Р. КУРАНТ и Д. ГИЛЬБЕРТ

МЕТОДЫ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

ТОМ ПЕРВЫЙ

ГТТИ • 1933

Глава I.

Алгебра линейных преобразований и квадратичных форм.

Многочисленные вопросы математического анализа, с которыми нам придется иметь дело в этом томе, самым тесным образом связаны как в смысле аналогии, так и в смысле внутренней зависимости с теорией линейных преобразований и квадратичных форм. Поэтому мы рассмотрим сначала со всей возможной краткостью именно эту область с точки зрения, имеющей для нас здесь важное значение, причем мы предполагаем у читателя некоторое знакомство с затронутыми вопросами.

§ 1. Линейные уравнения и линейные преобразования.

1. Векторы. Для того чтобы иметь возможность кратко формулировать известные факты из теории линейных уравнений, целесообразно ввести простейшие обозначения векторного исчисления[1]). Систему n действительных чисел $x_1, \ldots, x_n$ мы называем *n-мерным вектором* или вектором в пространстве n измерений и сокращенно обозначаем соответствующей немецкой буквой $\mathfrak{x}$. Числа x_i $(i=1, \ldots, n)$ называются *компонентами* вектора $\mathfrak{x}$. Если все компоненты обращаются в нуль, то мы говорим о *нулевом векторе*. При $n=2$ или при $n=3$ простейшим геометрическим истолкованием вектора является, как известно, „радиус-вектор", идущий из начала координат прямоугольной системы к точке с прямоугольными координатами x_i. При $n>3$, правда, геометрически наглядного образа нет, но пользоваться геометрическим способом выражения по сути дела удобно и в этом случае.

Если заданы два произвольных действительных числа λ и μ, то под вектором $\lambda\mathfrak{x}+\mu\mathfrak{y}=\mathfrak{z}$ мы разумеем вектор, компоненты z_i которого линейно составлены из компонент x_i и y_i векторов $\mathfrak{x}$ и $\mathfrak{y}$ по формуле

$$z_i=\lambda x_i+\mu y_i.$$

Тем самым в частности определена сумма и разность двух векторов. *Скалярным произведением* $(\mathfrak{x}\mathfrak{y})$ векторов $\mathfrak{x}$ и $\mathfrak{y}$ мы называем число

$$(\mathfrak{x}\mathfrak{y})=x_1y_1+\ldots+x_ny_n=y_1x_1+\ldots+y_nx_n=(\mathfrak{y}\mathfrak{x}). \tag{1}$$

Иногда мы будем называть скалярное произведение $(\mathfrak{x}\mathfrak{y})$ *компонентой вектора $\mathfrak{y}$ относительно вектора $\mathfrak{x}$*, или же наоборот.

---

[1]) При этом здесь речь идет только о сокращенном способе обозначений, а не об изложении собственно векторного анализа или его обобщения на случай измерений, где основным пунктом исследования является вопрос об известных инвариантах.

Если скалярное произведение $(\mathfrak{x}\mathfrak{y})$ обращается в нуль, то векторы $\mathfrak{x}$ и $\mathfrak{y}$ мы будем называть перпендикулярными или *ортогональными* друг к другу; при $n=2$ и $n=3$ этот способ выражения имеет непосредственное наглядное значение. Особенное значение имеет скалярное произведение $N\mathfrak{x}=(\mathfrak{x}\mathfrak{x})=\mathfrak{x}^2$ вектора на самого себя, которое мы называем *нормом* вектора. Положительный квадратный корень из $\mathfrak{x}^2$ называют *абсолютным значением* или *длиной* вектора $\mathfrak{x}$ и обозначают так:

$$|\mathfrak{x}|=\sqrt{\mathfrak{x}^2}.$$

Вектор, длина которого равна единице, называется *нормированным* вектором или *единичным* вектором.

Скалярное произведение $(\mathfrak{a}\mathfrak{b})$ двух векторов $\mathfrak{a}=(a_1, \ldots, a_n)$ и $\mathfrak{b}=(b_1, \ldots, b_n)$ удовлетворяет следующему неравенству:

$$(\mathfrak{a}\mathfrak{b})^2 \leqslant \mathfrak{a}^2\mathfrak{b}^2,$$

или, не пользуясь векторными обозначениями:

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right),$$

причем знак равенства имеет место в том и только в том случае, когда числа a_i пропорциональны числам b_i, т. е. имеет место соотношение:

$$\lambda\mathfrak{a}+\mu\mathfrak{b}=0.$$

Доказательство этого *неравенства Шварца* [1]) вытекает из замечания, что квадратное уравнение

$$\sum_{i=1}^{n}(a_i x+b_i)^2=x^2\sum_{i=1}^{n} a_i^2+2x\sum_{i=1}^{n} a_i b_i+\sum_{i=1}^{n} b_i^2=0$$

не может иметь двух различных действительных корней, а имеет, за исключением случая пропорциональности чисел a_i и чисел b_i, мнимые корни. Неравенство Шварца представляет соответствующее этому факту соотношение для дискриминанта квадратного уравнения. Другое доказательство неравенства Шварца непосредственно вытекает из тождества:

$$\sum_{i=1}^{n} a_i^2\sum_{i=1}^{n} b_i^2-\left(\sum_{i=1}^{n} a_i b_i\right)^2=\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}(a_j b_k-a_k b_j)^2.$$

Векторы $\mathfrak{x}_1, \ldots, \mathfrak{x}_m$ называются *линейно независимыми*, если невозможно найти числа $\lambda_1, \ldots, \lambda_m$, которые не равнялись бы одновременно нулю, так чтобы имело место векторное равенство:

$$\lambda_1\mathfrak{x}_1+\ldots+\lambda_m\mathfrak{x}_m=0,$$

т. е. чтобы все компоненты вектора, стоящие в левой части, равнялись нулю. В противном случае векторы называются *линейно зависимыми*.

---

1) Впрочем, этим соотношением пользовался еще Коши.

В n-мерном пространстве n векторов $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$, компоненты которых по порядку задаются строками схемы

$$\begin{matrix} 1, 0, \ldots, 0 \\ 0, 1, \ldots, 0 \\ \ldots\ldots\ldots\ldots \\ 0, 0, \ldots, 1, \end{matrix}$$

образуют систему n линейно независимых векторов.

В самом деле, если бы имело место соотношение $\lambda_1\mathfrak{e}_1 + \lambda_2\mathfrak{e}_2 + \ldots \ldots + \lambda_n\mathfrak{e}_n = 0$, то, умножив его скалярно на $\mathfrak{e}_h$, мы получили бы тотчас же $\lambda_h = 0$, так как $\mathfrak{e}_h^2 = 1$, а $(\mathfrak{e}_h\mathfrak{e}_k) = 0$ $(h \neq k)$.

В то время как безусловно существуют системы n линейно независимых векторов, между $n+1$ векторами $\mathfrak{u}_1, \ldots, \mathfrak{u}_{n+1}$ непременно должно иметь место, по крайней мере, одно линейное уравнение:

$$\mu_1\mathfrak{u}_1 + \ldots + \mu_{n+1}\mathfrak{u}_{n+1} = 0,$$

в котором не все коэфициенты равны нулю, так как система n однородных линейных уравнений

$$\sum_{i=1}^{n+1} u_{ik}\mu_i = 0 \quad (k = 1, \ldots, n)$$

относительно $n+1$ неизвестных $\mu_1, \ldots, \mu_{n+1}$ всегда имеет нетривиальное решение (см. п. 3).

2. Ортогональные системы векторов. Полнота системы. Предыдущие *„координатные векторы"* $\mathfrak{e}_i$ представляют специальную систему *„ортогональных единичных векторов"*. Мы разумеем под единичным вектором, как и выше, вектор, длина которого равна единице, а под системой n ортогональных единичных векторов $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ — такую систему, когда выполнены условия:

$$\mathfrak{e}_h^2 = 1, \quad (\mathfrak{e}_h\mathfrak{e}_k) = 0 \quad (h \neq k)$$

при $h, k = 1, 2, \ldots, n$. Так же как и раньше, можно заключить, что n векторов $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ линейно независимы.

Если имеем произвольный вектор $\mathfrak{x}$, то в силу линейной зависимости $n+1$ векторов должно иметь место соотношение:

$$c_0\mathfrak{x} - c_1\mathfrak{e}_1 - \ldots - c_n\mathfrak{e}_n = 0,$$

в котором не все постоянные c равны нулю; при этом c_0 не может равняться нулю, так как векторы $\mathfrak{e}_i$ линейно независимы, и потому можно принять c_0 равным единице. Всякий вектор $\mathfrak{x}$ можно, следовательно, представить при помощи системы ортогональных единичных векторов в виде:

$$\mathfrak{x} = c_1\mathfrak{e}_1 + c_2\mathfrak{e}_2 + \ldots + c_n\mathfrak{e}_n. \tag{2}$$

Значение коэфициентов c_i *компонент вектора* $\mathfrak{x}$ *относительно системы* $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ находим скалярным умножением равенства (2) на $\mathfrak{e}_i$, а именно:

$$c_i = (\mathfrak{x}\mathfrak{e}_i).$$

Из произвольной системы m линейно независимых векторов $\mathfrak{v}_1, \ldots, \mathfrak{v}_m$ мы можем получить систему m ортогональных единичных векторов $\mathfrak{e}_1, \ldots, \mathfrak{e}_m$ с помощью следующего *процесса ортогонализации*.

Полагаем $\mathfrak{e}_1 = \mathfrak{v}_1/|\mathfrak{v}_1|$. Затем мы выбираем два не обращающиеся одновременно в нуль числа c_1' и c_2' так, чтобы вектор $c_1'\mathfrak{e}_1 + c_2'\mathfrak{v}_2$ был ортогонален к $\mathfrak{e}_1$, т. е. чтобы $c_1' + c_2'(\mathfrak{e}_1\mathfrak{v}_2) = 0$. Вследствие линейной независимости векторов $\mathfrak{v}_1$ и $\mathfrak{v}_2$, а вместе с тем и $\mathfrak{e}_1$ и $\mathfrak{v}_2$ вектор $c_1'\mathfrak{e}_1 + c_2'\mathfrak{v}_2$ не равен нулю; разделив этот вектор на его длину, получим единичный вектор $\mathfrak{e}_2$, ортогональный к $\mathfrak{e}_1$. Далее, выбираем три не обращающиеся одновременно в нуль числа c_1'', c_2'', c_3'' так, чтобы вектор $c_1''\mathfrak{e}_1 + c_2''\mathfrak{e}_2 + c_3''\mathfrak{v}_3$ был ортогонален к векторам $\mathfrak{e}_1$ и $\mathfrak{e}_2$, т. е. чтобы $c_1'' + c_3''(\mathfrak{v}_3\mathfrak{e}_1) = 0$ и $\mathfrak{e}_2'' + c_3''(\mathfrak{v}_3\mathfrak{e}_2) = 0$. Вектор $c_1''\mathfrak{e}_1 + c_2''\mathfrak{e}_2 + c_3''\mathfrak{v}_3$ опять отличен от нуля и потому может быть нормирован, и мы получаем таким образом $\mathfrak{e}_3$.

Путем продолжения этого процесса мы приходим к искомой ортогональной системе.

При $m < n$ говорят о неполной, а при $m = n$ о *полной ортогональной системе*. Если обозначим компоненты вектора $\mathfrak{x}$ относительно $\mathfrak{e}_1, \ldots, \mathfrak{e}_m$ опять через $c, \ldots, c_m$, то из очевидного соотношения:

$$(\mathfrak{x} - c_1\mathfrak{e}_1 - \ldots - c_m\mathfrak{e}_m)^2 \geqslant 0$$

следует, если разложить квадрат вектора в левой части по имеющим место и здесь правилам алгебры, что

$$\mathfrak{x}^2 - 2(\mathfrak{x}\sum_{i=1}^{m} c_i\mathfrak{e}_i) + \sum_{i=1}^{m} c_i^2 = \mathfrak{x}^2 - 2\sum_{i=1}^{m} c_i^2 + \sum_{i=1}^{m} c_i^2 \geqslant 0$$

или

$$\mathfrak{x}^2 \geqslant \sum_{i=1}^{m} c_i^2, \tag{3}$$

причем $m \leqslant n$ и $c_i = (\mathfrak{x}\mathfrak{e}_i)$. При $m = n$ имеет место знак равенства:

$$\mathfrak{x}^2 = \sum_{i=1}^{m} c_i^2. \tag{4}$$

Оба последних соотношения, которые содержат векторное выражение теоремы Пифагора и при $n \leqslant 3$ имеют непосредственное геометрическое истолкование, обычно называют первое — *неравенством Бесселя*, а второе — *условием полноты*. Действительно, равенство (4) выражает в том случае, когда оно справедливо для любого вектора, что заданная ортогональная система является полной системой. В самом деле, равенство (4) не могло бы иметь места для нормированного вектора, ортогонального ко всем векторам $\mathfrak{e}_1, \ldots, \mathfrak{e}_m$, а такой вектор непременно существует, если $m < n$.

Можно, впрочем, записать условие полноты в более общем виде:

$$(\mathfrak{x}\mathfrak{x}') = \sum_{i=1}^{m} c_i c'_i, \tag{5}$$

который легко получается из ортогональности векторов $\mathfrak{e}_i$.

Все эти алгебраические соотношения имеют преимущественно формальный характер. Они приобретают более глубокое значение благодаря тому, что они снова встречаются формально совершенно аналогичным образом в трансцендентных вопросах анализа, где они уже не являются тривиальными.

3. Линейные преобразования, матрицы. При помощи системы n линейных уравнений:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = y_2, \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = y_n \end{array}\right\} \tag{6}$$

с заданными коэфициентами a_{ik} мы каждой системе значений $x_1, x_2, \ldots, x_n$ однозначно относим систему значений $y_1, y_2, \ldots, y_n$. Эту операцию называют *линейным преобразованием* системы значений $x_1, x_2, \ldots, x_n$ в систему $y_1, y_2, \ldots, y_n$ или короче, вектора $\mathfrak{x}$ в вектор $\mathfrak{y}$. Линейный характер преобразования выражается в том, что вектору $\lambda_1\mathfrak{x}_1 + \lambda_2\mathfrak{x}_2$ соответствует вектор $\lambda_1\mathfrak{y}_1 + \lambda_2\mathfrak{y}_2$.

Самой важной из задач, встречающихся при линейных преобразованиях, является задача об их обращении, иными словами: вопрос о решении системы линейных уравнений. Ответ на этот вопрос дает следующая основная теорема из теории линейных уравнений, доказательство которой мы можем предполагать известным.

Система уравнений

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = y_2, \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = y_n \end{array}$$

или, короче,

$$\sum_{n=1}^{n} a_{ik}x_k = y_i \quad (i = 1, \ldots, n) \tag{7}$$

либо имеет при заданных значениях a_{ik} для любого, произвольно заданного вектора $\mathfrak{y}$ одно и только одно решение $\mathfrak{x}$, в частности при $\mathfrak{y} = 0$ решение $\mathfrak{x} = 0$; либо же однородные уравнения, получающиеся из системы (7) *при $\mathfrak{y} = 0$, имеют положительное число ρ „нетривиальных“, т. е. отличных от нулевого вектора, линейно независимых между собой решений $\mathfrak{x}_1, \mathfrak{x}_2 \ldots, \mathfrak{x}_\rho$, которые мы можем считать нормированными. В этом случае и „транспонированная“ система уравнений*

$$\sum_{k=1}^{n} a'_{ik}x'_k = 0 \quad (i = 1, \ldots, n), \tag{8}$$

где $a'_{ik} = a_{ki}$, *имеет* ρ *линейно независимых, нетривиальных решений* $\mathfrak{x}'_1, \ldots, \mathfrak{x}'_\rho$. *Для тех, и только для тех векторов, которые удовлетворяют* ρ *соотношениям:* $(\mathfrak{y}\mathfrak{x}'_1) = 0, \ldots, (\mathfrak{y}\mathfrak{x}'_\rho) = 0$, *т. е. ортогональны к* $\mathfrak{x}'_1, \ldots, \mathfrak{x}'_\rho$, *может быть решена и неоднородная система* (7), *причем это решение определяется только с точностью до произвольного аддитивно входящего решения однородной системы уравнений.*

При этой формулировке основной теоремы мы не ссылаемся на теорию определителей. Определители нужны лишь для того, чтобы представить решение системы уравнений в явном виде, что мы сейчас и сделаем.

Самое существенное в таком линейном преобразовании дается таблицей коэфициентов или *матрицей* уравнений (7):

$$A = (a_{ik}) = \begin{pmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1} a_{n2} \ldots a_{nn} \end{pmatrix} \tag{9}$$

с определителем

$$\mathrm{A} = |a_{ik}| = \begin{vmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1} a_{n2} \ldots a_{nn} \end{vmatrix}.$$

В иных случаях целесообразно обозначить само преобразование или, как говорят также, тензор или оператор отдельной буквой $\mathfrak{A}$. Элементы a_{ik} матрицы A называются компонентами тензора. Мы можем рассматривать линейное преобразование (7) как „умножение“ тензора $\mathfrak{A}$ на вектор $\mathfrak{x}$ и символически записать в виде:

$$\mathfrak{A}\mathfrak{x} = \mathfrak{y}.$$

Многие положения „линейной алгебры“ особенно удобно выражаются, если их формулировать как теоремы о матрицах, пользуясь при этом некоторыми простыми определениями и правилами, которые носят название: *исчисление матриц*. Прежде всего мы приходим к понятию об умножении матриц, рассматривая вектор $\mathfrak{x}$, который требуется преобразовать при помощи предыдущих уравнений (7), в свою очередь, как произведение другого тензора $\mathfrak{B}$ с компонентами b_{ik} на вектор $\mathfrak{w}$; итак, пусть $\mathfrak{x}$ и $\mathfrak{w}$ связаны между собой системой линейных уравнений:

$$\sum_{k=1}^{n} b_{ik} w_k = x_i \quad (i = 1, \ldots, n)$$

с матрицей $B = (b_{ik})$, тогда и вектор $\mathfrak{y}$ получается из вектора $\mathfrak{w}$ путем умножения на тензор $\mathfrak{C}$. Матрица C этого тензора получается из A и B по правилам „*умножения матриц*“:

$$C = AB,$$

т. е. элемент c_{ik} представляет скалярное произведение i-й строки A и k-го столбца B:

$$c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk} \quad (i,\ k=1,\ \ldots,\ n). \tag{10}$$

Тензор или преобразование $\mathfrak{C}$ называют скалярным произведением или, короче, *произведением тензоров или преобразований* $\mathfrak{A}$ и $\mathfrak{B}$. В дальнейшем мы вместо тензоров всегда будем говорить об эквивалентных им матрицах. Мы видим, что произведение матриц обладает *свойством сочетательности:*

$$(AB)\,C=A\,(BC),$$

так что произведение $A_1A_2\ldots A_h$ произвольного числа матриц, взятых в известном порядке, имеет вполне определенное значение. Если $A_1=A_2=\ldots=A_h=A$, то это произведение записывают в виде h-й *степени* A^h матрицы A.

Однако следует заметить, что *закон переместительности* произведения, вообще говоря, *не* имеет места, так что приходится различать между умножением *слева* и умножением *справа* матриц A и B, причем в общем случае AB не равно BA. Наконец, под матрицей $\lambda A+\mu B$ мы согласно определению разумеем матрицу с компонентами $\lambda a_{ik}+\mu b_{ik}$, под нулевой матрицей — матрицу, все компоненты которой равны нулю [1]). Между прочим, можно непосредственно убедиться в справедливости распределительного закона:

$$(A+B)\,C=AC+BC.$$

Во многих случаях необходимо ввести „единичную матрицу“.

$$E=(e_{ik})=\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Она обладает тем свойством, что для любой матрицы A имеет место равенство:

$$AE=EA=A.$$

Единичной матрице соответствует тождественное преобразование, которое задается уравнениями:

$$x_i=y_i \quad (i=1,\ \ldots,\ n).$$

Нулевую степень любой матрицы A мы согласно определению считаем равной единичной матрице:

$$A^0=E.$$

---

[1]) При оперировании над матрицами необходимо заметить, что из равенства $AB=(0)$ ни в коем случае не вытекает равенство нулю одной из матриц A или B, как это видно из примера $A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$.

Установив определение степени A^k матрицы, можем также дать определение многочленов, аргументом которых является матрица. Если

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$$

есть многочлен m-й степени относительно x, то равенство

$$f(A) = a_0 + a_1 A + \ldots + a_m A^m$$

определяет матрицу $f(A)$ как целую рациональную функцию матрицы A. Определение матрицы как функции $f(A)$ от A можно иногда распространить и на случаи, когда выражение с помощью многочлена уже невозможно. Так, например, матрицу e^A определяют при помощи равенства:

$$B = e^A = E + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{\nu=0}^{\infty} \frac{A^\nu}{\nu!}.$$

При этом такой ряд имеет следующий смысл: сперва нужно взять сумму первых N членов, а затем надо исследовать, сходится ли каждый из n^2 элементов полученной таким путем матрицы при неограниченном возрастании N; матрица, составленная из n^2 предельных значений, принимается в этом случае за значение ряда. В данном частном случае матрицы e^A ряд, как будет далее доказано, всегда сходится.

Особенно важное соотношение между матрицами получается, если выбрать за функцию $f(A)$ „геометрический ряд". Пусть

$$S_m = E + A + A^2 + \ldots + A^m.$$

Умножая это равенство, определяющее матрицу S_m, на A, мы приходим к равенству:

$$S_m A + E = S_m + A^{m+1},$$

из которого следует, что

$$S_m(E - A) = E - A^{m+1}.$$

Если с возрастанием m матрица S_m стремится к определенному пределу S и, следовательно, матрица A^{m+1} стремится к нулевой матрице, то для определенной при помощи бесконечного геометрического ряда

$$S = E + A + A^2 + \ldots = \sum_{\nu=0}^{\infty} A^\nu = f(A)$$

матрицы S имеет место соотношение:

$$S(E - A) = E.$$

Вопрос о том, когда бесконечный геометрический ряд или, как его иногда незывают, *ряд Неймана,* составленный из матриц, сходится, мы рассмотрим в следующем параграфе.

Над многочленами из матриц можно оперировать совершенно таким же образом, как и над обыкновенными многочленами относительно x. Например, из тождества двух многочленов относительно x вытекает соответствующее тождество многочленов для произвольной матрицы A. Так, тождеству

$$x^3+2x^2+3x+4=(x^2+1)(x+2)+(2x+2)$$

соответствует справедливое для всякой матрицы A соотношение:

$$A^3+2A^2+3A+4=(A^2+E)(A+2E)+(2A+2E).$$

Подобным же образом разложению на множители

$$f(x)=a_0+a_1x+\ldots+a_mx^m=a_m(x-x_1)(x-x_2)\ldots(x-x_m),$$

где $x_1, x_2, \ldots, x_m$ суть корни многочлена $f(x)$, соответствует матричное равенство:

$$f(A)=a_0E+a_1A+\ldots+a_mA^m=a_0(A-x_1E)(A-x_2E)\ldots(A-x_mE),$$

имеющее место для любой матрицы A.

Обычно к каждой матрице A с компонентами a_{ik} относят другие матрицы. Элементы матрицы могут иметь и комплексные значения. Если $\overline{a}_{ik}$ является комплексным числом, сопряженным с a_{ik}, то матрицу $\overline{A}=(\overline{a}_{ik})$ называют *сопряженной* матрицей; далее, матрицу $A'=(a_{ki})$, получающуюся из A заменой столбцов и строк, называют *транспонированной* матрицей; наконец, $A^*=\overline{A}'=(\overline{a}_{kl})$ называют *сопутствующей* (begleitende) матрицей; она получается, следовательно, переходом к сопряженным комплексным величинам и заменой строк и столбцов.

Всегда имеет место равенство:

$$(AB)'=B'A',$$

которое легко непосредственно проверить. Матрица называется *симметричной*, если $A=A'$; действительная матрица, для которой

$$AA'=E,$$

называется *ортогональной*, наконец, вообще комплексная матрица называется *унитарной*, если

$$AA^*=E.$$

Обращение линейного преобразования (7), как известно из теории определителей, возможно при произвольных y_i в том и только в том случае, когда определитель $\mathrm{A}=|a_{ik}|$ не равен нулю. В этом случае решение однозначно определяется и выражается соответствующей системой уравнений

$$x_i=\sum\breve{a}_{ik}y_k \quad (i=1, \ldots, n). \tag{11}$$

При этом, как известно,

$$\breve{a}_{ik}=\frac{\mathrm{A}_{kl}}{\mathrm{A}}, \tag{12}$$

где A_{kl} означает минор, соответствующий элементу a_{kl} определителя A. Матрица $\breve{A}=(\breve{a}_{ik})$ называется *взаимной* или *обратной* матрице A и отличается тем, что

$$\breve{A}A=A\breve{A}=E.$$

Однозначно определенную матрицу $\breve{A}$ обозначим теперь A^{-1}; определитель этой матрицы имеет значение A^{-1}. На языке исчисления матриц мы можем, следовательно, охарактеризовать решение системы уравнений с матрицей A, определитель которой не равен нулю, при помощи матрицы $B = A^{-1}$, удовлетворяющей соотношениям:

$$AB = BA = E.$$

4. **Билинейные формы, квадратичные и эрмитовы формы.** Для того чтобы представить в компактной и наглядной форме линейные уравнения (7), можно воспользоваться эквивалентной системе уравнений *билинейной формой*, принадлежащей матрице A. Эта билинейная форма

$$A(u, x) = \sum_{i, k=1}^{n} a_{ik} u_i x_k \tag{13}$$

получается, если умножить линейные относительно $x_1, \ldots, x_n$ формы, стоящие в левых частях уравнений (7), на неопределенные величины $u_1, \ldots, u_n$ и затем сложить. Таким путем вместо системы уравнений (7) получаем одно тождественное относительно величин u уравнение:

$$A(u, x) = E(u, y), \tag{14}$$

где $E(u, y) = \sum_{i=1}^{n} u_i y_i$ представляет билинейную форму единичной матрицы или *билинейную единичную форму*. Под символическим произведением двух билинейных форм $A(u, x)$ и $B(u, x)$ с матрицами A и B разумеют билинейную форму $C(u, x)$ с матрицей $C = AB$. Степень $A^h(u, x)$ называют также h-й итерированной формой. *Обратная билинейная форма* $A^{-1}(u, x)$, имеющая матрицу A^{-1}, может быть представлена на основании теорем теории определителей в виде:

$$A^{-1}(u, x) = -\frac{\mathrm{A}(u, x)}{\mathrm{A}}, \tag{15}$$

где

$$\mathrm{A}(u, x) = \begin{vmatrix} 0 & u_1 & \ldots & u_n \\ x_1 & a_{11} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ x_n & a_{n1} & \ldots & a_{nn} \end{vmatrix} = -\sum_{i, k=1}^{n} \mathrm{A}_{ik} x_i u_k.$$

Особый интерес представляют *симметрические линейные преобразования*, которые характеризуются условием $a_{ki} = a_{ik}$. Для исследования таковых достаточно рассмотреть *квадратичную форму*:

$$A(x, x) = \sum_{i, k=1}^{n} a_{ik} x_i x_k,$$

которая получается из билинейной формы, если положить $u_i = x_i$. В самом деле, из квадратичной формы $A(x, x)$ получаем симметричную билинейную форму:

$$\sum_{i,k=1}^{n} a_{ik}u_i x_k = \frac{1}{2}\sum_{i=1}^{n} u_i \frac{\partial A(x, x)}{\partial x_i} = \frac{A(x+u, x+u) - A(x, x) - A(u, u)}{2},$$

которая называется *полярной формой*, соответствующей квадратичной форме $A(x, x)$.

Если $A(u, x) = \sum a_{ik}u_i x_k$ — произвольная несимметричная билинейная форма (с действительными коэфициентами), то $AA'(u, x)$ и $A'A(u, x)$ во всяком случае представляют симметричные билинейные формы; в самом деле:

$$AA'(u, x) = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_{ik}x_i \sum_{i=1}^{n} a_{ik}u_i\right)$$

и

$$A'A(u, x) = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}x_k \sum_{k=1}^{n} a_{ik}u_k\right).$$

Мы можем, следовательно, образовать также квадратичные формы:

$$AA'(x, x) = \sum_{k=1}\left(\sum_{i=1}^{n} a_{ik}x_i\right)^2$$

и

$$A'A(x, x) = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}x_k\right)^2.$$

Эти квадратичные формы, будучи суммами квадратов, обладают тем свойством, что могут принимать только неотрицательные значения. Такого рода квадратичные формы называются *определенными положительными* квадратичными формами.

Важным обобщением квадратичных форм являются *эрмитовы формы*. Это — билинейные формы:

$$A(u, x) = \sum_{i,k=1}^{n} a_{ik}u_i x_k,$$

коэфициенты которых a_{ik} могут иметь комплексные значения и должны удовлетворять соотношениям:

$$a_{ik} = \overline{a}_{ki}.$$

Следовательно, эрмитова форма принимает действительные значен я если придать u_i значения, комплексно сопряженные с x_i. В этом случае обыкновенно записывают эрмитову форму в виде:

$$H(x, \overline{x}) = \sum_{i,k=1}^{n} a_{ik}x_i\overline{x}_k = \sum_{i,k=1}^{n} a_{ki}\overline{x}_i x_k.$$

Произвольной билинейной форме

$$A(u, x) = \sum_{i, k=1}^{n} a_{ik} u_i x_k$$

с комплексными коэфициентами соответствуют эрмитовы формы:

$$AA^*(x, \bar{x}) = A\bar{A}'(x, \bar{x}) = \sum_{k=1}^{n} \left| \sum_{i=1}^{n} a_{ik} x_i \right|^2$$

и

$$A^*A(x, \bar{x}) = \bar{A}'A(x, \bar{x}) = \sum_{i=1}^{n} \left| \sum_{k=1}^{n} a_{ik} \bar{x}_k \right|^2.$$

Если в билинейной форме

$$A(x, y) = \sum_{i, k=1}^{n} a_{ik} x_i y_k$$

подвергнуть переменные линейным преобразованиям

$$x_i = \sum_{k=1}^{n} c_{ik} \zeta_k \quad \text{и} \quad y_i = \sum_{k=1}^{n} b_{ik} \eta_k,$$

матрицы которых C и B, то

$$A(x, y) = \sum_{i, k=1}^{n} a_{ik} x_i y_k = \sum_{i, j, k, l=1}^{n} a_{ik} c_{ij} b_{kl} \zeta_j \eta_l = \sum_{j, l=1}^{n} p_{jl} \zeta_j \eta_l,$$

$$p_{jl} = \sum_{i, k=1}^{n} a_{ik} c_{ij} b_{kl}.$$

Таким образом получаем из A преобразованную билинейную форму с матрицей

$$(p_{jl}) = C'AB,$$

определитель которой по теореме умножения определителей равен **АВГ**. Если в частности имеем дело с квадратичной формой:

$$K(x, x) = \sum_{p, q=1}^{n} k_{pq} x_p x_q$$

с симметрической матрицей $K = (k_{pq})$ и определителем $\text{К} = |k_{pq}|$, то следует взять $C = B$, и при преобразовании переменных x получается симметрическая матрица $C'KC$ с определителем КГ^2.

5. Ортогональные и унитарные преобразования. Мы ставим себе теперь задачу — найти линейные преобразования $\mathfrak{L}$:

$$x_p = \sum_{q=1}^{n} l_{pq} y_q = L_p(y) \qquad (p = 1, \ldots, n), \tag{16}$$

с матрицей $L=(l_{pq})$ и определителем $\Lambda=|l_{pq}|$, которые были бы *ортогональны*, т. е. переводили бы *единичную квадратичную* форму

$$E(x, x)=\sum_{p=1}^{n} x_p^2$$

в себя самое, удовлетворяя тождественно относительно y_p соотношению

$$E(x, x)=E(y, y). \tag{17}$$

Если применить наши правила преобразования к квадратичной форме $A(x, x)=E(x, x)$, то требование, выражаемое равенством (17), дает в качестве необходимого и достаточного условия ортогональности преобразования $\mathfrak{L}$ равенства:

$$L'EL=L'L=LL'=E, \quad L'=L^{-1}, \tag{18}$$

т. е. транспонированная матрица ортогонального преобразования совпадает с обратной матрицей, так что уравнения (16) разрешаются также с помощью ортогональной системы:

$$y_p=\sum_{q=1}^{n} l_{qp}x_q=L'_p(x). \tag{19}$$

Мы видим, таким образом, что ортогональные преобразования задаются ортогональными матрицами, определение которых дано уже на стр. 9. Подробно условия ортогональности записываются следующим образом:

$$\sum_{\nu=1}^{n} l_{\nu p}^2=1, \quad \sum_{\nu=1}^{n} l_{\nu p}l_{\nu q}=0 \qquad (p\neq q) \tag{20}$$

или

$$\sum_{\nu=1}^{n} l_{p\nu}^2=1, \quad \sum_{\nu=1}^{n} l_{p\nu}l_{q\nu}=0 \qquad (p\neq q) \tag{21}$$

Переходя к определителям, мы из соотношений (18) прежде всего видим, что $\Lambda^2=1$, т. е. определитель ортогонального преобразования равен $+1$ или -1; далее, мы видим, что определитель любой квадратичной формы инвариантен по отношению к ортогональным преобразованиям.

Соотношение $L'(AB)L=(L'AL)(L'BL)$ между матрицами A и B двух произвольных квадратичных форм и матрицей L ортогонального преобразования, вытекающее из (18), показывает, что можно ортогонально преобразовать символическое произведение квадратичных форм путем ортогонального преобразования каждого множителя в отдельности. Отсюда в частности следует, что квадратичные формы, полученные ортогональным преобразованием двух обратных квадратичных форм, также обратны друг другу.

Обобщение предыдущих рассмотрений на случай эрмитовых форм

$$H(x, \bar{x}) = \sum_{p,\, q=1}^{n} h_{pq} x_p \bar{x}_q$$

приводит к *унитарным преобразованиям*. Под унитарным преобразованием

$$x_p = \sum_{q=1}^{n} l_{pq} y_q$$

разумеют такое линейное преобразование (с комплексными коэфициентами l_{pq}), которое переводит единичную эрмитову форму

$$\sum_{p=1}^{n} |x_p|^2 = \sum_{p=1}^{n} x_p \bar{x}_p$$

снова в единичную форму, т. е.

$$\sum_{p=1}^{n} |x_p|^2 = \sum_{p=1}^{n} |y_p|^2.$$

Путем, вполне аналогичным предыдущему, находим в качестве необходимого и достаточного условия унитарности преобразования с матрицей L матричное равенство:

$$LL^* = L^*L = E,$$

где $L^* = \bar{L}'$ — матрица, сопутствующая L. Следовательно, L согласно определению на стр. 9 должно быть унитарной матрицей. *Условия унитарности* можно записать подробно следующим образом:

$$\sum_{\nu=1}^{n} |l_{\nu p}|^2 = 1, \quad \sum l_{\nu p} \bar{l}_{\nu q} = 0 \qquad (p \neq q) \tag{22}$$

и

$$\sum_{\nu=1}^{n} |l_{p\nu}|^2 = 1, \quad \sum l_{p\nu} \bar{l}_{q\nu} = 0 \qquad (q \neq p). \tag{23}$$

Определитель унитарного преобразования по абсолютному значению равен единице, что также непосредственно следует из равенства

$$LL^* = E.$$

§ 2. Линейные преобразования с линейным параметром.

Во многих вопросах система уравнений линейного преобразования представляется в следующем виде:

$$x_i - \lambda \sum_{k=1}^{n} t_{ik} x_k = y_i \quad (i = 1, \ldots, n), \tag{24}$$

где λ — параметр (который может принимать и комплексные значения). Соответствующая билинейная форма имеет вид:

$$E(u, x) - \lambda T(u, x),$$

причем $T(u, x)$ имеет матрицу $T = (t_{ik})$. Решение этой системы уравнений на основании предыдущего параграфа эквивалентно разысканию обратной билинейной формы $R(u, y; \lambda)$ с матрицей R, удовлетворяющей уравнению $(E - \lambda T) R = E$. Мы знаем, что эта обратная матрица существует в том и только в том случае, когда определитель $|E - \lambda T|$ не равен нулю. Так как этот определитель представляет целую рациональную функцию от λ, степени не выше n, то может быть только конечное число значений λ, для которых обратная форма R не существует, именно при значениях λ_i, являющихся корнями этой целой рациональной функции. Эти значения λ_i, *собственные значения* T относительно матрицы E, образуют так называемый спектр матрицы T. (Часто также называют спектром совокупность значений $\varkappa_i = \frac{1}{\lambda_i}$, при которых не существует матрицы, обратной $\varkappa E - T$.)

Вид уравнений (24) наводит на мысль попытаться решить их методом последовательных приближений, подставляя в уравнение

$$x_i = y_i + \lambda \sum_{k=1}^{n} t_{ik} x_k$$

вместо величин x_k снова значения

$$y_k + \lambda \sum_{j=1}^{n} t_{kj} x_j$$

и продолжая поступать таким образом неограниченно. Нагляднее всего представляется этот процесс с помощью соотношения $R = E + \lambda TR$, из которого мы последовательно получаем:

$$R = E + \lambda TR = E + \lambda T + \lambda^2 T^2 R = E + \lambda T + \lambda^2 T^2 + \lambda^3 T^3 R = \ldots$$

Если этот процесс сходится, то мы получаем выражение для R с помощью бесконечного ряда:

$$R = E + \lambda T + \lambda^2 T^2 + \lambda^3 T^3 + \ldots,$$

который действительно (в случае сходимости ряда) дает матрицу, обратную матрице $E - \lambda T$. Чтобы в этом убедиться, достаточно только умножить ряд на $E - \lambda T$, заметив, что в случае сходимости можно выполнить символическое умножение почленно. Непосредственно ясно, что выражение:

$$R = (E - \lambda T)^{-1} = E + \lambda T + \lambda^2 T^2 + \ldots$$

формально вполне совпадает с обыкновенным геометрическим рядом (см. рассуждения на стр. 8, где надо только положить $A = \lambda T$, чтобы получить полное совпадение).

Если представить нашу первоначальную систему уравнений не с помощью матриц, а с помощью соответствующих им билинейных форм в виде:

$$E(u, x) - \lambda T(u, x) = E(u, y),$$

то можно тотчас же записать решение во вполне симметричном к предыдущему виде:

$$E(u, y) + \lambda \mathrm{T}(u, y; \lambda) = E(u, x),$$

если положить

$$\mathrm{T}(u, y; \lambda) = T + \lambda T^2 + \lambda^2 T^3 + \ldots = \frac{R(u, y; \lambda) - E(u, y)}{\lambda}.$$

Билинейную форму Т называют *резольвентой* формы T. Легко доказать сходимость предыдущих рядов Неймана для R и Т при достаточно малых значениях $|\lambda|$.

Пусть M означает верхнюю грань абсолютных значений чисел t_{ik}, тогда для абсолютных значений коэфициентов форм $T^2, T^3, \ldots, T^h$ мы непосредственно получаем верхние грани $nM^2, n^2M^3, \ldots, n^{h-1}M^h$, следовательно, выражение

$$(M + \lambda n M^2 + \lambda^2 n^2 M^3 + \ldots)(|u_1| + \ldots + |u_n|)(|y_1| + \ldots + |y_n|)$$

представляет мажоранту для ряда Неймана. Но эта мажоранта непременно сходится при $|\lambda| < \frac{1}{nM}$. Следовательно, ряд Неймана также сходится при достаточно малых значениях $|\lambda|$ и действительно представляет резольвенту формы $T(u, x)$[1]).

[1]) Сходимость предыдущей мажоранты, очевидно, с возрастанием n все ухудшается. Следует, однако, заметить, что можно при помощи небольшого уточнения легко найти оценку границы сходимости, не зависящую от n, которую можно поэтому применять и при обобщениях на случай бесконечно большого числа переменных. Обозначим элементы матрицы T^{ν} через $t_{ik}^{(\nu)}$ и положим

$$\sum_{\alpha=1}^{n} |t_{p\alpha}^{(1)}| = z_p.$$

Если M представляет верхнюю грань всех n чисел z_p то, как мы покажем методом полной индукции, $\sum_{q=1}^{n} |t_{pq}^{(\nu)}| \leq M^{\nu}$, следовательно, и подавно

$$|t_{pq}^{(\nu)}| \leq M^{\nu}$$

при $p, q = 1, \ldots, n$ и любом значении ν. Отсюда непосредственно следует, что наш ряд Неймана сходится при $|\lambda| < \frac{1}{M}$. Таким образом получена грань, в которую число n явно не входит.

Чтобы доказать предыдущее неравенство для любого ν, будем считать его доказанным для $\nu - 1$.

Заметим, между прочим, что только что произведенная нами оценка показывает, что мы можем во всякий повсюду сходящийся степенной ряд $f(x)=\sum_{\nu=0}^{\infty} c_\nu x^\nu$ подставить вместо x произвольную матрицу A и получить таким путем новую матрицу $f(A)=\sum_{\nu=1}^{\infty} c_\nu A^\nu$. В частности, следовательно, всегда существует матрица e^A.

Полученное нами выражение для R или Т сходится только при достаточно малых $|\lambda|$. Между тем формула (15) предыдущего параграфа дает нам выражение для обратной формы или матрицы $R=(E-\lambda T)^{-1}$, имеющее смысл и вне области сходимости ряда. В самом деле, отождествляя форму $E-\lambda T$ с формой $A(u, x)$, получаем для обратной формы выражение:

$$R(u, y; \lambda)=-\frac{\Delta(u, y; \lambda)}{\Delta(\lambda)},$$

а для резольвенты Т выражение:

$$\mathrm{T}(u, y; \lambda)=-\frac{\Delta(u, y; \lambda)}{\lambda\Delta(\lambda)}-\frac{1}{\lambda}E(u, y),$$

причем

$$\Delta(u, y; \lambda)=\begin{vmatrix} 0 & u_1 & \dots & u_n \\ y_1 & 1-\lambda t_{11} & \dots & -\lambda t_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ y_n & -\lambda t_{n1} & \dots & 1-\lambda t_{nn} \end{vmatrix}$$

есть целая рациональная функция от λ степени $n-1$, а

$$\Delta(\lambda)=\begin{vmatrix} 1-\lambda t_{11} & -\lambda t_{12} & \dots & -\lambda t_{1n} \\ -\lambda t_{21} & 1-\lambda t_{22} & \dots & -\lambda t_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ -\lambda t_{n1} & -\lambda t_{n2} & \dots & 1-\lambda t_{nn} \end{vmatrix}$$

целая рациональная функция степени n. Корни многочлена $\Delta(\lambda)$ образуют, следовательно, определенный выше спектр формы T, т. е. совокупность тех значений λ, для которых форма $E-\lambda T$ не имеет обратной формы.

Тогда

$$\sum_{q=1}^{n}|t_{pq}^{(\nu)}|=\sum_{q=1}^{n}\left|\sum_{\alpha=1}^{n}t_{p\alpha}^{(1)}t_{\alpha q}^{(\nu-1)}\right|\leqslant\sum_{q=1}^{n}\sum_{\alpha=1}^{n}|t_{p\alpha}^{(1)}||t_{\alpha_q}^{(\nu-1)}|=$$
$$=\sum_{\alpha=1}^{n}|t_{p\alpha}^{(1)}|\left(\sum_{q=1}^{n}|t_{\alpha q}^{(\nu-1)}|\right)\leqslant M^{\nu-1}\sum_{\alpha=1}^{n}|t_{p\alpha}^{(1)}|\leqslant M^{\nu}.$$

Так как неравенство справедливо при $\nu=1$, то тем самым оно доказано для любого индекса ν.

При помощи формулы:

$$T+\lambda T^2+\lambda^2 T^3+\ldots=-\frac{\Delta(u,\ y;\ \lambda)}{\lambda\Delta(\lambda)}-\frac{1}{\lambda}E(u,\ y)$$

находящийся в левой части ряд, характер которого непосредственно усмотреть нельзя и который сходится не при всех значениях λ, аналитически продолжен на всю плоскость переменного λ. Обратная форма R, как и резольвента T, является рациональной функцией λ, полюсы которой представляют спектр формы T.

Разлагая определители $\Delta(u,\ y;\ \lambda)$ и $\Delta(\lambda)$ по правилам теории определителей по степеням λ, получаем:

$$\Delta(u,\ y;\ \lambda)=\Delta_1(u,\ y)-\lambda\Delta_2(u,\ y)+\lambda^2\Delta_3(u,\ y)-\ldots$$
$$\ldots+(-1)^{n-1}\lambda^{n-1}\Delta_n(u,\ y),$$
$$\Delta(\lambda)=1-\lambda\Delta_1+\lambda^2\Delta_2-\ldots+(-1)^n\lambda^n\Delta_n,$$

причем

$$\Delta_h(u,\ y)=\sum\begin{vmatrix} 0 & u_{p_1} & \ldots & u_{p_h} \\ y_{p_1} & t_{p_1p_1} & \ldots & t_{p_1p_h} \\ \cdot & \cdot & \cdot & \cdot \\ y_{p_h} & t_{p_hp_1} & \ldots & t_{p_hp_h} \end{vmatrix},$$

$$\Delta_h=\sum\begin{vmatrix} t_{p_1p_1} & t_{p_1p_2} & \ldots & t_{p_1p_h} \\ t_{p_2p_1} & t_{p_2p_2} & \ldots & t_{p_2p_h} \\ \cdot & \cdot & \cdot & \cdot \\ t_{p_hp_1} & t_{p_hp_2} & \ldots & t_{p_hp_h} \end{vmatrix}.$$

При этом суммирование распространяется на все целочисленные значения $p_1,\ p_2,\ \ldots,\ p_h$, от 1 до n, где $p_1<p_2<\ldots<p_h$.

Часто бывает удобно ввести вместо параметра λ обратное значение $\varkappa=\frac{1}{\lambda}$. В этом случае целесообразно рассматривать форму $\varkappa E-T$, с определителем

$$\begin{vmatrix} \varkappa-t_{11} & -t_{12} & \ldots & -t_{1n} \\ -t_{21} & \varkappa-t_{22} & \ldots & -t_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ -t_{n1} & -t_{n2} & \ldots & \varkappa-t_{nn} \end{vmatrix}=\varphi(\varkappa),$$

представляющим целую рациональную функцию степени n от $\varkappa$, корни которой $\varkappa_1,\ \ldots,\ \varkappa_n$, представляют величины, обратные корням $\Delta(\lambda)$, т. е. собственным значениям формы T. Обратная форма $(\varkappa E-T)^{-1}$, которая существует для всех значений $\varkappa$, отличных от $\varkappa_1,\ \ldots,\ \varkappa_n$, может быть представлена при достаточно больших значениях $|\varkappa|$ с помощью ряда Неймана:

$$(\varkappa E-T)^{-1}=\frac{E}{\varkappa}+\frac{T}{\varkappa^2}+\frac{T^2}{\varkappa^3}+\ldots$$

Из этого разложения можно получить интересное следствие. Так как левая часть, на основании предыдущего, является рациональной функцией от $\varkappa$ с знаменателем $\varphi(\varkappa)$, то форма $\varphi(\varkappa)(\varkappa E - T)^{-1}$ должна быть целой рациональной функцией от $\varkappa$, и следовательно, ее разложение по степеням $\varkappa$ не может содержать отрицательных степеней. Поэтому, если мы умножим предыдущее равенство на $\varphi(\varkappa) = \varkappa^n + c_1\varkappa^{n-1} + \ldots$ $\ldots + c_n$, то в правой части должно будет получиться выражение, в котором все коэфициенты при отрицательных степенях $\varkappa$ равны нулю. Коэфициент при $\varkappa^{-1}$, как легко видеть, равен выражению $T^n + c_1 T^{n-1} + \ldots + c_n E$, и мы получаем, таким образом, теорему Кэли (Cayley):

Если $\varphi(\varkappa)$ *есть определитель матрицы* $\varkappa E - T$, *то матрица* T *удовлетворяет соотношению:*

$$\varphi(T) = 0.$$

Другой важный факт относительно спектра, представленного с помощью *характеристических чисел* $\varkappa_1 \ldots \varkappa_n$, выражает следующая теорема:

Если $\varkappa_1, \ldots, \varkappa_n$ *представляют характеристические числа матрицы* T, *а* $g(x)$ — *произвольная целая рациональная функция от* x, *то характеристическими числами матрицы* $g(T)$ *будут*

$$g(\varkappa_1), \; g(\varkappa_2), \ldots, \; g(\varkappa_n).$$

Для доказательства мы исходим из тождества:

$$|\varkappa E - T| = \varphi(\varkappa) = \prod_{\nu=1}^{n} (\varkappa - \varkappa_\nu).$$

Целью нашей является доказать соотношение:

$$|\varkappa E - g(T)| = \prod_{\nu=1}^{n} [\varkappa - g(\varkappa_\nu)].$$

Пусть $h(x)$ — произвольная целая рациональная функция степени r, которая с помощью ее корней $x_1, \ldots, x_r$ представлена в следующем виде:

$$h(x) = a \prod_{\rho=1}^{r} (x - x_\rho).$$

Тогда для произвольной матрицы T имеет место тождество:

$$h(T) = a \prod_{\rho=1}^{r} (T - x_\rho E).$$

Переходя к определителям, имеем:

$$|h(T)| = a^n \prod_{\rho=1}^{r} |T - x_\rho E| = (-1)^{nr} a^n \prod_{\rho=1}^{r} |x_\rho E - T| =$$

$$= (-1)^{nr} a^n \prod_{\rho=1}^{r} \varphi(x_\rho) = (-1)^{nr} a^n \prod_{\rho=1}^{r} \left(\prod_{\nu=1}^{n} (x_\rho - \varkappa_\nu) \right) =$$

$$= (-1)^{nr} (-1)^{nr} a^n \prod_{\nu=1}^{n} \left(\prod_{\rho=1}^{r} (\varkappa_\nu - x_\rho) \right) = \prod_{\nu=1}^{n} h(\varkappa_\nu)].$$

Подставляя теперь вместо $h(T)$ функцию $\varkappa E - g(T)$, непосредственно получаем искомое равенство:

$$|\varkappa E - g(T)| = \prod_{\nu=1}^{n} [\varkappa - g(\varkappa_\nu)].$$

§ 3. Преобразование к главным осям квадратичных и эрмитовых форм.

Особенно важной главой алгебры является теория линейного преобразования квадратичной формы $K(x, x) = \sum_{p, q=1}^{n} k_{pq} x_p x_q$ к виду:

$$K(x, x) = \sum_{p=1}^{n} \varkappa_p y_p^2,$$

т. е. к виду, при котором переменные входят только в виде квадратов. В первую очередь мы займемся задачей о преобразовании к такому виду квадратичной формы $K(x, x)$ с помощью ортогонального преобразования:

$$x_p = \sum_{q=1}^{n} l_{pq} y_q = L_p(y) \qquad (p = 1, \ldots, n).$$

Эта задача, к которой часто приводят вопросы геометрии, механики и теории колебаний, называется *задачей о преобразовании к главным осям*, а искомое линейное преобразование — *преобразованием к главным осям*.

1. Проведение преобразования к главным осям на основании принципа максимума. Теперь мы докажем, что заданную квадратичную форму $K(x, x)$ всегда возможно преобразовать к главным осям. При этом мы будем опираться на тот факт, что непрерывная функция от многих переменных, изменяющихся в некоторой

ограниченной замкнутой области, принимает в этой области наибольшее значение (*теорема Вейерштрасса*)[1]).

Ввиду этого существует такой единичный вектор $\mathfrak{l}_1$ с компонентами $l_{11}, \ldots, l_{1n}$, что $K(x, x)$ принимает наибольшее значение $\varkappa_1$, при $x_1 = l_{11}, \ldots, x_n = l_{1n}$, при добавочном условии:

$$\sum_{p=1}^{n} x_p^2 = 1. \tag{25}$$

Геометрически вектор $\mathfrak{l}_1$ определяет на „единичной сфере" (25) точку, которая в то же время лежит на одной из поверхностей семейства центральных поверхностей второго порядка $K(x, x) = \text{const}$, касающейся единичной сферы.

Далее, существует ортогональный к $\mathfrak{l}_1$ единичный вектор $\mathfrak{l}_2$ с компонентами $l_{21}, \ldots, l_{2n}$, такой, что $K(x, x)$ при $x_1 = l_{21}, \ldots, x_n = l_{2n}$ принимает наибольшее значение $\varkappa_2$, которое только возможно, если к условию (25) присоединить условие:

$$\sum_{p=1}^{n} l_{1p} x_p = 0. \tag{26}$$

И здесь непосредственно ясно, что для фигуры, получающейся от пересечения единичной сферы с „плоскостью", ортогональной к вектору $\mathfrak{l}_1$, можно решить ту же задачу, решение которой для всей единичной сферы дается вектором $\mathfrak{l}_1$.

Далее, существует такой единичный вектор $\mathfrak{l}_3$, ортогональный к векторам $\mathfrak{l}_1$ и $\mathfrak{l}_2$, с компонентами $l_{31}, \ldots, l_{3n}$, что $K(x, x)$ в конечной точке этого вектора принимает наибольшее значение $\varkappa_3$, которое возможно при дополнительных условиях (25), (26) и

$$\sum_{p=1}^{n} l_{2p} x_p = 0. \tag{27}$$

Продолжая поступать таким образом, мы придем к системе n взаимно ортогональных векторов $\mathfrak{l}_1, \ldots, \mathfrak{l}_q, \ldots, \mathfrak{l}_n$, которые мы будем называть *векторами главных осей* или *собственными векторами*; компо-

---

[1]) Преобразование к главным осям можно очень легко провести также непосредственно алгебраически, если искать такую ортогональную матрицу L, чтобы $L'KL = D$ была диагональной матрицей с диагональными элементами $\varkappa_1, \ldots \varkappa_i, \ldots, \varkappa_n$. Из условия $KL = LD$ получаем для коэфициентов преобразования l_{qi} уравнения:

$$\sum_q k_{pq} l_{qi} = l_{pi} \varkappa_i,$$

из которых прежде всего числа $\varkappa_i$ определяются как корни уравнения (30), которое будет позже выведено; далее, на основании простых алгебраических соображений получается ортогональная система n^2 величин l_{qi}. Приведенный в тексте метод доказательства имеет по сравнению с алгебраическим то важное для дальнейшего преимущество, что он применим к обширному классу трансцендентных проблем.

ненты l_{qp} этих векторов, в силу того, что они удовлетворяют соотношениям (21), определяют ортогональное преобразование:

$$x_p = \sum_{q=1}^{n} l_{qp}\, y_q \quad (p = 1, \ldots, n), \tag{28}$$

которое, как мы утверждаем, дает решение нашей задачи.

Решая систему уравнений (28), получаем:

$$y_p = \sum_{q=1}^{n} l_{pq}\, x_q \quad (p = 1, \ldots, n), \tag{29}$$

следовательно, равенство $\mathfrak{x} = \mathfrak{l}_p$ означает, что $y_p = 1$, а $y_q = 0$ при $q \neq p$. В частности, следовательно, максимум $\varkappa_1$ достигается при значениях $y_1 = 1$, $y_2 = 0, \ldots, y_n = 0$, поэтому в преобразованной форме:

$$C(y, y) = \sum_{p,\, q=1}^{n} c_{pq}\, y_p\, y_q,$$

первый коэфициент c_{11} равен $\varkappa_1$. В таком случае форма

$$H(y, y) = \sum_{p,\, q=1}^{n} h_{pq}\, y_p\, y_q = C(y, y) - \varkappa_1 (y_1^2 + \ldots + y_n^2),$$

не может иметь положительных значений. В самом деле, это утверждение несомненно справедливо при условии:

$$\sum_{p=1}^{n} x_p^2 = \sum_{p=1}^{n} y_p^2 = 1,$$

в силу того, что $\varkappa_1$ является при этом наибольшим значением формы $C(y, y)$, следовательно, оно во всяком случае справедливо, если заменить y_i через $\dfrac{y_i}{\sqrt{\sum y_p^2}}$; но умножая на $\sum y_p^2$ получаем, что вообще $H(y, y) \leqslant 0$. Если бы переменное y_1 входило еще в выражение $H(y, y)$, например, если бы коэфициент $h_{12} = h_{21}$ был отличен от нуля, то форма $H(y, y)$ при

$$y_1 = 1, \quad y_2 = \varepsilon, \quad y_3 = \ldots = y_n = 0$$

принимала бы значение:

$$2h_{12}\, \varepsilon + h_{22}\, \varepsilon^2 = \varepsilon\, (2h_{12} + h_{22}\, \varepsilon),$$

которое при соответствующем выборе ε могло бы быть положительным.

Тем самым доказано, что $K(x, x)$ принимает после преобразования вид:

$$C(y, y) = \varkappa_1 y_1^2 + C_1(y, y),$$

где $C_1(y, y)$ означает квадратичную форму от $n-1$ переменных $y_2, \ldots, y_n$. При добавочном условии $y_1=0$ преобразованная форма, следовательно, равна $C_1(y, y)$, и мы можем теперь совершенно таким же образом, как и раньше, заключить, что $C_1(y, y)$ имеет вид $\varkappa_2 y_2^2 + C_2(y, y)$, причем $C_2(y, y)$ зависит только от $n-2$ переменных $y_3, \ldots, y_n$ и т. д.

Тем самым полностью доказана возможность преобразования к главным осям:

$$\sum_{p, q=1}^{n} k_{pq} x_p x_q = \sum_{p=1}^{n} \varkappa_p y_p^2; \qquad \sum_{p=1}^{n} x_p^2 = \sum_{p=1}^{n} y_p^2.$$

Впрочем, для доказательства можно было бы с таким же успехом исходить из соответствующей задачи о минимуме, т. е. можно было бы искать наименьшее значение $K(x, x)$ при условии $E(x, x)=1$, и тогда получились бы числа $\varkappa_1, \ldots, \varkappa_n$ в обратном порядке. Можно было бы также фиксировать значение $K(x, x)$ и разыскать наибольшие или наименьшие значения $E(x, x)$. При этом получились бы значения λ_i, обратные $\varkappa_i$.

2. Характеристические числа и собственные значения. Теперь мы покажем, что числа $\varkappa_i$, определенные в п. 1 как последовательные наибольшие значения, тождественны с характеристическими числами, введенными на стр. 18.

Составим уравнение:

$$\varphi(\varkappa) = (\varkappa - \varkappa_1)(\varkappa - \varkappa_2) \ldots (\varkappa - \varkappa_n) = 0,$$

которое можно записать в виде:

$$\begin{vmatrix} \varkappa - \varkappa_1 & 0 & \ldots & 0 \\ 0 & \varkappa - \varkappa_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \varkappa - \varkappa_n \end{vmatrix} = 0.$$

Но этот определитель является определителем квадратичной формы:

$$\varkappa \sum_{p=1}^{n} y_p^2 - \sum_{p=1}^{h} \varkappa_p y_p^2,$$

получающейся путем ортогонального преобразования из квадратичной формы:

$$\varkappa \sum_{p=1}^{n} x_p^2 - K(x, x).$$

Поэтому имеет место тождественное относительно $\varkappa$ соотношение:

$$\begin{vmatrix} \varkappa - \varkappa_1 & 0 & \ldots & 0 \\ 0 & \varkappa - \varkappa_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \varkappa - \varkappa_n \end{vmatrix} = \begin{vmatrix} \varkappa - k_{11} & -k_{12} & \ldots & -k_{1n} \\ -k_{21} & \varkappa - k_{22} & \ldots & -k_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ -k_{n1} & -k_{n2} & \ldots & \varkappa - k_{nn} \end{vmatrix};$$

следовательно, числа $\varkappa_1, \ldots, \varkappa_n$ являются корнями алгебраического уравнения:

$$\begin{vmatrix} k_{11}-\varkappa & k_{12} & \ldots & k_{1n} \\ k_{21} & k_{22}-\varkappa & \ldots & k_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ k_{n1} & k_{n2} & \ldots & k_{nn}-\varkappa \end{vmatrix} = 0, \tag{30}$$

относительно неизвестной $\varkappa$, т. е. характеристическими числами.

Наше доказательство вместе с тем обнаруживает, что уравнение (30) всегда имеет только действительные корни, если k_{pq} — произвольные действительные числа, удовлетворяющие условию $k_{qp}=k_{pq}$ [1]. Заметим, еще, что абсолютные величины собственных значений, т. е. чисел, обратных характеристическим числам, геометрически означают квадраты длин главных полуосей центральной поверхности второго порядка $K(x, x)=1$. Если по крайней мере одно из характеристических чисел равно нулю, то форма называется „выродившейся"; она может быть представлена как форма от меньшего чем n числа переменных. Вследствие инвариантности определителя (см. стр. 13) это имеет место в том и только в том случае, когда $\mathrm{K}=|k_{pq}|=\varkappa_1\varkappa_2\ldots\varkappa_n$ обращается в нуль. Для того чтобы $K(x, x)$ было определенной положительной формой, необходимо и достаточно, чтобы $\varkappa_p>0$, $p=1, \ldots, n$.

Если форма $K(x, x)$ преобразована к главным осям:

$$K(x, x)=\sum_{p=1}^{n}\varkappa_p y_p^2,$$

то для квадратичной формы, стоящей в правой части, можно непосредственно образовать итерированные формы, а принимая во внимание сказанное ранее относительно ортогонального преобразования, получаем:

$$K^2(x, x)=\sum_{p=1}^{n}\varkappa_p^2 y_p^2, \quad K^3(x, x)=\sum_{p=1}^{n}\varkappa_p^3 y_p^2, \ldots$$

Отсюда мы видим, что h-я итерированная форма $K^h(x, x)$ имеет в качестве характеристических чисел h-е степени характеристических чисел формы $K(x, x)$, а собственные значения $K^h(x, x)$ являются h-ми степенями собственных значений $K(x, x)$, что, впрочем, непосредственно следует из теоремы на стр. 19. Далее ясно, что при четном h форма $K^h(x, x)$ является определенной положительной формой.

3. Обобщение на эрмитовы формы. Совершенно аналогичным образом можно провести преобразование к главным осям и

[1] Уравнение (30) обыкновенно называют „уравнением вековых возмущений" (Säkulargleichung), так как оно встречается в задаче вековых планетных возмущений. Для прямого доказательства теоремы о действительности корней см., например, соответствующее рассуждение в гл. III, § 4, п. 2.

эрмитовых форм. Эрмитову форму:

$$H(x, \bar{x}) = \sum_{p,q=1}^{n} h_{pq} x_p \bar{x}_q$$

с матрицей $H = \bar{H}'$ можно всегда преобразовать с помощью унитарного преобразования $\mathfrak{L}$:

$$x_p = \sum_{q=1}^{n} l_{qp} y_q \quad (p = 1, \ldots, n)$$

к виду:

$$H(x, \bar{x}) = \sum_{p=1}^{n} \varkappa_p y_p \bar{y}_p = \sum_{p=1}^{n} \varkappa_p |y_p|^2,$$

где все коэфициенты $\varkappa_p$ — действительные числа. Характеристические числа $\varkappa_m$ опять являются наибольшими значениями эрмитовой формы $H(x, \bar{x})$ при условии

$$\sum |x_p|^2 = 1$$

и

$$\sum_{p=1}^{n} l_{ip} \bar{x}_p = 0 \quad (i = 1, \ldots, m-1).$$

4. Закон инерции квадратичных форм. Если отказаться от требования ортогональности искомого линейного преобразования, то квадратичную форму можно различным образом привести к виду, при котором входят только квадраты переменных. Например, если выполнить данное ранее ортогональное преобразование и затем сделать любое преобразование подобия, т. е. преобразование, при котором каждая переменная умножается только на некоторый множитель пропорциональности, то получим опять выражение такого же вида. В частности, следовательно, можно достигнуть того, чтобы все (действительные) коэфициенты формы имели значения $+1$ или -1. При этом имеет место следующая теорема, называемая законом инерции квадратичных форм.

Число положительных и отрицательных коэфициентов, которые получаются при действительном и обратимом преобразовании квадратичной формы в выражение, состоящее только из квадратов, не зависит от частного характера этого преобразования.

Доказательство. Положительные и отрицательные коэфициенты могут быть, как было указано, сделаны равными соответственно $+1$ и -1. Если бы квадратичная форма $K(x, x)$ двумя различными действительными преобразованиями приводилась соответственно к виду

$$y_1^2 + \ldots + y_r^2 - y_{r+1}^2 - \ldots - y_n^2$$

и

$$z_1^2 + \ldots + z_s^2 - z_{s+1}^2 - \ldots - z_n^2,$$

где $r < s$, то в силу соотношения

$$y_1^2 + \ldots + y_r^2 - y_{r+1}^2 - \ldots - y_n^2 = z_1^2 + \ldots + z_s^2 - z_{s+1}^2 - \ldots - z_n^2$$

или

$$y_1^2+\ldots+y_r^2+z_{s+1}^2+\ldots+z_n^2=y_{r+1}^2+\ldots+y_n^2+z_1^2+\ldots+z_s^2$$

из системы уравнений $y_1=\ldots=y_r=z_{s+1}=\ldots=z_n=0$ вытекало бы обращение в нуль и остальных y_j. Так как система однородных уравнений содержит менее n уравнений, то непременно существует отличный от нуля вектор $\mathfrak{x}$, удовлетворяющий этой системе; но этот вектор не может удовлетворять системе n однородных уравнений $\mathfrak{y}=0$ с определителем, не равным нулю.

5. Выражение для резольвенты формы. Резольвента квадратичной формы $K(x, x)$ на основании изложенного в этом параграфе также легко может быть приведена к удобному виду, если согласно § 2 определить ее с помощью символического равенства:

$$\mathrm{K}(x, x; \lambda)=\frac{[E(x, x)-\lambda K(x, x)]^{-1}-E(x, x)}{\lambda}.$$

Представим себе, что форма $K(x, x)$ путем преобразования к главным осям приведена к виду:

$$K(x, x)=\sum_{p=1}^{n}\frac{y_p^2}{\lambda_p},$$

тогда резольвента формы

$$\sum_{p=1}^{n}\frac{y_p^2}{\lambda_p}$$

должна совпасть с резольвентой формы $K(x, x)$, так как $[E(x, x)-\lambda K(x, x)]^{-1}$ переходит при преобразовании в

$$\left[E(y, y)-\lambda\sum_{p=1}^{n}\frac{y_p^2}{\lambda_p}\right]^{-1}.$$

Далее, имеем:

$$\frac{1}{\lambda}\left[\left(\sum_{p=1}^{n}y_p^2-\lambda\sum_{p=1}^{n}\frac{y_p^2}{\lambda_p}\right)^{-1}-E(y, y)\right]=$$

$$=\frac{1}{\lambda}\left[\left(\sum_{p=1}^{n}\frac{\lambda_p-\lambda}{\lambda_p}y_p^2\right)^{-1}-E(y, y)\right]=$$

$$=\frac{1}{\lambda}\left[\sum\frac{\lambda_p}{\lambda_p-\lambda}y_p^2-E(y, y)\right]=\frac{1}{\lambda}\left[\sum\frac{\lambda_p}{\lambda_p-\lambda}y_p^2-\sum y_p^2\right]=\sum_{p=1}^{n}\frac{y_p^2}{\lambda_p-\lambda}.$$

Если преобразовать здесь снова к переменным x_p, то получим, пользуясь обозначениями (19), резольвенту $\mathrm{K}(x, x; \lambda)$ формы $K(x, x)$ в виде:

$$\mathrm{K}(x, x; \lambda)=\sum_{p=1}^{n}\frac{[L'_p(x)]^2}{\lambda_p-\lambda}, \tag{31}$$

или, возвращаясь к билинейной форме:

$$\mathrm{K}(u,\ x;\ \lambda)=\sum_{p=1}^{n}\frac{L'_p(u)\,L'_p(x)}{\lambda_p-\lambda}. \tag{32}$$

Между прочим, это выражение обнаруживает, что вычет рациональной относительно λ функции $\mathrm{K}(u,\ x;\ \lambda)$ в точке $\lambda=\lambda_p$ равен

$$-L'_p(u)\,L'_p(x),$$

если предполагать, что $\lambda_p\neq\lambda_q$ при $p\neq q$.

6. Решение системы линейных уравнений, соответствующей данной форме. В заключение выразим решение системы линейных уравнений:

$$x_p-\lambda\sum_{q=1}^{n}k_{pq}x_q=y_p\quad(p=1,\ \ldots,\ n), \tag{33}$$

соответствующей квадратичной форме:

$$K(x,\ x)=\sum_{p,\ q=1}^{n}k_{pq}x_px_q,$$

с помощью собственных векторов.

Применим к переменным x_l и y_l преобразование к главным осям:

$$x_p=\sum_{q=1}^{n}l_{qp}u_q,$$

$$y_p=\sum_{q=1}^{n}l_{qp}v_q,$$

причем форма $K(x,\ x)$ примет вид:

$$\sum_{q=1}^{n}\varkappa_q\,u_q^2;$$

тогда наша система уравнений (33) переходит в систему:

$$u_p-\lambda\varkappa_p u_p=v_p\quad(p=1,\ \ldots,\ n), \tag{34}$$

решением которой является:

$$u_p=\frac{v_p}{1-\lambda\varkappa_p}=\frac{v_p}{1-\dfrac{\lambda}{\lambda_p}}=\frac{\lambda_p}{\lambda_p-\lambda}\,v_p. \tag{35}$$

В первоначальных переменных решение выражается эквивалентной формулой:

$$\mathfrak{x}=\sum_{p=1}^{n}\frac{(\mathfrak{y}\mathfrak{l}_p)}{1-\dfrac{\lambda}{\lambda_p}}\,\mathfrak{l}_p, \tag{36}$$

в которой оно развернуто по собственным векторам $\mathfrak{l}_1, \ldots, \mathfrak{l}_n$ формы $K(x, x)$. При этом полагаем

$$(\mathfrak{y}\mathfrak{l}_p) = \sum_{q=1}^{n} l_{pq} y_q.$$

Собственный вектор $\mathfrak{l}_p$ является нормированным решением системы однородных уравнений:

$$x_q - \lambda_p \sum_{r=1}^{n} k_{qr} x_r = 0$$

или

$$u_q - \lambda_p \varkappa_q u_q = 0 \quad (q = 1, \ldots, n).$$

Если при $q \neq p$ все $\varkappa_q$ не равны $\varkappa_p = \frac{1}{\lambda_p}$, то имеется только одно нормированное решение:

$$u_p = 1, \quad u_q = 0 \quad (q \neq p)$$

или

$$\mathfrak{x} = \mathfrak{l}_p.$$

Если среди характеристических чисел имеются равные между собой, то собственные векторы определяются неоднозначно.

§ 4. Минимально-максимальное свойство собственных значений.

1. **Определение характеристических чисел с помощью задачи о наименьшем значении максимума.** Выше мы ввели характеристические числа с помощью ряда задач на нахождение максимума, причем в каждой предполагалось решение предыдущей. Теперь же, предполагая, что характеристические числа расположены в порядке убывания, покажем, как можно непосредственно охарактеризовать h-е характеристическое число как решение несколько иной задачи, в которой не приходится ссылаться на решение предыдущих задач.

Требуется найти наибольшее значение формы:

$$K(x, x) = \sum_{p, q=1}^{n} k_{pq} x_p x_q,$$

если кроме условия

$$\sum_{p=1}^{n} x_p^2 = 1 \tag{25}$$

должны выполняться еще $h-1$ уравнений:

$$\sum_{p=1}^{n} \alpha_{\nu p} x_p = 0 \quad (\nu = 1, \ldots, h-1;\ h \leqslant n). \tag{37}$$

Далее, требуется, чтобы этот максимум, представляющий функцию параметров $a_{\nu p}$, путем соответствующего выбора этих параметров принял наименьшее возможное значение.

Преобразование к главным осям приводит форму $K(x, x)$ к виду:

$$\sum_{p=1}^{n} \varkappa_p y_p^2 \quad (\varkappa_1 \geqslant \ldots \geqslant \varkappa_n),$$

условие (25) — к виду:

$$\sum_{p=1}^{n} y_p^2 = 1, \tag{38}$$

а уравнения (37) — к виду:

$$\sum_{p=1}^{n} \beta_{\nu p} y_p = 0, \tag{39}$$

где $\beta_{\nu p}$ — новые параметры. Выбрав

$$y_{h+1} = \ldots = y_n = 0,$$

получаем, каковы бы ни были $\beta_{\nu p}$, $h-1$ однородных уравнений (39) для h неизвестных $y_1, \ldots, y_h$; всегда можно подобрать систему значений, удовлетворяющую этим уравнениям и уравнению (38). Для этих значений имеем:

$$K(x, x) = \varkappa_1 y_1^2 + \ldots + \varkappa_h y_h^2 \geqslant \varkappa_h (y_1^2 + \ldots + y_n^2) = \varkappa_h.$$

Следовательно, искомый максимум при любой системе $\beta_{\nu p}$ не может быть меньше $\varkappa_h$. Но этот максимум как раз становится равным $\varkappa_h$, если за систему (37) взять уравнения:

$$y_1 = \ldots = y_{h-1} = 0.$$

Таким образом получаем:

h-е характеристическое число $\varkappa_h$ квадратичной формы $K(x, x)$ есть наименьшее значение, которое может иметь максимум $K(x, x)$, когда кроме условия

$$\sum_{p=1}^{n} x_p^2 = 1$$

заданы еще $h-1$ произвольных линейных однородных уравнений между числами x_p.

2. Применения. На основании этого свойства характеристических чисел особенно легко выяснить, как они изменяются, когда на переменные налагаются j независимых „связей“:

$$\sum_{p=1}^{n} \gamma_{sp} x_p = 0 \quad (s = 1, \ldots, j), \tag{40}$$

так что $K(x, x)$ сводится к квадратичной форме $\tilde{K}(x, x)$ от $n-j$

независимых переменных. h-е характеристическое число $\tilde{\varkappa}_h$ представляет решение той же задачи о минимуме максимума, что и $\varkappa_h$, но условием (40) совокупность допустимых систем значений $x_1, \ldots, x_n$ суживается.

Поэтому отдельные максимумы, а вместе с тем характеристическое число для $\tilde{K}(x, x)$ не может превосходить соответствующего числа для формы $K(x, x)$.

Далее, $\varkappa_{j+h}$ является наименьшим максимумом, который может принять форма $K(x, x)$, когда, кроме условия (25), заданы еще $h+j-1$ линейных однородных уравнений для x_p, и потому $\varkappa_{j+h}$ не может быть больше, чем $\tilde{\varkappa}_h$, для которого j из этих уравнений даны системой (40).

Следовательно, $\tilde{\varkappa}_h \geqslant \varkappa_h \geqslant \varkappa_{j+h}$, или словами: *если квадратичная форма $K(x, x)$ от n переменных при j независимых линейных однородных связях переходит в квадратичную форму $\tilde{K}(x, x)$ от $n-j$ переменных, то характеристические числа $\tilde{\varkappa}_1, \ldots, \tilde{\varkappa}_{n-j}$ не больше соответствующих чисел ряда $\varkappa_1, \ldots, \varkappa_{n-j}$ и не меньше соответствующих чисел ряда $\varkappa_{j+1}, \ldots, \varkappa_n$* [1].

Если в частности взять $j=1$ и выбрать в качестве линейной связи условие $x_n=0$, то форма $K(x, x)$ переходит в $(n-1)$-ю усеченную форму, и мы получаем теорему:

h-е характеристическое число $(n-1)$-й усеченной формы не больше h-го характеристического числа и не меньше $(h+1)$-го характеристического числа первоначальной формы.

Применяя нашу теорему к $(n-1)$-й усеченной форме, мы получаем соответствующую теорему для характеристических чисел $(n-2)$-й усеченной формы и т. д. Вообще мы видим, что характеристические числа двух последовательных усеченных форм данной квадратичной формы распределяются по величине указанным образом.

Подобным же образом имеем: *если к форме $K(x, x)$ прибавить форму, не принимающую никогда отрицательных значений, то характеристические числа суммы не меньше соответствующих чисел формы $K(x, x)$.*

Вместо того чтобы для определения характеристических чисел пользоваться задачей о наименьшем значении максимума, можно было бы с таким же успехом взять за исходный пункт задачу о наибольшем значении минимума. Тогда опять получаются числа $\varkappa_h$, только в обратном порядке.

Предоставляем читателю формулировать и доказать свойство наименьшего значения максимума для характеристических чисел эрмитовых форм.

---

[1] Для пояснения изложенных рассуждений сделаем следующее замечание. Если пересечь трехосный эллипсоид плоскостью, проходящей через его центр, то большая ось эллипса, полученного в сечении, заключается между большой и средней осью эллипсоида, а малая ось эллипса заключается между средней и малой осью эллипсоида.

§ 5. Дополнения и задачи к первой главе.

1. **Линейная независимость и определитель Грама.** Вопрос о линейной зависимости m данных векторов $\mathfrak{v}_1, \ldots, \mathfrak{v}_m$ можно формально очень просто решить, не прибегая, как это обычно делают, к установлению ранга матрицы компонент, следующим образом. Рассмотрим квадратичную форму:

$$G(x, x) = (x_1\mathfrak{v}_1 + \ldots + x_m\mathfrak{v}_m)^2 = \sum_{i,k=1}^{m} (\mathfrak{v}_i\mathfrak{v}_k)\, x_i x_k.$$

Несомненно, $G(x, x) \geqslant 0$, а векторы $\mathfrak{v}_i$ линейно зависимы в том и только в том случае, когда существует система значений переменных $x_1, \ldots, x_m$, удовлетворяющая условию:

$$\sum_{i=1}^{m} x_i^2 = 1 \tag{25'}$$

и при которой $G(x, x) = 0$. Следовательно, для того чтобы имела место линейная зависимость, минимум формы $G(x, x)$ при условии (25') непременно должен равняться нулю. Но этот минимум представляет наименьшее характеристическое число квадратичной формы $K(x, x)$, т. е. наименьший корень уравнения:

$$\begin{vmatrix} \mathfrak{v}_1^2 - \varkappa & (\mathfrak{v}_1\mathfrak{v}_2) & \ldots & (\mathfrak{v}_1\mathfrak{v}_m) \\ (\mathfrak{v}_2\mathfrak{v}_1) & \mathfrak{v}_2^2 - \varkappa & \ldots & (\mathfrak{v}_2\mathfrak{v}_m) \\ \cdots & \cdots & \cdots & \cdots \\ (\mathfrak{v}_m\mathfrak{v}_1) & (\mathfrak{v}_m\mathfrak{v}_2) & \ldots & \mathfrak{v}_m^2 - \varkappa \end{vmatrix} = 0. \tag{41}$$

Итак: *необходимым и достаточным условием линейной зависимости векторов $\mathfrak{v}_1, \ldots \mathfrak{v}_m$ является обращение в нуль „определителя Грама“*

$$\Gamma = \begin{vmatrix} \mathfrak{v}_1^2 & (\mathfrak{v}_1\mathfrak{v}_2) & \ldots & (\mathfrak{v}_1\mathfrak{v}_m) \\ (\mathfrak{v}_2\mathfrak{v}_1) & \mathfrak{v}_2^2 & \ldots & (\mathfrak{v}_2\mathfrak{v}_m) \\ \cdots & \cdots & \cdots & \cdots \\ (\mathfrak{v}_m\mathfrak{v}_1) & (\mathfrak{v}_m\mathfrak{v}_2) & \ldots & \mathfrak{v}_m^2 \end{vmatrix}. \tag{42}$$

Если расположить левую часть уравнения (41), которому удовлетворяют все (неотрицательные) характеристические числа $\varkappa_1, \ldots, \varkappa_m$ формы $K(x, x)$, по степеням $\varkappa$, то свободный от $\varkappa$ член равен Γ, а коэфициент при $\varkappa^m$ равен $(-1)^m$. По известной теореме алгебры имеет место, следовательно, соотношение:

$$\Gamma = \varkappa_1 \ldots \varkappa_m. \tag{43}$$

Итак, *определитель Грама любой системы m векторов не может иметь отрицательного значения*. Соотношение

$$\Gamma = |(\mathfrak{v}_i\mathfrak{v}_k)| \geqslant 0, \tag{44}$$

в котором знак равенства может иметь место лишь для линейно зависимых векторов $\mathfrak{v}_1, \ldots \mathfrak{v}_m$, представляет обобщение неравенства Шварца (стр. 2):

$$\mathfrak{v}_1^2\mathfrak{v}_2^2 - (\mathfrak{v}_1\mathfrak{v}_2)^2 = \begin{vmatrix} \mathfrak{v}_1^2 & (\mathfrak{v}_1\mathfrak{v}_2) \\ (\mathfrak{v}_2\mathfrak{v}_1) & \mathfrak{v}_2^2 \end{vmatrix} \geqslant 0.$$

Значение определителя Грама или также наименьшее характеристическое число $\varkappa_m$ формы $G(x, x)$ представляет меру линейной независимости векторов $\mathfrak{v}_1, \ldots, \mathfrak{v}_m$. Чем меньше это число, тем более „плоской" является m-мерная фигура, составленная этими векторами; если эта мера независимости равна нулю, то число измерений фигуры не более $m-1$. Впрочем, определителю Грама можно легко приписать геометрическое значение. Этот определитель равен квадрату $m!$-кратного объема m-мерной геометрической фигуры, образуемой векторами $\mathfrak{v}_1, \ldots, \mathfrak{v}_m$, следовательно, например, при $m=2$ равен квадрату удвоенной площади треугольника, построенного на векторах $\mathfrak{v}_1$ и $\mathfrak{v}_2$.

Разумеется, критерий Грама для линейной зависимости должен быть эквивалентен обычному критерию, гласящему, что векторы линейно зависимы в том и только в том случае, когда все определители m-го порядка, которые можно выделить из прямоугольной таблицы компонент

$$\begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ v_{m1} & v_{m2} & \ldots & v_{mn} \end{pmatrix},$$

равны нулю. В самом деле, по известной теореме теории определителей:

$$\Gamma = \sum \begin{vmatrix} v_{1s_1} & v_{1s_2} & \ldots & v_{1s_m} \\ v_{2s_1} & v_{2s_2} & \ldots & v_{2s_m} \\ \cdot & \cdot & \cdot & \cdot \\ v_{ms_1} & v_{ms_2} & \ldots & v_{ms_m} \end{vmatrix}^2, \tag{45}$$

где суммирование распространяется на все целые значения $s_1, s_2, \ldots, s_m$ от 1 до n, причем $s_1 < s_2 < \ldots < s_m$.

2. Теорема Адамара (Hadamard) об оценке определителя. Для любого определителя:

$$A = |a_{ik}| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

с действительными элементами a_{ik} справедлива оценка

$$A^2 \leqslant \prod_{i=1}^{n}\sum_{k=1}^{n} a_{ik}^2. \tag{46}$$

Доказательство. Будем изменять элементы a_{ik}, однако, так, чтобы суммы квадратов

$$\sum_{k=1}^{n} a_{ik}^2 = c_i^2 \quad (i = 1, \ldots, n)$$

оставались неизменными. Если $A^2_{\max}$ представляет наибольшее значение, которое может иметь функция A^2 элементов a_{ik} при этих n условиях (из теоремы Вейерштрасса, приведенной на стр. 21, непосредственно следует, что такое наибольшее значение непременно должно быть получено для некоторого определителя $A_{\max}$), то элементы $A_{\max}$ в каждой строке должны быть пропорциональны соответствующим минорам. В самом деле, при любом выборе h имеем:

$$A = a_{h1}A_{h1} + \ldots + a_{hn}A_{hn},$$

следовательно, на основании неравенства Шварца (стр. 2) находим:

$$A^2 \leqslant \sum_{k=1}^{n} a_{hk}^2 \sum_{k=1}^{n} A_{hk}^2;$$

если при этом числа a_{hk} не пропорциональны числам A_{hk}, то значение A^2, безусловно, не будет наибольшим, так как в этом случае имеет место знак неравенства, между тем путем изменения n величин a_{hk} $(k = 1, \ldots, n)$ мы можем, не меняя чисел A_{hk} и числа c_h, сделать квадрат определителя равным правой части.

Умножая определитель $A_{\max}$ на себя самого по правилу умножения определителей, мы получаем:

$$A^2_{\max} = \prod_{i=1}^{n} c_i^2,$$

так как скалярные произведения двух различных строк равны нулю в силу только что доказанной пропорциональности, как это следует из элементарных теорем теории определителей. Поэтому для первоначального определителя A справедливо соотношение:

$$A^2 \leqslant \prod_{i=1}^{n} c_i^2 = \prod_{i=1}^{n} \sum_{k=1}^{n} a_{ik}^2.$$

Геометрически теорема Адамара означает, что объем параллелепипеда, построенного в n-мерном пространстве на n векторах данной длины, будет наибольшим, когда векторы взаимно перпендикулярны.

Оценка Адамара справедлива и при комплексных значениях a_{ik}, если вместо A и a_{ik} подставить их абсолютные значения.

3. Одновременное преобразование двух квадратичных форм к каноническому виду. Преобразование к главным осям является частным случаем более общей задачи (которая сводится к предыдущей, но может быть столь же просто изучена и непосред-

ственно) об одновременном преобразовании двух заданных квадратичных форм:

$$K(x, x) = \sum_{p, q=1}^{n} k_{pq} x_p x_q;$$

$$H(x, x) = \sum_{p, q=1}^{n} h_{pq} x_p x_q$$

одна из которых, скажем $H(x, x)$, определенная положительная, с помощью линейного преобразования

$$x_p = \sum_{q=1}^{n} l_{pq} y_q$$

в выражения, содержащие только квадраты переменных:

$$K(x, x) = \sum_{p=1}^{n} \varkappa_p y_p^2,$$

$$H(x, x) = \sum_{p=1}^{n} \eta_p y_p^2.$$

При этом коэфициенты η_p положительной формы H могут быть сделаны равными $+1$. Отношения $\frac{\varkappa_p}{\eta_p} = \rho_p$ мы назовем характеристическими числами, числа $\frac{1}{\rho_p} = \lambda_p$ — собственными значениями формы $K(x, x)$ по отношению к $H(x, x)$.

Требуется непосредственно провести теорию преобразования и доказать следующее свойство характеристических чисел.

При $\rho_1 \geqslant \ldots \geqslant \rho_n$ число ρ_h является наименьшим значением, которое может иметь максимум функции $\frac{K(x, x)}{H(x, x)}$ при добавочных условиях:

$$\sum_{p=1}^{n} \alpha_{\nu p} x_p = 0 \qquad (\nu = 1, \ldots, h-1),$$

если рассматривать этот максимум как функцию параметров $\alpha_{\nu p}$.

Для нахождения искомого преобразования можно искать сперва максимум $\frac{K(x, x)}{H(x, x)}$ при условии $H(x, x) = 1$; пусть этот максимум достигается при $x_p = l_{1p}$ $(p = 1, \ldots, n)$. Затем присоединяем дальнейшее добавочное условие:

$$\sum_{p, q=1}^{n} h_{pq} l_{1p} x_q = 0$$

и т. д. Характеристические числа ρ_p являются корнями уравнения:

$$\begin{vmatrix} k_{11}-\rho h_{11} & k_{12}-\rho h_{12} & \dots & k_{1n}-\rho h_{1n} \\ k_{21}-\rho h_{21} & k_{22}-\rho h_{22} & \dots & k_{2n}-\rho h_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ k_{n1}-\rho h_{n1} & k_{n2}-\rho h_{n2} & \dots & k_{nn}-\rho h_{nn} \end{vmatrix} = 0.$$

Это уравнение можно получить также как условие разрешимости однородной системы уравнений:

$$\sum_{j=1}^{n} (k_{ij}-\rho h_{ij})\, x_j = 0.$$

Системы значений x_j, соответствующие отдельным характеристическим числам, дают после соответствующего нормирования компоненты „собственных векторов“, т. е. коэфициенты искомого преобразования.

4. Билинейные и квадратичные формы от бесконечно большого числа переменных. Наши теории остаются справедливыми и в том случае, когда число переменных неограниченно возрастает, если только при этом сделать некоторые предположения, во-первых, относительно коэфициентов билинейных или квадратичных форм, например, допустить, что сумма их квадратов сходится, а во-вторых, предполагать также сходимость суммы квадратов входящих переменных. Эта теория форм от бесконечно большого числа переменных, разработанная Гильбертом, может быть непосредственно применена к многочисленным проблемам анализа. Однако в этих проблемах мы можем быстрее притти к цели, если пойдем более прямым путем, соответствующим алгебраическому векторному и тензорному исчислению.

5. Бесконечно малые линейные преобразования. Бесконечно малым линейным преобразованием называют преобразование с матрицей:

$$A = E + (\varepsilon\alpha_{ik}) = \begin{pmatrix} 1+\varepsilon\alpha_{11} & \varepsilon\alpha_{12} & \dots & \varepsilon\alpha_{1n} \\ \varepsilon\alpha_{21} & 1+\varepsilon\alpha_{22} & \dots & \varepsilon\alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \varepsilon\alpha_{n1} & \varepsilon\alpha_{n2} & \dots & 1+\varepsilon\alpha_{nn} \end{pmatrix},$$

причем ε есть бесконечно малая величина первого порядка, т. е. такая величина, высшими степенями которой можно в рассматриваемом случае пренебречь по сравнению с низшими. Произведение двух бесконечно малых преобразований с матрицами $A = E + (\varepsilon\alpha_{ik})$ и $B = E + (\varepsilon\beta_{ik})$ имеет матрицу $C = AB = E + (\varepsilon\alpha_{ik} + \varepsilon\beta_{ik})$. Отсюда следует:

Бесконечно малые преобразования обладают свойством переместительности.

Далее:

Матрицей, обратной матрице $A = E + (\varepsilon\alpha_{ik})$, является $A^{-1} = E - (\varepsilon\alpha_{ik})$; определитель матрицы $A = E + (\varepsilon\alpha_{ik})$ равен $1 + \varepsilon(\alpha_{11} + \alpha_{22} + \dots + \alpha_{nn})$.

Если бесконечно малое преобразование ортогонально, то из условия $AA' = E$, где A' — транспонированная матрица, следует соотношение

$\alpha_{ik}+\alpha_{ki}=0$, т. е. *необходимым и достаточным условием ортогональности бескочечно малого преобразования является требование, чтобы матрица его, если отнять от нее матрицу E, была косо симметрической.*

Бесконечно малое преобразование общего вида $C=E+(\varepsilon\gamma_{ik})$ можно с помощью величин:

$$\alpha_{ik}=\frac{1}{2}(\gamma_{ik}-\gamma_{ki}),$$

$$\beta_{ik}=\frac{1}{2}(\gamma_{ik}+\gamma_{ki})$$

представить в виде произведения ортогонального преобразования с матрицей $A=E+(\varepsilon\alpha_{ik})$ и симметрического преобразования с матрицей $B=E+(\varepsilon\beta_{ik})$.

Симметрическое, хотя бы и не бесконечно малое преобразование $x_i=\sum_k s_{ik}x_k$ с матрицей $S=(s_{ik})$ геометрически означает растяжение в n взаимно перпендикулярных направлениях. В самом деле, введем новые переменные u_i и v_i при помощи преобразования к главным осям квадратичной формы $S(x, x)$, так что

$$\sum_{i,k=1}^{n} s_{ik}x_ix_k=\sum_{i=1}^{n}\varkappa_i u_i^2,$$

тогда уравнения примут вид:

$$v_i=\varkappa_i u_i,$$

что аналитически выражает „неравномерное" растяжение, отнесенное к главным осям. Отношение приращения объема к первоначальному объему, „объемное расширение", выражается, очевидно, разностью $\varkappa_1\ldots\varkappa_n-1$, вместо которой мы можем также писать $|s_{ik}|-1$. Если в частности преобразование бесконечно малое, т. е. $(s_{ik})=E+(\varepsilon\beta_{ik})$, то

$$\varkappa_1\ldots\varkappa_n-1=\varepsilon(\beta_{11}+\ldots+\beta_{nn}).$$

Так как ортогональное преобразование означает вращение, то можно резюмируя сказать:

Бесконечно малое преобразование с матрицей $E+(\varepsilon\gamma_{ik})$ можно представить в виде произведения вращения и растяжения; объемное расширение равно $\varepsilon\sum_{i=1}^{n}\gamma_{ii}$.

6. Варьированные системы. Для теории малых колебаний имеет важное значение вопрос: как изменяются собственные значения и собственные векторы квадратичной формы

$$K(x, x)=\sum_{i,k=1}^{n} b_{ik}x_ix_k,$$

если варьировать и единичную форму $E(x, x)$ и форму $K(x, x)$, т. е.

если одновременно преобразовать к каноническому виду формы $E(x, x)+$ $+\varepsilon A(x, x)$ и $E(x, x)+\varepsilon B(x, x)$ (см. п. 3), где

$$A(x, x)=\sum_{i,k=1}^{n} \alpha_{ik} x_i x_k, \quad B(x, x)=\sum_{i,k=1}^{n} \beta_{ik} x_i x_k,$$

а ε — параметр.

Полагаем

$$K(x, x)+\varepsilon B(x, x)=\sum_{i,k=1}^{n} b'_{ik} x_i x_k, \qquad E(x, x)+\varepsilon A(x, x)=\sum_{i,k=1}^{n} a'_{ik} x_i x_k,$$

тогда уравнения, служащие для определения компонент собственных векторов (см. стр. 35), имеют вид:

$$\sum_{k=1}^{n} (b'_{ik}-\rho' a'_{ik}) x'_k=0 \quad (i=1, \ldots, n),$$

причем ρ' определяется из уравнения, получающегося приравниванием нулю соответствующего определителя и имеющего n действительных корней.

Обозначим характеристические числа $K(x, x)$ через $\rho_1, \rho_2, \ldots \rho_n$, причем предполагаем, что все они между собой различны; соответствующие значения для вариированной системы обозначим через $\rho'_1, \rho'_2, \ldots \rho'_n$. Мы можем тогда принять, что исходная форма $K(x, x)$ задана в виде суммы квадратов:

$$K(x, x)=\rho_1 x_1^2+\rho_2 x_2^2+\ldots+\rho_n x_n^2.$$

Величины ρ'_i, являясь простыми корнями алгебраического уравнения, представляют в окрестности точки $\varepsilon=0$ однозначные аналитические функции от ε; это же справедливо и для компонент x'_{hk} вариированных собственных векторов, соответствующих характеристическим числам ρ'_h. Следовательно, величины ρ'_i и x'_{hk} можно искать в виде степенных рядов, расположенных по степеням ε, свободные члены которых, конечно, соответственно равны ρ_i и x_{hk}. Для того чтобы последовательно определить коэфициенты при $\varepsilon, \varepsilon^2, \ldots$, подставляем эти ряды в уравнения

$$\sum_{k=1}^{n} (b'_{ik}-\rho'_h a'_{ik}) x'_{hk}=0 \quad (i, h=1, 2, \ldots, n),$$

в которых, кроме того, надо положить $b'_{ik}=\delta_{ik}+\varepsilon\beta_{ik}$, $a'_{ik}=e_{ik}+\varepsilon\alpha_{ik}$, где $\delta_{ii}=\rho_i$, $\delta_{ik}=0$ $(i\neq k)$, $e_{ii}=1$, $e_{ik}=0$ $(i\neq k)$. Располагая по степеням ε и приравнивая нулю коэфициенты при этих степенях, получаем бесконечную последовательность уравнений. Вполне эквивалентным, но несколько более удобным является следующее рассуждение, в котором ε рассматривают как величину бесконечно малую. Именно, рассмотрим из предыдущих уравнений для компонент h-го собственного вектора

h-е уравнение $(i=h)$, в нем все слагаемые с точностью до бесконечно-малых второго порядка относительно ε равны нулю:

$$\rho'_h = \frac{\rho_h + \varepsilon\beta_{hh}}{1+\varepsilon\alpha_{hh}} = \rho_h - \varepsilon\rho_h\alpha_{hh} + \varepsilon\beta_{hh}.$$

Рассмотрение уравнений при $i \neq h$, в которых все члены, кроме двух, представляют бесконечно-малые второго порядка, дает с точностью до бесконечно-малых второго порядка:

$$x'_{hh} = 1; \quad x'_{hi} = -\varepsilon\frac{\alpha_{ih}\rho_h - \beta_{ih}}{\rho_h - \rho_i},$$

так как сумма $\sum x'^2_{hk} = 1$.

Пользуясь этими значениями компонент собственных векторов, мы можем очень легко вычислить собственные значения с точностью до бесконечно-малых третьего порядка. Рассмотрим опять уравнение для компонент h-го собственного вектора, получающееся при $i=h$:

$$\sum_{k=1}^{n} \left(b'_{hk} - \rho'_h a'_{hk}\right) x'_{hk} = 0.$$

Отбрасывая в левой части величины третьего порядка относительно ε и уединяя член, для которого $k=h$, получаем:

$$b'_{hh} - \rho'_h a'_{hh} = \sum_k{}' \varepsilon\left(b'_{hk} - \rho'_h a'_{hk}\right)\frac{\alpha_{kh}\rho_h - \beta_{kh}}{\rho_h - \rho_k} = -\varepsilon^2 \sum{}' \frac{(\alpha_{kh}\rho_h - \beta_{kh})^2}{\rho_h - \rho_k};$$

отсюда следует, что

$$\rho'_h = \rho_h - \varepsilon\rho_h\alpha_{hh} + \varepsilon\beta_{hh} - \varepsilon^2\alpha_{hh}(\beta_{hh} - \rho_h\alpha_{hh}) + \varepsilon^2\sum_k{}' \frac{(\alpha_{kh}\rho_h - \beta_{kh})^2}{\rho_h - \rho_k}.$$

При этом штрих, поставленный у знака суммы, означает, что суммирование распространяется на все значения k от 1 до n, кроме значения $k=h$.

7. Наложение связи.

$$\gamma_1 x_1 + \ldots + \gamma_n x_n = 0$$

и связанное с этим уменьшение числа переменных квадратичной формы

$$K(x,\ x) = \sum_{p,q=1}^{n} k_{pq} x_p x_q$$

можно представить себе осуществленным при помощи непрерывного процесса, а именно рассматриваем квадратичную форму

$$K(x,\ x) + t(\gamma_1 x_1 + \ldots + \gamma_n x_n)^2,$$

где t — положительный параметр. Если t неограниченно возрастает, то и каждое из характеристических чисел монотонно возрастает. В частности наибольшее из характеристических чисел возрастает неограниченно, между тем как остальные переходят в характеристические числа квадратичной формы, получающейся из $K(x, x)$ исключением одной переменной при помощи соотношения $\gamma_1 x_1 + \ldots + \gamma_n x_n = 0$.

8. Элементарные делители матрицы или билинейной формы. Пусть $\mathfrak{A}$ тензор, $A = (a_{ik})$ — соответствующая матрица. Тогда полином

$$|\varkappa E - A| = \begin{vmatrix} \varkappa - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \varkappa - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & \varkappa - a_{nn} \end{vmatrix}$$

разлагается по известным правилам, которые мы здесь приводить не станем, на „элементарные делители“:

$$(\varkappa - r_1)^{e_1}, \quad (\varkappa - r_2)^{e_2}, \ldots, (\varkappa - r_h)^{e_h},$$

причем среди чисел $r_1, r_2, \ldots, r_h$ могут быть и равные между собой. Каждому элементарному делителю $(\varkappa - r_\nu)^{e_\nu}$ соответствуют e_ν векторов $\mathfrak{f}_1^{(\nu)}, \mathfrak{f}_2^{(\nu)}, \ldots, \mathfrak{f}_{e_\nu}^{(\nu)}$, для которых имеют место равенства:

$$\mathfrak{A}\mathfrak{f}_1^{(\nu)} = r_\nu \mathfrak{f}_1^{(\nu)}; \quad \mathfrak{A}\mathfrak{f}_2^{(\nu)} = r_\nu \mathfrak{f}_2^{(\nu)} + \mathfrak{f}_1^{(\nu)}, \ldots, \mathfrak{A}\mathfrak{f}_{e_\nu}^{(\nu)} = r_\nu \mathfrak{f}_{e_\nu}^{(\nu)} + \mathfrak{f}_{e_\nu - 1}^{(\nu)}.$$

При этом n векторов

$$\mathfrak{f}_1^{(1)}, \ldots, \mathfrak{f}_{e_1}^{(1)}; \mathfrak{f}_1^{(2)}, \ldots, \mathfrak{f}_{e_2}^{(2)}; \ldots; \mathfrak{f}_1^{(h)}, \ldots \mathfrak{f}_{e_h}^{(h)}$$

линейно независимы. Если ввести новые переменные $x_1^{(1)}, \ldots, x_{e_h}^{(h)}$, приняв эти векторы за единичные векторы новой системы, то матрица A превратится в матрицу:

$$\begin{pmatrix} A_1 & 0 & \ldots & 0 & 0 \\ 0 & A_2 & \ldots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \ldots & 0 & A_h \end{pmatrix},$$

где $A_1, A_2, \ldots, A_h$ в свою очередь являются матрицами, среди которых могут оказаться матрицы, состоящие только из одного элемента, а именно A_ν представляет матрицу порядка e_ν:

$$A_\nu = \begin{pmatrix} r_\nu & 0 & 0 & \ldots & 0 & 0 \\ 1 & r_\nu & 0 & \ldots & 0 & 0 \\ 0 & 1 & r_\nu & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & r_\nu \end{pmatrix}.$$

9. Спектр унитарной матрицы. Докажем, что спектр унитарной матрицы лежит на окружности радиуса 1.

Для доказательства заметим, что все элементы унитарной матрицы по абсолютному значению не больше единицы, а потому абсолютные значения коэфициентов характеристических уравнений всех унитарных матриц n-го порядка меньше некоторой грани, не зависящей от частного вида матрицы. Так как первый и последний коэфициенты характеристического уравнения по абсолютному значению равны единице, то тем самым доказано существование независящих от частного вида матрицы верхней и положительной нижней грани для абсолютных значений характеристических чисел. С другой стороны, все степени A^m унитарной матрицы являются унитарными, а их характеристическими числами являются m-е степени характеристических чисел матрицы A. Эти степени и обратные им величины могут только в том случае оставаться по абсолютному значению меньше некоторой грани, не зависящей от m, если абсолютное значение λ_i равно единице.

Другой метод доказательства, который может быть распространен и на бесконечные матрицы, получается путем изучения сходимости ряда Неймана для $(E-\lambda A)^{-1}$.

В самом деле, ряд

$$(E-\lambda A)^{-1}=E+\lambda A+\lambda^2A^2+\lambda^3A^3+\ldots,$$

где A — унитарная матрица, непременно сходится, пока $|\lambda|<1$, так как элементы матрицы A^m все по абсолютному значению не превосходят единицы, и потому для элементов, стоящих в правой части, геометрический ряд является мажорантой. Таким образом нули $|E-\lambda A|$ не могут лежать внутри круга радиуса 1. С другой стороны, ввиду того, что $A\bar{A}'=E$, имеем:

$$(E-\lambda A)^{-1}=-\frac{1}{\lambda}\bar{A}'\left(E+\frac{1}{\lambda}\bar{A}'+\frac{1}{\lambda^2}\bar{A}'^2+\ldots\right),$$

геометрический ряд в правой части сходится, если $\left|\frac{1}{\lambda}\right|<1$, так как и $\bar{A}'$ есть унитарная матрица. Вместе с тем нули $|E-\lambda A|$ не могут лежать и вне круга радиуса 1. Следовательно, эти нули должны лежать на окружности радиуса 1, как мы и утверждали.

Литература к главе I.

Учебники.

Kowalewski G., Einführung in die Determinantentheorie, Leipzig 1909.
Bôcher M., Einführung in die höhere Algebra, Leipzig und Berlin 1910. (Русский перевод. *Бохер*, Введение в высшую алгебру, ГТТИ, 1933 г.)

Монографии и статьи.

Hilbert D., Grundzüge einer allgemeinen Theorie der linearen-Integralgleichungen (в особенности первый и четвертый отдел). Leipzig und Berlin 1912.
Fischer E., Ueber quadratische Formen mit reellen Koeffizienten, Monatsh. f. Math. u. Phys., т. 16, стр. 234—249, 1905. Там, пожалуй, впервые указано максимально-минимальное свойство собственных значений.
Courant R., Zur Theorie der kleinen Schwingungen, Ztschr. f. angew. Math. и Mech., т. 2, стр. 278—285, 1922.
Wintner A., Spektraltheorie der unendlichen Matrizen, Leipzig 1929.

Глава II.

Задача о разложении в ряд произвольных функций.

Многие из соотношений, рассмотренных в предыдущей главе, находят далеко идущую аналогию, если вместо векторов в n-мерном пространстве рассматривать функции от одной или многих переменных, определенные в данной *основной области* G. Так, например, тому факту, что в пространстве n измерений всякий вектор может быть линейно представлен с помощью n произвольно выбранных независимых векторов, соответствует задача о выражении более или менее произвольно взятой функции, определенной в основной области G, в виде линейной комбинации заданных функций. (Множество заданных функций должно быть бесконечным, в чем мы непосредственно убедимся в дальнейшем.) Мы говорим тогда *о задаче разложения произвольно взятой функции по заданной системе функций.*

В настоящей главе мы рассмотрим с общей точки зрения этот вопрос, встречающийся в самых разнообразных формах в задачах математической физики.

При этом мы ограничиваемся *кусочно-непрерывными функциями*, т. е. такими функциями, для которых основная область G может быть разбита на конечное число частичных областей так, чтобы функция внутри каждой из них была непрерывна и стремилась при произвольном приближении изнутри к границе частичной области к определенному конечному пределу. Для более удобной записи мы сначала будем предполагать, что мы имеем дело с функциями только от одного переменного x, основной областью G которых является конечный отрезок оси x.

Если речь будет итти о функциях от многих переменных, например от двух переменных x и y, то мы будем предполагать, что основная область G ограничена конечным числом дуг кривых, с непрерывно вращающейся касательной. Когда мы будем считать точки границы принадлежащими основной области, то мы будем говорить о „замкнутой области“, если только это не вытекает из самого текста.

Далее, мы часто будем предполагать относительно рассматриваемых функций, что они *кусочно-гладкие* т. е. что они кусочно-непрерывны и имеют кусочно-непрерывные первые производные. Мы предполагаем, что наши функции имеют действительные значения в том случае, когда не оговорено противоположное.

§ 1. Ортогональные системы функций.

1. Определения. Под *скалярным* или *внутренним произведением* (f, g) *или* (fg) *двух функций* $f(x)$ и $g(x)$ мы разумеем взятый по основной области интеграл[1]).

$$(f, g) = \int fg\,dx. \tag{1}$$

Это произведение удовлетворяет неравенству Шварца:

$$(f, g)^2 \leqslant (f, f)(g, g), \tag{2}$$

которое так же, как и в случае векторов, либо вытекает из определенного положительного характера квадратичной относительно λ функции

$$\int (\lambda f + g)^2 dx,$$

либо же следует непосредственно из тождества:

$$(f, g)^2 = (f, f)(g, g) - \frac{1}{2}\iint [f(x)g(\xi) - f(\xi)g(x)]^2 dx\,d\xi;$$

знак равенства имеет место в том и только в том случае, если f и g пропорциональны. Две функции, скалярное произведение которых равно нулю, будем называть *ортогональными*. Скалярное произведение функции $f(x)$ на самое себя будем называть *нормом* этой функции и писать так:

$$Nf = (f, f) = \int f^2 dx; \tag{3}$$

функцию, норм которой равен единице, назовем *нормированной* функцией. Систему нормированных функций $\varphi_1(x)$, $\varphi_2(x), \ldots$, в которой каждые две различные функции взаимно ортогональны, будем называть *ортогональной нормированной системой*, а характеризующие ее соотношения

$$(\varphi_\nu, \varphi_\mu) = e_{\nu\mu} \quad (e_{\nu\nu} = 1, \quad e_{\nu\mu} = 0 \text{ при } \mu \neq \nu)$$

назовем *соотношениями ортогональности*.

Пример ортогональной нормированной системы функций в интервале $0 \leqslant x \leqslant 2\pi$ или вообще в любом интервале длины 2π представляют функции

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos x}{\sqrt{\pi}}, \quad \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \quad \frac{\sin x}{\sqrt{\pi}}, \quad \frac{\sin 2x}{\sqrt{\pi}}, \ldots$$

Для функций действительного переменного, принимающих комплексные значения, удобно ввести обобщение понятия ортогональности.

---

[1]) Мы в дальнейшем опускаем границы интегрирования там, где это не может привести к недоразумениям.

Две таких функции $f(x)$ и $g(x)$ называются ортогональными, если имеют место соотношения:

$$(f, \overline{g}) = (\overline{f}, g) = 0,$$

причем $\overline{f}$ и $\overline{g}$ означают, как это обычно принято, сопряженные комплексные функции по отношению к f и g.

Функция $f(x)$ называется *нормированной*, если

$$Nf = \int |f|^2 dx = 1.$$

Простейший пример комплексной ортогональной системы представляют в интервале $-\pi \leqslant x \leqslant \pi$ показательные функции:

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{e^{ix}}{\sqrt{2\pi}}, \quad \frac{e^{2ix}}{\sqrt{2\pi}}, \dots,$$

что непосредственно видно из „*соотношений ортогональности*“:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(\mu-\nu)x} dx = e_{\mu\nu} \quad (e_{\nu\nu} = 1, \quad e_{\mu\nu} = 0 \text{ при } \mu \neq \nu). \tag{4}$$

Функции $f_1, \dots, f_r$ называются *линейно зависимыми*, если они удовлетворяют тождественно относительно x однородному линейному соотношению

$$\sum_{i=1}^{r} c_i f_i = 0$$

с постоянными коэфициентами c_i $(i = 1, \dots, r)$, которые не все равны нулю. В противном случае эти r функций называются *линейно независимыми*. Важно заметить, что функции ортогональной системы всегда линейно независимы. В самом деле, из соотношения

$$c_1\varphi_1 + c_2\varphi_2 + \dots + c_n\varphi_n = 0$$

следовало бы, если его умножить на φ_ν и интегрировать, что $c_\nu = 0$.

2. Ортогонализация функций. Из заданной бесконечной системы функций $v_1, v_2, \dots$, обладающей тем свойством, что при любом r каждые r произвольно выбранных функций линейно независимы, можно при помощи „процесса ортогонализации“ получить ортогональную нормированную систему функций $\varphi_1, \varphi_2, \dots$, выбирая φ_n как соответствующую линейную комбинацию функций $v_1, \dots, v_n$. Сначала полагаем

$$\varphi_1 = \frac{v_1}{\sqrt{Nv_1}}.$$

Затем определяем любые два не обращающихся одновременно в нуль числа c_1 и c_2 так, чтобы функция $\varphi_2' = c_1\varphi_1 + c_2 v_2$ была ортогональной к φ_1, т. е. чтобы имело место равенство:

$$c_1 + c_2(\varphi_1, v_2) = 0.$$

Функция φ_2' в силу линейной независимости v_1 и v_2, а следовательно, и функций φ_1 и v_2 не может тождественно равняться нулю.

Таким образом

$$\varphi_2 = \frac{\varphi_2'}{\sqrt{N\varphi_2'}}$$

представляет нормированную функцию, ортогональную к φ_1.

Далее, образуем функцию

$$\varphi_3' = c_1^*\varphi_1 + c_2^*\varphi_2 + c_3^*v_3,$$

выбирая три не равных одновременно нулю числа c_1^*, c_2^*, c_3^*, удовлетворяющие двум линейным однородным уравнениям:

$$(\varphi_3'\varphi_1) = c_1^* + c_3^*(\varphi_1 v_3) = 0, \quad (\varphi_3'\varphi_2) = c_2^* + c_3^*(\varphi_2 v_3) = 0.$$

В силу линейной независимости v_1, v_2, v_3, а вместе с тем и $\varphi_1, \varphi_2, v_3$ функция φ_3' не может равняться нулю тождественно, и потому

$$\varphi_3 = \frac{\varphi_3'}{\sqrt{N\varphi_3'}}$$

представляет нормированную функцию, ортогональную к φ_1 и φ_2.

Продолжая неограниченно этот процесс, мы получаем искомую ортогональную систему функций с помощью рекуррентной формулы:

$$\varphi_{n+1} = \frac{\varphi'_{n+1}}{\sqrt{N\varphi'_{n+1}}}, \quad \varphi'_{n+1} = v_{n+1} - \sum_{h=1}^{n} \varphi_h(\varphi_h v_{n+1}).$$

Когда мы в дальнейшем будем говорить об ортогонализации, то мы всегда будем разуметь только что указанный процесс, который одновременно с ортогонализацией дает и нормирование, если только не будет явно указано нечто другое.

3. Неравенство Бесселя. Условие полноты системы. Аппроксимирование в среднем. Если дана ортогональная нормированная система функций φ_1, φ_2, ... и любая функция f, то числа

$$c_\nu = (f\varphi_\nu) \quad (\nu = 1, 2, \ldots) \tag{5}$$

называются *коэфициентами разложения* или *компонентами* функции f относительно заданной ортогональной системы[1]).

Из непосредственно очевидного соотношения

$$\int \left(f - \sum_{\nu=1}^{n} c_\nu\varphi_\nu\right)^2 dx \geqslant 0 \tag{6}$$

[1]) В связи с теорией рядов Фурье иногда употребляют также выражение „коэфициенты Фурье".

путем возведения в квадрат и почленного интегрирования получаем:

$$0 \leqslant \int f^2 dx - 2\sum_{\nu=1}^{n} c_\nu \int \varphi_\nu f\,dx + \sum_{\nu=1}^{n} c_\nu^2 = Nf - 2\sum_{\nu=1}^{n} c_\nu^2 + \sum_{\nu=1}^{n} c_\nu^2,$$

откуда

$$\sum_{\nu=1}^{n} c_\nu^2 \leqslant Nf; \tag{7}$$

так как в правой части находится постоянное, не зависящее от n число, то

$$\sum_{\nu=1}^{\infty} c_\nu^2 \leqslant Nf. \tag{8}$$

Это основное неравенство, *неравенство Бесселя*, справедливо для любой ортогональной нормированной системы. Это неравенство доказывает сходимость ряда с неотрицательными членами, составленного из квадратов коэфициентов разложения, находящегося в левой части соотношения (8).

Для системы функций, принимающих комплексные значения, справедливо соответствующее соотношение:

$$\sum_{\nu=1}^{\infty} |c_\nu|^2 \leqslant Nf = (f, \bar{f}), \tag{8'}$$

если под c_ν разуметь коэфициент разложения $c_\nu = (f, \bar{\varphi}_\nu)$.

Оно вытекает аналогично случаю действительных функций из неравенства:

$$\int \left| f(x) - \sum_{\nu=1}^{n} c_\nu \varphi_\nu \right|^2 dx = Nf - \sum_{\nu=1}^{n} |c_\nu|^2 \geqslant 0.$$

Интегральное выражение, стоящее в левой части формулы (6) получается совершенно естественно, если поставить себе задачу *аппроксимировать в смысле метода наименьших квадратов, данную функцию $f(x)$ с помощью линейного агрегата*

$$\sum_{\nu=1}^{n} \gamma_\nu \varphi_\nu$$

с постоянными коэфициентами γ_ν и фиксированным числом слагаемых и так, чтобы „средняя квадратичная ошибка"

$$M = \int (f - \sum \gamma_\nu \varphi_\nu)^2 dx$$

была возможно меньше. Действительно, путем простого преобразования интеграла получаем тождество:

$$M = \int \left(f - \sum_{\nu=1}^{n} \gamma_\nu \varphi_\nu\right)^2 dx = \int f^2 dx + \sum_{1}^{n} (\gamma_\nu - c_\nu)^2 - \sum_{1}^{n} c_\nu^2,$$

из которого непосредственно следует, что минимум M достигается при $\gamma_\nu = c_\nu$.

Если для любой кусочно-непрерывной функции f можно сделать наименьшую среднюю квадратичную ошибку

$$\int \left(f - \sum_{\nu=1}^{n} c_\nu \varphi_n\right)^2 dx$$

путем соответствующего выбора n меньше сколь угодно малого положительного числа, т. е. если можно каждую такую функцию аппроксимировать в смысле способа наименьших квадратов или, как мы будем говорить, *в „среднем“* с произвольной точностью, с помощью линейного аггрегата $\sum\limits_{\nu=1}^{n} c_\nu \varphi_\nu$ с достаточно большим числом членов, то систему функций $\varphi_1, \varphi_2, \ldots$ мы будем называть *„полной ортогональной системой функций“*.

На основании предыдущих рассуждений коэфициенты разложения $c_\nu = (f\varphi_\nu)$ любой функции f удовлетворяют соотношению:

$$\sum_{\nu=1}^{\infty} c_\nu^2 = Nf, \tag{9}$$

которое мы будем называть *„условием полноты“*.

Это условие можно записать в более общем виде:

$$\sum_{\nu=1}^{\infty} c_\nu d_\nu = (f, g), \tag{9'}$$

где

$$c_\nu = (f\varphi_\nu), \quad d_\nu = (g\varphi_\nu),$$

который получается, если применить формулу (9) к функции $f+g$:

$$N(f+g) = Nf + Ng + 2(f, g) = \sum_{\nu=1}^{\infty} (c_\nu + d_\nu)^2 = \sum_{\nu=1}^{\infty} (c_\nu^2 + d_\nu^2 + 2c_\nu d_\nu)$$

и затем вычесть соответствующие равенства для f и g.

Впрочем, для полноты системы функций $\varphi_1, \varphi_2, \ldots$ достаточно, чтобы условие полноты (9) было выполнено для всех непрерывных функций f. В самом деле, всякая кусочно-непрерывная функция g может быть аппроксимирована с помощью непрерывной функции f так, чтобы интеграл $\int (f-g)^2 dx$ имел сколь угодно малое значение. Такую функцию f можно, например, построить следующим образом: представим себе кривую, изображающую функцию g; около каждой точки разрыва x_i этой функции отметим на кривой две точки с абсциссами $x_i - \delta$ и $x_i + \delta$, где δ можно выбирать сколь угодно малым; эту пару точек соединим прямолинейным отрезком и этим отрезком заменяем нашу кривую в каждом таком интервале [1]). Если $a_1, a_2, \ldots$ представляют коэфициенты разложения функции g, а $c_1, c_2 \ldots$ — коэфициенты разложения функции f, то из того, что интеграл $\int\left(f-\sum c_\nu \varphi_\nu\right)^2 dx$ может быть сделан путем соответствующего выбора n сколь угодно малым, следует на основании неравенства Шварца справедливость аналогичного утверждения для интеграла:

$$M' = \int\left(g-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right)^2 dx = \int\left[(g-f)+\left(f-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right)\right]^2 dx.$$

В самом деле,

$$M' = N(g-f) + N\left(f-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right) + 2\left(g-f, f-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right),$$

следовательно,

$$M' \leqslant N(g-f) + N\left(f-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right) + 2\sqrt{N(g-f)\,N\left(f-\sum c_\nu \varphi_\nu\right)}.$$

Но

$$M = \int\left(g-\sum a_\nu \varphi_\nu\right)^2 dx \leqslant M',$$

так как коэфициенты разложения a_ν для g дают наименьшую квадратичную ошибку. Таким образом условие полноты доказано и для функции f.

Следует обратить внимание на то, что из полноты системы функций $\varphi_1, \varphi_2, \ldots$, т. е. из равенства

$$\lim_{n\to\infty}\int\left(f-\sum_{\nu=1}^{n} c_\nu \varphi_\nu\right)^2 dx = 0$$

ни в коем случае нельзя делать заключение, что $f = \sum_{\nu=1}^{\infty} c_\nu \varphi_\nu$, т. е. что функция f разлагается в ряд по функциям φ_ν. Однако, *если ряд* $\sum_{\nu=1}^{\infty} c_\nu \varphi_\nu$ *равномерно сходится и мы можем поэтому сделать переход к предельной функции под знаком интеграла, то разложимость функции f очевидна*. Полнота данной системы $\varphi_1, \varphi_2, \ldots$ является при этом, конечно, необходимым условием; действительно, выделив, например, из полной системы одну функцию, мы видим, что все компоненты ее относительно остающейся неполной системы равны нулю. Но и для полной системы

[1]) В самом деле, пусть M означает верхнюю грань $|g(x)|$, а q — число точек разрыва функции $g(x)$ в промежутке интегрирования, тогда в неравенстве

$$\int (f-g)^2 dx < 8 M^2 \delta q$$

правую часть можно сделать сколь угодно малой при соответствующем выборе δ.

функций $\varphi_1, \varphi_2, \ldots$ вопрос о разложимости функции f требует более подробного исследования, которое мы в дальнейшем (гл. V и VI) будем еще проводить в различных случаях.

Содержание предыдущего предельного равенства мы будем также выражать следующим образом: *последовательность функций* $\sum_{\nu=1}^{n} c_\nu \varphi_\nu$ *сходится в среднем к функции* f.

Далее, приводим теорему: *кусочно-непрерывная функция однозначно определяется своими коэфициентами разложения по заданной полной ортогональной системе, т. е. две кусочно-непрерывные функции тождественны между собой, если их коэфициенты разложения соответственно равны.* Действительно, разность двух таких функций с равными коэфициентами имеет коэфициенты, равные нулю, и следовательно, в силу условия полноты и норм ее равен нулю; следовательно, эта разность сама должна тождественно равняться нулю. Таким образом функция однозначно характеризуется своим разложением по полной ортогональной системе функций и в том случае, когда разложение сходится не в обычном смысле, а только в среднем. Во многих рассуждениях достаточно бывает этой сходимости в среднем.

Понятие полноты системы функций сохраняет смысл и в том случае, когда система не ортогональна и не нормирована. Вообще мы будем называть систему функций полной системой, если любая кусочно-непрерывная функция может быть аппроксимирована в среднем с любой точностью при помощи линейного агрегата этих функций. Полнота такой системы переносится также на ортогональную систему, получающуюся из нее путем ортогонализации.

4. Ортогональные и унитарные преобразования бесконечно большого числа переменных. Ортогональные нормированные системы функций аналогичны во многих отношениях ортогональным системам нормированных векторов в n-мерном пространстве; компоненты $c_\nu = (f\varphi_\nu)$ функции f можно рассматривать как прямоугольные координаты функции f в системе координат, определенной при помощи „координатных функций" $\varphi_1, \varphi_2 \ldots$ в пространстве бесконечно большого числа измерений.

Если $\psi_1, \psi_2, \ldots$ — другая ортогональная нормированная система, в которой компонентами функции f служат $d_\nu = (f\psi_\nu)$, и если обе системы являются полными, то применение условия полноты (9') к функции f и функциям φ_i по отношению к системе $\psi_1, \psi_2, \ldots$, непосредственно дает бесконечную систему равенств:

$$c_i = \sum_{k=1}^{\infty} a_{ik} d_k, \quad a_{ik} = (\varphi_i \psi_k) \quad (i = 1, 2, \ldots). \tag{10}$$

Соответственным образом получаем обратную систему равенств:

$$d_i = \sum_{k=1}^{\infty} a_{ki} c_k, \quad a_{ki} = (\psi_i \varphi_k) \quad (i = 1, 2, \ldots). \tag{10'}$$

Коэфициенты удовлетворяют условиям:

$$\sum_{k=1}^{\infty} a_{ik} a_{jk} = (\varphi_i \varphi_j) = \begin{cases} 0 \text{ при } i \neq j \\ 1 \text{ при } i = j \end{cases} \tag{11}$$

$$\sum_{k=1}^{\infty} a_{ki} a_{kj} = (\psi_i \psi_j) = \begin{cases} 0 \text{ при } j \neq i \\ 1 \text{ при } j = i \end{cases} \tag{11'}$$

представляющим обобщение условий ортогональности в пространстве n измерений (гл. I, § 1) на пространство бесконечно большого числа измерений. Поэтому преобразование (10), удовлетворяющее условиям (11) и (11'), называют *ортогональным преобразованием бесконечно большого числа переменных.*

Аналогично устанавливается связь между коэфициентами двух комплексных ортогональных систем при помощи *унитарного преобразования бесконечно большого числа переменных.*

5. Справедливость результатов в случае нескольких независимых переменных. Расширение предпосылок. Все установленные нами понятия и рассуждения остаются справедливыми, если вместо функций от одной переменной x рассматривать функции от нескольких переменных, например от x и y, причем переменные изменяются в заданной конечной области G, элемент площади которой обозначим через dG. Внутреннее произведение (f, g) двух функций $f(x, y)$ и $g(x, y)$, определенных в этой области G, мы определяем равенством $(f, g) = \int_G fg\,dG$, и тогда в обозначениях и доказательствах этого параграфа не приходится делать никаких существенных изменений.

Далее, все установленные понятия и факты сохраняются и в том случае, если считать основную область бесконечной и допустить, что все встречающиеся функции вместе с их квадратами интегрируемы во всей основной области.

Наконец, заметим, что наши понятия сохраняют смысл, если функция f обращается в бесконечность в основной области так, что квадрат ее интегрируем во всей основной области.

6. Построение полных систем функций от многих переменных. Если известны полные системы функций от одной переменной, то можно построить полные системы функций от двух и большего числа переменных на основании следующей теоремы: если система функций

$$\varphi_1(s),\ \varphi_2(s), \ldots$$

представляет полную ортогональную нормированную систему функций в интервале $a \leqslant s \leqslant b$ и если для любого $i = 1, 2 \ldots$ в интервале $c \leqslant t \leqslant d$ функции

$$\psi_{1i}(t),\ \psi_{2i}(t); \ldots$$

образуют такую же систему, то функции

$$\omega_{ik}(s, t) = \varphi_i(s)\, \psi_{ki}(t)$$

образуют полную ортогональную систему функций от s и t в прямоугольнике $a \leqslant s \leqslant b$, $c \leqslant t \leqslant d$. В частности система функций $\varphi_i(s)\varphi_h(t)$ является ортогональной и полной системой в области квадрата $a \leqslant s \leqslant b$, $a \leqslant t \leqslant b$.

Если $f(s, t)$ является непрерывной функцией от s и t в этом прямоугольнике, то имеет место условие полноты:

$$\iint f^2(s, t)\,ds\,dt = \sum_{i,k=1}^{n}\left(\iint f(s, t)\,\omega_{ik}(s, t)\,ds\,dt\right)^2.$$

Для доказательства исходим из соотношения:

$$\int f^2(s, t)\,ds = \sum_{i=1}^{\infty} g_i^2(t),$$

где $g_i(t) = \int f(s,t)\,\varphi_i(s)\,ds$, выражающего полноту системы функций φ_i. Так как ряд в правой части сходится равномерно[1]), то мы имеем право почленно интегрировать по t и получаем:

$$\iint f^2(s,t)\,ds\,dt = \sum_{i=1}^{\infty}\int g_i^2(t)\,dt.$$

К i-му члену в правой части применяем условие полноты по отношению к системе функций $\psi_{ki}(t)$ $(k = 1, 2, \ldots)$ и непосредственно получаем искомое соотношение.

§ 2. Принцип предельных точек в функциональном пространстве.

1. **Сходимость в функциональном пространстве.** Аналогия между функциями и векторами в n-мерном пространстве часто

---

[1]) Это следует из теоремы Дини (Dini): если ряд $\sum_{\nu=1}^{\infty} f_\nu(t)$ *положительных* непрерывных функций, сходящихся в замкнутой области G, представляет непрерывную функцию $S(t)$, то этот ряд сходится равномерно. Наметим здесь доказательство в самых общих чертах. Полагаем $S_n(t) = \sum_{\nu=1}^{n} f_\nu(t)$, $S(t) = S_n(t) + R_n(t)$. Если бы утверждение было неправильным, то существовали бы положительное число α, неограниченно возрастающая последовательность чисел $n_1, n_2, n_3, \ldots$ и соответствующие значения $t_1, t_2, t_3, \ldots$ в области G такие, что $R_{n_i}(t_i) \geqslant \alpha$, и следовательно, $S_{n_i}(t_i) \leqslant S(t_i) - \alpha$. При этом мы можем допустить, что значения t_i стремятся к некоторому пределу t из области G. Пусть теперь N означает определенно выбранное число, тогда при $n_i \geqslant N$ и $S_{n_i}(t_i) \geqslant S_N(t_i)$, следовательно, $S_N(t_i) \leqslant S(t_i) - \alpha$. Здесь мы неограниченно увеличиваем i и получаем в силу наших предположений о непрерывности:

$$S_N(t) \leqslant S(t) - \alpha,$$

что при достаточно большом N, конечно, невозможно.