Обращение прямоугольных матриц

Если A — квадратная невырожденная матрица, то для нее существует обратная матрица A^{-1}. Если же A — не квадратная, а прямоугольная $(m \times n)$-матрица $(m \neq n)$ или квадратная, но вырожденная матрица, то она не имеет обратной матрицы и символ A^{-1} не имеет смысла. В то же время, как будет показано далее, для произвольной прямоугольной матрицы A существует псевдообратная матрица A^{+}, которая обладает некоторыми свойствами обратной матрицы и имеет важные применения при решении систем линейных алгебраических уравнений. В случае, если A — квадратная невырожденная матрица, псевдообратная матрица A^{+} совпадает с обратной матрицей A^{-1}.

Представление $(m \times n)$-матрицы A ранга r в виде произведения

$$A = B\,C \tag{1}$$

с $(m \times r)$-матрицей B ранга r и $(r \times n)$-матрицей C ранга r называют **скелетным** разложением матрицы A.

В общем случае при конструировании скелетного разложения $(m \times n)$-матрицы A ранга r можно руководствоваться следующим правилом:

Для построения скелетного разложения $A = B\,C$ $(m \times n)$-матрицы A ранга r можно в качестве сомножителя B взять матрицу, составленную из каких-либо r линейно независимых столбцов матрицы A, а элементы j-го столбца матрицы C: $(c_{1j}, c_{2j}, \ldots, c_{rj})^{\mathrm{T}}$, $j = 1, 2, ..., n$, найти из равенства

$$c_{1j} \cdot B_1 + c_{2j} \cdot B_2 + \ldots + c_{rj} \cdot B_r = A_j, \tag{2}$$

где $B_1, B_2, \ldots, B_r$ — столбцы матрицы B; A_j — j-й столбец матрицы A.

Это правило обусловлено тем, что по теореме о ранге матрицы все столбцы матрицы A линейно выражаются через столбцы матрицы B, т. е. можно записать равенство (2) при $j = 1, 2, ..., n$. Но матричной записью этих равенств служит равенство (1).

Из приведенного правила следует, что матрица A может обладать многими скелетными разложениями, поскольку при применении этого правила в качестве столбцов матрицы B можно брать любую максимальную линейно независимую подсистему системы столбцов матрицы A. Более того, матрицу B можно строить из линейно независимых линейных комбинаций столбцов матрицы A, даже не являющихся столбцами матрицы A. Это обстоятельство увеличивает свободу выбора матрицы B. В то же время при выбранном сомножителе B скелетного разложения $A = BC$ сомножитель C определяется однозначно.

Пример 1. Построить скелетное разложение для следующих матриц:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 1 & 3 & -3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Решение. Легко убедиться в том, что ранг матрицы A равен двум и первые два ее столбца линейно независимые. Поэтому в качестве первого сомножителя в скелетном разложении $A = BC$ можно принять матрицу

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix}.$$

Элементы j-го столбца $(c_{1j},\ c_{2j})^{\mathrm{T}}$ матрицы C, $j = 1, 2, 3$, можно найти из равенства

$$c_{1j} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + c_{2j} \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix} = A_j, \qquad (3)$$

где A_j — j-й столбец матрицы A.

При $j=1$ соотношение (3) приводит к системе уравнений

$$\begin{cases} c_{11}+\ 2c_{21}=1, \\ 2c_{11}+\ \ c_{21}=2, \\ c_{11}+\ 3c_{21}=1, \\ c_{11}+\ \ c_{21}=1, \end{cases}$$

из которой находим $c_{11}=1,\ c_{21}=0$. Поэтому первым столбцом матрицы C будет столбец $(1,\ 0)^{\mathrm{T}}$. Аналогично при $j=2$ и $j=3$ найдем второй и третий столбцы $(0,\ 1)^{\mathrm{T}}$ и $(3,\ -2)^{\mathrm{T}}$ матрицы C. Следовательно,

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

и матрица A имеет скелетное разложение

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

Определение. *Матрица A^{+} размера $(n\times m)$ называется псевдообратной к действительной или комплексной $(m\times n)$-матрице A, если выполняются соотношения*

$$A\,A^{+}A=A,\qquad A^{+}=U\,A^{*}=A^{*}V\,,$$

где U и V — некоторые матрицы.

Теорема 1. *Для любой $(m\times n)$-матрицы A существует, и притом единственная, псевдообратная матрица A^{+}.*

Если ранг матрицы A совпадает с количеством ее столбцов (матрица A имеет максимальный столбцовый ранг), то

$$A^{+}=(A^{*}A)^{-1}\cdot A^{*}. \tag{4}$$

Если ранг матрицы A совпадает с количеством ее строк (матрица A имеет максимальный строчный ранг), то

$$A^{+}=A^{*}\cdot(A\,A^{*})^{-1}. \tag{5}$$

В общем случае, если матрица A *имеет скелетное разложение* $A = B\,C$ **,** *то*

$$A^{+} = C^{+}B^{+} = C^{*}\cdot(C\cdot C^{*})^{-1}(B^{*}\cdot B)^{-1}B^{*}. \qquad (6)$$

Псевдообратная матрица обладает следующими свойствами:

Свойство 1. $(A^{*})^{+} = (A^{+})^{*}$.

Свойство 2. $(A^{+})^{+} = A$.

Свойство 3. $(A\,A^{+})^{*} = A\,A^{+}, \quad (A\,A^{+})^{2} = A\,A^{+}$.

Свойство 4. $(A^{+}A)^{*} = A^{+}A, \quad (A^{+}A)^{2} = A^{+}A$.

Теорема 2. *Если* $A = U\Sigma V^{*}$ *— сингулярное разложение* $(m\times n)$ *-матрицы* A *с* $(m\times n)$ *-матрицей*

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & 0 \\ & & \ddots & \\ & 0 & & \sigma_r \\ & & & & 0 \end{pmatrix},$$

то

$$A^{+} = V\,\Sigma^{+}U^{*},$$

где $(n\times m)$ *-матрица*

$$\Sigma^{+} = \begin{pmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & 0 \\ & & \ddots & \\ & 0 & & \sigma_r^{-1} \\ & & & & 0 \end{pmatrix}.$$

Эта теорема и формулы (4)–(6) дают удобный способ вычисления псевдообратной матрицы.

Пример 2. Найти псевдообратную матрицу A^+ для матрицы

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 1 & 3 & -3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Решение Для матрицы A построено скелетное разложение

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

Поэтому для нахождения псевдообратной матрицы A^+ применим формулу (6). Для этого сначала вычислим

$$B^+ = (B^*B)^{-1}B^* = \frac{1}{41}\begin{pmatrix} -1 & 22 & -9 & 7 \\ 6 & -9 & 13 & -1 \end{pmatrix},$$

$$C^+ = C^*(CC^*)^{-1} = \frac{1}{14}\begin{pmatrix} 5 & 6 \\ 6 & 10 \\ 3 & -2 \end{pmatrix}$$

Теперь по формуле (6) получаем

$$A^+ = C^+B^+ = \frac{1}{574}\begin{pmatrix} 31 & 56 & 33 & 29 \\ 54 & 42 & 76 & 32 \\ -15 & 84 & -53 & 23 \end{pmatrix}.$$

Пример 3. Найти псевдообратную матрицу A^+ для матрицы

$$A = \begin{pmatrix} 7 & 1 & 2 \\ 5 & 5 & -2 \\ 3 & 3 & 6 \\ 1 & 7 & 2 \end{pmatrix}.$$

Решение Для матрицы A построено сингулярное разложение

$$A = U\Sigma V^* = \frac{1}{2}\cdot\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix}\cdot\frac{1}{3}\cdot\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Поэтому можно применить формулу (6), согласно которой

$$A^+ = V\Sigma^+U^* = \frac{1}{3}\cdot\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}\cdot\begin{pmatrix} \frac{1}{12} & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 \end{pmatrix}\times$$

$$\times\frac{1}{2}\cdot\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} = \frac{1}{72}\cdot\begin{pmatrix} 8 & 4 & 0 & -4 \\ -4 & 4 & 0 & 8 \\ 1 & -7 & 9 & 1 \end{pmatrix}.$$

Для построения псевдообратной матрицы A^+ к $(m \times n)$-матрице A можно пользоваться *методом Гревилля* [1, 3, 12], не требующим знания скелетного разложения матрицы A и вычисления определителей. Для этого вводят обозначения:

a_k — k-й столбец матрицы A, $k = 1, 2, \ldots, n$;

$A_k = (a_1, a_2, \ldots, a_k)$ — матрица, образованная первыми k столбцами матрицы A, $k = 1, 2, \ldots, n$;

b_k — последняя строка матрицы A_k^+, $k = 1, 2, \ldots, n$,

и применяют формулы:

$$A_1^+ = a_1^+ = (a_1^* a_1)^{-1} a_1^* \text{ и при } k = 2, 3, \ldots, n \quad A_k^+ = \begin{pmatrix} B_k \\ b_k \end{pmatrix},$$

где $B_k = A_{k-1}^+ - d_k b_k$, $d_k = A_{k-1}^+ a_k$, причем

при $c_k = a_k - A_{k-1} d_k \neq 0$, $\quad b_k = c_k^+ = (c_k^* c_k)^{-1} c_k^*$,

при $c_k = 0$, $\quad b_k = (E + d_k^* d_k)^{-1} d_k^* A_{k-1}^+$.

Матрица A_k^+, построенная по указанным формулам, является псевдообратной к матрице A_k, $k = 1, 2, \ldots, n$. В частности, $A_n^+ = A^+$.

Метод Гревилля особенно удобен, если в процессе исследования с течением времени данные пополняются и расчеты, ранее проведенные с матрицей A_k, нужно повторить с матрицей A_{k+1}, полученной из матрицы A_k добавлением $(k+1)$-го столбца. Если добавляются строки, то этот же метод применяется к транспонированной матрице.

Пример 4. Методом Гревилля вычислить псевдообратные матрицы к матрице

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

Решение. Построим последовательность матриц

$$A_k^+, \ k = 1, 2, ..., n.$$

При $k = 1$ получаем $A_1^+ = (a_1^* a_1)^{-1} a_1^* =$

$$= \left(\begin{pmatrix} 1 & -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & -1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & 0 \end{pmatrix}.$$

При $k = 2$

$$d_2 = A_1^+ a_2 = \begin{pmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -3 \\ 1 \end{pmatrix} = -\frac{3}{2},$$

$$c_2 = a_2 - A_1 d_2 = \begin{pmatrix} -1 \\ 2 \\ -3 \\ 1 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \neq 0.$$

Поэтому $b_2 = c_2^+ = (c_2^* c_2)^{-1} c_2^* =$

$$= \left(\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$$

и $B_2 = A_1^+ - d_2 b_2 =$

$$= \begin{pmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & 0 \end{pmatrix} + \frac{3}{2} \cdot \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{pmatrix}.$$

Следовательно,

$$A_2^+ = \begin{pmatrix} B_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}.$$

При $k = 3$

$$d_3 = A_2^+ a_3 = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$c_3 = a_3 - A_2 d_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ -1 & 2 \\ 2 & -3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0.$$

Поэтому $b_3 = (E + d_3^* d_3)^{-1} d_3^* A_2^+ =$

$$= \left(E + \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{9} & \frac{1}{9} & \frac{5}{9} \end{pmatrix}$$

и $B_3 = A_2^+ - d_3 b_3 =$

$$= \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{9} & \frac{1}{9} & \frac{5}{9} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{9} & \frac{2}{9} & \frac{4}{9} \\ 0 & \frac{1}{9} & -\frac{1}{9} & \frac{1}{9} \end{pmatrix}.$$

Следовательно,

$$A_3^+ = \begin{pmatrix} B_3 \\ b_3 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 3 & 1 & 2 & 4 \\ 0 & 1 & -1 & 1 \\ 3 & 2 & 1 & 5 \end{pmatrix} = A^+ .$$

Решение систем линейных алгебраических уравнений методом наименьших квадратов

1. *Постановка задачи.*

Пусть дана действительная или комплексная система линейных уравнений

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

или, что то же самое, в матричной форме

$$Ax = b. \qquad (2)$$

В общем случае эта система может быть и несовместной. Тогда решать ее точно не имеет смысла. Поэтому систему уравнений (1), совместная она или несовместная, решают методом наименьших квадратов (МНК), т. е. находят такие ее приближенные решения (*псевдорешения*) $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, при которых квадратичное отклонение

$$F(x) = \left| Ax - b \right|^2 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j - b_i \right|^2 \qquad (3)$$

имеет наименьшее значение.

Очевидно, что если система уравнений $Ax = b$ является совместной, то ее псевдорешения представляют собой обычные решения, и, наоборот, *любое решение этой системы является ее псевдорешением.*

Система может иметь не единственное псевдорешение. Среди всех псевдорешений системы линейных алгебраических уравнений псевдорешение x^0, имеющее наименьшую длину, называют ***нормальным псевдорешением***. На таком псевдорешении x^0 функция

$$\Phi(x) = \left| x \right|^2 = \sum_{i=1}^{n} \left| x_i \right|^2 \qquad (4)$$

принимает наименьшее значение.

Этим свойством нормального псевдорешения обычно пользуются при его выделении из общего псевдорешения. Другой способ выделения нормального псевдорешения системы $Ax = b$ основывается на свойстве ортогональности нормального псевдорешения к каждому решению однородной системы уравнений $Ax = 0$, в частности, к каждому решению, входящему в какую-либо фундаментальную систему решений этой однородной системы.

Практическая необходимость такого подхода возникает, например, если в эмпирической формуле

$$x_1 a_1 + x_2 a_2 + \ldots + x_n a_n = b$$

нужно найти коэффициенты $x_1,\ x_2, \ldots,\ x_n$ по результатам наблюдений (измерений) величин $a_1,\ a_2, \ldots,\ a_n, b$.

Пусть имеем следующие результаты наблюдений.

Номер наблюдения	a_1	a_2	$\cdots$	a_n	b
1	a_{11}	a_{12}	$\cdots$	a_{1n}	b_1
2	a_{21}	a_{22}	$\cdots$	a_{2n}	b_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
m	a_{m1}	a_{m2}	$\cdots$	a_{mn}	b_m

Подставляя данные каждого наблюдения в рассматриваемое выражение, придем к системе $Ax = b$. Эта система вследствие ошибок измерений и неучтенных факторов, как правило, бывает несовместной. Поэтому приходится ее решать методом наименьших квадратов.

Исходя из определения псевдорешений, укажем способ их нахождения. Для этого замечаем, что левая часть i-го уравнения системы (1) при любом наборе значений $x_1,\ x_2, \ldots,\ x_n$ отличается от его правой части на некоторую невязку. Поэтому вместо системы (1) будем рассматривать систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + z_1 = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + z_2 = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + z_m = b_m \end{cases}$$

в матричной записи имеющую вид $Ax + z = b$.

Для того чтобы функция $F(x) = \left| Ax - b \right|^2$ имела наименьшее значение, длина вектора z должна быть минимальной. Это возможно, лишь когда в равенстве $Ax + z = b$ вектор

$$Ax = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$$

является ортогональной проекцией вектора b на пространство $L(A)$ столбцов $a_1, a_2, \ldots, a_n$ матрицы A, а вектор z — ортогональной составляющей вектора b. Поэтому вектор z должен быть ортогональным к каждому вектору-столбцу a_j матрицы A. Следовательно, $(z, a_j) = 0$, $j = 1, 2, \ldots, n$. В матричной форме это означает, что $A^* z = 0$. Учитывая это равенство при умножении слева равенства $Ax + z = b$ на матрицу A^*, получим систему

$$A^* A x = A^* b. \tag{5}$$

В действительном (вещественном) случае к системе (5) можно прийти также, используя методы математического анализа, а именно, приравняв к нулю дифференциал

$$dF(x) = dx^* A^* (Ax - b) + (Ax - b)^* A dx = 2 \cdot dx^* (A^* A x - A^* b)$$

функции

$$F(x) = \left| Ax - b \right|^2 = (Ax - b)^* (Ax - b),$$

или, что то же самое, приравнивая к нулю первые частные производные по $x_1, x_2, \ldots, x_n$ функции $F(x)$.

Замечание. При вычислении дифференциала $dF(x)$ использовано то обстоятельство, что вещественное выражение $(Ax-b)^* A\,dx$, являясь матрицей первого порядка, не меняется при транспонировании и комплексном сопряжении, и потому совпадает с выражением $dx^* A^*(Ax-b)$.

Систему $A^* A x = A^* b$ называют *системой нормальных уравнений*. Ее матрица $A^* A$ квадратная n-го порядка. Определителем этой матрицы является определитель матрицы Грама системы столбцов матрицы A.

Теорема 1. *Система нормальных уравнений* $A^* A x = A^* b$ *всегда совместна, причем ее решения, и только они, являются псевдорешениями системы* $Ax = b$.

Доказательство. Для доказательства совместности системы $A^* A x = A^* b$ убедимся в том, что одним из ее решений является вектор $A^+ b$, где A^+ — псевдообратная к A матрица. Для этого запишем какое-либо скелетное разложение $A = BC$ матрицы A. Тогда

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$$

и имеет место цепочка равенств

$$A^* A\,(A^+ b) = C^* B^* B C C^*(CC^*)^{-1}(B^*B)^{-1}B^* b =$$

$$= C^* B^* B\,(B^*B)^{-1}B^* b = C^* B^* b = A^* b.$$

Следовательно, $A^* A\,(A^+ b) = A^* b$, т. е. вектор $A^+ b$ является решением системы нормальных уравнений, а сама эта система совместная.

Перейдем к доказательству второй части утверждения теоремы.

Пусть x_0 — произвольное решение системы $A^* A x = A^* b$. Покажем, что x_0 является псевдорешением системы $Ax = b$. Для этого рассмотрим значение функции $F(x) = |Ax-b|^2$ при $x = x_0 + z$, где z — произвольный вектор. Тогда получим

$$F(x_0+z)=\left|A(x_0+z)-b\right|^2=\left|(Ax_0-b)+Az\right|^2=$$

$$=((Ax_0-b)+Az;\,(Ax_0-b)+Az)=(Ax_0-b;Ax_0-b)+$$

$$+(Az,Ax_0-b)+(Ax_0-b,Az)+(Az,Az)=$$

$$=\left|Ax_0-b\right|^2+\overline{z^*\underbrace{(A^*Ax_0-A^*b)}_{=0}}+z^*\underbrace{(A^*Ax_0-A^*b)}_{=0}+\left|Az\right|^2=$$

$$=F(x_0)+\left|Az\right|^2.$$

Таким образом, для $x=x_0+z$ при любом z выполняется соотношение $F(x)\geq F(x_0)$, причем равенство возможно лишь при $z=0$. Следовательно, x_0 является псевдорешением системы $Ax=b$.

Следствие Общее псевдорешение системы $Ax=b$ представимо в виде

$$x=x_0+x_{\text{одн}},$$

где x_0 — какое-либо частное решение системы $A^*Ax=A^*b$, $x_{\text{одн}}$ — общее решение системы $A^*Ax=0$.

Лемма 1. *Однородные системы* $Ax=0$ *и* $A^*Ax=0$ *эквивалентны.*

Доказательство. Если $Ax=0$, то $A^*Ax=A^*\cdot 0=0$, т. е. все решения системы $Ax=0$ являются и решениями системы $A^*Ax=0$. Предположим, что столбец x является решением системы $A^*Ax=0$. Тогда $x^*A^*Ax=x^*\cdot 0=0$. Но

$$0=x^*A^*Ax=(Ax)^*Ax=\left|Ax\right|^2.$$

Таким образом, из равенства $A^*Ax=0$ вытекает, что столбец Ax имеет нулевую длину, а значит, он равен нулю, т. е. $Ax=0$ и x — решение этой системы.

Следствие. *Общее псевдорешение системы* $Ax = b$ *представимо в виде*

$$x = x_0 + x_{\text{одн}},$$

где x_0 *— какое-либо частное псевдорешение системы* $Ax = b$;

$x_{\text{одн}}$ *— общее решение однородной системы* $Ax = 0$.

Итак, чтобы решить систему $Ax = b$ методом наименьших квадратов, нужно найти общее решение системы $A^* A\, x = A^* b$ и, если требуется, из множества решений выделить нормальное решение, т. е. то, которое имеет наименьшую длину.

Пример 5. Пользуясь методами математического анализа, найти общее и нормальное псевдорешения системы линейных уравнений $Ax = b$ при

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 9 \\ 18 \\ -9 \\ 18 \end{pmatrix}.$$

Решение. Для решения данной системы указанным способом вычислим следующие матрицы:

$$A^* A = \begin{pmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & -9 & -3 \\ -9 & 15 & 6 \\ -3 & 6 & 3 \end{pmatrix},$$

$$A^* b = \begin{pmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 9 \\ 18 \\ -9 \\ 18 \end{pmatrix} = \begin{pmatrix} -27 \\ 72 \\ 45 \end{pmatrix},$$

выпишем систему нормальных уравнений $A^* A\, x = A^* b$, которая в данном случае имеет вид

$$\begin{cases} 6x_1 - 9x_2 - 3x_3 = -27, \\ -9x_1 + 15x_2 + 6x_3 = 72, \\ -3x_1 + 6x_2 + 3x_3 = 45. \end{cases}$$

Решив эту систему, получим искомое общее псевдорешение

$$x = (27 - x_3;\ 21 - x_3;\ x_3)^{\mathrm{T}}.$$

Чтобы из общего псевдорешения x выделить нормальное псевдорешение x^0, составим функцию

$$\Phi(x) = |x|^2 = (27 - x_3)^2 + (21 - x_3)^2 + (x_3)^2,$$

найдем ее первую частную производную по x_3 и приравняем ее к нулю:

$$-2 \cdot (27 - x_3) - 2 \cdot (21 - x_3) + 2 \cdot x_3 = 0.$$

Отсюда находим $x_3 = 16$. Поэтому нормальным псевдорешением исходной системы уравнений является

$$x^0 = (27 - 16,\ 21 - 16,\ 16)^{\mathrm{T}} = (11,\ 5,\ 16)^{\mathrm{T}}.$$

Пример 6. Найти многочлен

$$p(t) = x_1 + x_2 \cdot t + x_3 \cdot t^2,$$

который с наименьшей квадратичной погрешностью аппроксимирует на отрезке $[-2, 2]$ функцию $f(t)$, заданную табл.

t	-2	-1	1	2
$f(t)$	9	2	5	16

Решение. Подставляя поочередно данные из таблицы в многочлен $p(t) = x_1 + x_2 \cdot t + x_3 \cdot t^2$, придем к системе линейных алгебраических уравнений

$$\begin{cases} x_1 - 2x_2 + 4x_3 = 9, \\ x_1 - x_2 + x_3 = 2, \\ x_1 + x_2 + x_3 = 5, \\ x_1 + 2x_2 + 4x_3 = 16 \end{cases}$$

относительно неизвестных коэффициентов x_1, x_2, x_3 искомого многочлена $p(t)$.

Полученная система является несовместной. Будем решать ее методом наименьших квадратов, пользуясь средствами математического анализа. Для этого выпишем матрицу коэффициентов при неизвестных и вектор свободных членов

$$A=\begin{pmatrix}1 & -2 & 4\\ 1 & -1 & 1\\ 1 & 1 & 1\\ 1 & 2 & 4\end{pmatrix},\qquad b=\begin{pmatrix}9\\ 2\\ 5\\ 16\end{pmatrix}.$$

Вычислим матрицы

$$A^*A=\begin{pmatrix}1 & 1 & 1 & 1\\ -2 & -1 & 1 & 2\\ 4 & 1 & 1 & 4\end{pmatrix}\cdot\begin{pmatrix}1 & -2 & 4\\ 1 & -1 & 1\\ 1 & 1 & 1\\ 1 & 2 & 4\end{pmatrix}=\begin{pmatrix}4 & 0 & 10\\ 0 & 10 & 0\\ 10 & 0 & 34\end{pmatrix},$$

$$A^*b=\begin{pmatrix}1 & 1 & 1 & 1\\ -2 & -1 & 1 & 2\\ 4 & 1 & 1 & 4\end{pmatrix}\cdot\begin{pmatrix}9\\ 2\\ 5\\ 16\end{pmatrix}=\begin{pmatrix}32\\ 17\\ 107\end{pmatrix}.$$

Составим систему нормальных уравнений $A^*A\,x=A^*b$, которая в данном случае имеет вид

$$\begin{cases}4x_1 & & +\ 10x_3 & = & 32,\\ & 10x_2 & & = & 17,\\ 10x_1 & & +\ 34x_3 & = & 107.\end{cases}$$

Решив эту систему, найдем значения неизвестных $x_1=0,5$; $x_2=1,7$; $x_3=3$.

Таким образом, искомый многочлен, который с наименьшей квадратичной погрешностью аппроксимирует на отрезке $[-2,2]$ таблично заданную функцию $f(t)$, имеет вид

$$p(t)=0,5+1,7t+3t^2.$$

2. *Применение псевдообратной матрицы к решению систем линейных уравнений методом наименьших квадратов*

Решать системы линейных уравнений методом наименьших квадратов можно также с помощью псевдообратной матрицы.

Лемма 2. *Общее решение однородной системы* $Ax = 0$ *представимо в виде*

$$x_{\text{одн}} = (E - A^{+}A)\cdot c\,, \qquad (1)$$

где E *— единичная матрица соответствующего порядка* n; $c = (c_1, c_2, \ \ldots, \ c_n)^{\mathrm{T}}$ *— вектор произвольных постоянных.*

Доказательство. При любом векторе c столбец $(E - A^{+}A)\cdot c$ является решением системы $Ax = 0$, поскольку

$$A\cdot(E - A^{+}A)\cdot c = (A - A\,A^{+}A)\cdot c = (A - A)\cdot c = 0\,.$$

В то же время любое решение x_0 системы $Ax = 0$ можно представить в виде $x_0 = (E - A^{+}A)\cdot x_0$, так как

$$(E - A^{+}A)\cdot x_0 = x_0 - A^{+}A\,x_0 = x_0 - A^{+}\cdot 0 = x_0\,.$$

Это значит, что формула (1) содержит все решения системы $Ax = 0$.

Теорема 2. *Общее псевдорешение системы* $Ax = b$ *представимо в виде*

$$x = A^{+}b + (E - A^{+}A)\cdot c\,, \qquad (2)$$

где $c = (c_1, c_2, \ \ldots, \ c_n)^{\mathrm{T}}$ *— вектор произвольных постоянных.*

Доказательство. Общее псевдорешение системы $Ax = b$ совпадает с общим решением всегда совместной системы $A^{*}A\,x = A^{*}b$. Одним из решений системы $A^{*}A\,x = A^{*}b$ является столбец $A^{+}b$ (см. доказательство теоремы 1). Общее решение соответствующей однородной системы $A^{*}Ax = 0$ может быть представлено в виде (1). Значит, общее решение неоднородной системы $A^{*}A\,x = A^{*}b$ имеет вид (2).

Теорема 3. *Нормальное псевдорешение системы* $Ax = b$ *есть столбец*

$$x^0 = A^+ b. \tag{3}$$

Доказательство. Известно (см. доказательство теоремы 1 п. 4.2.1), что $x^0 = A^+ b$ является псевдорешением системы $Ax = b$. Пусть x — произвольное псевдорешение этой системы. В силу следствия из леммы 1 п. 4.2.1 это решение можно представить в виде $x = x^0 + z$, где $x^0 = A^+ b$ и z — произвольное решение системы $Ax = 0$. Тогда

$$|x|^2 = |x^0 + z|^2 = (x^0 + z)^*(x^0 + z) = |x^0|^2 + (x^0)^* z + z^* x^0 + |z|^2.$$

Поскольку столбец z является решением системы $Ax = 0$, а псевдообратная матрица A^+ может быть представлена в виде $A^+ = A^* V$, то

$$z^* x^0 = z^* A^+ b = z^* A^* V b = (Az)^* V b = 0 \cdot V b = 0.$$

При этом и

$$(x^0)^* z = (z^* x^0)^* = 0^* = 0.$$

Значит,

$$|x|^2 = |x^0|^2 + |z|^2$$

и наименьшее значение модуля псевдорешения $x = x^0 + z$ достигается при $z = 0$, т. е. $x^0 = A^+ b$ и есть нормальное псевдорешение системы $Ax = b$.

Следствие. Нормальное псевдорешение x^0 системы $Ax = b$ ортогонально к каждому решению системы $Ax = 0$.

Формула (3) является естественным обобщением формулы $x = A^{-1} b$, представляющей решение системы $Ax = b$ с квадратной невырожденной матрицей.

Учитывая доказанную теорему и следствие из леммы приходим к следующему ***утверждению***.

Теорема 4. *Общее псевдорешение системы* $Ax = b$ *представимо в виде*

$$x = A^{+} b + x_{\text{общ}},$$

где $x_{\text{общ}}$ *— общее решение однородной системы* $Ax = 0$, *которому, в частности, можно придать вид* (1).

Пример 7. Найти нормальное и общее псевдорешения системы уравнений $Ax = b$ при

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 1 & 3 & -3 \\ 1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 7 \\ 1 \\ 3 \end{pmatrix}.$$

Решение. В примере 2 для матрицы A найдена псевдообратная матрица

$$A^{+} = \frac{1}{574} \cdot \begin{pmatrix} 31 & 56 & 33 & 29 \\ 54 & 42 & 76 & 32 \\ -15 & 84 & -53 & 23 \end{pmatrix}.$$

Поэтому искомое нормальное псевдорешение

$$x^{0} = A^{+} b = \frac{1}{574} \cdot \begin{pmatrix} 31 & 56 & 33 & 29 \\ 54 & 42 & 76 & 32 \\ -15 & 84 & -53 & 23 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 7 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

У однородной системы уравнений $Ax = 0$, т. е. у системы

$$\begin{cases} x_1 + 2x_2 - x_3 = 0, \\ 2x_1 + x_2 + 4x_3 = 0, \\ x_1 + 3x_2 - 3x_3 = 0, \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

общее решение

$$x_{\text{одн}} = (-3x_3, 2x_3, x_3)^{\text{T}}.$$

Поэтому общим псевдорешением рассматриваемой системы является

$$x = x^0 + x_{\text{одн}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -3x_3 \\ 2x_3 \\ x_3 \end{pmatrix},$$

где переменная x_3 принимает любые числовые значения.

Пример 8. Найти нормальное псевдорешение системы уравнений $Ax = b$ при

$$A = \begin{pmatrix} 7 & 1 & 2 \\ 5 & 5 & -2 \\ 3 & 3 & 6 \\ 1 & 7 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 2 \\ -6 \\ -8 \end{pmatrix}.$$

Решение. В примере 3 для матрицы A найдена псевдообратная матрица

$$A^+ = V\,\Sigma^+\,U^* = \frac{1}{72}\cdot\begin{pmatrix} 8 & 4 & 0 & -4 \\ -4 & 4 & 0 & 8 \\ 1 & -7 & 9 & 1 \end{pmatrix}.$$

Поэтому искомое нормальное псевдорешение

$$x^0 = A^+ b = \frac{1}{72}\cdot\begin{pmatrix} 8 & 4 & 0 & -4 \\ -4 & 4 & 0 & 8 \\ 1 & -7 & 9 & 1 \end{pmatrix}\cdot\begin{pmatrix} 4 \\ 2 \\ -6 \\ -8 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$