ЛЕКЦИЯ 7

РАНГ МАТРИЦЫ

КРИТЕРИЙ СОВМЕСТНОСТИ

МАТРИЦЫ И ОТОБРАЖЕНИЯ

РАНГ МАТРИЦЫ

В векторном пространстве $\mathbb{R}^m$ столбцов высоты m рассмотрим n векторов

$$A^{(j)} = [a_{1j}, a_{2j}, \ldots, a_{mj}], \quad j = 1, 2, \ldots, n,$$

и их линейную оболочку

$$V = \langle A^{(1)}, A^{(2)}, \ldots, A^{(n)} \rangle.$$

Пусть дан еще один вектор $B = [b_1, b_2, \ldots, b_m]$. Спрашивается, принадлежит ли вектор B линейной оболочке $V \subset \mathbb{R}^m$, а если принадлежит, то каким образом его координаты $b_1, \ldots, b_m$ выражаются через координаты векторов $A^{(j)}$?

Мы берем линейную комбинацию векторов $A^{(j)}$ с произвольными (неизвестными) коэффициентами x_j и составляем уравнение

$$x_1 A^{(1)} + \cdots + x_n A^{(n)} = B.$$

Наглядный вид этого уравнения:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

В матричном виде это не что иное как запись системы из m линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

Именно такой системой мы занимались на первых лекциях. Там же были введены простая и расширенная матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ и } (A|B) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Назовем *пространством столбцов* прямоугольной матрицы A размера $m \times n$ введенную выше линейную оболочку

$$V = \langle A^{(1)}, A^{(2)}, \ldots, A^{(n)} \rangle.$$

Будем пока наше пространство V обозначать через $V_{vert}(A)$ или просто через V_{vert}.

Размерность

$$r_{vert}(A) = \dim V_{vert}$$

назовем *рангом по столбцам* матрицы A.

Аналогично вводится *ранг по строкам* матрицы A:

$$r_{horiz}(A) = \dim V_{horiz},$$

где

$$V_{horiz} = \langle A_{(1)}, A_{(2)}, \ldots, A_{(m)} \rangle$$

— пространство строк матрицы A, то есть линейная оболочка в $\mathbb{R}^n$, натянутая на векторы-строки

$$A_{(i)} = \langle a_{i1}, a_{i2}, \ldots, a_{in} \rangle, \quad i = 1, 2, \ldots, m.$$

Таким образом, мы ввели ранги систем векторов-столбцов и векторов-строк соответственно.

Мы уже доказывали, что такие ранги определены корректно.

Будем говорить, что матрица A' получена из A элементарным преобразованием типа (I), если переход от одной матрицы к другой осуществляется просто переменой местами двух строк. Будем говорить, что от A к A' мы переходим элементарным преобразованием типа (II), если все строки, кроме одной, остаются неизменными, а к одной из строк прибавляется некоторая другая, умноженная на число λ.

Заметим, что элементарные преобразования обоих типов обратимы.

Лемма 1. *Если матрица A' получена из прямоугольной матрицы A путем применения конечной последовательности элементарных преобразований над строками, то имеют место равенства:*

(1) $r_{horiz}(A') = r_{horiz}(A)$;

(2) $r_{vert}(A') = r_{vert}(A)$.

Доказательство. Достаточно рассмотреть тот случай, когда A' получена из A путем применения одного элементарного преобразования.

(1) Так как

$$\langle A_{(1)}, \ldots, A_{(s)}, \ldots, A_{(t)}, \ldots, A_{(m)} \rangle = \\ = \langle A_{(1)}, \ldots, A_{(t)}, \ldots, A_{(s)}, \ldots, A_{(m)} \rangle,$$

то элементарное преобразование типа (I) не меняет ранг $r_{horiz}(A)$.

Далее,

$$A'_{(s)} = A_{(s)} + \lambda A_{(t)} \Longrightarrow A_{(s)} = A'_{(s)} - \lambda A_{(t)},$$

и, следовательно,

$$\langle A_{(1)}, \dots, A_{(s)} + \lambda A_{(t)}, \dots, A_{(t)}, \dots, A_{(m)} \rangle = \\ = \langle A_{(1)}, \dots, A_{(s)}, \dots, A_{(t)}, \dots, A_{(m)} \rangle,$$

так что $r_{horiz}(A)$ не меняется при элементарных преобразованиях типа (II).

(2) Пусть $A'^{(j)}$, $1 \leqslant j \leqslant n$, — столбцы матрицы A'. Докажем, что

$$\sum_{j=1}^{n} \lambda_j A^{(j)} = 0 \Longleftrightarrow \sum_{j=1}^{n} \lambda_j A'^{(j)} = 0.$$

С этой целью рассмотрим две линейные однородные системы $\mathcal{A}$ и $\mathcal{A}'$ с матрицами A и A' соответственно, записанные так:

$$\mathcal{A} : \sum_{j=1}^{n} x_j A^{(j)} = 0, \qquad \mathcal{A}' : \sum_{j=1}^{n} x_j A'^{(j)} = 0.$$

Матрицы A и A' у нас таковы, что $\mathcal{A}'$ получается из $\mathcal{A}$ элементарным преобразованием типа (I) или (II). Мы доказывали, что в таком случае однородные системы эквивалентны, то есть всякое решение $(\lambda_1, \dots, \lambda_n)$ одной системы является решением другой, и наоборот. Это и есть та равносильность, которую мы доказывали.

Таким образом, всякой, в том числе и максимальной, независимой системе столбцов одной матрицы будет отвечать независимая система столбцов с теми же номерами другой матрицы, чем и устанавливается необходимое равенство $r_{vert}(A) = r_{vert}(A')$. $\square$

Основным результатом этой лекции является следующее утверждение:

Теорема 1. *Для любой прямоугольной $m \times n$-матрицы A справедливо равенство $r_{vert}(A) = r_{horiz}(A)$ (это число называется рангом матрицы и обозначается $\operatorname{rank} A$).*

Доказательство. Мы уже доказывали, что конечным числом элементарных преобразований типа (I) и (II), совершенных над строками $A_{(i)}$, матрицу A можно привести к ступенчатому виду

$$\overline{A} = \begin{pmatrix} \overline{a}_{11} & \dots & \overline{a}_{1k} & \dots & \overline{a}_{1l} & \dots & \overline{a}_{1s} & \dots & \overline{a}_{1n} \\ 0 & \dots & \overline{a}_{2k} & \dots & \overline{a}_{2l} & \dots & \overline{a}_{2s} & \dots & \overline{a}_{2n} \\ 0 & \dots & 0 & \dots & \overline{a}_{3l} & \dots & \overline{a}_{3s} & \dots & \overline{a}_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & \overline{a}_{rs} & \dots & \overline{a}_{rn} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

с $\overline{a}_{11}\overline{a}_{2k}\overline{a}_{3l}\dots\overline{a}_{rs} \neq 0$.

Согласно предыдущей лемме

$$r_{vert}(A) = r_{vert}(\overline{A}), \qquad r_{horiz}(A) = r_{horiz}(\overline{A}),$$

так что нам достаточно доказать только равенство

$$r_{vert}(\overline{A}) = r_{horiz}(\overline{A}).$$

Столбцы матриц A и $\overline{A}$ с номерами $1, k, l, \ldots, s$, отвечающими главным неизвестным $x_1, x_k, x_l, \ldots, x_s$ нашей линейной системы, будем называть *базисными столбцами.*

Эта терминология вполне оправдана. Преположив наличие соотношения

$$\lambda_1 \overline{A}^{(1)} + \lambda_r \overline{A}^{(k)} + \lambda_l \overline{A}^{(l)} + \cdots + \lambda_s \overline{A}^{(s)} = 0,$$

связывающего векторы-столбцы

$$\overline{A}^{(1)} = [\overline{a}_{11}, 0, \ldots, 0], \quad \overline{A}^{(k)} = [\overline{a}_{1k}, \overline{a}_{2k}, 0, \ldots, 0],$$
$$\ldots, \quad \overline{A}^{(s)} = [\overline{a}_{1s}, \overline{a}_{2s}, \ldots, \overline{a}_{rs}, 0, \ldots, 0]$$

нашей матрицы, получим

$$\lambda_s \overline{a}_{rs} = 0, \quad \ldots, \quad \lambda_l \overline{a}_{3l} = 0, \quad \lambda_k \overline{a}_{2k} = 0, \quad \lambda_1 \overline{a}_{11} = 0,$$

а так как $\overline{a}_{11}\overline{a}_{2k}\overline{a}_{3l}\ldots\overline{a}_{rs} \neq 0$, то $\lambda_1 = \lambda_k = \lambda_l = \cdots = \lambda_s = 0$.

Значит, $\operatorname{rank}\{\overline{A}^{(1)}, \overline{A}^{(k)}, \overline{A}^{(l)}, \ldots, \overline{A}^{(s)}\} = r$ и $r_{vert}(\overline{A}) \geqslant r$.

Но пространство $\overline{V}_{vert}$, порожденное столбцами матрицы $\overline{A}$, отождествляется с пространством столбцов матрицы, которая получается из $\overline{A}$ удалением последних $m - r$ нулевых строк.

Поэтому

$$r_{vert}(\overline{A}) = \dim \overline{V}_{vert} \leqslant \mathbb{R}^r = r.$$

Сопоставление двух неравенств показывает, что $r_{vert}(\overline{A}) = r$.

С другой стороны, все ненулевые строки матрицы $\overline{A}$ линейно независимы: любое гипотетическое соотношение

$$\lambda_1 \overline{A}_{(1)} + \lambda_2 \overline{A}_{(2)} + \cdots + \lambda_r \overline{A}_{(r)} = 0, \quad \lambda_i \in \mathbb{R},$$

как и в случае со столбцами, дает последовательно

$$\lambda_1 \overline{a}_{11} = 0, \quad \lambda_2 \overline{a}_{2k} = 0, \quad \ldots, \quad \lambda_r \overline{a}_{rs} = 0,$$

откуда

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0.$$

Следовательно,

$$r_{horiz}(\overline{A}) = r = r_{vert}(\overline{A}).$$

$\square$

КРИТЕРИЙ СОВМЕСТНОСТИ

Ступенчатый вид матрицы A, дающий ответ на ряд вопросов относительно линейных систем, содержит элементы произвола: выбор базисных столбцов, выбор главных неизвестных и т.п.

В то же время из предыдущей теоремы сразу можно получить

Следствие 1. *Число главных неизвестных линейной системы не зависит от способа приведения ее к ступенчатому виду и равно* $\operatorname{rank} A$*, где* A *— матрица системы.*

Доказательство. Действительно, в процессе доказательств предыдущих утверждений мы видели, что число главных неизвестных равно числу ненулевых строк матрицы $\overline{A}$, совпадающему с рангом матрицы A. Ранг определялся нами совершенно инвариантным образом, то есть не зависел никак от способа приведения к ступенчатому виду. $\square$

Теперь сформулируем и докажем важную теорему.

Теорема 2 (Кронекер–Капелли). *Система линейных уравнений совместна тогда и только тогда, когда ранг ее матрицы совпадает с рангом расширенной матрицы.*

Доказательство. Совместность линейной системы можно трактовать (как уже говорилось в начале лекции) как вопрос о представлении вектора-столбца B свободных членов в виде линейной комбинации векторов-столбцов $A^{(j)}$ матрицы A.

Если такое представление возможно (то есть система совместна), то

$$B \in \langle A^{(1)}, \ldots, A^{(n)} \rangle$$

и

$$\operatorname{rank}\{A^{(1)}, \ldots, A^{(n)}\} = \operatorname{rank}\{A^{(1)}, \ldots, A^{(n)}, B\},$$

откуда

$$\operatorname{rank} A = r_{vert}(A) = r_{vert}(A|B) = \operatorname{rank}(A|B).$$

Обратно, если ранги матриц A и $A|B$ совпадают и

$$\{A^{(j_1)}, \ldots, A^{(j_r)}\}$$

— какая-то максимальная линейно независимая подсистема столбцов матрицы A, то расширенная система

$$\{A^{(j_1)}, \ldots, A^{(j_r)}, B\}$$

будет линейно зависимой, а это означает, что B — линейная комбинация базисных столбцов $A^{(j)}$.

Значит, система совместна. □

МАТРИЦЫ И ОТОБРАЖЕНИЯ

Пусть $\mathbb{R}^n$ и $\mathbb{R}^m$ — векторные пространства столбцов высоты n и m соответственно. Пусть, далее, $A = (a_{ij})$ — матрица размера $m \times n$.

Определим отображение

$$\varphi_A : \mathbb{R}^n \to \mathbb{R}^m,$$

полагая для любого

$$X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n,$$

что

$$\varphi_A(X) = x_1 A^{(1)} + x_2 A^{(2)} + \cdots + x_n A^{(n)},$$

где $A^{(1)}, A^{(2)}, \ldots, A^{(n)}$ — столбцы матрицы A. Так как они имеют высоту m, то в правой части стоит вектор-столбец

$$Y = [y_1, y_2, \ldots, y_m] \in \mathbb{R}^m.$$

Более подробно это выражение переписывается в виде

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \ldots, m.$$

Если

$$X = X' + X'' = [x_1' + x_1'', x_2' + x_2'', \dots, x_n' + x_n''],$$

то

$$\varphi_A(X' + X'') = \sum_{i=1}^{n}(x_i' + x_i'')A^{(i)} = \sum_{i=1}^{n} x_i' A^{(i)} + \sum_{i=1}^{n} x_i'' A^{(i)} = \\ = \varphi_A(X') + \varphi_A(X'').$$

Аналогично,

$$\varphi_A(\lambda X) = \sum_{i=1}^{n} \lambda x_i A^{(i)} = \lambda \sum_{i=1}^{n} A^{(i)} = \lambda \varphi_A(X), \quad \lambda \in \mathbb{R}.$$

Таким образом, для введенного отображения φ_A всегда выполняются два свойства:

(1) $\varphi(X' + X'') = \varphi(X') + \varphi(X'')$ для всех $X', X'' \in \mathbb{R}^n$;

(2) $\varphi(\lambda X)\lambda\varphi(X)$ для всех $X \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Определение 1. Отображение $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, обладающее свойствами (1) и (2), называется *линейным отображением* из $\mathbb{R}^n$ в $\mathbb{R}^m$.

Как мы знаем, $\mathbb{R}^n = \langle E^{(1)}, \ldots, E^{(n)} \rangle$ — линейная оболочка стандартных базисных столбцов, так что

$$X = [x_1, x_2, \ldots, x_n] = \sum_{j=1}^{n} x_j E^{(j)}.$$

Из свойств линейности имеем

$$\varphi(X) = \varphi \left(\sum_{j=1}^{n} x_j E^{(j)} \right) = \sum_{j=1}^{n} x_j \varphi(E^{(j)}).$$

Это соотношение показывает, что отображение φ полностью определяется своими значениями на базисных векторах-столбцах.

Положив

$$\varphi(E^{(j)}) = [a_{1j}, a_{2j}, \ldots, a_{ml}] = A^{(j)} \in \mathbb{R}^m,$$

видим, что задание φ равносильно заданию прямоугольной матрицы $A = (a_{ij})$ размера $m \times n$ со столбцами $A^{(1)}, \ldots, A^{(n)}$. Значит, можно положить $\varphi = \varphi_A$.

Будем называть такую соответствующую отображению матрицу *матрицей линейного отображения* φ_A.

Пусть φ_A и $\varphi_{A'}$ — два линейных отображения $\mathbb{R}^n \to \mathbb{R}^m$ с матрицами $A = (a_{ij})$ и $A' = (a'_{ij})$. Тогда равенство $\varphi_A = \varphi_{A'}$ равносильно совпадению значений

$$\varphi_A(X) = \varphi_{A'}(X) \text{ для всех } X \in \mathbb{R}^n.$$

В частности,

$$A'^{(j)} = \varphi_{A'}(E^{(j)}) = \varphi_A(E^{(j)}) = A^{(j)}, \quad 1 \leqslant j \leqslant n,$$

откуда $a'_{ij} = a_{ij}$ и $A' = A$.

Резюмируем полученные результаты:

Теорема 3. *Между линейными отображениями $\mathbb{R}^n$ в $\mathbb{R}^m$ и матрица размера $m \times n$ существует взаимно однозначное соответствие.*

Обратим внимание на специальный случай $m = 1$, когда линейное отображение $\mathbb{R}^n \to \mathbb{R}$, обычно называемое *линейной функцией* от n переменных, задается n скалярами $a_1, a_2, \ldots, a_n$:

$$\varphi(X) = \varphi(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

Линейные функции и произвольные линейные отображения $\mathbb{R}^n$ в $\mathbb{R}^m$ при фиксированных n и m можно складывать и умножать на скаляры.

В самом деле, пусть $\varphi_A, \varphi_B : \mathbb{R}^n \to \mathbb{R}^m$ — два линейных отображения. Отображение

$$\varphi = \alpha\varphi_A + \beta\varphi_B : \mathbb{R}^n \to \mathbb{R}^m, \quad \alpha, \beta \in \mathbb{R},$$

определяется своими значениями

$$\varphi(X) = \alpha\varphi_A(X) + \beta\varphi_B(X).$$

В правой части стоит обычная линейная комбинация векторов-столбцов.

Так как

$$\begin{gathered}\varphi(X' + X'') = \alpha\varphi_A(X' + X'') + \beta\varphi_B(X' + X'') = \\ = \alpha(\varphi_A(X') + \varphi_A(X'')) + \beta(\varphi_B(X') + \varphi_B(X'')) = \\ = (\alpha\varphi_A(X') + \beta\varphi_B(X')) + (\alpha\varphi_A(X'') + \beta\varphi_B(X'')) = \varphi(X') + \varphi(X'')\end{gathered}$$

и

$$\begin{multlined}\varphi(\lambda X) = \alpha\varphi_A(\lambda X) + \beta\varphi_B(\lambda X) = \alpha\lambda\varphi_A(X) + \beta\lambda\varphi_B(X) = \\ = \lambda(\alpha\varphi_A(X) + \beta\varphi_B(X)) = \lambda\varphi(X),\end{multlined}$$

то φ — линейное отображение. Значит, можно говорить о его матрице C: $\varphi = \varphi_C$.