УДК 512.643

Численное решение дискретного ВНН-уравнения в нормальном случае

Х.Д. Икрамов[1], Ю.О. Воронцов[2]

[1]Московский государственный университет им. М.В. Ломоносова, Ленинские горы, 1, Москва, 119991

[2]ООО "Глобус Медиа", 1-й Нагатинский пр-д, д. 10, Москва, 115230

E-mails: ikramov@cs.msu.su (Икрамов Х.Д.), vv@cs.msu.su (Воронцов Ю.О.)

Икрамов Х.Д., Воронцов Ю.О. Численное решение дискретного ВНН-уравнения в нормальном случае // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2018. — Т. 21, № 4. — С. 367–373.

Известно, что решение полулинейного матричного уравнения $X - A\overline{X}B = C$ можно свести к решению классического уравнения Стейна. Нормальный случай означает, что коэффициенты левой части полученного уравнения суть нормальные матрицы. Предлагается способ решения исходного полулинейного уравнения в нормальном случае, позволяющий для уравнений порядка $n = 3000$ почти вдвое сократить время вычислений по сравнению с библиотечной функцией dlyap, решающей уравнения Стейна в системе Matlab.

DOI: 10.15372/SJNM20180402

Ключевые слова: *непрерывное и дискретное уравнения Сильвестра, ВНН-уравнения, форма Шура, сопряженно-нормальная матрица, функция Matlab'a dlyap.*

Ikramov Kh.D., Vorontsov Yu.O. Numerical solution of the discrete BHH-equation in the normal case // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2018. — Vol. 21, № 4. — P. 367–373.

It is known that the solution of the semilinear matrix equation $X - A\overline{X}B = C$ can be reduced to solving the classical Stein equation. The normal case means that the coefficients on the left-hand side of the resulting equation are normal matrices. We propose a method for solving the original semilinear equation in the normal case that permits to almost halve the execution time for equations of order $n = 3000$ compared to the library function dlyap, which solves Stein equations in Matlab.

Keywords: *continuous- and discrete-time Sylvester equations, BHH-equations, Schur form, conjugate-normal matrix, Matlab function dlyap.*

1. Введение

В литературе по вычислительной линейной алгебре и теории управления хорошо известны линейные матричные уравнения:

$$AX + XB = C \tag{1}$$

и

$$X - AXB = C, \tag{2}$$

называемые соответственно непрерывным и дискретным уравнениями Сильвестра (или уравнениями Сильвестра соответственно для непрерывного и дискретного времени). Второе из них называют еще уравнением Стейна. Полулинейное уравнение

$$AX + \overline{X}B = C \tag{3}$$

было, насколько нам известно, впервые исследовано в статье [1]. В честь авторов этой статьи, а также учитывая внешнее сходство уравнений (1) и (3), мы называем последнее *непрерывным ВНН-уравнением.*

В настоящей работе мы обсуждаем численное решение матричного уравнения

$$X - A\overline{X}B = C. \tag{4}$$

Оно является полулинейным аналогом уравнения (2) и потому мы называем его *дискретным ВНН-уравнением.* В общем случае матричные коэффициенты A и B этого уравнения суть квадратные матрицы, быть может, различных порядков m и n. Правая часть C и искомая матрица X — это матрицы размера $m \times n$.

В статье [2] показано, что решение уравнения (4) может быть сведено к решению уравнения Стейна.

Теорема 1. *Уравнение* (4) *разрешимо единственным образом тогда и только тогда, когда единственным образом разрешимо уравнение Стейна*

$$X - (A\overline{A})X(\overline{B}B) = C + A\overline{C}B. \tag{5}$$

При этом оба уравнения имеют одно и то же решение X.

Уравнения Стейна умеренного порядка могут быть численно решены посредством одного из двух хорошо известных ортогональных методов: Бартелса–Стьюарта (BS) и Голуба–Нэша–Ван Лоана (GNL). Краткое обсуждение этих методов дано ниже в пункте 2. Более подробное их описание можно найти, например, в [3]. Второй из названных методов — и именно в применении к уравнениям Стейна — реализован библиотечной функцией dlyap системы Matlab.

Теорема 1 указывает следующий способ численного решения дискретных ВНН-уравнений:

1) замена уравнения (4) уравнением (5);

2) применение к уравнению (5) процедуры dlyap.

Этот способ вполне эффективен в общем случае. Здесь мы хотим рассмотреть частную ситуацию, когда коэффициенты A и B уравнения (4) суть сопряженно-нормальные матрицы. Напомним, что сопряженно-нормальной называется квадратная матрица A такая, что

$$AA^* = \overline{A^*A}. \tag{6}$$

Сопряженно-нормальные матрицы выполняют в теории унитарных конгруэнций такую же роль, какую обычные нормальные матрицы играют в теории унитарных подобий.

По причинам, объясняемым в п. 3, мы называем указанную ситуацию нормальным случаем уравнения (4). Цель данной статьи — предложить модификацию описанного выше алгоритма для нормального случая. Эта модификация позволяет (даже при вполне любительской реализации на языке Matlab) почти вдвое сократить время решения ВНН-уравнений порядка $n = 3000$. Численные эксперименты, подтверждающие сказанное, обсуждаются в заключительном п. 4.

2. Ортогональные методы численного решения уравнений Стейна

Как уже отмечено во введении, имеются два ортогональных метода: метод Бартелса–Стьюарта (BS) и метод Голуба–Нэша–Ван Лоана (GNL). Остановимся более подробно на первом из них. Его описание удобно разделить на четыре этапа.

1. *Приведение матриц A и B к форме Шура.* На этом этапе вычисляются унитарные матрицы U и V такие, что матрицы

$$R = U^*AU, \qquad S = V^*BV$$

являются (верхними или нижними) треугольными. Выбор конкретной комбинации треугольных форм (из четырех возможных вариантов) определяет организацию этапа 3.

2. *Преобразование правой части.* На этом этапе вычисляется матрица

$$D = U^*CV. \tag{7}$$

Результатом первых двух этапов является замена исходного уравнения (2) новым уравнением Стейна

$$Y - RYS = D. \tag{8}$$

Неизвестная матрица Y связана с X соотношением

$$Y = U^*XV. \tag{9}$$

3. *Решение уравнения* (8). Это матричное уравнение можно рассматривать как систему из mn линейных уравнений относительно mn коэффициентов y_{ij} матрицы Y. При подходящем упорядочении этих уравнений система будет треугольной (по желанию, верхней или нижней). Если бы эта система не имела никакой специфики, то ее решение потребовало бы $O((mn)^2)$ арифметических операций. Однако специфика здесь имеется, и ее остроумное использование, предложенное Бартелсом и Стьюартом, снижает сложность этого этапа до $O(mn(m+n))$ операций. (О подробностях реализации этого этапа можно прочесть, например, в [4, §7.6].)

4. *Возврат к исходной матрице X.* В соответствии с (9), искомое решение X связано с уже вычисленной матрицей Y соотношением

$$X = UYV^*. \tag{10}$$

Алгоритм GNL имеет сходную структуру. Наиболее существенное его отличие от алгоритма BS касается этапа 1, где теперь одна из матриц A или B приводится к форме Хессенберга, а не к форме Шура. Как следствие, изменяется процесс решения системы линейных уравнений на этапе 3. Однако этот этап по-прежнему имеет сложность $O(mn(m+n))$ операций.

3. Нормальный случай

Нормальные матрицы отличаются от всех остальных следующим замечательным свойством.

Теорема 2. *Форма Шура нормальной матрицы A есть диагональная матрица ее собственных значений.*

Приведение к форме Шура осуществляется в системе Matlab функцией schur. QR-алгоритм, реализованный этой функцией, не способен извлечь пользы из информации о нормальности коэффициентов A и B. В частности, форма Шура, вычисляемая для нормальной матрицы A функцией schur, будет треугольной, а не диагональной матрицей. Однако теоретический факт, формулируемый теоремой 2, проявляется в том, что внедиагональные элементы этой формы, возникающие вследствие округлений, очень малы. Пусть, например, D — диагональная матрица порядка 2000, а N — нормальная матрица, полученная из D унитарным подобием с трансформирующей матрицей U. Если D' — треугольная матрица, полученная в арифметике двойной точности применением к N функции schur, то для типичной матрицы D с элементами, не превосходящими по модулю 10, справедливо неравенство $\|D-D'\|_F < 10^{-11}$ (норма Фробениуса, называемая также евклидовой матричной нормой).

С левой частью каждого из уравнений (2) и (4) можно связать оператор, действующий в пространстве $M_{m,n}$ комплексных матриц размера $m\times n$. Этот оператор будет линейным в случае уравнения Стейна и лишь полулинейным для ВНН-уравнения. Вводя в $M_{m,n}$ естественное скалярное произведение

$$\langle Y, Z\rangle = \mathrm{tr}\, Z^*Y,$$

можем поставить вопрос об условиях нормальности матричного оператора Стейна. Этот вопрос был разрешен в статье [5]. Аналогичный вопрос относительно ВНН-операторов значительно сложнее и требует интерпретации $M_{m,n}$ как вещественного евклидова пространства удвоенной размерности $2mn$. Именно такой подход был применен в статье [6], посвященной выводу условий нормальности для операторов типа Стейна. Как мы сейчас увидим, нормальный случай уравнения (4) соответствует одной из ситуаций, в которых матричный оператор соответствующего уравнения (5) оказывается нормальным.

Теорема 3. *В нормальном случае матрицы $A\overline{A}$ и $\overline{B}B$ в уравнении* (5) *являются нормальными в обычном смысле.*

Доказательство. Пусть A — сопряженно-нормальная матрица, т. е. выполнено (6). Перепишем это соотношение в виде

$$AA^* = A^T\overline{A}. \tag{11}$$

Беря поэлементное сопряжение от обеих матриц, получаем

$$\overline{A}A^T = A^*A. \tag{12}$$

Положим $B = A\overline{A}$. Имеем

$$BB^* = (A\overline{A})(A^TA^*) = A(\overline{A}A^T)A^* = A(A^*A)A^* = (AA^*)^2.$$

В этих выкладках использовано равенство (12). Далее,

$$B^*B = (A^TA^*)(A\overline{A}) = A^T(A^*A)\overline{A} = A^T(\overline{A}A^T)\overline{A} = (A^T\overline{A})^2 = (AA^*)^2.$$

Здесь использованы оба равенства (11) и (12). Таким образом, $BB^* = B^*B$.

Аналогичным образом можно проверить нормальность матрицы $\overline{A}A$.

Теорема 3 подсказывает следующий подход к решению уравнения (4) в нормальном случае. Как и положено в алгоритме BS, матрицы $A\overline{A}$ и $\overline{B}B$ уравнения (5) приводятся к формам Шура: $R = U^*(A\overline{A})U$ и $S = V^*(\overline{B}B)V$. Пересчитывается и правая часть уравнения:

$$C + A\overline{C}B \longrightarrow D = U^*(C + A\overline{C}B)V.$$

Затем матрицы R и S заменяются их диагональными частями:

$$R \longrightarrow \Lambda = \mathrm{diag}(r_{11}, \ldots, r_{mm}), \quad S \longrightarrow M = \mathrm{diag}(s_{11}, \ldots, s_{nn}).$$

В результате этап 3, бывший ранее наиболее трудоемкой частью всего алгоритма, сводится к тривиальному поэлементному вычислению матрицы Y:

$$y_{ij} = \frac{d_{ij}}{1 - r_{ii}s_{jj}}, \qquad i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n. \tag{13}$$

Предполагается, что выполнены условия однозначной разрешимости уравнений (4) и (5), поэтому деления в формулах (13) возможны.

Четвертый этап алгоритма остается неизменным.

4. Численные результаты

Метод численного решения уравнения (4) в нормальном случае, предложенный в предыдущем разделе, был реализован в виде функции SteinBarNormal языка Matlab. Мы провели сравнение времени работы этой функции с временем работы процедуры dlyap, примененной к соответствующему уравнению Стейна.

При формировании матричных уравнений для наших тестов использовался следующий важный факт из теории сопряженно-нормальных матриц: всякая матрица A из этого класса посредством преобразования унитарной конгруэнции

$$A \longrightarrow N = Q^T AQ, \qquad Q^*Q = I,$$

может быть преобразована в вещественную нормальную матрицу N, которую всегда можно выбрать в блочно-диагональном виде с диагональными блоками порядков 1 и 2. Наоборот, выполняя с такими блочно-диагональными матрицами конгруэнции, порождаемые случайными унитарными матрицами, будем получать плотные сопряженно-нормальные матрицы. Именно таким образом конструировались матричные коэффициенты A и B тестовых уравнений. Правая часть C для них выбиралась как матрица с псевдослучайными элементами, равномерно распределенными в круге радиуса 10. При этом мы полагали $m = n$.

Для каждого значения n решались десять ВНН-уравнений, используя вначале dlyap, а затем SteinBarNormal. Время решения, усредненное по этим десяти уравнениям, обозначим через $t_1(n)$ в случае dlyap и через $t_2(n)$ для SteinBarNormal. Начальное значение порядка n равнялось 50. В дальнейшем n возрастало с шагом 100 до значения 2950. На рисунке отношение $t_1(n)/t_2(n)$ показано как функция от порядка n.

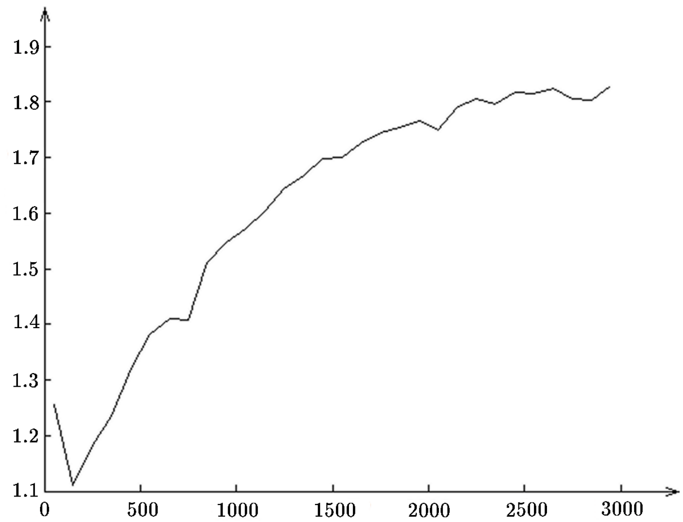

Рис.

Точность вычисленных решений оценивалась с помощью (евклидовой) нормы матричной невязки $R(\tilde{X}) = \tilde{X} - A\overline{\tilde{X}}B - C$. С ростом n значение $\|R(\tilde{X})\|$ возрастает, при $n = 3000$ не превосходя 10^{-5} для обоих сравниваемых алгоритмов. Этот уровень невязки мы считаем вполне приемлемым.

Литература

1. **Bevis J.H., Hall F.J., Hartwig R.H.** Consimilarity and the matrix equation $AX - XB = C$ // Proc. Third Auburn Matrix Theory Conference. — Amsterdam: North-Holland, 1987. — P. 51–64.
2. **Zhou B., Lam J., Duan G.-R.** Toward solution of matrix equation $X = Af(X)B + C$ // Linear Algebra Appl. — 2011. — Vol. 435, № 6. — P. 1370–1398.
3. **Икрамов Х.Д.** Численное решение матричных уравнений. — М.: Наука, 1984.
4. **Голуб Дж., Ван Лоун Ч.** Матричные вычисления. — М.: Мир, 1999; Перевод: Golub G.H., Van Loan C.F. Matrix Computations. — Baltimore, Maryland: The Johns Hopkins University Press, 1983.
5. **Икрамов Х.Д.** Условия нормальности линейных матричных операторов типа Стейна // Доклады академии наук. — 2015. — Т. 460, № 3. — С. 269–271; Перевод: Ikramov Kh.D. Normality conditions for linear matrix equations of the Stein type // Doklady Mathematics. — 2015. — Vol. 91, № 1. — P. 50–52.
6. **Икрамов Х.Д.** Условия нормальности полулинейных матричных операторов типа Стейна // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние.— Новосибирск, 2015. — Т. 18, № 4. — С. 369–375; Перевод: Ikramov Kh.D. Normality conditions for semilinear matrix operators of the Stein type // Num. Anal. and Appl. — 2015. — Vol. 8, № 4. — P. 299–303.

Поступила в редакцию 21 декабря 2017 г.,
в окончательном варианте 20 июня 2018 г.

Литература в транслитерации

1. **Bevis J.H., Hall F.J., Hartwig R.H.** Consimilarity and the matrix equation $AX - XB = C$ // Proc. Third Auburn Matrix Theory Conference. — Amsterdam: North-Holland, 1987. — P. 51–64.

2. **Zhou B., Lam J., Duan G.-R.** Toward solution of matrix equation $X = Af(X)B + C$ // Linear Algebra Appl. — 2011. — Vol. 435, № 6. — P. 1370–1398.
3. **Ikramov H.D.** Chislennoe reshenie matrichnyh uravneniy. — M.: Nauka, 1984.
4. **Golub Dzh., Van Loun Ch.** Matrichnye vychisleniya. — M.: Mir, 1999; Perevod: Golub G.H., Van Loan C.F. Matrix Computations. — Baltimore, Maryland: The Johns Hopkins University Press, 1983.
5. **Ikramov H.D.** Usloviya normal'nosti lineynyh matrichnyh operatorov tipa Steyna // Doklady akademii nauk. — 2015. — T. 460, № 3. — S. 269–271; Perevod: Ikramov Kh.D. Normality conditions for linear matrix equations of the Stein type // Doklady Mathematics. — 2015. — Vol. 91, № 1. — P. 50–52.
6. **Ikramov H.D.** Usloviya normal'nosti polulineynyh matrichnyh operatorov tipa Steyna // Sib. zhurn. vychisl. matematiki / RAN. Sib. otd-nie.— Novosibirsk, 2015. — T. 18, № 4. — S. 369–375; Perevod: Ikramov Kh.D. Normality conditions for semilinear matrix operators of the Stein type // Num. Anal. and Appl. — 2015. — Vol. 8, № 4. — P. 299–303.