Тема V: Линейные операторы

§ 5. Системы линейных уравнений

Б.М.Верников   М.В.Волков

Уральский федеральный университет

Институт естественных наук и математики

кафедра алгебры и фундаментальной информатики

2021/2022 учебный год

Постановка задачи

Мы установили необходимое и достаточное условие *совместности* системы линейных уравнений $A\mathbf{x} = \mathbf{b}$. Теперь обсудим, как *решать* такие системы.

Эта задача делится на две подзадачи:

- решение совместных систем;
- решение несовместных систем.

Сегодня займемся первой подзадачей; вторую (более сложную и практически более важную подзадачу) обсудим немного позднее.

Система линейных уравнений называется *однородной*, если свободные члены всех уравнений системы нулевые:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = 0. \end{cases}$$

Однородная система всегда совместна. Сведем нахождение решений произвольной совместной системы к нахождению решений однородной системы, а затем изучим строение решений однородной системы.

Пусть $A\mathbf{x} = \mathbf{b}$ – произвольная совместная система. Соответствующая ей однородная система $A\mathbf{x} = \mathbf{0}$ получается, если заменить столбец свободных членов нулевым столбцом.

Замечание

Если $\mathbf{x}_0$ – некоторое решение системы $A\mathbf{x} = \mathbf{b}$, то вектор-столбец $\mathbf{x}_1$ будет решением системы $A\mathbf{x} = \mathbf{b}$ тогда и только тогда, когда $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}$, где $\mathbf{y}$ – решение соответствующей однородной системы $A\mathbf{x} = \mathbf{0}$.

Доказательство. Если $\mathbf{x}_1$ – решение системы $A\mathbf{x} = \mathbf{b}$, положим $\mathbf{y} := \mathbf{x}_1 - \mathbf{x}_0$. Тогда

$$A\mathbf{y} = A(\mathbf{x}_1 - \mathbf{x}_0) = A\mathbf{x}_1 - A\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Итак, $\mathbf{y}$ – решение однородной системы $A\mathbf{x} = \mathbf{0}$ и $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}$.

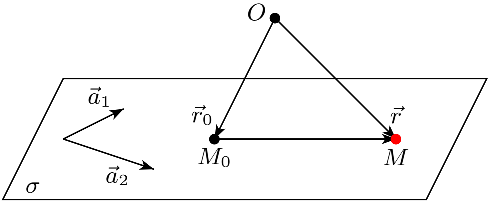
Обратно, если $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}$, где $\mathbf{y}$ – решение однородной системы, то

$$A\mathbf{x}_1 = A(\mathbf{x}_0 + \mathbf{y}) = A\mathbf{x}_0 + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Отсюда $\mathbf{x}_1$ – решение системы $A\mathbf{x} = \mathbf{b}$. $\square$

Доказанное выше замечание показывает, что если научиться находить всевозможные решения однородных систем, то для нахождения *всех* решений данной системы $A\mathbf{x} = \mathbf{b}$ достаточно найти какое-нибудь *одно* решение этой системы. Эту мысль часто выражают так: *общее решение системы $A\mathbf{x} = \mathbf{b}$ равно сумме какого-то частного решения этой системы и общего решения соответствующей однородной системы*.

Отметим еще геометрическую интерпретацию. В обычном трехмерном пространстве системы линейных уравнений задают прямые или плоскости, а однородные системы – прямые или плоскости, проходящие через начало координат, т.е. одномерные или двумерные подпространства. Замечание говорит, что любую точку прямой или плоскости можно получить, отложив от какой-то начальной точки этой прямой или плоскости подходящий вектор из направляющего подпространства этой прямой или плоскости.



Предложение

Множество решений однородной системы $A\mathbf{x} = \mathbf{0}$ *образует подпространство в пространстве столбцов.*

Доказательство. Если A – $k \times n$-матрица, то правило $\mathcal{A}(\mathbf{x}) := A\mathbf{x}$ определяет линейный оператор $\mathcal{A}\colon V_1 \to V_2$ из пространства столбцов высоты n в пространство столбцов высоты k. При этом матрица A будет матрицей этого оператора $\mathcal{A}$ (в стандартных базисах пространств V_1 и V_2), а множество решений системы $A\mathbf{x} = \mathbf{0}$ будет ядром оператора $\mathcal{A}$. Ядро линейного оператора является подпространством. $\square$

Если пространство решений однородной системы ненулевое, то любой базис этого пространства называется *фундаментальной системой решений*. Если $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_d$ – фундаментальная система решений системы $A\mathbf{x} = \mathbf{0}$, то любое решение $\mathbf{y}$ этой системы однозначно представимо в виде

$$\mathbf{y} = c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_d\mathbf{y}_d, \qquad (*)$$

где $c_1, c_2, \ldots, c_d$ – некоторые скаляры. Выражение $(*)$ принято называть *общим решением* системы $A\mathbf{x} = \mathbf{0}$.

Итак, *решить однородную систему линейных уравнений – значит построить для нее фундаментальную систему решений*.
Как это сделать?
Прежде всего, ответим на вопрос, как узнать, сколько решений входит в фундаментальную систему.

Теорема о размерности пространства решений однородной системы

Размерность пространства решений системы $A\mathbf{x} = \mathbf{0}$ *равна* $n - r$*, где* n *– число неизвестных в системе, а* r *– ранг матрицы* A.

Доказательство. Снова рассмотрим линейный оператор $\mathcal{A}$ из пространства столбцов высоты n в пространство столбцов высоты k, определяемый как умножение вектора-столбца на матрицу A слева, и применим к $\mathcal{A}$ теорему о ранге и дефекте. По этой теореме сумма ранга (размерности образа $\mathcal{A}$) и дефекта (размерности ядра $\mathcal{A}$) равна размерности пространства столбцов высоты n, т.е. n. Так как ранг линейного оператора совпадает с рангом его матрицы, ранг оператора $\mathcal{A}$ равен r. Ядро оператора $\mathcal{A}$ – это пространство решений системы, поэтому размерность последнего равна $n - r$. $\square$

Рассмотрим произвольную однородную систему.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = 0. \end{cases}$$

Пусть ранг ее матрицы A равен $r < n$. (В случае $r = n$ у системы есть единственное решение $\mathbf{0}$, и фундаментальной системы решений нет.)В силу теоремы о ранге в A есть r линейно независимых строк, а любой набор из более, чем r ее строк линейно зависим. Переставляя уравнения, можно считать, что первые r строк матрицы A линейно независимы, а все последующие строки линейно выражаются через первые r строк. Следовательно, все уравнения нашей системы, начиная с $(r+1)$-го, являются следствиями первых r уравнений. Вычеркнув все уравнения, начиная с $(r+1)$-го, получим систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rn}x_n = 0, \end{cases}$$

равносильную исходной. Обозначим матрицу этой системы через A'; ясно, что ранг A' равен r. Поэтому у A' есть r линейно независимых столбцов.

Построение фундаментальной системы решений (2)

Переставляя столбцы матрицы A' и переименовывая переменные, получим систему с матрицей, в которой первые r столбцов линейно независимы:

$$\begin{cases} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1r}y_r + b_{1\,r+1}y_{r+1} + \cdots + b_{1n}y_n = 0, \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2r}y_r + b_{2\,r+1}y_{r+1} + \cdots + b_{2n}y_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ b_{r1}y_1 + b_{r2}y_2 + \cdots + b_{rr}y_r + b_{r\,r+1}y_{r+1} + \cdots + b_{rn}y_n = 0; \end{cases}$$

здесь $\{y_1, \dots, y_n\} = \{x_1, \dots, x_n\}$, а матрица $(b_{ij})_{r\times n}$ получается из матрицы A' перестановкой столбцов. Перенеся слагаемые, содержащие неизвестные $y_{r+1}, \dots, y_n$, в правую часть, получим систему

$$\begin{cases} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1r}y_r = -b_{1\,r+1}y_{r+1} - \cdots - b_{1n}y_n, \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2r}y_r = -b_{2\,r+1}y_{r+1} - \cdots - b_{2n}y_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ b_{r1}y_1 + b_{r2}y_2 + \cdots + b_{rr}y_r = -b_{r\,r+1}y_{r+1} - \cdots - b_{rn}y_n. \end{cases} \qquad (\dagger)$$

Неизвестные $y_{r+1}, \dots, y_n$ называются *свободными*, а неизвестные $y_1, \dots, y_r$ – *связанными*. Будем смотреть на систему $(\dagger)$ как на (неоднородную) систему r линейных уравнений относительно r неизвестных $y_1, \dots, y_r$.

Построение фундаментальной системы решений (3)

В матричном виде систему (†)

$$\begin{cases} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1r}y_r = -b_{1\,r+1}y_{r+1} - \cdots - b_{1n}y_n, \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2r}y_r = -b_{2\,r+1}y_{r+1} - \cdots - b_{2n}y_n, \\ \dotfill \\ b_{r1}y_1 + b_{r2}y_2 + \cdots + b_{rr}y_r = -b_{r\,r+1}y_{r+1} - \cdots - b_{rn}y_n. \end{cases} \qquad (\dagger)$$

можно записать как $B\mathbf{y} = \mathbf{c}$, где $B = (b_{ij})_{r\times r}$ – матрица

из коэффициентов при неизвестных $y_1, \ldots, y_r$, $\mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix}$,

а $\mathbf{c} := \begin{pmatrix} -b_{1\,r+1}y_{r+1} - \cdots - b_{1n}y_n, \\ \vdots \\ -b_{r\,r+1}y_{r+1} - \cdots - b_{rn}y_n \end{pmatrix}$. Матрица B обратима, поскольку ее столбцы линейно независимы. Умножая равенство $B\mathbf{y} = \mathbf{c}$ слева на B^{-1}, получаем единственное решение системы (†) в виде $\mathbf{y} = B^{-1}\mathbf{c}$.

Переходя к координатам, имеем $\begin{cases} y_1 = c_{1\,r+1}y_{r+1} + \cdots + c_{1n}y_n, \\ y_2 = c_{2\,r+1}y_{r+1} + \cdots + c_{2n}y_n, \\ \dotfill \\ y_r = c_{r\,r+1}y_{r+1} + \cdots + c_{rn}y_n. \end{cases}$

Построение фундаментальной системы решений (4)

Из формул

$$\begin{cases} y_1 = c_{1\,r+1}y_{r+1} + \cdots + c_{1n}y_n, \\ y_2 = c_{2\,r+1}y_{r+1} + \cdots + c_{2n}y_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_r = c_{r\,r+1}y_{r+1} + \cdots + c_{rn}y_n. \end{cases} \qquad (\star)$$

можно извлекать решения исходной системы, придавая свободным неизвестным $y_{r+1}, \dots, y_n$ *произвольные* значения (поэтому они и названы «свободными»), вычисляя соответствующие значения связанных неизвестных $y_1, \dots, y_r$ и переходя к исходным неизвестным $x_1, \dots, x_n$.

Для каждого $i = r+1, \dots, n$ придадим свободной неизвестной y_i значение 1, а всем остальным свободным неизвестным – значение 0. Вычислив по формулам $(\star)$ соответствующие значения связанных неизвестных, получим $n-r$ решений $\mathbf{y}_1 := \begin{pmatrix} c_{1\,r+1} \\ \vdots \\ c_{r\,r+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{y}_{n-r} := \begin{pmatrix} c_{1n} \\ \vdots \\ c_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.

Утверждается, что эти решения образуют фундаментальную систему решений для системы

$$\begin{cases} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1r}y_r + b_{1\,r+1}y_{r+1} + \cdots + b_{1n}y_n = 0, \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2r}y_r + b_{2\,r+1}y_{r+1} + \cdots + b_{2n}y_n = 0, \\ \dotfill \\ b_{r1}y_1 + b_{r2}y_2 + \cdots + b_{rr}y_r + b_{r\,r+1}y_{r+1} + \cdots + b_{rn}y_n = 0. \end{cases}$$

Действительно, их число равно размерности $n-r$ пространства решений системы, а сами эти решения линейно независимы, что сразу видно, если

составить из них $n\times(n-r)$-матрицу :
$$\begin{array}{cccc} \mathbf{y}_1 & \mathbf{y}_2 & \ldots & \mathbf{y}_{n-r} \\ \end{array}$$
$$\begin{pmatrix} c_{1\,r+1} & c_{1\,r+2} & \ldots & c_{1n} \\ c_{2\,r+1} & c_{2\,r+2} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{r\,r+1} & c_{r\,r+2} & \ldots & c_{rn} \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Нижние $n-r$ строк линейно независимы, откуда ранг равен $n-r$.

Переход к исходным неизвестным $x_1, \ldots, x_n$ означает перестановку строк матрицы

$$\begin{array}{cccc}
\mathbf{y}_1 & \mathbf{y}_2 & \ldots & \mathbf{y}_{n-r} \\
\end{array}$$
$$\begin{pmatrix}
c_{1\,r+1} & c_{1\,r+2} & \ldots & c_{1n} \\
c_{2\,r+1} & c_{2\,r+2} & \ldots & c_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
c_{r\,r+1} & c_{r\,r+2} & \ldots & c_{rn} \\
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{pmatrix}.$$

Столбцы полученной при этой перестановке строк матрицы останутся линейно независимыми и потому образуют фундаментальную систему решений для исходной системы.

На практике фундаментальную систему решений ищут, приводя систему к ступенчатому виду. Разберем пример:

$$\begin{cases} x_1+3x_2+x_3-2x_4-2x_5 &= 0 \\ -x_1-7x_2-2x_3+x_4+2x_5 &= 0 \\ -x_1-11x_2-3x_3+2x_5 &= 0 \\ -2x_1-2x_2-x_3+5x_4+4x_5 &= 0 \end{cases}$$

$$\begin{pmatrix} 1 & 3 & 1 & -2 & -2 \\ -1 & -7 & -2 & 1 & 2 \\ -1 & -11 & -3 & 0 & 2 \\ -2 & -2 & -1 & 5 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 1 & -2 & -2 \\ 0 & -4 & -1 & -1 & 0 \\ 0 & -8 & -2 & -2 & 0 \\ 0 & 4 & 1 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 1 & -2 & -2 \\ 0 & 4 & 1 & 1 & 0 \end{pmatrix}$$

Получили ступенчатую систему $\begin{cases} x_1+3x_2+x_3-2x_4-2x_5 &= 0 \\ \phantom{x_1+{}}4x_2+x_3+x_4 &= 0 \end{cases}$.

Свободные неизвестные – x_3, x_4, x_5, связанные – x_1, x_2. Фундаментальная система решений состоит из $\mathbf{x}_1 = \begin{pmatrix} -0,25 \\ -0,25 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 2,75 \\ -0,25 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{x}_3 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

Процедура решения совместной системы линейных уравнений

Пусть дана система линейных уравнений $A\mathbf{x} = \mathbf{b}$ с n неизвестными.

1. Элементарными преобразованиями строк приводим матрицу $A|b$ к ступенчатому виду.
2. Если ранг r матрицы A меньше ранга матрицы $A|b$, система $A\mathbf{x} = \mathbf{b}$ несовместна. Если ранги равны, находим частное решение $\mathbf{x}_0$ этой системы.
3. Находим фундаментальную систему решений $\mathbf{x}_1, \dots, \mathbf{x}_{n-r}$ соответствующей однородной системы $A\mathbf{x} = \mathbf{0}$.
4. Выражение

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_{n-r}\mathbf{x}_{n-r},$$

   где $c_1, c_2, \dots, c_{n-r}$ – произвольные скаляры, дает *общее решение* системы $A\mathbf{x} = \mathbf{b}$. Каждое решение системы получается из общего при некотором (однозначно определяемом) наборе $c_1, c_2, \dots, c_{n-r}$.

Трудоемкость описанной процедуры $O(m^3)$, где m – число уравнений.

Приложение – извлечение данных

Задача

Дано: множество E (огромного!) размера m и функция $f\colon E \to \{0,1\}$.
Требуется: Структура данных R, которая по $y \in E$ возвращает $f(y)$.

Требования на R по времени и памяти

Память: $(1+\varepsilon)m$ бит для некоторой маленькой константы ε.
Время: константа (не зависит от m).

Заметим, что каждый элемент $y \in E$ может быть относительно большим, поэтому хранить массив всех пар $(y, f(y))$ не является решением (требует $m(\max\{|y|\}+1)$ бит и не допускает быстрого извлечения).

Пример: Результаты тестов на ковид

$$
\begin{array}{lccccccc}
E = \{ & \texttt{Ana}, & \texttt{Bea}, & \texttt{Cal}, & \texttt{Dan}, & \texttt{Eli}, & \texttt{Fen}, & \ldots \quad \} \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
f: & 1 & 1 & 0 & 0 & 0 & 1 &
\end{array}
$$

Извлечение данных трудно – ошибка Ватсона

Watson – это знаменитый суперкомьютер фирмы IBM.
В 2011 г. Watson выиграл в интеллектуальном шоу «Jeopardy!» у двух самых знаменитых белковых чемпионов.

Однако Watson допускал ошибки. Например, он не угадал ответ на такой вопрос из категории «Города США»: «Его крупнейший аэропорт назван в честь героя второй мировой войны, а его второй по величине аэропорт – в честь битвы этой войны». Ответ – Чикаго, аэропорты О'Хара и Мидуэй.

Важно понимать, что Watson имел все необходимые данные в своей 15-терабайтной базе знаний! Неудача была связана не с отсутствием данных, а с неспособностью быстро извлечь нужное из огромной базы!

Возьмем $n = (1+\varepsilon)m$ и хэш-функцию $h\colon E \to \{1,\dots,n\}^3$.
Функция h должна быть легко вычислимой и взаимно однозначной.
Это дает $m \times n$-матрицу A над двухэлементным полем $\mathbb{F} = \{0,1\}$:

Input	Hash Values
`Ana`	$h(\texttt{Ana}) = (1,3,9)$
`Bea`	$h(\texttt{Bea}) = (2,3,4)$
`Cal`	$h(\texttt{Cal}) = (3,6,8)$
`Dan`	$h(\texttt{Dan}) = (5,8,9)$
`Eli`	$h(\texttt{Eve}) = (2,8,9)$
`Fen`	$h(\texttt{Fen}) = (1,5,6)$

$$\begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array} \cdot \vec{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Структура $R = (h, \vec{x})$

query$(y) := \sum_{i=1}^{3} \vec{x}[h_i(y)]$ – **константное время** (сумма 3 бит)

Требуемая память: $1.09m$ бит (или $1.23m$ бит при конструкции с $O(m)$).