство — это линейный функционал на пространстве V.

Упражнение 1. Докажите, что все линейные отображения из пространства V в пространство W образуют векторное пространство. Выразите его размерность через размерности пространств V и W.

Линейное отображение достаточно знать на базисе: для вектора $v = \sum_{i=1}^{n} x_i e_i$ верно равенство $A(v) = \sum_{i=1}^{n} x_i A(e_i)$. Линейные отображения часто называют также *линейными операторами*, особенно, когда они действуют из пространства в то же самое пространство.

Определение 3. Если фиксировать базисы в пространствах V, W, то каждому вектору отвечает столбец его координат, например,

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

В этой ситуации оператору A отвечает матрица размера $m \times n$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & \dots & \dots & a_{m,n} \end{pmatrix},$$

называемая *матрицей оператора A*, для которой

$$A(v) = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & \dots & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Тогда вопрос о разрешимости системы линейных уравнений равносилен вопросу о том, принадлежит ли вектор b образу линейного оператора из K^n в K^m, задаваемого матрицей A (предполагается, что в пространствах K^n и K^m выбраны стандартные базисы из векторов e_i: единственная ненулевая координата вектора e_i имеет номер i и равна единице).

Замечание 5. Обратите внимание, что оператор можно задать и в отсутствие выбора базиса, а матрица оператора имеет смысл только если базисы даны. Например, оператор отражения переменной на пространстве многочленов, переводящий многочлен $f(x)$ в многочлен $f(-x)$, задан безотносительно конкретного выбора базиса. При этом матрица этого оператора существенно зависит от базиса в пространстве многочленов.

Замечание 6. Для того, чтобы выписать матрицу оператора, нужно разложить образы базисных векторов пространства V по базису пространства W и записать их координаты в столбцы матрицы.

Теорема 3. Пусть даны линейные операторы $A\colon V_1 \to V_2$, $B\colon V_2 \to V_3$. Рассмотрим линейный оператор $B \circ A\colon V_1 \to V_3$. Тогда для любого выбора базисов $\{e_i\}, \{e'_j\}, \{e''_k\}$ в пространствах V_1, V_2, V_3 соответственно матрица оператора $B \circ A$ равна произведению матриц операторов B и A.

Действительно,

$$B \circ A \left(\sum_{i=1}^{n_1} x_i e_i \right) = B \left(\sum_{k=1}^{n_2} \left(\sum_{i=1}^{n_1} a_{ki} x_i \right) e'_k \right) = \sum_{j=1}^{n_3} \left(\sum_{i=1}^{n_1} \left(\sum_{k=1}^{n_2} b_{jk} a_{ki} \right) x_i \right) e''_j.$$

Теорема 4. Пусть дан линейный оператор $A\colon V_1 \to V_2$, и пусть известна его матрица $A_{e,f}$ для базисов $e_1, \ldots, e_n$, $f_1, \ldots, f_m$ этих пространств. Пусть $e'_1, \ldots, e'_n$, $f'_1, \ldots, f'_m$ — другие базисы. Обозначим через B, C матрицы перехода от базисов $\{e_i\}$, $\{f_j\}$ к базисам $\{e'_i\}$, $\{f'_j\}$ соответственно. Тогда матрицу $A_{e',f'}$ этого оператора для новых базисов можно вычислить по формуле $A_{e',f'} = C^{-1} A_{e,f} B$.

Действительно, столбец вектора в старом базисе получается умножением матрицы перехода на столбец координат в новом базисе. Положим $w = A(v)$. Если для столбцов координат в старом базисе $w_{\text{ст}} = A_{e,f} v_{\text{ст}}$, то $C w_{\text{нов}} = A_{e,f} B v_{\text{нов}}$, откуда $A_{e',f'} v_{\text{нов}} = w_{\text{нов}} = C^{-1} A_{e,f} B v_{\text{нов}}$, что и требовалось.

Следствие. Пусть C — матрица перехода к новому базису пространства V, A — матрица в старом базисе линейного оператора на этом пространстве. Тогда матрица этого оператора в новом базисе есть $C^{-1} A C$.

В конце прошлой лекции было дано определение аффинного пространства. Напомним его.

Определение 4. *Аффинным пространством* называется множество T, на котором действует некоторое векторное пространство V, т.е. имеется операция $T \times V \to T$, $(t, v) \mapsto t + v$, причём

- $\forall t \in T, v, w \in V \ t + (v + w) = (t + v) + w$;
- $\forall t \in T \ t + 0 = t$;
- $\exists t \in T \ t + v = t \Rightarrow v = 0$;
- $\forall t_1, t_2 \in T \ \exists v \in V \ t_1 + v = t_2$.

(первые два свойства говорят, что имеется гомоморфизм абелевой группы векторов пространства V в группу преобразований множества T, третье свойство отражает «свободность» действия, а четвёртое — «транзитивность»).

Тогда для каждого линейного оператора решения соответствующей однородной системы уравнений (*ядро оператора*, прообраз нуля) образуют, очевидно, линейное пространство, а решения системы уравнений с данной правой частью b — аффинное пространство, на котором это линейное пространство действует (проверьте!).