E. L. Rabkin, Full investigation of the matrix equation $AX + XB = C$ and specifically of the equation $AX - XA = C$,
Algebra i Analiz, 2014, Volume 26, Issue 1, 165–184

Алгебра и анализ
Том 26 (2014), №1

ПОЛНОЕ ИССЛЕДОВАНИЕ МАТРИЧНОГО УРАВНЕНИЯ $AX + XB = C$ И, В ЧАСТНОСТИ, УРАВНЕНИЯ $AX - XA = C$

© Е. Л. РАБКИН

В статье полностью исследуются и решаются, если решение существует, матричные уравнения $AX - XA = C$ и $AX + XB = C$ (уравнение Ляпунова). Как частные случаи, получаются точные выражения для резольвенты уравнения $(E - A)X = 0$ для конечномерного оператора A и полное исследование уравнения Фредгольма второго рода в конечномерном случае. Выясняется, какой должна быть матрица с неотрицательными элементами C, чтобы она могла служить коммутатором между данной матрицей с неотрицательными элементами A и некоторой другой матрицей с неотрицательными элементами X.

Введение

1. Фробениус [1] полностью решил вопрос: какой вид должна иметь квадратная матрица X, чтобы она коммутировала с данной квадратной матрицей A порядка n? Иными словами, он нашел все решения уравнения

$$AX - XA = 0. \tag{1}$$

(На самом деле он решил уравнение

$$AX - XB = 0, \tag{2}$$

где B — любая другая квадратная матрица не обязательно того же порядка, а решение уравнения (1) получил как частный случай решения уравнения (2).) Естественно, возник вопрос: какова должна быть матрица C того же порядка, чтобы матричное уравнение

$$AX - XA = C \tag{3}$$

имело решение? Иными словами, какие матрицы могут быть коммутаторами? Уравнение (3) рассматривалось во многих работах. Последняя из них, известная автору, [4], датирована 2008 г., но, как отмечают сами авторы в своем реферате, ,,получены явные формулы для общего решения уравнения $AX - XA = C$ в том случае, когда A является матрицей

Ключевые слова: матрица, коммутатор, уравнение Ляпунова.

простой структуры, а также для общего самосопряженного решения этого уравнения в случае нормальной матрицы A. Построенные формулы представляют собой полиномы относительно A и C с коэффициентами, выражающимися через коэффициенты минимального многочлена матрицы“. (Таким образом, в [4] изучается лишь случай, когда жорданова форма матрицы A вырождается в диагональную матрицу.) Здесь дается полный ответ на этот классический вопрос для любой матрицы A, опирающийся на идеи Фробениуса (разумеется, автор не читал работу [1], а основывался на превосходном ее изложении в [2, с. 178–186], и в данной работе полностью опирается на это изложение).

Ответ таков: матрица C является коммутатором для матрицы A и некоторой матрицы X для тех и только тех матриц C, для которых матрица $C_0 = T^{-1}CT$ удовлетворяет условиям (10), где T — матрица, приводящая A к жордановой форме, т.е. такая, что $A_0 = T^{-1}AT$ — жорданова матрица (иными словами, оператор, задаваемый матрицей A, рассматривается в базисе из собственных и присоединенных векторов этого оператора).

В частности, если матрица A — диагональная и все ее собственные числа различны, то единственное условие, при котором матрица C является коммутатором A с некоторой матрицей X, заключается в том, что все элементы главной диагонали матрицы C должны равняться нулю, и тогда все элементы матрицы X определяются однозначно по формуле (17*), за исключением элементов главной диагонали, которые могут быть любыми числами (ср. с [4]).

2. В теории Ляпунова (устойчивости решений дифференциальных уравнений) исключительную роль играет более общее матричное уравнение

$$AX + XB = C, \tag{4}$$

которое иногда называют „уравнением Ляпунова“ (очевидно, что (4) превращается в (3) при $B = -A$). В монографии [3] доказывается, что уравнение (4) имеет единственное решение тогда и только тогда, когда собственные числа λ_α матрицы A и собственные числа μ_β матрицы B таковы, что $\lambda_\alpha + \mu_\beta \neq 0$ при всех возможных α и β. В этой монографии автор исходит из того, что решением матричного дифференциального уравнения

$$\frac{dX}{dt} = AX + XB, \quad X(0) = C, \tag{5}$$

является матрица $X(t) = e^{At}Ce^{Bt}$, что проверяется непосредственно подстановкой. При интегрировании уравнения (5) в пределах от 0 до ∞

получается решение уравнения (32) в форме

$$X = -\int_0^\infty e^{At} C e^{Bt}\, dt, \tag{6}$$

причем доказывается, что сходимость такого интеграла обеспечивается указанным выше условием (для доказательства приходится применять кронекеровское произведение матриц и другую нетривиальную математическую технику). Здесь будет показано (см. теорему 2), что идеи Фробениуса позволяют исследовать уравнение (4) во всех случаях, причем элементарными методами. По существу здесь выясняется, что в случае единственности решения уравнения (4) интеграл (6) является „берущимся“ и вычисляется по формуле

$$X = -U^{-1}\widetilde{X}V^{-1},$$

где U и V — матрицы, приводящие матрицы A и B к жордановой форме, а $\widetilde{X}$ — блочная матрица, состоящая из блоков $X_{\alpha\beta}$, соответствующих пересечениям жордановых ящиков матриц A и B и вычисляемых по формуле (40):

$$\widetilde{X}_{\alpha\beta} = \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha + \mu_\beta)^{i+1}} \sum_{k=0}^{i} C_i^k H_\alpha^k \widetilde{C}_{\alpha\beta} H_\beta^{i-k}$$

($\widetilde{C}_{\alpha\beta}$ — соответствующий блок матрицы $\widetilde{C} = UCV$. Кроме того, здесь выясняются необходимые и достаточные условия, при которых уравнение (4) не имеет решений, и когда решение не единственно (в этом случае их бесконечно много, и все они находятся по формулам (24) и (25)).

Интересно, что все рассуждения, приведенные ниже при решении уравнения (32) для матриц одинакового порядка, справедливы и в случае, когда матрица A имеет порядок n, B — порядок m, а C и X — $n \times m$.

3. Будем называть матрицы, все элементы которых неотрицательны (положительны), „неотрицательными (положительными) матрицами“ и обозначать эти факты символом $A \geqslant 0$ ($A > 0$). После публикации нашей совместной статьи [5], где доказывается, что если коммутатор двух неотрицательных матриц также неотрицателен, то он (коммутатор) обязательно нильпотентен, автора заинтересовал вопрос, который и побудил его провести данное исследование: если задана неотрицательная квадратная матрица A, то какой должна быть матрица C, чтобы она служила коммутатором A с некоторой неотрицательной же матрицей X. Для неразложимой матрицы A ответ такой: C — любая матрица, удовлетворяющая условиям (10). Действительно, если X_0 — какое-то решение уравнения (3), то и любая матрица вида $X = X_0 + f(A)$, где f — любая функция, тоже является решением этого уравнения (таким образом, все матрицы вида $X = X_0 + f(A)$ тоже входят в множество решений, задаваемых формулами (24) и (25)). В [2, с. 322] доказывается, что если A — неразложимая неотрицательная матрица, то матрица $(E+A)^{n-1}$, где n — порядок матрицы A, является матрицей со всеми строго положительными элементами. Поэтому матрица $X = X_0 + k(E + A)^{n-1}$ является решением (3) и при достаточно большом $k > 0$ тоже является матрицей с положительными элементами.

Для разложимых матриц вопрос остается открытым.

Следствие. *Для того, чтобы неотрицательная матрица C являлась коммутатором для данной неотрицательной неразложимой матрицы A и некоторой неотрицательной матрицы X, необходимо и достаточно, чтобы в жордановом базисе матрицы A матрица C являлась верхнетреугольной, т.е. чтобы в этом базисе выполнялись все равенства $c_{ij} = 0$ при $i \geqslant j$ (при этом C автоматически оказывается нильпотентной).*

4. Если матрицы A, B, C вещественны, а решение уравнения (3) или (4) комплексное: $X = X_1 + X_2 i$, то, отделяя вещественные и мнимые части, убедимся, что X_1 — вещественное решение рассматриваемого уравнения, а для уравнения (3) X_2 — любая матрица, коммутирующая с A. Таким образом, если матрицы A, B, C вещественны, то уравнения (3) или (4) обязательно имеют чисто вещественные решения. В частности, если решение единственно, то оно обязательно вещественно (впрочем, для уравнения (4) это следует и из формулы (6)).

5. Разумеется, уравнение (3) есть частный случай уравнения (4), но они здесь рассматриваются по отдельности, так как результаты исследования записываются в разных формах, и эти уравнения имеют разные сферы приложений.

6. При $B = -E$ уравнение (4) превращается в уравнение

$$(E - A)X = -C. \tag{4*}$$

Стандартный способ его решения — метод последовательных приближений при помощи разложения резольвенты в степенной ряд

$$E + A + A^2 + \dots + A^n + \dots$$

В [6, с. 463, теорема 3] указано, что если спектр оператора (в нашем случае все собственные числа матрицы A) лежит в круге $|\lambda| < 1$, то метод последовательных приближений сходится к искомому решению для любой правой части (в нашем случае для любой матрицы C) при любом начальном приближении X_0. Формула (40) при $\mu = 1$ и при замене c_{ij} на $-c_{ij}$

дает полное решение уравнения (4^*) для любой матрицы A (но в отличие от теоремы 3 в [6] только в конечномерном пространстве). Таким образом, (40) является результатом вычисления резольвенты и дает сумму упомянутого выше ряда.

7. Уравнение (4) при $B = -E$, т.е. (4^*), можно рассматривать как уравнение Фредгольма второго рода в пространстве матриц порядка n, так как конечномерный оператор заведомо является вполне непрерывным. Для выяснения альтернативы Фредгольма требуется описать пространство линейных функционалов в пространстве матриц. Проще всего это сделать при помощи введения скалярного произведения: если $X = (x_{ij})$, $Y = (y_{ij})$, то скалярным произведением этих матриц будем считать число $< X, Y >= \sum_{i,j=1}^n x_{ij}\bar{y}_{ij}$ и, следовательно, нормой — число $\ll X \gg = \sqrt{\sum_{i,j=1}^n x_{ij}^2 \bar{y}_{ij}^2}$ (т.е. в пространстве конечномерных операторов введем норму Гильберта–Шмидта). При этом пространство матриц превращается в нормированное кольцо. Правда, в этом кольце $\ll E \gg = \sqrt{n}$, но, как указано в [7, с. 15, теорема 1], для каждого нормированного кольца можно найти топологически и алгебраически изоморфное ему кольцо, в котором норма единичного элемента равна единице, а именно можно рассматривать каждый элемент как оператор умножения: $A_X Y = XY$, а его нормой считать норму этого оператора. При этом оказывается, что новая норма матрицы совпадает со старой, а сопряженный оператор — тот же, что и раньше. Действительно, для матриц $A = (a_{ij})$ и $A^*(\bar{a}_{ij})$ и любых матриц X и Y обе части равенства $< AX, Y >=< X, A^*Y >$ равны одному и тому же числу $\sum\limits_{i,j,k}^{n} a_{ik}x_{kj}\bar{y}_{ij}$. При этом сопряженный оператор в смысле введенного скалярного произведения матриц совпадает с классическим сопряженным, т.е. в обоих случаях справедливо равенство $X^* = \overline{X}^T$. Собственными элементами оператора A являются матрицы $X \neq 0$, удовлетворяющие равенству $AX = \lambda X$. Если λ — собственное число для A в обычном смысле, которому соответствует собственный элемент X, то это число является собственным и в матричном смысле, причем соответствующим собственным элементом является матрица, каждый столбец которой равен столбцу X.

Введенное скалярное произведение делает линейное пространство матриц гильбертовым пространством, в котором линейные функционалы определяются скалярным произведением на фиксированные матрицы. Однородным уравнением Фредгольма, сопряженным к уравнению (4^*), является уравнение

$$(E - A^*)F = 0, \tag{4**}$$

где F — матрица, определяющая неизвестный функционал.

В рассматриваемом случае A есть жордановый блок, соответствующий собственному числу 1 некоторого порядка p. Поэтому собственный и присоединенный элементы оператора $E-A^*$, т.е. функционалы, определяемые уравнением $(E-A^*)^kF=0$, $k=1,2,\dots,p$, суть E, H, $H^2,\dots,H^{p-1}$, где H матрица порядка p, определяемая равенством

$$H=\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Теорема об альтернативе Фредгольма утверждает, что либо уравнение (4^{**}) имеет единственное (нулевое) решение, и тогда уравнение (4^*) тоже имеет единственное решение, либо (4^{**}) имеет конечное число линейно независимых нетривиальных решений, и тогда (4^*) имеет столько же решений только при условии, что его правая часть ортогональна всем нетривиальным решениям уравнения (4^{**}). В рассматриваемом случае это значит, что либо 1 не является собственным числом матрицы A, и тогда решение уравнения (4^*) единственно (что следует и из теоремы 2, причем в этой теореме указывается и вид этого решения — оно дается формулой (40)), либо является, и тогда решения уравнения (4^*) существуют только при условии, что его правая часть ортогональна решениям уравнения (4^{**}), т.е. матрицам E, H, $H^2,\dots,H^{p-1}$, что также следует из теоремы 2, причем условия ортогональности совпадают с условиями (10) (или (44), что в данном случае то же самое); в этой теореме, кроме того, указывается, как найти все эти решения. Таким образом, теорема 2 является уточнением теоремы об альтернативе Фредгольма для пространства матриц.

Вопрос о возможности распространения теоремы 2 на бесконечномерные пространства остается открытым.

8. В [6, с. 466, теорема 5] указывается вид разложения резольвенты B_λ вполне непрерывного оператора U в окрестности его собственного числа λ_0:

$$B_\lambda=\sum_{n=-r}^{\infty}U_n(\lambda-\lambda_0)^n,$$

где U_{-r}, $U_{-r+1,\dots,U_{-1}}$ — некоторые конечномерные операторы, причем $U_{-r}\neq 0$, r — кратность собственного числа λ_0. Формула (44) является детализацией этой теоремы для случая матриц.

Вопрос о возможности обобщения этой формулы на бесконечномерный случай остается открытым.

§1. Уравнение $AX - XA = C$

Для простоты записей считаем, что матрицы A и C сразу заданы в жордановом базисе матрицы A, и, следовательно, считаем матрицу A разбитой на жордановы блоки

$$A = \begin{pmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ 0 & 0 & J_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & J_m \end{pmatrix}. \tag{7}$$

Каждому блоку J_k соответствует некоторое собственное число λ_k матрицы A (собственные числа могут быть одинаковыми для разных блоков). Матрицу C считаем разбитой на такие же блоки:

$$C = \begin{pmatrix} C_{11} & C_{12} & C_{13} & \dots & C_{1m} \\ C_{21} & C_{22} & C_{23} & \dots & C_{2m} \\ C_{31} & C_{32} & C_{33} & \dots & C_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ C_{m1} & C_{m2} & C_{m3} & \dots & C_{mm} \end{pmatrix}. \tag{8}$$

Искомую матрицу X считаем также разбитой на аналогичные блоки:

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \dots & X_{2m} \\ X_{31} & X_{32} & X_{33} & \dots & X_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ X_{m1} & X_{m2} & X_{m3} & \dots & X_{mm} \end{pmatrix}. \tag{9}$$

Теорема 1. *Тогда для того, чтобы уравнение* (3) *имело решение, необходимо и достаточно, чтобы в каждом блоке*

$$C_{\alpha\beta} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1q} \\ C_{21} & c_{22} & \dots & c_{2q} \\ \dots & \dots & \dots & \dots \\ C_{p1} & c_{p2} & \dots & c_{pq} \end{pmatrix}$$

размером $p \times q$, соответствующем одному собственному числу или двум равным собственным числам матрицы A, сумма элементов, стоящих на *каждой поддиагонали, равнялась нулю*:

$$\sum_{k=0}^{s-i} c_{i+k,k+1} = 0 \quad (i = 1, 2, \ldots, s, \ \ s = \min(p, q)). \tag{10}$$

(*В частности, если блок квадратный* $p = q = s$, *то* $c_{p1} = 0$, *а след этого блока должен равняться нулю, это условие получается при* $i = 1$; *таким образом, одно из условий, необходимых для разрешимости уравнения* (3), *заключается в том, что должно выполняться равенство* $\operatorname{tr} C = 0$, *что очевидно и непосредственно.*)

На блоки матрицы C, соответствующие различным собственным числам матрицы A, никакие ограничения не накладываются.

Доказательство. При перемножении блочных матриц (7) и (9) уравнение (3) принимает вид

$$AX - XA = \begin{pmatrix} J_{11}X_{11} - X_{11}J_{11} & J_{11}X_{12} - X_{12}J_{22} & \ldots & J_{11}X_{1m} - X_{1m}J_{mm} \\ J_{22}X_{21} - X_{21}J_{11} & J_{22}X_{22} - X_{22}J_{22} & \ldots & J_{22}X_{2m} - X_{2m}J_{mm} \\ \ldots & \ldots & \ldots & \ldots \\ J_{mm}X_{m1} - X_{m1}J_{11} & J_{mm}X_{m2} - X_{m2}J_{22} & \ldots & J_{mm}X_{mm} - X_{mm}J_{mm} \end{pmatrix} = C. \tag{11}$$

Каждый жорданов блок $J_{\alpha\alpha}$ имеет вид $J_{\alpha\alpha} = \lambda_\alpha I + H_\alpha$, где I — единичная матрица соответствующего размера, а H_α — матрица того же размера, все элементы которой равны 0, кроме элементов, стоящих непосредственно над главной диагональю, которые равны 1:

$$H_\alpha = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}. \tag{12}$$

Поэтому соответствующий блок равенства (11), стоящий в блочной строке с номером α и блочном столбце с номером β, имеет вид $(\lambda_\alpha I + H_\alpha)X_{\alpha\beta} - X_{\alpha\beta}(\lambda_\beta I + H_\beta) = C_{\alpha\beta}$, т.е.

$$(\lambda_\alpha - \lambda_\beta)X_{\alpha\beta} = X_{\alpha\beta}H_\beta - H_\alpha X_{\alpha\beta} + C_{\alpha\beta}. \tag{13}$$

Таким образом, задача решения матричного уравнения (3) равносильна решению уравнений (13) при всех α и β. При этом неизвестные элементы каждого блока свои, они не пересекаются, поэтому каждое уравнение (13) можно решать независимо от остальных.

Для решения этих уравнений, следуя Фробениусу [1], произведем „итерации“ равенства (13), т.е. будем умножать это равенство на $(\lambda_\alpha - \lambda_\beta)$

и каждый раз заменять $(\lambda_\alpha - \lambda_\beta)X_{\alpha\beta}$ на его выражение по формуле (13). В результате первой итерации получим

$$\begin{aligned}(\lambda_\alpha - \lambda_\beta)^2 X_{\alpha\beta} &= (\lambda_\alpha - \lambda_\beta)X_{\alpha\beta}H_\beta - H_\alpha(\lambda_\alpha - \lambda_\beta)X_{\alpha\beta} + (\lambda_\alpha - \lambda_\beta)C_{\alpha\beta} \\ &= X_{\alpha\beta}H_\beta^2 - 2H_\alpha X_{\alpha\beta}H_\beta + H_\alpha^2 X_{\alpha\beta} + C_{\alpha\beta}H_\beta - H_\alpha C_{\alpha\beta} + (\lambda_\alpha - \lambda_\beta)C_{\alpha\beta}.\end{aligned}$$

Аналогично проводятся 2-я, 3-я и следующие итерации. В результате итерации с номером r получаем

$$\begin{aligned}(\lambda_\alpha - \lambda_\beta)^{r+1}X_{\alpha\beta} = \sum_{i=0}^{r}(\lambda_\alpha - \lambda_\beta)^{r-i}\sum_{k=0}^{i}(-1)^k H_\alpha^k C_i^k C_{\alpha\beta}H_\beta^{i-k} \\ + \sum_{i=0}^{r+1} C_{r+1}^i (-1)^i H_\alpha^i X_{\alpha\beta} H_\beta^{r+1-i}.\end{aligned} \tag{14}$$

Эта формула легко проверяется индукцией (индукция проведена для более общего уравнения (26) при доказательстве теоремы 2, см. ниже).

Заметим, что матрица H нильпотентна: если ее порядок равен p, то $H^p = 0$. Поэтому если размеры блока $X_{\alpha\beta}$ равны $p \times q$, $p = p(\alpha,\beta)$, $q = q(\alpha,\beta)$, то при $r > i > p + q - 2$ либо $i > p - 1$, либо $r + 1 - i > q - 1$, и, следовательно, либо $H_\alpha^i = 0$, либо $H_\beta^{r+1-i} = 0$. Значит, при $r > p + q - 2$ все слагаемые в последней сумме формулы (14) равны 0. Таким образом, при $r > p + q - 2$ формула (14) принимает вид

$$(\lambda_\alpha - \lambda_\beta)^{r+1}X_{\alpha\beta} = \sum_{i=0}^{r}(\lambda_\alpha - \lambda_\beta)^{r-i}\sum_{k=0}^{i}(-1)^k C_i^k H_\alpha^k C_{\alpha\beta}H_\beta^{i-k}. \tag{15}$$

1) Если $\lambda_\alpha \neq \lambda_\beta$, то, деля обе части равенства (15) на $(\lambda_\alpha - \lambda_\beta)^{r+1}$, получаем равенство

$$X_{\alpha\beta} = \sum_{i=0}^{r}\frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}}\sum_{k=0}^{i}(-1)^k C_i^k H_\alpha^k C_{\alpha\beta}H_\beta^{i-k}. \tag{16}$$

Формально $X_{\alpha\beta}$ в формуле (16) зависит от r. Но мы уже выяснили, что при $o > p+q-2$ все слагаемые внутренней суммы в (16) равны 0, и поэтому в (16) можно считать, что $r = p + q - 1$, т.е. справедливо равенство

$$X_{\alpha\beta} = \sum_{i=0}^{p+q-1}\frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}}\sum_{k=0}^{i}(-1)^k C_i^k H_\alpha^k C_{\alpha\beta}H_\beta^{i-k}. \tag{17}$$

В частности, если A диагональная и все ее собственные числа λ_i различны, то (17) превращается в (ср. с [4])

$$x_{ij} = \frac{c_{ij}}{\lambda_i - \lambda_j}. \tag{17*}$$

Итак, если решение уравнения (13) при $\lambda_\alpha \neq \lambda_\beta$ существует, то оно единственно и определяется формулой (17). Остается проверить, что (17) действительно решает уравнение (13). Для этого подставим (17) в (13). Правая часть этого уравнения примет вид

$$
\begin{aligned}
&X_{\alpha\beta}H_\beta - H_\alpha X_{\alpha\beta} + C_{\alpha\beta} = \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta^{i+1}} \sum_{k=0}^{i} (-1)^k C_i^k H_\alpha^k C_{\alpha\beta} H_\beta^{i-k+1} \\
&- \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} \sum_{k=0}^{i} (-1)^k C_i^k H_\alpha^{k+1} C_{\alpha\beta} H_\beta^{i-k} + C_{\alpha\beta} \\
&= \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} - \sum_{k=-1}^{i-1} (-1)^{k+1} C_i^{k+1} H_\alpha^{k+1} C_{\alpha\beta} H_\beta^{i-k} \\
&+ \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} \left(\sum_{k=0}^{i-1} (-1)^{k+1} C_i^k H_\alpha^{k+1} C_{\alpha\beta} H_\beta^{i-k} + (-1)^{i+1} H_\alpha^{i+1} C_{\alpha\beta} \right) \\
&\quad + C_{\alpha\beta} \\
&= \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} \Biggl(C_{\alpha\beta} H_\beta^{i+1} \\
&\quad + \sum_{k=0}^{i-1} (-1)^{k+1} (C_i^k + C_i^{k+1}) H_\alpha^{k+1} C_{\alpha\beta} H_\beta^{i-k} + (-1)^{i+1} H_\alpha^{i+1} C_{\alpha\beta} \Biggr) + C_{\alpha\beta} \\
&= \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} \Biggl(C_{\alpha\beta} H_\beta^{i+1} \\
&\quad + \sum_{k=1}^{i} (-1)^k C_{i+1}^k H_\alpha^k C_{\alpha\beta} H_\beta^{i-k+1} + (-1)^{i+1} H_\alpha^{i+1} C_{\alpha\beta} \Biggr) + C_{\alpha\beta} \\
&= \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i+1}} \left(\sum_{k=0}^{i+1} (-1)^k C_{i+1}^k H_\alpha^k C_{\alpha\beta} H_\beta^{i-k+1} \right) + C_{\alpha\beta} \\
&= \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha - \lambda_\beta)^{i}} \sum_{k=0}^{i} (-1)^k C_i^k H_\alpha^k C_{\alpha\beta} H_\beta^{i-k} = (\lambda_\alpha - \lambda_\beta) X_{\alpha\beta},
\end{aligned}
\tag{18}
$$

и то, что (17) удовлетворяет (13), доказано. (В последней сумме в (18) индекс i должен меняться от 1 до $p+q$, но при $i = 0$ получается слагаемое

$C_{\alpha\beta}$, имеющееся в скобках, а при $i = p + q$, как уже было отмечено, все слагаемые в сумме по k равны 0).

Итак, все блоки матрицы X, у которых $\lambda_\alpha \neq \lambda_\beta$, определяются однозначно равенством (17) при любых блоках $C_{\alpha\beta}$ матрицы C.

Замечание. Если матрицы A и C вещественны, то собственные числа матрицы A могут быть комплексными, и поэтому формула (17) формально дает комплексное решение рассматриваемой системы уравнний. Но, как уже было отмечено в п. 4 введения, если в рассматриваемом случае X является решением уравнения (3), то и $\operatorname{Re} X$ тоже является его решением. Так как решение единственно, то (17) в этом случае обязательно дает вещественное решение.

2) Пусть теперь $\lambda_\alpha = \lambda_\beta$. Тогда исходное уравнение (13) принимает вид

$$H_\alpha X_{\alpha\beta} - X_{\alpha\beta} H_\beta = C_{\alpha\beta}. \tag{19}$$

Если обозначить неизвестные элементы блока $X_{\alpha\beta}$ через x_{ij}, а блока $C_{\alpha\beta}$ через c_{ij} $(i = 1, 2, \ldots, p; j = 1, 2, \ldots, q)$, то уравнение (19) примет вид

$$\begin{pmatrix} x_{21} & x_{22}-x_{11} & x_{23}-x_{12} & x_{24}-x_{13} & \ldots & x_{2q}-x_{1,q-1} \\ x_{31} & x_{32}-x_{21} & x_{33}-x_{22} & x_{34}-x_{23} & \ldots & x_{3q}-x_{2,q-1} \\ x_{41} & x_{42}-x_{31} & x_{43}-x_{32} & x_{44}-x_{33} & \ldots & x_{4q}-x_{3,q-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ x_{p1} & x_{p2}-x_{p-1,1} & x_{p3}-x_{p-1,2} & x_{p4}-x_{p-1,3} & \ldots & x_{pq}-x_{p-1,q-1} \\ 0 & -x_{p1} & -x_{p2} & -x_{p3} & \ldots & -x_{p,q-1} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & \ldots & c_{1q} \\ c_{21} & c_{22} & c_{23} & c_{24} & \ldots & c_{2q} \\ c_{31} & c_{32} & c_{33} & c_{34} & \ldots & c_{3q} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{p-1,1} & c_{p-1,2} & c_{p-1,3} & c_{p-1,4} & \ldots & c_{p-1,q} \\ c_{p1} & c_{p2} & c_{p3} & c_{p4} & \ldots & c_{pq} \end{pmatrix}. \tag{20}$$

Равенство (20) является системой pq линейных алгебраических уравнений с pq неизвестными x_{ij}: $x_{i+1,1} = c_{i1}$, $x_{i+1j} - x_{ij-1} = c_{ij}$, $i = 1, 2, \ldots, p-1$, $j = 2, 3, \ldots, q$; $-x_{pj-1} = c_{pj}$. Из (20) видно, что первым необходимым условием разрешимости системы является равенство $c_{p1} = 0$. Далее, складывая элементы главной диагонали, а также каждой параллельной ей поддиагонали, получим при $q \geqslant p$ равенства

$$x_{i+1,1} + \sum_{k=1}^{p-i-1} (x_{i+k+1,k+1} - x_{i+k,k}) - x_{p,p-i} = 0, \tag{21}$$

значит, матрица $C_{\alpha\beta}$ тоже должна удовлетворять этим условиям, т. е. должны выполняться равенства

$$\sum_{k=0}^{s-i} c_{i+k,k+1} = 0 \quad (i = 1, 2, \ldots, s, \;\; s = \min(p, q)). \tag{22}$$

Необходимость условий (10) (они же (22)) для решения (3), задекларированная в формулировке теоремы, доказана. Осталось доказать их достаточность.

Заметим, что при $q > p$ выполняются не только равенства (21), но и равенства

$$\sum_{k=1}^{p-1} (x_{k+1,j+k} - x_{k,j+k-1}) - x_{p,p+j-1} = \sum_{k=1}^{p} c_{k,j+k}, \quad j = 1, 2, \ldots, q-p. \tag{23}$$

Так как в сумме в левой части (23) все слагаемые, кроме одного, взаимно уничтожаются, то из (23) получаем, что в системе (20), если она имеет решение, неизвестные x_{1j} однозначно определяются формулой

$$x_{1j} = -\sum_{k=1}^{p} c_{k,j+k}, \quad j = 1, 2, \ldots, q-p. \tag{24}$$

На остальные неизвестные x_{1j}, $j = q-p+1$, $q-p+2, \ldots, q-2$, $q-1$, q, система (20) никаких ограничений не накладывает, и их можно выбирать произвольными числами. Покажем, что при фиксированном выборе значений этих неизвестных остальные неизвестные определяются однозначно, и, таким образом, при выполнении (10) общее решение системы (20) содержит p произвольных постоянных.

Условимся считать, что $x_{1j} = c_{0j}$ и что $c_{ij} = 0$, если $i < 0$. Тогда, приравнивая первые столбцы в (20), находим: $x_{i1} = c_{i-1,1}$, $i = 2, 3, \ldots, p$. Приравнивая вторые столбцы, найдем: $x_{i2} = x_{i-1,1} + c_{i-1,2}$, откуда ($i = 2, p, \ldots, p$) $x_{i2} = c_{i-2,1} + c_{i-1,2}$. Аналогично $x_{i3} = x_{i-1,2} + c_{i-1,3}$ ($i = 2, 3, \ldots, p$), т. е. $x_{i3} = c_{i-3,1} + c_{i-2,2} + c_{i-1,3}$. Если $p \leqslant q$, то, повторяя этот процесс $p-1$ раз, получим

$$x_{ij} = \sum_{k=0}^{j-1} c_{i-j+k,k+1}, \quad i = 2, 3, \ldots, p, \;\; j = 1, 2, \ldots, q. \tag{25}$$

В частности, при $i = p$ равенство (25) дает $x_{pj} = \sum_{k=0}^{j-1} c_{p-j+k,k+1}$. Но из последней строки (20) обнаруживаем, что $x_{pj} = -c_{pj+1}$. Однако противоречия здесь нет, так как в силу условия (10) эти равенства дают одно и то же число. То, что (24) и (25) действительно дают решение системы

(20) при любых x_{1j}, $j = q-p+1,\ q-p+2, \ldots, q-2,\ q-1,\ q$, проверяется непосредственной подстановкой. Таким образом, в случае $p \leqslant q$ достаточность доказана.

В случае $q \leqslant p$ аналогичный результат можно получить за счет транспонирования системы (20), но можно решать ее и непосредственно (естественно, при выполнении условия (10)), исходя из равенства последних строк равенства (20). Как и в случае $q > p$, заметим предварительно, что выполняются не только равенства (22), но и равенства

$$x_{i+1,1} + \sum_{k=1}^{q-1}(x_{i+k+1,k+1} - x_{i+k,k}) = \sum_{k=1}^{q} c_{i+k-1,k}, \quad i = 1, 2, \ldots, p-q. \tag{26}$$

Так как в сумме в левой части (26) все слагаемые, кроме одного, взаимно уничтожаются, то из (26) получаем, что в системе (20), если она имеет решение, неизвестные x_{iq} однозначно определяются формулой

$$x_{iq} = \sum_{k=1}^{q} c_{k,i+k}, \quad i = 1, 2, \ldots, q-p. \tag{27}$$

На неизвестные $x_{iq} = c_{i,q+1}$, $1 = 1, 2, \ldots, q$, система (20) никаких ограничений не налагает, и мы можем считать их произвольными числами (которые обозначены $x_{iq} = c_{i,q+1}$). При $j > q+1$ считаем, что $c_{ij} = 0$.

Приравнивая элементы последних строк матриц из (20), находим

$$x_{pj} = -c_{p,j+1}, \quad j = 1, 2, \ldots, q-1. \tag{28}$$

Из равенства предпоследних строк (20) и из (28) получаем, что

$$x_{p-1j} = -c_{p-l,j+1} - c_{p,j+2}, \quad j = 1, 2, \ldots, q-1. \tag{29}$$

Из равенства третьих снизу строк (20) и из (29) получаем, что

$$x_{p-2j} = -c_{p-2,j+1} - c_{p-l,j+2} - c_{p,j+3}, \quad j = 1, 2, \ldots, q-1. \tag{30}$$

Продолжая этот процесс, приходим к равенствам

$$x_{ij} = -\sum_{k=0}^{p-i} c_{i+k,j+k+1}, \quad i = 1, 2, 3, \ldots, p,\ \ j = 1, 2, \ldots, q-1. \tag{31}$$

В частности, при $j = 1$ (31) дает $x_{i1} = -\sum_{k=0}^{p-i} c_{i+k,k+2}$, а из первого столбца (20) следует, что $x_{i1} = c_{i-1,1}$, однако из условий (10) следует, что эти числа равны между собой. То, что (27) и (31) действительно дают решение системы (20) при любых $x_{i,q}$, $i = 1, 2, \ldots, q$, проверяется непосредственной подстановкой. Таким образом, в случае $q \leqslant p$ достаточность доказана.

Теорема доказана полностью. □

Замечания.

1) Ясно, что в случае $i, j = 1, 2, \ldots, s$, $s = \min(p, q)$, формулы (25) и (31) дают одно и то же значение для неизвестных x_{ij} в силу условий (10).

2) Фробениус [1] нашел общий вид матриц, перестановочных с данной матрицей A, т. е. решил уравнение (3) в том частном случае, когда $C = 0$. Очевидно, что в этом случае условия (10) заведомо выполняются, и поэтому задача имеет бесчисленное множество решений. При этом блоки $X_{\alpha\beta}$, соответствующие случаю $\lambda_\alpha \neq \lambda_\beta$, тождественно равны 0, что получается и из формулы (17) при $C = 0$, а блоки $X_{\alpha\beta}$, соответствующие случаю $\lambda_\alpha = \lambda_\beta$, содержат, как и в общем случае, $s = \min(p, q)$ произвольных параметров, причем равенства (25) и (31) в этом случае вырождаются в равенство соответствующих неизвестных x_{ij} произвольным числам или нулю.

§2. Уравнение $AX + XB = C$

Рассмотрим уравнение

$$AX + XB = C. \tag{32}$$

Пусть матрицы A, B, C одинакового порядка n таковы, что A приводится к жордановой форме матрицей U, а матрица B — матрицей V, т.е. выполняются равенства $A = U^{-1}\tilde{A}U$, $B = V^{-1}\tilde{B}V$, где $\tilde{A}$ и $\tilde{B}$ — жордановы формы матриц A и B. Тогда уравнение (32) перепишется в виде

$$U^{-1}\tilde{A}UX + XV\tilde{B}V^{-1} = C. \tag{33}$$

Следуя Фробениусу, умножим (32) слева на U, справа на V. Получим

$$\tilde{A}UXV + UXV\tilde{B} = UCV.$$

Опять по Фробениусу обозначим через

$$\tilde{X} = UXV, \quad \tilde{C} = UCV. \tag{34}$$

Уравнение (32) принимает вид

$$\tilde{A}\tilde{X} + \tilde{X}\tilde{B} = \tilde{C}. \tag{35}$$

Это уравнение имеет тот же вид, что и (32), но матрицы $\tilde{A}$ и $\tilde{B}$ имеют жорданову форму, т.е. состоят из жордановых блоков. Рассмотрим пересечение этих разбиений на блоки, т.е. будем считать, что матрицы $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{X}$ разбиты на блоки $\tilde{A}_{\alpha\beta}$, $\tilde{B}_{\alpha\beta}$, $\tilde{C}_{\alpha\beta}$, $\tilde{X}_{\alpha\beta}$, каждый из которых соответствует только одному собственному числу λ_α матрицы A и только одному собственному числу μ_β матрицы B.

По правилу перемножения блочных матриц, как и в случае уравнения (3), обе части уравнения (35) состоят из блоков, аналогичных блокам в (13), приравнивая которые получаем равносильную (30) систему блочных равенств

$$(\lambda_\alpha I + H_\alpha)\tilde{X}_{\alpha\beta} + \tilde{X}_{\alpha\beta}(\mu_\beta I + H_\beta) = \tilde{C}_{\alpha\beta},$$

т.е.

$$(\lambda_\alpha + \mu_\beta)\tilde{X}_{\alpha\beta} = \tilde{C}_{\alpha\beta} - \tilde{X}_{\alpha\beta}H_\beta - H_\alpha\tilde{X}_{\alpha\beta}. \tag{36}$$

Таким образом, задача решения матричного уравнения (32) равносильна решению уравнений (36) при всех α и β. Для решения этих уравнений опять произведем ,,итерации Фробениуса“ равенства (36), т.е. будем умножать это равенство на $(\lambda_\alpha+\mu_\beta)$ и каждый раз заменять $(\lambda_\alpha+\mu_\beta)\tilde{X}_{\alpha\beta}$ на его выражение по формуле (36).

В результате первой итерации получим

$$\begin{aligned}(\lambda_\alpha + \mu_\beta)^2\tilde{X}_{\alpha\beta} &= (\lambda_\alpha + \mu_\beta)\tilde{C}_{\alpha\beta} - (\lambda_\alpha + \mu_\beta)\tilde{X}_{\alpha\beta}H_\beta - H_\alpha(\lambda_\alpha + \mu_\beta)\tilde{X}_{\alpha\beta}\\ &= (\lambda_\alpha + \mu_\beta)\tilde{C}_{\alpha\beta} - \tilde{C}_{\alpha\beta}H_\beta\\ &\quad - H_\alpha\tilde{C}_{\alpha\beta} + \tilde{X}_{\alpha\beta}H_\beta^2 + 2H_\alpha\tilde{X}_{\alpha\beta}H_\beta + H_\alpha^2\tilde{X}_{\alpha\beta}.\end{aligned}$$

Аналогично проводятся 2-я, 3-я и следующие итерации. В результате итерации с номером r получаем

$$\begin{aligned}(\lambda_\alpha + \mu_\beta)^{r+1}\tilde{X}_{\alpha\beta} &= \sum_{i=0}^{r}(-1)^i(\lambda_\alpha + \mu_\beta)^{r-i}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k}\\ &\quad + (-1)^{r+1}\sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+1-i}.\end{aligned} \tag{37}$$

Докажем это равенство по индукции. Предположим, что (37) уже доказано для некоторого r. Докажем, что тогда оно верно и для $r+1$. Для этого умножим (32) на $(\lambda_\alpha + \mu_\beta)$ и воспользуемся (36). Получим

$$
\begin{aligned}
(\lambda_\alpha + \mu_\beta)^{r+2}\tilde{X}_{\alpha\beta} &= \sum_{i=0}^{r}(-1)^i(\lambda_\alpha + \mu_\beta)^{r-i+1} \\
&\times \sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k} + (-1)^{r+1}\sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i (\lambda_\alpha + \mu_\beta)\tilde{X}_{\alpha\beta} H^{r+1-i} \\
&= \sum_{i=0}^{r}(-1)^i(\lambda_\alpha + \mu_\beta)^{r-i+1}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k} + (-1)^{r+1} \\
&\times \sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i (\tilde{C}_{\alpha\beta} - \tilde{X}_{\alpha\beta}H_\beta - H_\alpha \tilde{X}_{\alpha\beta}) H_\beta^{r+1-i} \\
&= \sum_{i=0}^{r+1}(-1)^i(\lambda_\alpha + \mu_\beta)^{(r+1)-i}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k} \\
&+ (-1)^{r+2}\left(\sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} + \sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^{i+1}\tilde{X}_{\alpha\beta} H_\beta^{r+1-i}\right).
\end{aligned}
$$

В последней сумме сделаем замену индекса i на $i+1$:

$$
\begin{aligned}
&\sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} + \sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^{i+1}\tilde{X}_{\alpha\beta} H_\beta^{r+1-i} \\
&\quad = \sum_{i=0}^{r+1} C_{r+1}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} + \sum_{i=1}^{r+2} C_{r+1}^{i-1} H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} \\
&\quad = C_{r+1}^0 H_\alpha^0 \tilde{X}_{\alpha\beta} H_\beta^{r+2} + \sum_{i=1}^{r+1} C_{r+1}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} \\
&\quad + \sum_{i=1}^{r+1} C_{r+1}^{i-1} H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{r+2-i} + C_{r+1}^{r+1} H_\beta^{r+2}\tilde{X}_{\alpha\beta} H_\beta^{(r+2)-(r+2)} \\
&\quad = \sum_{i=0}^{r+2} C_{r+2}^i H_\alpha^i \tilde{X}_{\alpha\beta} H_\beta^{(r+2)-i}.
\end{aligned}
$$

Подставляя это в предыдущую формулу, получим результат применения $r+1$-й итерации равенства (36), и формула (37) доказана. (Замечание. Одновременно доказана и формула (14), так как она — частный случай (37) при $B = -A$.)

Как и в случае уравнения (3), если размеры блока $\tilde{X}_{\alpha\beta}$ равны $p \times q$, $p = p(\alpha,\beta)$, $q = q(\alpha,\beta)$, то в силу нильпотентности H при $r > p+q-2$ все слагаемые в последней сумме формулы (37) равны 0, и в этом случае равенство (37) приобретает вид

$$(\lambda_\alpha + \mu_\beta)^{r+1}\tilde{X}_{\alpha\beta} = \sum_{i=0}^{r}(-1)^i(\lambda_\alpha+\mu_\beta)^{r-i}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k}. \tag{38}$$

1) Если $\lambda_\alpha + \mu_\beta \neq 0$, то делим (38) на $(\lambda_\alpha+\mu_\beta)^{r+1}$. Находим, что

$$\tilde{X}_{\alpha\beta} = \sum_{i=0}^{r}(-1)^i(\lambda_\alpha+\mu_\beta)^{-i-1}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k}. \tag{39}$$

Как и при решении уравнения (3), при $i > p+q-2$ все слагаемые внутренней суммы в (39) равны 0, и поэтому в (39) можно считать, что $r = p+q-1$, т.е. справедливо равенство

$$\tilde{X}_{\alpha\beta} = \sum_{i=0}^{p+q-1}\frac{1}{(\lambda_\alpha+\mu_\beta)^{i+1}}\sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k}. \tag{40}$$

Итак, если решение уравнения (36) при $\lambda_\alpha + \mu_\beta \neq 0$ существует, то оно единственно и определяется формулой (40). То, что (40) действительно является решением (36), как и ранее, проверяется подстановкой.

2) Если $\lambda_\alpha + \mu_\beta = 0$, то (36) превращается в

$$H_\alpha \tilde{X}_{\alpha\beta} + \tilde{X}_{\alpha\beta} H_\beta = \tilde{C}_{\alpha\beta}. \tag{41}$$

Если обозначить неизвестные элементы блока $\tilde{X}_{\alpha\beta}$ через $\tilde{x}_{ij}$, а блока $\tilde{C}_{\alpha\beta}$ через $\tilde{c}_{ij}$ $(i = 1,2,\ldots,p; j = 1,2,\ldots,q)$, то уравнение (41) примет вид

$$\begin{pmatrix} \tilde{x}_{21} & \tilde{x}_{22}+\tilde{x}_{11} & \tilde{x}_{23}+\tilde{x}_{12} & \tilde{x}_{24}+\tilde{x}_{13} & \ldots & \tilde{x}_{2q}+\tilde{x}_{1,q-1} \\ \tilde{x}_{31} & \tilde{x}_{32}+\tilde{x}_{21} & \tilde{x}_{33}+\tilde{x}_{22} & \tilde{x}_{34}+\tilde{x}_{23} & \ldots & \tilde{x}_{3q}+\tilde{x}_{2,q-1} \\ \tilde{x}_{41} & \tilde{x}_{42}+\tilde{x}_{31} & \tilde{x}_{43}+\tilde{x}_{32} & \tilde{x}_{44}+\tilde{x}_{33} & \ldots & \tilde{x}_{4q}+\tilde{x}_{3,q-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \tilde{x}_{p1} & \tilde{x}_{p2}+\tilde{x}_{p-1,1} & \tilde{x}_{p3}+\tilde{x}_{p-1,2} & \tilde{x}_{p4}+\tilde{x}_{p-1,3} & \ldots & \tilde{x}_{pq}+\tilde{x}_{p-1,q-1} \\ 0 & \tilde{x}_{p1} & \tilde{x}_{p2} & \tilde{x}_{p3} & \ldots & \tilde{x}_{p,q-1} \end{pmatrix} = \begin{pmatrix} \tilde{c}_{11} & \tilde{c}_{12} & \tilde{c}_{13} & \tilde{c}_{14} & \ldots & \tilde{c}_{1q} \\ \tilde{c}_{21} & \tilde{c}_{22} & \tilde{c}_{23} & \tilde{c}_{24} & \ldots & \tilde{c}_{2q} \\ \tilde{c}_{31} & \tilde{c}_{32} & \tilde{c}_{33} & \tilde{c}_{34} & \ldots & \tilde{c}_{3q} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \tilde{c}_{p-1,1} & \tilde{c}_{p-1,2} & \tilde{c}_{p-1,3} & \tilde{c}_{p-1,4} & \ldots & \tilde{c}_{p-1,q} \\ \tilde{c}_{p1} & \tilde{c}_{p2} & \tilde{c}_{p3} & \tilde{c}_{p4} & \ldots & \tilde{c}_{pq} \end{pmatrix}. \tag{42}$$

Как и ранее, складывая с чередующимися знаками элементы главной диагонали, а также каждой параллельной ей поддиагонали, получим при $q \geqslant p$ равенства

$$\tilde{x}_{i+1,1} + \sum_{k=1}^{p-i-1} (-1)^k (\tilde{x}_{i+k+1,k+1} + \tilde{x}_{i+k,k}) + (-1)^{p-i} \tilde{x}_{p,p-i} = 0, \tag{43}$$

значит, матрица $\tilde{C}_{\alpha\beta}$ тоже должна удовлетворять этим условиям, т. е. должны выполняться равенства

$$\sum_{k=0}^{s-i} (-1)^k \tilde{c}_{i+k,k+1} = 0 \quad (i = 1, 2, \ldots, s, s = \min(p, q)). \tag{44}$$

Таким образом, найдены необходимые условия разрешимости системы (42). Осталось доказать их достаточность.

Рассуждая аналогично предыдущему, заметим, что при $q > p$ выполняются не только равенства (43), но и равенства

$$\begin{aligned} &\sum_{k=1}^{p-1} (-1)^k (\tilde{x}_{k+1,j+k} + \tilde{x}_{k,j+k-1}) + (-1)^p \tilde{x}_{p,p+j-1} \\ &= \sum_{k=1}^{p} (-1)^k \tilde{c}_{k,j+k}, \quad j = 1, 2, \ldots, q-p. \end{aligned} \tag{45}$$

Так как в сумме в левой части (45) все слагаемые, кроме одного, взаимно уничтожаются, то из (45) получаем, что в системе (42), если она имеет решение, неизвестные $\tilde{x}_{1j}$ однозначно определяются формулой

$$\tilde{x}_{1j} = \sum_{k=1}^{p} (-1)^k c_{k,j+k}, \quad j = 1, 2, \ldots, q-p. \tag{46}$$

На остальные неизвестные $\tilde{x}_{1j}$, $j = q-p+1, q-p+2, \ldots, q-2, \ldots, q-1, q$, система (42) никаких ограничений не накладывает, и их можно выбирать произвольными числами. Покажем, что при фиксированном выборе значений этих неизвестных остальные неизвестные определяются однозначно, и, таким образом, при выполнении (43) общее решение системы (42) содержит p произвольных постоянных.

Как и в первой части, условимся считать, что $x_{1j} = c_{0j}$ и что $c_{ij} = 0$, если $i < 0$. Рассуждая аналогично (т.е. приравнивая столбцы в (42) и повторяя этот процесс $p-1$ раз), получим

$$\tilde{x}_{ij} = \sum_{k=0}^{j-1} (-1)^k \tilde{c}_{i-j+k,k+1}, \quad i = 2, 3, \ldots, p, \;\; j = 1, 2, \ldots, q. \tag{47}$$

И опять здесь нет противоречия с последней строкой равенства (42) в силу условия (44). То, что (46) и (47) действительно дают решение системы (42) при любых x_{1j}, $j = q-p+1,\ q-p+2, \ldots, q-2,\ q-1,\ q$, проверяется непосредственной подстановкой. Таким образом, в случае $p \leqslant q$ достаточность доказана. Аналогично рассматривается и случай $q \leqslant p$.

Таким образом, доказана следующая теорема.

Теорема 2. *Для того, чтобы матричное уравнение $AX+XB=C$ имело единственное решение, необходимо и достаточно, чтобы все собственные числа λ_α матрицы A и все собственные числа μ_β матрицы B удовлетворяли условию $\lambda_\alpha + \mu_\beta \neq 0$, при этом решение этого уравнения определяется равенством*

$$X = U^{-1}\tilde{X}V^{-1},$$

где U и V — матрицы, приводящие матрицы A и B к жордановой форме, а $\tilde{X}$ — блочная матрица, состоящая из блоков $X_{\alpha\beta}$, соответствующих пересечениям жордановых ящиков матриц A и B и вычисляемых по формуле (40):

$$\tilde{X}_{\alpha\beta} = \sum_{i=0}^{p+q-1} \frac{1}{(\lambda_\alpha + \mu_\beta)^{i+1}} \sum_{k=0}^{i} C_i^k H_\alpha^k \tilde{C}_{\alpha\beta} H_\beta^{i-k},$$

где $\tilde{C}_{\alpha\beta}$ — соответствующий блок матрицы $\tilde{C} = UCV$. Если же хотя бы для одного блока выполняется равенство $\lambda_\alpha + \mu_\beta = 0$, то данное уравнение либо не имеет решений, либо имеет их бесконечно много. Для того, чтобы у рассматриваемого уравнения существовали решения, необходимо и достаточно, чтобы в каждом блоке $\tilde{C}_{\alpha\beta}$, для которого $\lambda_\alpha + \mu_\beta = 0$, выполнялись равенства (44). *Для каждого такого блока, имеющего p строк и q столбцов, $s = \min(p,q)$ неизвестных элементов блока $\tilde{X}_{\alpha\beta}$ можно выбрать произвольно, а остальные элементы найдутся через них однозначно по формулам* (46) *и* (47) (*в случае $q \geqslant p$, в противоположном случае результат и формулы для решения аналогичны*).

Замечание. Общее число произвольных значений неизвестных при решении уравнения (32) в случае $\lambda_\alpha + \mu_\beta = 0$ и выполнении (44) такое же, как и при решении уравнения $AX = XB$. Оно подсчитано еще Фробениусом и указано в [2]. Разумеется, в качестве произвольных можно использовать разные неизвестные (не только использованные выше), их количество от этого не изменится. Здесь использовались наиболее удобные для наших целей.

Список литературы

[1] Frobenius G., *Ueber die vertauschbare Matrizen*, Sitz.-Ber. Akad. Wiss. Phys.-Math. Klasse, Berlin, 1896, 601–614.
[2] Гантмахер Ф. Р., *Теория матриц*, ГИТТЛ, М., 1953.
[3] Беллман Р., *Введение в теорию матриц*, Наука, М., 1969.
[4] Деменчук А. К., Макаров Е. К., *Полиномиальная формула для самосопряженных решений матричного уравнения $AX - XA = C$ с нормальными коэффициентами*, Докл. НАН Беларуси **52** (2008), №2, 5–10.
[5] Bračič J., Drnovšek R., Farforovskaya Yu. B., Rabkin E. L., Zemanek J., *On positive commutators*, Positivity **14** (2010), no. 3, 431–439.
[6] Канторович Л. В., Акилов Г. П., *Функциональный анализ в нормированных пространствах*, Физматгиз, М., 1959.
[7] Гельфанд И. М., Райков Д. А., Шилов Г. Е., *Коммутативные нормированные кольца*, Физматгиз, М., 1960.

С.-Петербургский
государственный университет
телекоммуникаций
им. проф. М. А. Бонч-Бруевича
Россия
E-mail: rabk@sut.ru