Лекция 11

Симметрические билинейные функционалы. — Квадратичные функционалы и квадратичные формы. — Теорема Лагранжа.

По аналогии с кососимметрическими функционалами *симметрические полилинейные функционалы* определяются как функционалы B из $\mathsf{T}_p(\mathscr{V})$ (или из $\mathsf{T}^p(\mathscr{V})$), значения которых не меняются при любой перестановке аргументов. Однако в общем случае теория таких функционалов оказывается очень сложной и о них до сих пор мало что известно. Исключением является лишь случай $p=2$, т. е. случай билинейных функционалов. Этими функционалами мы и займемся. Для определенности мы будем рассматривать функционалы от векторов (т. е. из $\mathsf{T}_2(\mathscr{V})$).

При этом мы будем предполагать, что $\operatorname{char} \mathsf{K} \neq 2$.

Согласно общим результатам лекции 5, при заданном базисе $e_1, \ldots, e_n$ пространства $\mathscr{V}$ каждый симметрический билинейный функционал B выражается через ковекторы $e^1, \ldots, e^n$ сопряженного базиса пространства $\mathscr{V}'$ посредством формулы

$$B = b_{ij} e^i \otimes e^j,$$

коэффициенты $b_{ij} = B(e_i, e_j)$ которой составляют матрицу

$$B = \begin{Vmatrix} b_{11} & \ldots & b_{1n} \\ \cdot & \cdot & \cdot \\ b_{n1} & \ldots & b_{nn} \end{Vmatrix}, \tag{1}$$

называемую матрицей функционала B в базисе $e_1, \ldots, e_n$ (и ранг которой — независимо от выбора базиса $e_1, \ldots, e_n$ — называется рангом функционала B).

При этом для значения $B(\boldsymbol{x}, \boldsymbol{y})$ функционала B на векторах $\boldsymbol{x} = x^i e_i$ и $\boldsymbol{y} = y^j e_j$ имеет место формула

$$B(\boldsymbol{x}, \boldsymbol{y}) = b_{ij} x^i y^j, \tag{2}$$

означающая, что это значение является билинейной формой от координат векторов $\boldsymbol{x}$ и $\boldsymbol{y}$. В матричных обозначениях формула (2) имеет вид

$$B(\boldsymbol{x}, \boldsymbol{y}) = x^{\mathsf{T}} B y,$$

где x и y — столбцы координат векторов $\boldsymbol{x}$ и $\boldsymbol{y}$.

Симметричность билинейного функционала B по определению означает, что

$$B(y, x) = B(x, y)$$

для любых векторов $x, y \in \mathscr{V}$. При $x = e_i$ и $y = e_j$ отсюда в частности следует, что $b_{ij} = b_{ji}$, т. е. что матрица (1) симметрична.

Обратно, если матрица (1) симметрична, то для любых векторов x и y имеет место равенство

$$\begin{aligned} B(y, x) &= b_{ij} y^i x^j \\ &= b_{ji} y^j x^i \quad \text{(переименовали индексы)} \\ &= b_{ij} x^i y^j \quad \text{(использовали симметричность)} \\ &= B(x, y), \end{aligned}$$

показывающее, что функционал B симметричен (см. соответствующее рассуждение в лекции 8 для кососимметрических функционалов).

Таким образом, *билинейный функционал B тогда и только тогда симметричен, когда симметрична его матрица.*

Замечание 1. Так же как и для кососимметрических функционалов, аналогичное утверждение справедливо для симметрических функционалов любой степени, т. е. функционал тогда и только тогда симметричен, когда его коэффициенты не меняются при любой перестановке индексов.

Определение 1. *Функционал* Q: $x \longmapsto Q(x) \in \mathbb{K}$ называется *квадратичным*, если существует такой билинейный функционал B, что

$$Q(x) = B(x, x) \tag{3}$$

для любого вектора $x \in \mathscr{V}$.

Ясно, что все квадратичные функционалы составляют линейное пространство (являющееся подпространством линейного пространства всевозможных функций $\mathscr{V} \to \mathbb{K}$).

Представив B в виде суммы симметрического и кососимметрического функционалов (см. формулу (12) лекции 5) и приняв во внимание, что $A(x, x) = 0$ для любого кососимметрического функционала A, мы немедленно получим, что без ограничения общности функционал B в формуле (3) можно считать симметрическим.

Легко видеть, что тогда функционал B однозначно восстанавливается по функционалу Q, т. е., другими словами, *соответствие*

$$B \longmapsto Q$$

является биективным соответствием между линейными пространствами симметрических билинейных и квадратичных функционалов. Действительно, ясно, что если $Q(x) = B(x, x)$ и функционал B симметричен, то

$$\frac{Q(x+y) - Q(x) - Q(y)}{2} = B(x, y)$$

для любых векторов $x, y \in \mathscr{V}$ (напомним, что по условию $\operatorname{char} \mathbb{K} \neq 2$). □

Поэтому в принципе совершенно безразлично, рассматривать симметрические билинейные или квадратичные функционалы: любое утверждение о квадратичных функционалах можно переформулировать как утверждение о симметрических билинейных функционалах, и наоборот. В качестве основных мы выберем квадратичные функционалы, переформулировку же утверждений о них на язык симметрических билинейных функционалов мы оставим читателю.

Для упрощения обозначений симметрический билинейный функционал, соответствующий квадратичному функционалу Q, мы будем обозначать тем же символом Q. Его ранг мы будем называть *рангом* квадратичного функционала Q.

В каждом базисе $e_1, \ldots, e_n$ пространства $\mathscr{V}$ квадратичный функционал Q задается его матрицей

$$\left\| \begin{matrix} q_{11} & \cdots & q_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ q_{n1} & \cdots & q_{nn} \end{matrix} \right\|, \tag{4}$$

элементы которой определяются формулой

$$q_{ij} = Q(e_i, e_j), \qquad i, j = 1, \ldots n.$$

Матрица (4) является квадратной симметрической матрицей порядка n, и соответствие

«функционал» $\longmapsto$ «его матрица»

представляет собой биективное соответствие между множеством всех квадратичных функционалов на $\mathscr{V}$ и множеством всех симметрических матриц порядка n с элементами из поля $\mathbb{K}$. Это соответствие зависит от выбора

базиса: в другом базисе матрица (4) умножится слева и справа на матрицы C^{T} и C, где C — матрица перехода. Считая базис фиксированным, мы, чтобы не вводить новых букв, будем обозначать матрицу (4) тем же символом Q, что и квадратичный функционал.

Поскольку ковекторы из $\mathscr{V}'$ являются, по определению, $\mathbb{K}$-значными функциями на $\mathscr{V}$, для любых двух ковекторов $\xi, \eta \in \mathscr{V}'$ формула

$$(\xi\eta)(x) = \xi(x)\eta(x), \qquad x \in \mathscr{V},$$

определяет их *произведение* $\xi\eta$: $\mathscr{V} \to \mathbb{K}$. Сравнив определения, мы немедленно получим, что *функционал* $\xi\eta$ *является квадратичным функционалом, отвечающим* (*не симметрическому*!) *билинейному функционалу* $\xi \otimes \eta$.

В частности, билинейному функционалу $e^i \otimes e^j$ отвечает квадратичный функционал $e^i e^j$ и, значит, билинейному функционалу $q_{ij} e^i \otimes e^j$ — квадратичный функционал $q_{ij} e^i e^j$. Этим доказано, что *любой квадратичный функционал* Q *выражается через ковекторы* $e^1, \ldots, e^n$ *сопряженного базиса по формуле*

$$\begin{aligned} Q = q_{ij} e^i e^j = \\ = q_{11}(e^1)^2 + 2q_{12}e^1e^2 + \ldots + 2q_{1n}e^1e^n + \\ + q_{22}(e^2)^2 + \ldots + 2q_{2n}e^2e^n + \\ \cdots\cdots\cdots\cdots \\ + q_{nn}(e^n)^2, \end{aligned}$$

где q_{ij} — *элементы матрицы* (4).

Для значения $Q(x)$ функционала Q на произвольном векторе $x = x^i e_i$ пространства $\mathscr{V}$ отсюда вытекает формула

$$\begin{aligned} Q(x) = q_{ij}x^ix^j = \\ = q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \\ + q_{22}(x^2)^2 + \ldots + 2q_{2n}x^2x^n + \\ \cdots\cdots\cdots\cdots \\ + q_{nn}(x^n)^2, \end{aligned}$$

или, в матричных обозначениях, — формула

$$Q(x) = x^{\mathrm{T}} Q x, \tag{5}$$

где x — столбец координат вектора x, а Q — матрица (4).

Определение 2. Многочлен $Q(x^1, \ldots, x^n)$ от переменных $x^1, \ldots, x^n$ называется *квадратичной формой*, если он однороден (все его члены имеют одну и ту же степень) и его степень равна 2. (См. лекцию I. 12.)

Любая квадратичная форма имеет вид

$$
\begin{aligned}
Q(x^1, \ldots, x^n) = q_{ij}x^ix^j = \\
= q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \\
+ q_{22}(x^2)^2 + \ldots + 2q_{2n}x^2x^n + \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
+ q_{nn}(x^n)^2
\end{aligned}
$$

и потому однозначно определяется матрицей

$$
\begin{Vmatrix} q_{11} & q_{12} & \ldots & q_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ q_{n1} & q_{n2} & \ldots & q_{nn} \end{Vmatrix}, \tag{6}
$$

которая называется *матрицей* этой формы.

Таким образом, мы видим, что *значение $Q(x)$ произвольного квадратичного функционала Q на векторе $x \in \mathscr{V}$ выражается квадратичной формой*

$$
Q(x) = Q(x^1, \ldots, x^n)
$$

от координат $x^1, \ldots, x^n$ этого вектора.

Это устанавливает (зависящее от выбора базиса) биективное соответствие между квадратичными функционалами и квадратичными формами.

Определение 3. Две квадратичные формы называются *эквивалентными*, если они соответствуют одному и тому же квадратичному функционалу в различных базисах.

Можно также сказать, что две квадратичные формы эквивалентны, если для их матриц Q_1 и Q_2 имеет место равенство вида

$$
Q_2 = C^{\mathrm{T}}Q_1C,
$$

где C — некоторая невырожденная матрица.

Если же ввести в рассмотрение однородные линейные преобразования

$$
\begin{aligned}
y^1 = c_1^1x^1 + \ldots + c_n^1x^n, \\
\cdots\cdots\cdots\cdots \\
y^n = c_1^nx^1 + \ldots + c_n^nx^n
\end{aligned} \tag{7}
$$

с невырожденными матрицами

$$
\begin{vmatrix} c_1^1 & \ldots & c_n^1 \\ \cdot & \cdot & \cdot \\ c_1^n & \ldots & c_n^n \end{vmatrix} \neq 0,
$$

то можно будет сказать, что форма $Q_1(x^1, \ldots, x^n)$ эквивалентна форме $Q_2(x^1, \ldots, x^n)$, если существует такое

преобразование (7), что, обозначив переменные формы Q_1 символами $y^1, \ldots, y^n$ и подставив вместо них выражения (7), мы получим форму Q_2.

В соответствии с общими определениями лекции 5 векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ пространства $\mathscr{V}$ называются *Q-ортогональными*, если $Q(\boldsymbol{x}, \boldsymbol{y}) = 0$.

Базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ пространства $\mathscr{V}$ мы будем называть *Q-ортогональным*, если

$$Q(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0 \quad \text{при} \quad i \neq j. \tag{8}$$

Теорема 1 (теорема Лагранжа). *Для любого квадратичного функционала Q на $\mathscr{V}$ существует Q-ортогональный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ пространства $\mathscr{V}$.*

Доказательство. Мы не только докажем эту теорему, но и укажем практический алгоритм, позволяющий произвольный базис пространства $\mathscr{V}$ преобразовать в базис, обладающий свойством (8).

Этот алгоритм называется алгоритмом Лагранжа. Он состоит в последовательном применении трех элементарных преобразований, одно из которых мы назовем *основным*, а два остальных — *вспомогательными*.

Основное преобразование Лагранжа. Это преобразование применяется к базису $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, если

$$q_{11} = Q(\boldsymbol{e}_1) \neq 0.$$

Оно переводит этот базис в базис

$$\begin{aligned} \boldsymbol{e}'_1 &= \boldsymbol{e}_1, \\ \boldsymbol{e}'_2 &= -\frac{q_{12}}{q_{11}}\boldsymbol{e}_1 + \boldsymbol{e}_2, \\ &\cdots\cdots\cdots \\ \boldsymbol{e}'_n &= -\frac{q_{1n}}{q_{11}}\boldsymbol{e}_1 + \boldsymbol{e}_n. \end{aligned} \tag{9}$$

Полученный базис обладает тем свойством, что его первый вектор Q-ортогонален всем остальным:

$$Q(\boldsymbol{e}'_1, \boldsymbol{e}'_i) = 0 \quad \text{при} \quad i > 1.$$

Действительно,

$$Q(\boldsymbol{e}'_1, \boldsymbol{e}'_i) = Q\left(\boldsymbol{e}_1, -\frac{q_{1i}}{q_{11}}\boldsymbol{e}_2 + \boldsymbol{e}_i\right) = -\frac{q_{1i}}{q_{11}}q_{11} + q_{1i} = 0.$$

Если теперь $q'_{22} = Q(e'_2, e'_2) \neq 0$, то применяя к векторам $e'_2, \ldots, e'_n$ (т. е., точнее, к ограничению функционала Q на подпространстве $[e'_2, \ldots, e'_n]$) то же преобразование, мы получим базис $e''_1, e''_2, \ldots, e''_n$, первые два вектора которого e''_1 и e''_2 будут Q-ортогональны друг другу и остальным векторам, и т. д.

В случае, когда это построение до самого конца не останавливается, т. е. каждый раз (пока мы не исчерпаем базис или не получится нулевой функционал) основное преобразование применимо, теорема 1 окажется тем самым доказанной. Этот случай называется *регулярным*.

Если же на некотором этапе основное преобразование (9) оказывается неприменимым, то следует сделать вспомогательные преобразования, в результате которых получается базис, к которому преобразование (9) уже применимо.

Первое вспомогательное преобразование. Это преобразование применяется в случае, когда $q_{11} = 0$, но существует такой индекс i_0, что $q_{i_0 i_0} \neq 0$. Оно состоит в перестановке i_0-го вектора базиса на первое место:

$$e'_1 = e_{i_0}, \quad e'_{i_0} = e_1,$$

$$e'_i = e_i, \text{ если } i \neq 1, i_0.$$

Очевидно, что в новом базисе $q'_{11} \neq 0$.

Второе вспомогательное преобразование. Это преобразование применяется в случае, когда $q_{ii} = 0$ для всех $i = 1, \ldots, n$, но функционал Q не нулевой и потому существуют такие индексы i_0 и j_0, что $q_{i_0 j_0} \neq 0$. Если, например, $q_{12} \neq 0$ (это предположение общности, конечно, не ограничивает), то рассматриваемое преобразование задается формулами

$$e'_1 = e_1 + e_2,$$

$$e'_i = e_i, \text{ если } i \geqslant 2.$$

Тогда

$$q'_{11} = Q(e'_1) = Q(e_1 + e_2, e_1 + e_2) = 2q_{12} \neq 0$$

и можно применять основное преобразование.

Ни одно из этих преобразований не применимо только тогда, когда все коэффициенты q_{ij} равны нулю, т. е.

когда $Q=0$. Но в этом случае любой базис, очевидно, Q-ортогонален, и потому делать с ним ничего не надо.

Следовательно, применяя в нужной последовательности наши преобразования, мы рано или поздно получим Q-ортогональный базис. □

В Q-ортогональном базисе матрица формы Q, очевидно, диагональна, т. е. имеет вид

$$\left\| \begin{matrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{matrix} \right\|, \tag{10}$$

и, значит,

$$Q(x)=\lambda_1(x^1)^2+\ldots+\lambda_n(x^n)^2$$

для любого вектора x. Поэтому на языке квадратичных форм теорема Лагранжа утверждает, что *любая квадратичная форма* $Q(x^1, \ldots, x^n)$ *эквивалентна форме вида*

$$\lambda_1(x^1)^2+\ldots+\lambda_n(x^n)^2. \tag{11}$$

О форме (11) говорят, что она имеет *нормальный вид*. Таким образом, мы видим, что *любая квадратичная форма* $Q(x^1, \ldots, x^n)$ *невырожденным линейным преобразованием* (7) *может быть приведена к нормальному виду* (11).

Последнее утверждение, также известное как теорема Лагранжа, полностью относится к алгебре, и в нем исчезли все следы его геометрического происхождения. Поэтому оно применимо к квадратичным формам, возникающим в любых вопросах (скажем, в механике), априори никак не связанных с геометрией квадратичных функционалов.

На практике приведение квадратичной формы $Q(x^1, \ldots, x^n)$ к нормальному виду следует проводить, последовательно «выделяя квадраты», т. е. пользуясь тождеством

$$Q(x^1, \ldots, x^n)=\frac{1}{q_{11}}(q_{11}x^1+\ldots+q_{1n}x^n)^2+Q'(x^2, \ldots, x^n),$$

где форма Q' уже не содержит, как легко видеть, переменной x^1. Это тождество соответствует основному преобразованию Лагранжа. В нерегулярном случае, кроме

того, приходится перенумеровывать переменные и пользоваться преобразованиями вида

$$y_1 = x_1 - x_2,$$
$$y_i = x_i, \text{ если } i \geqslant 2.$$

Часть коэффициентов $\lambda_1, \ldots, \lambda_n$ (или все они) формы (11) может быть равна нулю. Ясно, что число r отличных от нуля этих коэффициентов равно рангу матрицы (6) и, следовательно, рангу функционала Q. Переставляя, если нужно, элементы базиса, мы всегда можем добиться того, чтобы были отличны от нуля первые коэффициенты $\lambda_1, \ldots, \lambda_r$. Поскольку членов с коэффициентами, равными нулю, писать не нужно, мы окончательно получаем, что *нормальным видом квадратичной формы ранга r является форма*

$$\lambda_1 (x^1)^2 + \ldots + \lambda_r (x^r)^2, \text{ где } \lambda_1 \neq 0, \ldots, \lambda_r \neq 0.$$