Институт механики Академии Наук Союза ССР
Прикладная математика и механика. Том XVI, 1952

О РАВНОСИЛЬНОСТИ И СОВМЕСТНОСТИ СИСТЕМ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

Б. В. Булгаков

(Москва)

В этой статье имеется в виду получить общие теоремы о равносильности и совместности систем, указанных в заглавии, предполагая число уравнений m и число неизвестных n любыми, хотя бы и не равными числами. Для случая $m = n$ этот вопрос, казалось бы, столь простой, рассматривался в одном исследовании Н. Н. Лузина [1], имеющем прикладное значение, где были даны строгий вывод условий равносильности и две теоремы о совместности. Однако вторая из них, понимаемая в том смысле, который ей придавал автор, неверна, в чем можно убедиться на противоречащем примере в конце настоящей статьи. Цель работы заключается в том, чтобы, сохраняя идею покойного математика, исправить его теорему, что он, несомненно, сделал бы сам. Основным средством для этого будет измененное определение ранга расширенной матрицы.

§ 1. Ранг расширенной матрицы. Пусть имеем матричное уравнение

$$f(D)\,y = w(t) \tag{1.1}$$

где

$$f(D) = \| f_{jk}(D) \| = \| a_{jk}^{(0)} D^p + a_{jk}^{(1)} D^{p-1} + \cdots + a_{jk}^{(p)} \| \qquad (m \times n) \tag{1.2}$$

$$D = d \,/\, dt \tag{1.3}$$

$$y = \left\| \begin{matrix} y_1 \\ \vdots \\ y_n \end{matrix} \right\| \quad (n \times 1), \qquad w(t) = \left\| \begin{matrix} w_1(t) \\ \vdots \\ w_m(t) \end{matrix} \right\| \quad (m \times 1) \tag{1.4}$$

Матричная символика и терминология здесь и в дальнейшем либо общеприняты, либо соответствуют [2]. Матрица $f(D)$ называется операционной, матрица

$$\| f(D),\ w(t) \| \tag{1.5}$$

расширенной. Уравнение (1) может быть записано с помощью расширенной матрицы в виде

$$\| f(D),\ w(t) \| \times \left\| \begin{matrix} y \\ -1 \end{matrix} \right\| = 0 \tag{1.6}$$

Все миноры операционной матрицы и те миноры расширенной, в которые не входит столбец свободных членов, являются полиномами от D. Те миноры расширенной матрицы, в которые входят $i - 1$ столбцов операционной матрицы и столбец свободных членов, будем всегда сокращать на общий наибольший делитель всех миноров $(i-1)$-го порядка, содержащихся в данных $i-1$ столбцах, и называть в этой форме измененными минорами; они должны рассматриваться как функции от t, получаемые путем развертывания по элементам последнего столбца и выполнения над ними дифференцирований, соответствующих различным степеням D.

Ранги операционной и расширенной матриц определяются обычным образом как наивысшие порядки их неисчезающих миноров; но исчезание, о котором идет речь, должно пониматься как тождественное по D для миноров, в которые не входит столбец свободных членов, и как тождественное по t для измененных миноров.

§ 2. Элементарные преобразования матриц. Расширяя обычный список, будем называть элементарными преобразованиями квадратной или прямоугольной полиномной матрицы следующие.

1. Перестановка двух строк (столбцов).

2. Умножение всех элементов какой-либо строки (столбца) на одну и ту же, отличную от нуля постоянную.

3. Прибавление к элементам строки (столбца) соответственных элементов другой строки (столбца), умноженных на один и тот же полином.

4. Присоединение или отбрасывание последней строки (столбца), состоящего из нулей.

Последнее преобразование позволяет менять число строк и столбцов. В отличие от других для его выполнимости в случае отбрасывания не безразлично, каковы элементы матрицы, к которой оно применяется. Однако оно разделяет с остальными известное свойство, которое выражает следующая теорема.

Теорема I. Каждое из элементарных преобразований равносильно умножению данной матрицы слева или справа на некоторую элементарную полиномную матрицу.

Например, при $m=3$ такими элементарными матрицами для четвертого преобразования над строками будут

$$\left\|\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{matrix}\right\|, \qquad \left\|\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\end{matrix}\right\| \tag{2.1}$$

Полиномные матрицы, получаемые друг из друга путем конечного числа элементарных преобразований, называются эквивалентными. Детерминантные делители эквивалентных матриц одинаковы. Для того чтобы две полиномные матрицы были эквивалентны, необходимо и достаточно, чтобы они имели один и тот же ранг и одни и те же элементарные делители. Если элементарные преобразования производятся только над строками или только над столбцами, то мы будем говорить соответственно об эквивалентности слева или справа. Так как нас главным образом будут интересовать элементарные преобразования над строками, то мы для них и будем формулировать следующие теоремы.

Теорема II. Полиномные матрицы, эквивалентные слева, могут быть получены друг из друга путем умножения слева на некоторые другие полиномные матрицы.

В дальнейшем будем предполагать, что если ранг данной матрицы равен r, то уже в первых r столбцах можно найти неисчезающий минор r-го порядка. Это всегда может быть достигнуто, если нужно, при помощи первого элементарного преобразования над столбцами,

Теорема III. Среди всех полиномных матриц, эквивалентных между собой слева, имеется единственная матрица, имеющая каноническую форму:

$$f^*(\lambda) = \begin{Vmatrix} f_{11}^*(\lambda) & f_{12}^*(\lambda) \ldots & f_{1r}^*(\lambda) \ldots & f_{1n}^*(\lambda) \\ 0 & f_{22}^*(\lambda) \ldots & f_{2r}^*(\lambda) \ldots & f_{2n}^*(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 \ \ldots & f_{rr}^*(\lambda) \ldots & f_{rn}^*(\lambda) \end{Vmatrix} \tag{2.2}$$

которая отличается тем, что коэффициент старшего члена в каждом элементе $f_{jj}^*(\lambda)$ $(j = 1, \ldots, r)$ равен единице и что этот полином является старшим в своем столбце.

Чтобы доказать самое существование такой матрицы, можно исходить от любой из данных эквивалентных матриц $f(\lambda)$ и, так же как у Лузина[1] (стр. 8, 18), привести ее с помощью первых трех элементарных преобразований над строками к эквивалентной матрице, отличающейся от $f^*(\lambda)$, быть может, несколькими лишними строками, состоящими из нулей. Эти строки можно отбросить с помощью четвертого элементарного преобразования.

Если теперь $f^*(\lambda)$, $f_a^*(\lambda)$ — две канонические формы, полученные описанным путем, то, будучи эквивалентны слева, они могут быть получены друг из друга умножением слева на полиномные матрицы:

$$f_a^*(\lambda) = u(\lambda) f^*(\lambda)$$

Рассуждая далее так же, как у Лузина[1] (стр. 12), убедимся, что $u(\lambda) = E_r$ и что $f^*(\lambda) = f_a^*(\lambda)$.

Следующая теорема является обратной для теоремы II.

Теорема IV. Полиномные матрицы, получающиеся друг из друга умножением слева на некоторые другие полиномные, эквивалентны слева.

В самом деле, первое из двух соотношений

$$f_a(\lambda) = u(\lambda) f(\lambda), \qquad f(\lambda) = v(\lambda) f_a(\lambda) \tag{2.3}$$

позволяет утверждать на основании известной теоремы о ранге произведения двух матриц, что ранг $f_a(\lambda)$ не превосходит ранга $f(\lambda)$; из второго соотношения следует, что ранг $f(\lambda)$ не превосходит ранга $f_a(\lambda)$. Поэтому эти ранги одинаковы, и канонические формы $f^*(\lambda)$, $f_a^*(\lambda)$ будут одного и того же типа $r \times r$. Эти канонические формы могут быть получены из $f(\lambda)$, $f_a(\lambda)$, а значит, и друг из друга умножением слева на полиномные матрицы:

$$f_a^*(\lambda) = u^*(\lambda) f^*(\lambda)$$

Рассуждая опять так же, как у Лузина[1] (стр. 12), убеждаемся, что $u^*(\lambda) = E_r$ и $f^*(\lambda) = f_a^*(\lambda)$. Отсюда вытекает эквивалентность матриц $f(\lambda)$ и $f_a(\lambda)$ слева.

Элементарные преобразования над строками могут быть применены также к расширенным матрицам и приводят к матрицам той же структуры, т. е. содержащим в первых n столбцах полиномы от D, а в столбце $n+1$ — функции от t.

Расширенные матрицы, эквивалентные слева, имеют один и тот же ранг; в этом легко убедиться, применяя обычное доказательство, но пользуясь измененными минорами.

Теорема V. Расширенные матрицы, эквивалентные слева, могут быть получены друг из друга путем умножения слева на некоторые полиномные матрицы.

§ 3. Линейные преобразования матричного уравнения. Умножая обе части уравнения (1.1) слева на полиномную матрицу $u(D)$ типа $m' \times m$, где вообще $m \neq m'$, получим преобразованное уравнение

$$f_a(D)\, y = w_a(t) \tag{3.1}$$

с расширенной матрицей

$$\| f_a(D),\ w_a(t) \| = u(D)\ \| f(D),\ w(t) \| \qquad (m' \times n+1) \tag{3.2}$$

Согласно определению умножения матриц связь между скалярными системами, соответствующими первоначальному и преобразованному матричным уравнениям, заключается в следующем: j-е уравнение второй системы $(j = 1, \ldots, m')$ получается путем умножения уравнений первой на $u_{j1}(D), \ldots, u_{jm}(D)$ и сложения результатов. Поэтому каждое решение первого матричного уравнения удовлетворяет и второму. Обратное будет верно, если, наряду с (3.2), имеет место соотношение

$$\| f(D),\ w(t) \| = v(D)\ \| f_a(D),\ w_a(t) \| \qquad (m \times n+1) \tag{3.3}$$

где $v(D)$ — полиномная матрица типа $m \times m'$. Таким образом, верна следующая теорема.

Теорема VI. Если расширенные матрицы двух уравнений получаются друг из друга путем умножения слева на некоторые полиномные матрицы, то уравнения равносильны, т. е. каждое решение одного есть также решение другого.

Непосредственно очевидна также равносильность уравнений, расширенные матрицы которых эквивалентны слева. С помощью теоремы V этот случай можно также свести к предыдущему. Если известна цепь элементарных преобразований, приводящих от одной матрицы к другой, то, перемножая соответствующие элементарные матрицы, тотчас получим $u(D), v(D)$. Если, в частности, применялись только первые три преобразования, то $u(D), v(D)$ будут квадратными, неособыми и обратными друг другу.

§ 4. Условие существования решений. *Теорема VII.* Для существования решений матричного линейного уравнения необходимо и достаточно, чтобы операционная и расширенная матрицы имели один и тот же ранг.

Для доказательства обозначим через r ранг операционной матрицы и перенумеруем неизвестные таким образом, чтобы по крайней мере один из неисчезающих миноров r-го порядка принадлежал первым r столбцам. Так же как и при доказательстве теоремы III, приведем операционную матрицу при помощи первых трех элементарных преобразований к форме, отличающейся от $f(D)$ только $m-r$ последними строками, состоящими из нулей. Если те же преобразования произвести и над расширенной матрицей, то скалярная система примет вид:

$$\begin{array}{l}
f_{11}^*(D)y_1 + f_{12}^*(D)y_2 + \cdots + f_{1r}^*(D)y_r + \cdots + f_{1n}^*(D)y_n = w_1^*(t) \\
\quad\star \qquad f_{22}^*(D)y_2 + \cdots + f_{2r}^*(D)y_r + \cdots + f_{2n}^*(D)y_n = w_2^*(t) \\
\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (4.1) \\
\quad\star \qquad \star \qquad\qquad f_{rr}^*(D)y_r + \cdots + f_{rn}^*(D)y_n = w_r^*(t) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 = w_{r+1}^*(t) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdots\cdots \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 = w_m^*(t)
\end{array}$$

Первые r строк операционной матрицы этой системы заключают некоторое число неисчезающих миноров r-го порядка, причем каждый будет единственным таким минором в субматрице, образованной из r входящих в него столбцов. Поэтому он может только постоянным множителем отличаться от общего наибольшего делителя миноров r-го порядка этой субматрицы. Неисчезающими минорами $(r+1)$-го порядка расширенной матрицы, если они существуют, могут быть только измененные миноры, образованные из первых r и одной из остальных строк и из r столбцов одной из рассмотренных субматриц и столбца свободных членов. Каждый такой минор будет произведением постоянной на одну из функций $w_{r+1}^*(t), \ldots, w_m^*(t)$. Поэтому равенство рангов операционной и расширенной матриц представляет условие, необходимое и достаточное для того, чтобы

$$w_{r+1}^*(t) \equiv \cdots \equiv w_m^*(t) \equiv 0 \qquad (4.2)$$

Эти тождества в свою очередь представляют условия, необходимые и достаточные для того, чтобы скалярная система (4.1) была непротиворечива и допускала решения, так как если все они не имеют места, то последним $m-r$ уравнениям или по крайней мере некоторым из них нельзя удовлетворить, если же все тождества выполняются, то последние $m-r$ уравнений получают вид $0=0$, а остальные могут быть последовательно разрешены, начиная с r-го, относительно y_r, затем относительно y_{r-1} и т. д. При этом неизвестные функции $y_{r+1}, \ldots, y_n$ остаются произвольными. Этим заканчивается доказательство.

Вместо рангов операционной и расширенной матриц можно было бы в тексте теоремы говорить только о рангах субматрицы, составленной из первых r столбцов, и расширенной субматрицы, получаемой добавлением столбца свободных членов.

Разрешая относительно $y_1, \ldots, y_r$ первые r уравнений (4.1), мы имеем усеченную систему, порядок которой равен сумме степеней полиномов $f_{11}^*(D), \ldots, f_{rr}^*(D)$ или же степени общего наибольшего делителя миноров r-го порядка той субматрицы первоначальной системы, о которой только что была речь; этим общим наибольшим делителем приходится пользоваться при вычислении измененных миноров.

Доказанная теорема представляет вторую теорему Лузина о совместности, исправленную при помощи определения ранга расширенной матрицы, основанного на измененных минорах.

Укажем еще две теоремы о расширенных матрицах, при формулировке и доказательстве которых попрежнему предполагается, что по крайней мере один неисчезающий минор порядка, равного рангу r, заключается в первых r столбцах операционной матрицы.

Теорема VIII. Среди всех расширенных матриц, имеющих тот же ранг, что и принадлежащие им операционные, и эквивалентных между собой слева, имеется единственная, имеющая каноническую форму $\| f^*(D), w^*(t) \|$, где $f^*(\lambda)$ определяется согласно теореме III.

Теорема IX. Расширенные матрицы, имеющие тот же ранг, что и принадлежащие им операционные, и получающиеся друг из друга умножением слева на некоторые другие операционные, эквивалентны слева.

Доказательство легко провести по образцу аналогичных теорем о полиномных матрицах.

§ 5. **Пример**. Пусть имеем систему

$$\begin{aligned} -y_1 \qquad\quad - y_2 \qquad + \dot{y}_3 - 4y_3 &= -z(t) \\ \dot{y}_1 - 2y_1 \qquad + y_2 \qquad + \dot{y}_3 - 2y_3 &= z(t) \\ \ddot{y}_1 - 2\dot{y}_1 - y_1 + \dot{y}_2 - y_2 + \ddot{y}_3 - \dot{y}_3 - 4y_3 &= -2z(t) \end{aligned} \tag{5.1}$$

где $z(t)$ — заданная функция. Расширенная матрица есть

$$\| f(D) \;\; w(t) \| = \begin{Vmatrix} -1 & -1 & D-4 & -z(t) \\ D-2 & 1 & D-2 & z(t) \\ D^2-2D-1 & D-1 & D^2-D-4 & -2z(t) \end{Vmatrix} \tag{5.2}$$

Ранг операционной матрицы равен двум. Для трех неизмененных миноров третьего порядка расширенной матрицы, содержащих столбец свободных членов, получаем выражения

$$2(D-3)(D+1)z(t), \qquad (D-2)(D-3)(D+1)z(t), \qquad -(D-3)(D+1)z(t) \tag{5.3}$$

Принимая во внимание, что общие наибольшие делители миноров второго порядка, заключающихся соответственно во 2-м и 3-м, в 3-м и 1-м и, наконец, в 1-м и 2-м столбцах, суть

$$D-3, \qquad (D-2)(D-3), \qquad D-3 \tag{5.4}$$

находим измененные миноры третьего порядка:

$$2(D+1)z(t), \qquad (D+1)z(t), \qquad -(D+1)z(t) \tag{5.5}$$

Приравнивая их нулю, получаем условие существования решений

$$(D+1)\,z(t)=0 \tag{5.6}$$

или

$$z(t)=ce^{-t} \tag{5.7}$$

где c — постоянная.

Умножая $\| f(D),\ w(t) \|$ слева на матрицу

$$u_1(D)=\begin{Vmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & -D & 1 \end{Vmatrix}$$

получим преобразованную расширенную матрицу

$$\| f_a(D),\ w_a(t) \| = \begin{Vmatrix} D-3 & 0 & 2(D-3) & 0 \\ 1 & 1 & -(D-4) & z(t) \\ 0 & 0 & 0 & -(D+1)\,z(t) \end{Vmatrix}$$

причем первоначальная расширенная матрица в свою очередь получается из преобразованной умножением ее слева на

$$v_1(D)=\begin{Vmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ D & D-1 & 1 \end{Vmatrix}$$

Первым двум строкам матрицы $\| f_a(D),\ w_a(t) \|$ соответствует уравнение

$$\begin{Vmatrix} D-3 & 0 \\ 1 & 1 \end{Vmatrix} \times \begin{Vmatrix} y_1 \\ y_2 \end{Vmatrix} = \begin{Vmatrix} -2(D-3) \\ D-4 \end{Vmatrix} \varphi(t) + \begin{Vmatrix} 0 \\ z(t) \end{Vmatrix} \tag{5.8}$$

в котором $y_3=\varphi(t)$ есть произвольная функция, а последней строке соответствует уравнение

$$0=-(D+1)\,z(t) \tag{5.9}$$

Если условие (5.7) выполняется, то первое уравнение дает

$$\begin{Vmatrix} y_1 \\ y_2 \end{Vmatrix} = \begin{Vmatrix} -2 \\ D-2 \end{Vmatrix} \varphi(t) + \begin{Vmatrix} 0 \\ 1 \end{Vmatrix} ce^{-t} + \begin{Vmatrix} 1 \\ -1 \end{Vmatrix} Ce^{3t} \tag{5.10}$$

а второе обращается в тождество.

В матрице $\| f_a(D),\ w_a(t) \|$ можно отбросить последнюю строку или, что то же, умножить эту матрицу слева на

$$u_2(D)=\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}$$

В результате получим каноническую форму

$$\| f^*(D),\ w^*(t) \| = \begin{Vmatrix} D-3 & 0 & 2(D-3) & 0 \\ 1 & 1 & -(D-4) & z(t) \end{Vmatrix} \tag{5.11}$$

связанную с первоначальной соотношениями

$$\begin{aligned} \| f^*(D),\ w^*(t) \| &= u(D)\, \| f(D) \quad w(t) \| \\ \| f(D),\ w(t) \| &= v(D)\, \| f^*(D) \quad w^*(t) \| \end{aligned} \tag{5.12}$$

где $u(D)$, $v(D)$ — полиномные матрицы:

$$u(D) = \begin{Vmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \end{Vmatrix}, \qquad v(D) = \begin{Vmatrix} 0 & -1 \\ 1 & 1 \\ D & D-1 \end{Vmatrix} \tag{5.13}$$

Уравнение (5.8) и соответствует канонической форме.

На этом примере можно убедиться, что формулировать теорему о совместности при помощи неизмененных миноров нельзя. В самом деле, все неизмененные миноры (5.3) обращаются в нуль при $z(t) = ce^{3t}$; очевидно, однако, что скалярная система при этом несовместна, так как уравнение (5.9) не удовлетворяется.

Если же «... обе буквы, D и t, рассматриваются как независимые переменные, никак не связанные между собой» [1] (стр. 16), то условиям совместности нельзя удовлетворить ни при каком $z(t)$, не равным тождественно нулю, что также неверно.

В качестве последней иллюстрации тех подводных камней, на которые приходится наталкиваться, укажем, что вообще непозволительно пользоваться «практическим правилом» [1] (стр. 20), состоящим в том, что в самой заданной системе выбираются r уравнений и разрешаются относительно r неизвестных, хотя бы операционная матрица получившейся усеченной системы была неособой.

Возьмем, например, первое и третье из заданных уравнений (5.1) и положим в них $y_3 = \varphi(t)$, $z(t) = ce^{-t}$. Имеем

$$\begin{gathered} -y_1 - y_2 = -\dot{\varphi}(t) + 4\varphi(t) - ce^{-t} \\ \ddot{y}_1 - 2\dot{y}_1 - y_1 + \dot{y}_2 - y_2 = -\ddot{\varphi}(t) + \dot{\varphi}(t) + 4\varphi(t) - 2ce^{-t} \end{gathered}$$

Отсюда

$$\begin{gathered} y_1 = -2\varphi(t) + Ce^{3t} + C' \\ y_2 = \dot{\varphi}(t) - 2\varphi(t) + ce^{-t} - Ce^{3t} - C' \end{gathered}$$

но при $C' \neq 0$ это решение неверно, в чем легко убедиться, подставляя во второе из заданных уравнений.

Поступила 22 X 1951

Институт механики
Академии Наук СССР

ЛИТЕРАТУРА

1. Лузин Н. Н. К изучению матричной теории дифференциальных уравнений. Автоматика и телемеханика. 1940. № 5. Стр. 4—66.
2. Булгаков Б. В. Колебания. М.—Л. 1949. Т. I. Гл. 1.