ГЛАВА 14. КВАДРАТИЧНЫЕ ФОРМЫ

§1. КАНОНИЧЕСКИЙ ВИД КВАДРАТИЧНОЙ ФОРМЫ

Квадратичной формой называют вещественную функцию F от n вещественных переменных x_1, x_2, $\ldots$, x_n вида

$$F(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j.$$

Заданные вещественные числа a_{ij} называют ***коэффициентами*** квадратичной формы.

Коэффициенты квадратичной формы можно считать удовлетворяющими условиям симметрии

$$a_{ij} = a_{ji}, \quad i,\ j = 1,\ 2,\ \ldots,\ n,$$

поскольку, слагаемые в квадратичной форме, содержащие коэффициенты a_{ij}, a_{ji}, можно записать так:

$$a_{ij}x_ix_j + a_{ji}x_jx_i = \frac{a_{ij} + a_{ji}}{2}x_ix_j + \frac{a_{ij} + a_{ji}}{2}x_jx_i.$$

Запишем квадратичную форму в более компактном виде. Пусть A есть симметричная матрица с элементами

$$a_{ij}, \quad i,\ j = 1,\ 2,\ \ldots\ n.$$

Вектор

$$x = (x_1,\ x_2,\ \ldots,\ x_n)$$

будем считать элементом пространства $\mathbb{R}^n$. Тогда

$$F(x) = (Ax, x).$$

Здесь и всюду на протяжении данной главы скобки обозначают стандартное скалярное произведение в пространстве $\mathbb{R}^n$.

.

Пусть в квадратичной форме $F(x) = (Ax, x)$ выполнена замена переменных, т. е. введены новые переменные $y = (y_1, y_2, \ldots, y_n)$, связанные со старыми переменными $x = (x_1, x_2, \ldots, x_n)$ соотношением

$$x = Qy,$$

где Q — невырожденная ***матрица преобразования переменных***.

.

Выполнив замену переменных $x = Qy$, получим

$$(Ax, x) = (AQy, Qy) = (Q^T AQy, y) = (By, y) = \sum_{i,j=1}^{n} b_{ij} y_i y_j,$$

где,

$$B = Q^T AQ.$$

.

Говорят, что преобразование переменных

$$x = Qy,$$

приводит квадратичную форму к ***канонческому виду***, если матрица

$$B = Q^T AQ$$

диагональна, т. е.

$$(Ax, x) = \sum_{i=1}^{n} b_{ii} y_i^2.$$

Можно сказать также, что форма ***приведена к сумме квадратов***.

.

Всякую квадратичную форму невырожденным преобразованием переменных можно привести к каноническому виду.

Действительно, поскольку A — симметричная матрица, существует ортогональная матрица Q такая, что

$$Q^T A Q = \Lambda,$$

где Λ — диагональная матрица, по диагонали которой расположены все собственные числа матрицы A.

.

При таком выборе матрицы Q преобразование переменных приводит квадратичную форму к виду

$$(Ax, x) = \sum_{i=1}^{n} \lambda_i y_i^2,$$

где λ_1, λ_2, $\ldots$, λ_n — собственные числа матрицы A.

Существуют и другие способы приведения квадратичной формы к каноническому виду. Опишем, например, *метод Лагранжа*, или *метод выделения полных квадратов*, приведения квадратичной формы к каноническому виду.

В ходе описания этого метода, фактически, будет дано еще одно, независимое, доказательство возможности приведения любой квадратичной формы к каноническому виду.

.

Будем различать два случая: 1) в квадратичной форме

$$F(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j.$$

коэффициент при квадрате какой-либо переменной отличен от нуля, 2) коэффициенты при квадратах всех переменных — нули.

.

Рассмотрим сначала первый случай и пусть

$$a_{11} \neq 0.$$

Если это не так, придется ввести другую нумерацию неизвестных.

.

Запишем квадратичную форму

$$F(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

в виде

$$F = a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^2 + G,$$

где

$$G = F - a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^2.$$

Нетрудно убедиться, что G не содержит x_1, а является квадратичной формой только от переменных $x_2, x_3, \ldots, x_n$.

.

Положим

$$y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \quad y_2 = x_2, \quad \ldots, \quad y_n = x_n.$$

Тогда

$$F = a_{11}^{-1}y_1^2 + G(y_2, \ldots, y_n),$$

где $G(y_2, \ldots, y_n)$ — квадратичная форма от переменных $y_2, \ldots, y_n$.

.

Матрица замены переменных

$$y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \quad y_2 = x_2, \quad \ldots, \quad y_n = x_n.$$

невырождена, так как она имеет определитель, равный a_{11}, а по предположению $a_{11} \neq 0$.

Пусть теперь все коэффициенты при квадратах переменных в

$$F(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

равны нулю. Тогда будем считать, что хотя бы один коэффициент при произведениях переменных отличен от нуля, иначе квадратичная форма тождественно равна нулю и она имеет ***тривиальный канонический вид***: все коэффициенты при квадратах неизвестных равны нулю.

.

Итак, примем для определенности, что

$$a_{12} \neq 0,$$

и выполним преобразование переменных по формулам

$$x_1 = z_1 - z_2, \quad x_2 = z_1 + z_2, \quad x_3 = z_3, \quad \ldots, \quad x_n = z_n.$$

Заметим, во-первых, что определитель матрицы этого преобразования равен двум, а во-вторых, что

$$2a_{12}x_1x_2 = 2a_{12}z_1^2 - 2a_{12}z_2^2.$$

.

Следовательно, в квадратичной форме появились слагаемые, содержащие квадраты переменных, поэтому, повторяя рассуждения предыдущего случая при помощи невырожденной замены переменных приведем квадратичную форму к виду

$$F = \alpha y_1^2 + G(y_2, \ldots, y_n).$$

Таким образом, выполняя одно или два последовательных невырожденных преобразования переменных, квадратичную форму можно привести к виду

$$F = \alpha y_1^2 + G(y_2, \ldots, y_n).$$

.

Аналогичными преобразованиями переменных в квадратичной форме

$$G(y_2, \ldots, y_n)$$

выделим квадрат переменной y_2. Продолжая преобразования, в конце концов приведем квадратичную форму к сумме квадратов.

.

<u>Пример.</u> Приведем к каноническому виду квадратичную форму

$$F(x_1, x_2, x_3) = 2x_1x_2 - 6x_2x_3 + 2x_3x_1.$$

.

Поскольку в форме

$$F(x_1, x_2, x_3) = 2x_1x_2 - 6x_2x_3 + 2x_3x_1$$

отсутствуют квадраты переменных, выполним сначала преобразование переменных

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3.$$

Получим

$$F = 2y_1^2 - 4y_1y_3 - 2y_2^2 - 8y_2y_3.$$

.

Положим в

$$F = 2y_1^2 - 4y_1y_3 - 2y_2^2 - 8y_2y_3$$

теперь

$$z_1 = y_1 - y_3, \quad z_2 = y_2, \quad z_3 = y_3.$$

Тогда

$$F = 2z_1^2 - 2z_2^2 - 8z_2z_3 - 2z_3^2 = 2z_1^2 - 2(z_2^2 + 4z_2z_3) - 2z_3^2.$$

.

Из

$$F = 2z_1^2 - 2(z_2^2 + 4z_2z_3) - 2z_3^2.$$

после замены переменных

$$t_1 = z_1, \quad t_2 = z_2 + 2z_3, \quad t_3 = z_3$$

получаем

$$F = 2t_1^2 - 2t_2^2 + 6t_3^2,$$

т. е. в переменных t_1, t_2, t_3 квадратичная форма принимает канонический вид.

Очевидно, каждое из выполненных нами преобразований переменных имеет невырожденную матрицу. Результирующее преобразование переменных, как нетрудно проверить, имеет вид

$$t_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - x_3, \quad t_2 = -\frac{1}{2}x_1 + \frac{1}{2}x_2 + 2x_3, \quad t_3 = x_3,$$

откуда

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

§2. ЗАКОН ИНЕРЦИИ КВАДРАТИЧНЫХ ФОРМ

Среди коэффициентов b_{ii} канонического вида квадратичной формы могу быть положительные, отрицательные, а также нулевые числа. Нумеруя соответствующим образом переменные, запишем

$$(Ax, x) = (By, y) = \sum_{i=1}^{n_+} b_{ii} y_i^2 + \sum_{i=n_+ +1}^{n_-} b_{ii} y_i^2.$$

Считаем при этом, что числа b_{ii} положительны при $i = 1, 2, \ldots, n_+$ и отрицательны при $i = n_+ + 1, \ldots, n_-$.

Поскольку, как мы уже убедились, приведение квадратичной формы к каноническому виду может быть выполнено различными способами, то естественно поставить вопрос, зависят ли числа

$$n_+, \quad n_-$$

от способа приведения квадратичной формы к каноническому виду.

При исследовании этого вопроса будут полезны следующие определения.

Симметричные матрицы A и B называют ***конгруэнтными***, если существует невырожденная матрица C такая, что

$$B = C^T AC.$$

С каждой симметричной матрицей A свяжем три целых числа:

$$n_0(A)$$

количество нулевых характеристических чисел матрицы A,

$$n_+(A)$$

количество положительных характеристических чисел,

$$n_-(A)$$

количество отрицательных характеристических чисел (характеристические числа подсчитываются с учетом их кратности).

Тройка чисел

$$n_0(A), \quad n_+(A), \quad n_-(A)$$

называется *инерцией матрицы* A или *инерцией*, соответствующей ей *квадратичной формы*.

Теорема. Для того, чтобы матрицы A и B были конгруэнтными необходимо и достаточно, чтобы их инерции совпадали.

.

Доказательство. ***Достаточность.*** Как было показано выше, для всякой симметричной матрицы A можно указать ортогональную матрицу Q такую, что

$$(Ax, x) = (Q^T AQy, y) = \sum_{i=1}^{n} \lambda_i y_i^2,$$

где λ_1, λ_2, $\ldots$, λ_n — собственные числа матрицы A.

.

Заметим, что

$$\sum_{i=1}^{n} \lambda_i y_i^2 = \sum_{i=1}^{n} \operatorname{sgn}(\lambda_i)|\lambda_i| y_i^2 = \sum_{i=1}^{n} \operatorname{sgn}(\lambda_i)(\sqrt{|\lambda_i|} y_i)^2 = \sum_{i=1}^{n_+(A)} t_i^2 - \sum_{i=1}^{n_-(A)} t_i^2.$$

Эти преобразования можно трактовать как невырожденную замену переменных:

$$t_i = \sqrt{|\lambda_i|} y_i, \quad \textbf{если} \quad \lambda_i \neq 0$$

и

$$t_i = y_i, \quad \textbf{если} \quad \lambda_i = 0.$$

Таким образом, установлено, что всякая симметричная матрица A конгруэнтна диагональной матрице, у которой на диагонали $n_+(A)$ единиц, $n_-(A)$ минус единиц, остальные элементы главной диагонали — нули.

Если симметричная матрица B имеет инерцию, равную инерции матрицы A, то она конгруэнтна точно такой же диагональной матрице. Отношение конгруэнтности, как нетрудно убедиться, транзитивно, следовательно, матрицы A и B конгруэнтны.

Необходимость. Заметим, прежде всего, что у когруэнтных матриц ранги, очевидно, совпадают. Кроме того, для любой симметричной матрицы A справедливо равенство

$$\operatorname{rank}(A) = n_+(A) + n_-(A).$$

Действительно, всякая симметричная матрица A подобна диагональной матрице, у которой по диагонали расположены все собственные числа матрицы A.

.

Из этих рассуждений вытекает, что если матрицы A и B конгруэнты, то

$$n_+(A) + n_-(A) = n_+(B) + n_-(B).$$

.

Таким образом, для завершения доказательства теоремы достаточно установить, что если матрицы A, B конгруэнтны, то

$$n_+(A) = n_+(B).$$

.

Пусть

$$0 < \lambda_{n-n_+ +1} \leqslant \ \ldots \ \leqslant \lambda_n$$

есть положительные собственные числа матрицы A,

$$e^{n-n_+ +1}, \quad \ldots, \quad e^n$$

есть соответствующие им ортонормированные собственные векторы матрицы A.

.

По предположению теоремы

$$B = C^T A C,$$

где C — невырожденная матрица, или

$$A = D^T B D,$$

где

$$D = C^{-1}.$$

Поскольку матрица D невырождена, векторы

$$De^{n-n_++1}, \quad \dots, \quad De^n$$

линейно независимы и подпространство

$$S_{n_+},$$

натянутое на эти векторы, имеет размерность n_+.

.

Пусть $x \in S_{n_+}$. **Тогда**

$$x = \alpha_{n-n_++1} De^{n-n_++1} + \cdots + \alpha_n De^n = Dy,$$

где

$$y = \alpha_{n-n_++1} e^{n-n_++1} + \cdots + \alpha_n e^n,$$

следовательно,

$$(Bx, x) = (D^T BDy, y) = (Ay, y) \geqslant \lambda_{n-n_++1}(y, y).$$

.

Заметим теперь, что

$$(y, y) = (Cx, Cx) = (C^T Cx, x).$$

Матрица C невырождена, поэтому матрица $C^T C$ положительно определена, следовательно,

$$(y, y) = (C^T Cx, x) \geqslant \lambda_1(C^T C)(x, x),$$

причем

$$\lambda_1(C^T C) > 0.$$

.

Из

$$(Bx, x) \geqslant \lambda_{n-n_{+}+1}(y, y),$$

$$(y, y) \geqslant \lambda_1(C^T C)(x, x),$$

вытекает, что

$$(Bx, x) \geqslant \lambda_{n-n_{+}+1}\lambda_1(C^T C)(x, x)$$

и

$$\min_{x \in S_{n_{+}},\ x \neq 0} \frac{(Bx, x)}{(x, x)} \geqslant \lambda_{n-n_{+}+1}\lambda_1(C^T C) > 0,$$

Итак,

$$\min_{x\in S_{n_+},\ x\neq 0} \frac{(Bx,x)}{(x,x)} > 0,$$

но

$$\lambda_{n-n_+ +1}(B) = \max_{S_{n_+}} \min_{x\in S_{n_+},\ x\neq 0} \frac{(Bx,x)}{(x,x)}.$$

поэтому

$$\lambda_{n-n_+ +1}(B) > 0.$$

.

Неравенство

$$\lambda_{n-n_{+}+1}(B) > 0$$

означает, что у матрицы B не меньше чем n_{+} положительных характеристических чисел, иначе говоря,

$$n_{+}(B) \geqslant n_{+}(A).$$

Итак,

$$n_+(B) \geqslant n_+(A).$$

В выполненных рассуждениях матрицы A и B можно поменять местами:

$$n_+(B) \leqslant n_+(A).$$

Таким образом, приходим к заключению, что

$$n_+(A) = n_+(B). \quad \square$$

.

<u>Следствие (закон инерции квадратичных форм)</u>. Количества положительных и отрицательных слагаемых в

$$(Ax, x) = (By, y) = \sum_{i=1}^{n_+} b_{ii} y_i^2 + \sum_{i=n_++1}^{n_-} b_{ii} y_i^2$$

не зависят от способа приведения невырожденным линейным преобразованием переменных квадратичной формы к каноническому виду

<u>Доказательство</u>. Коэффициенты b_{ii} в

$$(Ax, x) = (By, y) = \sum_{i=1}^{n_+} b_{ii} y_i^2 + \sum_{i=n_+ +1}^{n_-} b_{ii} y_i^2$$

это характеристические числа диагональной матрицы

$$B = Q^T A Q,$$

конгруэнтной матрице A, поэтому количества положительных и отрицательных слагаемых определяются инерцией матрицы A и не зависят от способа приведения невырожденным линейным преобразованием переменных квадратичной формы к каноническому виду. $\square$

§3. ПОЛОЖИТЕЛЬНО ОПРЕДЕЛЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ

Квадратичная форма называется *положительно определенной*, если соответствующая ей матрица A положительно определена, т. е.

$$(Ax, x) > 0 \quad \text{для всех не равных нулю} \quad x \in \mathbb{R}^n.$$

Как известно, для того чтобы матрица A была положительно определена необходимо и достаточно, чтобы все ее собственные числа были положительны.

.

Отметим, что поскольку определитель матрицы равен произведению ее собственных чисел, то определитель положительно определенной матрицы положителен.

Полезный признак положительной определенности дает

Теорема (критерий Сильвестра). Для того, чтобы квадратичная форма была положительно определена, необходимо и достаточно, чтобы все главные миноры матрицы A были положительны.

.

<u>Доказательство</u>. ***Необходимость.*** Пусть

$$(Ax, x) > 0 \quad \text{для всех не равных нулю} \quad x \in \mathbb{R}^n.$$

Выберем здесь в качестве вектора x вектор вида

$$x = (x_1, \ldots, x_k, 0, \ldots, 0) = (y, 0, \ldots, 0),$$

где y можно считать произвольным вектором пространства $\mathbb{R}^k$ для

$$1 \leqslant k \leqslant n.$$

Тогда

$$(Ax, x) = (A_k y, y),$$

где A_k — матрица, соответствующая главному минору порядка k матрицы A.

.

Из условия

$$(Ax, x) > 0 \quad \text{для всех не равных нулю} \;\; x \in \mathbb{R}^n.$$

и

$$(Ax, x) = (A_k y, y),$$

очевидно, вытекает, что

$$(A_k y, y) > 0 \quad \text{для всех не равных нулю} \;\; y \in \mathbb{R}^k.$$

т. е. матрица A_k положительно определена, следовательно ее определитель (главный минор порядка k матрицы A) положителен.

Достаточность. Покажем, что если все главные миноры матрицы A положительны, то положительны все ее собственные числа. Тогда положительная определенность матрицы A будет установлена. На самом деле, мы докажем большее, мы покажем, что собственные числа всех главных миноров матрицы A положительны.

Для минора первого порядка, т. е. для a_{11}, это выполняется тривиальным образом. Предположим, что у матрицы A_k, соответствующей главному минору порядка k, все собственные числа

$$\lambda_1 \leqslant \ldots \leqslant \lambda_k$$

положительны и покажем, что тогда и у матрицы A_{k+1} все собственные числа

$$\widehat{\lambda}_1 \leqslant \ldots \leqslant \widehat{\lambda}_{k+1}$$

положительны.

.

Как мы знаем, выполнены неравенства

$$\widehat{\lambda}_1 \leqslant \lambda_1 \leqslant \widehat{\lambda}_2 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k \leqslant \widehat{\lambda}_{k+1}.$$

Т. к.

$$\lambda_1 > 0$$

отсюда вытекает, что

$$\widehat{\lambda}_2, \ldots, \widehat{\lambda}_{k+1} > 0.$$

Поскольку по условию

$$\det(A_{k+1}) > 0,$$

а

$$\det(A_{k+1}) = \widehat{\lambda}_1 \widehat{\lambda}_2 \cdots \widehat{\lambda}_{k+1}, \quad \textbf{то и} \quad \widehat{\lambda}_1 > 0. \ \square$$

§4. КВАДРАТИЧНАЯ ФУНКЦИЯ И ЕЕ ИНВАРИАНТЫ

Пусть фиксированы

$$A \in \mathbb{R}^{n\times n}, \quad a \in \mathbb{R}^n, \quad a_0 \in \mathbb{R}$$

Определенная на пространстве $\mathbb{R}^n$ вещественная функция вида

$$F(x) = (Ax, x) + 2(a, x) + a_0$$

называется ***квадратичной***. Не ограничивая общности можно считать, что

$$A^T = A.$$

Понятно, что теория квадратичных функций может строиться как некоторое обобщение теории квадратичных форм.

Свяжем с каждой квадратичной функцией

$$F(x) = (Ax, x) + 2(a, x) + a_0$$

симметричную матрицу

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}.$$

Здесь a трактуется как вектор столбец.

.

Выполним так называемое *аффинное преобразование* переменных, т. е. положим

$$x = x^0 + Ty,$$

где (фиксированые)

$$x^0 \in \mathbb{R}^n,$$

$$T \in \mathbb{R}^{n \times n}, \quad \det(T) \neq 0.$$

Иногда замену переменных

$$x = x^0 + Ty$$

удобнее записывать в виде

$$x = T(\widehat{x}^0 + y),$$

где

$$\widehat{x}^0 = T^{-1}x^0.$$

Покажем, что

$$F(x^0 + Ty) \equiv \widehat{F}(y) = (\widehat{A}y, y) + 2(\widehat{a}, y) + \widehat{a}_0,$$

где

$$\widehat{A} = T^T AT,$$

$$\widehat{a} = T^T a + \widehat{A}\widehat{x}^0, \quad \widehat{a}_0 = a_0 + 2(T^T a, \widehat{x}^0) + (\widehat{A}\widehat{x}^0, \widehat{x}^0).$$

Действительно,

$$F(x^0 + Ty) = F(T(\widehat{x}^0 + y)) =$$

$$= (AT(\widehat{x}^0 + y), T(\widehat{x}^0 + y)) + 2(a, T(\widehat{x}^0 + y)) + a_0 =$$

$$= (AT\widehat{x}^0, T\widehat{x}^0) + \underline{(AT\widehat{x}^0, Ty)} + \underline{(ATy, T\widehat{x}^0)} + \underline{\underline{(ATy, Ty)}} +$$

$$+2(a, T\widehat{x}^0) + \underline{2(a, Ty)} + a_0 =$$

$$= \underline{\underline{(T^T ATy, y)}} +$$

$$+\underline{2(T^T a, y)} + \underline{2(T^T AT\widehat{x}^0, y)} +$$

$$+a_0 + 2(T^T a, \widehat{x}^0) + (T^T AT\widehat{x}^0, \widehat{x}^0) =$$

$$= \underline{\underline{(\widehat{A}y, y)}} + \underline{2(\widehat{a}, y)} + \widehat{a}_0 \equiv \widehat{F}(y),$$

где

$$\widehat{A} = T^T AT, \quad \widehat{a} = T^T a + \widehat{A}\widehat{x}^0, \quad \widehat{a}_0 = a_0 + 2(T^T a, \widehat{x}^0) + (\widehat{A}\widehat{x}^0, \widehat{x}^0).$$

Таким образом, любое аффинное преобразование переменных переводит квадратичную функцию в квадратичную:

$$F(x^0 + Ty) = \widehat{F}(y).$$

Введем в рассмотрение квадратные матрицы

$$Q = \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix}, \quad U = \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix}$$

порядка $n+1$. Здесь 0 — вектор столбец длины n, I — единичная матрица порядка n. Ясно, что

$$\det(Q) = \det(T),$$

$$\det(U) = 1,$$

т. е. матрицы Q, U невырождены.

Покажем, что

$$\widehat{B} \equiv \begin{pmatrix} \widehat{A} & \widehat{a} \\ \widehat{a}^T & \widehat{a}_0 \end{pmatrix} = (QU)^T B (QU).$$

.

Действительно, с одной стороны,

$$\begin{pmatrix} \widehat{A} & \widehat{a} \\ \widehat{a}^T & \widehat{a}_0 \end{pmatrix} = \begin{pmatrix} T^T A T & T^T a + T^T A T \widehat{x}^0 \\ (T^T a + T^T A T \widehat{x}^0)^T & a_0 + 2(T^T a, \widehat{x}^0) + (T^T A T \widehat{x}^0, \widehat{x}^0) \end{pmatrix} =$$

$$= \begin{pmatrix} T^T A T & T^T a + T^T A T \widehat{x}^0 \\ a^T T + (\widehat{x}^0)^T T^T A T & a_0 + 2(T^T a, \widehat{x}^0) + (T^T A T \widehat{x}^0, \widehat{x}^0) \end{pmatrix}.$$

.

С другой стороны,

$$(QU)^T B(QU) = U^T Q^T BQU,$$

$$Q^T BQ = \begin{pmatrix} T^T & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} T^T A & T^T a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} T^T AT & T^T a \\ a^T T & a_0 \end{pmatrix}.$$

.

Далее,

$$U^T \underline{Q^T BQ}U = \begin{pmatrix} I & 0 \\ (\widehat{x}^0)^T & 1 \end{pmatrix} \begin{pmatrix} T^T AT & T^T a \\ a^T T & a_0 \end{pmatrix} \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} T^T AT & T^T a \\ (\widehat{x}^0)^T T^T AT + a^T T & (\widehat{x}^0)^T T^T a + a_0 \end{pmatrix} \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} T^T AT & T^T AT\widehat{x}^0 + T^T a \\ (\widehat{x}^0)^T T^T AT + a^T T & (\widehat{x}^0)^T T^T AT\widehat{x}^0 + a^T T\widehat{x}^0 + (\widehat{x}^0)^T T^T a + a_0 \end{pmatrix}.$$

Итак,

$$\widehat{B} =$$

$$= \begin{pmatrix} T^T A T & T^T a + T^T A T \widehat{x}^0 \\ a^T T + (\widehat{x}^0)^T T^T A T & a_0 + \underline{2(T^T a, \widehat{x}^0)} + \underline{\underline{(T^T A T \widehat{x}^0, \widehat{x}^0)}} \end{pmatrix},$$

$$(QU)^T B (QU) =$$

$$= \begin{pmatrix} T^T A T & T^T a + T^T A T \widehat{x}^0 \\ a^T T + (\widehat{x}^0)^T T^T A T & a_0 + \underline{(\widehat{x}^0)^T T^T a + a^T T \widehat{x}^0} + \underline{\underline{(\widehat{x}^0)^T T^T A T \widehat{x}^0}} \end{pmatrix}.$$

Отсюда в силу $(x, y) = (y, x) = x^T y = y^T x$ **окончательно имеем**

$$\widehat{B} = (QU)^T B (QU).$$

Из соотношений

$$\widehat{A} = T^T A T,$$

$$\widehat{B} = (QU)^T B (QU).$$

вытекает, что матрицы A и $\widehat{A}$, B и $\widehat{B}$, соответственно, конгруэнтны, поэтому их инерции совпадают. Можно сказать, таким образом, что инерции матриц A, B являются ***аффинными инвариантами*** квадратичной функции.

.

Будем считать теперь, что матрица T ортогональна, т. е.

$$T^T T = I.$$

Тогда матрица Q, очевидно, также ортогональна:

$$Q^T Q = \begin{pmatrix} T^T & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} T^T T & 0 \\ 0^T & 1 \end{pmatrix} = I.$$

.

Из

$$\widehat{A} = T^T A T,$$

и

$$T^T = T^{-1}$$

вытекает, что матрицы A и $\widehat{A}$ подобны:

$$\widehat{A} = T^{-1} A T.$$

Следовательно, их собственные числа совпадают.

.

Из

$$\widehat{B} = U^T Q^T B Q U,$$

$$\det(T^T) = \det(T)^{-1},$$

$$\det(Q) = \det(T), \quad \det(Q^T) = \det(T^T)$$

$$\det(U) = 1, \quad \det(U^T) = 1$$

вытекает, что

$$\det(\widehat{B}) = \det(B).$$

Таким образом, собственные числа матрицы A, инерция, а следовательно, и ранг матрицы B, а также определитель матрицы B могут быть названы *ортогональными инвариантами* квадратичной функции

$$F(x) = (Ax, x) + 2(a, x) + a_0.$$

Они не меняются при любом преобразовании переменных

$$x = x^0 + Ty$$

с ортогональной матрицей T.

§5. ПРИВЕДЕННАЯ ФОРМА КВАДРАТИЧНОЙ ФУНКЦИИ

Покажем, что, выбирая в преобразовании переменных

$$x = x^0 + Ty$$

соответствующим образом ортогональную матрицу T и вектор x^0, любую квадратичную функцию

$$F(x) = (Ax, x) + 2(a, x) + a_0$$

можно преобразовать к простейшему ***приведенному виду***.

Матрица A симметрична, поэтому существует ортонормированный базис

$$e^1,\ e^2,\ \ldots,\ e^n$$

пространства $\mathbb{R}^n$, составленный из собственных векторов матрицы A. Обозначим через

$$\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$$

соответствующие им собственные числа матрицы A.

Будем считать что первые r собственных чисел матрицы A отличны от нуля,

$$\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_r \neq 0,$$

остальные — нули:

$$\lambda_{r+1} = \cdots = \lambda_n = 0.$$

Обозначим через T ортогональную матрицу, столбцы которой образованы векторами

$$e^1,\ e^2,\ \ldots,\ e^n.$$

.

Тогда, как мы знаем,

$$T^T A T = \Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Отметим, что последние $n - r$ столбцов матрицы T принадлежат ядру матрицы A:

$$Ae^k = \lambda_k e^k = 0e^k = 0, \quad k = r+1, \ \ldots, \ n.$$

.

Выполним замену переменных в функции

$$F(x) = (Ax, x) + 2(a, x) + a_0,$$

полагая

$$x = Tu.$$

Тогда

$$F(Tu) = (T^T ATu, u) + 2(\widehat{a}, u) + a_0 =$$

$$= (\Lambda u, u) + 2(\widehat{a}, u) + a_0 =$$

$$= \lambda_1 u_1^2 + \lambda_2 u_2^2 + \cdots + \lambda_r u_r^2 +$$

$$+2(\widehat{a}_1 u_1 + \widehat{a}_2 u_2 + \cdots + \widehat{a}_r u_r) +$$

$$+2(\widehat{a}_{r+1} u_{r+1} + \widehat{a}_{r+2} u_{r+2} + \cdots + \widehat{a}_n u_n) + a_0.$$

Заметим, что

$$\lambda_k u_k^2 + 2\widehat{a}_k u_k = \lambda_k \left(u_k + \frac{\widehat{a}_k}{\lambda_k} \right)^2 - \frac{\widehat{a}_k^2}{\lambda_k}, \quad k = 1,\ 2,\ \ldots,\ r.$$

Действительно,

$$\lambda_k \left(u_k + \frac{\widehat{a}_k}{\lambda_k} \right)^2 - \frac{\widehat{a}_k^2}{\lambda_k} = \lambda_k \left(u_k^2 + 2u_k \frac{\widehat{a}_k}{\lambda_k} + \frac{\widehat{a}_k^2}{\lambda_k^2} \right) - \frac{\widehat{a}_k^2}{\lambda_k} =$$

$$= \lambda_k u_k^2 + 2u_k \widehat{a}_k + \frac{\widehat{a}_k^2}{\lambda_k} - \frac{\widehat{a}_k^2}{\lambda_k} = \lambda_k u_k^2 + 2\widehat{a}_k u_k.$$

.

Подставим в

$$F(Tu) = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \cdots + \lambda_r u_r^2 + 2(\widehat{a}_1 u_1 + \widehat{a}_2 u_2 + \cdots + \widehat{a}_r u_r) +$$

$$+2(\widehat{a}_{r+1} u_{r+1} + \widehat{a}_{r+2} u_{r+2} + \cdots + \widehat{a}_n u_n) + a_0$$

выражения

$$\lambda_k u_k^2 + 2\widehat{a}_k u_k = \lambda_k \left(u_k + \frac{\widehat{a}_k}{\lambda_k} \right)^2 - \frac{\widehat{a}_k^2}{\lambda_k}, \quad k = 1,\ 2,\ \ldots,\ r.$$

Получим

$$F(Tu) = \lambda_1 \left(u_1 + \frac{\widehat{a}_1}{\lambda_1} \right)^2 + \lambda_2 \left(u_2 + \frac{\widehat{a}_2}{\lambda_2} \right)^2 + \cdots + \lambda_r \left(u_r + \frac{\widehat{a}_r}{\lambda_r} \right)^2 - \sum_{k=1}^{r} \frac{\widehat{a}_k^2}{\lambda_k} +$$

$$+2(\widehat{a}_{r+1} u_{r+1} + \widehat{a}_{r+2} u_{r+2} + \cdots + \widehat{a}_n u_n) + a_0.$$

Выражение

$$F(Tu) = \lambda_1 \left(u_1 + \frac{\widehat{a}_1}{\lambda_1}\right)^2 + \lambda_2 \left(u_2 + \frac{\widehat{a}_2}{\lambda_2}\right)^2 + \cdots + \lambda_r \left(u_r + \frac{\widehat{a}_r}{\lambda_r}\right)^2 +$$

$$+2(\widehat{a}_{r+1}u_{r+1} + \widehat{a}_{r+2}u_{r+2} + \cdots + \widehat{a}_n u_n) + a_0 - \sum_{k=1}^{r} \frac{\widehat{a}_k^2}{\lambda_k}$$

запишем в виде

$$F(Tu) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2(b, \widetilde{y}) + \widehat{a}_0,$$

где

$$y_k = u_k + \frac{\widehat{a}_k}{\lambda_k}, \quad k = 1,\ 2,\ \ldots,\ r,$$

$$b = (\widehat{a}_{r+1}, \widehat{a}_{r+2}, \ldots, \widehat{a}_n), \quad \widetilde{y} = (u_{r+1}, u_{r+2}, \ldots, u_n) \in \mathbb{R}^{n-r},$$

$$\widehat{a}_0 = a_0 - \sum_{k=1}^{r} \frac{\widehat{a}_k^2}{\lambda_k}.$$

Далее будем различать два случая. Предположим сначала, что

$$b = 0.$$

Тогда

$$F(Tu) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2(b, \widetilde{y}) + \widehat{a}_0 =$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0.$$

.

Пусть

$$\widehat{x}_k^0 = -\frac{\widehat{a}_k}{\lambda_k}, \quad k = 1,\ 2,\ \ldots,\ r,$$

$$\widehat{x}_k^0 = 0, \quad k = r+1,\ r+2,\ \ldots,\ n.$$

Тогда из

$$u_k = y_k - \frac{\widehat{a}_k}{\lambda_k}, \quad k = 1,\ 2,\ \ldots,\ r,$$

получаем

$$u = y + \widehat{x}^0,$$

$$Tu = Ty + T\widehat{x}^0.$$

.

Вследствие

$$Tu = T\widehat{x}^0 + Ty$$

равенство

$$F(Tu) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0$$

принимает вид

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0,$$

где

$$x^0 = T\widehat{x}^0.$$

.

Пусть теперь

$$0 \neq b \in \mathbb{R}^{n-r}.$$

Построим симметричную ортогональную матрицу отражения

$$R \in \mathbb{R}^{(n-r)\cdot(n-r)}$$

такую, что

$$Rb = |b|(1, 0, \ \ldots, 0).$$

Выполним в

$$F(Tu) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2(b, \widetilde{y}) + \widehat{a}_0$$

замену переменных $y = \widetilde{R}v$, **где**

$$\widetilde{R} = \begin{pmatrix} I_r & 0 \\ 0^T & R \end{pmatrix},$$

I_r — **единичная матрица порядка** r,

$$Rb = |b|(1, 0, \ \ldots, 0).$$

Тогда

$$F(Tu) = \lambda_1 v_1^2 + \lambda_2 v_2^2 + \cdots + \lambda_r v_r^2 + 2(b, Rv) + \widehat{a}_0 =$$

$$= \lambda_1 v_1^2 + \lambda_2 v_2^2 + \cdots + \lambda_r v_r^2 + 2b_{r+1}v_{r+1} + \widehat{a}_0,$$

где $b_{r+1} = |b|$.

Заметим, что

$$2b_{r+1}v_{r+1}+\widehat{a}_0=2b_{r+1}\left(v_{r+1}+\frac{\widehat{a}_0}{2b_{r+1}}\right).$$

Поэтому, полагая

$$w=v+x^1,$$

где

$$x^1_{r+1}=\frac{\widehat{a}_0}{2b_{r+1}},$$

$$x^1_i=0,\quad i=1,\ 2,\ \ldots,\ n,\quad i\neq r+1,$$

из

$$F(Tu)=\lambda_1v_1^2+\lambda_2v_2^2+\cdots+\lambda_rv_r^2+2b_{r+1}v_{r+1}+\widehat{a}_0$$

получим

$$F(Tu)=\lambda_1w_1^2+\lambda_2w_2^2+\cdots+\lambda_rw_r^2+2b_{r+1}w_{r+1}.$$

Учитывая

$$y = \widetilde{R}v, \quad T\widehat{x}^0 = x^0, \quad v = w - x^1, \quad Tu = Ty + T\widehat{x}^0,$$

получаем

$$Tu = Ty + T\widehat{x}^0 = T\widetilde{R}v + x^0 = T\widetilde{R}(w - x^1) + x^0 =$$

$$= x^0 - T\widetilde{R}x^1 + T\widetilde{R}w = \widetilde{x}^0 + \widetilde{T}w,$$

где

$$\widetilde{T} = T\widetilde{R}, \quad \widetilde{x}^0 = x^0 - \widetilde{T}x^1.$$

.

Итак,

$$F(Tu) = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_r w_r^2 + 2b_{r+1} w_{r+1},$$

$$Tu = \widetilde{x}^0 + \widetilde{T} w,$$

следовательно,

$$F(\widetilde{x}^0 + \widetilde{T} w) = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_r w_r^2 + 2b_{r+1} w_{r+1}.$$

.

Матрица $\widetilde{T}$ ортогональна, поскольку является произведением ортогональных матриц:

$$\widetilde{T} = T \begin{pmatrix} I_r & 0 \\ 0^T & R \end{pmatrix}.$$

.

Формула

$$\widetilde{T} = T \begin{pmatrix} I_r & 0 \\ 0^T & R \end{pmatrix}$$

показывает, что первые r столбцов матрицы $\widetilde{T}$ совпадают с соответствующими столбцами матрицы T.

Кроме того, формула

$$\widetilde{T} = T \begin{pmatrix} I_r & 0 \\ 0^T & R \end{pmatrix}$$

показывает, что последние $n-r$ столбцов матрицы $\widetilde{T}$ являются линейными комбинациями последних $n-r$ столбцов матрицы матрицы T и потому принадлежат ядру матрицы A.

Таким образом, доказано, что для любой квадратичной функции

$$F(x) = (Ax, x) + 2(a, x) + a_0,$$

найдутся матрица T, столбцы которой есть векторы ортонормированного базиса пространства $\mathbb{R}^n$, образованного собственными векторами матрицы A, и вектор x^0 такие, что либо

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0,$$

либо

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2b_{r+1} y_{r+1}.$$

Здесь λ_1, λ_2, $\ldots$, λ_r — все ненулевые собственные числа матрицы A,

$$b_{r+1} > 0.$$

Эти представления называются ***приведенными формами*** квадратичной функции.

Квадратичной функции

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0$$

соответствует матрица

$$\widehat{B} = \begin{pmatrix} \Lambda & 0 \\ 0^T & \widehat{a}_0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \lambda_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & \widehat{a}_0 \end{pmatrix},$$

$$\mathrm{rank}(\widehat{B}) = r, \quad \textbf{если} \quad \widehat{a}_0 = 0,$$

$$\mathrm{rank}(\widehat{B}) = r + 1, \quad \textbf{если} \quad \widehat{a}_0 \neq 0.$$

Квадратичной функции

$$F(x^0+Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2b_{r+1}y_{r+1}$$

соответствует матрица

$$\widehat{B} = \begin{pmatrix} \Lambda & b \\ b^T & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \lambda_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & b_{r+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_{r+1} & \dots & 0 \end{pmatrix}.$$

Ясно, что

$$\operatorname{rank}(\widehat{B}) = r + 2.$$

Собственные числа матрицы A и ранг матрицы

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}$$

инвариантны по отношению к замене переменных $x = x^0 + Ty$ с любой ортогональной матрицей T и любым вектором x^0. Поэтому любой квадратичной функции

$$F(x) = (Ax, x) + 2(a, x) + a_0$$

однозначно соответствует либо приведенная форма вида

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + \widehat{a}_0,$$

либо приведенная форма вида

$$F(x^0 + Ty) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + 2b_{r+1} y_{r+1}.$$

Покажем, что коэффициенты приведенной формы квадратичной функции F однозначно определяются по элементам матрицы B. Они не зависят от выбора вектора x^0 и ортогональной матрицы T в преобразовании переменных, дающем приведенную форму квадратичной функции.

.

Нам потребуются в дальнейшем некоторые вспомогательные результаты.

<u>Лемма.</u> **Пусть**

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}, \quad A = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn}) \in \mathbb{R}^{n\cdot n},$$

$$a_{11},\ a_{22},\ \ldots,\ a_{rr} \neq 0, \quad a_{r+1 r+1} = \cdots = a_{nn} = 0, \quad r \leqslant n-1,$$

$$a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n.$$

Тогда

$$\mathcal{I}_{r+2}(B) = -a_{11}a_{22}\cdots a_{rr}(a_{r+1}^2 + \cdots + a_n^2).$$

Здесь

$$\mathcal{I}_{r+2}(B) = \sum_{1\leqslant i_1<i_2<\cdots<i_{r+2}\leqslant n+1} \begin{vmatrix} b_{i_1,i_1} & b_{i_1,i_2} & \ldots & b_{i_1,i_k} \\ b_{i_2,i_1} & b_{i_2,i_2} & \ldots & b_{i_2,i_k} \\ \ldots & \ldots & \ldots & \ldots \\ b_{i_k,i_1} & b_{i_k,i_2} & \ldots & b_{i_k,i_k} \end{vmatrix}.$$

<u>Доказательство.</u> Легко видеть, что среди диагональных миноров порядка $r+2$ матрицы

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}, \quad a_{11},\, a_{22},\, \ldots,\, a_{rr} \neq 0,$$

лишь миноры вида

$$\Delta_{r,m} = \begin{vmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{vmatrix}, \quad D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}),$$

$$D_{12} = \begin{pmatrix} 0 & a_1 \\ \vdots & \vdots \\ 0 & a_r \end{pmatrix}, \quad D_{22} = \begin{pmatrix} 0 & a_m \\ a_m & a_0 \end{pmatrix}, \quad m = r+1,\ r+2,\ \ldots,\ n,$$

отличны от нуля. Все остальные диагональные миноры порядка $r+2$ содержат хотя бы одну нулевую строку (и столбец).

.

Таким образом,

$$\mathcal{I}_{r+2}(B) = \sum_{m=r+1}^{n} \Delta_{r,m}.$$

.

Мы знаем, что

$$\Delta_{r,m} = \begin{vmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{vmatrix} = |D_{11}| \left| D_{22} - D_{12}^T D_{11}^{-1} D_{12} \right| .$$

.

Имеем

$$D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}) \quad \Longrightarrow \quad |D_{11}| = a_{11}a_{22}\cdots a_{rr}.$$

.

Далее,

$$D_{11}^{-1} = \mathrm{diag}(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{rr}^{-1}), \qquad D_{12} = \begin{pmatrix} 0 & a_1 \\ \vdots & \vdots \\ 0 & a_r \end{pmatrix},$$

следовательно,

$$D_{11}^{-1} D_{12} = \begin{pmatrix} 0 & d_1 \\ \vdots & \vdots \\ 0 & d_r \end{pmatrix}.$$

.

Теперь имеем

$$D_{12}^T = \begin{pmatrix} 0 & \dots & 0 \\ a_1 & \dots & a_r \end{pmatrix}, \quad D_{11}^{-1} D_{12} = \begin{pmatrix} 0 & d_1 \\ \vdots & \vdots \\ 0 & d_r \end{pmatrix},$$

значит,

$$D_{12}^T D_{11}^{-1} D_{12} = \begin{pmatrix} 0 & 0 \\ 0 & d_{22} \end{pmatrix}.$$

.

Из

$$D_{22} = \begin{pmatrix} 0 & a_m \\ a_m & a_0 \end{pmatrix}, \quad D_{12}^T D_{11}^{-1} D_{12} = \begin{pmatrix} 0 & 0 \\ 0 & d_{22} \end{pmatrix}$$

вытекает, что

$$\left| D_{22} - D_{12}^T D_{11}^{-1} D_{12} \right| = \begin{vmatrix} 0 & a_m \\ a_m & a_0 - d_{22} \end{vmatrix} = -a_m^2.$$

.

Итак,

$$\Delta_{r,m} = \begin{vmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{vmatrix} = |D_{11}| \left| D_{22} - D_{12}^T D_{11}^{-1} D_{12} \right|,$$

$$|D_{11}| = \underline{a_{11} a_{22} \cdots a_{rr}},$$

$$\left| D_{22} - D_{12}^T D_{11}^{-1} D_{12} \right| = -a_m^2.$$

Следовательно,

$$\Delta_{r,m} = -\underline{a_{11} a_{22} \cdots a_{rr}} a_m^2$$

.

Суммируя теперь все миноры вида

$$\Delta_{r,m} = -\underline{a_{11}a_{22}\cdots a_{rr}}a_m^2, \quad m = r+1,\, r+2,\, \ldots,\, n,$$

приходим к

$$\mathcal{I}_{r+2}(B) = \sum_{m=r+1}^{n} \Delta_{r,m} = -\underline{a_{11}a_{22}\cdots a_{rr}}(a_{r+1}^2 + \cdots + a_n^2). \;\Box$$

.

<u>Лемма.</u> **Пусть выполнены условия предыдущей леммы,**

$$\operatorname{rank}(B) = r + 1.$$

Тогда

$$\mathcal{I}_{r+1}(B) = \begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix},$$

где

$$D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}),$$

а

$$d = (a_1, a_2, \ldots, a_r)$$

есть вектор столбец.

.

<u>Доказательство.</u> Вследствие условия

$$\operatorname{rank}(B) = r + 1$$

все миноры порядка $r + 2$ матрицы B равны нулю. Поэтому

$$\mathcal{I}_{r+2}(B) = 0,$$

откуда вследствие

$$\mathcal{I}_{r+2}(B) = -a_{11}a_{22}\cdots a_{rr}(a_{r+1}^2 + \cdots + a_n^2),$$

$$a_{11},\, a_{22},\, \cdots,\, a_{rr} \neq 0$$

вытекает, что

$$a_{r+1} = a_{r+2} = \cdots = a_n = 0.$$

.

Из условия

$$a_{r+1} = a_{r+2} = \cdots = a_n = 0$$

следует, что все диагональные миноры порядка $r+1$ **матрицы**

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}, \quad a_{11},\, a_{22},\, \ldots,\, a_{rr} \neq 0,$$

кроме минора вида

$$\mathcal{I}_{r+1}(B) = \begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix}, \quad D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}), \quad d = (a_1, a_2, \ldots, a_r),$$

содержат хотя бы одну нулевую строку. $\square$

<u>Лемма.</u> **Пусть выполнены условия первой леммы, кроме того,**

$$U = \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix}, \quad \widehat{x}^0 = T^{-1}x^0, \quad x = x^0 + Ty,$$

$$\widetilde{B} = U^T BU, \quad B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}, \quad a_{11},\, a_{22},\, \ldots,\, a_{rr} \neq 0.$$

Тогда

$$\mathcal{I}_{r+2}(\widetilde{B}) = \mathcal{I}_{r+2}(B).$$

Если при этом

$$\mathrm{rank}(B) = r + 1,$$

то

$$\mathcal{I}_{r+1}(\widetilde{B}) = \mathcal{I}_{r+1}(B).$$

.

<u>Доказательство.</u> Заметим, что

$$BU = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} A & \widetilde{a} \\ a^T & \widetilde{a}_0 \end{pmatrix},$$

где

$$\widetilde{a} = A\widehat{x}^0 + a,$$

подробнее,

$$\widetilde{a}_i = a_{ii}\widehat{x}_i^0 + a_i, \quad i = 1, 2, \ldots, r, \quad \widetilde{a}_i = a_i, \quad i = r+1, \ldots, n,$$

кроме того,

$$\widetilde{a}_0 = a^T\widehat{x}^0 + a_0 = \sum_{i=1}^{n} a_i\widehat{x}_i^0 + a_0.$$

.

Вследствие

$$\mathcal{I}_{r+2}(B) = -a_{11}a_{22}\cdots a_{rr}(a_{r+1}^2 + \cdots + a_n^2),$$

$$BU = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} A & \widetilde{a} \\ a^T & \widetilde{a}_0 \end{pmatrix},$$

$$\widetilde{a}_i = a_i, \quad i = r+1, \ldots, n,$$

справедливо равенство

$$\mathcal{I}_{r+2}(BU) = \mathcal{I}_{r+2}(B).$$

Точно так же проверяется, что умножение матрицы вида

$$BU = \begin{pmatrix} A & \widetilde{a} \\ a^T & \widetilde{a}_0 \end{pmatrix}$$

слева на матрицу

$$U^T = \begin{pmatrix} I & 0 \\ (\widehat{x}^0)^T & 1 \end{pmatrix}$$

не меняет инварианта

$$\mathcal{I}_{r+2}(BU) = \mathcal{I}_{r+2}(B),$$

т. е.

$$\mathcal{I}_{r+2}(U^T BU) = \mathcal{I}_{r+2}(BU).$$

Отсюда, очевидно, вытекает равенство

$$\mathcal{I}_{r+2}(U^T BU) = \mathcal{I}_{r+2}(B).$$

Предположим теперь, что выполнено условие

$$\operatorname{rank}(B) = r + 1.$$

Значит, все миноры порядка $r+2$ матрицы B равны нулю. Поэтому

$$\mathcal{I}_{r+2}(B) = 0,$$

откуда вследствие

$$\mathcal{I}_{r+2}(B) = -a_{11}a_{22}\cdots a_{rr}(a_{r+1}^2 + \cdots + a_n^2),$$

$$a_{11},\, a_{22},\, \cdots,\, a_{rr} \neq 0$$

вытекает, что

$$a_{r+1} = a_{r+2} = \cdots = a_n = 0.$$

.

Итак,

$$a_{r+1} = a_{r+2} = \cdots = a_n = 0.$$

Отсюда следует, что все диагональные миноры порядка $r+1$ матрицы

$$B = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix}, \quad a_{11},\, a_{22},\, \ldots,\, a_{rr} \neq 0,$$

кроме минора вида

$$\mathcal{I}_{r+1}(B) = \begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix}, \quad D_{11} = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{rr}), \quad d = (a_1, a_2, \ldots, a_r),$$

содержат хотя бы одну нулевую строку.

.

Покажем, что определитель

$$\mathcal{I}_{r+1}(B) = \begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix},$$

$$D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}), \quad d = (a_1, a_2, \ldots, a_r),$$

не меняется при переходе от матрицы B к матрице

$$\widetilde{B} = U^T B U.$$

Тем самым будет доказано требуемое равенство:

$$\mathcal{I}_{r+1}(\widetilde{B}) = \mathcal{I}_{r+1}(B).$$

Имеем

$$BU = \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} A & A\widehat{x}^0 + a \\ a^T & a^T\widehat{x}^0 + a_0 \end{pmatrix},$$

$$U^T BU = \begin{pmatrix} I & 0 \\ (\widehat{x}^0)^T & 1 \end{pmatrix} \begin{pmatrix} A & A\widehat{x}^0 + a \\ a^T & a^T\widehat{x}^0 + a_0 \end{pmatrix} =$$

$$= \begin{pmatrix} A & A\widehat{x}^0 + a \\ (\widehat{x}^0)^T A + a^T & (\widehat{x}^0)^T A\widehat{x}^0 + (\widehat{x}^0)^T a + a^T\widehat{x}^0 + a_0 \end{pmatrix}.$$

Итак,

$$\mathcal{I}_{r+1}(B) = \begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix} = |D_{11}|(a_0 - d^T D_{11}^{-1} d),$$

$$D_{11} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{rr}), \qquad d = (a_1, a_2, \ldots, a_r),$$

$$\widetilde{B} = U^T B U = \begin{pmatrix} A & A\widehat{x}^0 + a \\ (\widehat{x}^0)^T A + a^T & (\widehat{x}^0)^T A \widehat{x}^0 + (\widehat{x}^0)^T a + a^T \widehat{x}^0 + a_0 \end{pmatrix}.$$

Тогда

$$\mathcal{I}_{r+1}(\widetilde{B}) = \begin{vmatrix} D_{11} & D_{11} d^0 + d \\ (d^0)^T D_{11} + d^T & (d^0)^T D_{11} d^0 + (d^0)^T d + d^T d^0 + a_0 \end{vmatrix},$$

где

$$d^0 = (\widehat{x}_1^0, \widehat{x}_2^0, \ldots, \widehat{x}_r^0).$$

Вычислим минор

.

$$\mathcal{I}_{r+1}(\widetilde{B}) = \begin{vmatrix} D_{11} & D_{11}d^0 + d \\ (d^0)^T D_{11} + d^T & (d^0)^T D_{11} d^0 + (d^0)^T d + d^T d^0 + a_0 \end{vmatrix}$$

по формуле

$$\begin{vmatrix} D_{11} & d \\ d^T & a_0 \end{vmatrix} = |D_{11}|(a_0 - d^T D_{11}^{-1} d) :$$

$$\mathcal{I}_{r+1}(\widetilde{B}) = |D_{11}| \Big[(d^0)^T D_{11} d^0 + (d^0)^T d + d^T d^0 + a_0 -$$

$$- \left((d^0)^T D_{11} + d^T \right) D_{11}^{-1} \left(D_{11} d^0 + d \right) \Big] = |D_{11}| \cdot$$

$$\cdot \Big[\underline{(d^0)^T D_{11} d^0} + \underline{\underline{(d^0)^T d}} + \underline{\underline{\underline{d^T d^0}}} + a_0 - \underline{(d^0)^T D_{11} d^0} - \underline{\underline{(d^0)^T d}} - \underline{\underline{\underline{d^T d^0}}} - d^T D_{11}^{-1} d \Big] =$$

$$= |D_{11}|(a_0 - d^T D_{11}^{-1} d).$$

.

Итак,

$$\mathcal{I}_{r+1}(B) = |D_{11}|(a_0 - d^T D_{11}^{-1} d)$$

и

$$\mathcal{I}_{r+1}(\widetilde{B}) = |D_{11}|(a_0 - d^T D_{11}^{-1} d),$$

т. е.

$$\mathcal{I}_{r+1}(\widetilde{B}) = \mathcal{I}_{r+1}(B). \ \square$$

.

<u>Теорема.</u> Пусть ранг матрицы A квадратичной функции F равен r, ранг матрицы B не превосходит $r+1$. Пусть при помощи замены переменных $x = x^0 + Ty$ с ортогональной матрицей T квадратичная функция F приведена к виду

$$\widehat{F}(y) = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \cdots + \alpha_m y_m^2 + \widehat{a}_0.$$

Тогда:

1) $m = r$, $\alpha_i = \lambda_i$, где λ_i, $i = 1, 2, \ \ldots, r$, — все ненулевые собственные числа матрицы A;

2) справедливо равенство

$$\widehat{a}_0 = \frac{\mathcal{I}_{r+1}(B)}{\mathcal{I}_r(A)}.$$

.

<u>Доказательство.</u> В общем случае

$$F(x^0 + Ty) = \widehat{F}(y) = (\widehat{A}y, y) + 2(\widehat{a}, y) + \widehat{a}_0,$$

где

$$\widehat{A} = T^T AT.$$

В рассматриваемом в теореме случае

$$\widehat{F}(y) = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \cdots + \alpha_m y_m^2 + \widehat{a}_0,$$

т. е.

$$\widehat{a} = 0,$$

и

$$\widehat{A} = \mathrm{diag}(\alpha_1,\ \alpha_2,\ \ldots,\alpha_m).$$

Поскольку матрица T ортогональна,

$$T^T = T^{-1},$$

а

$$\widehat{A} = T^T AT,$$

то матрицы A и $\widehat{A}$ подобны:

$$\widehat{A} = T^{-1}AT.$$

Спектры подобных матриц совпадают, и из

$$\widehat{A} = \mathrm{diag}(\alpha_1,\, \alpha_2,\, \ldots, \alpha_m)$$

теперь вытекает справедливость утверждения 1):

$$m = r, \quad \alpha_i = \lambda_i(A), \quad i = 1,\, 2,\, \ldots,\, r.$$

Используем теперь тот факт, что, в рассматриваемом случае

$$\widehat{a} = 0.$$

Имеем

$$\widehat{B} = \begin{pmatrix} \widehat{A} & \widehat{a} \\ \widehat{a}^T & \widehat{a}_0 \end{pmatrix} = \begin{pmatrix} \widehat{A} & 0 \\ 0^T & \widehat{a}_0 \end{pmatrix}, \quad \widehat{A} = \mathrm{diag}(\lambda_1,\, \lambda_2,\, \ldots,\, \lambda_m).$$

Поэтому по второй лемме получаем, что

$$\mathcal{I}_{r+1}(\widehat{B}) = \lambda_1 \lambda_2 \cdots \lambda_r \widehat{a}_0.$$

Мы знаем, что $\mathcal{I}_r(A)$ — есть сумма всевозможных произведений r различных собственных чисел матрицы A. Она имеет ровно r ненулевых собственных чисел, следовательно,

$$\mathcal{I}_r(A) = \lambda_1 \lambda_2 \cdots \lambda_r.$$

.

Итак,

$$\mathcal{I}_{r+1}(\widehat{B}) = \lambda_1\lambda_2\cdots\lambda_r\widehat{a}_0,$$

и

$$\mathcal{I}_r(A) = \lambda_1\lambda_2\cdots\lambda_r.$$

Таким образом,

$$\widehat{a}_0 = \frac{\mathcal{I}_{r+1}(\widehat{B})}{\mathcal{I}_r(A)}.$$

Вспомним теперь, что

$$\widehat{B} = (QU)^T B(QU) = U^T (Q^T BQ)U,$$

где

$$Q = \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix}, \quad U = \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix}.$$

Кроме того, матрица Q ортогональна,

$$Q^T = Q^{-1},$$

и матрица $Q^T BQ$ подобна матрице B:

$$Q^T BQ = Q^{-1} BQ.$$

Поэтому все их инварианты совпадают:

$$\mathcal{I}_k(Q^T BQ) = \mathcal{I}_k(B), \quad k = 1,\, 2,\, \ldots,\, n+1.$$

.

Заметим теперь, что матрица

$$Q^T B Q$$

удовлетворяет условиям третьей леммы:

$$Q^T B Q = \begin{pmatrix} T^T & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} T^T A & T^T a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} T^T A T & T^T a \\ a^T T & a_0 \end{pmatrix} = \begin{pmatrix} \widehat{A} & T^T a \\ a^T T & a_0 \end{pmatrix},$$

где

$$\widehat{A} = \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\lambda_m).$$

.

Итак, имеем

$$\widehat{B} = U^T(Q^T BQ)U,$$

и по третьей лемме

$$\mathcal{I}_{r+1}(\widehat{B}) = \mathcal{I}_{r+1}(Q^T BQ),$$

но по доказанному только что

$$\mathcal{I}_{r+1}(Q^T BQ) = \mathcal{I}_{r+1}(B),$$

т. е.

$$\mathcal{I}_{r+1}(\widehat{B}) = \mathcal{I}_{r+1}(B),$$

и утверждение 2) теоремы также доказано:

$$\widehat{a}_0 = \frac{\mathcal{I}_{r+1}(\widehat{B})}{\mathcal{I}_r(A)} = \frac{\mathcal{I}_{r+1}(B)}{\mathcal{I}_r(A)}. \ \square$$

.

<u>Теорема.</u> Пусть ранг матрицы A квадратичной функции F равен r, $r \leqslant n-1$, ранг матрицы B равен $r+2$. Пусть при помощи замены переменных $x = x^0 + Ty$ с ортогональной матрицей T квадратичная функция F приведена к виду

$$\widehat{F}(y) = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \cdots + \alpha_m y_m^2 + 2b y_{m+1}.$$

Тогда:

1) $m = r$, $\alpha_i = \lambda_i$, где λ_i, $i = 1, 2, \ldots, r$, — все ненулевые собственные числа матрицы A;

2) выполнено равенство

$$b^2 = -\frac{\mathcal{I}_{r+2}(B)}{\mathcal{I}_r(A)}.$$

.

<u>Доказательство.</u> В общем случае

$$F(x^0 + Ty) = \widehat{F}(y) = (\widehat{A}y, y) + 2(\widehat{a}, y) + \widehat{a}_0,$$

где

$$\widehat{A} = T^T AT.$$

В рассматриваемом в теореме случае

$$\widehat{F}(y) = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \cdots + \alpha_m y_m^2 + 2b y_{m+1},$$

т. е.

$$\widehat{a}_{m+1} = b, \quad \widehat{a}_k = 0, \quad k \neq m+1, \quad \widehat{a}_0 = 0,$$

и

$$\widehat{A} = \mathrm{diag}(\alpha_1,\, \alpha_2,\, \ldots, \alpha_m).$$

Поскольку матрица T ортогональна,

$$T^T = T^{-1},$$

а

$$\widehat{A} = T^T A T,$$

то матрицы A и $\widehat{A}$ подобны:

$$\widehat{A} = T^{-1} A T.$$

Спектры подобных матриц совпадают, и из

$$\widehat{A} = \mathrm{diag}(\alpha_1,\, \alpha_2,\, \ldots, \alpha_m)$$

теперь вытекает справедливость утверждения 1):

$$m = r, \quad \alpha_i = \lambda_i(A), \quad i = 1,\, 2,\, \ldots,\, r.$$

.

Имеем

$$\widehat{B} = \begin{pmatrix} \widehat{A} & \widehat{a} \\ \widehat{a}^T & 0 \end{pmatrix},$$

$$\widehat{A} = \mathrm{diag}(\lambda_1,\ \lambda_2,\ \ldots,\lambda_r),$$

$$\widehat{a}_{r+1} = b, \quad \widehat{a}_k = 0, \quad k \neq r+1.$$

Поэтому по первой лемме

$$\mathcal{I}_{r+2}(\widehat{B}) = -\lambda_1\lambda_2\cdots\lambda_r b^2.$$

.

Мы установили ранее, что

$$\mathcal{I}_r(A) = \lambda_1 \lambda_2 \cdots \lambda_r.$$

.

Итак,

$$\mathcal{I}_{r+2}(\widehat{B}) = -\lambda_1\lambda_2\cdots\lambda_r b^2,$$

и

$$\mathcal{I}_r(A) = \lambda_1\lambda_2\cdots\lambda_r.$$

Таким образом,

$$b^2 = -\frac{\mathcal{I}_{r+2}(\widehat{B})}{\mathcal{I}_r(A)}.$$

Вспомним теперь, что

$$\widehat{B} = (QU)^T B(QU) = U^T (Q^T BQ)U,$$

где

$$Q = \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix}, \quad U = \begin{pmatrix} I & \widehat{x}^0 \\ 0^T & 1 \end{pmatrix}.$$

Кроме того, матрица Q ортогональна,

$$Q^T = Q^{-1},$$

и матрица $Q^T BQ$ подобна матрице B:

$$Q^T BQ = Q^{-1} BQ.$$

Поэтому все их инварианты совпадают:

$$\mathcal{I}_k(Q^T BQ) = \mathcal{I}_k(B), \quad k = 1,\, 2,\, \ldots,\, n+1.$$

.

Заметим теперь, что матрица

$$Q^T B Q$$

удовлетворяет условиям третьей леммы:

$$Q^T B Q = \begin{pmatrix} T^T & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} A & a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} T^T A & T^T a \\ a^T & a_0 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} T^T A T & T^T a \\ a^T T & a_0 \end{pmatrix} = \begin{pmatrix} \widehat{A} & T^T a \\ a^T T & a_0 \end{pmatrix},$$

где

$$\widehat{A} = \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\lambda_m).$$

.

Итак, имеем

$$\widehat{B} = U^T (Q^T B Q) U,$$

и по третьей лемме

$$\mathcal{I}_{r+2}(\widehat{B}) = \mathcal{I}_{r+2}(Q^T B Q),$$

но по доказанному только что

$$\mathcal{I}_{r+2}(Q^T B Q) = \mathcal{I}_{r+2}(B),$$

т. е.

$$\mathcal{I}_{r+2}(\widehat{B}) = \mathcal{I}_{r+2}(B),$$

и утверждение 2) теоремы также доказано:

$$b^2 = -\frac{\mathcal{I}_{r+2}(\widehat{B})}{\mathcal{I}_r(A)} = -\frac{\mathcal{I}_{r+2}(B)}{\mathcal{I}_r(A)}. \ \square$$