

МГТУ ФН-12 МГТУ ФН-12 МГТУ ФН-12

Московский государственный технический университет
имени Н.Э. Баумана

Факультет «Фундаментальные науки»
Кафедра «Математическое моделирование»

А.Н. Канатников

**ИЗБРАННЫЕ ЛЕКЦИИ
ПО АЛГЕБРЕ**

Конспект лекций

Для студентов факультета ФН

Москва
2009

МГТУ ФН-12 МГТУ ФН-12 МГТУ ФН-12

4. ПСЕВДОРЕШЕНИЯ И ПСЕВДООБРАТНАЯ МАТРИЦА

4.1. Метод наименьших квадратов

Рассмотрим систему $Ax = b$ с матрицей A типа $m \times n$. Здесь x и b обозначают матрицы столбцы, которые мы будем рассматривать как элементы арифметических пространств $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно. Мы будем полагать, что в этих линейных пространствах задано стандартное скалярное произведение (сумма попарных произведений компонент), так что мы имеем дело с евклидовыми пространствами. Длину вектора x в евклидовом пространстве будем обозначать через $|x|$ (т.е./ как абсолютную величину числа). Матричное произведение Ax мы, как правило, будем интерпретировать как матричную запись линейной комбинации столбцов $a_1, \dots, a_n$ матрицы A , коэффициентами которой являются компоненты вектор-столбца x , т.е.

$$Ax = x_1 a_1 + x_2 a_2 + \dots + x_n a_n.$$

Выбрав вектор x , мы для выражения $b - Ax$ получим значение, которое равно нулю, если выбранный вектор есть решение системы. В общем случае величина $|b - Ax|$ характеризует, насколько вектор неизвестных x близок к решению системы. Компоненты вектора $b - Ax$ называют **невязками** уравнений системы. Сам вектор мы будем называть **вектором невязок**, а его длину — **нормой невязки**.

Если система несовместна, то в качестве решения можно выбрать такой вектор x , что величина $|b - Ax|^2$ примет наименьшее значение. Квадрат нормы в данном случае выбран потому, что он без радикала выражается через скалярное произведение. Описанный подход к решению системы линейных уравнений называют **методом наименьших квадратов**.

Задача поиска решения несовместной системы по методу наименьших квадратов по своему классу относится к задачам оптимизации: речь идет о поиске таких значений неизвестных x_i , при которых функция

$$f(x_1, \dots, x_n) = \sum_{i=1}^n (b_i - (a_{i1}x_1 + \dots + a_{in}x_n))^2$$

принимает наименьшее значение. Такую задачу можно решать общими методами исследования функций многих переменных на экстремум. В данном случае, однако, решение можно получить чисто геометрическими методами.

Рассмотрим подпространство $H = \text{span}(a_1, \dots, a_n) \subset \mathbb{R}^m$ — линейную оболочку столбцов матрицы A . Это подпространство представляет собой множество всевозможных линейных комбинаций Ax столбцов матрицы A с вектором коэффициентов x . Решение системы $Ax = b$ по методу наименьших квадратов означает выбор вектора $h \in H$, для которого величина $|b - h|$ принимает наименьшее значение, а также разложение этого вектора по системе столбцов $a_1, \dots, a_n$. Если система имеет решение, то $b \in H$. Иначе $b \notin H$.

Величина $\min_{h \in H} |b - h|$ называется **расстоянием от вектора b до подпространства H** . Геометрически (в V_2) вектор h , обеспечивающий минимум $|b - h|$, определяется основанием перпендикуляра, опущенного из точки на плоскость или прямую, т.е. $|b - h|^2$ имеет наименьшее значение, когда $b - h \perp H$. Покажем, что это условие выполняется в общем случае. Разложим вектор b на ортогональную проекцию b_0 на H и ортогональную $b^\perp$, т.е. представим в виде

$b = b_0 + b^\perp$, где $b_0 \in H$, $b^\perp \in H^\perp$. Тогда по теореме Пифагора

$$|b - h|^2 = |b_0 + b^\perp - h|^2 = |(b_0 - h) + b^\perp|^2 = |b_0 - h|^2 + |b^\perp|^2.$$

Отсюда следует, что $|b - h|^2$ принимает минимальное значение тогда, когда $b_0 - h = 0$, т.е. вектор h совпадает с проекцией вектора b на подпространство H . В этом случае $b - h = b^\perp \perp H$. Само минимальное значение совпадает с длиной ортогональной составляющей вектора b .

Теорема 4.1. Следующие условия эквивалентны:

- 1) величина $|b - Ax|$ минимальна для данного вектора $x \in \mathbb{R}^m$;
- 2) вектор x является решением системы $Ax = \text{пр}_H b$;
- 3) вектор x является решением системы $A^T Ax = A^T b$.

◀ Согласно проведенным рассуждениям величина $|b - Ax|$ минимальна тогда и только тогда, когда вектор $h = Ax$ является проекцией $\text{пр}_H b$ вектора b на H . Это будет в том случае, когда вектор $x \in \mathbb{R}^m$ является решением системы $Ax = \text{пр}_H b$. Тем самым доказана эквивалентность первых двух из трех условий.

Условие $Ax = \text{пр}_H b$ означает, что вектор $b - Ax$ принадлежит $H^\perp$. Это равносильно ортогональности вектора $b - Ax$ каждому вектору a_j , $j = \overline{1, n}$. Записав в матричном виде условия ортогональности, получим $a_j^T (b - Ax) = 0$, $j = \overline{1, n}$, или $A^T (b - Ax) = 0$. Последнее равенство эквивалентно равенству $A^T b = A^T Ax$, т.е. третьему условию в формулировке теоремы. ▶

Замечание. Матрица $A^T A$ представляет собой матрицу скалярных произведений $a_i^T a_j$ столбцов a_j , $j = \overline{1, n}$, матрицы A , называемую **матрицей Грама** системы векторов $a_1, \dots, a_n$.

Из доказанной теоремы вытекает ряд следствий:

1. Системы $Ax = 0$ и $A^T Ax = 0$ имеют одно и то же множество решений.
2. $\text{rang } A^T A = \text{rang } A$.
3. Матрица $A^T A$ невырождена тогда и только тогда, когда столбцы матрицы A линейно независимы.
4. Система векторов $a_1, a_2, \dots, a_k$ в евклидовом пространстве L линейно независима тогда и только тогда, когда матрица Грама этой системы невырождена.
5. Система $A^T Ax = A^T b$ всегда совместна; она является определенной тогда и только тогда, когда столбцы матрицы A линейно независимы.

Первое из сформулированных утверждений вытекает из того, что системы $Ax = \text{пр}_H b$ и $A^T Ax = A^T b$ имеют одно и то же множество решений (следовательно, и соответствующие однородные системы имеют одно и то же множество решений). Впрочем, это можно доказать и непосредственно (доказательство, представленное в учебнике). Третье следствие представляет собой частный случай второго следствия, когда системы имеют единственное решение. Однородная система имеет единственное решение в том и только в том случае, когда ранг матрицы системы совпадает с количеством неизвестных (количеством столбцов матрицы). Для квадратной матрицы $A^T A$ это равносильно невырожденности. Ну, и вообще равенство ранга матрицы количеству ее столбцов равносильно тому, что ее столбцы линейно независимы. Из третьего следствия вытекает пятое.

Чтобы обосновать четвертое следствие, достаточно записать координаты векторов в некотором ортонормированном базисе и составить из них матрицу A . Линейная независимость системы векторов равносильна линейной независимости столбцов матрицы A , что в силу следствия 3 равносильно невырожденности матрицы $A^T A$. Но эта матрица и есть матрица Грама системы векторов.

4.2. Псевдорешения

Любой столбец x , обеспечивающий минимальное значение нормы $|b - Ax|$ вектора невязки, называется **псевдорешением** системы $Ax = b$. Если система совместна, то минимальное значение нормы вектора невязки достигается на любом решении и равно нулю. В этом случае псевдорешения системы представляют собой решения системы.

Чтобы найти псевдорешения системы $Ax = b$, необходимо, согласно теореме 4.1 составить систему $A^T Ax = A^T b$, называемую **нормальной системой**, и найти все ее решения. Отметим, что система $Ax = b$ имеет единственное псевдорешение тогда и только тогда, когда у матрицы A столбцы линейно независимы или, иначе, ранг матрицы A равен количеству ее столбцов.

Если система $Ax = b$ имеет много псевдорешений, среди них выделяют то, которое имеет минимальную норму. Такое решение единственно и называется **нормальным псевдорешением**. В случае единственности псевдорешения понятия „псевдорешение“ и „нормальное псевдорешение“ имеют один и тот же смысл.

Теорема 4.2. Любая СЛАУ имеет нормальное псевдорешение, и притом единственное. Нормальное псевдорешение СЛАУ $Ax = b$ есть решение нормальной СЛАУ, являющееся линейной комбинацией столбцов матрицы A^T .

◀ Множество псевдорешений СЛАУ $Ax = b$ — это множество решений нормальной СЛАУ $A^T Ax = A^T b$, которое можно представить в виде $x = x_{\text{ч}} + x_0$, где $x_{\text{ч}}$ — частное решение нормальной СЛАУ, а x_0 — общее решение однородной СЛАУ $A^T Ax = 0$, совпадающее с общим решением СЛАУ $Ax = 0$. Нормальное псевдорешение определяется таким вектором x_0 , при котором норма $|x_{\text{ч}} + x_0|$ имеет минимальное значение. Из метода наименьших квадратов вытекает, что наименьшее значение норма $|x_{\text{ч}} + x_0|$ принимает, когда вектор $x_{\text{ч}} + x_0$ ортогонален подпространству $K \subset \mathbb{R}^m$ решений СЛАУ $Ax = 0$. Значит, единственный вектор вида $x_{\text{ч}} + x_0$, $x_0 \in K$, имеющий минимальную норму, является ортогональной составляющей x^* вектора $x_{\text{ч}}$ при его проектировании на K . Это доказывает, что нормальное псевдорешение существует и единственно, причем $x^* \in K^\perp$ может быть получен как ортогональная проекция на $K^\perp$ любого псевдорешения системы $Ax = b$.

Систему $Ax = 0$ можно интерпретировать как систему $(s_i, x) = 0$, $i = \overline{1, m}$, означающую, что K есть ортогональное дополнение к линейной оболочке системы $\{s_i\}$. Но тогда $K^\perp = \text{span}(s_i)$, а вектор $x^* \in K^\perp$ представим в виде линейной комбинации векторов s_i — столбцов матрицы A^T . ▶

Как найти нормальное псевдорешение? Можно поступить таким образом: записать нормальную СЛАУ $A^T Ax = A^T b$ и добавить условие $x \in K^\perp$, означающее, что $(x, f_j) = 0$, $j = \overline{1, k}$, где система векторов $\{f_j\}$ — фундаментальная система решений (ФСР) однородной СЛАУ $Ax = 0$. Последнее условие представляет собой систему линейных уравнений $F^T x = 0$, где F — матрица, составленная из вектор-столбцов f_j , $j = \overline{1, k}$. В результате получим систему уравнений

$$\begin{cases} A^T Ax = A^T b, \\ F^T x = 0, \end{cases} \quad \text{или} \quad \begin{pmatrix} A^T A \\ F^T \end{pmatrix} x = \begin{pmatrix} A^T b \\ 0 \end{pmatrix}.$$

Этот путь приведет к решению, но не всегда самый короткий. Можно указать другие способы решения.

Для псевдорешений сохраняются некоторые свойства обычных решений СЛАУ.

Теорема 4.3. Пусть x_j — псевдорешения СЛАУ $Ax = b_j$, $j = \overline{1, l}$. Тогда псевдорешением СЛАУ $Ax = b$, где $b = \lambda_1 b_1 + \dots + \lambda_l b_l$, является вектор $x = \lambda_1 x_1 + \dots + \lambda_l x_l$.

◀ Каждый вектор x_j , как нормальное псевдорешение СЛАУ $Ax = b_j$, является единственным решением СЛАУ

$$\begin{pmatrix} A^T A \\ F^T \end{pmatrix} x = \begin{pmatrix} A^T b_j \\ 0 \end{pmatrix}.$$

Непосредственный подсчет показывает, что если $x = \lambda_1 x_1 + \dots + \lambda_l x_l$, то

$$\begin{pmatrix} A^T A \\ F^T \end{pmatrix} x = \sum_{j=1}^l \begin{pmatrix} A^T A \\ F^T \end{pmatrix} x_j = \sum_{j=1}^l \begin{pmatrix} A^T b_j \\ 0 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^l A^T b_j \\ 0 \end{pmatrix} = \begin{pmatrix} A^T b \\ 0 \end{pmatrix}.$$

Полученное равенство означает, что x есть нормальное псевдорешение СЛАУ $Ax = b$. ▶

Находя нормальные псевдорешения СЛАУ $Ax = b$ с разными правыми частями b , можно поступить так. В $\mathbb{R}^m$ выбираем стандартный базис $e_1, \dots, e_m$. Находим нормальные псевдорешения g_i СЛАУ $Ax = e_i$, $i = \overline{1, m}$. Тогда $b = b_1 e_1 + \dots + b_m e_m$, где b_i — компоненты вектор-столбца b , и нормальное псевдорешение СЛАУ $Ax = b$ имеет вид $x = b_1 g_1 + \dots + b_m g_m$, что можно записать так: $x = (g_1 \dots g_m)b = A^+ b$, где матрица A^+ составлена из столбцов g_i .

Проведенное рассуждение показывает следующее. Для данной матрицы A существует такая матрица A^+ , что нормальное псевдорешение СЛАУ $Ax = b$ имеет вид $x = A^+ b$. Это очевидно в случае квадратной невырожденной матрицы A . В этом случае $A^+ = A^{-1}$ — обратная матрица.

Отметим, что матрица A^+ определена однозначно. Действительно, если $A^+ b$ — нормальное псевдорешение СЛАУ $Ax = b$, то $A^+ e_i = g_i$ — нормальное псевдорешение СЛАУ для правой части, являющейся вектором стандартного базиса в $\mathbb{R}^m$. Но произведение $A^+ e_i$ представляет собой i -й столбец матрицы A^+ . Значит, i -м столбцом матрицы A^+ является вектор-столбец g_i .

Матрицу A^+ называют **псевдообратной** по отношению к матрице A . Отметим, что если матрица A имеет тип $m \times n$, то псевдообратная матрица A^+ имеет тип nm .

4.3. Скелетное разложение

Рассмотрим линейную оболочку H столбцов a_i матрицы A . В этой оболочке выберем базис $u_1, \dots, u_k$. Тогда каждый вектор f_i можно представить в виде линейной комбинации векторов $u_1, \dots, u_k$, т.е. столбец a_i можно записать в виде $a_i = (u_1 \dots u_k)v_i = Uv_i$. Объединяя все эти представления, получим $A = UV$, где $U = (u_1 \dots u_k)$, а V — матрица координат векторов a_i в базисе $u_1, \dots, u_k$. Матрица U имеет ранг k , равный рангу матрицы A и совпадающий с количеством столбцов U , т.е. **максимальный столбцовый ранг**. Матрица V — матрица координат — имеет тот же ранг, что и система векторов $\{a_i\}$, т.е. тоже k , причем ранг совпадает с количеством строк матрицы. Значит, матрица V имеет **максимальный строчный ранг**. Представление $A = UV$, в котором у матрицы U линейно независимы столбцы, а у матрицы V линейно независимы строки, называют **скелетным разложением** матрицы A . В любом скелетном разложении матрицы A левый множитель можно интерпретировать как набор векторов базиса в линейной оболочке столбцов матрицы A , а правый множитель — как набор координат в этом базисе столбцов матрицы A .

Чтобы построить скелетное разложение матрицы A , нужно выбрать базис в линейной оболочке столбцов матрицы A и разложить по этому базису столбцы матрицы A . В принципе в качестве базиса можно брать любой. Перебирая все возможные варианты, получим все мыслимые скелетные разложения матрицы A . Это, кстати, показывает, что скелетное разложение не единственно. Однако в дальнейшем нам потребуется какое-либо одно такое разложение. Поэтому можно остановить выбор на том, которое строится проще всего. В качестве базиса $u_1, \dots, u_k$ можно выбрать любой набор базисных столбцов матрицы A , определяемый каким-нибудь базисным минором.

Предположим для простоты, что базисный минор матрицы A порядка r расположен в ее верхнем левом углу. Тогда в качестве базиса в линейной оболочке столбцов матрицы A можно выбрать ее первые r столбцов, т.е. Пусть $u_j = a_j$, $j = \overline{1, r}$. Запишем скелетное разложение $A = UV$, соответствующее выбранному базису. При одновременном выполнении элементарных преобразований в матрицах U и A равенство $UV = A$ не нарушается. С помощью таких преобразований приведем матрицу A к ступенчатому виду, выберем базисный минор в первых r столбцах (в соответствии с матрицей U) и, продолжая элементарные преобразования, сведем базисный минор к единичному. Тогда матричное равенство сведется к следующему:

$$\begin{pmatrix} E \\ 0 \end{pmatrix} V = \begin{pmatrix} E & M \\ 0 & 0 \end{pmatrix},$$

где M — некоторый блок, расположенный в последних столбцах преобразованной матрицы. Удалив нулевые строки в левой и правой частях, получим равенство $EV = (E \ M)$, откуда заключаем, что при указанном выборе базиса $u_1, \dots, u_r$ матрица V получается сведением матрицы A к ступенчатому виду с единичным базисным минором и последующим удалением нулевых строк.

Исходя из этих рассуждений можно сформулировать следующий метод построения скелетного разложения.

1. Матрицу A элементарными преобразованиями строк приводим к ступенчатому виду.
2. В матрице ступенчатого вида выбираем базисный минор (стандартно отмечаем все ненулевые строки и все столбцы, в которых расположены лидирующие элементы ненулевых строк) и удаляем все нулевые строки.
3. С помощью дополнительных элементарных преобразований строк приводим выбранный базисный минор к ступенчатому виду. Полученная при этом матрица будем матрицей V , а матрицу U составляем из тех столбцов исходной матрицы A , которым в матрицы ступенчатого вида соответствуют столбцы выбранного базисного минора.

Пример 4.1. Найдем скелетное разложение матрицы

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 1 & 5 \\ 2 & -2 & 6 \end{pmatrix}.$$

Выполняя элементарные преобразования строк матрицы, получаем

$$\begin{pmatrix} 1 & 3 & -1 \\ 3 & 1 & 5 \\ 2 & -2 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -1 \\ 0 & -8 & 8 \\ 0 & -8 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

В качестве базисного минора выберем минор M_{12}^{12} второго порядка в верхнем левом углу. В матрице ступенчатого вида удалим третью нулевую строку, а из первой строки вычтем утроенную вторую. Получим

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

Таким образом, матрицу U составляют первый и второй столбцы матрицы A , а матрица V — результат преобразований, т.е.

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \\ -3 & 7 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{pmatrix}.$$

Замечание. Рассуждая так же, как и при построении скелетного разложения, можно показать, что ранг произведения матриц не превышает минимального из рангов сомножителей, т.е. $\text{rang}(AB) \leq \min\{\text{rang } A, \text{rang } B\}$. Действительно, равенство $C = AB$ можно интерпретировать как представление столбцов матрицы C в виде линейных комбинаций столбцов матрицы A , причем коэффициенты линейных комбинаций, записанные по столбцам, составляют матрицу B . Эта интерпретация показывает, что все столбцы матрицы C принадлежат линейной оболочке столбцов матрицы A . Значит, линейная оболочка столбцов C содержится в линейной оболочке столбцов матрицы A , а ранг системы столбцов C не превышает ранга системы столбцов матрицы A , т.е. $\text{rang } C \leq \text{rang } A$.

Итак, доказано, что ранг произведения не превышает ранга левого сомножителя. Поскольку $C^T = B^T A^T$, заключаем, что $\text{rang } C = \text{rang } C^T \leq \text{rang } B^T = \text{rang } B$.

Отметим также, что ранг матрицы не изменяется при умножении ее на невырожденную матрицу. Действительно, если $C = AB$ и матрица A квадратная невырожденная, то, с одной стороны, $\text{rang } A \leq \text{rang } B$, а с другой стороны, в силу равенства $B = A^{-1}C$ имеем $\text{rang } B \leq \text{rang } A$. Аналогичен случай, когда невырожденная матрица является правым множителем.

4.4. Псевдообратная матрица

Псевдообратную матрицу можно строить, опираясь на понятие нормального псевдорешения. Поскольку каждый столбец a_i^+ псевдообратной матрицы A^+ есть нормальное псевдорешение системы $Ax = e_i$, т.е. решение системы

$$\begin{pmatrix} A^T A \\ F^T \end{pmatrix} x = \begin{pmatrix} A^T e_i \\ 0 \end{pmatrix}$$

(здесь F — матрица, составленная из столбцов фундаментальной системы решений однородной СЛАУ $Ax = 0$), то сама псевдообратная матрица есть единственное решение матричного уравнения

$$\begin{pmatrix} A^T A \\ F^T \end{pmatrix} X = \begin{pmatrix} A^T e_1 & A^T e_2 & \dots & A^T e_n \\ 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} A^T \\ 0 \end{pmatrix}. \quad (4.1)$$

Пример 4.2. Рассмотрим матрицу из предыдущего примера. Используя выполненное приведение к ступенчатому виду, находим фундаментальную систему решений СЛАУ $Ax = 0$. Она состоит из единственного вектора $\text{row}-2; 1; 1^T$. Далее вычисляем

$$A^T A = \begin{pmatrix} 14 & 2 & 26 \\ 2 & 14 & -10 \\ 26 & -10 & 62 \end{pmatrix}.$$

В результате приходим к матричному уравнению

$$\begin{pmatrix} 14 & 2 & 26 \\ 2 & 14 & -10 \\ 26 & -10 & 62 \\ -2 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 1 & -2 \\ -1 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Это уравнение будем решать методом элементарных преобразований. Записав блочную матрицу, приведем ее к ступенчатому виду:

$$\left(\begin{array}{ccc|ccc} 14 & 2 & 26 & 1 & 3 & 2 \\ 2 & 14 & -10 & 3 & 1 & -2 \\ 26 & -10 & 62 & 1 & 5 & 6 \\ -2 & 1 & 1 & 0 & 0 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} -2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 9 & 33 & 1 & 3 & 2 \\ 0 & 15 & -9 & 3 & 1 & -2 \\ 0 & 3 & 75 & -1 & 5 & 6 \end{array} \right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} -2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 75 & -1 & 5 & 6 \\ 0 & 0 & -192 & 4 & -12 & -16 \\ 0 & 0 & -384 & 8 & -24 & -32 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} -2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 75 & -1 & 5 & 6 \\ 0 & 0 & 48 & -1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Удалив последнюю нулевую строку в матрице ступенчатого вида, продолжим элементарные преобразования:

$$\left(\begin{array}{ccc|ccc} -2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 75 & -1 & 5 & 6 \\ 0 & 0 & 48 & -1 & 3 & 4 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} -96 & 48 & 0 & 1 & -3 & -4 \\ 0 & 48 & 0 & 9 & 5 & -4 \\ 0 & 0 & 48 & -1 & 3 & 4 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 48 & 0 & 0 & 4 & 4 & 0 \\ 0 & 48 & 0 & 9 & 5 & -4 \\ 0 & 0 & 48 & -1 & 3 & 4 \end{array} \right).$$

Таким образом,

$$A^+ = \frac{1}{48} \left(\begin{array}{ccc} 4 & 4 & 0 \\ 9 & 5 & -4 \\ -1 & 3 & 4 \end{array} \right) = \left(\begin{array}{ccc} \frac{1}{12} & \frac{1}{12} & 0 \\ \frac{3}{16} & \frac{5}{48} & -\frac{1}{12} \\ -\frac{1}{48} & \frac{1}{16} & \frac{1}{12} \end{array} \right).$$

Псевдообратную матрицу можно строить с помощью скелетного разложения. Прежде всего докажем один критерий для псевдообратной матрицы.

Теорема 4.4 (первый критерий псевдообратной матрицы). Псевдообратная матрица однозначно определяется равенствами

$$A^T A A^+ = A^T, \quad A^+ = A^T P. \quad (4.2)$$

◀ Псевдообратная матрица A^+ является решением матричного уравнения (4.1), из которого вытекает, что $A^T A A^+ = A^T$. Кроме того, столбцы a_i^+ матрицы A^+ , как нормальные псевдорешения систем $Ax = e_i$ (e_i — векторы стандартного базиса), являются линейными комбинациями столбцов матрицы A^T (теорема 4.2). Это можно записать в виде $a_i^+ = A^T p_i$, где столбец p_i составлен из коэффициентов линейной комбинации. Объединяя эти представления, получим

$$A^+ = (a_1^+ \ a_2^+ \ \dots \ a_m^+) = A^T (p_1 \ p_2 \ \dots \ p_m) = A^T P.$$

Итак, в матричном уравнении (4.1) первая часть $A^T A X = A^T$ совпадает с первым условием (4.2), а вторая часть $F^T A^+ = 0$, означающая, что столбцы матрицы A^+ принадлежат ортогональному дополнению $K^\perp$ к подпространству K решений системы $Ax = 0$, равносильна второму условию (4.2), поскольку подпространство $K^\perp$ есть линейная оболочка столбцов матрицы A^T . ▶

Из доказанной теоремы легко понять, что псевдообратная матрица к нулевой матрице получается транспонированием. Действительно, равенства $0^T 0 0^+ = 0^T$ тривиально (выполняется независимо от того, что является матрицей 0^+), а равенство $0^+ = 0^T P$ автоматически дает в качестве 0^+ нулевую матрицу. Остается согласовать типы матриц.

Теорема 4.5. Если U — матрица максимального столбцового ранга, то $U^+ = (U^T U)^{-1} U^T$. Если V — матрица максимального строчного ранга, то $V^+ = V^T (V V^T)^{-1}$. Наконец, для произвольной ненулевой матрицы A со скелетным разложением $A = UV$ имеем

$$A^+ = V^+ U^+ = V^T (V V^T)^{-1} (U^T U)^{-1} U^T. \quad (4.3)$$

◀ Будем опираться на доказанный критерий. Для матрицы U максимального столбцового ранга $U^T U$ — невырожденная матрица. Поэтому из равенства $U^T U U^+ = U^T$ находим $U^+ =$

$(U^T U)^{-1} U^T$. Условие $U^+ = U^T P$ в данном случае тривиально, так как ранг матрицы U^T совпадает с размерностью пространства всех столбцов, т.е. любой столбец разлагается в линейную комбинацию столбцов матрицы U^T .

Для матрицы V максимального строчного ранга матрица $V V^T$ невырождена. Поэтому из равенства $V^T V V^+ = V^T$ после домножения на V слева и сокращения на невырожденную матрицу $V V^T$, получим $V V^+ = E$. Подставляя представление $V^+ = V^T P$, получим уравнение $V V^T P = E$, из которого $P = (V V^T)^{-1}$. Значит, $V^+ = V^T P = V^T (V V^T)^{-1}$.

Рассмотрим общий случай $A = UV$. Равенство $A^T A A^+ = A^T$ превращается в равенство $V^T U^T U V A^+ = V^T U^T$. Домножим на V слева и сократим на невырожденную матрицу $V V^T$: $U^T U V A^+ = U^T$. В левой части матрица $U^T U$ невырожденная. Поэтому $V A^+ = (U^T U)^{-1} U^T = U^+$. Используя представление $A^+ = A^T P = V^T U^T P$, находим $V V^T U^T P = U^+$. Отсюда $U^T P = (V V^T)^{-1} U^+$ и $A^+ = A^T P = V^T U^T P = V^T (V V^T)^{-1} U^+ = V^+ U^+$. ►

Доказанная теорема позволяет вычислять псевдообратную матрицу, находя ее скелетное разложение. Отметим два важных тождества: $V V^+ = E$ и $U^+ U = E$. Первое из них было использовано в доказательстве.

Пример 4.3. Для матрицы A из предыдущего примера найдено скелетное разложение:

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

Для матрицы U вычисляем

$$U^T U = \begin{pmatrix} 14 & 2 \\ 21 & 4 \end{pmatrix}, \quad (U^T U)^{-1} = \frac{1}{192} \begin{pmatrix} 14 & -2 \\ -2 & 14 \end{pmatrix} = \frac{1}{96} \begin{pmatrix} 7 & -1 \\ -1 & 7 \end{pmatrix},$$

$$U^+ = \frac{1}{96} \begin{pmatrix} 7 & -1 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} 13 & 2 \\ 31 & -2 \end{pmatrix} = \frac{1}{24} \begin{pmatrix} 15 & 4 \\ 51 & -4 \end{pmatrix}.$$

Аналогичны действия для матрицы V :

$$V V^T = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}, \quad (V V^T)^{-1} = \frac{1}{6} \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}, \quad V^+ = \frac{1}{6} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 & 2 \\ 2 & 5 \\ 2 & -1 \end{pmatrix}.$$

Выполняя последнее умножение, находим

$$A^+ = V^+ U^+ = \frac{1}{48} \begin{pmatrix} 4 & 4 & 0 \\ 9 & 5 & -4 \\ -1 & 3 & 4 \end{pmatrix}.$$

Теорема 4.6. Псевдообратная матрица обладает следующими свойствами:

1. $\text{rang } A^+ = \text{rang } A$.
2. $(A^+)^+ = A$.
3. $(A^T)^+ = (A^+)^T$.
4. $(A^+ A)^T = A^+ A$; $(A^+ A)^2 = A^+ A$.
5. $(A A^+)^T = A A^+$; $(A A^+)^2 = A A^+$.

◀ Равенство рангов можно установить разными способами. Например, из критерия 4.4 и замечания о ранге произведения заключаем, что во-первых, $\text{rang } A = \text{rang } A^T = \text{rang}(A^T A A^+) \leq \leq \text{rang } A^+$, а во-вторых, $\text{rang } A^+ = \text{rang}(A^T P) \leq \text{rang } A^T = \text{rang } A$. Равенство $(A^+)^+ = A$ вытекает из того, что $A^+ = V^+ U^+$ — скелетное разложение матрицы A^+ . А то, что это скелетное

разложение, вытекает из первого доказанного свойства, поскольку $\text{rang } V^+ = \text{rang } V$ совпадает с количеством столбцов матрицы V^+ , а $\text{rang } U^+ = \text{rang } U$ — с количеством строк матрицы V^+ .

Поскольку $A^+ = V^+U^+$ — скелетное разложение матрицы A^+ , то $(A^+)^+ = (U^+)^+(V^+)^+$. Остается доказать второе свойство для матриц максимального ранга. Имеем

$$U^+(U^+)^T = (U^T U)^{-1} U^T U (U^T U)^{-1} = (U^T U)^{-1}.$$

Поэтому

$$(U^+)^+ = (U^+)^T (U^+(U^+)^T)^{-1} = U (U^T U)^{-1} U^T U = U.$$

Аналогичны выкладки для матрицы V .

Докажем третье свойство. Если $A = UV$ — скелетное разложение матрицы A , то $A^T = V^T U^T$ — скелетное разложение матрицы A^T . Поэтому

$$(A^T)^+ = (U^T)^+(V^T)^+.$$

Предположим, что третье свойство установлено для матриц максимального ранга. Тогда

$$(A^T)^+ = (U^T)^+(V^T)^+ = (U^+)^T (V^+)^T = (V^+ U^+)^T = (A^+)^T.$$

Проверить свойство для матриц максимального ранга можно, используя формулы вычисления.

Докажем четвертое свойство. Имеем $U^+ U = (U^T U)^{-1} U^T U = E$. Отсюда вытекает, что $A^+ A = V^+ U^+ UV = V^+ V$. Значит, достаточно доказать свойство для матриц максимального строчного ранга. Имеем

$$\begin{aligned} (V^+ V)^T &= V^T (V^+)^T = V^T (V V^T)^{-1} V = V^+ V, \\ (V^+ V)^2 &= V^+ V V^+ V = V^+ (V V^+) V = V^+ V. \end{aligned}$$

Здесь использовано тождество $V V^+ = E$, выполняющееся для матриц максимального строчного ранга.

Пятое свойство вытекает из четвертого, примененного к матрице A^+ . ►

Следующий критерий в ряде литературных источников лежит в основе определения псевдообратной матрицы. От первого критерия он отличается большей симметрией условий.

Теорема 4.7 (второй критерий псевдообратной матрицы). Псевдообратная матрица A^+ однозначно определяется равенствами

$$AA^+A = A, \quad A^+ = A^T P, \quad A^+ = Q A^T. \quad (4.4)$$

◀ Нетрудно показать, что для псевдообратной матрицы выполняются равенства (4.4). Например, транспонируя первое равенство (4.2), с учетом симметричности матрицы AA^+ (свойство 5 в теореме 4.6) получаем первое равенство (4.4). Второе равенство (4.4) дублирует второе равенство (4.2), а третье равенство (4.4) есть второе равенство (4.2), записанное для матрицы A^T .

Остается показать, что только псевдообратная матрица удовлетворяет условиям (4.4). Один из вариантов — доказать, что только одна матрица может удовлетворять равенствам (4.4). Мы проведем доказательство, решив систему матричных уравнений (4.4) относительно неизвестного A^+ . Решение основывается на следующем соображении. Из равенства $A^T A X = A^T A Y$ вытекает равенство $A X = A Y$, а из равенства $X A^T A = Y A^T A$ — равенство $X A^T = Y A^T$. Действительно, равенство $A^T A X = A^T A Y$ равносильно равенству $A^T A Z = 0$, где $Z = X - Y$. Равенство $A^T A Z = 0$ означает, что столбцы матрицы Z являются решениями системы $A^T A z = 0$, а потому — решениями системы $A z = 0$. Следовательно, $A Z = 0$. Транспонируя равенство

$XA^T A = Y A^T A$, получаем $A^T A X^T = A^T A Y^T$, откуда согласно уже доказанному $A X^T = A Y^T$. Еще раз транспонируя, находим $X A^T = Y A^T$.

Пусть некая матрица A^+ удовлетворяет равенствам (4.4). Тогда, домножив первое равенство на A^T и заменив в нем A^+ в соответствии с равенством $A^+ = Q A^T$, получим равенство $A^T A Q A^T A = A^T A$, в котором можно сократить правый множитель A . В результате получим $A^T A Q A^T = A^T$, или $A^T A A^+ = A^T$, поскольку $Q A^T = A^+$. Таким образом, из равенств (4.4) вытекает первое равенство (4.2). Второе равенство (4.2) дублирует второе равенство (4.4). Значит, из равенств (4.4) вытекают равенства (4.2), а матрица A^+ , удовлетворяющая равенствам (4.4), является псевдообратной. ►

4.5. Проектирование на подпространство

С помощью псевдообратных матриц можно вычислять проекции векторов на подпространство. Отметим, что для системы $Ax = b$ с произвольной правой частью b вектор $x = A^+b$ является псевдорешением. Значит, согласно теореме 4.1 выполняется равенство $AA^+b = \text{пр}_H b$, т.е. вектор AA^+b представляет собой ортогональную проекцию вектора b на подпространство $H = \text{span}(a_1, a_2, \dots, a_n)$ — линейную оболочку столбцов матрицы A .

Итак, если H — линейная оболочка системы векторов $a_1, a_2, \dots, a_n$, то чтобы вычислить проекцию $\text{пр}_H b$ вектора b на подпространство H , достаточно из координат векторов a_i составить матрицу A и вычислить AA^+b .

Матрица A^+A также осуществляет проекцию вектора на некоторое подпространство. Это подпространство представляет собой линейную оболочку столбцов матрицы A^+ . А как связать это подпространство непосредственно с матрицей A ? Отметим, что, согласно теореме 4.2 все столбцы матрицы A^+ , как нормальные псевдорешения систем с соответствующими правыми частями¹, являются линейными комбинациями столбцов матрицы A^T . Следовательно, линейная оболочка столбцов матрицы A^+ является подпространством в линейной оболочке столбцов матрицы A^T . Но поскольку размерности этих подпространств совпадают (ранги матриц A^+ и A^T одинаковы), то на самом деле линейная оболочка столбцов матрицы A^+ совпадает с линейной оболочкой столбцов матрицы A^T , которая в свою очередь представляет собой ортогональное дополнение к подпространству K решений однородной системы $Ax = 0$. Следовательно, вектор A^+Ab есть проекция вектора b на подпространство $K^\perp$.

Так как A^+Ab — проекция вектора b на подпространство $K^\perp$, вектор $b - A^+Ab = (E - A^+A)b$ представляет собой ортогональную составляющую вектора b при проектировании на $K^\perp$. Но этот же вектор есть ортогональная проекция вектора b на подпространство $(K^\perp)^\perp = K$.

Итак, с помощью псевдообратной матрицы A^+ можно вычислять ортогональную проекцию произвольного вектора b и на линейную оболочку столбцов матрицы A (это вектор AA^+b), и на множество решений однородной линейной системы $Ax = 0$ (это вектор $(E - A^+A)b$).

Теорема 4.8. Любое псевдорешение x СЛАУ $Ax = b$, где A — матрица типа $m \times n$, можно представить в виде

$$x = A^+b + (E - A^+A)t, \quad (4.5)$$

где $t \in \mathbb{R}^m$ — некоторый вектор.

◀ Множество всех псевдорешений можно представить в виде $x = x_{\text{ч}} + x_0$, где $x_{\text{ч}}$ — некоторое частное псевдорешение, а x_0 — общее решение однородной СЛАУ $Ax = 0$. В качестве частного псевдорешения можно выбрать нормальное, а множество все решений СЛАУ $Ax = 0$ можно получить, проектируя на подпространство K всех решений произвольные векторы $t \in \mathbb{R}^n$. ►

¹Тот же вывод можно сделать, интерпретируя соответствующим образом второе равенство (4.2).

Замечание. 1. Запись $(E - A^+A)t$ можно рассматривать как линейную комбинацию столбцов матрицы $E - A^+A$, коэффициенты в которой определяются вектором t . Это похоже на запись общего решения однородной системы через фундаментальную систему решений. Однако фундаментальная система решений есть базис в подпространстве K всех решений системы, а множество столбцов матрицы $E - A^+A$ базисом не является, поскольку в этой матрице количество столбцов равно размерности всего арифметического пространства $\mathbb{R}^n$ и превышает размерность пространства решений (исключение составляет только случай нулевой матрицы A). Можно лишь утверждать, что подпространство решений системы $Ax = 0$ является линейной оболочкой столбцов матрицы $E - A^+A$.

2. Равенство (4.5) представляет собой разложение произвольного псевдорешения x на ортогональную составляющую A^+b и ортогональную проекцию $(E - A^+A)t$ при проектировании на подпространство K всех решений. Кстати, в качестве вектора t во втором слагаемом разложения можно выбрать сам вектор x , т.е. для любого псевдорешения x системы $Ax = b$ верно разложение

$$x = A^+b + (E - A^+A)x.$$

ОГЛАВЛЕНИЕ

4. Псевдорешения и псевдообратная матрица	1
4.1. Метод наименьших квадратов	1
4.2. Псевдорешения	3
4.3. Скелетное разложение	4
4.4. Псевдообратная матрица	6
4.5. Проектирование на подпространство	10
9. Жорданова нормальная форма	12
9.1. Корневые подпространства	12
9.2. Жорданова нормальная форма	14
9.3. Комплексные корни	19
9.4. Теорема Кэли — Гамильтона	20
13. Операции над тензорами	22
13.1. Понятие тензора	22
13.2. Матричная запись тензоров	23
13.3. Преобразование тензоров, записанных в матричной форме	24
14. Множества и отношения	26
14.1. Алгебра множеств	26
14.2. Отображения и соответствия	30
14.3. Отношения и операции	32
14.4. Элементы математической логики	34
14.5. Мощность множеств	35
16. Кольца и поля	36
16.1. Кольца	36
16.2. Специальные типы колец	36
16.3. Гомоморфизмы колец и факторизация	36
16.4. Модули и алгебры	36
16.5. Алгебры на поле	40
17. Кольцо многочленов	43
17.1. Определение кольца многочленов	43
17.2. Деление с остатком и его свойства	45
17.3. Разложение на неприводимые множители	49
17.4. Использование делимости в теории шифрования	49
17.5. Кватернионы	50
19. Полукольца и булевы алгебры	51
19.1. Определение полукольца	51
19.2. Ряды в полукольцах	53
19.3. Замкнутые полукольца	57
19.4. Системы линейных уравнений в полукольцах	58
19.5. Симметричные полукольца	63
19.6. Решетки	67