Алгоритм нормального решения совместных систем линейных уравнений*

М. В. ЛЕЖНЁВ
Кубанский государственный технологический университет
e-mail: lezhnev@kubstu.ru

УДК 519.61+512.84

Ключевые слова: совместные системы линейных уравнений, неполный строчный ранг, нормальное решение, численный алгоритм.

Аннотация

Предлагается численный алгоритм нахождения нормального решения совместных систем линейных алгебраических уравнений неполного строчного ранга. Приведены результаты сравнения численной реализации предложенного алгоритма с некоторыми известными подпрограммами.

Abstract

M. V. Lezhnev, An algorithm for finding normal solutions of consistent systems of linear equations, Fundamentalnaya i prikladnaya matematika, vol. 14 (2008), no. 2, pp. 121—128.

We propose a numerical algorithm for finding the normal solution of consistent systems of linear algebraic equations of incomplete rank of rows. Results of a comparison of a numerical realization of the proposed algorithm with some known subroutines are given.

1. Диагонализация матрицы и нормальное решение системы

Рассмотрим совместную систему линейных алгебраических уравнений

$$Ax = f \tag{1}$$

с матрицей A размера $N \times M$, $\operatorname{rank} A = K \leqslant \min(N, M)$. Приведём её к виду

$$x_j + \sum_{i=K+1}^{M} b_{ji} x_i = g_j, \quad j = 1, \ldots, K, \tag{2}$$

*Работа выполнена в рамках проектов РФФИ 06-01-96645, 08-01-00907-а.

Фундаментальная и прикладная математика, 2008, том 14, № 2, с. 121—128.

который далее будем называть диагонализированным видом исходной системы. Полученные таким образом матрицу и вектор правой части обозначим через B и g соответственно, преобразованную систему запишем в матричном виде

$$Bx = g,$$

где минор $B(1:K, 1:K)$ — единичная матрица, минор вида $(1:K, K+1:M)$ матрицы B общего вида. Здесь и далее запись $B(K:L, M:N)$, или просто $(K:L, M:N)$, означает минор матрицы B, стоящий на пересечении строк от K до L и столбцов от M до N, включая их. Таким образом, мы перешли от совместной системы (1) к эквивалентной системе (2) диагонализированного вида [2].

В случае $K < N$ исходная система имеет неполный строчный ранг, для совместных систем ранг матрицы A совпадает с рангом расширенной матрицы (и равен K), т. е. в процессе приведения (1) к диагонализированному виду (2) линейно зависимые строки становятся равными нулю вместе с правыми частями. Заметим, что в случае несовместности исходной системы (1) системы уравнений (1) и (2) не эквивалентны друг другу, так как в процессе приведения (1) к виду (2) различными способами (например, методом Гаусса с выбором максимального элемента и без выбора максимального элемента) элементы правой части системы (2) могут получаться различными.

В частном случае при $K = M$ матрица $B = E$ единичная; таким образом, имеем решение $x = g$.

Одно из возможных решений системы (2) при $K < M$, отнюдь не минимальной нормы, имеет вид $x(1:K) = g(1:K)$, $x(K+1:M) = 0$, где запись $x(M:N)$ здесь и далее означает компоненты вектора x с номерами от M до N включительно.

Найдём решение системы (2) при $K < M$ с минимальной нормой. Квадрат евклидовой нормы вектора x при условии (2), т. е. квадрат расстояния от начала координат до M-мерной точки x, удовлетворяющей всем уравнениям диагонализированной системы (2), имеет вид

$$\|x\|^2 = \sum_{j=1}^{K}\left(g_j - \sum_{i=K+1}^{M} b_{ji}x_i\right)^2 + \sum_{i=K+1}^{M} x_i^2.$$

Задача минимизации квадрата нормы $\|x\|^2$, т. е. поиск нормального [2] решения системы (2), приводит к системе линейных уравнений

$$x_p + \sum_{i=K+1}^{M}\left(\sum_{j=1}^{K} b_{pj}^{\mathrm{T}} b_{ji}\right)x_i = \sum_{j=1}^{K} b_{pj}^{\mathrm{T}} g_j, \quad p = K+1, \ldots, M,$$

где b_{pj}^{T} — элемент матрицы B, стоящий на пересечении j-й строки и p-го столбца; значком T здесь и далее будем обозначать операцию транспонирования. В матричном виде имеем

$$(E + C^{\mathrm{T}}C)z = C^{\mathrm{T}}g, \tag{3}$$

где через C обозначен минор матрицы B,

$$C = B(1:K, K+1:M),$$

E — единичная матрица, $z = x(K+1:M)$ — часть искомого решения. Симметричная положительно определённая матрица $E + C^{\mathrm{T}}C$ имеет полный ранг $M - K$, система (3) может быть решена стандартными методами. Оставшаяся же часть решения $x(1:K)$ восстанавливается (обратный ход исключений Гаусса) непосредственно из (2),

$$x_j = g_j - \sum_{i=K+1}^{M} b_{ji}x_i, \quad j = 1, \ldots, K.$$

2. Замечания к численной реализации алгоритма

Процесс численного приведения системы (1) к виду (2) при использовании классических прямых исключений Гаусса удобнее реализовывать с выбором ведущего элемента по всей текущей подматрице (полный выбор), что, во-первых, в силу устойчивости последнего [3] предпочтительнее частичного выбора (по столбцу) и, во-вторых, ведёт к диагональному преобладанию в системе (3), так как элементы матрицы C по модулю не больше единицы (в силу выбора максимального элемента с индексами (j, j) по всей подматрице $(J:N, J:M)$ в процессе приведения (1) к виду (2)). Помимо того, полный выбор упрощает дальнейшую численную реализацию алгоритма.

Количество операций умножения и операций сложения для нахождения нормального решения методом диагонализации для недоопределённых систем полного строчного ранга оценено в [7]. С точностью до младших членов при больших N и M общее число флопов не превосходит

$$\frac{2}{3}M^3 - M^2N + 2MN^2 - N^3$$

и равно этому значению при $N = K$, количество операций сложения здесь равно количеству операций умножения. Если для решения (3) использовать метод Холецкого для симметричных положительно определённых матриц, что требует в два раза меньшего числа вычислительных затрат, чем метод Гаусса [3], то количество флопов станет ещё меньше.

Данный метод можно применять и к несовместным системам линейных алгебраических уравнений, при этом классические прямые исключения Гаусса предпочтительно заменить ортогональными, например вращениями слева.

Алгоритмы реализованы на языке программирования Фортран, вычисления проводились с двойной точностью, нулём считались числа, абсолютные значения которых не превосходили 10^{-14}.

Отметим здесь, что обычно «линейная задача наименьших квадратов заключается в разыскании вектора x, минимизирующего величину $\|Ax - f\|_2$» [4], т. е. невязку системы уравнений, а не само решение системы. Более того, метод наименьших квадратов чаще всего применяется для переопределённых систем [3, 4].

3. Численное сравнение алгоритма с некоторыми процедурами известных библиотечных пакетов

В [5] приведены результаты численного сравнения процедур библиотеки MS IMSL [1] с предложенным алгоритмом для недоопределённых систем полного строчного ранга, для реализации алгоритма использовались классические исключения Гаусса. Здесь мы приводим результаты численного эксперимента для недоопределённых совместных систем неполного ранга в сравнении с процедурами ASP1d и ASP5d библиотеки НИВЦ МГУ, взятыми на сайте `http://www.srcc.msu.su/`, и процедурами dLSQRR, dLSBRR и dLQRRV библиотеки MS IMSL.

Подпрограммы ASP1d и ASP5d находят нормальное решение (т. е. решение с минимальной нормой) недоопределённых систем линейных уравнений полного ранга приведением исходной матрицы к нижнему треугольному виду методами отражения и вращения соответственно.

Процедуры dLSQRR, dLSBRR и dLQRRV библиотеки MS IMSL основаны на следующих методах [1]. Первая из них использует QR-разложение исходной матрицы, определяет базис используемого k-элементного подмножества столбцов, $k \leqslant \min(N, M)$, и далее обращается к подпрограмме dLQRSL для решения задачи $Ax = b$ методом наименьших квадратов. В случае недоопределённых систем выделенный базис состоит из k столбцов, $k \leqslant N < M$, а компоненты решения, соответствующие столбцам, не включённым в базис (и линейно зависимым от столбцов базиса), приравниваются нулю. Таким образом, мы получаем одно из возможных решений, не обязательно минимальной нормы. Получаемые подпрограммами dLSQRR и dLSBRR решения зависят от некоторого вещественного скаляра tol, содержащего неотрицательный допуск для определения подмножества столбцов матрицы A, включаемых в решение; если $\mathrm{tol} = 0$, используются все $\min(N, M)$ столбцов [1]. Вторая из подпрограмм, dLSBRR, отличается от dLSQRR итерационным уточнением корней; третья, dLQRRV, решает задачу $Ax = f$ методом наименьших квадратов, используя блочные преобразования Хаусхолдера [1].

Для совместной системы с матрицей порядка 3×5 [6, пример 134]

$$\begin{pmatrix} 1 & -3 & 2 & 5 & -9 \\ 2 & 2 & -4 & 2 & -2 \\ 1 & 2 & -3 & 0 & 1 \end{pmatrix}$$

неполного ранга 2 (здесь первая строка является линейной комбинацией второй и третьей строк с коэффициентами 2,5 и -4 соответственно) и правой частью

$$f = (10,\ 29{,}6,\ 16)^{\mathrm{T}}$$

подпрограмма ASP1d дала решение

$$x = (-1{,}207,\ -3{,}317,\ -8{,}448,\ 0{,}923,\ -1{,}504)^{\mathrm{T}}$$

с $\|x\|^2 = 86{,}948$, подпрограмма ASP5d дала решение

$$x = (5{,}800,\ -6{,}127,\ -7{,}987,\ -2{,}354,\ -1{,}507)^{\mathrm{T}}$$

с $\|x\|^2 = 142{,}786$, представленный же алгоритм даёт решение

$$x = (1{,}850,\ 2{,}021,\ -3{,}871,\ 1{,}678,\ -1{,}507)^{\mathrm{T}}$$

с квадратом нормы 27,586. Для неё же подпрограммы dLSQRR и dLSBRR при значениях $\mathrm{tol} = 10^{-10}$ и $\mathrm{tol} = 10^{-4}$ дали одинаковые решения

$$x = (0,\ 0,\ -6{,}16,\ 0,\ -2{,}48)^{\mathrm{T}}$$

с $\|x\|^2 = 44{,}096$, при $\mathrm{tol} = 0$ — различные решения

$$x_1 = (0,\ 16{,}0,\ 3{,}44,\ 0,\ -5{,}08)^{\mathrm{T}}, \quad x_2 = (0,\ 22{,}636,\ 7{,}422,\ 0,\ -7{,}007)^{\mathrm{T}}$$

соответственно с $\|x_1\|^2 = 303{,}096$ и $\|x_2\|^2 = 616{,}597$, подпрограмма dLQRRV — решение

$$x = (13{,}6,\ 1{,}2,\ 0,\ 0,\ 0)$$

с $\|x\|^2 = 186{,}4$. Квадраты невязки $\|Ax - f\|^2$ везде здесь не более 10^{-28}. Видно, что все шесть тестируемых процедур дали решение, но лишь предлагаемый алгоритм — решение с минимальной нормой. Далее для совместных систем значения невязок приводить не будем.

Для совместной системы с расширенной матрицей порядка 6×11 неполного строчного ранга 4

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & -1 & 45 & -9 & 25 & -17 & 8 \\ 2 & 3 & 4 & 5 & 2 & 3 & 7 & -53 & -11 & 31 & 13 \\ 5 & 2 & 9 & 2 & -3 & 7 & 4 & 42 & 6 & -3 & 72 \\ 2 & 2 & -4 & 2 & -2 & 29 & 11 & 4 & -8 & 42 & 18 \\ -6 & -4 & -24 & -6 & 6 & 36 & 125 & 151 & 97 & -88 & -64 \\ -5 & 5 & -33 & 7 & -1 & 77 & -48 & -255 & -125 & 262 & -139 \end{pmatrix},$$

в которой первые четыре линейно независимые строки составлены случайно, две последние являются линейными комбинациями предыдущих четырёх с коэффициентами 3, −4, −1, 2 и −2, 3, −3, 3 соответственно, подпрограмма ASP1d дала решение

$$x = (2{,}358,\ 1{,}755,\ 3{,}858,\ -0{,}773,\ -0{,}535,\ -0{,}550,\ -0{,}251,\ 0{,}537,\ 1{,}334,\ 1{,}260)^{\mathrm{T}}$$

с $\|x\|^2 = 28{,}428$, подпрограмма ASP5d дала решение

$$x = (4{,}187,\ 2{,}056,\ 2{,}878,\ 1{,}042,\ -1{,}675,\ 0{,}412,\ -1{,}798,\ -0{,}033,\ 3{,}870,\ 1{,}203)^{\mathrm{T}}$$

с $\|x\|^2 = 53{,}765$, тогда как предложенный алгоритм даёт решение

$$x = (1{,}522,\ 0{,}876,\ 3{,}651,\ 1{,}231,\ -0{,}304,\ -0{,}732,\ 0{,}340,\ 0{,}801,\ -0{,}123,\ 0{,}905)^{\mathrm{T}}$$

с $\|x\|^2 = 20{,}1461$. Для неё же подпрограммы dLSQRR и dLSBRR при значениях $\mathrm{tol} = 10^{-10}$ и $\mathrm{tol} = 10^{-4}$ дали решение

$$x = (0,\ 0,\ 5{,}157,\ 0,\ 0,\ 0,\ 0{,}261,\ 0{,}641,\ 0,\ 0{,}790)^{\mathrm{T}}$$

с $\|x\|^2 = 27{,}703$, при $\mathrm{tol} = 0$ — одинаковые решения

$$x = (0,\ 0,\ 10{,}098,\ 0,\ 0,\ 12{,}370,\ 5{,}688,\ -1{,}160,\ -19{,}361,\ -12{,}218)^{\mathrm{T}}$$

с $\|x\|^2 = 812{,}816$, dLQRRV — решение

$$x = (25{,}034,\ -93{,}359,\ 0{,}752,\ 49{,}234,\ -5{,}570,\ 1{,}657,\ 0,\ 0,\ 0,\ 0)^{\mathrm{T}}$$

с квадратом решения 11800,8. Здесь, как и в предыдущем примере, наименьшее по норме решение дал предлагаемый в настоящей статье алгоритм.

Для несовместной системы, получаемой из приведённой выше матрицы порядка 3×5 ранга 2 [6, пример 134] заменой правой части на

$$f = (10,\ 11,\ 16)^{\mathrm{T}},$$

подпрограммы ASP1d и ASP5d дают решения с квадратами норм порядка 10^{32} и квадратами невязок соответственно 285,0 и 370,185, тогда как предложенный алгоритм — решение

$$x = (1{,}269,\ 1{,}274,\ -2{,}543,\ 1{,}263,\ -1{,}258)^{\mathrm{T}}$$

с $\|x\|^2 = 12{,}878$ и $\|Ax - f\|^2 = 120{,}275$. При этом для приведения системы (1) к виду (2) применялись прямые исключения Гаусса. Подпрограммы dLSQRR и dLSBRR для неё же при $\mathrm{tol} = 10^{-10}$ и $\mathrm{tol} = 10^{-4}$ дали решение

$$(0,\ 0,\ -3{,}2,\ 0,\ -1{,}6)^{\mathrm{T}}$$

с квадратами норм и невязки 12,8 и 93,0 соответственно, при значении $\mathrm{tol} = 0$ — одинаковые решения с квадратом нормы порядка 10^{32} и квадратом невязки 1629,0, dLQRRV — решение с квадратом порядка 10^{32} и квадратом невязки 413,0. При использовании вращений слева вместо прямых исключений Гаусса в процессе приведения (1) к виду (2) (QR-разложение недоопределённой матрицы) предлагаемый алгоритм даёт для этой системы нормальное псевдорешение [6]

$$x = (1,\ 1,\ -2,\ 1,\ -1)^{\mathrm{T}}$$

с ошибкой в последнем (15-м) значащем знаке с $\|x\|^2 = 8$ и $\|Ax - b\|^2 = 93$.

Приведём также результаты решения системы № 115 из [6] с полноранговой (т. е. $\operatorname{rank} A = \min(N, M)$) матрицей размера 3×4

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 0 & 0 & 3 \\ -1 & 0 & 2 & 1 \end{pmatrix}$$

и правой частью

$$(-1,\ 3,\ 1)^{\mathrm{T}}$$

указанными выше процедурами. Подпрограммы ASP1d, ASP5d и предлагаемый алгоритм дали нормальное решение

$$x = (0{,}125,\ -0{,}625,\ 0{,}125,\ 0{,}875)^{\mathrm{T}}$$

с квадратом решения $\|x\|^2 = 1{,}1875$, тогда как dLSQRR и dLSBRR при значениях $\mathrm{tol} = 0$, $\mathrm{tol} = 10^{-10}$ и $\mathrm{tol} = 10^{-4}$ — одинаковые решения

$$(-0{,}5,\ 0,\ -0{,}5,\ 1{,}5)^{\mathrm{T}}$$

со значением $\|x\|^2 = 2{,}75$, dLQRRV — решение

$$(1,\ -1{,}5,\ 1,\ 0)^{\mathrm{T}}$$

с $\|x\|^2 = 4{,}25$.

Отметим, что решения совместных систем неполного ранга, даваемые как подпрограммами ASP5d и dLQRRV, так и dLSQRR и dLSBRR (при значении допуска $\mathrm{tol} = 0$), зависят от перестановок строк исходной матрицы, чего лишён предлагаемый алгоритм.

Обратим внимание на то, что в [1] не приведены примеры численного решения систем с недоопределённой матрицей.

4. Замечание об обусловленности матриц

Недоопределённые матрицы полного строчного ранга (как и матрицы неполного строчного ранга со строгой линейной зависимостью некоторых строк) плохообусловленными [3] считать не следует. В частности, линейное уравнение

$$\sum_{i=1}^{M} a_i x_i = f$$

решается предлагаемым алгоритмом, хотя это же нормальное решение можно выписать и явно:

$$x_i = f\mu^2 a_i, \tag{4}$$

где $\mu = \|a\|^{-1}$ — нормирующий множитель, $a = (a_1, a_2, \ldots, a_M)$. Так, для уравнения

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 0{,}00001 \end{pmatrix} x = 2$$

процедуры ASP1d, ASP5d и предложенный алгоритм дают одинаковое решение

$$x = (0{,}06666667,\ 0{,}13333333,\ 0{,}2,\ 0{,}26666667,\ 0{,}00000067)^{\mathrm{T}}$$

с $\|x\|^2 = 0{,}13333333$, совпадающее с решением, получаемым по (4), подпрограммы же dLSQRR и dLSBRR при значениях $\mathrm{tol} = 0$, $\mathrm{tol} = 10^{-10}$ и $\mathrm{tol} = 10^{-4}$ — одинаковое (очевидное для данных алгоритмов) решение

$$x = (0,\ 0,\ 0,\ 0{,}5,\ 0)^{\mathrm{T}}$$

со значением $\|x\|^2 = 0{,}25$, dLQRRV же — тривиальное решение с квадратом невязки, равным 4. Под плохой обусловленностью имеет смысл понимать «почти линейную зависимость», под числом обусловленности — отношение самого большого сингулярного числа к самому малому из не равных нулю. В общем случае для матрицы $N \times M$ ранга $K \leqslant \min(N, M)$ число ненулевых сингулярных значений равно K, среди них могут быть близкие нулю. Число обусловленности в данном случае равно отношению максимального к минимальному из этих K чисел. Для одного линейного уравнения значение нормирующего множителя μ связано с единственным ненулевым сингулярным числом σ следующим образом: $\mu = \sigma^{-1}$, число обусловленности ненулевой строки равно единице. Сингулярное разложение строки имеет вид $(a_1, \ldots, a_M) = 1\sigma(\mu a_1, \ldots, \mu a_M)$, откуда для уравнения $\sum\limits_{i=1}^{M} a_i x_i = f$ сразу имеем решение $x = f\frac{\mu}{\sigma}a$, т. е. (4).

Таким образом, предлагаемый алгоритм находит нормальное решение (решение с минимальное нормой) систем линейных алгебраических уравнений полного строчного ранга, как и подпрограммы ASP1d и ASP5d библиотеки НИВЦ МГУ (но что не всегда делают подпрограммы dLSQRR, dLSBRR и dLQRRV библиотеки MS IMSL), и находит нормальное решение совместных систем неполного строчного ранга. Его также можно использовать для решения несовместных систем линейных уравнений.

Литература

[1] Бартенев О. В. Фортран для профессионалов. Математическая библиотека IMSL. Ч. 1. — М.: Диалог-МИФИ, 2000.

[2] Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. — М.: Наука, 1984.

[3] Голуб Дж., Ван Лоун Ч. Матричные вычисления. — М.: Мир, 1999.

[4] Деммель Дж. Вычислительная линейная алгебра. Теория и приложения. — М.: Мир, 2001.

[5] Лежнёв М. В. Минимальное решение систем линейных уравнений полного строчного ранга // Сб. тр. X Всероссийской школы-семинара «Современные проблемы математического моделирования». — Ростов-на-Дону: Изд-во РГУ, 2004. — С. 151—155.

[6] Фаддеева В. Н., Колотилина Л. Ю. Вычислительные методы линейной алгебры. Набор матриц для тестирования. — Вып. I—III. — Ленинград: ЛОМИ АН СССР, 1982.

[7] Lezhnev M. V. A diagonalization method for solution of rectangular systems of linear equations // Lectures of Instructors and Abstracts of Young Scientists of the Int. Summer School «Iterative Methods and Matrix Computations». — Rostov-on-Don: RSU, 2002. — P. 431—433.