ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ДИНАМИКИ СИСТЕМ И ТЕОРИИ УПРАВЛЕНИЯ
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК

Серия: «Неклассические задачи динамики и управления»

Выпуск 2

А.А. Щеглова

ВВЕДЕНИЕ В ТЕОРИЮ АЛГЕБРО-ДИФФЕРЕНЦИАЛЬНЫХ СИСТЕМ

Учебное пособие

Иркутск — 2013

УДК 517.922, 517.977.1, 517.926.4
Рекомендовано к изданию Ученым советом ИДСТУ СО РАН

Серия «Неклассические задачи динамики и управления» основана в 2013 году

Научный редактор серии: д-р физ.-мат. наук, чл.-к. РАН А.А. Толстоногов

Рецензент: д-р физ.-мат. наук, чл.-к. РАН А.А. Толстоногов

Щеглова А.А. Введение в теорию алгебро-дифференциальных симстем. — Иркутск: ИДСТУ СО РАН, 2013. — 60 с. — (Серия «Неклассические задачи динамики и управления»; вып. 2).

Учебное пособие подготовлено в рамках читаемого студентам Института математики, экономики и информатики Иркутского госуниверситета спецкурса «Специальные вопросы теории дифференциальных уравнений» и посвящено структурной теории и сязанной с ней проблеме существования решений систем линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами с необратимой матрицей при производной. В пособии содержатся выходящие за рамки университетского курса линейной алгебры необходимые сведения, касающиеся свойств линейных операторов и структуры матричных пучков. Пособие ориентировано на студентов старших курсов, аспирантов и научных работников, интересующихся теорией систем дифференциальных уравнений, не разрешенных относительно старшей производной.

Библиогр. 13 назв.

Введение

Спецкурс будет посвящен новому интересному направлению в теории обыкновенных дифференциальных уравнений (ОДУ) – системам вида

$$F(t, x(t), x'(t)) = 0, \ \ t \in T = [\alpha, \beta) \tag{1}$$

$\Big((t, x, x') \in \mathcal{D} = \{(t, x, y) \in \mathbf{R}^{2n+1} : \ t \in T, \ \|x - x_0\| < K_0, \ \|y - y_0\| < K_1\}, \ F : \mathcal{D} \to \mathbf{R}^n \Big)$, не разрешенным относительно производной искомой вектор-функции $x : \ T \to \mathbf{R}^n$ и тождественно вырожденным в области определения:

$$\det \frac{\partial F(t, x, x')}{\partial x'} \equiv 0 \ \ \forall (t, x, x') \in \mathcal{D}\,. \tag{2}$$

Системы вида (1) с условием (2) принято называть *алгебро-дифференциальными* (АДС), поскольку в линейном случае

$$A(t)x'(t) + B(t)x(t) = f(t), \ \ t \in T, \tag{3}$$

где $A(t)$ и $B(t)$ — $(n \times n)$-матрицы,

$$\det A(t) \equiv 0 \ \ \forall t \in T, \tag{4}$$

такая система представляет собой систему взаимосвязаных дифференциальных и алгебраических уравнений.

Исследование систем вида (3), (4) началось независимо друг от друга группами математиков в нашей стране (Ю.Е. Бояринцев) и США (C.W. Gear, S.L. Campbell, L.R. Petzold), несколько позднее — в Германии (R. Maerz, E. Griepentrog, R. Lamour), Швейцарии (E. Hairer, Ch. Lubich) и других странах. Хотя теория и в особенности численные методы решения АДС интенсивно развиваются в последние три десятилетия, отдельные результаты были получены значительно ранее. Большая часть из них относится к системам с постоянными коэфициентами (матрицы A и B не зависят от t). Большое влияние на последующее развитие теории АДС оказала работа Ф.Р. Гантмахера о приложении теории матричных пучков к системам с постоянными коэффициентами.

1 Некоторые сведения из теории систем обыкновенных дифференциальных уравнений

Пусть на интервале $T = [\alpha, \beta)$ определены и непрерывны функции $a_{ij}(t)$, $f_i(t)$ $(i, j = 1, 2, \ldots, n)$ со значениями в $\mathbf{R} = (-\infty, +\infty)$. Система линейных дифференциальных уравнений с коэффициентами $a_{ij}(t)$ и правыми частями $f_i(t)$ задается равенствами

$$x_i'(t) + \sum_{j=1}^{n} a_{ij}(t) x_i(t) = f_i(t), \quad i = 1, 2, \ldots, n. \tag{5}$$

Определим матрицу

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \ldots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \ldots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \ldots & a_{nn}(t) \end{pmatrix},$$

и вектор-функцию

$$f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix},$$

тогда систему (5) можно представить в матричной форме

$$x'(t) + A(t)x(t) = f(t), \ \ t \in T, \tag{6}$$

где

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Определение 1. *Решением* системы (6) называется n-мерная вектор-функция $x(t) \in \mathbf{C}^1(T)$, поточечно удовлетворяющая на T уравнению (6).

Теорема 1 (о существовании и единственности решения). *Пусть в (6) $A(t), f(t) \in C(T)$. Тогда при любых $x_0 \in \mathbf{R}^n$, $t_0 \in T$ система (6) имеет единственное решение, подчиняющееся условию*

$$x(t_0) = x_0. \tag{7}$$

Определение 2. Задача поиска решения системы (6), удовлетворяющего условию (7), называется *задачей Коши.*

Теорема 2 (о непрерывной зависимости решений от начальных данных). *Пусть в (6)* $A(t), f(t) \in C(T)$. *Тогда для любого* $\varepsilon > 0$ *найдется* $\delta > 0$, *такое, что как только*

$$\|x_0 - \tilde{x}_0\|_{\mathbf{R}^n} < \delta,$$

так

$$\|x(t) - \tilde{x}(t)\|_{\mathbf{C}^1(T)} < \varepsilon,$$

где $x(t)$ *и* $\tilde{x}(t)$ *— решения системы (6), удовлетворяющие соответственно условиям*

$$x(t_0) = x_0, \ \ \tilde{x}(t_0) = \tilde{x}_0$$

для любых произвольных $x_0, \tilde{x}_0 \in \mathbf{R}^n, \ t_0 \in T$.

Теорема 3 (о непрерывной зависимости решений от входных данных). *Пусть* $A(t), \tilde{A}(t), f(t), \tilde{f}(t) \in C(T)$. *Тогда для любого* $\varepsilon > 0$ *найдется* $\delta > 0$, *такое, что как только*

$$\|A(t) - \tilde{A}(t)\|_{\mathbf{C}(T)} < \delta, \ \ \|f(t) - \tilde{f}(t)\|_{\mathbf{C}(T)} < \delta,$$

так

$$\|x(t) - \tilde{x}(t)\|_{\mathbf{C}^1(T)} < \varepsilon,$$

где $x(t)$ *— решение задачи Коши (6), (7), а* $\tilde{x}(t)$ *— решение системы*

$$\tilde{x}'(t) + \tilde{A}(t)\tilde{x}(t) = \tilde{f}(t), \ \ t \in T,$$

подчиняющиеся условию (7).

2 Специфические особенности АДС

По своим свойствам АДС существенно отличаются от систем ОДУ, разрешенных относительно производной (в нормальной форме). Эта специфика обусловливает не только необходимость поиска принципиально новых теоретических подходов, но и переосмысления многих базовых понятий классической теории ОДУ, таких как устойчивость, управляемость, наблюдаемость и т.п.

Особенности АДС проиллюстрируем простейшими примерами.

1. *Нет непрерывной зависимости решения от входных данных.*

Пример 1. Система

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x'(t) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}, \quad t \in [0,1), \tag{3}$$

имеет единственное решение $x(t) = \begin{pmatrix} f_1(t) - f_2'(t) \\ f_2(t) \end{pmatrix}$. В частности, при $f_1(t) = 0$, $f_2(t) = 1$

$$x(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Возмущенная система

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x_\varepsilon'(t) + \begin{pmatrix} 1 & 0 \\ \varepsilon & 1 \end{pmatrix} x_\varepsilon(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \varepsilon > 0,$$

имеет однопараметрическое семейство решений

$$x_\varepsilon(t,c) = e^{t/\varepsilon} \begin{pmatrix} 1 \\ -\varepsilon \end{pmatrix} c + \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

где c – произвольное число из $\mathbf{R}$. Очевидно, что если $c \neq 0$, то $\|x_\varepsilon(t) - x(t)\|_{\mathbf{C}(T)} \to \infty$ при $\varepsilon \to 0$.

2. *АДС может иметь бесконечное число различных решений.*

Пример 2. Вектор-функция $x(t) = ((1+t)t^i;\ t^i)^T$ при любом $i = 0, 1, 2, \ldots$ является решением системы

$$\begin{pmatrix} 1+t & -(1+t)^2 \\ 1 & -(1+t) \end{pmatrix} x'(t) - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x(t) = 0, \quad t \in T. \tag{4}$$

3. *Неоднородная система может оказаться несовместной на отрезке T.* Если задать правую часть для системы (4) в виде $(f_1(t),\ f_2(t))^T$, то полученная неоднородная АДС будет иметь решение только в случае, когда компоненты вектора правой части связаны равенством

$$\frac{d}{dt}[(1+t)f_2(t) - f_1(t)] = f_2(t), \quad t \in T.$$

4. *Поставленная задача Коши*

$$x(t_0) = x_0 \tag{5}$$

может быть неразрешима для заданного вектора $x_0 \in \mathbf{R}^n$. В частности, задача (3),(5) будет иметь решение лишь при одном значении

$$x_0 = \begin{pmatrix} f_1(t_0) - f_2'(t_0) \\ f_2(t_0) \end{pmatrix}. \tag{7}$$

Условия, которым должны удовлетворять входные данные вырожденной задачи для того, чтобы последняя была разрешима, называются *условиями согласования*. Равенство (7) в данном случае представляет собой условие согласования для задачи (5),(3).

5. *Для существания решения в пространстве* $\mathbf{C}^1(T)$ *может оказаться необходима дифференцируемость входных данных системы вплоть до порядка* n *включительно*. В частности, компоненты правой части системы (3) должны удовлетворять включениям $f_1(t) \in \mathbf{C}^1(T)$, $f_2(t) \in \mathbf{C}^2(T)$.

Часть I
Структура общего решения линейной алгебро-дифференциальной системы с постоянными коэффициентами

3 Матрицант

Рассмотрим систему дифференциальных уравнений

$$X'(t) = A(t)X(t), \ \ t \in T, \tag{8}$$

где $A(t) \in C(T)$, $X(t)$ — искомая $(n \times n)$-матрица с элементами из пространства $\mathbf{C}^1(T)$.

Воспользуемся методом последовательных приближений для определения нормированного решения системы (8), т.е. удовлетворяющего условию

$$X(t_0) = E, \ \ t_0 \in T.$$

В качестве первого приближения возьмем единичную матрицу

$$X_0(t) \equiv E, \ \ t \in T.$$

Последовательные приближения $X_k(t)$ $(k = 1, 2, \ldots)$ будем находить из рекуррентных соотношений

$$X_k'(t) = A(t)X_{k-1}(t).$$

Полагая $X_k(t_0) = E$, мы можем представить $X_k(t)$ в виде

$$X_k(t) = E + \int\limits_{t_0}^{t} A(\tau)X_{k-1}(\tau)d\tau, \ \ k = 1, 2, \ldots.$$

Таким образом,

$$X_0 = E, \quad X_1 = E + \int\limits_{t_0}^{t} A(\tau)d\tau,$$

$$X_2 = E + \int\limits_{t_0}^{t} A(\tau)d\tau + \int\limits_{t_0}^{t} A(\tau)d\tau \int\limits_{t_0}^{\tau} A(\tau_1)d\tau_1, \ \ldots,$$

т.е. $X_k(t)$ есть сумма первых $k+1$ членов матричного ряда

$$E + \int\limits_{t_0}^{t} A(\tau)d\tau + \int\limits_{t_0}^{t} A(\tau)d\tau \int\limits_{t_0}^{\tau} A(\tau_1)d\tau_1 + \ldots. \tag{9}$$

Для того чтобы доказать, что этот ряд абсолютно и равномерно сходится в любой замкнутой части интервала T и определяет искомое решение уравнения (8), мы построим мажорантный ряд.

Определим неотрицательные функции $g(t)$ и $h(t)$ в интервале T равенствами

$$g(t) = \max\left\{|a_{11}(t)|, |a_{12}(t)|, \ldots, |a_{nn}(t)|\right\}, \quad h(t) = \left|\int\limits_{t_0}^{t} g(\tau)d\tau\right|.$$

По определению значение функции $g(t)$ при каком-то из значений t равно наибольшему из n^2 модулей значений $a_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ при том же значении t.

Легко проверяется, что функции $g(t)$, а следовательно, и $h(t)$ непрерывны в интервале T. Непрерывность функции $g(t)$ в любой точке $t_1 \in T$ следует из того, что разность $g(t) - g(t_1)$ при t, достаточно близком к t_1, всегда совпадает с одной из n^2 разностей $|a_{ik}(t)| - |a_{ik}(t_1)|$ $(i, k = 1, 2, \ldots, n)$.

Каждый из n^2 скалярных рядов, на которые распадается матричный ряд (9), мажорируется рядом

$$1 + h(t) + \frac{nh^2(t)}{2!} + \frac{n^2h^3(t)}{3!} + \ldots \tag{10}$$

Действительно,

$$\left|\left\{\int\limits_{t_0}^{t} A(\tau)d\tau\right\}_{i,k}\right| = \left|\int\limits_{t_0}^{t} a_{ik}(\tau)d\tau\right| \le \left|\int\limits_{t_0}^{t} g(\tau)d\tau\right| = h(t),$$

$$\left|\left\{\int\limits_{t_0}^{t} A(\tau)d\tau \int\limits_{t_0}^{\tau} A(\tau_1)d\tau_1\right\}_{i,k}\right| = \left|\sum_{j=1}^{n}\int\limits_{t_0}^{t} a_{ij}(\tau)\int\limits_{t_0}^{\tau} a_{jk}(\tau_1)d\tau_1\right| \le$$

$$\le n\left|\int\limits_{t_0}^{t} g(\tau)d\tau \int\limits_{t_0}^{\tau} g(\tau_1)d\tau_1\right| = \frac{nh^2(t)}{2}$$

и т.д.

Ряд (10) сходится в интервале T, причем сходится равномерно в любой замкнутой части этого интервала. Отсюда вытекает, что и матричный ряд (9) сходится в T и притом абсолютно и равномерно в любом замкнутом интервале, входящем в T.

Почленным дифференцированием проверяем, что сумма ряда (9) представляет собой решение уравнения (8); это решение обращается в E при $t = t_0$. Почленное дифференцирование ряда (9) допустимо, поскольку ряд, получающийся после дифференцирования, отличается множителем $A(t)$ от ряда (9) и, следовательно, как и ряд (9), является равномерно сходящимся в любой замкнутой части интервала T.

Таким образом, нами доказана теорема о существовании нормированного решения задачи (8). Это решение будем обозначать $\Omega(t, t_0)$.

Определение 3. Нормированное решение $\Omega(t, t_0)$ уравнения (8) часто называют *матрицантом.*

Мы показали, что матрицант представим в виде ряда

$$\Omega(t, t_0) = E + \int\limits_{t_0}^{t} A(\tau)d\tau + \int\limits_{t_0}^{t} A(\tau)d\tau \int\limits_{t_0}^{\tau} A(\tau_1)d\tau_1 + \ldots, \qquad (11)$$

который сходится абсолютно и равномерно в любом замкнутом интервале, в котором матричнозначная функция $A(t)$ непрерывна.

Свойства матрицанта

1. $\det \Omega(t, t_0) \neq 0 \ \ \forall t \in T$.

Продифференцируем определитель матрицанта, дифференцируя последовательно строки определителя и используя при этом дифференциальные соотношения

$$x'_{i,j}(t) = \sum_{k=1}^{n} a_{ik}(t)x_{kj}(t), \ \ i, j = 1, 2, \ldots, n.$$

Тогда получим

$$(\det\Omega(t,t_0))' = (a_{11}(t) + a_{22}(t) + \ldots + a_{nn}(t))\det\Omega(t,t_0).$$

Отсюда следует известное *тождество Якоби*

$$\det\Omega(t,t_0) = ce^{\int\limits_{t_0}^{t}\operatorname{Sp}A(t)dt},$$

где c — постоянная, а

$$\operatorname{Sp}A(t) = a_{11}(t) + a_{22}(t) + \ldots + a_{nn}(t)$$

— след матрицы $A(t)$.

Так как $\det\Omega(t,t_0)$ не может тождественно равняться нулю, то $c \neq 0$. Но тогда из тождества Якоби следует, что определитель $\Omega(t,t_0)$ при любом значении аргумента отличен от нуля.

2. Произвольное решение уравнения (8) имеет вид

$$X(t) = \Omega(t,t_0)C,$$

где C — произвольная постоянная матрица.

Легко убедиться, что матрица $\Omega(t,t_0)C$ является решением уравнения (8). С другой стороны, если $X(t)$ произвольное решение уравнения (8), то

$$X'(t) = \Big(\Omega(t,t_0)\Omega^{-1}(t,t_0)X(t)\Big)' =$$

$$= \Omega(t,t_0)'\Omega^{-1}(t,t_0)X(t) + \Omega(t,t_0)\left(\Omega^{-1}(t,t_0)X(t)\right)' =$$

$$= A(t)X(t) + \Omega(t,t_0)\left(\Omega^{-1}(t,t_0)X(t)\right)',$$

откуда в силу (8)

$$\left(\Omega^{-1}(t,t_0)X(t)\right)' = 0$$

и

$$\Omega^{-1}(t,t_0)X(t) = \text{const} = C.$$

3. $\Omega(t,t_0) = \Omega(t,t_1)\Omega(t_1,t_0)$, $t,t_0,t_1 \in T$.

Дуйствительно, поскольку $\Omega(t,t_0)$ и $\Omega(t,t_1)$ — два решения уравнения (8), то

$$\Omega(t,t_0) = \Omega(t,t_1)C,$$

где C — постоянная матрица. Полагая $t = t_1$, получим $C = \Omega(t_1, t_0)$.

4. Если $A = \mathrm{const}$, то

$$\Omega(t, t_0) = e^{A(t-t_0)},$$

где $e^{A(t-t_0)}$ называется матричной экспонентой, и

$$e^{A(t-t_0)} = \sum_{i=0}^{\infty} \frac{\left(A(t-t_0)\right)^i}{i!}$$

для произвольной матрицы A.

Покажем теперь, как при помощи матрицанта выражается общее решение системы линейных дифференциальных уравнений с правыми частями

$$x'(t) = A(t)x(t) + f(t), \quad t \in T.$$

Будем искать решение этого уравнения в виде

$$x(t) = \Omega(t, t_0)z(t), \tag{12}$$

где $z(t)$ — неизвестный столбец. Подставляя (12) в систему, получим

$$A(t)\Omega(t, t_0)z(t) + \Omega(t, t_0)z'(t) = A(t)\Omega(t, t_0)z(t) + f(t),$$

откуда

$$z'(t) = \left(\Omega(t, t_0)\right)^{-1} f(t).$$

Интегрируя, находим

$$z(t) = \int\limits_{t_0}^{t} \left(\Omega(\tau, t_0)\right)^{-1} f(\tau)d\tau + c,$$

где c — произвольный постоянный вектор. Подставим это выражение в (12), получим

$$x(t) = \Omega(t, t_0) \int\limits_{t_0}^{t} \left(\Omega(\tau, t_0)\right)^{-1} f(\tau)d\tau + \Omega(t, t_0)c.$$

Давая t значение t_0, найдем $x(t_0) = c$. Поэтому формула для $x(t)$ принимает вид

$$x(t) = \Omega(t, t_0) x(t_0) + \int_{t_0}^{t} K(t, \tau) f(\tau) d\tau,$$

где

$$K(t, \tau) = \Omega(t, t_0) \left(\Omega(\tau, t_0) \right)^{-1}$$

— так называемая матрица Коши.

4 Понятие о матричных пучках

Рассмотрим линейную АДС с постоянными матрицами коэффициентов

$$Ax'(t) + Bx(t) = f(t), \quad t \in T = [\alpha, \beta]. \tag{a1}$$

На свойства решений этой системы влияет не только матрица при производных, их определяет структура всего матричного пучка $\lambda A + B$ (λ — скалярный параметр). Это очевидно из следующих примеров, в которых матрица при производных одна и та же

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x'(t) + \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x(t) = \begin{pmatrix} 0 \\ 0 \\ f(t) \end{pmatrix}, \tag{a2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x'(t) + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x(t) = \begin{pmatrix} 0 \\ 0 \\ f(t) \end{pmatrix}. \tag{a3}$$

Решение системы (a2)

$$x_*(t) = \begin{pmatrix} -f'(t) \\ f(t) \\ f''(t) \end{pmatrix},$$

и решение системы (a3)

$$x_{**}(t) = \begin{pmatrix} c_1 e^{-t} \\ c_2 e^{-t} \\ f(t) \end{pmatrix}$$

по своим свойствам значительно отличаются друг от друга: $x_*(t)$ единственно, но зависит от производных правой части системы $f(t)$, а $x_{**}(t)$ не зависит от производных, зато включает в себя две константы, для нахождения которых требуются два дополнительных условия на компоненты $x_1(t)$ и $x_2(t)$ (начальные условия).

В связи с тем, что структура общего решения АДС однозначно определяется внутренней структурой пучка матриц $\lambda A + B$ в спецкурсе значительное место уделено выходящим за рамки курса линейной алгебры сведениям, касающимся свойств матричных пучков и многочленных матриц вообще. В частности, на основе теории элементарных делителей строится каноническая форма матричного пучка в самом общем случае.

4.1 Понятие λ-матрицы. Эквивалентность λ-матриц

Определение 4. *Многочленной матрицей* или λ-*матрицей* называется $(m \times n)$-матрица $A(\lambda)$, элементы которой суть многочлены от λ

$$A(\lambda) = \|a_{ik}(\lambda)\|_{i=\overline{1,m},\ k=\overline{1,n}} = \left\|a_{ik}^{(0)}\lambda^l + a_{ik}^{(1)}\lambda^{l-1} + \ldots + a_{ik}^{(l)}\right\|_{i=\overline{1,m},\ k=\overline{1,n}},$$

l — наибольшая из степеней многочленов $a_{ik}(\lambda)$.

Обозначим $A_j = \left\|a_{ik}^{(j)}\right\|_{i=\overline{1,m},\ k=\overline{1,n}}$, $j = \overline{0,l}$, тогда

$$A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \ldots + A_{l-1}\lambda + A_l.$$

Введем в рассмотрение *элементарные операции* над многочленной матрицей $A(\lambda)$:

1. Умножение i-ой строки на число $c \neq 0$.
2. Прибавление к i-ой строке j-ой строки, умноженной на произвольный многочлен $b(\lambda)$.
3. Перестановка местами i-ой и j-ой строки.

Убедиться *самостоятельно*, что операции 1.–3. равносильны умножению многочленной матрицы $A(\lambda)$ слева соответственно на квадратные матрицы порядка m

$$
S' = \begin{array}{c} \begin{array}{ccccc} & & (i) & & \end{array} \\ \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & c & \ldots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 1 \end{pmatrix} \end{array}, \quad S'' = \begin{array}{c} \begin{array}{ccccccc} & & (i) & & (j) & & \end{array} \\ \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & b(\lambda) & \ldots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{pmatrix} \end{array},
$$

$$
S''' = \begin{array}{c} \begin{array}{ccccccc} & & (i) & & (j) & & \end{array} \\ \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{pmatrix} \end{array}.
$$

Операции 1.–3. называются *левыми элементарными операциями.*

Совершенно аналогично определяются *правые элементарные операции* (эти операции производятся над столбцами многочленной матрицы) и соответствующие им квадратные матрицы порядка n

$$
T' = \begin{array}{cc} \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & c & \ldots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 1 \end{pmatrix} & \begin{array}{c} \\ \\ (i) \\ \\ \\ \end{array} \end{array}, \quad T'' = \begin{array}{cc} \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & b(\lambda) & \ldots & 0 \\ \vdots & & & & \\ 0 & \ldots & \ldots & \ldots & 1 \end{pmatrix} & \begin{array}{c} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{array} \end{array},
$$

$$
T''' = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{array}{c} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{array} .
$$

Матрицы типа S', S'', S''' (или T', T'', T''') называются *элементарными матрицами*. Определитель элементарной матрицы не зависит от λ и отличен от нуля (проверить *самостоятельно*). Поэтому для каждой левой (правой) элементарной операции существует обратная операция, которая также является левой (соответственно правой) элементарной операцией.

Определение. Две многочленные матрицы $A(\lambda)$ и $B(\lambda)$ называются: 1) *левоэквивалентными*, 2) *правоэквивалентными*, 3) *эквивалентными*, если одна из них получается из другой путем применения соответственно 1) левых, 2) правых, 3) левых и правых элементарных операций.

Пусть матрица $B(\lambda)$ получается из $A(\lambda)$ при помощи левых элементарных операций, соответствующих матрицам $S_1, S_2, \dots, S_p$

$$
B(\lambda) = S_p, S_{p-1}, \dots, S_2, S_1 A(\lambda) = P(\lambda) A(\lambda), \tag{13}
$$

где $P(\lambda) = S_p, S_{p-1}, \dots, S_2, S_1$, как и каждая из матриц S_i, $i = \overline{1,p}$, имеет отличный от нуля и не зависящий от λ определитель.

Теорема [5, с. 131]. *Любая квадратная λ-матрица с постоянным отличным от нуля определителем может быть представлена в виде произведения элементарных матриц.*

Доказательство теоремы опущено.

В силу сформулированной теоремы равенство (13) означает левую эквивалентность матриц $A(\lambda)$ и $B(\lambda)$.

В случае правой эквивалентности многочленных матриц будем иметь

$$
B(\lambda) = A(\lambda) Q(\lambda),
$$

и в случае двусторонней эквивалентности —

$$
B(\lambda) = P(\lambda) A(\lambda) Q(\lambda),
$$

где $P(\lambda)$ и $Q(\lambda)$ — матрицы с отличным от нуля и не зависящим от λ определителями.

Таким образом, сформулированное выше определение можно заменить равносильным.

Определение. Две прямоугольные λ-матрицы $A(\lambda)$ и $B(\lambda)$ называются 1) *левоэквивалентными*, 2) *правоэквивалентными*, 3) *эквивалентными*, если соответственно: 1) $B(\lambda) = P(\lambda)A(\lambda)$, 2) $B(\lambda) = A(\lambda)Q(\lambda)$, 3) $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$, где $\det P(\lambda) = \mathrm{const} \neq 0$, $\det Q(\lambda) = \mathrm{const} \neq 0$.

4.2 Каноническая форма многочленной матрицы

В многочленной $(m \times n)$-матрице $A(\lambda) = \|a_{ij}(\lambda)\|_{i=\overline{1,m},j=\overline{1,n}}$ выберем элемент $a_{ij}(\lambda) \neq 0$ наименьшей степени по λ и перестановкой строк и столбцов сделаем его элементом $a_{11}(\lambda)$.

После этого найдем частные и остатки от деления многочленов $a_{i1}(\lambda)$ и $a_{1j}(\lambda)$ на $a_{11}(\lambda)$

$$a_{i1}(\lambda) = a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda),$$

$$a_{1j}(\lambda) = a_{11}(\lambda)q_{1j}(\lambda) + r_{1j}(\lambda) \quad (i = \overline{2,m},\ j = \overline{2,n}).$$

Если найдется хотя бы один из остатков $r_{i1}(\lambda)$ или $r_{1j}(\lambda)$, например, $r_{1j}(\lambda)$, на равный тождественно нулю, то вычтем из j-ого столбца первый, умноженный на $q_{1j}(\lambda)$. Тем самым мы заменим элемент $a_{1j}(\lambda)$ остатком $r_{1j}(\lambda)$, который имеет меньшую степень по λ, нежели $a_{11}(\lambda)$. Теперь мы имеем возможность снова уменьшить степень элемента, стоящего в левом верхнем углу матрицы, поместив на это место $r_{1j}(\lambda)$.

Если же все $r_{i1}(\lambda)$, $r_{1j}(\lambda) \equiv 0$, то вычитая из i-ой строки первую, умноженную на $q_{i1}(\lambda)$ $(i = \overline{2,m})$, а из j-ого столбца — первый, предварительно умноженный на $q_{1j}(\lambda)$ $(j = \overline{2,n})$, мы приведем нашу многочленную матрицу к виду

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \dots & 0 \\ 0 & a_{22}(\lambda) & \dots & a_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2}(\lambda) & \dots & a_{mn}(\lambda) \end{pmatrix}.$$

Если при этом какой-нибудь элемент $a_{ij}(\lambda)$ $(i = \overline{2,m},\ j = \overline{2,n})$ не делится нацело на $a_{11}(\lambda)$, то прибавим к 1-му столбцу тот столбец, где содержится этот элемент. Тем самым мы придем к предыдущему случаю и, следовательно, снова можем заменить элемент $a_{11}(\lambda)$ многочленом меньшей степени.

Таким образом, после конечного числа (*почему?*) элементарных операций мы получим матрицу вида

$$\begin{pmatrix} a_1(\lambda) & 0 & \dots & 0 \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{m2}(\lambda) & \dots & b_{mn}(\lambda) \end{pmatrix},$$

причем все $b_{ij}(\lambda)$ делятся без остатка на $a_1(\lambda)$.

Продолжая процесс, т.е. производя аналогичные операции над матрицей $B(\lambda) = \|b_{ij}(\lambda)\|_{i=\overline{2,m},\ j=\overline{2,n}}$ и т.д., мы в конце концов придем к матрице вида

$$\begin{pmatrix} a_1(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_r(\lambda) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}, \qquad (14)$$

где многочлены $a_1(\lambda),\ a_2(\lambda), \dots, a_r(\lambda)$ $(r \le m, n)$ не равны тождественно нулю, и каждый делится без остатка на предыдущий.

Умножая первые r строк на соответствующие отличные от нуля числовые множители, мы можем добиться того, чтобы старшие коэффициенты многочленов $a_1(\lambda),\ a_2(\lambda), \dots, a_r(\lambda)$ были равны единице.

Определение. *Многочленная* $(m \times n)$*-матрица называется* *канонической диагональной*, если она имеет вид (14), где
1) многочлены $a_i(\lambda)$, $\ i = \overline{1,r}$, не равны тождественно нулю,
2) каждый последующий многочлен делится на предыдущий без остатка.
При этом предполагается, что старшие коэффициенты многочленов равны единице.

Таким образом, мы доказали теорему.

Теорема. *Произвольная прямоугольная многочленная матрица* $A(\lambda)$ *эквивалентна некоторой канонической диагональной.*

Самостоятельно, пользуясь представленным выше алгоритмом, привести к канонической диагональной форме матрицу

$$\begin{pmatrix} 2\lambda^2+\lambda+3 & 0 & \lambda+1 \\ 3\lambda-1 & \lambda^2-2\lambda & 0 \\ 0 & \lambda+2 & 4 \end{pmatrix}.$$

4.3 Инвариантные многочлены и элементарные делители многочленной матрицы

Пусть $\operatorname{rank} A(\lambda) = r$, т.е. в этой матрице имеются не равные тождественно нулю миноры r-го порядка, в то время как все миноры порядка больше, чем r, тождественно относительно λ равны нулю. Обозначим через $D_j(\lambda)$ наибольший общий делитель всех миноров j-го порядка матрицы $A(\lambda)$. В каждом $D_j(\lambda)$ берем старший коэффициент равный единице. Тогда в ряду

$$D_r(\lambda),\ D_{r-1}(\lambda),\ \ldots,\ D_1(\lambda),\ D_0(\lambda) \equiv 1$$

каждый многочлен делится без остатка на последующий (*почему?*). Соответствующие частные обозначим

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \ldots, \quad i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} = D_1(\lambda).$$

Определение. Многочлены $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ называются *инвариантными многочленами* матрицы $A(\lambda)$.

Пусть матрицы $A(\lambda)$ и $B(\lambda)$ эквивалентны, т.е. приводятся одна к другой элементарными операциями, которые не меняют ни ранга, ни многочленов $D_j(\lambda)$ (*почему?*), тогда при переходе от матрицы $A(\lambda)$ к матрице $B(\lambda)$ многочлены $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ остаются неизменными — "инвариантными".

Если многочленная матрица имеет канонический диагональный вид (14), то

$$D_1(\lambda) = a_1(\lambda), \quad D_2(\lambda) = a_1(\lambda)a_2(\lambda), \quad \ldots, \quad D_r(\lambda) = a_1(\lambda)a_2(\lambda)\ldots a_r(\lambda).$$

Тогда

$$i_1(\lambda) = a_r(\lambda), \quad i_2(\lambda) = a_{r-1}(\lambda), \quad \ldots, \quad i_r(\lambda) = a_1(\lambda). \tag{15}$$

Пусть $A(\lambda)$ эквивалентна некоторой многочленной матрице вида (14), тогда (15) будут инвариантными многочленами матрицы $A(\lambda)$.

Полученные результаты могут быть сформулированы в виде теоремы

Теорема. *Многочленная $(m \times n)$-матрица $A(\lambda)$ всегда эквивалентна канонической диагональной матрице*

$$\begin{pmatrix} i_r(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & i_{r-1}(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & i_1(\lambda) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

где $r = \operatorname{rank} A(\lambda)$, $i_j(\lambda)$ $(j = \overline{1, r})$ — инвариантные многочлены матрицы $A(\lambda)$.

Следствие. *Для того чтобы две прямоугольные матрицы одинаковых размеров были эквивалентными, необходимо и достаточно, чтобы они имели одни и те же инвариантные многочлены.*

Доказательство. Необходимость доказана выше. Достаточность следует из того, что матрицы , имеющие одни и те же инвариантные многочлены , эквивалентны одной и той же канонической диагональной матрице и, следовательно, эквивалентны между собой. $\square$

Следствие. *В ряду инвариантных многочленов каждый многочлен, начиная со второго, является делителем предыдущего.*

Это утверждение следует из того, что многочлены $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ совпадают с соответственно с многочленами $a_r(\lambda), a_{r-1}(\lambda), \dots, a_1(\lambda)$ канонической диагональной матрицы (14).

Теорема. *Если в квазидиагональной прямоугольной матрице*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & O \\ O & B(\lambda) \end{pmatrix}$$

любой инвариантный многочлен матрицы $A(\lambda)$ является делителем любого инвариантного многочлена матрицы $B(\lambda)$, то совокупность инвариантных многочленов матрицы $C(\lambda)$ получается объединением инвариантных многочленов матриц $A(\lambda)$ и $B(\lambda)$.

Доказательство опущено.

Разложим инвариантные многочлены $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ на неприводимые в данном числовом поле K множители

$$\begin{aligned} i_1(\lambda) &= (\phi_1(\lambda))^{c_1} (\phi_2(\lambda))^{c_2} \ldots (\phi_s(\lambda))^{c_s}, \\ i_2(\lambda) &= (\phi_1(\lambda))^{d_1} (\phi_2(\lambda))^{d_2} \ldots (\phi_s(\lambda))^{d_s}, \\ &\ldots \ \ldots \ \ldots \ \ldots \\ i_r(\lambda) &= (\phi_1(\lambda))^{l_1} (\phi_2(\lambda))^{l_2} \ldots (\phi_s(\lambda))^{l_s}, \end{aligned} \tag{16}$$

где $c_j \geq d_j \geq \ldots \geq l_j$, $\quad j = \overline{1, s}$; $\phi_1(\lambda), \phi_2(\lambda), \ldots, \phi_s(\lambda)$ — все различные неприводимые в числовом поле K многочлены (со старшими коэффициентами равными единице), входящие в состав $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$.

Определение. Все отличные от единицы степени среди $(\phi_1(\lambda))^{c_1}, \ldots,$ $(\phi_s(\lambda))^{c_s}$ в разложении (16) называются *элементарными делителями* матрицы $A(\lambda)$.

Теорема. *Совокупность элементарных делителей квазидиагональной матрицы*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & O \\ O & B(\lambda) \end{pmatrix}$$

всегда получается объединением элементарных делителей матрицы $A(\lambda)$ *с элементарными делителями матрицы* $B(\lambda)$.

Доказательство опущено (см. [5, с. 136]).

Пусть теперь дана квадратная матрица $A = \|a_{ij}\|_{i,j=\overline{1,n}}$ с элементами из поля K. Составим для нее характеристическую матрицу

$$\lambda E - A = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{pmatrix}.$$

Характеристическая матрица является λ-матрицей ранга n. Ее инвариантные многочлены называются *инвариантными многочленами* матрицы A, а соответствующие элементарные делители в поле K — *элементарными делителями в поле* K *матрицы* A.

Пример. Найдем элементарные делители матрицы

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix}.$$

В характеристической матрице

$$\lambda E - A = \begin{pmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14 & 5 & 1 & \lambda \end{pmatrix}$$

прибавим к четвертой строке третью, предварительно умноженную на λ,

$$\begin{pmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{pmatrix}.$$

Теперь, прибавляя к первым трем столбцам четвертый, предварительно умноженный соответственно на -6, -1, $\lambda - 2$, получим

$$\begin{pmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{pmatrix}.$$

К первому столбцу прибавим второй, умноженный на $\lambda - 3$

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{pmatrix}.$$

Прибавим ко второй и четвертой строкам первую, умноженную соответственно на $\lambda + 1$ и $5 - \lambda$, тем самым найдем матрицу

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 0 & \lambda^2-2\lambda+1 & 0 \end{pmatrix}.$$

Прибавим к четвертой строке вторую, затем умножим первую и третью строки на -1. После соответствующей перестановки строк и столбцов получим

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (\lambda-1)^2 & 0 \\ 0 & 0 & 0 & (\lambda-1)^2 \end{pmatrix}.$$

Итак, матрица A имеет два элементарных делителя: $(\lambda-1)^2$ и $(\lambda-1)^2$.

Следующая теорема показывает, что инвариантные многочлены определяют исходную матрицу с точностью до преобразования подобия.

Теорема. *Для того чтобы две матрицы* $A = \|a_{ij}\|_{i,j=\overline{1,n}}$ *и* $B = \|b_{ij}\|_{i,j=\overline{1,n}}$ *были подобны* ($B = T^{-1}AT$) *необходимо и достаточно, чтобы они имели одни и те же инвариантные многочлены, или, что то же, одни и те же элементарные делители в поле* K.

Рассмотрим две квадратные λ-матрицы n-го порядка, у которых все элементы имеют степень не выше единицы относительно λ. Эти многочленные матрицы могут быть представлены в виде матричных двучленов

$$A(\lambda) = A_0\lambda + A_1, \quad B(\lambda) = B_0\lambda + B_1.$$

Будем предполагать, что $|A_0| \neq 0, \quad |B_0| \neq 0$.

Следующая теорема устанавливает критерий эквивалентности таких двучленов.

Теорема. *Если двучлены первой степени* $A_0\lambda+A_1$ *и* $B_0\lambda+B_1$ ($|A_0| \neq 0, \quad |B_0| \neq 0$) *эквивалентны, то они строго эквивалентны, т.е. в тождестве*

$$B_0\lambda + B_1 = P(\lambda)\,(A_0\lambda + A_1)\,Q(\lambda)$$

можно $P(\lambda)$ *и* $Q(\lambda)$ *— матрицы с постоянными и отличными от нуля определителями — заменить не зависящими от* λ *невырожденными матрицами* P *и* Q:

$$B_0\lambda + B_1 = P\,(A_0\lambda + A_1)\,Q.$$

Замечание. Элементарные делители матрицы A, в отличие от инвариантных многочленов, существенно связаны с данным числовым полем K. Если вместо исходного поля K возьмем другое числовое поле (которому также принадлежат элементы данной матрицы A), то элементарные делители могут измениться.

Так, например, пусть дана матрица A с вещественными элементами. Характеристический многочлен этой матрицы будет иметь вещественные коэффициенты, в то же время этот многочлен может иметь комплексные корни. Если K — поле вещественных чисел, то среди элементарных делителей могут быть и степени неприводимых квадратных трехчленов с вещественными коэффициентами. Если K — поле комплексных чисел, то каждый элементарный делитель имеет вид $(\lambda - \lambda_0)^p$.

Самостоятельно привести пример квадратной матрицы третьего порядка для иллюстрации этого замечания.

5 Жорданова форма матрицы

Пусть числовое поле K содержит не только элементарные делители матрицы A, но и все ее характеристические числа. Тогда элементарные делители матрицы A имеют вид

$$(\lambda-\lambda_1)^{p_1}\,,\ \ (\lambda-\lambda_2)^{p_2}\,,\ \ \ldots,\ \ (\lambda-\lambda_u)^{p_u}\quad (p_1+p_2+\ldots+p_u=n). \qquad (17)$$

Среди чисел λ_1, λ_2, ..., λ_u могут быть и равные между собой.

Рассмотрим один из таких элементарных делителей $(\lambda-\lambda_0)^p$ и поставим ему в соответствие матрицу порядка p

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_0 \end{pmatrix} = \lambda_0 E_p + H_p, \qquad (18)$$

где E_p — единичная матрица порядка p,

$$H_p = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Определение. Матрица N размеров $n \times n$ называется *нильпотентной индекса нильпотентности* l, если $N^l = O$, но $N^{l-1} \neq O$, при некотором целом $l > 0$.

Очевидно, что $\det N = 0$ (*почему?*).

Отметим, что H_p — нильпотентная матрица индекса p (проверить *самостоятельно*).

Матрица $\lambda_0 E_p + H_p$ имеет только один элементарный делитель $(\lambda - \lambda_0)^p$ (*почему?*). Матрицу (18) будем называть *жордановой клеткой*, соответствующей элементарному делителю $(\lambda-\lambda_0)^p$.

Жордановы клетки, соответствующие элеменатрным делителям (17), обозначим через $J_1, J_2, \ldots, J_u$. Тогда квазидиагональная матрица

$$J = \operatorname{diag}\{J_1, J_2, \ldots, J_u\} = \begin{pmatrix} J_1 & O & \ldots & O \\ O & J_2 & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & J_u \end{pmatrix}$$

имеет своими элементарными делителями степени (17) (*почему?*).

Матрицу J можно записать еще и так

$$J = \operatorname{diag}\{\lambda_1 E_{p_1} + H_{p_1}, \lambda_2 E_{p_2} + H_{p_2}, \ldots, \lambda_u E_{p_u} + H_{p_u}\}$$

или $J = J_0 + J_N$, где

$$J_0 = \begin{pmatrix} \lambda_1 E_{p_1} & O & \ldots & O \\ O & \lambda_2 E_{p_2} & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & \lambda_u E_{p_u} \end{pmatrix}, \quad J_N = \begin{pmatrix} H_{p_1} & O & \ldots & O \\ O & H_{p_2} & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & H_{p_u} \end{pmatrix},$$

$\det J_0 \neq 0$, J_N называется *жордановой нильпотентной матрицей*, ее инжекс нильпотентности p определяется как $\max\{p_1, p_2, \ldots, p_u\}$ (проверить *самостоятельно*).

Поскольку матрицы A и J имеют одни и те же элементарные делители, они подобны между собой, т.е. существует такая невырожденная матрица T $(\det T \neq 0)$, что

$$A = TJT^{-1} = T \operatorname{diag}\{\lambda_1 E_{p_1} + H_{p_1}, \lambda_2 E_{p_2} + H_{p_2}, \ldots, \lambda_u E_{p_u} + H_{p_u}\} T^{-1}.$$

Определение. Матрица J называется *жордановой формой* матрицы A.

Пример. *Самостоятельно* выписать жорданову форму для матрицы, имеющей следующие элементарные делители:

$$(\lambda - \lambda_1)^2, \ (\lambda - \lambda_2)^3, \ \lambda - \lambda_3, \ (\lambda - \lambda_4)^2.$$

Если все элементарные делители матрицы A первой степени, то жорданова форма является диагональной матрицей

$$A = T \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} T^{-1}.$$

Определение. Матрица подобная диагональной матрице называется *матрицей простой структуры.*

Таким образом, *матрица A имеет простую стуктуру в том и только том случае , когда все ее элементарные делители имеют первую степень.*

6 Понятие об эквивалентных матричных пучках

Даны четыре матрицы $A,\ B,\ A_1,\ B_1$ одинакового размера $m \times n$ с элементами из числового поля K. Требуется выяснить, при каких условиях существуют две квадратные невырожденные матрицы P и Q соответственно порядков m и n такие, что одновременно

$$PAQ = A_1, \ \ PBQ = B_1.$$

Вводя в рассмотрение *пучки матриц* (матричные двучлены вида $A+\lambda B$ и $A_1 + \lambda B_1$), эти два матричных равенства можно заменить одним

$$P(A + \lambda B)Q = A_1 + \lambda B_1. \tag{19}$$

Определение. Два пучка прямоугольных матриц $A+\lambda B$ и $A_1+\lambda B_1$ одного и того же размера $m \times n$, связанные равенством (19) (в котором P и Q не зависят от λ, имеют соответственно порядки m и n, и $\det P \neq 0$, $\det Q \neq 0$), называют *строго эквивалентными.*

Согласно общему определению эквивалентности λ-матриц, пучки $A+\lambda B$ и $A_1 + \lambda B_1$ являются эквивалентными, если имеет место равенство (19), в котором P и Q — две квадратные λ-матрицы с постоянными и отличными от нуля определителями. При строгой эквивалентности требуется дополнительно, чтобы матрицы P и Q не зависели от λ.

Критерий "нестрогой" эквивалентности пучков $A+\lambda B$ и $A_1+\lambda B_1$ следует из общего критерия эквивалентности λ-матриц и состоит в совпадении инвариантных многочленов или, что то же, элементарных делителей пучков $A + \lambda B$ и $A_1 + \lambda B_1$.

Все пучки $(m \times n)$-матриц $A + \lambda B$ подразделяются на два основных типа: регулярные и сингулярные.

Определение. Пучок матриц $A + \lambda B$ называется *регулярным*, если:
1) A и B — квадратные матрицы одного и того же порядка n;

2) определитель $\det(A + \lambda B)$ не равен тождественно нулю.
Во всех остальных случаях ($m \neq n$ или $m = n$, но $\det(A+\lambda B) \equiv 0$) пучок называется *сингулярным.*

Самостоятельно привести примеры регулярного и сингулярного матричных пучков.

7 Бесконечные элементарные делители пучка. Критерий строгой эквивалентности пучков

Рассмотрим частный случай, когда $\det B \neq 0$ и $\det B_1 \neq 0$, $m = n$. При этих условиях, как было показано выше (*где именно?*), два понятия "эквивалентность" и "строгая эквивалентность" пучков $A + \lambda B$ и $A_1 + \lambda B_1$ совпадают. Поэтому, применяя к пучкам общий критерий эквивалентности λ-матриц, приходим к теореме.

Теорема. *Два пучка квадратных матриц одного и того же порядка $A + \lambda B$ и $A_1 + \lambda B_1$, у которых $\det B \neq 0$, $\det B_1 \neq 0$, являются строго эквивалентными в том и только том случае, когда эти пучки имеют одни и те же элементарные делители в поле K.*

Для того чтобы выяснить, имеет ли место эта теорема в случае, когда $\det B = \det B_1 = 0$, рассмотрим следующий пример.

Пример.

$$A + \lambda B = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 5 \\ 3 & 2 & 6 \end{pmatrix} + \lambda \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix},$$

$$A_1 + \lambda B_1 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Самостоятельно проверить, что каждый из пучков $A+\lambda B$ и $A_1+\lambda B_1$ имеет только один элементарный делитель $\lambda + 1$.

В то же время эти пучки не являются строго эквивалентными, так как матрицы B и B_1 имеют соответственно ранги 2 и 1, а из равенства (19), если бы оно имело место, следовало бы, что ранги матриц B и B_1 равны между собой. При этом пучки являются регулярными, так как

$$\det(A + \lambda B) = \det(A_1 + \lambda B_1) \equiv \lambda + 1.$$

Т.е. теорема не верна для случая $\det B = 0$, $\det B_1 = 0$.

Для пучков такого сорта вводится понятие "бесконечных" элементарных делителей. Будем пучок $A + \lambda B$ задавать при помощи "однородных" параметров λ, μ : $\mu A + \lambda B$. Тогда определитель $\det(\mu A + \lambda B)$ будет однородной функцией параметров λ и μ. Определяя наибольший общий делитель $D_k(\lambda, \mu)$ всех миноров k-го порядка матрицы $\mu A + \lambda B$ ($k = \overline{1, n}$), получим инвариантные многочлены по известным формулам

$$i_1(\lambda, \mu) = \frac{D_n(\lambda, \mu)}{D_{n-1}(\lambda, \mu)}, \quad i_2(\lambda, \mu) = \frac{D_{n-1}(\lambda, \mu)}{D_{n-2}(\lambda, \mu)}, \quad \dots$$

при этом все $D_k(\lambda, \mu)$ и $i_j(\lambda, \mu)$ — однородные относительно λ и μ многочлены. Разлагая инвариантные многочлены на степени неприводимых в поле K однородных многочленов, получим элементарные делители $e_l(\lambda, \mu)$ $(l = 1, 2, \dots)$ пучка $\mu A + \lambda B$ в поле K.

Очевидно, что, полагая $\mu = 1$ в $e_l(\lambda, \mu)$, мы вернемся к элементарным делителям $e_l(\lambda)$ пучка $A + \lambda B$. Обратно, из каждого элементарного делителя $e_l(\lambda)$ степени q пучка $A + \lambda B$ мы получаем соответствующий элементарный делитель $e_l(\lambda, \mu)$ по формуле $e_l(\lambda, \mu) = \mu^q e_l\left(\frac{\lambda}{\mu}\right)$ (*почему?*). Таким способом могут быть получены все элементарные делители пучка $\mu A + \lambda B$, за исключением элементарных делителей вида μ^q.

Элементарные делители вида μ^q существуют в том и только том случае, когда $\det B = 0$, и носят название *"бесконечных" элементарных делителей пучка* $A + \lambda B$.

Поскольку из строгой эквивалентности пучков $A + \lambda B$ и $A_1 + \lambda B_1$ следует строгая эквивалентность пучков $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$, то у строго эквивалентных пучков $A + \lambda B$ и $A_1 + \lambda B_1$ должны совпадать не только "конечные", но и "бесконечные" элементарные делители.

Пусть в пучках $A + \lambda B$ и $A_1 + \lambda B_1$ совпадают все (в том числе и "бесконечные") элементарные делители. Введем однородные параметры $\mu A + \lambda B$, $\mu A_1 + \lambda B_1$. Преобразуем параметры: $\lambda = \alpha_1 \tilde{\lambda} + \alpha_2 \tilde{\mu}$, $\mu = \beta_1 \tilde{\lambda} + \beta_2 \tilde{\mu}$ $(\alpha_1 \beta_2 - \alpha_2 \beta_1 \neq 0)$. В новых параметрах пучки запишутся так: $\tilde{\mu} \tilde{A} + \tilde{\lambda} \tilde{B}$, $\tilde{\mu} \tilde{A}_1 + \tilde{\lambda} \tilde{B}_1$, где $\tilde{B} = \beta_1 A + \alpha_1 B$, $\tilde{B}_1 = \beta_1 A_1 + \alpha_1 B_1$. Из регулярности пучков $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ вытекает, что можно подобрать числа α_1 и β_1 так, чтобы $\det \tilde{B} \neq 0$ и $\det \tilde{B}_1 \neq 0$. Поэтому по теореме выше пучки $\tilde{\mu} \tilde{A} + \tilde{\lambda} \tilde{B}$ и $\tilde{\mu} \tilde{A}_1 + \tilde{\lambda} \tilde{B}_1$, а следовательно, и исходные пучки $\mu A + \lambda B$ и $\mu A_1 + \lambda B_1$ (или, что то же, $A + \lambda B$ и $A_1 + \lambda B_1$) строго эквивалентны. Таким образом, мы пришли к следующей теореме

Теорема. *Для того чтобы два регулярных пучка $A + \lambda B$ и $A_1 + \lambda B_1$ были строго эквивалентны, необходимо и достаточно, чтобы эти пучки имели одни и те же ("конечные" и "бесконечные") элементарные делители.*

В разобранном примере пучки отличались "бесконечными" элементарными делителями (первый пучок имеет один "бесконечный" элементарный делитель μ^2, а второй — два: μ, μ). Поэтому пучки и не оказались строго эквивалентными.

8 Каноническая форма регулярного матричного пучка

Пусть дан произвольный пучок $A + \lambda B$. Тогда существует число c, что $\det(A + cB) \neq 0$. Представим данный пучок в виде

$$A + \lambda B = A + cB - cB + \lambda B = A_1 + (\lambda - c)B, \tag{20}$$

где $A_1 = A + cB$, и поэтому $\det A_1 \neq 0$. Умножим (20) слева на A_1^{-1} —

$$E + (\lambda - c)A_1^{-1}B.$$

Матрицу A_1^{-1} приведем к жордановой форме с помощью невырожденного преобразования T

$$T\left(E + (\lambda - c)A_1^{-1}B\right)T^{-1} = E + (\lambda - c)TA_1^{-1}BT^{-1} =$$

$$= \begin{pmatrix} E + (\lambda - c)J_0 & O \\ O & E + (\lambda - c)J_1 \end{pmatrix}, \tag{21}$$

J_0 — жорданова нильпотентная матрица, а $\det J_1 \neq 0$.

В силу структуры матрицы J_0 существует $(E - cJ_0)^{-1}$ (пояснить *самостоятельно*). Умножая (21) справа на матрицу $\begin{pmatrix} (E - cJ_0)^{-1} & O \\ O & E \end{pmatrix}$, получим

$$\begin{pmatrix} E + \lambda (E - cJ_0)^{-1} J_0 & O \\ O & E + (\lambda - c)J_1 \end{pmatrix}. \tag{22}$$

Здесь $(E - cJ_0)^{-1} J_0$ нильпотентная матрица, ее жорданова форма будет представлять собой жорданову нильпотентную матрицу N, т.е. существует матрица T_1, $\det T_1 \neq 0$, такая, что

$$T_1\left(E + \lambda (E - cJ_0)^{-1} J_0\right)T_0^{-1} = E + \lambda T_1\left((E - cJ_0)^{-1} J_0\right)T_1^{-1} = E + \lambda N =$$

$$= E + \lambda \begin{pmatrix} N_{u_1} & O & \dots & O \\ O & N_{u_2} & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & N_{u_s} \end{pmatrix},$$

где матрица

$$N_{u_i} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

порядка u_i. Теперь умножим (22) слева на матрицу $\begin{pmatrix} T_1 & O \\ O & J_1^{-1} \end{pmatrix}$ и справа на $\begin{pmatrix} T_1 & O \\ O & E \end{pmatrix}$, получим

$$\begin{pmatrix} E + \lambda N & O \\ O & J_1^{-1} + (\lambda - c)E \end{pmatrix},$$

а затем преобразованием подобия приведем матрицу $J_1^{-1} - cE$ к жордановой форме

$$T_2 \left(J_1^{-1} + (\lambda - c)E \right) T_2^{-1} = \lambda E + T_2 \left(J_1^{-1} - cE \right) T_2^{-1} = \lambda E + J,$$

где J — жорданова форма матрицы $J_1^{-1} - cE$. Мы пришли к теореме.

Теорема. *Произвольный регулярный пучок $A+\lambda B$ может быть приведен к строго эквивалентному виду*

$$\begin{pmatrix} E + \lambda N & O \\ O & J + \lambda E \end{pmatrix} =$$

$$= \begin{pmatrix} E_{u_1} + \lambda N_{u_1} & O & \dots & O & O \\ O & E_{u_2} + \lambda N_{u_2} & \dots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \dots & E_{u_s} + \lambda N_{u_s} & O \\ O & O & \dots & O & J + \lambda E \end{pmatrix},$$

где первые s диагональнвх блоков соответствуют бесконечным элементарным делителям $\mu^{u_1}, \mu^{u_2}, \dots, \mu^{u_s}$ пучка $A + \lambda B$, а жорданова форма последнего диагонального блока $J + \lambda E$ однозначно определяется конечными элементарными делителями данного пучка.

9 Сингулярные пучки

Переходим к рассмотрению сингулярного пучка матриц $A + \lambda B$ размера $m \times n$. Обозначим через r ранг пучка, т.е. наибольший из порядков миноров, не равных тождественно нулю. Из сингулярности пучка следует, что всегда имеет место по крайней мере одно из неравенств $r < n$ или $r < m$. Пусть для определенности $r < n$. Тогда столбцы λ-матрицы линейно зависимы, т.е. уравнение

$$(A + \lambda B)\, x = 0, \tag{23}$$

где x — искомый столбец, имеет ненулевое решение. Каждое ненулевое решение этого уравнения определяет некоторую линейную зависимость между столбцами λ-матрицы $A+\lambda B$. Среди множества решений системы (23) можно выбрать базисные линейно независимые решения. Поскольку коэффициенты линейной однородной системы алгебраических уравнений линейно зависят от λ, то базисные решения x всегда могут быть выбраны так, чтобы их элементами были многочлены от λ. Среди этих решений возьмем решение наименьшей степени ε

$$x(\lambda) = x_0 - \lambda x_1 + \lambda^2 x_2 - \ldots + (-1)^\varepsilon \lambda^\varepsilon x_\varepsilon \quad (x_\varepsilon \neq 0)$$

(ε при этом называется *минимальным индексом пучка.* подставляя это решение в (23) и приравнивая к нулю коэффициенты при одинаковых степенях λ, получим

$$Ax_0 = O, \;\; Bx_0 - Ax_1 = O, \;\; Bx_1 - Ax_2 = O, \;\; \ldots,$$
$$Bx_{\varepsilon-1} - Ax_\varepsilon = O, \;\; Bx_\varepsilon = O.$$

Рассматривая эту систему равенств как систему линейных однородных уравнений относительно элементов столбцов

$$x_0, \; -x_1, \; x_2, \; \ldots, \; (-1)^\varepsilon x_\varepsilon,$$

заключаем, что матрица коэффициентов этой системы

$$M_\varepsilon = M_\varepsilon\,[A + \lambda B] = \begin{array}{c} \varepsilon+1 \\ \ldots\ldots\ldots\ldots \\ \begin{pmatrix} A & O & \ldots & O \\ B & A & \ldots & O \\ O & B & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & A \\ O & O & \ldots & B \end{pmatrix} \end{array}$$

имеет ранг $\rho_\varepsilon < (\varepsilon + 1)n$. В то же врем, поскольку $x(\lambda)$ было решение наименьшей степени ϵ, отличное от нуля, то все решения степени $0, 1, \ldots, \varepsilon - 1$ должны быть нулевыми, поэтому для матриц

$$M_0 = \begin{pmatrix} A \\ B \end{pmatrix}, \quad M_1 = \begin{pmatrix} A & O \\ B & A \\ O & B \end{pmatrix}, \quad \ldots, \quad M_{\varepsilon-1} = \underbrace{\begin{pmatrix} A & O & O & \ldots & O & O \\ B & A & O & \ldots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \ldots & B & A \\ O & O & O & \ldots & O & B \end{pmatrix}}_{\varepsilon}$$

имеют место равенства

$$\rho_0 = n, \quad \rho_1 = 2n, \quad \ldots, \quad \rho_{\varepsilon-1} = \varepsilon n,$$

т.е. эти матрицы имеют полный ранг по столбцам.

Таким образом, *число ε есть наименьшее значение индекса k, при котором в соотношении $\rho_k \leq (k+1)n$ имеет место знак $<$.*

Сформулируем следующую фундаментальную теорему (док-во см. [5, с. 316]).

Теорема (о приведении). *Если уравнение (23) имеет решение минимальной степени ε и $\varepsilon > 0$, то данный пучок эквивалентен пучку вида*

$$\begin{pmatrix} L_\varepsilon & O \\ O & \hat{A} + \lambda \hat{B} \end{pmatrix},$$

где

$$L_\varepsilon = \underbrace{\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \end{pmatrix}}_{\varepsilon+1} \left.\vphantom{\begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}}\right\} \varepsilon \quad ,$$

а $\hat{A} + \lambda \hat{B}$ — пучок матриц, для которых уравнение аналогичное уравнению (23), не имеет решений степени меньше ε.

10 Каноническая форма сингулярного пучка матриц

Пусть дан произвольный сингулярный пучок матриц $A+\lambda B$ размера $m\times n$. Допустим сначала, что как между столбцами, так и между строками этого пучка нет линейной зависимости с постоянными коэффициентами, т.е. $\varepsilon \neq 0$.

Для определенности будем предполагать линейную зависимость столбцов: $r < n$. В этом случае уравнение $(A+\lambda B)\,x = 0$ имеет ненулевое решение минимальной степени $\varepsilon_1 > 0$. По теореме о приведении данный пучок строго эквивалентен пучку

$$\begin{pmatrix} L_{\varepsilon_1} & O \\ O & A_1+\lambda B_1 \end{pmatrix},$$

где уравнение $(A_1+\lambda B_1)\,x^{(1)} = 0$ не имеет решений $x^{(1)}$ степени $< \varepsilon_1$.

Если это уравнение имеет ненулевое решение минимальной степени ε_2 (при этом непременно $\varepsilon_2 \geq \varepsilon_1$), то, применяя к пучку $A_1+\lambda B_1$ теорему о приведении, мы данный пучок преобразуем к виду

$$\begin{pmatrix} L_{\varepsilon_1} & O & O \\ O & L_{\varepsilon_2} & O \\ O & O & A_2+\lambda B_2 \end{pmatrix}.$$

Продолжая этот процесс далее, мы приведем данный пучок к квазидиагональному виду

$$\begin{pmatrix} L_{\varepsilon_1} & O & \ldots & O & O \\ O & L_{\varepsilon_2} & \ldots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \ldots & L_{\varepsilon_p} & O \\ O & O & \ldots & O & A_p+\lambda B_p \end{pmatrix},$$

где $0 < \varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_p$, а уравнение $(A_p+\lambda B_p)\,x^{(p)} = 0$ не имеет ненулевых решений, т.е. столбцы матрицы $A_p+\lambda B_p$ линейно независимы.

Но в этой матрице могут быть линейно зависимы строки, тогда по той же теореме транспонированный пучок $A^\top + \lambda B^\top$ может быть преобразован к аналогичному квазадиагональному виду, где вместо чисел

$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p$ будут фигурировать числа $(0 <)\ \eta_1 \le \eta_2 \le \ldots \le \eta_q$. Тогда исходный пучок $A + \lambda B$ окажется преобразованным к виду

$$\begin{pmatrix} L_{\varepsilon_1} & O & \ldots & O & O & O & \ldots & O & O \\ O & L_{\varepsilon_2} & \ldots & O & O & O & \ldots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ O & O & \ldots & L_{\varepsilon_p} & O & O & \ldots & O & O \\ O & O & \ldots & O & L_{\eta_1}^{\top} & O & \ldots & O & O \\ O & O & \ldots & O & O & L_{\eta_2}^{\top} & \ldots & O & O \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \ldots & O & O & O & \ldots & L_{\eta_q}^{\top} & O \\ O & O & \ldots & O & O & O & \ldots & O & A_0 + \lambda B_0 \end{pmatrix} \tag{24}$$

$(0 < \varepsilon_1 \le \varepsilon_2 \le \ldots \le \varepsilon_p,\ \ 0 < \eta_1 \le \eta_2 \le \ldots \le \eta_q)$, где у пучка $A_0 + \lambda B_0$ как столбцы, так и строки линейно независимы, т.е. $A_0 + \lambda B_0$ — регулярный пучок.

Рассмотрим теперь общий случай, когда строки и столбцы пучка $A + \lambda B$ могут быть связаны линейными зависимостями с постоянными коэффициентами. Обозначим максимальное число постоянных линейно независимых решений уравнений

$$(A + \lambda B)\, x = 0, \qquad \left(A^{\top} + \lambda B^{\top}\right) y = 0$$

соответственно через g и h. Вместо первого из этих уравнений рассмотрим соответствующее операторное уравнение $(\mathcal{A} + \lambda \mathcal{B})\, x = 0$ ($\mathcal{A}$, $\mathcal{B}$ — линейные операторы из $\mathbf{R}^n$ в $\mathbf{R}^m$). Линейно независимые постоянные решения этого уравнения обозначим через $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_g$ и примем их за первые g базисных векторов в $\mathbf{R}^n$. Тогда в соответствующей матрице $\tilde{A} + \lambda \tilde{B}$ первые g столбцов будут состоять из нулей (пояснить *самостоятельно*)

$$\begin{pmatrix} 0 & \ldots & 0 & \\ \vdots & \ldots & \vdots & \tilde{A}_1 + \lambda \tilde{B}_1 \\ 0 & \ldots & 0 & \end{pmatrix}.$$

Замечание. Каждому линейному оператору $\mathcal{L} : \mathbf{R}^n \to \mathbf{R}^m$ с выбранными базисами: $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ (в $\mathbf{R}^n$) и $\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\}$ (в $\mathbf{R}^m$); соответствует матрица L, составленная из координат векторов $\mathcal{L}[\mathbf{e}_1]$, $\mathcal{L}[\mathbf{e}_2]$, …, $\mathcal{L}[\mathbf{e}_n]$ при разложении их по базису $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m$. Матрицы соответствующие одному и тому же линейному оператору в разных базисах связаны преобразованием подобия: $L = T^{-1}\tilde{L}T$.

Совершенно также в пучке $\tilde{A}_1 + \lambda\tilde{B}_1$ первые h строк можно сделать нулевыми. Тогда исходный пучок примет вид

$$\begin{array}{cc} & \overset{g}{\overbrace{}} \\ h\{ & \begin{pmatrix} O & O \\ O & A_0 + \lambda B_0 \end{pmatrix} \end{array},$$

где строки и столбцы пучка $A_0+\lambda B_0$ уже не связаны линейными зависимостями с постоянными коэффициентами. К пучку $A_0+\lambda B_0$ применимо представление (24). Таким образом, в самом общем случае пучок $A+\lambda B$ всегда может быть приведен к каноническому квазидиагональному виду

$$\mathrm{diag}\left\{O,\ L_{\varepsilon_1},\ \ldots,\ L_{\varepsilon_p}, L_{\eta_1}^{\top},\ \ldots,\ L_{\eta_q}^{\top},\ A_0+\lambda B_0\right\},$$

где нулевой блок имеет размеры $h \times g$.

Заменяя фигурирующий здесь регулярный пучок $A_0+\lambda B_0$ его канонической формой, получим окончательно следующую квазидиагональную матрицу

$$\begin{gathered}\mathrm{diag}\left\{O,\ L_{\varepsilon_1},\ \ldots,\ L_{\varepsilon_p}, L_{\eta_1}^{\top},\ \ldots,\ L_{\eta_q}^{\top},\right. \\ \left. E_{u_1}+\lambda N_{u_1},\ \ldots,\ E_{u_s}+\lambda N_{u_s},\ J+\lambda E\right\},\end{gathered} \tag{25}$$

где матрица J имеет жорданову форму.

Матрица (25) представляет собой каноническую форму пучка $A+\lambda B$ в самом общем случае.

11 Решение систем дифференциальных уравнений

Рассмотрим приложение полученных результатов к интегрированию системы m линейных обыкновенных дифференциальных уравнений первого порядка с постоянными коэффициентами относительно n неизвестных функций

$$Ax(t) + B\frac{dx(t)}{dt} = f(t),\quad t \in T = [a, b), \tag{26}$$

$A = \|a_{ij}\|_{i=\overline{1,m},\, j=\overline{1,n}}, \quad B = \|b_{ij}\|_{i=\overline{1,m},\, j=\overline{1,n}},$

$$x(t) = \operatorname{colon}(x_1(t), x_2(t), \ldots, x_n(t)) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

$$\frac{dx(t)}{dt} = \operatorname{colon}\left(\frac{dx_1(t)}{dt}, \frac{dx_2(t)}{dt}, \ldots, \frac{dx_n(t)}{dt}\right).$$

Для системы (26) введем замену переменных $x(t) = Qz(t)$, где $Q = \|q_{ij}\|_{i,j=\overline{1,n}}$, $\det Q \neq 0$, $z = \operatorname{colon}(z_1(t), z_2(t), \ldots, z_n(t))$. А затем умножим полученную систему слева на матрицу $P = \|p_{ij}\|_{i,j=\overline{1,m}}$, $\det P \neq 0$. В результате получим

$$\tilde{A}z(t) + \tilde{B}\frac{dz(t)}{dt} = \tilde{f}(t), \quad t \in T, \tag{27}$$

где $\tilde{A} = PAQ$, $\tilde{B} = PBQ$, $\tilde{f} = Pf$.

При этом пучки матриц $A + \lambda B$ и $\tilde{A} + \lambda\tilde{B}$ строго эквивалентны друг другу:

$$\tilde{A} + \lambda\tilde{B} = P\,(A + \lambda B)\,Q.$$

Выберем матрицы P и Q так, чтобы пучок $\tilde{A}+\lambda\tilde{B}$ имел каноническую квазидиагональную форму

$$\tilde{A} + \lambda\tilde{B} = \operatorname{diag}\Big\{O,\ L_{\varepsilon_1},\ \ldots,\ L_{\varepsilon_p}, L^{\top}_{\eta_1},\ \ldots,\ L^{\top}_{\eta_q},$$

$$E_{u_1} + \lambda N_{u_1},\ \ldots,\ E_{u_s} + \lambda N_{u_s},\ J + \lambda E\Big\},$$

где нулевой блок состоит из g столбцов и h строк.

В соответствии с диагональными блоками система дифференциальных уравнений (26) распадается на несколько отдельных систем следующих типов:

$$(I) \qquad Oz^{(1)}(t) = \tilde{f}^{(1)}(t),$$

$z^{(1)}(t) = \operatorname{colon}\left(z_1^{(1)}(t), z_2^{(1)}(t), \ldots, z_g^{(1)}(t)\right)$, O — $(h \times g)$–нулевая матрица;

$$(II) \qquad L_\varepsilon\left(\frac{d}{dt}\right) z^{(2)}(t) = \tilde{f}^{(2)}(t),$$

$$L_\varepsilon = \underbrace{\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \end{pmatrix}}_{\varepsilon+1} \Bigg\} \ \varepsilon;$$

(III) $$L_\eta^\top\left(\frac{d}{dt}\right) z^{(3)}(t) = \tilde f^{(3)}(t),$$

$$L_\eta^\top = \underbrace{\begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ 1 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \lambda \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}}_{\eta};$$

(IV) $$z^{(4)}(t) + N_u \frac{dz^{(4)}(t)}{dt} = \tilde f^{(4)}(t),$$

$$N_u = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix};$$

(V) $$Jz^{(5)}(t) + \frac{dz^{(5)}(t)}{dt} = \tilde f^{(5)}(t).$$

Здесь использовано обозначение: $\Lambda\left(\frac{d}{dt}\right) = A + B\frac{d}{dt}$, если $\Lambda(\lambda) = A + \lambda B$.

Таким образом, в самом общем случае интегрирование системы (26) сводится к интегрированию частных систем (I)–(V) такого же типа.

Проанализируем полученные системы.

(I). Для того чтобы система (27) не была противоречива (т.е. имела решение при заданной правой части $\tilde{f}(t)$ необходимо и достаточно, чтобы

$$\tilde{f}^{(1)}(t) \equiv 0. \tag{28}$$

В этом случае компонентами вектора $z^{(1)}(t)$ могут быть произвольные функции аргумента t.

(II). В подробной записи система (II) имеет вид

$$\frac{dz_1^{(2)}(t)}{dt} + z_2^{(2)}(t) = \tilde{f}_1^{(2)}(t),$$

$$\frac{dz_2^{(2)}(t)}{dt} + z_3^{(2)}(t) = \tilde{f}_2^{(2)}(t),$$

$$\dots \quad \dots \quad \dots \quad \dots$$

$$\frac{dz_\varepsilon^{(2)}(t)}{dt} + z_{\varepsilon+1}^{(2)}(t) = \tilde{f}_\varepsilon^{(2)}(t).$$

Такая система всегда совместна. Если в качестве $z_{\varepsilon+1}^{(2)}(t)$ взять произвольную функцию аргумента t, то последовательными интегрированиями из (28) определятся все остальные неизвестные функции $z_\varepsilon^{(2)}(t), z_{\varepsilon-1}^{(2)}(t), \dots,$ $z_1^{(2)}(t)$.

(III). Система (III) есть система вида

$$\frac{dz_1^{(3)}(t)}{dt} = \tilde{f}_1^{(3)}(t),$$

$$\frac{dz_2^{(3)}(t)}{dt} + z_1^{(3)}(t) = \tilde{f}_2^{(3)}(t),$$

$$\dots \quad \dots \quad \dots \quad \dots$$

$$\frac{dz_\eta^{(3)}(t)}{dt} + z_{\eta-1}^{(3)}(t) = \tilde{f}_\eta^{(3)}(t),$$

$$z_\eta^{(3)}(t) = \tilde{f}_{\eta+1}^{(3)}(t).$$

Из этих уравнений, кроме первого, мы однозначно определяем $z_\eta^{(3)}(t), z_{\eta-1}^{(3)}(t),$ $\dots, z_1^{(3)}(t)$:

$$z_\eta^{(3)}(t) = \tilde{f}_{\eta+1}^{(3)}(t), \quad z_{\eta-1}^{(3)}(t) = \tilde{f}_\eta^{(3)}(t) - \frac{d\tilde{f}_{\eta+1}^{(3)}(t)}{dt}, \quad \dots,$$

$$z_1^{(3)}(t) = \tilde{f}_2^{(3)}(t) - \frac{d\tilde{f}_3^{(3)}(t)}{dt} + \ldots + (-1)^{\eta-1}\frac{d^{\eta-1}\tilde{f}_{\eta+1}^{(3)}(t)}{(dt)^{\eta-1}}.$$

Подставляя полученное выражение для $z_1^{(3)}(t)$ в первое уравнение, получаем условие совместности системы (27)

$$\tilde{f}_1^{(3)}(t) - \frac{d\tilde{f}_2^{(3)}(t)}{dt} + \frac{d^2\tilde{f}_3^{(3)}(t)}{dt} + \ldots + (-1)^{\eta}\frac{d^{\eta}\tilde{f}_{\eta+1}^{(3)}(t)}{(dt)^{\eta}} = 0. \qquad (29)$$

(IV). Система (IV) в подробной записи имеет вид

$$z_1^{(4)}(t) + \frac{dz_2^{(4)}(t)}{dt} = \tilde{f}_1^{(4)}(t),$$

$$z_2^{(4)}(t) + \frac{dz_3^{(4)}(t)}{dt} = \tilde{f}_2^{(4)}(t),$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots$$

$$z_u^{(4)}(t) = \tilde{f}_u^{(4)}(t),$$

отсюда последовательно однозначно определяем решение

$$z_u^{(4)}(t) = \tilde{f}_u^{(4)}(t), \;\; z_{u-1}^{(4)}(t) = \tilde{f}_{u-1}^{(4)}(t) - \frac{d\tilde{f}_u^{(4)}(t)}{dt}, \;\; \ldots,$$

$$z_1^{(4)}(t) = \tilde{f}_1^{(4)}(t) - \frac{d\tilde{f}_2^{(4)}(t)}{dt} + \frac{d^2\tilde{f}_3^{(4)}(t)}{(dt)^2} - \ldots + (-1)^{u-1}\frac{d^{u-1}\tilde{f}_u^{(4)}(t)}{(dt)^{u-1}}.$$

(V). Общее решение системы (V), разрешенной относительно производной, имеет вид

$$z^{(5)}(t) = \Omega(t)c + \int_a^t \Omega(t)\Omega^{-1}(\tau)\tilde{f}^{(5)}(\tau)d\tau, \;\; t \in T = [a, b),$$

где c — произвольный числовой вектор, $\Omega(t)$ — матрицант системы (V), т.е. матрица, являющаяся решением задачи Коши

$$\Omega'(t) + J\Omega(t) = O, \;\; \Omega(a) = E.$$

Обратный переход от решения системы (27) к решению системы (26) осуществляется преобразованием

$$x(t) = Qz(t).$$

Проведенный анализ показывает, что *для совместности системы (26) в общем случае должны выполняться некоторые определенные линейные алгебраические и дифференциальные зависимости (с постоянными коэффициентами) между правыми частями уравнений типа (28), (29), где* $\tilde{f}(t) = Pf(t)$.

Если эти условия выполнены, то общее решение системы (в общем случае) содержит линейно как произвольные постоянные, так и произвольные функции.

Характер условий совместности и структура решения (в частности, количество произвольных констант и произвольных функций) определяются минимальными индексами и элементарными делителями пучка $A + \lambda B$, *поскольку от этих индексов и делителей зависит каноническая форма системы дифференциальных уравнений.*

Часть II

Применение обобщенных обратных матриц для решения линейных алгебро-дифференциальных систем

1 Полуобратные матрицы

В этом разделе излагаются результаты, касающиеся так называемых обобщенных обратных матриц (полуобратных, псевдообратных матриц и обратных матриц Дразина), в совокупности представляющие собой алгебраический аппарат, который используется в дальнейшем при изучении линейных алгебро-дифференциальных систем.

Определение. *Полуобратной матрицей* к $(m \times n)$-матрице A называется $(n \times m)$-матрица X, удовлетворяющая уравнению

$$AXA = A. \tag{1}$$

Полуобратная матрица существует для любой $(m \times n)$-матрицы. Однако она, вообще говоря, не единственна.

В дальнейшем любую из полуобратных матриц к матрице A будем обозначать A^-, так что

$$AA^-A = A. \tag{2}$$

Описание множества полуобратных матриц к матрице A дают следующие леммы.

Лемма 1. *Общее решение матричного уравнения (1) представляется в виде*

$$X = A^- + \left(E - A^-A\right)U + V\left(E - AA^-\right), \tag{3}$$

где A^- — любая полуобратная матрица к матрице A, а U и V — произвольные матрицы подходящих размеров.

Иначе говоря, зная одну из полуобратных матриц для матрицы A, по формуле (3) можно получить и все остальные.

Доказательство. Очевидно, что матрица, имеющая вид (3), удовлетворяет уравнению (1) (*показать*). Обратно, если матрица X удовлетворяет уравнению (1), то она представляется в виде (3) при $U = X$ и $V = A^-AX - A^-$. Этим и завершается доказательство леммы.

Лемма 2. *Пусть $A_1 = PAQ$, где матрицы P и Q неособенные. Тогда, если матрица X пробегает все полуобратные матрицы к матрице A, то матрица $X_1 = Q^{-1}XP^{-1}$ пробегает все полуобратные матрицы к матрице A_1.*

Доказательство. Пусть A^- — полуобратная матрица к матрице A. Тогда матрица $A_1^- = Q^{-1}A^-P^{-1}$ есть полуобратная матрица к матрице A_1. В самом деле,

$$A_1A_1^-A_1 = PAQQ^{-1}A^-P^{-1}PAQ = PAA^-AQ = PAQ = A_1.$$

Если же матрица A_1^- является полуобратной к матрице A_1, то она необходимо представляется в виде $A_1^- = Q^{-1}A^-P^{-1}$, где матрица $A^- = QA_1^-P$ вследствие равенств

$$AA^-A = AQA_1^-PA = P^{-1}PAQA_1^-PAQQ^{-1} = P^{-1}A_1A_1^-A_1Q^{-1} =$$

$$= P^{-1}A_1Q^{-1} = P^{-1}PAQQ^{-1} = A$$

есть полуобратная к матрице A. Лемма доказана.

Лемма 3. *Пусть в блочной матрице $K = \begin{pmatrix} A \\ B \end{pmatrix}$ подматрица A неособенная. Тогда блочная матрица вида $\begin{pmatrix} A^{-1} & 0 \end{pmatrix}$ является полуобратной к матрице K.*

Доказать самостоятельно.

Лемма 4. *Пусть в блочной матрице $K = \begin{pmatrix} A & B \end{pmatrix}$ подматрица A неособенная. Тогда блочная матрица $\begin{pmatrix} A^{-1} \\ O \end{pmatrix}$ является полуобратной к матрице K.*

Доказать самостоятельно.

Определение. *Инверсной полуобратной матрицей* к матрице A называется матрица $A^{\sim}$, которая удовлетворяет двум матричным уравнениям

$$AA^{\sim}A = A, \quad A^{\sim}AA^{\sim} = A^{\sim}. \tag{4}$$

Лемма 5. *Общее решение системы матричных уравнений (4) представляется в виде*

$$A^{\sim} = A^{-}AA^{-} + A^{-}AU\left(E - AA^{-}\right) + \left(E - A^{-}A\right)UAA^{-}+$$

$$+\left(E - A^{-}A\right)UAU\left(E - AA^{-}\right), \tag{5}$$

где A^{-} *— любая полуобратная матрица к матрице* A*, а* U *— произвольная матрица подходящих размеров. Иначе говоря, по формуле (5) можно получить все инверсные полуобратные матрицы к матрице* A*.*

Доказательство. С помощью непосредственной подстановки легко проверить, что матрица, имеющая вид (5), удовлетворяет системе (4). Обратно, если матрица X удовлетворяет системе (4), то она представима в виде (5) при $U = X - A^{-}AA^{-}$. Этим и завершается доказательство леммы.

Лемма 6. *Когда матрица* X *пробежит все полуобратные матрицы к матрице* A*, матрица* $Y = XAX$ *пробежит все инверсные полуобратные матрицы к матрице* A*.*

Доказательство. Пусть X — полуобратная матрица к матрице A, т.е. $AXA = A$. Тогда

$$AYA = AXAXA = AXA = A,$$

$$YAY = X(AXA)XAX = X(AXA)X = XAX = Y.$$

Таким образом, матрица $Y = XAX$ удовлетворяет системе (4) и потому является инверсной полуобратной матрицей к матрице A. Если же Y — произвольная инверсная полуобратная матрица к матрице A, то в силу второго уравнения системы (4) она представляется в виде $Y = XAX$ при значении полуобратной матрицы X к матрице A, равном матрице Y. Лемма доказана.

1.1 Псевдообратная матрица

Среди инверсных полуобратных матриц особое место занимает так называемая псевдообратная матрица. Исключительность псевдообратной матрицы связана с тем, что она для каждой матрицы существует, единственна и ее приближенное представление легко находится.

Определение. Матрица A^{+} размеров $(n \times m)$ называется *псевдообратной* к матрице A, если выполнены равенства:

1) $AA^+A = A$,
2) $A^+AA^+ = A^+$,
3) $(AA^+)^* = AA^+$,
4) $(A^+A)^* = A^+A$,
которые называются уравнениями Мура-Пенроуза.

P^* обозначает матрицу сопряженную к матрице P.

Известно, что произвольная прямоугольная $(m \times n)$-матрица A, имеющая ранг r, представляется в виде произведения

$$A = BC$$

матриц B и C соответственно размеров $(m \times r)$ и $(r \times n)$. Это представление называется *скелетным разложением матрицы A*. Оно не единственно, однако матрицы B^*B и CC^* неособенны и имеет место равенство рангов

$$\operatorname{rank} B = \operatorname{rank} C = \operatorname{rank} A.$$

Теорема 1. *Псевдообратная матрица к произвольной матрице A существует, единственна и имеет представление*

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*, \tag{6}$$

где B и C — компоненты скелетного разложения $A = BC$ матрицы A.

Доказательство. То, что псевдообратная матрица к матрице существует и представляется в виде (6), доказывается путем непосредственной проверки выполнения равенств 1)–4) для матрицы (6). Остается доказать, что для данной матрицы A не может существовать двух различных псевдообратных матриц A_1^+ и A_2^+.

В самом деле, пусть

$$D = A_1^+ - A_2^+.$$

Тогда первое равенство 1) дает

$$ADA = O, \tag{7}$$

а 3) и 4) —

$$(DA)^* = DA, \quad (AD)^* = AD. \tag{8}$$

Из равенств (7) и (8) вытекают равенства

$$(DA)^*DA = (DA)^2 = DADA = O,$$

$$(AD)^*AD = (AD)^2 = ADAD = O$$

и, следовательно, $AD = O$ и $DA = O$, так как для любой матрицы M из $M^*M = 0$ следует $M = O$. Но тогда

$$A_1^+A = A_2^+A, \quad AA_1^+ = AA_2^+$$

с учетом 2) из 1)–4)

$$A_1^+ = A_1^+AA_1^+ = A_2^+AA_1^+ = A_2^+AA_2^+ = A_2^+.$$

Теорема доказана.

Следствие 1. *Если столбцы $(m \times n)$-матрицы A линейно независимы, т.е.* $\operatorname{rank} A = n$, *то*

$$A^+ = (A^*A)^{-1} A^*. \tag{9}$$

Доказательство. В условиях следствия скелетное разложение выглядит как $A = AE$. Но тогда из (6) следует (9).

Следствие 2. *Если строки $(m \times n)$-матрицы A линейно независимы, т.е.* $\operatorname{rank} A = m$, *то*

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

Доказательство аналогично предыдущему.

Одним из преимуществ псевдообратной матрицы является легкость ее приближенного вычисления.

Теорема 2.

$$A^+ = \lim_{\delta \to 0} \left(\delta^2 E + A^*A\right)^{-1} A^* = \lim_{\delta \to 0} A^* \left(\delta^2 E + AA^*\right)^{-1} \tag{10}$$

и имеет место оценка

$$\|A^+ - \left(\delta^2 E + A^*A\right)^{-1} A^*\| \le \delta^2 \|A^+ \left(A^+\right)^* A^+\|, \tag{11}$$

где переменная δ — вещественная; $\|A\|$ — спектральная норма матрицы, которая определяется как максимум из модулей собственных значений матрицы (спектральная норма матрицы согласована с евклидовой нормой векторов).

Доказательство. Так как $A^+ = A^+AA^+$, $(A^+A)^* = A^+A$, то $(A^+)^* = (A^+)^* A^+A$ и, следовательно,

$$A^+ = A^* \left(A^+\right)^* A^+. \tag{12}$$

Кроме того, поскольку $AA^+A = A$, $(AA^+)^* = AA^+$, имеем

$$\left(A^* \left(AA^+ - E\right)\right)^* = \left(AA^+ - E\right) A = O,$$

а потому

$$A^* = A^*AA^+. \tag{13}$$

Подставив теперь (13) в (12), получим

$$A^+ = A^*AA^+ \left(A^+\right)^* A^+. \tag{14}$$

Но тогда с учетом (13), (14)

$$A^+ - \left(\delta^2 E + A^*A\right)^{-1} A^* = \left(\delta^2 E + A^*A\right)^{-1} \left(\delta^2 A^+ + A^*AA^+ - A^*\right) =$$

$$= \delta^2 \left(\delta^2 E + A^*A\right)^{-1} A^*AA^+ \left(A^+\right)^* A^+.$$

Для спектральной нормы хорошо известна оценка

$$\| \left(\delta^2 E + A^*A\right)^{-1} A^*A\| \leq 1,$$

с учетом которой получается неравенство (11), а значит, и первое равенство из (10). Второе равенство (10) также справедливо, поскольку легко убедиться , что $(\delta^2 E + A^*A)^{-1} A^* = A^* (\delta^2 E + AA^*)^{-1}$.

1.2 Обратная матрица Дразина

Обратная матрица Дразина определяется только для квадратных матриц.

Определение. *Обратной матрицей Дразина* к матрице A называется матрица A^D, которая удовлетворяет системе уравнений
1^0. $AA^D = A^DA$,
2^0. $A^DAA^D = A^D$,
3^0. $\left(E - A^DA\right) A^k = O$,

где k — индекс матрицы A, т.е. наименьшее из неотрицательных целых чисел, для которых $\operatorname{rank} A^{k+1} = \operatorname{rank} A^k$.

Равенство $k = 0$ означает, что матрица A неособенная. В этом случае системе 1^0–3^0 удовлетворяет матрица $A^D = A^{-1}$. При $k = 1$ матрица A^D в силу 1^0–3^0 совпадает с одной из инверсных полуобратных матриц $A^{\tilde{}}$. Если же матрица A самосопряженная, то $A^D = A^+$.

Простой смысл обратной матрицы Дразина вскрывает следующая теорема.

Теорема 3. *Обратная матрица Дразина A^D к произвольной $(n \times n)$- матрице A существует, единственна и допускает представление, которое строится так: пусть*

$$A = P \begin{pmatrix} J_0 & O \\ O & J_1 \end{pmatrix} P^{-1} \tag{15}$$

есть жорданова форма матрицы A, где J_0 состоит из нильпотентных блоков, а J_1 — из неособенных. Тогда

$$A^D = P \begin{pmatrix} O & O \\ O & J_1^{-1} \end{pmatrix} P^{-1}. \tag{16}$$

Доказательство. Сначала заметим, что индекс k матрицы A равен порядку l наибольшего нильпотентного блока в ее жордановом представлении (15), т.к. $\operatorname{rank} A^{i+1} < \operatorname{rank} A^i$ при $i < l$ и $\operatorname{rank} A^{l+1} < \operatorname{rank} A^l$. Теперь, учитывая, что $J_0^k = O$, сделаем подстановку представлений (15), (16) в равенства 1^0–3^0. В результате убедимся, что равенства 1^0–3^0 удовлетворяются.

Так доказывается существование обратной матрицы Дразина A^D, имеющей представление (16).

Докажем теперь, что для данной матрицы A не может существовать двух различных обратных матриц Дразина A_1^D и A_2^D.

Пусть $R = A_1^D - A_2^D$. Первое и третье равенства из 1^0–3^0 дают

$$RA = AR, \quad RA^{k+1} = O, \tag{17}$$

а второе —

$$\left(AA_1^D\right)^{k+1} = AA_1^D AA_1^D \left(AA_1^D\right)^{k-1} = AA_1^D \left(AA_!^D\right)^{k-1} =$$

$$= \left(AA_1^D\right)^k = \ldots = AA_1^D, \tag{18}$$

и аналогично

$$\left(A_2^D A\right)^{k+1} = A_2^D A. \tag{19}$$

Из (17)– (19) с учетом первого и второго равенств 1^0–3^0 следует, что

$$A_1^D = A_1^D A A_1^D = \left(A_1^D - A_2^D\right) A A_1^D + A_2^D A \left(A_1^D - A_2^D\right) + A_2^D =$$

$$= RAA_1^D + A_2 AR + A_2^D = RA^{k+1} \left(A_1^D\right)^{k+1} +$$

$$+ \left(A_2^D\right)^{k+1} A^{k+1} R + A_2^D = A_2^D.$$

Теорема доказана.

Теорема 4. *Если матрица P неособенная и*

$$A = PBP^{-1}, \tag{20}$$

то

$$A^D = PB^D P^{-1}. \tag{21}$$

Доказательство можно осуществить подстановкой (20), (21) в равенства 1^0–3^0 с учетом того, что вследствие неособенности матрицы P ранги, а следовательно, и индексы матриц A и B совпадают.

Теорема 5. *Если $A(t), A^D(t) \in \mathbf{C}^1[\alpha, \beta]$, то*

$$\left(A^D\right)' = -A^D A' A^D + \left(E - A^D A\right) \left(A^k\right)' \left(A^D\right)^{k+1} + \left(A^D\right)^{k+1} \left(A^k\right)' \left(E - A^D A\right),$$

$$\left(A^D A\right)' = \left(E - A^D A\right) \left(A^k\right)' \left(A^D\right)^k + \left(A^D\right)^k \left(A^k\right)' \left(E - A^D A\right).$$

Наконец, приведем еще одну теорему, которая может служить основой при построении приближенных методов вычисления матрицы A^D.

Теорема 6. *Справедливо равенство*

$$A^D = \lim_{\delta \to 0} \left(\delta^2 E + A^{k+1}\right)^{-1} A^k, \tag{22}$$

где k — индекс матрицы A, причем имеет место оценка

$$\|A^D - \left(\delta^2 E + A^{k+1}\right)^{-1} A^k\| \le \delta^2 \|A^D\|^{k+2} \left(1 - \delta^2 \|A^D\|^{k+1}\right)^{-1}, \tag{23}$$

где переменная δ — вещественная; $\|A\|$ — спектральная норма матрицы.

Доказательство. Поскольку индекс матрицы A^{k+1} равен единице, с помощью известной формулы

$$(zE - A)^{-1} = \sum_{i=1}^{k} z^{-i} A^{i-1} \left(E - A^D A\right) - \left(E - zA^D\right)^{-1} A^D$$

и равенств 1^0–3^0 и

$$\left(A^D\right)^{k+1} = \left(A^{k+1}\right)^D$$

можно написать

$$\left(\delta^2 E + A^{k+1}\right)^{-1} A^k = \left(E + \delta^2 \left(A^D\right)^{k+1}\right)^{-1} A^D.$$

Поэтому

$$A^D - \left(\delta^2 E + A^{k+1}\right)^{-1} A^k = \delta^2 \left(E + \delta^2 \left(A^D\right)^{k+1}\right)^{-1} \left(A^D\right)^{k+2},$$

откуда при достаточно малых δ следует оценка (23) и, следовательно, равенство (22). Теорема доказана.

2 Регулярные АДС с переменными коэффициентами

Определение. Пара матриц $(A(t), B(t))$ в системе

$$A(t)x'(t) = B(t)x(t) + f(t), \quad t \in T = [t_0, t_1], \tag{1}$$

называется *регулярной на отрезке* T, если существует константа c такая, что при всех $t \in T$ матрица $B(t) - cA(t)$ обратима. При этом матрицы $A(t)$ и $B(t)$ должны быть, очевидно, квадратными.

Если пара матриц в системе (1) регулярна, то с помощью замены переменной

$$x(t) = e^{c(t-t_0)} y(t)$$

эта система приводится к виду

$$\tilde{A}(t)y'(t) = y(t) + \tilde{f}(t), \quad t \in T, \tag{2}$$

где

$$\tilde{A}(t) = (B(t) - cA(t))^{-1} A(t), \quad \tilde{f}(t) = e^{-c(t-t_0)} (B(t) - cA(t))^{-1} f(t).$$

Следовательно, при изучении систем (1) с регулярной парой матриц можно ограничиться рассмотрением систем вида

$$A(t)x'(t) = x(t) + f(t), \quad t \in T. \tag{3}$$

Индекс неразрешенности системы (2) во многом зависит от величины индекса матрицы $\tilde{A}(t)$. Поэтому возникает вопрос о зависимости индекса матрицы $\tilde{A}(t)$ от входящей в нее константы c. На этот счет справедлива теорема.

Теорема 1. *Структура матрицы*

$$(B(t) - cA(t))^{-1} A(t) \tag{4}$$

(т.е. количество и размеры нильпотентных и неособенных блоков в ее жордановом представлении) не зависит от чисел c, при которых матрица $B(t) - cA(t)$ обратима.

Доказательство. Пусть числа c и c_1 таковы, что матрицы $B - cA$ и $B - c_1A$ обратимы. Тогда

$$(B - c_1A)^{-1} A = ((B - cA) - (c_1 - c)\, A)^{-1} A =$$

$$= \left(E - (c_1 - c)\,(B - cA)^{-1} A\right)^{-1} (B - cA)^{-1} A \tag{5}$$

(указание зависимости от t, если это не вызывает недоразумений, здесь и в дальнейнем будем опускать). Следовательно, матрица (5) является функцией от матрицы (4), а именно

$$(B - c_1A)^{-1} A = \phi\left((B - cA)^{-1} A\right), \tag{6}$$

где

$$\phi(\lambda) = \frac{\lambda}{1 - (c_1 - c)\lambda}.$$

Так как производная

$$\phi'(\lambda) = \frac{1}{(1 - (c_1 - c)\lambda)^2},$$

очевидно, не равна нулю для тех λ, которые являются кратными корнями минимального многочлена матрицы (4), то в силу замечания ниже количество и размеры нильпотентных и неособенных блоков матриц (4) и (6) совпадают. Этим и исчерпывается доказательство теоремы.

Замечание. Если производная функции $\phi'(\lambda)$ не равна нулю для тех λ, которые являются кратными корнями минимального многочлена матрицы A, то при переходе от матрицы A к матрице $\phi(A)$ элементарные делители не "расщепляются", т.е. если матрица A имеет элементарные делители

$$(\lambda-\lambda_1)^{p_1},\ (\lambda-\lambda_2)^{p_2},\ \ldots,\ (\lambda-\lambda_n)^{p_n},$$

то матрица $\phi(A)$ имеет элементарные делители

$$(\lambda-\phi(\lambda_1))^{p_1},\ (\lambda-\phi(\lambda_2))^{p_2},\ \ldots,\ (\lambda-\phi(\lambda_n))^{p_n}$$

[5, с. 148].

Согласно теореме 1 индекс матрицы $\tilde{A}(t)$ в системе (2) не зависит от c, так что переход от общей системы (1) к системе вида (3) действительно оправдан.

3 Применение обратной матрицы Дразина для решения АДС с переменными коэффициентами

Рассмотрим следующую задачу Коши

$$A(t)x'(t)=x(t)+f(t),\ \ t\in T=[t_0,t_1], \tag{1}$$

$$x(t_0)=a. \tag{2}$$

Относительно матриц и векторов, входящих в задачу (1), (2), для простоты изложения предположим, что они обладают гладкостью достаточной для последующих рассуждений.

Кроме того, для упрощения выкладок введем в рассмотрение оператор $\mathcal{M}$, действие которого на произвольную достаточно гладкую функцию $g(t)$ определим по формуле

$$\mathcal{M}[g(t)]=A(t)g'(t),\ \ \mathcal{M}^0[g(t)]=g(t), \tag{3}$$

так что для степеней оператора $\mathcal{M}$ будем иметь

$$\mathcal{M}^2[g(t)] = AA'g' + A^2g'',$$

$$\mathcal{M}^3[g(t)] = \left(A\left(A'\right)^2 + A^2A''\right)g' + \left(2A^2A' + AA'A\right)g'' + A^3g''' \qquad (4)$$

и т.д. В частности, если матрица A постоянна, то при любом неотрицательном числе s

$$\mathcal{M}^s[g(t)] = A^s g^{(s)}(t), \qquad (5)$$

а если к тому же $A^k = O$ при некотором $k \geq 0$, т.е. матрица A нильпотентная индекса нильпотентности $\operatorname{ind} A \leq k$, то

$$\mathcal{M}^k[g] = 0. \qquad (6)$$

Следует обратить внимание на то, что в равенстве (6) можно положить $k = n$, где n — порядок матрицы A, так как $\operatorname{ind} A \leq n$.

Таким образом, если $(n \times n)$-матрица A является нильпотентной и постоянной, то имеет место равенство

$$\mathcal{M}^n[g(t)] = 0. \qquad (7)$$

Обобщение этого результата на случай переменной матрицы $A(t)$ получается, например, тогда, когда существует постоянная матрица P, при которой матрица $J(t) = PA(t)P^{-1}$ имеет верхнюю (или нижнюю) треугольную форму с нулевой главной диагональю. Это следует из формул (4), если учесть, что произведение n любых верхних (нижних) треугольных матриц n-ого порядка с нулевой главной диагональю равно нулевой матрице.

Если матрица $A(t)$ не является нильпотентной, но, тем не менее, приводится с помощью неособенной постоянной матрицы P к виду

$$\begin{pmatrix} J_0(t) & O \\ O & J_1(t) \end{pmatrix} = PA(t)P^{-1},$$

где $J_0(t)$ — треугольная матрица с нулевой главной диагональю, а матрица $J_1(t)$ при всех $t \in T$ неособенная, то равенство (7) заменяется на равенство

$$\left(E - A^D A\right)\mathcal{M}^k[g(t)] = 0, \qquad (8)$$

где k — порядок блока $J_0(t)$.

В самом деле, поскольку по теореме 3 из раздела 1

$$A^D(t) = P \begin{pmatrix} O & O \\ O & J_1^{-1}(t) \end{pmatrix} P^{-1},$$

для матрицы $A^D A$ имеем

$$A^D A = P \begin{pmatrix} O & O \\ O & E \end{pmatrix} P^{-1},$$

т.е. матрица $A^D A$ является постоянной. Кроме того, при любом целом $s > 0$ для любой квадратной матрицы A справедливо равенство $\left(E - A^D A\right) = \left(E - A^D A\right)^s$. Поэтому, например, для случая $k = 3$ с учетом формул (4) можно написать

$$\left(E - A^D A\right) \mathcal{M}^3[g] = \left(E - A^D A\right)^3 \mathcal{M}^3[g] =$$

$$= \left(B\left(B'\right)^2 + B^2 B''\right) g' + \left(2B^2 B' + BB'B\right) g'' + B^3 g''', \qquad (9)$$

где $B = \left(E - A^D A\right) A$. Однако матрица B имеет вид

$$B = P \begin{pmatrix} J_0 & O \\ O & O \end{pmatrix} P^{-1},$$

т.е. она с помощью постоянной матрицы P приводится к треугольной матрице с нулевой главной диагональю. Но тогда, как уже говорилось, правая часть равенств (9) обращается в нуль и поэтому

$$\left(E - A^D A\right) \mathcal{M}^3[g] = 0.$$

При произвольном k доказательство проводится аналогично.

Другая возможность распространения равенства (7) на случай переменных матриц осуществляется при постоянстве матрицы $\left(E - A^D A\right) A$. Дело в том, что если эта матрица постоянна, то при любом целом числе $s > 0$ выполнено равенство

$$\left(E - A^D A\right) \mathcal{M}^s[g] = \left(E - A^D A\right) A^s g^{(s)}, \qquad (10)$$

правая часть которого при $s = k = \sup\limits_{t_0 \le t \le t_1} \operatorname{ind} A(t)$ вследствие свойства 3^0 обратной матрицы Дразина (см. раздел 1) обращается в нуль, и поэтому

$$\left(E - A^D A\right) \mathcal{M}^k[g] = 0. \qquad (11)$$

Что касается равенства (10), то его можно доказать индукцией по s, а именно: при $s = 0$ и $s = 1$ равенство (10) очевидно, а предположение о его справедливости при некотором $s > 1$ приводит (с учетом постоянства матриц

$$\left(\left(E - A^D A\right) A\right)^i = \left(E - A^D A\right) A^i$$

при всех $i = 1, 2, \ldots$) к цепочке равенств

$$\left(E - A^D A\right) \mathcal{M}^{s+1}[g] = \left(E - A^D A\right) \mathcal{M}^s\left[\mathcal{M}[g]\right] =$$

$$= \left(E - A^D A\right) A^s \left(\mathcal{M}[g]\right)^{(s)} = \left(\left(E - A^D A\right) A^s \left(\mathcal{M}[g]\right)\right)^{(s)} =$$

$$= \left(\left(E - A^D A\right) A^{s+1} g'\right)^{(s)} = \left(E - A^D A\right) A^{s+1} g^{(s+1)},$$

откуда следует, что

$$\left(E - A^D A\right) \mathcal{M}^{s+1}[g] = \left(E - A^D A\right) A^{s+1} g^{(s+1)},$$

чем и исчерпывается доказательство равенства (10). Впрочем доказательство можно провести и непосредственно, используя формулы (4).

Это свойство оператора $\mathcal{M}$ позволяет сформулировать следующие теоремы.

Теорема 2. *Пусть для матрицы $A(t)$ из системы (1) при некоторой постоянной неособенной матрице P имеет место представление*

$$A(t) = P \begin{pmatrix} J_0(t) & O \\ O & J_1(t) \end{pmatrix} P^{-1}, \tag{12}$$

где $J_0(t)$ — нильпотентная треугольная матрица порядка k, а матрица $J_1(t)$ является неособенной при всех $t \in T$. Тогда если решение задачи (1), (2) существует, то оно единственно и представимо в виде

$$x(t) = \eta(t) - \left(E - A^D(t)A(t)\right) \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)], \tag{13}$$

где $\eta(t)$ — решение разрешенной относительно производной задачи Коши

$$\eta'(t) = A^D(t)\eta(t) + A^D(t)f(t), \tag{14}$$

$$\eta(t_0) = A^D(t_0)A(t_0)a. \tag{15}$$

Доказательство. Пусть задача (1), (2) разрешима и $x(t)$ — ее решение. Используя оператор $\mathcal{M}$, запишем равенство (1) в виде

$$x(t) = \mathcal{M}[x(t)] - f(t). \tag{16}$$

Из равенства (16), очевидно, следует, что при любом $s \geq 0$

$$\mathcal{M}^s[x(t)] = \mathcal{M}^{s+1}[x(t)] - \mathcal{M}^s[f(t)],$$

и поэтому $x(t)$ удовлетворяет системе

$$\begin{gathered} x(t) = \mathcal{M}[x(t)] - f(t), \\ \mathcal{M}[x(t)] = \mathcal{M}^2[x(t)] - \mathcal{M}[f(t)], \\ \mathcal{M}^2[x(t)] = \mathcal{M}^3[x(t)] - \mathcal{M}^2[f(t)], \\ \cdots\cdots\cdots\cdots \\ \mathcal{M}^{k-1}[x(t)] = \mathcal{M}^k[x(t)] - \mathcal{M}^{k-1}[f(t)]. \end{gathered} \tag{17}$$

Сложив равенства системы (17), получим

$$x(t) = \mathcal{M}^k[x(t)] - \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)], \tag{18}$$

что после умножения на матрицу $E - A^D A$ (с учетом свойства (8) оператора $\mathcal{M}$) дает

$$\Big(E - A^D(t)A(t)\Big)\, x(t) = -\Big(E - A^D(t)A(t)\Big) \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)],$$

и, следовательно,

$$x(t) = \eta(t) - \Big(E - A^D(t)A(t)\Big) \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)]$$

(ср. с (13)), где $\eta(t) = A^D(t)A(t)x(t)$. Осталось показать, что $\eta(t)$ удовлетворяет системе (14), (15). Для этого умножим равенство (1) на матрицу $A^D(t)$ слева. В результате будем иметь

$$A^D(t)A(t)x'(t) = A^D(t)x(t) + A^D(t)f(t),$$

что вследствие равенства $A^D = A^D A^D A$, постоянства матрицы $A^D A$ и обозначения $\eta(t) = A^D(t)A(t)x(t)$ легко записать в виде (14). Что касается условия (15), то справедливость его для вектора $\eta(t)$ очевидна. Единственность решения задачи (1), (2) следует теперь из единствеености решения задачи (14), (15). Теорема доказана.

Чтобы найти условия, при которых существует решение задачи (1), (2), вычислим невязки

$$r(t) = A(t)x'(t) - x(t) - f(t), \quad s = x(t_0) - a,$$

но прежде заметим, что решение задачи (14), (15) удовлетворяет равенству

$$\left(E - A^D(t)A(t)\right)\eta'(t) = 0.$$

В самом деле, умножив уравнение (14) на матрицу $E - A^D A$, из постоянства матрицы $E - A^D A$ и равенства $\left(E - A^D A\right)A^D = O$ получим

$$\left(\left(E - A^D A\right)\eta\right)' = O,$$

откуда следует, что при всех $t \in T$

$$\left(E - A^D(t)A(t)\right)\eta(t) = 0, \tag{19}$$

поскольку при $t = t_0$ в силу условия (15) равенство (19) справедливо.

Учитывая этот факт, постоянство матрицы $A^D A$ и вводя обозначение

$$\xi(t) = \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)], \tag{20}$$

приходим к следующей цепочке

$$r = Ax' - x - f = A\eta' - \eta - f - \left(E - A^D A\right)A\xi' + \left(E - A^D A\right)\xi =$$

$$= A^D A\eta - \eta - \left(E - A^D A\right)f - \left(E - A^D A\right)A\xi' + \left(E - A^D A\right)\xi =$$

$$= -\left(E - A^D A\right)A\xi' - \left(E - A^D A\right)f + \left(E - A^D A\right)\xi.$$

Однако

$$A\xi' = \sum_{s=0}^{k-1} \mathcal{M}^{s+1}[f(t)] = \mathcal{M}^k[f(t)] + \sum_{s=0}^{k-1} \mathcal{M}^s[f(t)] + \xi(t) - f(t),$$

поэтому, как легко проверить

$$r = -\left(E - A^D A\right)\mathcal{M}^k[f(t)] \tag{21}$$

и, следовательно, в силу свойства (8) $r = 0$.

Далее,

$$s = x(t_0) - a = \eta(t_0) - \left(E - A^D A\right)\xi(t_0) - a = -\left(E - A^D A\right)(a + \xi(t_0)),$$

т.е.

$$s = \left(E - A^D A\right)\left(-a - \sum_{s=0}^{k-1}\mathcal{M}^s[f(t)]\right)\Bigg|_{t=t_0}. \tag{22}$$

Таким образом, в рассматриваемом случае единственным условием существования решения является справедливость равенства

$$s = 0 \tag{23}$$

(см. (22)). При этом очевидно, что если даже условие (23) не выполнено, формула (13) по заданному a дает "ближайшее"начальное данное $x(t_0)$, при котором задача (1), (2) разрешима.

Теперь заметим, что при вычислении невязок (21), (22) представление (12) не понадобилось: достаточно было постоянства матрицы $A^D A$. Поэтому справедлива также следующая теорема.

Теорема 3. *Если матрица $A^D(t)A(t)$ постоянна и при некотором целом числе $k \geq 0$ выполнено равенство*

$$\left(E - A^D(t)A(t)\right)\mathcal{M}^k[f(t)] = 0,$$

то формула (13) дает решение задачи (1), (2) с начальным данным $x(t_0)$, вычисленным по заданному a по той же формуле (13), причем если $s = 0$ (см. (22), (23)), то $x(t_0) = a$.

Теперь посмотрим, что получается, когда постоянной является матрица $\left(E - A^D(t)A(t)\right)A(t)$.

Теорема 4. *Если матрица $\left(E - A^D(t)A(t)\right)A(t)$ постоянна и* $\operatorname{ind} A(t) \leq k$, *то решение задачи (1), (2), если оно существует, единственно и представляется в виде*

$$x(t) = \eta(t) - \left(E - A^D(t)A(t)\right)\xi(t), \tag{37}$$

где

$$\xi(t) = \sum_{s=0}^{k-1} A^s(t) f^{(s)}(t), \tag{38}$$

а вектор $\eta(t)$ *есть решение задачи Коши*

$$\eta'(t) = \left(A^D(t) + \left(E - A^D(t)A(t)\right) A'(t)A^D(t)\right) \eta(t) +$$

$$+ A^D(t)A'(t) \left(E - A^D(t)A(t)\right) \xi(t) + A^D(t)f(t), \tag{39}$$

$$\eta(t_0) = A^D(t_0)A(t_0)a. \tag{40}$$

Доказательство. Учитывая равенства (10), (11), (38), так же, как при доказательстве теоремы 2, легко показать, что

$$\left(E - A^D A\right) x = - \left(E - A^D A\right) \xi. \tag{41}$$

Поэтому решение задачи (1), (2) представляется в виде (37) при значении $\eta(t) = A^D(t)A(t)x(t)$, и остается убедиться , что $\eta(t) = A^D(t)A(t)x(t)$ является решением задачи (39), (40).

Умножив уравнение (1) на матрицу $A^D(t)$, получим

$$\left(A^D(t)A(t)x(t)\right)' = A^D(t) \left(A^D(t)A(t)x(t)\right) +$$

$$+ \left(A^D(t)A(t)\right)' x(t) + A^D(t)f(t), \tag{42}$$

где

$$\left(A^D A\right)' = \left(E - A^D A\right) A' A^D + A^D A' \left(E - A^D A\right) \tag{43}$$

в силу известной формулы для производной [6, с. 43]

$$\left(A^D A\right)' = \left(A^D A\right)' A^D A + A^D A \left(A^D A\right)'.$$

Подстановка (43) в (42) с учетом (41) приводит к равенству (39). Условие (40) для ветора $\eta(t)$, очевидно, выполняется. Единственность решения задачи (1), (2) следует из единственности решения задачи (39), (40). Теорема доказана.

Вопрос о существовании решения в этом случае исследуется также как при доказательстве теоремы 3 путем вычисления соответствующих невязок.

К сожалению, далеко не все переменные матрицы $A(t)$ обладают свойствами, допускающими применение теорем 2–4. Рассмотрим, например, матрицу

$$A(t) = \begin{pmatrix} t & 1 \\ -t^2 & -t \end{pmatrix}.$$

Для нее $A^2(t) = O$, $A(t) \not\equiv O$, поэтому $\operatorname{ind} A(t) = 2$. Отсюда в силу определения обратной матрицы Дразина $A^D(t) \equiv O$. Значит, матрица $\left(E - A^D A\right) A = A$ не является постоянной и теорема 4 неприменима.

Список литературы

[1] *Воеводин В.В.* Линейная алгебра. — М.: Наука, 1974.

[2] *Ильин В.А., Позняк Э.Г.* Линейная алгебра. — М.: Наука, 1974.

[3] *Гайшун И.В.* Введение в теорию линейных нестационарных систем. — Минск: Изд-во Ин-та мат. НАН Беларуси, 1999.

[4] *Петровский И.Г.* Лекции по теории обыкновенных дифференциальных уравнений. — М.–Л.: Гос. изд-во технико-теоретической литературы, 1949.

[5] *Гантмахер Ф.Р.* Теория матриц. — М.: Наука, 1988.

[6] *Бояринцев Ю.Е.* Регулярные и сингулярные системы линейных обыкновенных дифференциальных уравнений. — Новосибирск: Наука, 1980.

[7] *Griepentrog E., Maerz R.* Differential-algebraic equations and their numerical treatment. — Leipzig: BSB B. G. Teubner Verlag gesellschaft, 1986.

[8] *Бояринцев Ю.Е.* Методы решения вырожденных систем обыкновенных дифференциальных уравнений. — Новосибирск: Наука, 1988.

[9] *Dai L.* Singular control systems. Lecture Notes in Control and Information Sciences. V. 118. — Berlin: Springer-Verlag, 1989.

[10] *Чистяков В.Ф.* Алгебро-дифференциальные операторы с конечномерным ядром. — Новосибирск: "Наука", 1996.

[11] *Brenan K.E., Campbell S.L., Petzold L.D.* Numerical solution of initial–value problems in differential–algebraic equations (classics in applied mathematics; 14). — Philadelphia: SIAM, 1996.

[12] *Rabier P.J., Rheinboldt W.C.* Theoretical and numerical analysis of differential-algebraic equations. Handbook of Numerical Analysis. V. VIII. — Amsterdam, 2002.

[13] *Чистяков В.Ф., Щеглова А.А.* Избранные главы теории алгебро–дифференциальных систем. — Новосибирск: Наука, 2003.