Тема IV. Квадратичные формы

§ 2. Критерий Сильвестра

М.В.Волков

Уральский федеральный университет
Институт естественных наук и математики
кафедра алгебры и фундаментальной информатики

2020/2021 учебный год

Определение

Квадратичная форма над $\mathbb{R}$, которая положительна при любом ненулевом наборе значений переменных, называется *положительно определенной*.
Квадратичная форма над $\mathbb{R}$, которая отрицательна при любом ненулевом наборе значений переменных, называется *отрицательно определенной*.

Из закона инерции следует, что форма $q(x_1, x_2, \ldots, x_n)$ положительно определена, если и только если она приводится к каноническому виду

$$\alpha_1 y_1^2 + \alpha_2 y_2^2 + \cdots + \alpha_n y_n^2, \qquad (*)$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n > 0$. Аналогично, форма $q(x_1, x_2, \ldots, x_n)$ отрицательно определена тогда и только тогда, когда она приводится к каноническому виду $(*)$ с $\alpha_1, \alpha_2, \ldots, \alpha_n < 0$. Поэтому положительную/отрицательную определенность можно распознать, приведя форму к каноническому виду.

Однако иногда удобны условия положительной/отрицательной определенности, выраженные в терминах матрицы исходной формы $q(x_1, x_2, \ldots, x_n)$.

Мы выведем критерий положительной определенности формы из одного полезного разложения квадратных матриц над произвольными полями. Нам понадобятся два новых определения.

Определение

Верхнетреугольная (нижнетреугольная) матрица называется *верхней (нижней) унитреугольной*, если элементы ее главной диагонали равны 1.

Типичная верхняя унитреугольная матрица:
$$\begin{pmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 1 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Отметим, что определитель любой унитреугольной матрицы равен 1.

Определение

Миноры $n \times n$-матрицы, расположенные в ее первых k строках и первых k столбцах ($k = 1, 2, \dots, n$) называются *угловыми минорами*.

k-й угловой минор обозначим Δ_k. Если $A = (a_{ij})$, то $\Delta_1 = a_{11}$ и $\Delta_n = |A|$.

Теорема (LDU-разложение, Тадеуш Банахевич, 1938)

Квадратная матрица $A = (a_{ij})$ над произвольным полем, все угловые миноры которой отличны от 0, однозначно представима в виде $A = LDU$, где матрица L нижняя унитреугольная, матрица D диагональная, а матрица U верхняя унитреугольная.

Доказательство. Обозначим через A_k матрицу на пересечении первых k строк и первых k столбцов матрицы A, т.е. ту матрицу, определителем которой служит угловой минор Δ_k. Индукцией по k докажем, что она однозначно представима в виде $A_k = L_k D_k U_k$, где L_k – нижняя унитреугольная, D_k – диагональная, а U_k – верхняя унитреугольная $k \times k$-матрицы. При k, равном размеру A, получим утверждение теоремы.

База индукции. При $k = 1$ имеем $A_1 = (a_{11})$, и единственно возможное LDU-разложение для A_1 есть $(a_{11}) = (1) \cdot (a_{11}) \cdot (1)$.

Шаг индукции. Допустим, что доказываемое утверждение верно для матрицы A_k, и проверим, что тогда оно верно и для матрицы A_{k+1}.

Матрицу A_{k+1} запишем в виде $A_{k+1} = \begin{pmatrix} A_k & \mathbf{v}^T \\ \mathbf{u} & a_{k+1\,k+1} \end{pmatrix}$,

где $\mathbf{u} = (a_{k+1\,1}, \dots, a_{k+1\,k})$, а $\mathbf{v} = (a_{1\,k+1}, \dots, a_{k\,k+1})$.

Будем искать представление $A_{k+1} = L_{k+1}D_{k+1}U_{k+1}$, где L_{k+1} – нижняя унитреугольная, D_{k+1} – диагональная, а U_{k+1} – верхняя унитреугольная $(k+1) \times (k+1)$-матрицы. Разобьём каждый множитель на четыре блока:

- $L_{k+1} = \begin{pmatrix} L & O_k^T \\ \mathbf{x} & 1 \end{pmatrix}$, где L – нижняя унитреугольная $k \times k$-матрица,

$\mathbf{x}$ – строка длины k, O_k^T – нулевой столбец высоты k;

- $U_{k+1} = \begin{pmatrix} U & \mathbf{y}^T \\ O_k & 1 \end{pmatrix}$, где U – верхняя унитреугольная $k \times k$-матрица,

$\mathbf{y}^T$ – столбец высоты k, O_k – нулевая строка длины k;

- $D_{k+1} = \begin{pmatrix} D & O_k^T \\ O_k & d_{k+1} \end{pmatrix}$, где D – диагональная $k \times k$-матрица.

Перемножая блочные матрицы, получаем

$$\begin{pmatrix} A_k & \mathbf{v}^T \\ \mathbf{u} & a_{k+1\,k+1} \end{pmatrix} = \begin{pmatrix} L & O_k^T \\ \mathbf{x} & 1 \end{pmatrix} \begin{pmatrix} D & O_k^T \\ O_k & d_{k+1} \end{pmatrix} \begin{pmatrix} U & \mathbf{y}^T \\ O_k & 1 \end{pmatrix} = \begin{pmatrix} LDU & LD\mathbf{y}^T \\ \mathbf{x}DU & \mathbf{x}D\mathbf{y}^T + d_{k+1} \end{pmatrix}.$$

Приравнивая северо-западные блоки, заключаем, что $A_k = LDU$.
По предположению индукции матрица A_k **однозначно** представима как $A_k = L_kD_kU_k$, где L_k и U_k – нижняя и верхняя унитреугольные, а D_k – диагональная $k \times k$-матрицы. Отсюда $L = L_k$, $U = U_k$ и $D = D_k$.

Итак,

$$\begin{pmatrix} A_k & \mathbf{v}^T \\ \mathbf{u} & a_{k+1\,k+1} \end{pmatrix} = \begin{pmatrix} L_k D_k U_k & L_k D_k \mathbf{y}^T \\ \mathbf{x} D_k U_k & \mathbf{x} D_k \mathbf{y}^T + d_{k+1} \end{pmatrix}.$$

Приравняв северо-восточные блоки, получим равенство $L_k D_k \mathbf{y}^T = \mathbf{v}^T$, которое дает систему линейных уравнений для координат столбца $\mathbf{y}^T$. Определитель этой системы $\det L_k D_k = \det L_k D_k \det U_k$ (так как $\det U_k = 1$) $= \det L_k D_k U_k = \det A_k = \Delta_k \neq 0$ по условию теоремы. Значит, система крамеровская, и столбец $\mathbf{y}^T$ существует и единствен. Аналогично, из равенства $\mathbf{x} D_k U_k = \mathbf{u}$ однозначно определяется строка $\mathbf{x}$. Наконец, зная $\mathbf{x}$ и $\mathbf{y}^T$, из равенства $\mathbf{x} D_k \mathbf{y}^T + d_{k+1} = a_{k+1\,k+1}$ однозначно определяем элемент d_{k+1}. Итак, $(k+1) \times (k+1)$-матрицы L_{k+1}, D_{k+1} и U_{k+1}, такие, что L_{k+1} – нижняя унитреугольная, D_{k+1} – диагональная, а U_{k+1} – верхняя унитреугольная и $A_{k+1} = L_{k+1} D_{k+1} U_{k+1}$, существуют и однозначно определяются по матрице A_{k+1}. $\square$

Следствие (LDU-разложение симметрической матрицы)

Если A – симметрическая матрица, все угловые миноры которой отличны от 0, и $A = LDU$ – ее LDU*-разложение, то $L = U^T$.*

Доказательство. Транспонируя обе части равенства $A = LDU$, получим $A = U^T DL^T$ (поскольку $A = A^T$ и $D = D^T$). Матрица U^T нижняя унитреугольная, а матрица L^T верхняя унитреугольная, т.е. равенство $A = U^T DL^T$ есть LDU-разложение матрицы A. Но LDU-разложение единственно, откуда $L = U^T$. □

Комментарии. Доказательство теоремы о LDU-разложении конструктивно и позволяет строить такие разложения. В действительности, можно проверить, что для симметрической матрицы условие, что все угловые миноры отличны от 0, означает в точности то, что при приведении квадратичной формы с этой матрицей к каноническому виду методом Лагранжа всегда встретится только первый случай (и потому приведение сводится к последовательному выделению полных квадратов).

Метод Лагранжа в этой ситуации дает LDU-разложение. В примере на метод Лагранжа из прошлой лекции встретилась именно такая ситуация.

Мы привели форму

$$f = 4x_1^2 + 2x_2^2 + x_3^2 + x_4^2 - 4x_1x_2 - 4x_1x_3 + 4x_1x_4 + 4x_2x_3 - 4x_3x_4.$$

к каноническому виду $4y_1^2 + y_2^2 - y_3^2 + 3y_4^2$, и результирующая замена выглядела так:

$$\begin{cases} y_1 = x_1 - \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_4 \\ y_2 = x_2 + x_3 + x_4 \\ y_3 = x_3 + 2x_4 \\ y_4 = x_4 \end{cases}$$

Видно, что матрица этой замены верхняя унитреугольная и по существу мы построили LDU-разложение матрицы формы f:

$$\begin{pmatrix} 4 & -2 & -2 & 2 \\ -2 & 2 & 2 & 0 \\ -2 & 2 & 1 & -2 \\ 2 & 0 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 1 & 0 \\ \frac{1}{2} & 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Теорема (Сильвестр, 1852)

Квадратичная форма над $\mathbb{R}$ положительно определена тогда и только тогда, когда все угловые миноры ее матрицы положительны.

Доказательство. Необходимость. Пусть форма $q(x_1, x_2, \dots, x_n) = X^T AX$ положительно определена. Тогда из нее невырожденной линейной заменой переменных можно получить форму

$$\alpha_1 y_1^2 + \alpha_2 y_2^2 + \dots + \alpha_n y_n^2, \tag{$*$}$$

где $\alpha_1, \alpha_2, \dots, \alpha_n > 0$. Матрица формы $(*)$ диагональна и ее определитель равен $\alpha_1\alpha_2 \dots \alpha_n > 0$. На прошлой лекции мы отмечали, что если форма g получена из формы q невырожденной линейной заменой переменных, то определители матриц форм q и g либо оба положительны, либо оба отрицательны, либо оба равны 0. В нашем случае определитель матрицы формы $(*)$ положителен, откуда и определитель $|A| = \Delta_n$ положителен.

Осталось заметить, что если форма $q(x_1, x_2, \dots, x_n)$ положительно определена, то такова и форма от k переменных $q(x_1, x_2, \dots, x_k, \underbrace{0, \dots, 0}_{n-k})$.

Матрица этой формы есть A_k, откуда $|A_k| = \Delta_k > 0$.

Достаточность. Предположим, что $\Delta_1, \Delta_2, \ldots, \Delta_n > 0$. По следствию о LDU-разложении симметрической матрицы $A = U^T DU$ для некоторой пары из верхней унитреугольной матрицы U и диагональной матрицы D. Сделаем в форме $q = X^T AX$ замену переменных $Y = UX$. Тогда $q = X^T AX = X^T U^T DUX = (UX)^T DUX = YDY$, т.е. эта замена приводит форму к каноническому виду

$$\delta_1 y_1^2 + \delta_2 y_2^2 + \cdots + \delta_n y_n^2,$$

где $D = \begin{pmatrix} \delta_1 & 0 & \ldots & 0 \\ 0 & \delta_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \delta_n \end{pmatrix}$. Поскольку $A_k = U_k^T D_k U_k$ и определители унитреугольных матриц равны 1, имеем $\Delta_k = |A_k| = |D_k| = \delta_1\delta_2 \cdots \delta_k$. Отсюда $\delta_1 = \Delta_1 > 0$ и $\delta_i = \dfrac{\Delta_i}{\Delta_{i-1}} > 0$ для всех $i = 2, \ldots, n$. Поэтому форма q положительно определена. □

Понятно, что форма q отрицательно определена тогда и только тогда, когда форма $-q$ положительно определена. Поэтому из доказательства критерия Сильвестра немедленно следует критерий отрицательной определенности:

Следствие

Квадратичная форма над $\mathbb{R}$ отрицательно определена, если и только если знаки угловых миноров ее матрицы чередуются, причем $\Delta_1 < 0$.

Упражнение. Докажите *критерий Якоби*: *квадратичная форма над $\mathbb{R}$ положительно определена, если и только если у характеристического многочлена ее матрицы все коэффициенты ненулевые и их знаки чередуются.*