

УДК 372.851:512.644

СИСТЕМЫ НОРМАЛЬНЫХ УРАВНЕНИЙ И СПОСОБЫ ИХ РЕШЕНИЯ В КУРСЕ ЛИНЕЙНОЙ АЛГЕБРЫ

В. В. БАРДУШКИН, А. М. РЕВЯКИН

Национальный исследовательский университет «МИЭТ»
Москва, Зеленоград, Россия

Многие задачи, связанные с измерениями, приводят к системам вида $Ax = b$, где число m уравнений больше, чем число неизвестных n . Здесь m – число измерений. Интуитивно ясно, что увеличение m позволяет точнее найти значения величин, интересующих исследователя, проводящего эти измерения. Однако, поскольку измерения проводятся приблизительно, то система $Ax = b$, как правило, несовместна.

Например, пусть известно, что $y = x_1 + x_2 t$ и для нахождения параметров x_1 и x_2 ($n = 2$) проведено m измерений (t_i, y_i) , где $i = 1, 2, \dots, m$. Тогда для определения x_1 и x_2 имеем систему

$$\begin{cases} x_1 + x_2 t_1 = y_1, \\ x_1 + x_2 t_2 = y_2, \\ \dots \\ x_1 + x_2 t_m = y_m \end{cases} \Leftrightarrow \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ 1 & t_m \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{pmatrix}.$$

Любые два из этих уравнений имеют решения, а при $m > 2$ система, как правило, несовместна.

Пусть A – матрица коэффициентов системы $Ax = b$, а b – ее вектор-столбец свободных членов. Для произвольного вектора $x = (x_1, x_2, \dots, x_n)^T$ вектор $\delta = Ax - b$ называется вектором невязки.

Под псевдорешением системы $Ax = b$ понимают любой вектор $\tilde{x}$, для которого норма вектора невязки $\|Ax - b\|$ достигает наименьшего значения. Метод наименьших квадратов основывается на использовании евклидовой нормы для вычисления длины вектора невязки $\delta = Ax - b$, т. е. минимизирует $\|\delta\| = \sqrt{\delta^T \delta}$.

Псевдорешение $\tilde{x}$ системы $Ax = b$ методом наименьших квадратов находится как решение системы нормальных уравнений $A^T A \tilde{x} = A^T b$ [1–3]. Если ранг матрицы A равен n , то нормальное псевдорешение является обычным решением квадратной системы нормальных уравнений, т. е. $\tilde{x} = (A^T A)^{-1} A^T b$.

Для изучения студентами способов решения систем нормальных уравнений авторами разработан банк матриц и вектор-столбцов свободных членов, полу-

ченных при решении задач регрессионного анализа, прогнозирования временных рядов и аппроксимации функций степенными рядами [2, 4, 5]. Число обусловленности каждой из матриц больше 100, а вектор-столбцы заданы с точностью до 10^{-4} . Далее студентам предлагается вначале округлить значения коэффициентов столбца свободных членов до 10^{-2} , затем просто «обрезать» значения коэффициентов этого же столбца, оставив только первые две значащие цифры после запятой в десятичной записи этих коэффициентов, получив две системы уравнений с разной правой частью. После этого предлагается решить две полученные системы нормальных уравнений любым известным способом (методом Гаусса, по правилу Крамера, обращением матрицы или итерационным методом). Далее для найденных псевдорешений вычислить нормы векторов невязки. Результаты всегда вызывают живой интерес и желание разобраться в существе проблемы.

Отметим, что метод наименьших квадратов часто приводит к плохо обусловленным матрицам. Если матрица A не является плохо обусловленной, можно устранить большую чувствительность решений к малым возмущениям при решении систем нормальных уравнений. Большая погрешность при решении систем нормальных уравнений $A^T A \tilde{x} = A^T b$ часто связана с тем, что число обусловленности $c(A^T A)$ матрицы $A^T A$ равняется $c^2(A)$, где $c(A)$ – число обусловленности матрицы A . Матрица $A^T A$ – симметрическая и неотрицательно определенная. Поэтому уменьшить погрешности при решении $A^T A \tilde{x} = A^T b$ можно двумя способами, воспользовавшись предварительными разложениями либо матрицы A в произведение ортогональной матрицы Q и верхней треугольной матрицы R , либо матрицы $A^T A$ в произведение $R_1^T R_1$ (разложение Холецкого), где R_1 – верхняя треугольная матрица [1–3]. Оба способа эффективны, поскольку $c(A) = c(R) = c(R_1) = c(R_1^T)$ и $c(Q) = 1$.

QR -разложение матрицы A с линейно независимыми столбцами получается без каких-либо дополнительных вычислений в процессе ортогонализации вектор-столбцов матрицы A (аналогично тому, как это делается при доказательстве теоремы о том, что во всяком n -мерном евклидовом пространстве имеются ортонормированные базисы [1–3, 6]).

Пусть $A = A_{m \times n} = Q_{m \times n} \cdot R_{n \times n}$, где $Q_{m \times n} = (q_1, q_2, \dots, q_n)$, где q_i – нормированные и попарно ортогональные вектор-столбцы. Заметим, что матрица $Q_{m \times n}$ уже не является ортогональной (она даже не квадратная), но обладает свойством, что $Q_{m \times n}^T \cdot Q_{m \times n} = E_{n \times n}$. Тогда

$$A^T A \tilde{x} = A^T b \Leftrightarrow R_{n \times n} \tilde{x} = Q_{m \times n}^T b.$$

Погрешность решения $\tilde{x}$, полученного из системы $R_{n \times n} \tilde{x} = Q_{m \times n}^T b$, опреде-

ляется условием $c(A) = c(R_{n \times n})$. Привлекательность QR -разложения объясняется тем, что умножение матрицы A или вектора b на ортогональную матрицу Q не увеличивает ошибок округления или неопределенностей. Следовательно, для нахождения решения $\tilde{x}$ требуется лишь найти произведение $Q^T b$ и выполнить «обратный» ход метода Гаусса в системе с треугольной матрицей R . Предварительно выполненное разложение матрицы A избавляет от необходимости строить матрицу $A^T A$ и решать систему нормальных уравнений, тем самым не ухудшается обусловленность системы [1–3].

Удивительно, но факт, что довольно часто норма вектора невязки решения, полученного таким способом, меньше нормы вектора невязки решения, полученного с помощью встроенной процедуры пакета MATLAB [7].

Далее, пусть имеется разложение Холецкого $A^T A = R_1^T R_1$, где R_1 – квадратная верхняя треугольная матрица порядка n с ненулевыми элементами на главной диагонали [2, 3]. Отсюда

$$A^T A \tilde{x} = A^T b \Leftrightarrow \begin{cases} R^T y = A^T b, \\ R \tilde{x} = y. \end{cases}$$

Решив две системы с треугольными матрицами, получим псевдорешение точнее, чем при решении системы нормальных уравнений.

Разложение Холецкого для симметрической положительно определенной матрицы легко получить методом неопределенных коэффициентов. Рекомендуем давать его, как один из критериев положительной определенности симметрической матрицы [2, 3].

Сравнивая нормы векторов невязки псевдорешений систем $A^T A \tilde{x} = A^T b$, $R_{n \times n} \tilde{x} = Q_{m \times n}^T b$ и $\begin{cases} R^T y = A^T b, \\ R \tilde{x} = y, \end{cases}$ студенты сами делают выводы, как им в дальнейшем решать системы нормальных уравнений. Опыт авторов показывает, что лучше пользоваться QR -разложением, но проще (с небольшой потерей точности) разложением Холецкого.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Стренг, Г. Линейная алгебра и ее применения / Г. Стренг. – М. : Мир, 1980. – 456 с.
2. Ревякин, А. М. Высшая алгебра : учеб. пособие для экон. специальностей / А. М. Ревякин. – М. : МИЭТ, 2007. – 504 с.
3. Бардушкин, В. В. Линейная алгебра для экономистов : учеб. пособие / В. В. Бардушкин, А. М. Ревякин. – М. : МИЭТ, 2019. – 252 с.
4. Бардушкина, И. В. Сборник заданий для самостоятельной работы студентов по линейной алгебре и математическому анализу / И. В. Бардушкина, А. М. Ревякин. – М. : МИЭТ, 2021. – 200 с.

5. **Ревякин, А. М.** Теория вероятностей и математическая статистика: учеб. пособие: в 2 ч. Ч. 2: Математическая статистика / А. М. Ревякин, В. В. Бардушкин, И. В. Бардушкина. – М. : МИЭТ, 2017. – 224 с.

6. **Бардушкин, В. В.** Линейная алгебра и аналитическая геометрия : учеб. пособие / В. В. Бардушкин, С. Г. Кальней, А. М. Ревякин. – М. : МИЭТ, 2018. – 268 с.

7. **Ревякин, А. М.** MATLAB для изучения вычислительных методов линейной алгебры / А. М. Ревякин // Проектирование инженерных и научных приложений в среде MATLAB : материалы V Междунар. науч. конф., Харьков, 11–13 мая 2011 г. – Харьков : БЭТ, 2011. – С. 546–607.