Тема 5. Элементы теории матриц

Предполагается, что студенты имеют знания в рамках стандартного вузовского курса линейной алгебры. В частности, предполагается, что известно понятие определителя матрицы и знакомы его свойства. Здесь напомним лишь некоторые известные сведения, а затем уже дополним их новыми. Более подробно можно ознакомиться с ними в Приложении 3 учебного пособия.

Матрицей называется совокупность $n \times m$ скаляров a_{ij}, образующих прямоугольную таблицу из n строк и m столбцов

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nm} \end{pmatrix}$$

Скаляры a_{ij} (i=1,2,...,n, j=1,2,....,m) называются элементами или компонентами матрицы. Первый индекс всегда означает номер строки или число строк, а второй - номер столбца или число столбцов.

Определение матрицы, как единого алгебраического объекта **A**, принадлежит английскому математику Артуру Кели (Arthur Cayley). Примером математических рассуждений, где появляется такой алгебраический объект, может служить преобразование координат вектора при изменении базиса линейного пространства.

Известные алгебраические объекты, такие как скаляры и векторы можно считать частными случаями матриц: для скаляров n=m=1, для векторов-строк n=1, для векторов-столбцов m=1. Матрица **0** является нулевой, если все ее элементы представлены нулями. При m=n матрица становится квадратной. Элементы a_{11}, a_{22}, a_{33}, ..., a_{nn} образуют главную диагональ, элементы $a_{1,n}$, $a_{2,n\text{-}1}$, $a_{3,n\text{-}2}$, ..., $a_{n,1}$ - побочную. Сумма диагональных элементов образует *след матрицы*:

$$Tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \; .$$

Если все элементы матрицы, кроме тех, что на главной диагонали, равны нулю, матрица называется *диагональной* и обозначается **D**=diag(**A**)=diag(a_{11},a_{22},....,a_{nn}). Диагональная матрица с компонентами равными единице является *единичной матрицей* и обозначается буквами **E** или **I**. Если все элементы матрицы выше главной диагонали равны нулю, то такая матрица называется *левой треугольной* или *нижней треугольной*. Аналогично, если все элементы матрицы ниже главной диагонали равны нулю, то такая матрица называется *правой треугольной* или *верхней треугольной*.

Если рассматривать матрицу как сложный объект, состоящий из скаляров, то она может быть подвергнута некоторым преобразованием, в частности, перестановке местами строк и столбцов:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nm} \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{21} & \dots & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & \dots & a_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & \dots & a_{nm} \end{pmatrix}.$$

Новая матрица называется транспонированной, а сама операция – транспонированием: $\mathbf{A} \Rightarrow \mathbf{A}^{T}$. Ясно, что $(\mathbf{A}^{T})^{T}=\mathbf{A}$. Важным классом матриц являются симметричные, а именно те, что не меняются при транспонировании, т.е. когда $\mathbf{A}^{T}=\mathbf{A}$.

На множестве матриц, прежде всего, вводятся операции сложения элементов матриц и умножения элементов матриц на скаляры. Обе операции определяются покомпонентно, т.е. при сложении складываются элементы с одинаковыми индексами (и это накладывает ограничение на операцию в том смысле, что могут складываться только матрицы с одинаковым числом строк и столбцов), а при умножении на скаляр умножаются все компоненты.

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & \dots & b_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & \dots & b_{nm} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & \dots & a_{2m}+b_{2m} \\ & & \dots & \dots & \\ & & \dots & \dots & \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & \dots & a_{nm}+b_{nm} \end{pmatrix},$$

$$\alpha\mathbf{A} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \dots & \alpha a_{1m} \\ \alpha a_{21} & \alpha a_{22} & \dots & \dots & \alpha a_{2m} \\ & & \dots & \dots & \\ & & \dots & \dots & \\ \alpha a_{n1} & \alpha a_{n2} & \dots & \dots & \alpha a_{nm} \end{pmatrix}.$$

Обе операции являются дистрибутивными.

$$(\alpha+\beta)\mathbf{A} = \alpha\mathbf{A}+\beta\mathbf{A}$$

$$\alpha(\mathbf{A}+\mathbf{B}) = \alpha\mathbf{A}+\alpha\mathbf{B}.$$

Следующая операция – умножение матриц. Компоненты c_{ij} матрицы произведения $\mathbf{C}_{np}$ вычисляются как скалярные произведения строк левой матрицы $\mathbf{A}_{\mathbf{nm}}$ на столбцы правой матрицы $\mathbf{B}_{mp}$:

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Легко убедиться, что введенная операция ассоциативна: $(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}=\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$. Но она не коммутативна. Убедиться в некоммутативности можно, даже не выполняя никаких операций. Например, при умножении матрицы $\mathbf{A}_{24}$ на матрицу $\mathbf{B}_{42}$ получится матрица $\mathbf{C}_{22}$, а при умножении $\mathbf{B}_{42}\cdot\mathbf{A}_{24}=\mathbf{C}_{44}$. Более того, для матриц $\mathbf{A}_{24}$ и $\mathbf{B}_{43}$ вторую матрицу на первую умножать вообще нельзя.

Частным случаем является умножение матрицы на вектор:

$$\mathbf{Ax}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} a_{11}x_1+a_{12}x_2 \\ a_{21}x_1+a_{22}x_2 \end{pmatrix}.$$

Здесь и в дальнейшем будем употреблять термин «вектор» применительно к вектору-столбцу.

Теперь обратимся к обратному элементу для операции умножения матриц. Речь может идти лишь о квадратных матрицах. Обратной матрицей, обозначаемой $\mathbf{A}^{-1}$, называется матрица $\mathbf{X} = \mathbf{A}^{-1}$, удовлетворяющая уравнениям

$$\mathbf{A}\cdot\mathbf{X}=\mathbf{E}, \qquad \mathbf{X}\cdot\mathbf{A}=\mathbf{E}.$$

Для нахождения обратной матрицы может быть использован, например, метод Гаусса. Пусть $\mathbf{x_k}$ – это столбец матрицы $\mathbf{X}$ с номером k, а $\mathbf{e_k}$ – это столбец матрицы $\mathbf{E}$ с номером k. Тогда необходимо m раз решить систему уравнений

$$\mathbf{A}\cdot\mathbf{x_k} = \mathbf{e_k}$$

для различных k ($m\times m$ – размерность матрицы). Как известно, эта система имеет единственное решение, а, следовательно, обратная матрица существует и является единственной, если определитель матрицы не равен нулю ($\det(\mathbf{A})\neq 0$).Если определитель оказывается равным нулю, матрицу называют *особенной* или *вырожденной*, в противном случае – *неособенной* или *невырожденной*.

Пример. Пусть матрицы **A, B, C** квадратные и неособенные. Требуется найти матрицы **A** и **B**, если $\mathbf{C} = \mathbf{A}\cdot\mathbf{B}$.

$$\mathbf{B} = \mathbf{A}^{-1}\cdot\mathbf{C}; \qquad \mathbf{A} = \mathbf{C}\cdot\mathbf{B}^{-1}$$

Для эффективного вычисления обратной матрицы на практике используется **LU**-разложение. **LU**-разложением матрицы A называется представление матрицы в следующем виде

$$\mathbf{A} = \mathbf{L}\cdot\mathbf{U},$$

где **L** – нижняя треугольная матрица с единицами на главной диагонали, а **U** – верхняя треугольная матрица. Если **LU**-разложение построено, то решение системы

$$\mathbf{Ax} = \mathbf{b} \tag{*}$$

сводится к решению двух систем с треугольными матрицами

$$\mathbf{Ly} = \mathbf{b}, \qquad \mathbf{Ux} = \mathbf{y} \qquad (**)$$

Если трудоемкость решения системы (*) пропорциональна m^3, то трудоемкость решения двух систем (**) пропорциональна лишь m^2.

Для эффективного нахождения обратной матрицы могут быть использованы программы ***DECOMP*** и ***SOLVE***. Первая из них строит **LU**-разложение, а вторая решает две системы (**). Программы имеют следующие параметры:

DECOMP (***NDIM***, ***N***, ***A, COND, IPVT, WORK***),

SOLVE(***NDIM, N, A, B, IPVT***),

где ***NDIM*** – объявленная в описании строчная размерность массива, в котором располагается матрица **A**;

N – порядок системы уравнений;

A – матрица, подвергающаяся разложению (по окончании работы программы на ее месте располагаются матрицы **L** и **U**),

COND – оценка числа обусловленности;

IPVT – вектор индексов ведущих элементов (размерность ***N***);

WORK– рабочий одномерный массив (размерность ***N***),

B – вектор правых частей системы (*), где по окончании работы программы ***SOLVE*** размещается вектор решения **x**.

К этим программам и к процедуре построения **LU**-разложения вернемся на следующих лекциях, а сейчас рассмотрим схему эффективного использования программ для нахождения $\mathbf{A}^{-1}$.

<u>*Однократно*</u> вызывается программа ***DECOMP*** и строится **LU**-разложение матрицы **A**. Далее в цикле m раз вызывается программа ***SOLVE***:

Цикл по k от 1 до m

в вектор **b** *системы* (*) *записывается* $\mathbf{e_k}$;

вызывается программа ***SOLVE*** (*на выходе в векторе* **b** *располагается k-й столбец* $\mathbf{x_k}$ *обратной матрицы*);

содержимое вектора **b** *записывается в k-й столбец обратной матрицы.*

конец цикла

Собственные значения матрицы и собственные векторы

Рассмотрим линейное преобразование с квадратной матрицей **A** размерности $n \times n$

$$\mathbf{y} = \mathbf{Au}.$$

Поставим задачу: для заданной матрицы **A** найти такие векторы **u**, которые сохраняют свое направление после линейного преобразования. Другими словами, каковы должны быть векторы **u**, чтобы вектор **y** был пропорционален или коллинеарен вектору **u**, т.е. $\mathbf{y} = \lambda\,\mathbf{u}$. Тогда

$$\mathbf{Au} = \lambda\mathbf{u} \quad \text{или} \quad (\mathbf{A} - \lambda\mathbf{E})\mathbf{u} = \mathbf{0}. \tag{1}$$

Числа λ и векторы **u**, удовлетворяющие этому уравнению, получили название *собственные значения* и *собственные векторы* соответственно. Равенство (1) представляет собой однородную линейную алгебраическую систему относительно искомых компонентов вектора **u**. Известно, что, если определитель матрицы $\mathbf{A} - \lambda\mathbf{E}$ не равен нулю, то система имеет единственное тривиальное нулевое решение **u**=**0.** Чтобы система (1) имела нетривиальные решения, нужно потребовать, чтобы ее определитель был равен нулю: $\det(\mathbf{A} - \lambda\mathbf{E}) = 0$. Запишем это равенство в покомпонентном виде:

$$\det\left(\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & \dots & a_{2m} \\ & & \dots & \dots & \\ & & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & \dots & a_{nm} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & & \lambda \end{pmatrix}\right) = \, = \det\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & \dots & a_{1m} \\ a_{21} & a_{22}-\lambda & \dots & \dots & a_{2m} \\ & & \dots & \dots & \\ & & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & \dots & a_{nm}-\lambda \end{pmatrix} = 0$$

Вычисление этого определителя дает:

$$\det(\mathbf{A} - \lambda\mathbf{E}) = (-1)^n \lambda^n + b_1\lambda^{n-1} + \dots + b_{n-1}\lambda + b_n = 0. \tag{2}$$

Полином, стоящий в левой части этого уравнения, называется характеристическим полиномом, а само уравнение характеристическим уравнением. Оно имеет ровно n корней с учетом их кратности. Это и есть собственные значения матрицы $\lambda_1, \lambda_2, \dots, \lambda_n$. Они составляют спектр матрицы **A**, а величина $\rho(\mathbf{A}) = \max\limits_i |\lambda_i|$, $(i = 1, 2, \dots, n)$ называется *спектральным радиусом*.

Для каждого λ_i в соответствии с (1) можно найти решение $\mathbf{u}_i$ однородной системы $\mathbf{Au}_i = \lambda_i\,\mathbf{u}_i$, т.е. собственный вектор. Если все собственные значения различны, то $\mathbf{u}_i$ образуют систему из n различных векторов.

Поскольку собственные векторы находятся из решения однородной системы, то и известными они оказываются с точностью до постоянного (ненулевого) множителя, т.е. собственные векторы однозначно определены по направлению, но их длины (нормы) остаются произвольными. Часто бывает удобно приводить векторы к единичной длине, т.е. нормировать их.

Приведем пример вычисления собственных значений и собственных векторов матрицы.

$$\mathbf{A} = \begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix}$$

Собственные значения:

$$\det(A - \lambda E) = \det\begin{pmatrix} -51-\lambda & -49 \\ -50 & -50-\lambda \end{pmatrix} = \lambda^2 + 101\lambda + 100 = 0.$$

$$\lambda_1 = -100, \quad \lambda_2 = -1.$$

Собственные векторы:

$$\begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix}\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = -100\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}, \qquad \begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix}\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = -1\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

$$-49u_{12} = -49u_{11} \qquad -50u_{21} = 49u_{22}$$

$$u_{12} = u_{11} \qquad u_{21} = \left(\frac{-49}{50}\right)u_{22}$$

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}u_{11} \qquad \mathbf{u}_2 = \begin{pmatrix} \frac{-49}{50} \\ 1 \end{pmatrix}u_{22}$$

Нормируем собственные векторы:

$$\|\mathbf{u}_1\| = \sqrt{(\mathbf{u}_1, \mathbf{u}_1)} = \sqrt{2}u_{11} \qquad \|\mathbf{u}_2\| = \sqrt{(\mathbf{u}_2, \mathbf{u}_2)} = \sqrt{\left(\frac{49}{50}\right)^2 + 1}\ u_{21}$$

$$\mathbf{u}_1^{\text{норм}} = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathbf{u}_2^{\text{норм}} = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \sqrt{\frac{2500}{4901}}\begin{pmatrix} \frac{-49}{50} \\ 1 \end{pmatrix}$$

Сформулируем ряд свойств матриц различного вида. Их доказательство оставим для самостоятельных упражнений. В дальнейшем **D** – диагональная матрица.

Задачи на матрицы.

1. $\left(\mathbf{AB}\right)^T = \mathbf{B}^T\mathbf{A}^T$

2. $\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

3. $\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$

4. При умножении матрицы **A** на диагональную матрицу **D** *слева* **B**=**DA** все *строки* **A** умножаются на соответствующие диагональные элементы

5. При умножении матрицы **A** на диагональную матрицу **D** справа **B**=**AD** все *столбцы* **A** умножаются на соответствующие диагональные элементы. Как частный случай, имеем равенства **A**=**EA=AE**.

6. При умножении треугольных матриц одного вида (например, левых треугольных) получается матрица того же вида (т.е. левая треугольная).

7. Обратная матрица для треугольной является треугольной матрицей того вида.

8. Собственные значения диагональной матрицы равны ее диагональным элементам.

9. Собственные значения треугольной матрицы равны ее диагональным элементам.

10. Сумма собственных значений матрицы равна сумме ее диагональных элементов (*следу матрицы*: $Tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$), а произведение всех собственных значений матрицы равно ее определителю.