

Матричная алгебра и системы линейных уравнений

П.С. Петров

Дальневосточный федеральный университет

Тихоокеанский океанологический институт им. В.И. Ильичева ДВО РАН

E-mail: petrov@poi.dvo.ru

Введение

В этой части курса мы переходим к изучению собственно линейной алгебры. Можно считать, что материал предыдущих лекций готовил Вас к освоению этого предмета. Линейная алгебра возникла как теория систем линейных уравнений, хотя в дальнейшем стало понятно, что наиболее интересный объект в ней – матрица, или линейный оператор. Первая часть нашего курса исключительно конкретна: мы будем иметь дело в основном с системами линейных уравнений. В следующей части мы, оттолкнувшись от наработанных представлений, постараемся посмотреть на эти системы с более абстрактной точки зрения. Изучение рассмотренных здесь понятий для вас необходимо по целому ряду причин. Во-первых, эта теория важна сама по себе и позволяет решать многие задачи. Во-вторых, язык современной науки (в особенности физики) основан на понятиях, которые мы рассмотрим в этом курсе (векторы, линейные операторы, самосопряженность, метод наименьших квадратов – все это вещи, без которых невозможно представить себе, например, квантовой механики). В-третьих, линейная алгебра – самый простой из математических курсов (хотя поначалу вам, возможно, будет казаться иначе), и потому он должен служить очередной ступенью в развитии вашего мышления. Если вы пропустите эту ступень, то быстро обнаружите, что от следующей вас отделяет пропасть, через которую не так просто перебраться. Иными словами, вы будете совершенно не готовы воспринимать действительно сложные математические предметы на старших курсах.

1 Матрицы, умножение матриц, системы линейных уравнений в матричной форме

Предметом нашего исследования на этой лекции будут матрицы. Мы начнем с формального определения, установив, что матрицы суть просто объекты, которые выглядят так-то и над которыми можно совершать

такие-то операции, то есть подойдем к ним исключительно формально. Может быть непонятно, для чего нужно все это. Пожалуйста, наберитесь терпения – мы основательно разберемся в этом в течение этого семестра. Пока же постарайтесь воспринять это как своеобразный язык, на котором необходимо научиться говорить без запинки.

1.1 Матрицы: определения и примеры

Матрицей называют квадратную таблицу из m строк и n столбцов чисел:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}. \quad (1)$$

Как видите, это довольно простая вещь. Точки показывают, что так как числа m и n для разных матриц могут быть различными, мы не можем выписать явно все элементы, пока речь идет о “некоторой” матрице. Они показывают, что хотя на самом деле в этой таблице m строк, записаны лишь первая, вторая и последняя из них. Разумеется, числа в матрицах бывают, например, вещественными, рациональными или комплексными. В редких случаях интерес представляют даже матрицы с исключительно целыми числами. В зависимости от того, из какого числового множества берутся элементы матрицы, говорят о матрицах над полями $\mathbb{R}, \mathbb{Q}, \mathbb{C}$. Теперь давайте посмотрим на конкретную матрицу (обозначим ее A), скажем, такую:

$$A = \begin{pmatrix} 4 & 3 & 1 & \frac{3}{7} \\ 3.25 & 1 & 1 & 0 \\ 1+i & 0 & 0 & 0 \end{pmatrix}.$$

Во-первых, видно, что эта матрица содержит как целые, так и рациональные числа. Одно из чисел и вовсе комплексное. Если матрица имеет элементы из разных числовых полей, то выбирают, конечно, самое широкое из них, т.е. в данном случае это матрица над \mathbb{C} . Когда матрица содержит m строк и n столбцов, то говорят, что она лежит в классе матриц $Mat(m, n)$ – эти числа определяют размер матрицы. В частности, $A \in Mat(3, 4)$. Следующая важная вещь – индексация элементов матрицы. По определению, элемент $a_{i,j}$ – число, стоящее на пересечении i -ой строки (первый индекс!) и j -ого столбца (второй индекс!) матрицы A ¹. Иногда, когда это не вызывает путаницы, запятую между индекса-

¹Обратите внимание: матрицу обозначают прописной буквой A , а ее элементы той же самой буквой, но строчной.

ми опускают. Так, например, $a_{13} = 1$, $a_{21} = 3.25$, $a_{34} = 0$. Иногда удобнее разделять индексы на верхний и нижний. В этой случае элемент a_j^i означает число, стоящее на пересечении i -ой строки (верхний индекс) и j -ого столбца (нижний индекс) матрицы A .

Существуют матрицы специального вида, которые особенно важны. К ним относятся квадратные матрицы, векторы-строки и векторы столбцы. Матрица называется квадратной, если количество строк в ней совпадает с количеством столбцов. Множество квадратных матриц размера $n \times n$ над некоторым полем обозначают как $Mat(n)$. В случае, если в матрице только один столбец, ее называют вектором-столбцом (иногда, когда понятно, какой вектор имеется в виду, говорят просто “вектор”) и обозначают либо строчной буквой с чертой или стрелочкой (на письме), либо строчной буквой, выделенной жирным шрифтом (так чаще делают в книгах): \bar{a} , \vec{a} . Элементы вектора-столбца имеют только один индекс – номер строки²:

$$\bar{b} = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{pmatrix}.$$

Аналогичным образом определяется и вектор строка:

$$\bar{x} = (x_1, x_2, \dots, x_n).$$

Если непонятно, о каком векторе идет речь, следует считать, что подразумевается вектор-столбец. Обратите внимание, что при $n = 2$ и $n = 3$ векторы с n координатами (в строку или столбец – неважно) суть те самые векторы, что встречались вам в школьной геометрии и физике. Используя декартову систему координат, их можно изображать.

1.2 Операции над матрицами

Из всех операций, которые можно выполнять над матрицами, самые простые – **сложение** и **умножение на число**. Эти операции выполняются точно так же, как и соответствующие операции над векторами, т.е. **поэлементно**. Это означает очень простую вещь: складывать можно только матрицы одинакового размера, причем элемент, стоящий в i -ой строке и j -ом столбце матрицы-суммы получается сложением

²По аналогии с матрицами, в случае одного столбца все его элементы должны иметь вид a_i^1 , где i – номер строки. Нижний индекс, равный 1 для всех элементов, принято опускать.

элементов в i -ой строке и j -ом столбце у матриц-слагаемых, т.е. если $A, B \in \text{Mat}(m, n)$, то $C = A + B \in \text{Mat}(m, n)$ и, кроме того $c_j^i = a_j^i + b_j^i$. Рассмотрим пример:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{2} & -8 \\ \frac{3}{2} & 0.1 & 1 \\ 0 & -1 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1+0 & 2+\frac{1}{2} & 3-8 \\ 4+\frac{3}{2} & 5+0.1 & 6+1 \\ 7+0 & 8-1 & 9+0 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{2} & -5 \\ \frac{11}{2} & 5.1 & 7 \\ 7 & 7 & 9 \end{pmatrix}$$

Разумеется, сложение матриц обладает теми же свойствами, что и сложение обычных чисел. Прежде всего это ассоциативность (сочетательный закон: $A + (B + C) = (A + B) + C$) и коммутативность (переместительный закон $A + B = B + A$). Как и сложение, умножение матрицы на число производится поэлементно, т.е. для умножения матрицы на число нужно каждый ее элемент умножить на это число.

Самая интересная операция, которая и придает матричной алгебре всю ее значимость, есть умножение одной матрицы на другую. Перемножать можно не всякие матрицы. Более точно выполняется следующее правило: матрицу A можно умножить на матрицу B тогда и только тогда, когда количество столбцов A совпадает с количеством строк в B , причем если $A \in \text{Mat}(m, n)$, $B \in \text{Mat}(n, k)$, то $C = AB \in \text{Mat}(m, k)$ (т.е. количество строк произведения совпадает с количеством строк в первом сомножителе, а количество столбцов в произведении совпадает с количеством столбцов во втором сомножителе). Элементы матрицы произведения определяются с помощью следующего правила (правило “строка на столбец”):

$$c_j^i = \sum_{s=1}^n a_s^i b_j^s = a_1^i b_j^1 + a_2^i b_j^2 + \dots + a_n^i b_j^n.$$

Это правило можно схематически изобразить так:

$$i \rightarrow \begin{pmatrix} a_1^i & \dots & a_n^i \end{pmatrix} \times \begin{pmatrix} \downarrow j \\ b_j^1 \\ \vdots \\ b_j^n \end{pmatrix} = i \rightarrow \begin{pmatrix} \dots & \downarrow j \\ \dots & c_j^i & \dots \\ \dots & \vdots & \dots \end{pmatrix}.$$

На схеме показано, что для получения элемента в i -ой строке и j -м произведения необходимо попарно перемножить элементы i -ой строки матрицы A и j -ого столбца матрицы B , а затем результаты перемножений

сложить. Рассмотрим пример перемножения матриц:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 12 & 18 \\ 4 & 6 \\ 6 & 7 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 \cdot 12 + 2 \cdot 4 + 3 \cdot 6 & 1 \cdot 18 + 2 \cdot 6 + 3 \cdot 7 \\ 1 \cdot 12 + \frac{1}{2} \cdot 4 + \frac{1}{3} \cdot 6 & 1 \cdot 18 + \frac{1}{2} \cdot 6 + \frac{1}{3} \cdot 7 \end{pmatrix} = \begin{pmatrix} 38 & 51 \\ 16 & 23\frac{1}{3} \end{pmatrix}.$$

Эти же матрицы, согласно правилу, можно перемножить и в другом порядке:

$$\begin{pmatrix} 12 & 18 \\ 4 & 6 \\ 6 & 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} =$$

$$= \begin{pmatrix} 12 \cdot 1 + 18 \cdot 1 & 12 \cdot 2 + 18 \cdot \frac{1}{2} & 12 \cdot 3 + 18 \cdot \frac{1}{3} \\ 4 \cdot 1 + 6 \cdot 1 & 4 \cdot 2 + 6 \cdot \frac{1}{2} & 4 \cdot 3 + 6 \cdot \frac{1}{3} \\ 6 \cdot 1 + 7 \cdot 1 & 6 \cdot 2 + 7 \cdot \frac{1}{2} & 6 \cdot 3 + 7 \cdot \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 30 & 33 & 42 \\ 10 & 11 & 14 \\ 13 & 15\frac{1}{2} & 20\frac{1}{3} \end{pmatrix}.$$

Этот пример показывает, что произведение матриц меняется при перестановке сомножителей, причем очень существенно. При умножении матрицы 3×2 на матрицу 2×3 получается матрица 3×3 . Если переписать их, то результатом будет уже матрица 2×2 . Представьте себе матрицу A , скажем, 5×2 и матрицу B размера 2×3 . Тогда у произведение AB будет иметь размер 5×3 , а выражение BA просто не будет иметь смысла, т.к. в этом случае в ПЕРВОЙ матрице (B) всего 3 столбца, а во ВТОРОЙ матрице (A) 5 строк, и, согласно правилу, такие матрицы нельзя перемножать. Таким образом, для матриц, вообще говоря, не выполняется переместительный закон умножения³, т.е. $AB \neq BA$. В силе, впрочем, остаются сочетательный закон умножения (“ассоциативность”) и распределительный закон умножения относительно сложения (“дистрибутивность”). Они могут быть выражены соотношениями

$$(AB)C = A(BC),$$

$$A(B + C) = AB + AC.$$

Мы докажем эти соотношения в следующей лекции.

Еще одной важной операцией, которую можно применять к матрицам, является транспонирование. При транспонировании строки исходной матрицы становятся столбцами результирующей и наоборот. Транспонирование матрицы A обозначается как A^T . Математически правило транспонирования записывается следующим образом:

$$B = A^T \Rightarrow b_j^i = a_i^j.$$

³Математики говорят умную фразу: “Умножение матриц некоммутативно.”

Рассмотрим пример транспонирования матрицы:

$$\begin{pmatrix} 12 & 18 \\ 4 & 6 \\ 6 & 7 \end{pmatrix}^T = \begin{pmatrix} 12 & 4 & 6 \\ 18 & 6 & 7 \end{pmatrix}.$$

Двумя важными матрицами в множестве $Mat(n)$ являются нулевая матрица 0 и единичная матрица E :

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Нулевая матрица состоит из одних нулей, а в единичной матрице элементы, стоящие на главной диагонали (диагонали, соединяющей левый верхний и правый нижний угол), равны 1 (все остальные – нули). Число n может быть различным, и обычно его значение ясно из контекста. Единичная матрица обладает тем свойством, что при умножении на нее справа и слева любая другая матрица не изменяется, т.е.

$$AE = EA = A$$

для всех A , имеющих такие размеры, чтобы эти произведения могли существовать. Видно, что умножение A на E дает тот же эффект, что умножение A на число 1. Поэтому единичную матрицу часто отождествляют с этим числом.

Можно отметить, что квадратные матрицы одного размера $Mat(n)$ всегда можно умножать друг на друга. Так как их всегда можно складывать, то в совокупности множество $Mat(n)$ над произвольным полем чисел образует кольцо. Как было выяснено ранее, это кольцо некоммутативно. Кроме того, мы до сих пор не рассматривали вопроса о делении матриц. Разумеется, так же, как и в случае комплексных чисел, вместо деления мы будем использовать “умножение на обратный элемент”. Числа мы называем обратными друг к другу, если их произведение равно 1. Но произведение квадратных матриц – это всегда квадратная матрица. Как в этом случае сформулировать определение обратной матрицы? Вспомним о матрице E , которая в некотором смысле эквивалентна числу 1! Теперь, собственно, мы готовы к формулировке:

Определение. Предположим, что для квадратной матрицы A существует такая матрица A^{-1} , что $AA^{-1} = A^{-1}A = E$. Тогда матрица A является обратимой, а матрица A^{-1} называется обратной к A .

Рассмотрим пример: пусть $A = \begin{pmatrix} 5 & 6 \\ 6 & 7 \end{pmatrix}$, умножим A на матрицу $B = \begin{pmatrix} -7 & 6 \\ 6 & -5 \end{pmatrix}$:

$$\begin{pmatrix} 5 & 6 \\ 6 & 7 \end{pmatrix} \begin{pmatrix} -7 & 6 \\ 6 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Последняя матрица – единичная матрица второго порядка. Следовательно $B = A^{-1}$, а матрица A – обратима.

Не все квадратные матрицы обратимы. Этот печальный факт является основой всех сложностей с решением систем линейных уравнений. В задаче в конце лекции указывается путь доказательства этого утверждения. Обязательно решите ее.

1.3 Системы уравнений в матричной форме

Язык матричной алгебры очень хорошо подходит для описания и исследования систем линейных алгебраических уравнений. Мы начнем с простого примера, а затем запишем общий результат. Пусть дана система уравнений

$$\begin{cases} x + y + z = 1 \\ x - 2y + z = 4 \\ 2x + y - z = 0 \end{cases}$$

Разумеется, для нас прежде всего интересно нахождение неизвестных x, y, z . Мы имеем дело с тремя числовыми равенствами, которые должны выполняться одновременно. Это эквивалентно тому выполнению одного векторного равенства. Действительно, предположим, что у нас есть векторы-столбцы

$$\bar{l} = \begin{pmatrix} x + y + z \\ x - 2y + z \\ 2x + y - z \end{pmatrix}, \quad \bar{r} = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix},$$

где \bar{l} – вектор из левых частей всех равенств, а \bar{r} – из правых. Тогда система уравнений может быть заменена одним векторным равенством $\bar{l} = \bar{r}$ (два вектора равны тогда и только тогда, когда попарно равны все их координаты). Рассмотрим теперь подробнее вектор \bar{l} . Все его координаты представляют собой набор чисел x, y, z , сложенных с некоторыми коэффициентами. Согласно правилу умножения матриц, этот вектор-столбец

может быть представлен в виде произведения некоторой матрицы на вектор неизвестных. Эту матрицу легко определить:

$$\bar{l} = \begin{pmatrix} x + y + z \\ x - 2y + z \\ 2x + y - z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

. Проверьте перемножением, что это действительно так. Сформулируйте для себя правило, по которому можно получить матрицу из системы уравнений. Введем обозначения:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -1 \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix},$$

матрицу A мы будем называть основной матрицей системы уравнений, а вектор-столбец \bar{b} – столбцом свободных членов. Таким образом, любую систему уравнений можно записать в виде

$$A\bar{x} = \bar{b},$$

где A – некоторая матрица, \bar{b} – некоторый известный вектор-столбец, \bar{x} – столбец неизвестных. Предположим, что мы научились находить матрицы, обратные к данным, т.е. мы можем построить A^{-1} . Тогда следующая цепочка равенств

$$A\bar{x} = \bar{b} \Rightarrow A^{-1}A\bar{x} = A^{-1}\bar{b} \Rightarrow E\bar{x} = A^{-1}\bar{b} \Rightarrow \bar{x} = A^{-1}\bar{b}$$

позволяет найти вектор-столбец неизвестных \bar{x} . Для записанной выше системы матрица A^{-1} имеет вид

$$A^{-1} = \begin{pmatrix} \frac{1}{9} & \frac{2}{9} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{2}{9} & \frac{1}{9} & -\frac{1}{3} \end{pmatrix}.$$

Используйте ее, чтобы решить систему.

1.4 Примеры использования матриц

Давайте рассмотрим еще один пример использования матриц – преобразования поворота плоскости. Любой вектор \bar{x} на плоскости может быть представлен парой координат в некоторой декартовой системе

$\bar{x} = (x^1, x^2)^T$ ⁴. Тогда выполнено равенство

$$\bar{x} = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = x^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

где $\bar{e}_1 = (1, 0)^T$ и $\bar{e}_2 = (0, 1)^T$ – единичные координатные векторы (орты). Рассмотрим теперь матрицу

$$A_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}.$$

Умножим ее на координатные векторы

$$A_\phi \bar{e}_1 = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix},$$

$$A_\phi \bar{e}_2 = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}.$$

Из этих формул видно, что при умножении на A_ϕ оба координатных вектора совершили поворот на угол ϕ (изобразите на координатной плоскости векторы $\bar{e}_1, \bar{e}_2, A_\phi \bar{e}_1, A_\phi \bar{e}_2$ для угла $\phi = \pi/3$, вы сразу все поймете). Поскольку при этом с ними повернулись и координатные оси, то отсюда следует, что любой вектор \bar{x} при умножении на A_ϕ повернется на угол ϕ . Поэтому матрица A_ϕ называется матрицей поворота. Составив ее для некоторого угла ϕ вы легко сможете поворачивать любые векторы на угол ϕ просто умножая их на эту матрицу. Чтобы хорошо разобраться, как это делать, решите четвертую задачу.

1.5 Задачи

Решите задачи № 788, 790, 795, 796, 801, 802, 828, 829 из задачник И.В. Проскурякова [1].

1. Как было показано в лекции, умножение матриц некоммукативно, так как вообще говоря AB и BA могут быть даже разного размера. Предположим теперь, что AB и BA все-таки одинакового размера.

- а. Докажите, что A и B – квадратные матрицы одинакового размера.

⁴Обратите внимание: я хочу, чтобы \bar{x} был вектором-столбцом, однако в тексте столбец записывать неудобно. Поэтому я записываю его с строку (ее гораздо удобнее встраивать в текст) и применяю к ней транспонирование. Транспонированная строка есть столбец. Привыкните к этой записи, она часто встречается.

- b.** Выясните, всегда ли в этом случае $AB = BA$ (указание рассмотрите для начала матрицы небольшого размера, например, 2×2).
2. Пусть $A \in \text{Mat}(m, n)$, \bar{b} – вектор-строка, \bar{c} – вектор-столбец. Выясните, при каких размерах матрицы и векторов можно умножить A на \bar{b} , A на \bar{c} или \bar{c} на \bar{b} .
3. Постарайтесь доказать, что не все матрицы обратимы. Доказательство таких “отрицательных” утверждений требует построения контрпримера, т.е. примера того, что утверждение “все матрицы обратимы” неверно. Вот каким путем можно следовать:
- a.** Постройте две такие квадратные матрицы A, B , что $A \neq 0$, $B \neq 0$, но при этом $AB = 0$ (для чисел это невозможно, но с матрицами такой фокус получается). Возьмите квадратные матрицы небольшого размера, скажем 2×2 . Пусть большинство элементов в них будут нулевыми, подберите их так, чтобы их произведение давало 0.
- b.** Возьмите построенные матрицы A, B . Так как они ненулевые, то можно попытаться предположить, что у них есть обратные⁵. Умножьте равенство $AB = 0$ на матрицу, обратную к A . Что при этом получится?
4. Пусть одна из вершин правильного треугольника с центром в начале координат расположена в точке $(\sqrt{2}/2, \sqrt{2}/2)$. Найдите координаты остальных вершин. Для этого
- a.** Представьте себе равносторонний треугольник с одной из вершин на оси абсцисс (нарисуйте!). Найдите координаты его вершин.
- b.** Подумайте, как можно получить данный треугольник из треугольника п. **a**.
- c.** Запишите матрицу поворота, которая переводит треугольник из п. **a** в нужный вам.
- d.** Поверните все вершины найденные в п. **a**. помощью матрицы из п. **c**.

⁵Этот прием называется *reductio ad absurdum* – “приведение к абсурду или доказательство от противного. Вы предполагаете, что верно отрицание того утверждения, которое вам нужно доказать и стараетесь найти противоречие. Сейчас вам нужно доказать утверждение “не все матрицы обратимы”. Вы предполагаете обратное к нему – “все матрицы обратимы”, а затем ищете противоречие.

2 Дополнительные сведения о матрицах*

2.1 Обозначения Эйнштейна

Часто неудобно записывать суммы в различных выражениях с индексами (если сумм много, это выглядит чересчур громоздко), поэтому пользуются следующим соглашением. **Соглашение о суммировании.** Если в некотором выражении один и тот же буквенный индекс встречается вверху и внизу, то по нему подразумевается суммирование.

Пример. Пусть $A = (a_{i,j})$ – квадратная матрица размера 3×3 . Тогда выражение $a_1^i a_3^i$ означает следующую сумму $\sum_{i=1}^3 a_1^i a_3^i = a_1^1 a_3^1 + a_1^2 a_3^2 + a_1^3 a_3^3$.

Соглашение о суммировании впервые ввел А.Эйнштейн для записи формул из общей теории относительности, где зачастую встречаются выражения с 4 и 6 индексами.

2.2 Доказательство матричных тождеств

Обозначения Эйнштейна очень удобны для доказательства различных тождеств, относящихся к умножению матриц. Так, например, свойство ассоциативности умножения матриц $A(BC) = (AB)C$ может быть выведено следующим образом. Пусть $A \in Mat(m, n)$, $B \in Mat(n, k)$, $C \in Mat(k, l)$, тогда оба произведения существуют и имеют одинаковый размер. Две матрицы равны тогда и только тогда, когда они имеют одинаковый размер, и соответствующие их элементы равны. Таким образом, нам для доказательства тождества нам нужно установить, что $(A(BC))^i_j = ((AB)C)^i_j$. В обозначениях Эйнштейна это получается почти автоматически:

$$\begin{aligned} ((AB)C)^i_j &= (AB)^i_p C^p_j = (A^i_q B^q_p) C^p_j = A^i_q B^q_p C^p_j = \\ &= A^i_q (B^q_p C^p_j) = A^i_q (BC)^q_j = (A(BC))^i_j. \end{aligned}$$

2.3 Матрицы в MATLAB

Любая система компьютерной алгебры, разумеется, умеет перемножать матрицы. Язык умножения матриц исключительно универсален, так что очень многие задачи, возникающие в физике, математике, программировании могут быть решены с помощью матричной алгебры. Именно поэтому широкую популярность приобрели языки программирования, в основе которых лежит матричная алгебра. Наиболее попу-

лярным из них в настоящее время является, пожалуй, язык, встроенный в систему MATLAB (почти идентичен синтаксис у ее GNU-аналогов Scilab, Octave). В отличие от обычных языков программирования, где двумерные массивы чисел являются "безжизненными" хранилищами, в MATLAB двумерный массив по умолчанию считается матрицей. Соответственно, двумерные массивы можно перемножать (обычным оператором "*"), и перемножение по умолчанию осуществляется по правилу умножения матриц. Задать матрицу можно с помощью оператора []. Его можно использовать в виде $A = [a \ b \ c \ \dots]$ или $A = [a; b; c; \dots]$. В первом случае он формирует вектор-строку из элементов a, b, c , а во втором – вектор столбец. Соответственно, чтобы сформировать квадратную матрицу, нужно сделать несколько столбцов, а затем собрать их в одну строку. Запишем, например, в переменные A и B матрицы

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$$

соответственно. Начнем с первой матрицы. Сформируем векторы-строки:

```
» a1 = [1 2 3]
a1 = 1 2 3
```

```
» a2 = [4 5 6]
a2 = 4 5 6
```

```
» a3 = [7 8 9]
a3 = 7 8 9
```

Далее запишем их в столбец одну под другой:

```
» A = [a1; a2; a3]
A =
1 2 3
4 5 6
7 8 9
```

Теперь первая матрица хранится в переменной A. Запишем другую матрицу в переменную B, однако теперь выполним все одной командой:

```
» B = [[1 1 1]; [2 2 2]; [3 3 3]]
B =
1 1 1
2 2 2
3 3 3
```

Теперь мы можем перемножить матрицы, просто выполнив команду

```
A*B:
» A*B
ans =
14 14 14
32 32 32
50 50 50
```

Если Вам потребуется перемножить матрицы с рациональными или алгебраическими числами, такими как, скажем $\frac{1}{3}$, то лучше воспользоваться пакетом символьных вычислений, так как по умолчанию MATLAB запишет это число, используя плавающую точку:

```
» 1/3
ans =
0.3333
```

Все последующие вычисления будут выполняться приближенно. Чтобы производить вычисления точно, нужно создать матрицу символьного типа:

```
» C = sym([[1/1 1/2 1/3]; [1/2 1/2 1/2]; [2/3 3/2 1/2]])
C =
1, 1/2, 1/3
1/2, 1/2, 1/2
2/3, 3/2, 1/2
```

```
» B*C
ans =
13/6, 5/2, 4/3
13/3, 5, 8/3
13/2, 15/2, 4
```

2.4 Вложение поля комплексных чисел в кольцо матриц

Алгебраическая структура кольца матриц очень богата. В частности, матрицы содержат в себе комплексные числа. Алгебраически комплексные числа появляются при добавлении к вещественным числам корня уравнения $x^2 = -1$ ⁶. Вещественные числа мы отождествляем с матрица-

⁶Математики говорят, что поле комплексных чисел есть алгебраическое расширение поля вещественных чисел, полученное добавлением к ним корня алгебраического уравнения $x^2 = -1$.

ми вида

$$\alpha E = \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & \alpha & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha \end{pmatrix}.$$

При умножении произвольной матрицы B на матрицу αE все элементы B увеличиваются в α раз: $(\alpha E)B = \alpha EB = \alpha B$. Поэтому при матричном умножении матрицы вида αE действуют на другие матрицы в точности как обычные числа. Числу -1 соответствует матрица $-E$. Так что мы можем попробовать найти матричное решение уравнения $X^2 = -E$, и это решение будет обладать теми же свойствами, что число i . Решим это уравнение в $Mat(2, 2)$. Для этого положим $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$. Теперь мы можем получить 4 уравнения на x_{ij} , записав поэлементно матричное равенство:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Решив его (проделайте это!), можно получить, что $X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (на самом деле, конечно, уравнение имеет два решения). Эта матрица с вещественными коэффициентами соответствует числу i . Поэтому матрицы вида $aE + bX$ должны быть эквивалентны числам вида $a + bi$. Эта эквивалентность устанавливается помощью следующей инъективной функции⁷ из множества комплексных чисел в множество матриц:

$$\varphi : a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Это отображение сохраняет алгебраические операции, т.е. $\varphi(z_1)\varphi(z_2) = \varphi(z_1 z_2)$ и $\varphi(z_1) + \varphi(z_2) = \varphi(z_1 + z_2)$ (проверьте это! возьмите $z_1 = a + bi$, $z_2 = c + di$). Такое отображение (взаимно-однозначное и сохраняющее алгебраические операции) в математике называют изоморфизмом. Наличие изоморфизма указывает на то, что с алгебраической точки зрения два множества устроены одинаково, и узнав что-то об одном из них, мы автоматически можем использовать это знание для исследования другого. Мы уже умеем находить обратные к комплексным числам, однако не

⁷Функцию f называют инъективной, если она переводит различные элементы в различные, т.е. два разных элемента не могут склеиться при действии на них f . Это можно записать так: $a \neq b \Rightarrow f(a) \neq f(b)$. А можно записать и еще короче: $f(a) = f(b) \Rightarrow a = b$. Подумайте, почему эти две записи значат одно и то же.

умеем находить обратные матрицы. Но с алгебраической точки зрения множество матриц вида $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ устроено так же, как и множество \mathbb{C} . Поэтому мы можем воспользоваться следующей диаграммой:

$$\begin{array}{ccccc} A & & ? & & A^{-1} \\ \varphi^{-1} \downarrow & & & & \uparrow \varphi \\ z & & \rightarrow & & z^{-1} \end{array}$$

Она означает, что мы можем перевести матрицу $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ в комплексное число с помощью функции, обратной к φ , затем найти обратное для этого числа, а затем найти соответствующую ему матрицу:

$$\begin{aligned} \varphi^{-1} \left(\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \right) &= a + bi \\ (a + bi)^{-1} &= \frac{a - bi}{a^2 + b^2} \\ A^{-1} = \varphi \left(\frac{a - bi}{a^2 + b^2} \right) &= \begin{pmatrix} \frac{a}{a^2 + b^2} & \frac{-b}{a^2 + b^2} \\ \frac{b}{a^2 + b^2} & \frac{a}{a^2 + b^2} \end{pmatrix} \end{aligned}$$

Проверьте, что найденная матрица действительно обратна к A .

3 Метод Гаусса для решения систем линейных алгебраических уравнений

Рассмотрим решение простой системы линейных уравнений

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

Мы будем пользоваться методом последовательного исключения неизвестных. Выразим x из первого уравнения и подставим в два других:

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

Теперь два последних уравнения содержат две неизвестных, а про первое уравнение мы можем временно забыть. Теперь повторим эту же процедуру с последними уравнениями: выразим y из третьего и подставим во

второе:

$$\begin{cases} x + y + z = 3 \\ -y - 3z = -4 \\ -3(4 - 3z) + z = -2 \end{cases} \quad \begin{cases} x + y + z = 3 \\ -y - 3z = -4 \\ 10z = 10 \end{cases}$$

Наше последнее уравнение содержит одну неизвестную, и его можно решить: $10z = 10$, $z = 1$. Подставим найденное значение во второе уравнение: $-y - 1 \cdot 3 = -4$, $y = 1$. Подставим найденные z, y в первое уравнение: $x + 1 + 1 = 3$, $x = 1$. Рассмотренный метод решения состоит из двух шагов, на каждом из которых некоторая переменная выражается из одного из "рабочих" уравнений и подставляется во все остальные. В результате этого остается на одно рабочее уравнение меньше, причем в них становится меньше на одну неизвестную. Поэтому метод Гаусса также называется методом последовательного исключения неизвестных. Шаги производятся до тех пор, пока не останется одно рабочее уравнение. На этом завершается прямой ход метода Гаусса. Если в нем будет одна неизвестная, как в примере, рассмотренном выше, то ее можно найти, а затем, подставляя во все уравнения, начиная с последнего, найденные ранее неизвестные, можно найти значения всех переменных. Это называется обратным ходом метода Гаусса.

Метод Гаусса удобнее использовать в матричной форме. Для этого необходимо перейти от системы линейных уравнений к ее расширенной матрице, выписав коэффициенты. Пусть нам дана система уравнений вида

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \dots a_n^1 x^n = b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \dots a_n^2 x^n = b^2 \\ \vdots \\ a_1^m x^1 + a_2^m x^2 + \dots a_n^m x^n = b^m \end{cases}, \quad (2)$$

которую мы будем также записывать в матричной форме как $Ax = b$, где A – основная матрица системы:

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix}$$

$\bar{x} = (x^1, x^2, \dots, x^n)^T$ – вектор-столбец неизвестных, а $\bar{b} = (b^1, b^2, \dots, b^n)^T$ – вектор-столбец свободных членов. Расширенная матрица (2) запишется как

$$\left(\begin{array}{cccc|c} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{array} \right) \quad (3)$$

Рассмотрим, как преобразуется расширенная матрица системы линейных уравнений при последовательном исключении неизвестных на примере рассмотренной выше системы. Исходная матрица имеет вид

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -2 & 2 & 1 \\ 2 & 1 & -1 & 2 \end{array} \right)$$

После первого исключения неизвестных мы получаем

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -3 & 1 & -2 \\ 0 & -1 & -3 & -4 \end{array} \right),$$

а после второго

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -1 & -3 & -4 \\ 0 & 0 & 10 & 10 \end{array} \right).$$

Мы видим, что при реализации метода Гаусса основная матрица системы приводится к треугольному виду: ниже главной диагонали матрицы A появляются нули, в то время как на самой диагонали остаются ненулевые элементы. Обратно: очевидно, что если от любой строки расширенной матрицы отнять любую другую строку, то множество решений соответствующей матрицы системы не изменится (объясните, почему). Таким образом, мы можем производить метод Гаусса в матричной форме. Он состоит в приведении матрицы системы к треугольному виду с помощью элементарных преобразований – перестановки строк и вычитании из одной строки другой, умноженной на произвольное число. Обратите внимание, что при каждом исключении неизвестных мы делаем нулевыми элементы в ОДНОМ столбце под главной диагональю.

3.1 Общее описание алгоритма метода Гаусса.

Coming soon

3.2 Задачи.

Решите задачи №700-703, 717, 777 из задачника И.В.Проскурякова.

4 Кватернионы

Кватернионами называют числа вида $w = a + bi + cj + dk$, где $a, b, c, d \in \mathbb{R}$, а i, j, k суть единичные орты в трехмерном пространстве,

перемножаемые согласно правилам: $i^2 = j^2 = k^2 = ijk = -1$. При этом a называют скалярной частью кватерниона w (обозначается $a = \Re(w)$), а $bi + cj + dk$ – векторной, или мнимой частью $\Im(w)$ кватерниона w . Кватернион, у которого есть только скалярная часть, отождествляют с обычным числом, а кватернион, у которого есть только мнимая часть – с вектором в трехмерном пространстве. Кватернион общего вида можно также считать парой, состоящей из вектора и скаляра (a, \bar{v}) . В разных случаях удобны различные обозначения. Кватернионы $1, i, j, k$ называют базисными, так как любой другой кватернион может быть представлен в виде их комбинации с вещественными коэффициентами. Также, как комплексным числам соответствуют векторы на плоскости, кватернионам соответствуют векторы в четырехмерном пространстве.

4.1 Операции над кватернионами

На множестве кватернионов заданы три операции:

1. Кватернионы можно складывать, причем сложение осуществляется поэлементно: если $w_1 = a_1 + b_1i + c_1j + d_1k$, $w_2 = a_2 + b_2i + c_2j + d_2k$, то можно построить кватернион $w_1 + w_2$, который будет иметь вид $(a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$. Разумеется, кватернионы можно и вычитать, причем для любого кватерниона w существует обратный к нему ПО СЛОЖЕНИЮ кватернион $-w$. В векторной форме, очевидно, если $w_1 = (a_1, \bar{v}_1)$, $w_2 = (a_2, \bar{v}_2)$, то $w_1 + w_2 = (a_1 + a_2, \bar{v}_1 + \bar{v}_2)$.
2. Кватернионы можно умножать, причем их умножение дистрибутивно относительно сложения, а правила получаются из правил для базисных кватернионов: $i^2 = j^2 = k^2 = ijk = -1$. Из этих правил можно получить следствия, которые позволяют перемножить любую пару базисных кватернионов. Например, получим выражение для ij : домножим равенство $ijk = -1$ на $-k$, получим $ijk(-k) = -(-k)$, $-ijk^2 = k$, $-ij \cdot (-1) = k$, $ij = k$. Аналогичным образом можно получить $ji = -k$, $ki = j$, $ik = -j$, $jk = i$, $kj = -i$ (проверьте это). После того, как правила умножения базисных кватернионов установлены, легко получить произведение двух кватернионов общего вида: достаточно формально раскрыть скобки в произведении $(a_1 + b_1i + c_1j + d_1k)(a_2 + b_2i + c_2j + d_2k)$, а затем воспользоваться полученными правилами.

В общем виде формула перемножения кватернионов выглядит так:

если $w_1 = (a_1, \bar{v}_1)$, $w_2 = (a_2, \bar{v}_2)$, то

$$w_1 w_2 = (a_1 a_2 - (\bar{v}_1, \bar{v}_2)) + (a_1 \bar{v}_2 + a_2 \bar{v}_1 + [\bar{v}_1, \bar{v}_2]),$$

где (\bar{v}_1, \bar{v}_2) и $[\bar{v}_1, \bar{v}_2]$ – скалярное и векторное произведения \bar{v}_1 и \bar{v}_2 соответственно. Докажите эту формулу самостоятельно.

3. Также, как и на множестве комплексных чисел, на множестве кватернионов задано комплексное сопряжение, которое сохраняет скалярную часть кватерниона и меняет знак у векторной. Пусть $w = a + bi + cj + dk$, тогда

$$w^* = a - (bi + cj + dk)$$

Кватернион $w = a + bi + cj + dk$ естественно представлять себе как вектор (a, b, c, d) в четырехмерном пространстве. Длина такого вектора может быть найдена по теореме Пифагора:

$$|w| = a^2 + b^2 + c^2 + d^2.$$

Легко проверить, что "комплексное" соотношение $ww^* = |w|^2$ выполняется и для кватернионов. Это позволяет нам использовать метод нахождения обратных к комплексным числам и для кватернионов. Сформулируем результат в виде следующего утверждения:

Лемма. Для любого кватерниона $w = a + bi + cj + dk \neq 0$ существует обратный ПО УМНОЖЕНИЮ кватернион w^{-1} , такой что $ww^{-1} = w^{-1}w = 1$. Этот кватернион может быть найден по формуле:

$$w^{-1} = \frac{w^*}{|w|^2} = \frac{a - (bi + cj + dk)}{a^2 + b^2 + c^2 + d^2}.$$

Доказательство тривиально. На этом заканчивается "алгебраическая" часть описания кватернионов. Как мы видим, при сложении они ведут себя в точности так же, как матрицы, векторы или комплексные числа. Умножение их также не представляет сложностей, достаточно лишь раскрывать скобки и пользоваться свойствами умножения базисных кватернионов. Единственная особенность умножения кватернионов, которая принципиально отличает их от множества комплексных или вещественных чисел, состоит в его некоммутативности ($ab \neq ba$). Именно эта особенность не позволяет считать кватернионы числами: они не образуют поля. Множество, в котором выполняются все аксиомы поля, кроме коммутативности умножения в алгебре называют телом. Кватернионы –

пример тела, не являющегося полем. Для чего рассматривать некоммутативные тела? Ответ на этот вопрос мы получим в следующем разделе. Коротко его можно сформулировать так: кватернионы можно отождествить с вращениями трехмерного пространства. Конечно же, каждому такому вращению можно сопоставить обратное вращение, однако вращения не коммутируют друг с другом.

4.2 Представление вращений трехмерного пространства с помощью кватернионов

Теорема. Пусть дан вектор $\bar{v} = (x, y, z)$ в трехмерном пространстве \mathbb{R}^3 . Отождествим вектор с соответствующим ему кватернионом $\bar{v} = xi + yj + zk$. Тогда отображение φ , заданное по правилу

$$\varphi : v \mapsto qvq^{-1},$$

где $q = (\cos \theta, \sin \theta \bar{l})$ – единичный кватернион, поворачивает вектор \bar{v} на угол 2θ вокруг оси, заданной единичным вектором \bar{l} .

Доказательство. Для начала заметим, что, во-первых, действие единичного кватерниона q на вектор \bar{v} совпадает с действием любого кратного ему кватерниона αq , где $\alpha \in \mathbb{R}$ – число. Действительно,

$$\alpha q \bar{v} (\alpha q)^{-1} = \alpha q \bar{v} \alpha^{-1} q^{-1} = \alpha \alpha^{-1} q \bar{v} q^{-1},$$

так что на самом деле любой кватернион действует на вектор как оператор поворота.

Второе утверждение, которое нам потребуется, состоит в том, что результат действия φ на вектор тоже является вектором. Для произвольного кватерниона p , как и для комплексных чисел, выполняется соотношение $2\Re(p) = p + p^*$. Пусть $p = (\alpha, \bar{v})$. Тогда

$$\begin{aligned} 2\Re(\varphi(p)) &= 2\Re(qpq^{-1}) = 2\Re(qpq^*) = \\ &= qpq^* + (qpq^*)^* = qpq^* + qp^*q^* = q(p + p^*)q^* = q(2\Re(p))q^* = 2\Re(p), \end{aligned}$$

поэтому φ переводит $p = (\alpha, \bar{v})$ в кватернион с той же вещественной частью: $\varphi(p) = (\alpha, \bar{v}')$. Легко видеть, однако, что $|qpq^{-1}| = |p|$, так что отсюда следует, что непременно $|v| = |v'|$, причем если $\Re(p) = 0$, то и $\Re(\varphi(p)) = 0$. Следовательно отображение φ переводит единичные векторы в единичные векторы (кватернионы с нулевой вещественной частью).

Теперь перейдем к доказательству основного утверждения теоремы. Мы воспользуемся геометрической конструкцией на рис. 1. Пусть даны

два единичных вектора \bar{v}_0, \bar{v}_1 в плоскости, перпендикулярной к оси вращения \bar{l} . Пусть эти векторы расположены под углом θ друг к другу. Это означает, что, во-первых, ось вращения параллельна их векторному произведению (на самом деле $\bar{l} = [\bar{v}_0, \bar{v}_1]/|[\bar{v}_0, \bar{v}_1]|$), а во-вторых выполняется равенство $(\bar{v}_0, \bar{v}_1) = \cos \theta$. Тогда задающий вращение кватернион q может быть выражен как $q = \bar{v}_1 \bar{v}_0^*$ (объясните, почему). Так как векторы $\bar{v}_0, \bar{v}_1, \bar{l}$ некопланарны, то любой другой вектор можно единственным образом представить в виде их комбинации. Поэтому достаточно рассмотреть действие φ на этих трех векторах – тогда будет понятно, как отображение действует на произвольный вектор.

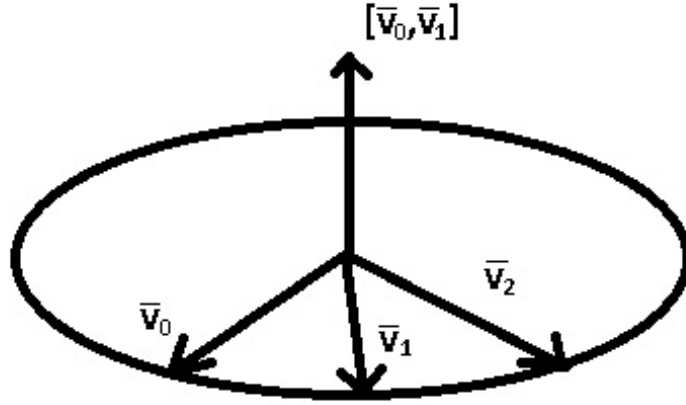


Рис. 1: К доказательству основной теоремы о кватернионах

Мы начнем с вектора \bar{v}_0 . Докажем, что

$$\bar{v}_1 \bar{v}_0^* = \varphi(\bar{v}_0) \bar{v}_1^*. \quad (4)$$

Проверяем непосредственно:

$$\varphi(\bar{v}_0) \bar{v}_1^* = q \bar{v}_0 q^* \bar{v}_1^* = q \bar{v}_0 (\bar{v}_1 \bar{v}_0^*)^* \bar{v}_1^* = q (\bar{v}_0 \bar{v}_0) (\bar{v}_1^* \bar{v}_1^*) = q(-1)(-1) = q = \bar{v}_1 \bar{v}_0^*.$$

Из равенства (4) непосредственно следует, что $\bar{v}_2 = \varphi(\bar{v}_0)$ лежит в той же плоскости, что и \bar{v}_0, \bar{v}_1 (это следует из векторной части (4): поскольку перпендикуляры к плоскостям, лежащим на парах векторов $\varphi(\bar{v}_0), \bar{v}_1$ и \bar{v}_0, \bar{v}_1 совпадают, то векторы $\varphi(\bar{v}_0)$ и \bar{v}_0 лежат в одной плоскости) и, кроме того, составляет с вектором \bar{v}_1 тот же угол, что и \bar{v}_1 составляет с \bar{v}_0 (это следует из скалярной части (4)). Это означает, что при отображении φ вектор \bar{v}_0 поворачивается вокруг оси \bar{l} на угол 2θ .

Теперь мы покажем, что

$$\bar{v}_1 \bar{v}_0^* = \varphi(\bar{v}_1) \varphi(\bar{v}_0)^*. \quad (5)$$

Это будет означать, что при отображении φ вектор \bar{v}_1 остается в плоскости \bar{v}_0, \bar{v}_1 и поворачивается на тот же угол, что и \bar{v}_0 . Проверяем (5):

$$\varphi(\bar{v}_1) \varphi(\bar{v}_0)^* = q \bar{v}_1 q^* (q \bar{v}_0 q^*)^* = q \bar{v}_1 q^* q \bar{v}_0^* q^* = q \bar{v}_1 \bar{v}_0^* q^* = q q q^* = q = \bar{v}_1 \bar{v}_0^*.$$

Последним шагом будет проверить, как φ действует на вектор \bar{l} . Здесь мы воспользуемся тем, что если кватернионы a, b таковы, что $\Im(a) \parallel \Im(b)$, то $ab = ba$. Мы знаем, что $\Im(q) \parallel \bar{l}$, поэтому

$$\varphi(\bar{l}) = q \bar{l} q^* = \bar{l} q q^*,$$

так что вектор \bar{l} инвариантен по отношению к φ . Таким образом, мы установили, что любой вектор в плоскости \bar{v}_0, \bar{v}_1 при отображении φ поворачивается на угол 2θ , а любой вектор, перпендикулярный к ней, остается неизменным. Таким образом, φ действительно представляет собой вращение вокруг оси \bar{l} на угол 2θ . ■

При доказательстве мы воспользовались тремя не доказанными ранее фактами:

1. $|ab| = |a||b|$ для любых кватернионов a, b ;
2. если $|q| = 1$, то $q^{-1} = q^*$;
3. если кватернионы a, b таковы, что $\Im(a) \parallel \Im(b)$, то $ab = ba$

Докажите эти утверждения.

4.3 Задачи

1. Вычислите

a $(1 + 2i + 3j + 4k)(4 + 3i + 2j + k)$ и $(4 + 3i + 2j + k)(1 + 2i + 3j + 4k)$

b $(2 + 4i + 6j)(1 + i + j + k)^{-1}$ и $(1 + i + j + k)^{-1}(2 + 4i + 6j)$

2. **a** Пространство повернули сначала вокруг оси z на $\pi/2$, а затем вокруг оси x на $\pi/3$. Определите результирующий поворот пространства, т.е. найдите его ось и угол. Что будет, если повороты применить в обратном порядке?

- б Пространство повернули сначала вокруг оси y на $\pi/4$, а затем вокруг оси z на $\pi/6$. Определите результирующий поворот пространства, т.е. найдите его ось и угол. Что будет, если повороты применить в обратном порядке?
3. а Даны точки $A = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ и $B = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$. Запишите кватернион кратчайшего поворота, переводящего точку A в точку B .
- б Даны точки $A = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}, \frac{3}{\sqrt{3}}\right)$ и $B = (1, 0, 0)$. Запишите кватернион кратчайшего поворота, переводящего точку A в точку B .
4. Решите задачу о тетраэдре из предыдущей лекции, используя кватернионы вместо матриц поворота. Какое решение проще?
5. Постройте изоморфизм тела кватернионов на
- а множество матриц с комплексными коэффициентами вида

$$\begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}$$

- б множество матриц с вещественными коэффициентами вида

$$\begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

5 Ранг матрицы. Теорема о разрешимости систем линейных уравнений

5.1 Ранг матрицы и основная теорема

Мы уже знаем, что в зависимости от строения основной матрицы система линейных уравнений может иметь единственное решение, иметь бесконечное множество решений или не иметь решений вообще. Классификация случаев, когда это происходит, получается очень простой, если ввести

Определение. Рангом матрицы называется количество ненулевых строк, которое остается после применения к ней прямого хода метода Гаусса.

До сих пор мы применяли метод Гаусса только к системам уравнений. Однако на прошлой лекции мы выяснили, что применять его удобнее в матричной форме, приводя расширенную матрицу системы к треугольному виду. Так как эту процедуру можно применить к произвольной матрице (вне зависимости от того, получена она из системы уравнений или нет), то наше определение имеет смысл. Ранг матрицы обозначается как $\text{rank}(A)$ или просто $r(A)$. В терминах рангов матриц мы можем сформулировать нашу основную теорему:

Теорема. Пусть система линейных алгебраических уравнений $A\bar{x} = \bar{b}$ с n неизвестными такова, что ранг ее основной матрицы A равен m , ранг расширенной матрицы $(A|\bar{b})$ равен k . Тогда выполняются следующие утверждения:

1. Система имеет решение тогда и только тогда, когда $m = k$.
2. Система имеет единственное решение тогда и только тогда, когда $m = k = n$.
3. В случае, если $m = k < n$, решение системы будет зависеть от $n - m$ свободных постоянных, каждая из которых может быть любым числом.

Доказательство. ■

5.2 Нахождение обратной матрицы методом Гаусса-Жордана

Наиболее эффективный способ нахождения обратной матрицы состоит в применении метода Гаусса-Жордана. Это название используется для обозначения аналога метода Гаусса, где в матричной форме производится не только прямой, но и обратный ход. На прямом ходу метода Гаусса мы приводим основную матрицу системы к нижнему квазитреугольному виду. Однако, как легко увидеть из следующего примера, после выполнения прямого хода можно, начав с нижнего правого угла квазитреугольной матрицы, выполнить метод Гаусса "наоборот сделав нулевыми также и все элементы ВЫШЕ главной диагонали. Рассмотрим ту же систему уравнений, с которой мы начали первую лекцию о методе Гаусса

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

Ее расширенная матрица имеет вид

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -2 & 2 & 1 \\ 2 & 1 & -1 & 2 \end{array} \right).$$

выполним прямой ход метода Гаусса, мы приведем ее к следующему виду

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -1 & -3 & -4 \\ 0 & 0 & 10 & 10 \end{array} \right).$$

Для этой матрицы мы выполним обратный ход метода

5.3 Связь с теорией определителей

Теорема. Ранг матрицы равен порядку наибольших ненулевых миноров этой матрицы.

Доказательство. Эту теорему мы докажем чуть позже, когда у нас будет возможность применить некоторые результаты теории линеалов. ■.

Теорема. Пусть дана квадратная матрица

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}.$$

Предположим, что A невырожденная, т.е. $\det(A) \neq 0$, или, что то же самое $\text{rank}(A) = n$. Тогда обратная к ней матрица может быть найдена по формуле:

$$A^{-1} = \frac{\hat{A}^T}{\det(A)},$$

где

$$\hat{A} = \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,n} \end{pmatrix}$$

– матрица, составленная из алгебраических дополнений к элементам матрицы A (такая матрица называется присоединенной к A). **Доказательство.** Напомню Вам формулу разложения определителя матрицы A по

i -ой строке:

$$\det(A) = \sum_{s=1}^n a_{i,s} A_{i,s}.$$

Пусть $X = \frac{\hat{A}^T}{\det(A)}$. Найдем общий элемент произведения AX :

$$(AX)_{i,j} = \sum_{s=1}^n \frac{1}{\det(A)} a_{i,s} (\hat{A}^T)_{s,j} = \frac{1}{\det(A)} \sum_{s=1}^n a_{i,s} (\hat{A})_{j,s} = \frac{1}{\det(A)} \sum_{s=1}^n a_{i,s} A_{j,s}.$$

Сумма в последнем выражении представляет собой формулу для разложения определителя матрицы A по j -ой строке, в которой вместо элементов $a_{j,s}$ стоят элементы $a_{i,s}$. В случае, когда $i = j$ эта сумма дает просто $\det(A)$ (получается предыдущая формула), а в случае, когда $i \neq j$, сумма представляет собой определитель матрицы, в которой две строки одинаковы (элементы $a_{i,s}$ стоят как в i -ой, так и в j -ой строках), и потому равна нулю. Итак, мы получаем, что

$$(AX)_{i,j} = \begin{cases} \frac{\det(A)}{\det(A)} = 1, & \text{если } i = j \\ 0, & \text{если } i \neq j \end{cases}$$

Это означает, что AX – единичная матрица. Итак, $AX = E$, следовательно $X = A^{-1}$. ■

Теорема (Правило Крамера). Пусть дана система линейных алгебраических уравнений вида $A\bar{x} = \bar{b}$ вида (2), причем матрица A является квадратной (размера $n \times n$) и невырожденной. Тогда ее решение может быть получено в виде

$$x^i = \frac{\det(A|_b^i)}{\det(A)},$$

где матрица $A|_b^i$ получается из матрицы A заменой i -ого столбца на столбец свободных членов \bar{b} .

Доказательство. В этот раз вспомним формулу разложения определителя матрицы A по i -ому столбцу:

$$\det(A) = \sum_{s=1}^n a_{s,i} A_{s,i}.$$

Так как $\bar{x} = A^{-1}\bar{b}$, то для $x^i = (\bar{x})_1^i$ мы получаем следующую цепочку равенств:

$$x^i = (\bar{x})_1^i = (A^{-1}\bar{b})_1^i = \sum_{s=1}^n (A^{-1})_s^i (\bar{b})_1^s = \sum_{s=1}^n \left(\frac{\hat{A}^T}{\det(A)} \right)_s^i b^s =$$

$$= \frac{1}{\det(A)} \sum_{s=1}^n (\hat{A}^T)_s^i b^s = \frac{1}{\det(A)} \sum_{s=1}^n (\hat{A})_i^s b^s = \frac{1}{\det(A)} \sum_{s=1}^n A_{s,i} b^s$$

Сумма в последнем выражении представляет собой формулу для разложения определителя матрицы A по i -ому столбцу, в котором вместо элементов $a_{s,i}$ стоят элементы b^s , т.е. в точности определитель матрицы $A|_b^i$. ■

6 LU-разложение, сложность метода Гаусса

6.1 Учебник

Обязательно прочитайте параграф §1.4 в учебнике Г.Стренга [1].

6.2 Задачи.

1. Объясните, для чего необходимо LU -разложение. Оцените, каким должен быть размер системы линейных уравнений, чтобы ее решение с десятью различными правыми частями с использованием LU -разложения занимало вдвое меньше времени, чем с помощью обычного метода Гаусса, применяемого десять раз.

2. Найдите количество умножений, необходимых для вычисления определителя

а с помощью разложения по строке;

б с помощью метода элементарных преобразований.

Оцените, каково будет соотношение в скорости расчетов для определителя порядков 20, 50.

3. Выполните LU -разложение матриц из задач №843,844 из задачника И.В.Проскурякова.

7 Дискретизация дифференциальных уравнений

В этой лекции мы рассмотрим важные примеры практических задач, в которых необходимо уметь решать системы уравнений с матрицами порядка, скажем, 2000×2000 . Эти задачи связаны с математическими

моделями физических явлений. Как правило, такие модели формулируются на языке теории дифференциальных уравнений, однако в большинстве случаев такие уравнения не могут быть решены аналитически (т.е. ответ нельзя записать в виде некоторой достаточно простой замкнутой формулы), причем это связано вовсе не с тем, что задачу сложно бывает сложно решить. Трудности здесь имеют принципиальный характер, связанный с тем, что многие интегралы не выражаются в элементарных функциях. Так, например, даже простой интеграл $\int \frac{\sin(x)}{x} dx$ не может быть выражен через “обычные” (элементарные) функции. Однако, как вам известно, любой интеграл можно вычислить приближенно, как предел римановых сумм. Аналогично любое дифференциальное уравнение, имеющее физический смысл, можно решить приближенно, произведя его ДИСКРЕТИЗАЦИЮ. Дискретизация означает переход от функций, заданных на бесконечном множестве точек некоторого интервала $[a, b]$, к функциям, заданным на конечном множестве точек, например, равномерно распределенных по $[a, b]$, т.е. таких точек x_i , что $x_0 = a$, $x_N = b$, $x_{i+1} - x_i = \Delta x = (b - a)/N$. Предположим, что нас интересуют лишь значения функции в этих точках, т.е. перейдем от функции $f(x)$, заданной при всех $x \in [a, b]$ к функции $f_i = f(x_i)$, заданной на конечном множестве $\{x_0, x_1, \dots, x_N\}$. При этом дифференциальное уравнение, связывающее значение функции в данной точке со значениями ее производных в этой точке, переходит, как мы увидим, в систему линейных уравнений, связывающих ее значения f_i в точках x_i .

8 Разностная аппроксимация производных

8.1 Разбиение отрезка и задание функции на сетке

Рассмотрим функцию $f(x)$, заданную на отрезке $x \in [a, b]$. Если функция f представляет собой значение некоторой физической величины, определенной во всех точках пространства x , то на практике ее значения могут быть, например, результатами измерения приборов, расположенных в некоторых точках отрезка $[a, b]$. Простейший вариант расположения N приборов на отрезке состоит в распределении их на равных расстояниях друг от друга, т.е. в точках $x_0 = a, x_1, x_2, \dots, x_N = b$, причем $x_i - x_{i-1} = \Delta x = (b - a)/N$. При таком измерении мы всегда знаем значения функции $f_i = f(x_i)$ лишь на конечном множестве точек. Между этими точками функцию можно считать, например, линейной. Если число N будет достаточно большим, то такая кусочно-линейная функция будет хорошим приближением для функции $f(x)$. Пример такого приближения

для функции $f(x) = \sin(x)$ на отрезке $[0, 5]$ для $N = 5$ показан на рис. 2. Пожалуй, наиболее важным достижением человеческой расы является

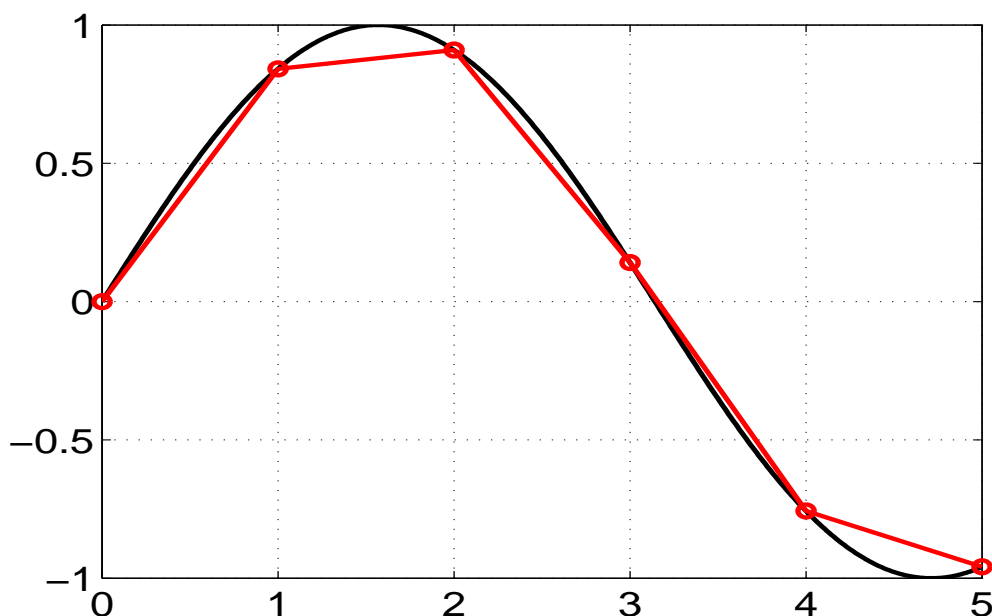


Рис. 2: Функция (черная линия) и ее кусочно-линейная аппроксимация для $N = 5$ (красная линия).

понимание того факта, что законы природы можно записывать на языке дифференциальных уравнений (т.е. уравнений, связывающих значения неизвестных функций и их производных). Решая эти дифференциальные уравнения, можно предсказывать и анализировать различные процессы в природе. Большинство дифференциальных уравнений невозможно решить аналитически (т.е. выписать функцию, которая является их решением), а даже в тех редких случаях, когда это сделать можно, полученные таким путем решения бывают громоздки и сложны для понимания. Поэтому дифференциальные уравнения часто решают приближенно с использованием компьютеров. В памяти компьютера, конечно, невозможно хранить значения функции на всем бесконечном множестве точек отрезка $[a, b]$, и потому всегда приходится ограничиться вычислением значений функции на каком-либо конечном множестве точек.⁸ Та-

⁸Для простоты мы будем рассматривать лишь случай, когда точки лежат на равных расстояниях друг от друга.

ким образом, с помощью компьютера на самом деле можно вычислять не сами функции, заданные на всем отрезке, а их кусочно-линейные аппроксимации, определяющиеся значениями в конечном множестве точек $x_0 = a, x_1, x_2, \dots, x_N = b$.

8.2 Аппроксимация производных

Чтобы записать дифференциальное уравнение для кусочно линейной функции, заданной своими значениями $f_i = f(x_i)$ на множестве $x_0 = a, x_1, x_2, \dots, x_N = b$ (причем $x_i - x_{i-1} = \Delta x = (a - b)/N$ для всех i), необходимо уметь вычислять производные по значениям f_i . Первую производную функции можно вычислить, используя определение производной:

$$f'(x_i) = \lim_{\Delta x \rightarrow 0} \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x} \approx \frac{f(x_{i+1}) - f(x_i)}{\Delta x} = \frac{f_{i+1} - f_i}{\Delta x}.$$

Приближенное равенство выполняется, если Δx достаточно мало, т.е. если число точек N выбрано достаточно большим. Общий способ получения аппроксимаций производных состоит в рассмотрении рядов Тейлора для функции $f(x)$ (конечно, для этого необходимо предположить, что f -дифференцируемая нужное число раз функция). Рассмотрим ряды Тейлора для $f(x)$ в точке $x = x_i$.

$$f(x_{i+1}) = f(x_i + \Delta x) = f(x_i) + \frac{f'(x_i)}{1!} \Delta x + \frac{f''(x_i)}{2!} \Delta x^2 + \frac{f^{(3)}(x_i)}{3!} \Delta x^3 + O(\Delta x^4),$$

$$f(x_{i-1}) = f(x_i - \Delta x) = f(x_i) - \frac{f'(x_i)}{1!} \Delta x + \frac{f''(x_i)}{2!} \Delta x^2 - \frac{f^{(3)}(x_i)}{3!} \Delta x^3 + O(\Delta x^4).$$

Записанные здесь выражения называются рядами Тейлора для функции $f(x)$ с остаточным членом в форме Пеано. В этой записи используются выражения типа $O(\Delta x^n)$, которое означает “бесконечно малые порядка Δx^n или более высокого порядка при $\Delta x \rightarrow 0$ ”. Прибавим к первому равенству второе:

$$f(x_{i+1}) + f(x_{i-1}) = 2f(x_i) + 2 \frac{f''(x_i)}{2!} \Delta x^2 + O(\Delta x^4),$$

или, что то же самое

$$\frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1}))}{\Delta x^2} = f''(x_i) + O(\Delta x^2). \quad (6)$$

Уравнение (6) представляет собой разностную аппроксимацию второго порядка для второй производной от функции $f(x)$ в точке x_i . То, что

аппроксимация имеет второй порядок, означает, что равенство выполняется с точностью до членов порядка Δx^2 . Рассмотрим теперь аппроксимацию первой производной. Для этого вычтем из первого ряда второй, получим

$$f(x_{i+1}) - f(x_{i-1}) = 2\frac{f'(x_i)}{1!}\Delta x + 2\frac{f^{(3)}(x_i)}{3!}\Delta x^3 + O(\Delta x^4),$$

$$\frac{f(x_{i+1}) - f(x_{i-1}))}{2\Delta x} = f'(x_i) + O(\Delta x^2). \quad (7)$$

Это – разностная аппроксимация второго порядка для первой производной $f'(x)$. В качестве упражнения проверьте, что полученная ранее аппроксимация первой производной имеет первый порядок. Чем выше порядок разностных аппроксимаций, тем точнее вычисления производных. Однако аппроксимации более высоких порядков требуют гораздо больших ресурсов при вычислении. Поэтому выбора порядка аппроксимации – это всегда компромисс между скоростью и точностью расчетов, приемлемый в данной задаче.

8.3 Матрицы разностных операторов

Так как набор значений f_0, f_1, \dots, f_N функции $f(x)$ в точках множества $x_0 = a, x_1, x_2, \dots, x_N = b$ можно считать вектором столбцом

$$\bar{f} = \begin{pmatrix} f_0 \\ f_1 \\ \dots \\ f_i \\ \dots \\ f_N \end{pmatrix},$$

то операторы разностного дифференцирования (т.е. приближенного нахождения первой и второй производных в точках $x_0 = a, x_1, x_2, \dots, x_N = b$ по данным значениям функции) можно задать матрицей. При нахождении разностной производной второго порядка (6), например, в результате умножения \bar{f} на матрицу оператора должен получаться вектор $D_2\bar{f}$, равный

$$D_2\bar{f} = \begin{pmatrix} \frac{f_1 - 2f_0 - f_{-1}}{\Delta x^2} \\ \frac{f_2 - 2f_1 + f_0}{\Delta x^2} \\ \dots \\ \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2} \\ \dots \\ \frac{f_{N+1} - 2f_N + f_{N-1}}{\Delta x^2} \end{pmatrix}.$$

Будем считать, что за границами отрезка функция равна нулю, т.е. что $f_{-1} = f_{N+1} = 0$. Тогда оператор, как легко понять, можно задать следующей матрицей:

$$\begin{pmatrix} \frac{f_1 - 2f_0 - f_{-1}}{\Delta x^2} \\ \frac{f_2 - 2f_1 - f_0}{\Delta x^2} \\ \dots \\ \frac{f_{i+1} - 2f_i - f_{i-1}}{\Delta x^2} \\ \dots \\ \frac{f_{N+1} - 2f_N - f_{N-1}}{\Delta x^2} \end{pmatrix} = \frac{1}{\Delta x^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \dots \\ f_i \\ \dots \\ f_N \end{pmatrix}$$

В полученной матрице элементы на главной диагонали равны -2 , элементы поддиагонали равны 1 , элементы наддиагонали равны 1 . Такая матрица называется трехдиагональной. Чтобы хранить такую матрицу, вообще говоря, достаточно хранить лишь $3n - 2$ элемента этих диагоналей, все n^2 чисел матрицы хранить необязательно. Метод Гаусса для такой матрицы также имеет очень простой вид и носит название прогонки. В качестве упражнения выпишите матричную форму оператора (7). Как видите, операторы дифференцирования можно приближенно заменить матрицами, если функция задана на конечном множестве точек.

8.4 Приложение: уравнение маятника

Для начала рассмотрим очень простой пример: уравнение гармонических колебаний вида

$$\frac{d^2x}{dt^2} + \omega^2 x = 0,$$

где $x(t)$ – функция, описывающая зависимость координаты тела x от времени t . Для простоты положим, что $\omega = 1$. Предположим, что при $t = 0$ тело находилось в точке $x = 0$, а при $t = \frac{3}{2}\pi$ тело оказалось в точке $x = -1$. Установим зависимость координаты тела от времени между этими двумя моментами. Довольно очевидно, что решением этой задачи является функция $x(t) = \sin(t)$. Мы с вами получим ее решение численно, сведя задачу к системе линейных уравнений с использованием матричного приближения для оператора дифференцирования. Для этого разобьем отрезок $[0, \frac{3}{2}\pi]$ на N интервалов длины $\Delta t = \frac{3}{2}\pi/N$ с помощью точек $t_0 = 0, t_1, t_2, \dots, t_N = \frac{3}{2}\pi$. Для определенности будем считать, что $N = 20$. Нам известно, что $x_0 = 0, x_{20} = \frac{3}{2}\pi$. Во всех остальных точках (т.е. точках t_1, \dots, t_{N-1}) дифференциальное уравнение с помощью (6) можно записать в виде:

$$\frac{x(t_{i+1}) - 2x(t_i) + x(t_{i-1}))}{\Delta t^2} + x(t_i) = 0,$$

$$\frac{x_{i+1} - 2x_i + x_{i-1}}{\Delta t^2} + x_i = 0.$$

В матричном виде это соотношение можно записать так:

$$\frac{1}{\Delta t^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_{19} \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_{19} \end{pmatrix} = \begin{pmatrix} 0 \\ \dots \\ 0 \\ \frac{1}{\Delta t^2} \end{pmatrix}.$$

($\frac{1}{\Delta t^2}$ появилось в правой части из соотношения $\frac{x_{20}-2x_{19}+x_{18}}{\Delta t^2} + x_{19} = 0$ с учетом того, что $x_{20} = -1$: $\frac{-2x_{19}+x_{18}}{\Delta t^2} + x_{19} = \frac{1}{\Delta t^2}$). Теперь можно внести второй вектор-столбец в матрицу:

$$\frac{1}{\Delta t^2} \begin{pmatrix} -2 + \Delta t^2 & 1 & 0 & \dots & 0 \\ 1 & -2 + \Delta t^2 & 1 & \dots & 0 \\ 0 & 1 & -2 + \Delta t^2 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -2 + \Delta t^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_{19} \end{pmatrix} = \begin{pmatrix} 0 \\ \dots \\ 0 \\ \frac{1}{\Delta t^2} \end{pmatrix}. \quad (8)$$

Матричное соотношение (8) представляет собой систему из 19 линейных уравнений с 19 неизвестными, имеющую единственное решение. Коэффициенты матрицы окончательно определяются путем подстановки в уравнение значения Δt . Систему (8) легко решить методом Гаусса. Естественно, это лучше делать на компьютере. Результат для $N = 10$ представлен на рис. 3. Видно, что даже для совсем небольшого числа точек результат получается очень близким к точному решению.

8.5 Приложение: уравнение Шредингера

Рассмотрим теперь более сложный пример использования матричных операторов разностного дифференцирования. Получим с их помощью решение уравнения Шредингера:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi, \quad (9)$$

где

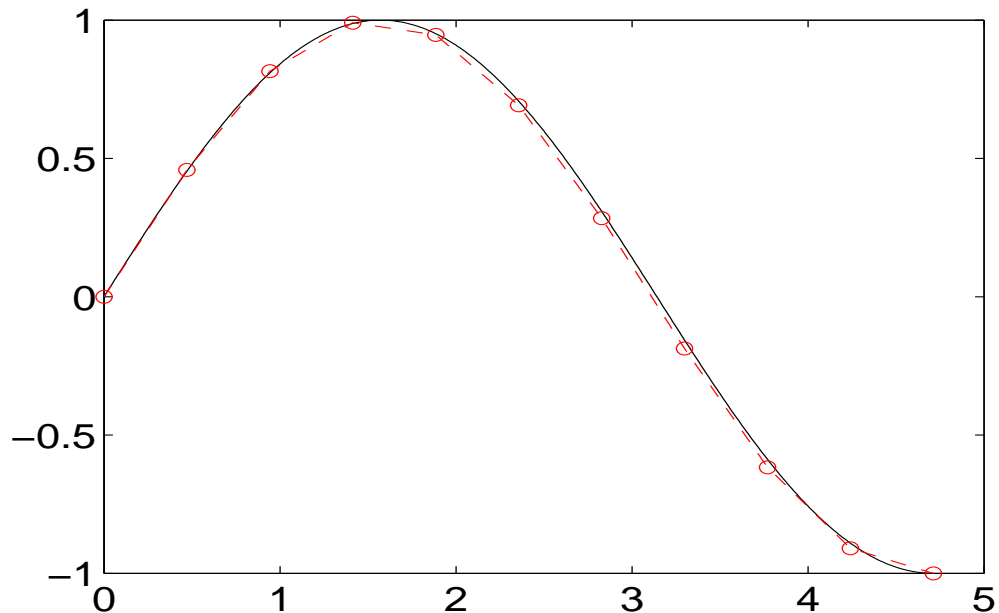


Рис. 3: Численное решение уравнения маятника для $N = 10$ (красная линия) и точное решение (черная линия).

8.6 Создание разреженных матриц и решение разностных уравнений в MATLAB

8.7 Задачи

1. Решите систему (8) вручную для $N = 3$, $N = 4$, используя калькулятор для выполнения шагов метода Гаусса. Ограничьтесь четырьмя цифрами после запятой.
2. Решите систему (8) с помощью MATLAB, составляя разреженные матрицы так, как это описано в последней части лекции. Задавайте различные значения переменной N и проследите, как меняется точность аппроксимации, рисуя на одном графике точное и приближенное решения.

Список литературы

1. Стренг Г. Линейная алгебра и ее применения. Москва: Мир, 1980. 459 с.