Матричные разложения

В вычислительных приложениях линейной алгебры нередко актуальным встает вопрос о существовании эффективных методов решения той или иной задачи. Наиболее частой задачей линейной алгебры, которая встречается в приложениях, является решение систем линейных алгебраических уравнений. Однако встречается и еще ряд полезных в приложениях задач, где также необходимо эффективно выполнять различные матричные операции. Предполагая некоторые условия на матрицу, можно построить ее представление в виде произведения двух (или трех) матриц, которые по-отдельности будут более простыми, чем исходная матрица. Рассмотрим ряд таких алгоритмов.

§1. LU-разложение

Первый класс методов разложения в произведения матриц рассмотрим именно в аспекте решения систем линейных алгебраических уравнений. Пусть дана система, записанная в матричном виде

$$Ax = b, \qquad A \in M_n(\mathbb{R}), \quad x, b \in \mathbb{R}^n,$$

где матрица A является невырожденной, т.е. $\det A \neq 0$. Согласно теореме Крамера, решение такой системы существует и единственно.

Для упрощения этой задачи в контексте уменьшения количества арифметических операций, мы могли бы предположить, что матрица A является, например, верхнетреугольной. Тогда система имеет вид

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \ldots + \alpha_{1n}x_n = b_1 \\ \alpha_{22}x_2 + \alpha_{23}x_3 + \ldots + \alpha_{2n}x_n = b_2 \\ \ldots = \ldots \\ \alpha_{n-1,n-1}x_{n-1} + \alpha_{n-1,n}x_n = b_{n-1} \\ \alpha_{nn}x_n = b_n \end{cases}$$

решение которой имеет следующий вид, если записывать его, начиная с нижней строки

$$\begin{cases} x_1 = \dfrac{1}{a_{11}}\left(b_1 - \displaystyle\sum_{i=2}^{n} \alpha_{2,i}x_i\right) \\ x_2 = \dfrac{1}{a_{22}}\left(b_2 - \displaystyle\sum_{i=3}^{n} \alpha_{2,i}x_i\right) \\ \ldots = \ldots \\ x_{n-1} = \dfrac{1}{a_{n-1,n-1}}\left(b_{n-1} - \alpha_{n-1,n}x_n\right) \\ x_n = \dfrac{b_n}{a_{nn}} \end{cases}$$

или

$$x_k = \frac{1}{a_{kk}} \left(b_k - \sum_{i=k+1}^{n} \alpha_{ki} x_i \right)$$

Иными словами, решение системы уже может быть представлено в виде конечных выражений без дополнительных вычислений. Однако, очевидно, нельзя предположить, что любая матрица системы является треугольной, но встает вопрос о существовании способы представления матрицы в виде произведения треугольных матриц.

Теорема 1.1. *Невырожденная матрица A с главными минорами, отличными от нуля, допускает разложение* ***(LU-разложение)*** *в виде*

$$A = LU,$$

где L — нижнетреугольная матрица с единицами на диагонали, а U — верхнетреугольная.

Доказательство. Воспользуемся методом Гаусса для решения системы $Ax = b$, но представим его в матричном виде. Действительно, в методе Гаусса основным преобразованием столбца для исключения элементов является добавление ведущей строки, умноженной на некоторый коэффициент. Это преобразование, применяемое к k-му столбцу можно представить в виде умножения матрицы A слева на матрицу вида

$$M^{(k)} = \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & -\mu_{k+1,k} & 1 & 0 & \dots & 0 \\ 0 & \dots & -\mu_{k+2,k} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -\mu_{n,k} & 0 & 0 & \dots & 1 \end{pmatrix},$$

где $\mu_{ik} = \frac{\alpha_{ik}}{\alpha_{kk}}$ — множители Гаусса. Внимательный читатель обратит внимание, что при выполнении элементарных преобразований именно на этот множитель домножается ведущая строка, чтобы обнулить элемент столбца, находящийся под текущим ведущим элементом.

Последовательно применим гауссовы преобразования, умножая на эти матрицы, с целью получить из матрицы верхнетреугольный вид

$$M^{(n-1)} M^{(n-2)} \dots M^{(2)} M^{(1)} A = U$$

Обозначим $L^{-1} = M^{(n-1)} M^{(n-2)} \dots M^{(2)} M^{(1)}$, и при помощи непосредственной проверки можем убедиться в том, что

$$L = \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mu_{k,1} & \dots & 1 & 0 & 0 & \dots & 0 \\ \mu_{k+1,1} & \dots & \mu_{k+1,k} & 1 & 0 & \dots & 0 \\ \mu_{k+2,1} & \dots & \mu_{k+2,k} & \mu_{k+2,k+1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mu_{n,1} & \dots & \mu_{n,k} & \mu_{n,k+1} & \mu_{n,k+2} & \dots & 1 \end{pmatrix},$$

Откуда получаем, что

$$A = LU,$$

где для получения этого представления необходимо, также как и ранее, выполнять элементарные преобразования, но при этом дополнительно запоминать множители Гаусса, которые появляются в процессе их выполнения, чтобы после приведения матрицы к верхнетреугольному виду составить нижнетреугольную матрицу L. $\square$

Замечание 1.1. Условие неравенства нулю главных миноров было необходимо для того, чтобы на каждом шаге a_{kk}, участвующее в множителях Гаусса, было ненулевым. В ином случае, если матрица невырождена, но не обладает таким свойством, всегда можно найти такую предварительную перестановку строк, что это условие может быть выполнено. Тогда получится схожее представление, которое называют LUP-разложением, где P — матрица перестановок строк.

Замечание 1.2. Естественно, данный способ нахождения LU-разложения не является единственным. Можно показать, что все основные алгоритмы LU-разложения могут быть сведены к шести вычислительным схемам, но мы не затрагиваем эти вопросы, т.к. они требуют строгих подходов в алгоритмизации, что выходит за пределы текущего курса.

Замечание 1.3. Треугольные матрицы с единицами на диагонали также называют унитреугольными матрицами. Для множества верхнетреугольных матриц такого вида принято обозначение $\mathrm{SUT}(\mathbb{K})$, а для нижнетреугольных — $\mathrm{SLT}(\mathbb{K})$. Можно также, к слову, заметить, что эти множества образуют группу относительно операции умножения. Проверьте это самостоятельно. Далее для краткости будем пользоваться этими обозначениями.

Покажем еще ряд разложений, которые напрямую следуют из LU-разложения, но для них мы просто покажем существование без конкретных вычислительных.

Теорема 1.2. *Невырожденная матрица A с главными минорами, отличными от нуля, допускает разложение* ***(LDU-разложение)*** *в виде*

$$A = LDU,$$

где $L \in \mathrm{SLT}(\mathbb{K})$, $U \in \mathrm{SUT}(\mathbb{K})$ и $D \in \mathrm{D}(\mathbb{K})$ — диагональная матрица.

Доказательство. Ранее мы показали, что матрица, обладающая требованиями из условия теоремы, допускает LU-разложение. Введем универхнетреугольную матрицу $\widetilde{U}$ такую, что $U = D\widetilde{U}$. Очевидно, что эта диагональная матрица составлена из диагональных элементов матрицы U. Связь между элементами этих равенств обеспечивается следующим равенством

$$\widetilde{u}_{ij} = \frac{u_{ij}}{u_{ii}}$$

Тогда мы имеем представление в виде

$$A = LU = LD\widetilde{U}$$

□

Теорема 1.3. *Невырожденная симметричная матрица A с главными минорами, отличными от нуля, допускает разложение (**LDL^T-разложение**) в виде*

$$A = LDL^T,$$

где $L \in \mathrm{SLT}(\mathbb{K})$ и $D \in \mathrm{D}(\mathbb{K})$ — диагональная матрица.

Доказательство. В силу предыдущей теоремы эта матрица допускает разложение вида

$$A = LDU,$$

При этом транспонирование матрицы A^T приведет к следующему виду относительно разложения

$$A^T = U^T DL^T$$

Учитывая симметричность, мы должны положить эти два разложения равными друг другу. Откуда

$$LDU = U^T DL^T, \qquad \Rightarrow \qquad U = L^T,$$

что приводит к утверждению теоремы. □

Лемма 1.1. *При LDL^T-разложении определитель матрицы A совпадает с определителем матрицы D.*

Доказательство. Действительно, любая унитреугольная матрица имеет единичный определитель, а значит

$$\det A = \det L \det D \det L^T = \det D$$

□

Теорема 1.4. *Симметричная положительно определенная матрица A, допускает разложение Холецкого (**LL^T-разложение**) в виде*

$$A = LL^T,$$

где $L \in \mathrm{LT}(\mathbb{K})$ — нижнетреугольная матрица.

Доказательство. Вновь воспользуемся результатом предыдущей теоремы, записывая разложение

$$A = LDL^T$$

Условие положительной определенности приводит к тому, что диагональные элементы d_i матрицы D должны быть строго положительны (см.предыдущую лемму и доказательство критерия Сильвестра). Следовательно мы можем ввести матрицу

$$D^{1/2} = \mathrm{diag}\{\sqrt{d_1}, \sqrt{d_2}, \dots, \sqrt{d_n}\}, \qquad \Rightarrow D = D^{1/2}D^{1/2}$$

и ввести новую матрицу $\widetilde{L} = LD^{1/2}$. Тогда в записанном разложении мы можем сделать следующие преобразования

$$A = (LD^{1/2})(LD^{1/2})^T = \widetilde{L}\widetilde{L}^T,$$

что соответствует разложению Холецкого. □

Замечание 1.4. Разложение Холецкого в ряде численных алгоритмов используется как вычислительный критерий положительной определенности матрицы. Если разложение выполнено без ошибок, то следовательно матрица была положительно определенной и, наоборот, при возникновении ошибок можно утверждать, что матрица не обладает этим свойством.

§2. QR-разложение

Данный вид разложения предполагает наличие заданного скалярного произведения, благодаря которому может быть определено условие ортогональности векторов, а также процесс ортогонализации Грама-Шмидта. Предположим, что эти необходимые условия выполняются, тогда справедлива следующая теорема. Приведем ее в случае квадратных матриц, хотя может быть построено обобщение и для прямоугольных.

Теорема 2.1. *Матрица $A \in M_n(\mathbb{K})$ может быть представлена в виде*

$$A = QR,$$

где $Q \in M_n(\mathbb{K})$ — унитарная (ортогональная) матрица, а $R \in M_n(\mathbb{K})$ — верхнетреугольная матрица. При этом, если матрица A невырождена, то QR-разложение единственно.

Доказательство. Пусть $\{a_i\}_{i=1}^n$ — система столбцов матрицы A. Наличие скалярного произведения позволяет провести ортогонализацию Грама-Шмидта. Получим с ее помощью преобразование системы столбцов матрицы в набор векторов

$$\begin{cases} q_1 = u_{11}a_1 \\ q_2 = u_{12}a_1 + u_{22}a_2 \\ \ldots \\ q_n = u_{1n}a_1 + u_{2n}a_2 + \ldots + u_{nn}a_n, \end{cases}$$

где векторы $\{q_i\}_{i=1}^n$ представляют собой ортонормированный базис, т.е. векторы, которые были дополнительно нормированы на единицу. Если система столбцов матрицы A изначально линейно независима, то набор ортонормированных векторов определяется однозначно в силу процесса ортогонализации Грама-Шмидта. Если же система $\{a_i\}_{i=1}^n$ линейно зависима, то всегда набор $\{q_i\}$ можно дополнить до базиса так, чтобы он выражался треугольным преобразованием из изначального набора столбцов.

Записанное преобразование между системами векторов можно представить в матричном виде

$$Q = AU,$$

где Q — унитарная ($Q\dagger = Q^{-1}$) матрица, если $\mathbb{K} = \mathbb{C}$ и ортогональная ($Q^T = Q^{-1}$), если $\mathbb{K} = \mathbb{R}$, а матрица U — верхнетреугольная матрица. Обозначая $R = U^{-1}$, получаем

$$A = QR$$

$\square$

Замечание 2.1. Существует множество вариаций QR-разложения для многих схожих случаев, но это остается на усмотрение читателя в любой литературе, посвященной линейной алгебре или численным методам линейной алгебры.