

Глава 4. ТРЕУГОЛЬНОЕ РАЗЛОЖЕНИЕ КВАДРАТНОЙ МАТРИЦЫ

4.1. LU-разложение

При решении различных задач матричной алгебры оказывается полезным представить заданную матрицу в виде произведения нескольких матриц специальной структуры. Один из способов *факторизации* (разложения на множители) матрицы мы сейчас рассмотрим.

Итак, пусть задана квадратная матрица A .

1-й шаг. Найдем ведущий (наибольший по модулю) элемент A . Переставляя (если нужно) строки и столбцы, поместим ведущий элемент в 1-ю строку и в 1-й столбец. Далее обнулим *поддиагональные* элементы 1-го столбца, прибавляя к m -й строке ($m = 2, \dots, n$ первую строку, умноженную на $\left(-\frac{a_{m1}}{a_{11}}\right)$).

Полученную матрицу обозначим $A^{(1)}$.

2-й шаг. Найдем ведущий элемент в подматрице, получающейся из $A^{(1)}$ вычеркиванием 1-го столбца и 1-й строки. Переставляя (если нужно) строки и столбцы, поместим ведущий элемент во 2-ю строку и во 2-й столбец. Далее обнулим поддиагональные элементы 2-го столбца, прибавляя к m -й строке ($m = 3, \dots, n$ вторую строку, умноженную на $\left(-\frac{a_{m2}}{a_{22}}\right)$).

Полученную матрицу обозначим $A^{(2)}$.

k -й шаг. Найдем ведущий элемент в подматрице, получающейся из $A^{(k-1)}$ вычеркиванием $(k-1)$ -го столбца и $(k-1)$ -й строки. Переставляя (если нужно) строки и столбцы, поместим ведущий элемент в k -ю строку и в k -й столбец. Далее обнулим поддиагональные элементы k -го столбца, прибавляя к m -й строке ($m = k+1, \dots, n$ вторую строку, умноженную на $\left(-\frac{a_{mk}}{a_{kk}}\right)$).

Если на каком-то шаге ведущий элемент окажется нулем, работа алгоритма заканчивается. Таким образом, после не более, чем $(n-1)$ шагов мы преобразуем матрицу A в верхнюю треугольную матрицу, которую принято обозначать U (от английского "Upper").

В рассмотренном выше алгоритме использовались два элементарных преобразования матрицы - прибавление к ее строке другой строки, умноженной на число, и перестановка строк (столбцов). Покажем, что эти элементарные преобразования матрицы равносильны умножению ее на матрицы специального вида.

1) Обозначим $E_{km}(\alpha)$ ($k \neq m$) квадратную матрицу, получаемую из единичной заменой нуля в k -й строке и m -м столбце числом $\alpha \neq 0$.

Пример. Матрица 4-го порядка

$$E_{31}(-2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Замечание. Если $k > m$, то $E_{km}(\alpha)$ - *нижняя* треугольная, если $k < m$ - *верхняя*.

Нетрудно убедиться, что умножив матрицу A на $E_{km}(\alpha)$ *слева*, мы прибавим к ее k -й строке m -ю строку, умноженную на α .

Пример.

$$\begin{aligned} E_{31}(-2) \cdot A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & \dots & a_{14} \\ a_{21} & \dots & a_{24} \\ a_{31} & \dots & a_{34} \\ a_{41} & \dots & a_{44} \end{bmatrix} = \\ &= \begin{bmatrix} a_{11} & \dots & a_{14} \\ a_{21} & \dots & a_{24} \\ a_{31} - 2a_{11} & \dots & a_{34} - 2a_{14} \\ a_{41} & \dots & a_{44} \end{bmatrix}. \end{aligned}$$

Заметим, что умножив матрицу A слева на $E_{km}(\alpha)$, а затем умножив полученное произведение слева на $E_{km}(-\alpha)$ (прибавив к k -й строке матрицы ее m -ю строку, умноженную на α , а затем прибавив к получившейся k -й строке ту же m -ю, умноженную на $-\alpha$), мы получим исходную матрицу, т.е.

$$E_{km}(\alpha) \cdot E_{km}(-\alpha) \cdot A = A \implies E_{km}(\alpha) \cdot E_{km}(-\alpha) = I. \quad (4.1.1)$$

Если бы можно было провести описанный выше процесс преобразования матрицы в верхнюю треугольную без выбора ведущих элементов (т.е. без перестановок строк и столбцов), то с учетом установленных свойств матриц $E_{km}(\alpha)$ можно было бы написать

$$E_{n,n-1}(\alpha_N) \cdot \dots \cdot E_{21}(\alpha_1) \cdot A = U. \quad (4.1.2)$$

Здесь U - верхняя треугольная матрица, а $N = \frac{n(n-1)}{2}$ - количество сомножителей слева от A (количество обнуляемых поддиагональных элементов матрицы n -го порядка).

Умножив (4.1.2) слева на $E_{21}(-\alpha_1) \cdot \dots \cdot E_{n,n-1}(-\alpha_N)$, получим

$$\begin{aligned} E_{21}(-\alpha_1) \cdot \dots \cdot E_{n,n-1}(-\alpha_N) \cdot E_{n,n-1}(\alpha_N) \cdot \dots \cdot E_{21}(\alpha_1) \cdot A = \\ = E_{21}(-\alpha_1) \cdot \dots \cdot E_{n,n-1}(-\alpha_N) \cdot U. \end{aligned} \quad (4.1.3)$$

В силу (4.1.1) $E_{21}(-\alpha_1) \cdot \dots \cdot E_{n,n-1}(-\alpha_N) \cdot E_{n,n-1}(\alpha_N) \cdot \dots \cdot E_{21}(\alpha_1) = I$.

Слева от матрицы U стоит произведение нижних треугольных матриц с единичными диагоналями, представляющее собой (как нетрудно показать) также нижнюю треугольную матрицу с единичной диагональю. Обозначив ее L (от английского "Lower"), перепишем (4.1.3) в виде

$$A = LU.$$

Однако в реальном алгоритме присутствуют еще перестановки строк и столбцов матрицы. Введем *матрицу элементарных перестановок* Π_{km} , получающуюся из единичной перемещением единицы, стоящей в k -й строке в m -й столбец, а единицы, стоящей в m -й строке, - в k -й столбец.

Пример.

$$\begin{aligned} \Pi_{13} \cdot A &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b & c & d \\ * & * & * & * \\ p & q & r & s \\ * & * & * & * \end{bmatrix} = \\ &= \begin{bmatrix} p & q & r & s \\ * & * & * & * \\ a & b & c & d \\ * & * & * & * \end{bmatrix} \\ A \cdot \Pi_{13} &= \begin{bmatrix} a & * & p & * \\ b & * & q & * \\ c & * & r & * \\ d & * & s & * \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} p & * & a & * \\ q & * & b & * \\ r & * & c & * \\ s & * & d & * \end{bmatrix}. \end{aligned}$$

Здесь звездочками обозначены элементы матриц, не меняющие своего положения.

Видно, что умножая матрицу A на Π_{km} *слева*, мы переставляем в A k -ю и m -ю *строки*, а умножая ее на Π_{km} *справа*, переставляем k -й и m -й *столбцы*.

Отметим свойства матрицы элементарных перестановок.

1. В каждой ее строке и каждом столбце ровно одна единица; остальные элементы - нули.
2. $\Pi_{km} \cdot \Pi_{km} = I$ (дважды поменяв местами одну и ту же пару столбцов или строк, получаем исходную матрицу).
3. $\det(\Pi_{km}) = -1$ (перестановка двух строк или двух столбцов матрицы приводит к умножению ее определителя на (-1)).

Произведение нескольких матриц элементарных перестановок есть матрица, меняющая местами уже несколько пар столбцов (строк). Такая матрица называется *матрицей перестановок* и обладает свойством 1. Ее определитель равен либо $(+1)$, либо (-1) . Матрица, обратная матрице перестановок, также есть матрица перестановок.

Теперь может быть сформулирована общая

Теорема. Для каждой квадратной матрицы A существуют такие матрицы перестановок Π_1 и Π_2 , что

$$\Pi_1 \cdot A \cdot \Pi_2 = L \cdot U, \quad (4.1.4)$$

где L – нижняя треугольная матрица с единичной диагональю, а U – верхняя треугольная матрица.

Формула (4.1.4) называется *LU-разложением* или *треугольным разложением* матрицы A .

Доказательство мы опускаем из-за его технической сложности. Отметим, что Π_1 и Π_2 (а, следовательно, L и U) определены не единственным образом. Это следует, например, из того, что при выборе ведущего элемента может встретиться несколько равных по модулю наибольших элементов и, отдав предпочтение одному из них, мы получим одно из возможных LU -разложений матрицы.

4.2. Некоторые применения LU -разложения

1. Вычисление определителей. Используя формулу (4.1.4) и свойства определителей, мы можем написать:

$$\begin{aligned} \det(A) &= \det(\Pi_1^{-1} \cdot L \cdot U \cdot \Pi_2^{-1}) = \det(\Pi_1^{-1}) \cdot \det(L) \cdot \det(U) \cdot \det(\Pi_2^{-1}) = \\ &= (\pm 1) \cdot u_{11} \cdot \dots \cdot u_{nn} \end{aligned} \quad (4.2.1)$$

(определители *треугольных* матриц L и U равны произведениям их диагональных элементов, а знак (+) или (-) в правой части зависит от количества элементарных перестановок при выполнении LU -разложения матрицы).

В п.3.1 было показано, что вычисление определителя матрицы "по определению" невозможно, поскольку требуемое для его реализации количество операций превышает $n!$ (n – порядок матрицы). Можно показать, что LU -разложение требует порядка $\frac{n^3}{3}$ операций. При $n = 30$ это составит около $3 \cdot 10^4$ (сравните с полученной ранее оценкой $3 \cdot 10^{23}$ для вычисления "по определению").

Технологию рассмотрим на примере так называемой матрицы Вандермонда¹

$$Z = \begin{bmatrix} 1 & z_1 & z_1^2 & \dots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \dots & z_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & z_n & z_n^2 & \dots & z_n^{n-1} \end{bmatrix}.$$

Обозначим определитель матрицы Вандермонда $V(z_1, z_2, \dots, z_n)$. Вычтем из n -го столбца матрицы ее $(n-1)$ -й столбец, умноженный на z_1 ; вычтем из $(n-1)$ -го столбца матрицы ее $(n-2)$ -й столбец, умноженный на z_1 ; ...; вычтем из 2-го столбца матрицы ее 1-й столбец, умноженный на z_1 . В силу свойства (4) определитель не изменится:

$$V(z_1, z_2, \dots, z_n) = \det \left(\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & z_2 - z_1 & z_2(z_2 - z_1) & \dots & z_2^{n-2}(z_2 - z_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & z_n - z_1 & z_n(z_n - z_1) & \dots & z_n^{n-2}(z_n - z_1) \end{bmatrix} \right).$$

Разложим определитель по 1-й строке:

$$V(z_1, z_2, \dots, z_n) = 1 \cdot \det \left(\begin{bmatrix} z_2 - z_1 & z_2(z_2 - z_1) & \dots & z_2^{n-2}(z_2 - z_1) \\ \dots & \dots & \dots & \dots \\ z_n - z_1 & z_n(z_n - z_1) & \dots & z_n^{n-2}(z_n - z_1) \end{bmatrix} \right).$$

В силу свойства (3) можно вынести за знак определителя общие множители строк $(z_2 - z_1), \dots, (z_n - z_1)$.

$$V(z_1, z_2, \dots, z_n) = (z_2 - z_1) \cdot \dots \cdot (z_n - z_1) \cdot \det \left(\begin{bmatrix} 1 & z_2 & \dots & z_2^{n-2} \\ \dots & \dots & \dots & \dots \\ 1 & z_n & \dots & z_n^{n-2} \end{bmatrix} \right) = V(z_2, \dots, z_n).$$

Повторяя эту операцию, понижающую порядок матрицы, в конце концов получим

$$\begin{aligned} V(z_1, z_2, \dots, z_n) &= \\ &= ((z_2 - z_1) \cdot \dots \cdot (z_n - z_1)) \cdot ((z_3 - z_2) \cdot \dots \cdot (z_n - z_2)) \cdot ((z_n - z_{n-1})) = \\ &= \prod_{m > k} (z_m - z_k). \end{aligned}$$

Заметим, что если z_1, z_2, \dots, z_n – *попарно различные* числа, то определитель матрицы Вандермонда *отличен от нуля*.

¹Александр Теофил ВАНДЕРМОНД (1735-1796) – французский математик, член Парижской АН.

2. Решение систем линейных уравнений. Если получено LU -разложение матрицы A ($\Pi_1 \cdot A \cdot \Pi_2 = L \cdot U$), то решение системы $Ax = b$ сводится к последовательному решению двух систем с треугольными матрицами:

$$Ly = \Pi_1 \cdot b, \quad U(\Pi_2 \cdot x) = y. \quad (4.2.2)$$

Решение системы с треугольной матрицей коэффициентов требует выполнения порядка n^2 операций (против n^3 для системы с заполненной матрицей коэффициентов). Правда, затраты времени на LU -разложение сравнимы с временем решения системы с заполненной матрицей и, на первый взгляд, выгоды не видно. Однако в приложениях весьма часто приходится решать несколько систем с одной и той же матрицей (см., например, гл.14). В этом случае факторизация производится один раз, что при матрицах большого порядка дает весьма значительный выигрыш во времени.

Отметим еще, что иногда вместо LU -разложения используют так называемое LDU -разложение матрицы.

Если A – невырожденная матрица, то в формуле (4.1.4) U – также невырожденная матрица. Положим $D = \text{diag}[u_{11}, \dots, u_{nn}]$ и $\tilde{U} = D^{-1}U$. Тогда \tilde{U} – верхняя треугольная матрица с единичной диагональю, и справедливо соотношение

$$\Pi_1 \cdot A \cdot \Pi_2 = LD\tilde{U},$$

которое и называется LDU -разложением матрицы A .