1. Линейное пространство: определение, примеры, простейшие следствия из аксиом.

Линейным (***векторным***) пространством называется множество V произвольных элементов, называемых ***векторами***, в котором определены операции сложения векторов и умножения вектора на число, т.е. любым двум векторам u и v поставлен в соответствие вектор $u+v$, называемый ***суммой векторов*** u и v, любому вектору v и любому числу λ из поля действительных чисел R поставлен в соответствие вектор λv, называемый ***произведением вектора*** v ***на число*** λ; так что выполняются следующие условия:

1. $u+v=v+u \quad \forall u,v\in V$ (коммутативность сложения);
2. $u+(v+w)=(u+v)+w \quad \forall u,v,w\in V$ (ассоциативность сложения);
3. существует такой элемент $o\in V$, называемый ***нулевым вектором***, что $v+o=v \quad \forall v\in V$;
4. для каждого вектора v существует такой вектор $(-v)\in V$, называемый ***противоположным*** вектору v, что $v+(-v)=o$;
5. $\lambda(u+v)=\lambda u+\lambda v \quad \forall u,v\in V,\ \forall\lambda\in R$;
6. $(\lambda+\mu)v=\lambda v+\mu v \quad \forall v\in V,\ \forall\lambda,\mu\in R$;
7. $\lambda(\mu v)=(\lambda\mu)v \quad \forall v\in V,\ \forall\lambda,\mu\in R$;
8. $1\cdot v=v \quad \forall v\in V$.

Условия 1–8 называются ***аксиомами линейного пространства***. Знак равенства, поставленный между векторами, означает, что в левой и правой частях равенства представлен один и тот же элемент множества V. Такие векторы называются ***равными***.

В определении линейного пространства операция умножения вектора на число введена для действительных чисел. Такое пространство называют ***линейным пространством над полем действительных*** (***вещественных***) ***чисел***, или, короче, ***вещественным линейным пространством***. Если в определении вместо поля R действительных чисел взять поле комплексных чисел, то получим ***линейное пространство над полем комплексных чисел***, или, короче, ***комплексное линейное пространство***. В качестве числового поля можно выбрать и поле Q рациональных чисел, при этом получим линейное пространство над полем рациональных чисел.

8.1.2. Простейшие следствия аксиом

1. В линейном пространстве существует единственный нулевой вектор.
2. В линейном пространстве для любого вектора $v\in V$ существует единственный противоположный вектор $(-v)\in V$.
3. Произведение произвольного вектора пространства на число нуль равно нулевому вектору, т.е. $0\cdot v=o \quad \forall v\in V$.
4. Произведение нулевого вектора на любое число равно нулевому вектору, т.е. $\lambda o=o$ для любого числа λ.
5. Вектор, противоположный данному вектору, равен произведению данного вектора на число (-1), т.е. $(-v)=(-1)v \quad \forall v\in V$.
6. В выражениях вида $a+b+\dots+z$ (сумма конечного числа векторов) или $\alpha\cdot\beta\cdot\ldots\cdot\omega\cdot v$ (произведение вектора на конечное число множителей) можно расставлять скобки в любом порядке, либо вообще не указывать.

Докажем, например, первые два свойства. Единственность нулевого вектора. Если o и o' – два нулевых вектора, то по аксиоме 3 получаем два равенства: $o'+o=o'$ или $o+o'=o$, левые части которых равны по аксиоме 1. Следовательно, равны и правые части, т.е. $o=o'$. Единственность противоположного вектора. Если вектор $v\in V$ имеет два противоположных вектора $(-v)$ и $(-v)'$, то по аксиомам 2, 3, 4 получаем их равенство:

$$(-v)'=(-v)'+\underbrace{v+(-v)}_{o}=\underbrace{(-v)'+v}_{o}+(-v)=(-v).$$

Остальные свойства доказываются аналогично.

Примеры:

1. Обозначим $\{o\}$ – множество, содержащее один нулевой вектор, с операциями $o+o=o$ и $\lambda o=o$. Для указанных операций аксиомы 1–8 выполняются. Следовательно, множество $\{o\}$ является линейным пространством над любым числовым полем. Это линейное пространство называется ***нулевым***.

2. Обозначим V_1, V_2, V_3 – множества векторов (направленных отрезков) на прямой, на плоскости, в пространстве соответственно с обычными операциями сложения векторов и умножения векторов на число. Выполнение аксиом 1–8 линейного пространства следует из курса элементарной геометрии. Следовательно, множества V_1, V_2, V_3 являются вещественными линейными пространствами. Вместо свободных векторов можно рассмотреть соответствующие множества радиус-векторов. Например, множество векторов на плоскости, имеющих общее начало, т.е. отложенных от одной фиксированной точки плоскости, является вещественным линейным пространством. Множество радиус-векторов единичной длины не образует линейное пространство, так как для любого из этих векторов сумма $v+v$ не принадлежит рассматриваемому множеству.

3. Обозначим R^n – множество матриц-столбцов размеров $n\times 1$ с операциями сложения матриц и умножения матриц на число. Аксиомы 1–8 линейного пространства для этого множества выполняются (см. разд. 1.2). Нулевым вектором в этом множестве служит нулевой столбец $o=(0 \ \cdots \ 0)^T$. Следовательно, множество R^n является вещественным линейным пространством. Аналогично, множество C^n столбцов размеров $n\times 1$ с комплексными элементами является комплексным линейным пространством. Множество матриц-столбцов с неотрицательными действительными элементами, напротив, не является линейным пространством, так как не содержит противоположных векторов.

2. Линейная зависимость и линейная независимость элементов линейного пространства. Свойства.

8.2.1. Понятие линейной зависимости и линейной независимости векторов

Для элементов линейного пространства были введены операции умножения вектора на число (из некоторого числового поля) и сложения векторов. При помощи этих операций можно составлять алгебраические выражения.

Вектор v называется ***линейной комбинацией*** векторов $v_1, v_2, \dots, v_k$, если

$$v=\alpha_1 v_1+\alpha_2 v_2+\dots+\alpha_k v_k, \tag{8.1}$$

где $\alpha_1, \alpha_2, \dots, \alpha_k$ – некоторые числа. В этом случае говорят, что ***вектор*** v ***разложен по векторам*** $v_1, v_2, \dots, v_k$ (***вектор*** v ***линейно выражается через векторы*** $v_1, v_2, \dots, v_k$), а числа $\alpha_1, \alpha_2, \dots, \alpha_k$ называют ***коэффициентами разложения***. Линейная комбинация с нулевыми коэффициентами $v=0\cdot v_1+0\cdot v_2+\dots+0\cdot v_k$ называется ***тривиальной***.

Набор векторов $v_1, v_2, \dots, v_k$ из V называется ***системой векторов***, а любая часть системы векторов – ***подсистемой***.

Система из k векторов $v_1, v_2, \dots, v_k$ называется ***линейно зависимой***, если существуют такие числа $\alpha_1, \alpha_2, \dots, \alpha_k$, не все равные нулю одновременно, что справедливо равенство

$$\alpha_1 v_1+\alpha_2 v_2+\dots+\alpha_k v_k=o, \tag{8.2}$$

т.е. линейная комбинация является нулевым вектором.

Система из k векторов $v_1, v_2, \dots, v_k$ называется ***линейно независимой***, если равенство (8.2) возможно только при $\alpha_1=\alpha_2=\dots=\alpha_k=0$, т.е. когда линейная комбинация в левой части (8.2) тривиальная.

Замечания 8.2.

1. Один вектор v_1 тоже образует систему: при $v_1=o$ – линейно зависимую, а при $v_1\ne o$ – линейно независимую.
2. Понятия линейной зависимости и линейной независимости для векторов определяются также, как для столбцов (см. разд. 3.1). Поэтому все свойства, рассмотренные в разд. 3.1, переносятся на векторы. Применение свойств, доказанных для векторов, к столбцам, можно делать без обоснования, так как множество столбцов является линейным пространством (см. п. 3 в разд. 8.1.3.).
3. ***Рангом системы векторов*** $v_1, v_2, \dots, v_k$ называется максимальное число линейно независимых векторов этой системы и обозначается $\operatorname{rg}(v_1, v_2, \dots, v_k)$.

8.2.2. Свойства линейно зависимых и линейно независимых векторов

1. *Если в систему векторов входит нулевой вектор, то она линейно зависима.*
2. *Если в системе векторов имеется два равных вектора, то она линейно зависима.*
3. *Если в системе векторов имеется два пропорциональных* (*коллинеарных*) *вектора* ($v_i=\lambda v_j$), *то она линейно зависима.*
4. *Система из $k>1$ векторов линейно зависима тогда и только тогда, когда хотя бы один из векторов есть линейная комбинация остальных.*
5. *Любые векторы, входящие в линейно независимую систему, образуют линейно независимую подсистему.*
6. *Система векторов, содержащая линейно зависимую подсистему, линейно зависима.*
7. *Если система векторов $v_1, v_2, \dots, v_k$ – линейно независима, а после присоединения к ней вектора v – оказывается линейно зависимой, то вектор v можно разложить по векторам $v_1, v_2, \dots, v_k$ и притом единственным образом, т.е. коэффициенты разложения* (8.1) *находятся однозначно.*
8. *Пусть каждый вектор системы $u_1, u_2, \dots, u_l$ может быть разложен по векторам системы $v_1, v_2, \dots, v_k$*, т.е. $u_i=\sum_{j=1}^{k}a_{ji}v_j$, $i=1,\dots,l$ (говорят, что ***система векторов*** $u_1, u_2, \dots, u_l$ ***линейно выражается через систему векторов*** $v_1, v_2, \dots, v_k$). *Тогда, если $l>k$, то система векторов $u_1, u_2, \dots, u_l$ – линейно зависима.*

Докажем, например, последнее свойство. Составим линейную комбинацию векторов $u_1, u_2, \dots, u_l$ с коэффициентами $x_1, x_2, \dots, x_l$ и приравняем ее нулевому вектору:

$$\sum_{i=1}^{l}x_i u_i=o. \tag{8.3}$$

Надо показать, что эта линейная комбинация может быть нетривиальной, т.е. среди коэффициентов x_i, $i=1,\dots,l$, можно взять числа, не равные нулю. Действительно, подставим в линейную комбинацию (8.3) разложения векторов u_i по векторам системы $v_1, v_2, \dots, v_k$:

$$o=\sum_{i=1}^{l}x_i u_i=\sum_{i=1}^{l}x_i\sum_{j=1}^{k}a_{ji}v_j=\sum_{j=1}^{k}\left(\sum_{i=1}^{l}a_{ji}x_i\right)v_j.$$

Чтобы это равенство выполнялось, достаточно потребовать, чтобы $\sum_{i=1}^{l}a_{ji}x_i=0$, $j=1,\dots,k$. Таким образом, получили однородную систему $Ax=o$ линейных уравнений с неизвестными $x_1, x_2, \dots, x_l$. Матрица $A=(a_{ji})$ системы имеет размеры $k\times l$, т.е. количество уравнений (k) меньше количества (l) неизвестных, так как $l>k$. Поэтому $\operatorname{rg}A\le k<l$, т.е. система имеет бесконечно много решений, в том числе и ненулевых (см. разд.5.5). Таким образом, линейная комбинация в (8.3) может быть нетривиальной, т.е. система векторов $u_1, u_2, \dots, u_l$ линейно зависима.

3. Размерность и базис линейного пространства. Примеры.

8.3. РАЗМЕРНОСТЬ И БАЗИС ЛИНЕЙНОГО ПРОСТРАНСТВА

8.3.1. Определения размерности и базиса

Линейное пространство V называется ***n-мерным***, если в нем существует система из n линейно независимых векторов, а любая система из большего количества векторов линейно зависима. Число n называется ***размерностью*** (***числом измерений***) линейного пространства V и обозначается $\dim V$. Другими словами, размерность пространства – это максимальное число линейно независимых векторов этого пространства. Если такое число существует, то пространство называется ***конечномерным***. Если же для любого натурального числа n в пространстве V найдется система, состоящая из n линейно независимых векторов, то такое пространство называют ***бесконечномерным*** (записывают: $\dim V=\infty$). Далее, если не оговорено противное, будут рассматриваться конечномерные пространства.

Базисом n-мерного линейного пространства называется упорядоченная совокупность n линейно независимых векторов (***базисных векторов***).

Замечания 8.4.

1. Базис линейного пространства определяется неоднозначно. Например, если $e_1, e_2, \dots, e_n$ – базис пространства V, то система векторов $\lambda e_1, \lambda e_2, \dots, \lambda e_n$ при любом $\lambda\ne 0$ также является базисом V. Количество базисных векторов в разных базисах одного и того же конечномерного пространства, разумеется, одно и то же, так как это количество равно размерности пространства.
2. В некоторых пространствах, часто встречающихся в приложениях, один из возможных базисов, наиболее удобный с практической точки зрения, называют ***стандартным***.
3. Теорема 8.1 позволяет говорить, что базис – это ***полная*** система элементов линейного пространства, в том смысле, что любой вектор пространства линейно выражается через базисные векторы.
4. Если множество L является линейной оболочкой $Lin(v_1, v_2, \dots, v_k)$, то векторы $v_1, v_2, \dots, v_k$ называют ***образующими множества*** L. Следствие 1 теоремы 8.1 в силу равенства $V=Lin(e_1, e_2, \dots, e_n)$ позволяет говорить, что базис – это *минимальная система образующих* линейного пространства V, так как нельзя уменьшить количество образующих (удалить хотя бы один вектор из набора $e_1, e_2, \dots, e_n$) без нарушения равенства $V=Lin(e_1, e_2, \dots, e_n)$.
5. Теорема 8.2 позволяет говорить, что базис – это *максимальная линейно независимая система векторов* линейного пространства, так как базис – это линейно независимая система векторов, и ее нельзя дополнить каким-либо вектором без потери линейной независимости.
6. Следствие 2 теоремы 8.1 удобно применять для нахождения базиса и размерности линейного пространства. В некоторых учебниках оно берется за определение базиса, а именно: *линейно независимая система $e_1, e_2, \dots, e_n$ векторов линейного пространства называется базисом, если любой вектор пространства линейно выражается через векторы $e_1, e_2, \dots, e_n$. Количество базисных векторов определяет размерность пространства.* Разумеется, что эти определения эквивалентны приведенным выше.

8.3.2. Примеры базисов линейных пространств

Укажем размерность и базис для примеров линейных пространств, рассмотренных в разд.8.1.3.

1. Нулевое линейное пространство $\{o\}$ не содержит линейно независимых векторов. Поэтому размерность этого пространства полагают равной нулю: $\dim\{o\}=0$. Это пространство не имеет базиса.

2. Пространства V_1, V_2, V_3 имеют размерности 1, 2, 3 соответственно. Действительно, любой ненулевой вектор пространства V_1 образует линейно независимую систему (см. п.1. замечаний 8.2), а любые два ненулевых вектора пространства V_1 коллинеарны, т.е. линейно зависимы (см. пример 8.1). Следовательно, $\dim V_1=1$, а базисом пространства V_1 является любой ненулевой вектор. Аналогично доказывается, что $\dim V_2=2$ и $\dim V_3=3$. Базисом пространства V_2 служат любые два неколлинеарных вектора, взятые в определенном порядке (один из них считается первым базисным вектором, другой – вторым). Базисом пространства V_3 являются любые три некомпланарных (не лежащих в одной или параллельных плоскостях) вектора, взятые в определенном порядке. Стандартным базисом в V_1 является единичный вектор $\vec{i}$ на прямой. Стандартным базисом в V_2 считается базис $\vec{i}, \vec{j}$, состоящий из двух взаимно перпендикулярных единичных векторов плоскости. Стандартным базисом в пространстве V_3 считается базис $\vec{i}, \vec{j}, \vec{k}$, составленный из трех единичных попарно перпендикулярных векторов, образующих правую тройку.

3. Пространство R^n содержит не более, чем n, линейно независимых векторов. В самом деле, возьмем k столбцов из R^n и составим из них матрицу размеров $n\times k$. Если $k>n$, то столбцы линейно зависимы по теореме 3.4 о ранге матрицы. Следовательно, $\dim R^n\le n$. В пространстве R^n нетрудно найти n линейно независимых столбцов. Например, столбцы единичной матрицы

$$e_1=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix},\ e_2=\begin{pmatrix}0\\1\\ \vdots\\0\end{pmatrix},\ \dots,\ e_n=\begin{pmatrix}0\\ \vdots\\0\\1\end{pmatrix}$$

линейно независимы. Следовательно, $\dim R^n=n$. Пространство R^n называется ***n-мерным вещественным арифметическим пространством***. Указанный набор векторов считается стандартным базисом пространства R^n. Аналогично доказывается, что $\dim C^n=n$, поэтому пространство C^n называют ***n-мерным комплексным арифметическим пространством***.

4. Напомним, что любое решение однородной системы $Ax=o$ можно представить в виде $x=C_1\varphi_1+C_2\varphi_2+\dots+C_{n-r}\varphi_{n-r}$, где $r=\operatorname{rg}A$, а $\varphi_1, \varphi_2, \dots, \varphi_{n-r}$ – фундаментальная система решений (см. разд. 5.5). Следовательно, $\{Ax=o\}=Lin(\varphi_1, \varphi_2, \dots, \varphi_{n-r})$, т.е. базисом пространства $\{Ax=o\}$ решений однородной системы служит ее фундаментальная система решений, а размерность пространства $\dim\{Ax=o\}=n-r$, где n – количество неизвестных, а r – ранг матрицы системы.

5. В пространстве $M_{2\times 3}$ матриц размеров 2×3 можно выбрать 6 матриц:

$$e_1=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix},\ e_2=\begin{pmatrix}0&1&0\\0&0&0\end{pmatrix},\ e_3=\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix},$$
$$e_4=\begin{pmatrix}0&0&0\\1&0&0\end{pmatrix},\ e_5=\begin{pmatrix}0&0&0\\0&1&0\end{pmatrix},\ e_6=\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix},$$

которые линейно независимы. Действительно, их линейная комбинация

$$\alpha_1\cdot e_1+\alpha_2\cdot e_2+\alpha_3\cdot e_3+\alpha_4\cdot e_4+\alpha_5\cdot e_5+\alpha_6\cdot e_6=\begin{pmatrix}\alpha_1&\alpha_2&\alpha_3\\ \alpha_4&\alpha_5&\alpha_6\end{pmatrix} \tag{8.5}$$

равна нулевой матрице только в тривиальном случае $\alpha_1=\alpha_2=\dots=\alpha_6=0$. Прочитав равенство (8.5) справа налево, заключаем, что любая матрица из $M_{2\times 3}$ линейным образом выражается через выбранные 6 матриц, т.е. $M_{2\times 3}=Lin(e_1, e_2, \dots, e_6)$. Следовательно, $\dim M_{2\times 3}=2\cdot 3=6$, а матрицы $e_1, e_2, \dots, e_6$ являются базисом (стандартным) этого пространства. Аналогично доказывается, что $\dim M_{m\times n}=m\cdot n$.

4. Теорема о разложении элементов линейного пространства по базису.

Теорема 8.1 (о разложении вектора по базису). *Если $e_1, e_2, \dots, e_n$ – базис n-мерного линейного пространства V, то любой вектор $v\in V$ может быть представлен в виде линейной комбинации базисных векторов:*

$$v=v_1e_1+v_2e_2+\dots+v_ne_n \tag{8.4}$$

и притом единственным образом, т.е. коэффициенты $v_1, v_2, \dots, v_n$ определяются однозначно. Другими словами, *любой вектор пространства может быть разложен по базису и притом единственным образом.*

Действительно, размерность пространства V равна n. Система векторов $e_1, e_2, \dots, e_n$ линейно независима (это базис). После присоединения к базису любого вектора v, получаем линейно зависимую систему $e_1, e_2, \dots, e_n, v$ (так как это система состоит из $(n+1)$ векторов n-мерного пространства). По свойству 7 линейно зависимых и линейно независимых векторов (см. разд.8.2.2) получаем заключение теоремы.

Следствие 1. *Если $e_1, e_2, \dots, e_n$ – базис пространства V, то* $V=Lin(e_1, e_2, \dots, e_n)$, т.е. *линейное пространство является линейной оболочкой базисных векторов.*

В самом деле, для доказательства равенства $V=Lin(e_1, e_2, \dots, e_n)$ двух множеств достаточно показать, что включения $V\subset Lin(e_1, e_2, \dots, e_n)$ и $Lin(e_1, e_2, \dots, e_n)\subset V$ выполняются одновременно. Действительно, с одной стороны, любая линейная комбинация векторов линейного пространства принадлежит самому линейному пространству, т.е. $Lin(e_1, e_2, \dots, e_n)\subset V$. С другой стороны, любой вектор пространства по теореме 8.1 можно представить в виде линейной комбинации базисных векторов, т.е. $V\subset Lin(e_1, e_2, \dots, e_n)$. Отсюда следует равенство рассматриваемых множеств.

Следствие 2. *Если $e_1, e_2, \dots, e_n$ – линейно независимая система векторов линейного пространства V и любой вектор $v\in V$ может быть представлен в виде линейной комбинации* (8.4): $v=v_1e_1+v_2e_2+\dots+v_ne_n$, *то пространство V имеет размерность n, а система $e_1, e_2, \dots, e_n$ является его базисом.*

В самом деле, в пространстве V имеется система n линейно независимых векторов, а любая система $u_1, u_2, \dots, u_k$ из большего количества векторов ($k>n$) линейно зависима (по свойству 8 в разд.8.2.2), поскольку каждый вектор из этой системы линейно выражается через векторы $e_1, e_2, \dots, e_n$. Значит, $\dim V=n$ и $e_1, e_2, \dots, e_n$ – базис V.

Примечание: (свойство 7)

7. *Если система векторов $v_1, v_2, \dots, v_k$ – линейно независима, а после присоединения к ней вектора v – оказывается линейно зависимой, то вектор v можно разложить по векторам $v_1, v_2, \dots, v_k$ и притом единственным образом, т.е. коэффициенты разложения* (8.1) *находятся однозначно.*

ОПРЕДЕЛЕНИЕ. *Коэффициенты в разложении вектора по базису называются координатами этого вектора в данном базисе.*

НАПРИМЕР.

1) Матрица $\mathbf{A}=\begin{pmatrix}1&-2\\-3&4\end{pmatrix}$ имеет в стандартном базисе $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4$ пространства $M(2\times 2,\mathbb{R})$ координаты $\{1;-2;-3;4\}$. Действительно,

$$\begin{pmatrix}1&-2\\-3&4\end{pmatrix}=1\begin{pmatrix}1&0\\0&0\end{pmatrix}-2\begin{pmatrix}0&1\\0&0\end{pmatrix}-3\begin{pmatrix}0&0\\1&0\end{pmatrix}+4\begin{pmatrix}0&0\\0&1\end{pmatrix},$$
$$\Rightarrow\ \mathbf{A}=\mathbf{E}_1-2\mathbf{E}_2-3\mathbf{E}_3+4\mathbf{E}_4.$$

2) n-мерный вектор $(1;-2;0)$ имеет в стандартном базисе пространства $\mathbb{R}^3$ координаты $\{1;-2;0\}$

3) Многочлен $f(x)=x^3+2x^2-x+4$ имеет в базисе $(x+1)^3$, $(x+1)^2$, $(x+1)$, 1 пространства $\mathbb{R}^4[x]$ координаты $\{1;-1;-2;6\}$

5. Линейная оболочка конечной системы векторов. Линейная оболочка подмножества линейного пространства. Свойства

Пусть дана система векторов $v_1, v_2, \dots, v_k$ вещественного линейного пространства V (т.е. над полем R). Множество линейных комбинаций векторов $v_1, v_2, \dots, v_k$ называется их ***линейной оболочкой*** и обозначается:

$$Lin(v_1, v_2, \dots, v_k)=\{v:\ v=\alpha_1v_1+\alpha_2v_2+\dots+\alpha_kv_k;\ \alpha_i\in R,\ i=1,\dots,k\}.$$

Векторы $v_1, v_2, \dots, v_k$ называются ***образующими линейной оболочки*** $Lin(v_1, v_2, \dots, v_k)$.

Линейная комбинация векторов $v_1, v_2, \dots, v_k$ называется ***аффинной***, если сумма ее коэффициентов равна единице. Множество аффинных комбинаций векторов $v_1, v_2, \dots, v_k$ называется их ***аффинной оболочкой*** и обозначается:

$$Aff(v_1, v_2, \dots, v_k)=\left\{v:\ v=\alpha_1v_1+\alpha_2v_2+\dots+\alpha_kv_k;\ \sum_{i=1}^{k}\alpha_i=1,\ \alpha_i\in R,\ i=1,\dots,k\right\}.$$

Линейная комбинация векторов $v_1, v_2, \dots, v_k$ называется ***неотрицательной***, если все ее коэффициенты – неотрицательные числа. Множество неотрицательных комбинаций векторов $v_1, v_2, \dots, v_k$ называется их ***конической оболочкой*** и обозначается:

$$Con(v_1, v_2, \dots, v_k)=\{v:\ v=\alpha_1v_1+\alpha_2v_2+\dots+\alpha_kv_k;\ \alpha_i\ge 0,\ i=1,\dots,k\}.$$

Линейная комбинация векторов $v_1, v_2, \dots, v_k$ называется ***выпуклой***, если все ее коэффициенты – неотрицательные числа, а их сумма равна единице. Множество выпуклых комбинаций векторов $v_1, v_2, \dots, v_k$ называется их ***выпуклой оболочкой*** и обозначается:

$$Conv(v_1, v_2, \dots, v_k)=\left\{v:\ v=\alpha_1v_1+\alpha_2v_2+\dots+\alpha_kv_k;\ \sum_{i=1}^{k}\alpha_i=1,\ \alpha_i\ge 0,\ i=1,\dots,k\right\}.$$

6. Теорема о дополнении системы векторов до базиса.

Теорема 8.2 (о дополнении системы векторов до базиса). *Всякую линейно независимую систему k векторов n-мерного линейного пространства* ($1\le k<n$) *можно дополнить до базиса пространства.*

В самом деле, пусть $e_1, e_2, \dots, e_k$ – линейно независимая система векторов n-мерного пространства V ($1\le k<n$). Рассмотрим линейную оболочку этих векторов: $L_k=Lin(e_1, e_2, \dots, e_k)$. Любой вектор $v\in L_k$ образует с векторами $e_1, e_2, \dots, e_k$ линейно зависимую систему $e_1, e_2, \dots, e_k, v$, так как вектор v линейно выражается через остальные (см. свойство 4 в разд.8.2.2). Поскольку в n-мерном пространстве существует n линейно независимых векторов, то $L_k\ne V$ и существует вектор $e_{k+1}\in V$, который не принадлежит L_k. Дополняя этим вектором линейно независимую систему $e_1, e_2, \dots, e_k$, получаем систему векторов $e_1, e_2, \dots, e_k, e_{k+1}$, которая также линейно независимая. Действительно, если бы она оказалась линейно зависимой, то из п.1 замечаний 8.3 следовало, что $e_{k+1}\in Lin(e_1, e_2, \dots, e_k)=L_k$, а это противоречит условию $e_{k+1}\notin L_k$. Итак, система векторов $e_1, e_2, \dots, e_k, e_{k+1}$ линейно независимая. Значит, первоначальную систему векторов удалось дополнить одним вектором без нарушения линейной независимости. Продолжаем аналогично. Рассмотрим линейную оболочку этих векторов: $L_{k+1}=Lin(e_1, e_2, \dots, e_k, e_{k+1})$. Если $L_{k+1}=V$, то $e_1, e_2, \dots, e_k, e_{k+1}$ – базис и теорема доказана. Если $L_{k+1}\ne V$, то дополняем систему $e_1, e_2, \dots, e_k, e_{k+1}$ вектором $e_{k+2}\notin L_{k+1}$ и т.д. Процесс дополнения обязательно закончится, так как пространство V конечномерное. В результате получим равенство $V=L_n=Lin(e_1, \dots, e_k, \dots, e_n)$, из которого следует, что $e_1, \dots, e_k, \dots, e_n$ – базис пространства V. Теорема доказана.

7. Матрица перехода от базиса к базису.

Пусть заданы два базиса пространства V: $(e)=(e_1, e_2, \dots, e_n)$ и $(e')=(e'_1, e'_2, \dots, e'_n)$. Базис (e) будем условно называть "старым", а базис (e') – "новым". Пусть известны разложения каждого вектора нового базиса по старому базису:

$$e'_i=s_{1i}\,e_1+s_{2i}\,e_2+\dots+s_{ni}\,e_n, \quad i=1,2,\dots,n. \tag{8.8}$$

Записывая по столбцам координаты векторов $(e'_1, e'_2, \dots, e'_n)$ в базисе (e), составляем матрицу:

$$S=\begin{pmatrix}s_{11}&\cdots&s_{1n}\\ \vdots&\ddots&\vdots\\ s_{n1}&\cdots&s_{nn}\end{pmatrix}. \tag{8.9}$$

Квадратная матрица S, составленная из координатных столбцов векторов нового базиса (e') в старом базисе (e), называется ***матрицей перехода*** от старого базиса к новому. При помощи матрицы перехода (8.9) формулы (8.8) можно записать в виде:

$$(e'_1 \ \cdots \ e'_n)=(e_1 \ \cdots \ e_n)\cdot S \quad \text{или, короче,} \quad (e')=(e)S. \tag{8.10}$$

Умножение символической матрицы-строки (e) на матрицу перехода S в (8.10) производится по правилам умножения матриц (см. разд. 1.3.1).

Пусть в базисе (e) вектор v имеет координаты $v_1, v_2, \dots, v_n$, а в базисе (e') – координаты $v'_1, v'_2, \dots, v'_n$, т.е.

$$v=v_1e_1+v_2e_2+\dots+v_ne_n=v'_1e'_1+v'_2e'_2+\dots+v'_ne'_n$$

или, короче, $v=(e)v=(e')v'$. Подставляя в правую часть последнего равенства выражение (8.10), получаем $v=(e)v=(e)Sv'$ – два разложения вектора v в одном и том же базисе (e). Коэффициенты этих разложений должны совпадать (по теореме 8.1), так как это координаты одного и того же вектора в одном базисе. Поэтому

$$\underset{(e)}{v}=S\underset{(e')}{v'} \quad \text{или, что то же самое,} \quad \begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\begin{pmatrix}s_{11}&\cdots&s_{1n}\\ \vdots&\ddots&\vdots\\ s_{n1}&\cdots&s_{nn}\end{pmatrix}\begin{pmatrix}v'_1\\ \vdots\\ v'_n\end{pmatrix} \tag{8.11}$$

Формула (8.11) устанавливает связь координат вектора в разных базисах: *координатный столбец вектора в старом базисе получается в результате умножения матрицы перехода на координатный столбец вектора в новом базисе.*

8. Связь координат вектора в разных базисах.

Пусть в базисе (e) вектор v имеет координаты $v_1, v_2, \dots, v_n$, а в базисе (e') – координаты $v'_1, v'_2, \dots, v'_n$, т.е.

$$v=v_1e_1+v_2e_2+\dots+v_ne_n=v'_1e'_1+v'_2e'_2+\dots+v'_ne'_n$$

или, короче, $v=(e)v=(e')v'$. Подставляя в правую часть последнего равенства выражение (8.10), получаем $v=(e)v=(e)Sv'$ – два разложения вектора v в одном и том же базисе (e). Коэффициенты этих разложений должны совпадать (по теореме 8.1), так как это координаты одного и того же вектора в одном базисе. Поэтому

$$\underset{(e)}{v}=S\underset{(e')}{v'} \quad \text{или, что то же самое,} \quad \begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\begin{pmatrix}s_{11}&\cdots&s_{1n}\\ \vdots&\ddots&\vdots\\ s_{n1}&\cdots&s_{nn}\end{pmatrix}\begin{pmatrix}v'_1\\ \vdots\\ v'_n\end{pmatrix} \tag{8.11}$$

Формула (8.11) устанавливает связь координат вектора в разных базисах: *координатный столбец вектора в старом базисе получается в результате умножения матрицы перехода на координатный столбец вектора в новом базисе.*

(См. билет № 7 для подробностей по этой теме)

Аналогично определению образующих линейной оболочки, векторы $v_1, v_2, \ldots, v_k$ называют ***образующими*** множеств $Aff(v_1, v_2, \ldots, v_k)$, $Con(v_1, v_2, \ldots, v_k)$, $Conv(v_1, v_2, \ldots, v_k)$ соответственно.

Понятия линейной, аффинной, конической и выпуклой оболочек, определенные для конечной системы векторов, можно обобщить.

Линейной оболочкой непустого подмножества M линейного пространства V ($M \subset V$, $M \neq \varnothing$) называется множество всевозможных линейных комбинаций векторов из M:

$$Lin(M) = \left\{ v : v = \sum_{i=1}^{k} \alpha_i v_i;\ k \in N,\ v_i \in M,\ \alpha_i \in R,\ i = 1, \ldots, k \right\}.$$

Аналогично определяются ***аффинная, коническая и выпуклая оболочки непустого подмножества*** M:

$$Aff(M) = \left\{ v : v = \sum_{i=1}^{k} \alpha_i v_i;\ k \in N,\ v_i \in M,\ \sum_{i=1}^{k} \alpha_i = 1,\ \alpha_i \in R,\ i = 1, \ldots, k \right\};$$

$$Con(M) = \left\{ v : v = \sum_{i=1}^{k} \alpha_i v_i;\ k \in N,\ v_i \in M,\ \alpha_i \geq 0,\ i = 1, \ldots, k \right\};$$

$$Conv(M) = \left\{ v : v = \sum_{i=1}^{k} \alpha_i v_i;\ k \in N,\ v_i \in M,\ \sum_{i=1}^{k} \alpha_i = 1,\ \alpha_i \geq 0,\ i = 1, \ldots, k \right\}.$$

Если множество M пустое ($M = \varnothing$), то по определению считается, что

$$Lin(M) = Aff(M) = Con(M) = Conv(M) = \{ o \}.$$

Из определений следуют включения:

$$M \subset Conv(M) \subset Aff(M) \subset Lin(M), \quad M \subset Conv(M) \subset Con(M) \subset Lin(M);$$

и равенство $Conv(M) = Con(M) \cap Aff(M)$.

Замечания 8.3.

1. Свойство 7 линейно зависимых и линейно независимых векторов можно сформулировать (без указания единственности разложения) следующим образом: если система векторов $v_1, v_2, \ldots, v_k$ линейно независима, а после присоединения к ней вектора v – оказывается линейно зависимой, то вектор $v \in Lin(v_1, v_2, \ldots, v_k)$.

2. Свойство 8 линейно зависимых и линейно независимых векторов можно сформулировать так: если каждый вектор системы $u_1, u_2, \ldots, u_l$ принадлежит линейной оболочке $Lin(v_1, v_2, \ldots, v_k)$ и $l > k$, то система векторов $u_1, u_2, \ldots, u_l$ – линейно зависима.

Свойства 7-8:

7. *Если система векторов $v_1, v_2, \ldots, v_k$ – линейно независима, а после присоединения к ней вектора v – оказывается линейно зависимой, то вектор v можно разложить по векторам $v_1, v_2, \ldots, v_k$ и притом единственным образом, т.е. коэффициенты разложения* (8.1) *находятся однозначно.*

8. *Пусть каждый вектор системы $u_1, u_2, \ldots, u_l$ может быть разложен по векторам системы $v_1, v_2, \ldots, v_k$*, т.е. $u_i = \sum_{j=1}^{k} a_{ji} v_j$, $i = 1, \ldots, l$ (говорят, что ***система векторов*** $u_1, u_2, \ldots, u_l$ ***линейно выражается через систему векторов*** $v_1, v_2, \ldots, v_k$). *Тогда, если $l > k$, то система векторов $u_1, u_2, \ldots, u_l$ – линейно зависима.*

Свойства матрицы перехода .

Пример 8.3. В пространстве $P_2(R)$ многочленов степени не выше второй даны две системы многочленов:

8.4.4. Свойства матрицы перехода от одного базиса к другому

1. *Пусть имеются три базиса $(e), (f), (g)$ пространства V и известны матрицы перехода: $\underset{(e)\to(f)}{S}$ от базиса (e) к базису (f); $\underset{(f)\to(g)}{S}$ от (f) к (g); $\underset{(e)\to(g)}{S}$ от (e) к (g). Тогда*

$$\underset{(e)\to(g)}{S} = \underset{(e)\to(f)}{S} \cdot \underset{(f)\to(g)}{S}. \qquad (8.12)$$

Действительно, запишем связь (8.10) для данных базисов:

$$(f) = (e) \cdot \underset{(e)\to(f)}{S}; \quad (g) = (f) \cdot \underset{(f)\to(g)}{S}; \quad (g) = (e) \cdot \underset{(e)\to(g)}{S}.$$

Подставляя первое выражение во второе равенство, получаем $(g) = (e) \cdot \underset{(e)\to(f)}{S} \cdot \underset{(f)\to(g)}{S}$. Сравнивая с третьим равенством, приходим к (8.12).

2. *Если S – матрица перехода от базиса (e) к базису (f), то матрица S обратима и обратная матрица S^{-1} является матрицей перехода от базиса (f) к базису (e). Координаты вектора v в базисах (e) и (f) связаны формулами:*

$$\underset{(e)}{v} = S \underset{(f)}{v}, \qquad \underset{(f)}{v} = S^{-1} \underset{(e)}{v}.$$

В самом деле, пусть T – матрица перехода от базиса (f) к базису (e). Учитывая, что матрица перехода от базиса (e) к базису (e) – единичная, применяем свойство 1 к трем базисам $(e), (f), (e)$: $E = ST$. Для трех базисов $(f), (e), (f)$ аналогично получаем: $E = TS$. Следовательно, $T = S^{-1}$ (см. разд. 4.1).

3. *Всякая обратимая квадратная матрица n-го порядка может служить матрицей перехода от одного базиса n-мерного линейного пространства к другому базису.*

Пример 8.4. В двумерном арифметическом пространстве R^2 даны два базиса: $f_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, $f_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ и $g_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $g_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Найти матрицу $\underset{(f)\to(g)}{S}$ перехода от базиса (f) к базису (g) и координаты вектора $v = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ в каждом из базисов.

□ Рассмотрим стандартный базис $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ пространства R^2 (см. п.3 в разд. 8.3.2). Находим координаты векторов f_1, f_2, g_1, g_2 в стандартном базисе. Раскладываем вектор f_1:

$$e_1 = 1, e_2 = x, e_3 = x^2 \quad \text{и} \quad f_1 = (x+1)^2, f_2 = (x-1)^2, f_3 = x^2.$$

Доказать, что каждая система является базисом пространства $P_2(R)$. Найти матрицу S перехода от базиса (e) к базису (f). Определить координаты квадратного трехчлена $p = x^2 - x + 1$ относительно базисов (e) и (f).

□ Система многочленов $e_1 = 1, e_2 = x, e_3 = x^2$ является стандартным базисом пространства $P_2(R)$ (см. п.6 в примерах базисов в разд. 8.3.2). Докажем, что система $f_1 = (x+1)^2, f_2 = (x-1)^2, f_3 = x^2$ является базисом. Поступим следующим образом. Найдем координатные столбцы f_1, f_2, f_3 этих многочленов в стандартном базисе. Раскладывая по базису (e), получаем

$$f_1 = (x+1)^2 = x^2 + 2x + 1 = 1 \cdot e_1 + 2 \cdot e_2 + 1 \cdot e_3, \text{ т.е. } f_1 = (1 \ \ 2 \ \ 1)^T;$$

$$f_2 = (x-1)^2 = x^2 - 2x + 1 = 1 \cdot e_1 - 2 \cdot e_2 + 1 \cdot e_3, \text{ т.е. } f_2 = (1 \ \ -2 \ \ 1)^T;$$

$$f_3 = x^2 = 0 \cdot e_1 + 0 \cdot e_2 + 1 \cdot e_3, \text{ т.е. } f_3 = (0 \ \ 0 \ \ 1)^T.$$

Составим из этих столбцов матрицу $S = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Ранг этой матрицы равен 3, так как $\det S = -4 \neq 0$. Следовательно, столбцы f_1, f_2, f_3 линейно независимы, тогда и многочлены f_1, f_2, f_3 линейно независимы (см. п.2 замечаний 8.5). Итак, многочлены f_1, f_2, f_3 являются базисом пространства $P_2(R)$, а матрица S – искомая матрица перехода от базиса (e) к базису (f). Осталось найти координаты многочлена $p = x^2 - x + 1$ в этих базисах. Раскладывая p по базисам, находим

$$p = x^2 - x + 1 = 1 \cdot e_1 - 1 \cdot e_2 + 1 \cdot e_3, \text{ т.е. } \underset{(e)}{p} = (1 \ \ -1 \ \ 1)^T;$$

$$p = x^2 - x + 1 = \frac{(x+1)^2 + 3(x-1)^2}{4} = \frac{1}{4} \cdot f_1 + \frac{3}{4} \cdot f_2 + 0 \cdot f_3, \text{ т.е. } \underset{(f)}{p} = \left(\tfrac{1}{4} \ \ \tfrac{3}{4} \ \ 0\right)^T.$$

Проверим результат, вычисляя $\underset{(e)}{p}$ по формуле (8.11):

$$\underset{(e)}{p} = S \underset{(f)}{p} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ \frac{3}{4} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Результаты совпадают. ■

$$f_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3 \cdot e_1 + 2 \cdot e_2, \text{ т.е. } f_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

В стандартном базисе (e) пространства R^2 координатный столбец f_1 совпадает с вектором f_1. Для других векторов аналогично получаем $f_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $g_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $g_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Из координатных столбцов составим матрицы перехода (8.9) от стандартного базиса (e) к данным базисам (f) и (g):

$$\underset{(e)\to(f)}{S} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}, \quad \underset{(e)\to(g)}{S} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

По свойству 1 матриц перехода имеем $\underset{(f)\to(g)}{S} = \underset{(f)\to(e)}{S} \cdot \underset{(e)\to(g)}{S}$. По свойству 2: $\underset{(f)\to(e)}{S} = \underset{(e)\to(f)}{S^{-1}}$. Поэтому

$$\underset{(f)\to(g)}{S} = \underset{(e)\to(f)}{S^{-1}} \cdot \underset{(e)\to(g)}{S} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 4 & 5 \end{pmatrix}.$$

В стандартном базисе (e) пространства R^2 координатный столбец $\underset{(e)}{v} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ совпадает с вектором v. Найдем координаты этого вектора в базисе (f) (по свойству 2 матрицы перехода):

$$\underset{(f)}{v} = \underset{(e)\to(f)}{S^{-1}} \underset{(e)}{v} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} -3 \\ 15 \end{pmatrix}.$$

В самом деле, справедливо разложение

$$v = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = -3 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 15 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -3 \cdot f_1 + 15 \cdot f_2.$$

Найдем координаты вектора v в базисе (g) двумя способами:

$$\underset{(g)}{v} = \underset{(f)\to(g)}{S^{-1}} \underset{(f)}{v} = \begin{pmatrix} -1 & -2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} -3 \\ 15 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 5 & 2 \\ -4 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 15 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix};$$

$$\underset{(g)}{v} = \underset{(e)\to(g)}{S^{-1}} \underset{(e)}{v} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Полученный результат подтверждает разложение:

$$v = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-1) \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 5 \cdot g_1 + (-1) \cdot g_2. \ ■$$

Связь между координатами вектора в разных базисах дает следующая теорема:

Пусть $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_n$ два базиса линейного пространства L. Причем имеют место равенства:

$$\begin{aligned} f_1 &= \tau_{11} e_1 + \tau_{21} e_2 + \ldots + \tau_{n1} e_n, \\ f_2 &= \tau_{12} e_1 + \tau_{22} e_2 + \ldots + \tau_{n2} e_n, \\ &\ldots \\ f_n &= \tau_{1n} e_1 + \tau_{2n} e_2 + \ldots + \tau_{nn} e_n. \end{aligned}$$

Если вектор a имеет в базисе $e_1, e_2, \ldots, e_n$ координаты $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, а в базисе $f_1, f_2, \ldots, f_n$ – координаты $\{\beta_1, \beta_2, \ldots, \beta_n\}$, то справедливо равенство $\mathbf{A} = \mathbf{TB}$,

где $\mathbf{A} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}$, $\mathbf{T} = \begin{pmatrix} \tau_{11} & \tau_{12} & \ldots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \ldots & \tau_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \tau_{n1} & \tau_{n2} & \ldots & \tau_{nn} \end{pmatrix}$ *(матрицу* $\mathbf{T}$ *называют матрицей перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $f_1, f_2, \ldots, f_n$).*

ДОКАЗАТЕЛЬСТВО

По условию $a = \beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_n f_n$.

Расписывая векторы $f_1, f_2, \ldots, f_n$ по базису $e_1, e_2, \ldots, e_n$, получим:

$$a = \beta_1(\tau_{11} e_1 + \tau_{21} e_2 + \ldots + \tau_{n1} e_n) + \beta_2(\tau_{12} e_1 + \tau_{22} e_2 + \ldots + \tau_{n2} e_n) + \ldots + \beta_n(\tau_{1n} e_1 + \tau_{2n} e_2 + \ldots + \tau_{nn} e_n).$$

Раскроем скобки и перегруппируем слагаемые:

$$\begin{aligned} a = &(\beta_1 \tau_{11} + \beta_2 \tau_{12} + \ldots + \beta_n \tau_{1n}) e_1 + \\ &+ (\beta_1 \tau_{21} + \beta_2 \tau_{22} + \ldots + \beta_n \tau_{2n}) e_2 + \\ &+ \ldots\ldots\ldots + \\ &+ (\beta_1 \tau_{n1} + \beta_2 \tau_{n2} + \ldots + \beta_n \tau_{nn}) e_n. \end{aligned} \qquad (1)$$

Так как по условию $a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$, то из (1) получаем:

$$\begin{aligned} \alpha_1 &= \beta_1 \tau_{11} + \beta_2 \tau_{12} + \ldots + \beta_n \tau_{1n}, \\ \alpha_2 &= \beta_1 \tau_{21} + \beta_2 \tau_{22} + \ldots + \beta_n \tau_{2n}, \\ &\ldots\ldots\ldots \\ \alpha_n &= \beta_1 \tau_{n1} + \beta_2 \tau_{n2} + \ldots + \beta_n \tau_{nn}, \end{aligned}$$

или в матричном виде $\mathbf{A} = \mathbf{TB}$.

9. Теорема об изоморфизме конечномерных линейных пространств.

Преамбула:

Говорят, что между элементами двух множеств U и V установлено ***взаимно однозначное соответствие***, если указано правило, которое каждому элементу $u \in U$ сопоставляет один и только один элемент $v \in V$, причем каждый элемент $v \in V$ оказывается сопоставленным одному и только одному элементу $u \in U$. Взаимно однозначное соответствие будем обозначать $U \leftrightarrow V$, а соответствующие элементы: $u \leftrightarrow v$.

Два линейных пространства U и V называются ***изоморфными***, если между их элементами можно установить такое взаимно однозначное соответствие, что выполняются условия:

1) сумме векторов пространства U соответствует сумма соответствующих векторов пространства V:

$$\left.\begin{aligned} u_1 \leftrightarrow v_1 \\ u_2 \leftrightarrow v_2 \end{aligned}\right\} \Rightarrow (u_1 + u_2) \leftrightarrow (v_1 + v_2);$$

2) произведению числа на вектор пространства U соответствует произведение того же числа на соответствующий вектор пространства V:

$$u \leftrightarrow v \Rightarrow \lambda u \leftrightarrow \lambda v.$$

Другими словами, ***изоморфизм*** – это взаимно однозначное соответствие, сохраняющее линейные операции.

Замечания 8.6:

10. Подпространства линейного пространства. Примеры

8.6.1. Определение линейного подпространства

Непустое подмножество L линейного пространства V называется ***линейным подпространством*** пространства V, если

1) $u + v \in L \quad \forall u, v \in L$ (подпространство *замкнуто по отношению к операции сложения*);

2) $\lambda v \in L \quad \forall v \in L$ и любого числа λ (подпространство *замкнуто по отношению к операции умножения вектора на число*).

Для указания линейного подпространства будем использовать обозначение $L \triangleleft V$, а слово "линейное" опускать для краткости.

Замечания 8.7.

1. Условия 1, 2 в определении можно заменить одним условием: $\lambda u + \mu v \in L \quad \forall u, v \in L$ и любых чисел λ и μ. Разумеется, что здесь и в определении речь идет о произвольных числах из того числового поля, над которым определено пространство V (см. разд.8.1.1).

2. В любом линейном пространстве V имеются два линейных подпространства:

а) само пространство V, т.е. $V \triangleleft V$;

б) нулевое подпространство $\{ o \}$, состоящее из одного нулевого вектора пространства V, т.е. $\{ o \} \triangleleft V$. Эти подпространства называются ***несобственными***, а все остальные – ***собственными***.

3. Любое подпространство L линейного пространства V является его подмножеством: $L \triangleleft V \Rightarrow L \subset V$, но не всякое подмножество $M \subset V$ является линейным подпространством, так как оно может оказаться незамкнутым по отношению к линейным операциям.

4. Подпространство L линейного пространства V само является линейным пространством с теми же операциями сложения векторов и умножения вектора на число, что и в пространстве V, поскольку для них выпол-

11. Пересечение и алгебраическая сумма подпространств. Примеры.

Пусть L_1 и L_2 – подпространства линейного пространства V.

Пересечением подпространств L_1 и L_2 называется множество $L_1 \cap L_2$ векторов, каждый из которых принадлежит L_1 и L_2 одновременно, т.е. пересечение подпространств определяется как обычное пересечение двух множеств.

Алгебраической суммой подпространств L_1 и L_2 называется множество векторов вида $v_1 + v_2$, где $v_1 \in L_1$, $v_2 \in L_2$. Алгебраическая сумма (короче просто сумма) подпространств обозначается $L_1 + L_2$:

$$L_1 + L_2 = \{ v \in V : v = v_1 + v_2, v_1 \in L_1, v_2 \in L_2 \}.$$

Представление вектора $v \in L_1 + L_2$ в виде $v = v_1 + v_2$, где $v_1 \in L_1$, $v_2 \subset L_2$, называется ***разложением вектора*** v ***по подпространствам*** L_1 и L_2.

Пример 8.6. В пространстве V_3 радиус-векторов с общим началом в точке O заданы подпространства: L_0, L_1 и L_2 – три множества радиус-векторов, принадлежащих пересекающимся в точке O прямым l_0, l_1 и l_2 соответственно; Π_1 и Π_2 – два множества радиус-векторов, принадлежащих пересекающимся плоскостям π_1 и π_2 соответственно; прямая l_1 принадлежит плоскости π_1, прямая l_2 принадлежит плоскости π_2, плоскости π_1 и π_2 пересекаются по прямой l_0 (рис. 8.2). Найти суммы и пересечения каждых двух из указанных пяти подпространств.

□ Найдем сумму $L_0 + L_1$. Складывая два вектора, принадлежащих L_0 и L_1 соответственно, получаем вектор, принадлежащий плоскости Π_1. Наоборот, любой вектор $\bar{v}$ (см. рис.8.2), принадлежащий Π_1, можно представить в виде $\bar{v}_0 + \bar{v}_1$, построив проекции $\bar{v}_0$ и $\bar{v}_1$ вектора $\bar{v}$ на прямые l_0 и l_1

12. Прямая сумма, алгебраическое дополнение.

1. При изоморфизме линейных пространств U и V:

– их нулевые элементы соответствуют друг другу ($o_U \leftrightarrow o_V$);

- противоположные элементы соответствуют друг другу.

Это следует из определения, если в условии 2 положить $\lambda = 0$ или $\lambda = -1$.

2. Линейной комбинации векторов пространства U соответствует линейная комбинация соответствующих векторов пространства V.

3. Линейно независимой (линейно зависимой) системе векторов пространства U соответствует линейно независимая (линейно зависимая) система векторов пространства V. Действительно, из п.1,2 следует, что равенства $\lambda_1 u_1 + \ldots + \lambda_n u_n = o_U$ и $\lambda_1 v_1 + \ldots + \lambda_n v_n = o_V$ равносильны. Если не все коэффициенты $\lambda_1, \ldots, \lambda_n$ равны нулю, то обе системы $u_1, u_2, \ldots, u_n$ и $v_1, v_2, \ldots, v_n$ линейно зависимы, в противном случае, обе системы линейно независимы.

4. Любое n-мерное линейное вещественное пространство V изоморфно n-мерному арифметическому пространству R^n, а n-мерное комплексное пространство изоморфно C^n

Это следует из п.4 замечаний 8.5, где установлено взаимно однозначное соответствие $V \leftrightarrow R^n$ между векторами и координатными столбцами. Линейные операции с векторами в координатной форме (см. разд.8.4.2) показывают, что это взаимно однозначное соответствие является изоморфизмом.

5. Если пространство U изоморфно пространству V, а V изоморфно пространству W, то пространства U и W также изоморфны.

В самом деле, имея взаимно однозначные соответствия $U \leftrightarrow V$ и $V \leftrightarrow W$, поставим в соответствие вектору u такой вектор w, что $u \leftrightarrow v \leftrightarrow w$. Такое "сквозное" соответствие $U \leftrightarrow W$ будет взаимно однозначным, сохраняющим линейные операции.

Теорема 8.3 (об изоморфизме линейных пространств). *Два конечномерных линейных пространства (над одним и тем же числовым полем) изоморфны тогда и только тогда, когда они имеют одну и ту же размерность.*

Действительно, если пространства изоморфны ($U \leftrightarrow V$), то базису $u_1, u_2, \ldots, u_n$ пространства U соответствует линейно независимая система векторов $v_1, v_2, \ldots, v_n$ пространства V (см. п.3 замечаний 8.6), которую в случае необходимости можно дополнить до базиса пространства V (см. теорему 8.2). Следовательно, $\dim U \le \dim V$. Аналогично получаем противоположное неравенство $\dim U \ge \dim V$. Таким образом, $\dim U = \dim V$ (необходимость доказана). Достаточность следует из п.4, 5 замечаний 8.6. Действительно, пусть U и V определены над полем R и $\dim U = \dim V = n$. Тогда, выбрав любые базисы в пространствах U и V, установим изоморфизмы $U \leftrightarrow R^n$ и $V \leftrightarrow R^n$, если U и V – вещественные пространства. Если пространства U и V определены над полем C комплексных чисел, то $U \leftrightarrow C^n$ и $V \leftrightarrow C^n$. В обоих случаях, согласно п.5 замечаний 8.6, пространства U и V изоморфны. Теорема доказана.

Следствие. *Изучение конечномерных линейных пространств сводится к изучению арифметических пространств той же размерности.*

нали: $a_{ij} = a_{ji} = 1$, $i = 1, \ldots, n$, $j = i, i+1, \ldots, n$. Всего в базисе будет $n + (n-1) + \ldots + 2 + 1 = \frac{n(n+1)}{2}$ матриц. Следовательно, $\dim M^{\text{сим}}_{n\times n} = \frac{n(n+1)}{2}$.

Аналогично получаем, что $M^{\text{кос}}_{n\times n} \lhd M_{n\times n}$ и $\dim M^{\text{кос}}_{n\times n} = \frac{n(n-1)}{2}$.

Множество вырожденных квадратных матриц n-го порядка не является подпространством $M_{n\times n}$, так как сумма двух вырожденных матриц может оказаться невырожденной матрицей, например, в пространстве $M_{2\times 2}$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

няются аксиомы 1–8 (см. разд.8.1). Поэтому можно говорить о размерности подпространства, его базисе и т.п.

5. *Размерность любого подпространства L линейного пространства V не превосходит размерности пространства V: $\dim L \le \dim V$. Если же размерность подпространства $L \lhd V$ равна размерности конечномерного пространства V ($\dim L = \dim V$), то подпространство совпадает с самим пространством: $L = V$.*

Это следует из теоремы 8.2 (о дополнении системы векторов до базиса). Действительно, взяв базис подпространства L, будем дополнять его до базиса пространства V. Если это возможно, то $\dim L < \dim V$. Если нельзя дополнить, т.е. базис подпространства L является базисом пространства V, то $\dim L = \dim V$. Учитывая, что пространство есть линейная оболочка базиса (см. следствие 1 теоремы 8.1), получаем $L = V$.

6. *Для любого подмножества M линейного пространства V линейная оболочка $Lin(M)$ является подпространством V и $M \subset Lin(M) \lhd V$.*

В самом деле, если $M = \varnothing$ (пустое множество), то по определению $Lin(M) = \{o\}$, т.е. является нулевым подпространством и $\varnothing \subset \{o\} \lhd V$. Пусть $M \ne \varnothing$. Нужно доказать, что множество $Lin(M)$ замкнуто по отношению к операциям сложения его элементов и умножения его элементов на число. Напомним, что элементами линейной оболочки $Lin(M)$ служат линейные комбинации векторов из M. Так как линейная комбинация линейных комбинаций векторов является их линейной комбинацией, то, учитывая п.1, делаем вывод, что $Lin(M)$ является подпространством V, т.е. $Lin(M) \lhd V$. Включение $M \subset Lin(M)$ – очевидное, так как любой вектор $v \in M$ можно представить как линейную комбинацию $1 \cdot v$, т.е. как элемент множества $Lin(M)$.

7. *Линейная оболочка $Lin(L)$ подпространства $L \lhd V$ совпадает с подпространством L, т.е. $Lin(L) = L$.*

Действительно, так как линейное подпространство L содержит всевозможные линейные комбинации своих векторов, то $Lin(L) \subset L$. Противоположное включение ($L \subset Lin(L)$) следует из п.6. Значит, $Lin(L) = L$.

8.6.2. Примеры линейных подпространств

Укажем некоторые подпространства линейных пространств, примеры которых рассматривались в разд. 8.1.3. Перечислить все подпространства линейного пространства невозможно, за исключением тривиальных случаев.

1. Пространство $\{o\}$, состоящее из одного нулевого вектора пространства V, является подпространством, т.е. $\{o\} \lhd V$.

2. Пусть, как и ранее, V_1, V_2, V_3 – множества векторов (направленных отрезков) на прямой, на плоскости, в пространстве соответственно. Если прямая принадлежит плоскости, то $V_1 \lhd V_2 \lhd V_3$. Напротив, множество единичных векторов не является линейным подпространством, так как при умножении вектора на число, не равное единице, получаем вектор, не принадлежащий множеству.

3. В n-мерном арифметическом пространстве R^n рассмотрим множество L "полунулевых" столбцов вида $x = (x_1 \ \cdots \ x_m \ 0 \ \cdots \ 0)^T$ с последними $(n-m)$ элементами, равными нулю. Сумма "полунулевых" столбцов является столбцом того же вида, т.е. операция сложения замкнута в L. Умножение "полунулевого" столбца на число дает "полунулевой" столбец, т.е. операция умножения на число замкнута в L. Поэтому $L \lhd R^n$, причем $\dim L = m$. Напротив, подмножество ненулевых столбцов R^n не является линейным подпространством, так как при умножении на нуль получается нулевой столбец, который не принадлежит рассматриваемому множеству. Примеры других подпространств R^n приводятся в следующем пункте.

4. Пространство $\{Ax = o\}$ решений однородной системы уравнений с n неизвестными является подпространством n-мерного арифметического пространства R^n. Размерность этого подпространства определяется матрицей системы:

$$\dim\{Ax = o\} = n - \operatorname{rg} A.$$

Множество $\{Ax = b\}$ решений неоднородной системы (при $b \ne o$) не является подпространством R^n, так как сумма двух решений неоднородной системы не будет решением той же системы.

5. В пространстве $M_{n\times n}$ квадратных матриц порядка n рассмотрим два подмножества: множество $M^{\text{сим}}_{n\times n}$ симметрических матриц и множество $M^{\text{кос}}_{n\times n}$ кососимметрических матриц (см. разд.1.4.1). Сумма симметрических матриц является симметрической матрицей, т.е. операция сложения замкнута в $M^{\text{сим}}_{n\times n}$. Умножение симметрической матрицы на число также не нарушает симметричность, т.е. операция умножения матрицы на число замкнута в $M^{\text{сим}}_{n\times n}$. Следовательно, множество симметрических матриц является подпространством пространства квадратных матриц, т.е $M^{\text{сим}}_{n\times n} \lhd M_{n\times n}$. Нетрудно найти размерность этого подпространства. Стандартный базис образуют n матриц с единственным ненулевым (равным единице) элементом на главной диагонали: $a_{ii} = 1$ $i = 1, \ldots, n$, а также матрицы с двумя ненулевыми (равными единице) элементами, симметричными относительно главной диаго-

...

соответственно. Значит, любой радиус-вектор плоскости Π_1 раскладывается по подпространствам L_0 и L_1, т.е. $L_0 + L_1 = \Pi_1$. Аналогично получаем, что $L_0 + L_2 = \Pi_2$, а $L_1 + L_2$ – множество радиус-векторов, принадлежащих плоскости, проходящей через прямые l_1 и l_2.

Найдем сумму $\Pi_1 + L_2$. Любой вектор $\overline{w}$ пространства V_3 можно разложить по подпространствам L_2 и Π_1. В самом деле, через конец радиус-вектора $\overline{w}$ проводим прямую, параллельную прямой l_2 (см. рис. 8.2), т.е. строим проекцию $\overline{u}$ вектора $\overline{w}$ на плоскость Π_1. Затем на L_2 откладываем вектор $\overline{v}_2$ так, чтобы $\overline{w} = \overline{u} + \overline{v}_2$. Следовательно, $\Pi_1 + L_2 = V_3$. Так как $L_2 \lhd \Pi_2$, то $\Pi_1 + \Pi_2 = V_3$. Аналогично получаем, что $L_1 + \Pi_2 = V_3$. Остальные суммы находятся просто: $L_0 + \Pi_1 = L_1 + \Pi_1 = \Pi_1$, $L_0 + \Pi_2 = L_2 + \Pi_2 = \Pi_2$. Заметим, что $L_0 + L_1 + L_2 = V_3$.

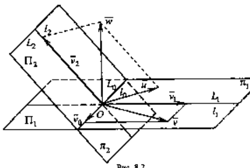

Рис. 8.2

Используя теорему 8.4, проверим, например, равенство $\Pi_1 + \Pi_2 = V_3$ по размерности. Подставляя $\dim \Pi_1 = \dim \Pi_2 = 2$ и $\dim(\Pi_1 \cap \Pi_2) = \dim L_0 = 1$ в формулу Грассмана, получаем $\dim(\Pi_1 + \Pi_2) = 2 + 2 - 1 = 3$, что и следовало ожидать, так как $\dim(\Pi_1 + \Pi_2) = \dim V_3 = 3$.

Пересечения подпространств находим по рис. 8.2, как пересечение геометрических фигур:

$$L_0 \cap L_1 = L_0 \cap L_2 = L_1 \cap L_2 = L_1 \cap \Pi_2 = L_2 \cap \Pi_1 = \{\overline{o}\},$$

$$L_0 \cap \Pi_1 = L_0,\ L_0 \cap \Pi_2 = L_0,\ L_1 \cap \Pi_1 = L_1,\ L_2 \cap \Pi_2 = L_2,\ \Pi_1 \cap \Pi_2 = L_0,$$

где $\overline{o}$ – нулевой радиус-вектор $\overline{OO}$. ■

Примечание:

Теорема 8.4 (о размерности суммы подпространств). *Если L_1 и L_2 подпространства конечномерного линейного пространства V, то размерность суммы подпространств равна сумме их размерностей без размерности их пересечения* (***формула Грассмана***):

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2). \qquad (8.13)$$

Алгебраическая сумма подпространств L_1 и L_2 линейного пространства V называется ***прямой суммой***, если пересечение подпространств состоит из одного нулевого вектора. Прямая сумма подпространств обозначается $L_1 \oplus L_2$ и обладает следующим свойством: *если $V = L_1 \oplus L_2$, то для каждого вектора $v \in V$ существует единственное представление в виде $v = v_1 + v_2$, где $v_1 \in L_1$, $v_2 \in L_2$.*

Действительно, если предположить противное, а именно существование двух разных разложений: $v = v_1 + v_2 = w_1 + w_2$, где $v_1, w_1 \in L_1$, $v_2, w_2 \in L_2$, $v_1 \ne w_1$, то получим противоречие: из равенства $v_1 - w_1 = w_2 - v_2$ следует, что ненулевой вектор $v_1 - w_1 \ne o$ принадлежит обоим подпространствам L_1 и L_2 одновременно, значит, принадлежит их пересечению, а по определению их пересечение состоит из одного нулевого вектора.

ПРИЗНАКИ ПРЯМЫХ СУММ ПОДПРОСТРАНСТВ

Сумма $V = L_1 + L_2$ является прямой суммой, если:

– существует вектор $v \in V$, который однозначно представляется в виде $v = v_1 + v_2$, где $v_1 \in L_1$, $v_2 \in L_2$;

– базис пространства V является объединением базисов подпространств L_1 и L_2;

– справедливо равенство $\dim V = \dim L_1 + \dim L_2$.

Замечания 8.9.

1. Понятие прямой суммы распространяется на любое конечное число слагаемых. Сумма $V = L_1 \oplus L_2 \oplus \ldots \oplus L_k$ называется ***прямой суммой подпространств***, если пересечение каждого из них с суммой остальных равно одному нулевому вектору: $L_i \cap (L_1 + \ldots + L_{i-1} + L_{i+1} + \ldots + L_k) = \{o\}$.

2. Свойства и признаки, указанные для прямой суммы двух подпространств, справедливы и для любого конечного числа слагаемых. Отметим еще одно свойство: если $e_1, e_2, \ldots, e_n$ – базис пространства V, то $V = Lin(e_1) \oplus \ldots \oplus Lin(e_n)$.

Пример 8.7. В примере 8.6 найдены алгебраические суммы подпространств. Какие суммы являются прямыми?

□ Так как $L_0 \cap L_1 = \overline{o}$, то сумма $L_0 \oplus L_1$ – прямая. Аналогично получаем, что суммы $L_0 \oplus L_2$, $L_1 \oplus L_2$, $\Pi_1 \oplus L_2$, $L_1 \oplus \Pi_2$, $L_0 \oplus L_1 \oplus L_2$ – прямые. Остальные суммы подпространств, найденные в примере 8.6, не являются прямыми: $L_0 + \Pi_1$, $L_0 + \Pi_2$, $L_1 + \Pi_1$, $L_2 + \Pi_2$, $\Pi_1 + \Pi_2$, поскольку их пересечение содержит не только нулевой вектор. Например, пересечение $\Pi_1 \cap \Pi_2 = L_0 \ne \{\overline{o}\}$. ■

АЛГЕБРАИЧЕСКИЕ ДОПОЛНЕНИЯ ПОДПРОСТРАНСТВ

Пусть L – подпространство конечномерного линейного пространства V. Подпространство $L^+ \lhd V$ называется ***алгебраическим дополнением*** подпространства L в пространстве V, если $V = L \oplus L^+$. Говорят, что L^+ дополняет (алгебраически) подпространство L до V.

Рассмотрим *свойства алгебраических дополнений.*

1. *Для любого подпространства $L \lhd V$ существует алгебраическое дополнение $L^+ \lhd V$.*

Действительно, если $L = \{o\}$, то $L^+ = V$. Если $L = V$, то $L^+ = \{o\}$. В остальных случаях базис $e_1, e_2, \ldots, e_s$ подпространства L можно дополнить по теореме 8.2 до базиса $e_1, \ldots, e_s, e_{s+1}, \ldots, e_n$ пространства V. Тогда $L^+ = Lin(e_{s+1}, \ldots, e_n)$. В примере 8.7 получено равенство $V_3 = L_1 \oplus \Pi_2$, т.е. подпространства L_1 и Π_2 дополняют друг друга до всего пространства.

2. *Базис любого подпространства $L \lhd V$ дополняется базисом алгебраического дополнения $L^+ \lhd V$ до базиса всего пространства.*

3. *Алгебраическое дополнение L^+ подпространства $L \lhd V$, кроме случаев $L = \{o\}$ или $L = V$, определяется неоднозначно.*

В примере 8.7 дополнением плоскости Π_2 в пространстве V_3 служит множество радиус-векторов, принадлежащих любой прямой, пересекающей плоскость Π_2 в точке O, в частности, подпространство L_1.

4. *Для любого подпространства $L \lhd V$: $L^+ \oplus (L^+)^+ = V$.*

Это равенство следует непосредственно из определения. Заметим, что равенство $L = (L^+)^+$ в силу неоднозначности определения алгебраического дополнения, вообще говоря, не справедливо.

5. *Если L_1 и L_2 – подпространства пространства V, то пересечение их алгебраических дополнений является алгебраическим дополнением суммы подпространств, и, наоборот, сумма алгебраических дополнений является алгебраическим дополнением пересечения подпространств:*

$$(L_1 + L_2) \oplus (L_1^+ \cap L_2^+) = V \quad \text{и} \quad (L_1 \cap L_2) \oplus (L_1^+ + L_2^+) = V. \qquad (8.15)$$

Заметим, что равенства $(L_1 + L_2)^+ = L_1^+ \cap L_2^+$ и $(L_1 \cap L_2)^+ = L_1^+ + L_2^+$ в силу неоднозначности определения алгебраического дополнения, вообще говоря, не справедливы.

Докажем последнее свойство. Как при доказательстве теоремы 8.4 построим базис суммы подпространств из трех наборов $(e), (e'), (e'')$ векторов: $L_1 + L_2 = Lin((e), (e'), (e''))$. Дополним теперь этот базис (по теореме 8.2) векторами $(f) = (f_1, \ldots, f_t)$ до базиса $(e), (e'), (e''), (f)$ пространства V. Так как $(e), (e')$ базис L_1, то по свойству 2 алгебраических дополнений заключаем, что $(e''), (f)$ – базис L_1^+. Аналогично получаем, что $(e'), (f)$ – базис L_2^+. Следовательно, (f) – базис пересечения $L_1^+ \cap L_2^+$. Таким образом, базис всего пространства V получается объединением базиса суммы $L_1 + L_2$ и базиса пересечения $L_1^+ \cap L_2^+$: $\underbrace{(e)(e')(e'')}_{L_1 + L_2}\ \underbrace{(f)}_{L_1^+ \cap L_2^+}$. Используя признак 2 прямой суммы подпространств, получаем $(L_1 + L_2) \oplus (L_1^+ \cap L_2^+) = V$. Равенство $(L_1 \cap L_2) \oplus (L_1^+ + L_2^+) = V$ следует аналогично из структуры $\underbrace{(e)}_{L_1 \cap L_2}\ \underbrace{(e')(e'')(f)}_{L_1^+ + L_2^+}$ базиса пространства V.

13. Теорема о размерности суммы подпространств.

Теорема 8.4 (о размерности суммы подпространств). *Если L_1 и L_2 подпространства конечномерного линейного пространства V, то размерность суммы подпространств равна сумме их размерностей без размерности их пересечения* (***формула Грассмана***).

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2). \qquad (8.13)$$

В самом деле, пусть $(e) = (e_1, \ldots, e_s)$ – базис пересечения $L_1 \cap L_2$ ($\dim(L_1 \cap L_2) = s$). Дополним его упорядоченным набором $(e') = (e'_{s+1}, \ldots, e'_{m_1})$ векторов до базиса $(e), (e')$ подпространства L_1 ($\dim L_1 = m_1$) и упорядоченным набором $(e'') = (e''_{s+1}, \ldots, e''_{m_2})$ векторов до базиса $(e), (e'')$ подпространства L_2 ($\dim L_2 = m_2$). Такое дополнение возможно по теореме 8.2. Из указанных трех наборов векторов составим упорядоченный набор $(e), (e'), (e'')$ векторов. Покажем, что эти векторы являются образующими пространства $L_1 + L_2$. Действительно, любой вектор v этого пространства представляется в виде линейной комбинации векторов из упорядоченного набора $(e), (e'), (e'')$:

$$v = v_1 + v_2 = \underbrace{\sum_{i=1}^{s} \alpha_i e_i + \sum_{i=s+1}^{m_1} \alpha'_i e'_i}_{v_1} + \underbrace{\sum_{i=1}^{s} \beta_i e_i + \sum_{i=s+1}^{m_2} \beta''_i e''_i}_{v_2}.$$

Следовательно, $L_1 + L_2 = Lin((e), (e'), (e''))$. Докажем, что образующие $(e), (e'), (e'')$ линейно независимы и поэтому они являются базисом пространства $L_1 + L_2$. Действительно, составим линейную комбинацию этих векторов и приравняем ее нулевому вектору:

$$\underbrace{\sum_{i=1}^{s} \alpha_i e_i + \sum_{i=s+1}^{m_1} \beta_i e'_i}_{w_1} + \underbrace{\sum_{i=s+1}^{m_2} \gamma_i e''_i}_{w_2} = o. \qquad (8.14)$$

14. Евклидово пространство: определение, примеры, простейшие следствия из аксиом.

8.8.1. Определение евклидова пространства

Вещественное линейное пространство E называется ***евклидовым***, если каждой паре элементов u, v этого пространства поставлено в соответствие действительное число (u, v), называемое ***скалярным произведением***, причем это соответствие удовлетворяет следующим условиям:

1. $(u, v) = (v, u) \quad \forall u, v \in E$;
2. $(u + v, w) = (u, w) + (v, w) \quad \forall u, v, w \in E$;
3. $(\lambda \cdot u, v) = \lambda \cdot (u, v) \quad \forall u, v \in E,\ \forall \lambda \in R$;
4. $(v, v) > 0 \quad \forall v \ne o$ и $(v, v) = 0 \ \Rightarrow \ v = o$.

В скалярном произведении (u, v) вектор u - первый, а вектор v второй сомножители. Скалярное произведение (v, v) вектора v на себя называется ***скалярным квадратом***. Условия 1–4 называются ***аксиомами скалярного произведения***. Аксиома 1 определяет ***симметричность*** скалярного произведения, аксиомы 2 и 3 – ***аддитивность*** и ***однородность по первому сомножителю***, аксиома 4 – ***неотрицательность скалярного квадрата*** (v, v).

Линейные операции над векторами евклидова пространства удовлетворяют аксиомам 1–8 линейного пространства, а операция скалярного умножения векторов удовлетворяет аксиомам 1–4 скалярного произведения. Можно сказать, что евклидово пространство – это вещественное линейное пространство со скалярным произведением. Поскольку евклидово пространство является линейным пространством, на него переносятся все понятия, определенные для линейного пространства, в частности, понятия размерности и базиса.

ПРОСТЕЙШИЕ СЛЕДСТВИЯ ИЗ АКСИОМ СКАЛЯРНОГО ПРОИЗВЕДЕНИЯ

1. Аксиомы 2 и 3 скалярного произведения можно заменить одним условием ***линейности скалярного произведения по первому сомножителю***:

$$(\alpha \cdot u + \beta \cdot v, w) = \alpha \cdot (u, w) + \beta \cdot (v, w) \quad \forall u, v, w \in E,\ \forall \alpha, \beta \in R.$$

2. Условие линейности скалярного произведения по первому сомножителю в силу симметричности (аксиома 1) справедливо и для второго сомножителя, т.е. скалярное произведение линейно по любому сомножителю.

3. Линейность скалярного произведения по любому сомножителю распространяется на линейные комбинации векторов:

$$\left(\sum_{i=1}^{m} \alpha_i u_i, \sum_{j=1}^{n} \beta_j v_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j (u_i, v_j)$$

для любых векторов u_i, v_j и действительных чисел α_i, β_j, $i = 1, \ldots, m$; $j = 1, \ldots, n$.

4. Если хотя бы один сомножитель - нулевой вектор, то скалярное произведение равно нулю:

$$(v, o) = (o, v) = 0 \quad \forall v \in E.$$

Действительно, представим нулевой вектор в виде $o = 0 \cdot u$, где u – произвольный вектор из E. Тогда из аксиомы 3 получаем:

$$(o, v) = (0 \cdot u, v) = 0 \cdot (u, v) = 0.$$

15. Основные метрические понятия . Неравенство Коши – Буняковского.

Длиной (нормой) вектора v в евклидовом пространстве E называется число $|v| = \sqrt{(v, v)}$.

Имея в виду обозначение, длину $|v|$ называют также ***модулем*** вектора. Рассматривается арифметическое значение квадратного корня, которое определено для любого вектора из-за неотрицательности подкоренного выражения (аксиома 4). Поэтому каждый вектор имеет положительную длину, за исключением нулевого, длина которого равна нулю: $|o| = 0$.

Углом между ненулевыми векторами u и v евклидова пространства E называется число

$$\varphi = \arccos \frac{(u, v)}{|u| \cdot |v|}, \text{ т.е. } \cos\varphi = \frac{(u, v)}{|u| \cdot |v|} \text{ и } 0 \le \varphi \le \pi.$$

Представив неравенство Коши–Буняковского (8.25) в виде

$$|(u, v)| \le |u| \cdot |v|,$$

можно сделать вывод, что абсолютное значение выражения $\frac{(u, v)}{|u| \cdot |v|}$ не превосходит единицы, т.е. величина угла определена для любой пары ненулевых векторов. Заметим, что угол между коллинеарными векторами равен нулю или π.

Длина вектора и угол между векторами называются ***основными метрическими понятиями*** [42].

НЕРАВЕНСТВО КОШИ–БУНЯКОВСКОГО

Для любых векторов u и v евклидова пространства E выполняется ***неравенство Коши–Буняковского***:

$$(u, v)^2 \le (u, u) \cdot (v, v). \qquad (8.25)$$

В самом деле, для любого действительного числа λ и любых векторов u и v справедливо неравенство

$$0 \le (u - \lambda v, u - \lambda v) = \lambda^2 (v, v) - 2\lambda (u, v) + (u, u).$$

Следовательно, дискриминант квадратного трехчлена (переменной λ) не больше нуля, т.е. $4 \cdot (u, v)^2 - 4 \cdot (u, u) \cdot (v, v) \le 0$. Отсюда следует (8.25). Заметим, что равенство нулю дискриминанта возможно только в случае существования такого корня λ, для которого $(u - \lambda v, u - \lambda v) = 0$. Это условие равносильно коллинеарности векторов u и v: $u = \lambda \cdot v$. Напомним, что ненулевые векторы u и v называются ***коллинеарными***, если существует такое число λ, что $u = \lambda \cdot v$. Нулевой вектор считается коллинеарным любому вектору. Неравенство Коши–Буняковского выполняется как равенство только для коллинеарных векторов и как строгое неравенство для неколлинеарных.

16. Неравенство треугольника . Теорема Пифагора.

Из неравенства Коши–Буняковского (8.25) следует *неравенство треугольника*:

$$\big||u| - |v|\big| \le |u + v| \le |u| + |v|.$$

Докажем последнее неравенство. Применяя оценку $(u, v) \le |u| \cdot |v|$, получаем

$$|u + v|^2 = (u + v, u + v) = (u, u) + 2(u, v) + (v, v) \le$$

$$\le |u|^2 + 2 \cdot |u| \cdot |v| + |v|^2 = (|u| + |v|)^2,$$

т.е. $|u + v|^2 \le (|u| + |v|)^2 \ \Leftrightarrow \ |u + v| \le |u| + |v|$.

6. *Если $v_1, v_2, \ldots, v_k$ – ортогональная система векторов, то*

$$|v_1 + v_2 + \ldots + v_k|^2 = |v_1|^2 + |v_2|^2 + \ldots + |v_k|^2.$$

Это утверждение является обобщением теоремы Пифагора.

Замечания 8.14.

1. Из теоремы Пифагора $|v|^2 = |l|^2 + |h|^2$ следуют неравенства: $|l| \le |v|$, $|h| \le |v|$. Равенства возможны только тогда, когда $v \in L$ или $v \perp L$ соответственно. В остальных случаях неравенства строгие, т.е. получаем утверждения, знакомые читателю из курса геометрии: $|l| < |v|$ – проекция меньше наклонной, $|h| < |v|$ – перпендикуляр есть кратчайшее расстояние от конца вектора v до подпространства L.

Первые две суммы обозначим w_1 – это некоторый вектор из L_1, последнюю сумму обозначим w_2 – это некоторый вектор из L_2. Равенство (8.14): $w_1 + w_2 = o$ означает, что вектор $w_2 = -w_1$ принадлежит также и пространству L_1. Значит, $w_2 \in L_1 \cap L_2$. Раскладывая этот вектор по базису (e), находим $w_2 = \sum_{i=1}^{s} \delta_i e_i$. Учитывая разложение этого вектора в (8.14), получаем

$$w_2 = \sum_{i=1}^{s} \delta_i e_i = \sum_{i=s+1}^{m_2} \gamma_i e_i'' \quad \Leftrightarrow \quad \sum_{i=1}^{s} \delta_i e_i - \sum_{i=s+1}^{m_2} \gamma_i e_i'' = o .$$

Последнее равенство можно рассматривать, как разложение нулевого вектора по базису $(e),(e'')$ подпространства L_2. Все коэффициенты такого разложения нулевые: $\delta_1 = \ldots = \delta_s = 0$ и $\gamma_{s+1} = \ldots = \gamma_{m_2} = 0$. Подставляя $\gamma_i = 0$ в (8.14), получаем $\sum_{i=1}^{s} \alpha_i e_i + \sum_{i=s+1}^{m_1} \beta_i e_i' = o$. Это возможно только в тривиальном случае $\alpha_1 = \ldots = \alpha_s = 0$ и $\beta_{s+1} = \ldots = \beta_{m_1} = 0$, так как система векторов $(e),(e')$ линейно независима (это базис подпространства L_1). Таким образом, равенство (8.14) выполнятся только в тривиальном случае, когда все коэффициенты равны нулю одновременно. Следовательно, совокупность векторов $(e),(e'),(e'')$ линейно независима, т.е. является базисом пространства $L_1 + L_2$. Подсчитаем размерность суммы подпространств:

$\dim(L_1 + L_2) = s + (m_1 - s) + (m_2 - s) = m_1 + m_2 - s = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2)$, что и требовалось доказать.

8.8.2. Примеры евклидовых пространств

Определяя для элементов линейного пространства операцию скалярного произведения, получаем евклидово пространство. Если скалярное произведение можно ввести разными способами в одном и том же линейном пространстве, то и получаемые евклидовы пространства будут разными. Приведем примеры евклидовых пространств, соответствующих рассмотренным в разд.8.1.3 примерам линейных пространств.

1. В нулевом линейном пространстве $\{o\}$ скалярное произведение можно определить единственным способом, положив $(o, o) = 0$. Аксиомы скалярного произведения при этом выполняются.

2. В пространствах V_1, V_2, V_3 векторы (свободные или радиус-векторы) рассматриваются как направленные отрезки. В курсе элементарной геометрии вводятся понятия длины вектора и величины угла между векторами, а затем определяется скалярное произведение: $(\vec{u}, \vec{v}) = |\vec{u}| \cdot |\vec{v}| \cdot \cos\varphi$. Аксиомы 1–4 для этого скалярного произведения выполняются. Поэтому пространства V_1, V_2, V_3 являются евклидовыми. Неравенство Коши–Буняковского в этом пространстве означает, что $|\cos\varphi| = \frac{|(\vec{u}, \vec{v})|}{|\vec{u}| \cdot |\vec{v}|} \le 1$. Геометрический смысл: длина проекции не превосходит длины наклонной (катет короче гипотенузы).

3. В пространстве R^n скалярное произведение столбцов $x = (x_1 \ \cdots \ x_n)^T$ и $y = (y_1 \ \cdots \ y_n)^T$ можно задать формулой:

$$(x, y) = x^T A y = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j , \tag{8.26}$$

где A – квадратная симметрическая положительно определенная матрица n-го порядка (см. разд.6.6.3). Проверим выполнение аксиом 1–4. Аксиома 1 (симметричность) выполняется в силу симметричности матрицы A: $(x, y) = x^T A y = y^T A^T x = y^T A x = (y, x)$, поскольку число при транспонировании не изменяется, т.е. $x^T A y = y^T A^T x$. Свойство линейности по первому сомножителю (см. п.1 простейших следствий из аксиом) для (8.26) выполняется: $(\alpha x + \beta y, z) = (\alpha x + \beta y)^T A z = \alpha x^T A z + \beta y^T A z = \alpha(x, z) + \beta(y, z)$. Значит, выполняются аксиомы 2 и 3. Аксиома 4 также выполняется, так как квадратичная форма $(x, x) = x^T A x$ положительно определенная (см. разд. 6.5.4). Таким образом, пространство R^n со скалярным произведением (8.26) является евклидовым пространством. В частности, если в качестве матрицы A взять единичную матрицу, формула (8.26) примет вид:

$$(x, y) = x^T y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n . \tag{8.27}$$

Это скалярное произведение считается ***стандартным*** в пространстве R^n. Неравенство (8.25) Коши–Буняковского в n-мерном арифметическом пространстве R^n со скалярным произведением (8.27) трансформируется в ***неравенство Коши***:

$$(x_1 y_1 + x_2 y_2 + \ldots + x_n y_n)^2 \le (x_1^2 + x_2^2 + \ldots + x_n^2) \cdot (y_1^2 + y_2^2 + \ldots + y_n^2).$$

Приведем примеры формул, которые не задают скалярного произведения в R^2:

$(x, y) = |x_1| \cdot |y_1| + |x_2| \cdot |y_2|$ – аксиомы 1, 4 выполняются, а аксиомы 2, 3 – нет;

$(x, y) = x_2 \cdot y_2$ – аксиомы 1, 2, 3 выполняются, а аксиома 4 – нет.

4. Пространство $\{Ax = o\}$ решений однородной системы $Ax = o$ линейных уравнений со скалярным произведением (8.27) является евклидовым пространством.

17. Изоморфизм евклидовых пространств.

Два евклидовых пространства E и E' называются ***изоморфными*** ($E \leftrightarrow E'$), если они изоморфны как линейные пространства (см. разд.8.5) и скалярные произведения соответствующих векторов равны:

$$\left.\begin{matrix} u \leftrightarrow u' \\ v \leftrightarrow v' \end{matrix}\right\} \quad \Rightarrow \quad (u, v) = (u', v')',$$

где $(\cdot,\cdot)$ и $(\cdot,\cdot)'$ – скалярные произведения в пространствах E и E' соответственно.

Напомним, что для изоморфизма конечномерных линейных пространств необходимо и достаточно, чтобы их размерности совпадали (см. теорему 8.3). Покажем, что это условие достаточно для изоморфизма евклидовых пространств (необходимость следует из определения). Как и при доказательстве теоремы 8.3, установим изоморфизм n-мерного евклидова пространства E с вещественным арифметическим пространством R^n со скалярным произведением (8.27). В самом деле, взяв в пространстве E какой-нибудь ортонормированный базис $(e) = (e_1, \ldots, e_n)$, поставим в соответствие каждому вектору $x \in E$ его координатный столбец $x \in R^n$ ($x \leftrightarrow x$). Это взаимно однозначное соответствие устанавливает изоморфизм линейных пространств: $E \leftrightarrow R^n$. В ортонормированном базисе скалярное произведение векторов x и y пространства E находится по формуле $(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$ (см. п.1 преимуществ ортонормированного базиса) Такое же выражение даст скалярное произведение (8.27) координатных столбцов x и y, т.е. скалярные произведения соответствующих элементов равны $(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = x^T y$. Следовательно, евклидовы пространства E и R^n изоморфны.

Таким образом, изучение конечномерных евклидовых пространств может быть сведено к исследованию вещественного арифметического пространства R^n со стандартным скалярным произведением (8.27).

Примечание:

$$(x, y) = x^T y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n .$$

18. Ортогональные векторы: определение, примеры, свойства.

Два вектора u и v евклидова пространства называются ***ортогональными*** (***перпендикулярными***), если их скалярное произведение равно нулю: $(u, v) = 0$.

Система векторов $v_1, v_2, \ldots, v_k$ называется ***ортогональной***, если все ее векторы попарно ортогональны, т.е. $(v_i, v_j) = 0$ при $i \ne j$. Система векторов $v_1, v_2, \ldots, v_k$ называется ***ортонормированной***, если все ее векторы попарно ортогональны и длина (норма) каждого вектора системы равна единице, т.е.

$$(v_i, v_j) = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases}$$

Говорят, что вектор v ***ортогонален*** (***перпендикулярен***) ***множеству*** M, если он ортогонален каждому вектору из M. Ортогональность векторов обозначается знаком перпендикуляра ($\perp$).

СВОЙСТВА ОРТОГОНАЛЬНЫХ ВЕКТОРОВ

1. *Нулевой вектор ортогонален каждому вектору пространства.*

2. *Взаимно ортогональные ненулевые векторы линейно независимы.*

В самом деле, пусть векторы $v_1, v_2, \ldots, v_k$ попарно ортогональны. Составим из них линейную комбинацию и приравняем ее нулевому вектору:

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = o .$$

Умножим обе части равенства скалярно на вектор v_1:

$$\lambda_1 (v_1, v_1) + \lambda_2 \underbrace{(v_1, v_2)}_{0} + \ldots + \lambda_k \underbrace{(v_1, v_k)}_{0} = \underbrace{(v_1, o)}_{0}.$$

Следовательно, $\lambda_1 \cdot |v_1|^2 = 0$. Так как $v_1 \ne o$, то $\lambda_1 = 0$. Аналогично доказываем, что $\lambda_2 = \ldots = \lambda_k = 0$, т.е рассматриваемая линейная комбинация тривиальная. Значит, ортогональная система векторов $v_1, v_2, \ldots, v_k$ линейно независима.

3. *Если сумма взаимно ортогональных векторов равна нулевому вектору, то каждое из слагаемых равно нулевому вектору.*

4. *Если вектор u ортогонален каждому вектору системы $v_1, v_2, \ldots, v_k$, то он также ортогонален и любой их линейной комбинации.* Другими словами, *если $u \perp v_i$, $i = 1, \ldots, k$, то $u \perp Lin(v_1, \ldots, v_k)$.*

5. *Если вектор u ортогонален подмножеству M евклидова пространства, то он ортогонален и линейной оболочке этого подмножества,* т.е. $u \perp M \Rightarrow u \perp Lin(M)$.

6. *Если $v_1, v_2, \ldots, v_k$ – ортогональная система векторов, то*

$$|v_1 + v_2 + \ldots + v_k|^2 = |v_1|^2 + |v_2|^2 + \ldots + |v_k|^2 .$$

Это утверждение является обобщением теоремы Пифагора.

19. Ортонормированный базис и его преимущества.

Так как евклидово пространство является линейным, на него переносятся все понятия и свойства, относящиеся к линейному пространству, в частности, понятия базиса и размерности.

Базис $e_1, \ldots, e_n$ евклидова пространства называется ***ортогональным***, если все образующие его векторы попарно ортогональны, т.е.

$$(e_i, e_j) = 0 \text{ при } i \ne j,\ i = 1, \ldots, n,\ j = 1, \ldots, n .$$

Базис $e_1, \ldots, e_n$ евклидова пространства называется ***ортонормированным***, если его векторы попарно ортогональны и длина каждого из них равна единице:

$$(e_i, e_j) = \begin{cases} 1, & i = j, \\ 0, & i \ne j, \end{cases} \quad i = 1, \ldots, n,\ j = 1, \ldots, n . \tag{8.31}$$

Теорема 8.5. *В конечномерном евклидовом пространстве любую систему ортогональных (ортонормированных) векторов можно дополнить до ортогонального (ортонормированного) базиса.*

В самом деле, по теореме 8.2 любую систему линейно независимых векторов, в частности, ортогональную (ортонормированную), можно дополнить до базиса. Применяя к этому базису процесс ортогонализации (см. разд. 8.8.5), получаем ортогональный базис. Нормируя векторы этого базиса (см. п.4 замечаний 8.11), получаем ортонормированный базис.

ПРЕИМУЩЕСТВА ОРТОНОРМИРОВАННОГО БАЗИСА

Для ортонормированного базиса $e_1, \ldots, e_n$ формула (8.32) упрощается, так как из условия (8.31) следует, что матрица Грама $G(e_1, \ldots, e_n)$ ортонормированной системы $e_1, \ldots, e_n$ равна единичной матрице: $G(e_1, \ldots, e_n) = E$.

1. *В ортонормированном базисе $e_1, \ldots, e_n$ скалярное произведение векторов x и y находится по формуле: $(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$, где $x_1, \ldots, x_n$ – координаты вектора x, а $y_1, \ldots, y_n$ – координаты вектора y.*

2. *В ортонормированном базисе $e_1, \ldots, e_n$ длина вектора x вычисляется по формуле $|x| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, где $x_1, \ldots, x_n$ – координаты вектора x.*

3. *Координаты $x_1, \ldots, x_n$ вектора x относительно ортонормированного базиса $e_1, \ldots, e_n$ находятся при помощи скалярного произведения по формулам: $x_1 = (x, e_1), \ldots, x_n = (x, e_n)$.*

В самом деле, умножая обе части равенства $x = x_1 e_1 + \ldots + x_n e_n$ на e_1, получаем $(x, e_1) = x_1 \underbrace{(e_1, e_1)}_{1} + x_2 \underbrace{(e_2, e_1)}_{0} + \ldots + x_n \underbrace{(e_n, e_1)}_{0}$, т.е. $x_1 = (x, e_1)$. Аналогично доказываются остальные формулы.

20. Ортогональные дополнения подмножеств: определения, примеры, свойства.

Ортогональным дополнением непустого подмножества M евклидова пространства E называется множество векторов, ортогональных каждому вектору из M. Ортогональное дополнение обозначается

$$M^{\perp} = \{ v : (v, w) = 0,\ \forall w \in M \}.$$

Рассмотрим примеры ортогональных дополнений.

1. Ортогональным дополнением нулевого подпространства $\{o\} \triangleleft E$ служит все пространство E: $\{o\}^{\perp} = E$. Ортогональным дополнением всего пространства является его нулевое подпространство $E^{\perp} = \{o\}$.

2. Пусть в пространстве V_3 радиус-векторов (с началом в точке O) заданы три взаимно перпендикулярных радиус-вектора $\overline{OA}$, $\overline{OB}$ и $\overline{OC}$. Тогда ортогональным дополнением вектора $\overline{OA}$ является множество радиус-векторов на плоскости, содержащей векторы $\overline{OB}$ и $\overline{OC}$, точнее, $\{\overline{OA}\}^{\perp} = Lin(\overline{OB}, \overline{OC})$. Ортогональным дополнением векторов $\overline{OA}$ и $\overline{OB}$ служит множество радиус-векторов на прямой, содержащей вектор $\overline{OC}$: $\{\overline{OA}, \overline{OB}\}^{\perp} = Lin(\overline{OC})$. Ортогональным дополнение трех заданных векторов служит нулевой радиус-вектор: $\{\overline{OA}, \overline{OB}, \overline{OC}\}^{\perp} = \{\overline{OO}\}$.

3. В пространстве $P_2(R)$ многочленов степени не выше второй со скалярным произведением (8.29) (см. п.6 в разд.8.8.2) задано подмножество $P_0(R)$ – многочленов нулевой степени. Найдем ортогональное дополнение этого подмножества. Для этого приравняем нулю скалярное произведение многочлена $p_2(x) = ax^2 + bx + c$ на постоянный многочлен $p_0(x) = d$: $(p_2(x), p_0(x)) = a \cdot 0 + b \cdot 0 + c \cdot d = 0$. Поскольку величина d произвольная, то $c = 0$. Следовательно, ортогональным дополнением подмножества $P_0(R)$ является множество многочленов из $P_2(R)$ с нулевым свободным членом.

СВОЙСТВА ОРТОГОНАЛЬНОГО ДОПОЛНЕНИЯ

Рассмотрим свойства ортогональных дополнений подмножеств n-мерного евклидова пространства E.

1. *Ортогональное дополнение $M^{\perp}$ непустого подмножества $M \subset E$ является линейным подпространством, т.е. $M^{\perp} \triangleleft E$, и справедливо включение $M \subset (M^{\perp})^{\perp}$.*

В самом деле, множество $M^{\perp}$ замкнуто по отношению к операциям сложения векторов и умножения вектора на число, так как сумма двух векторов, ортогональных M, ортогональна M, и произведение вектора, ортогонального M, на любое число является вектором, ортогональным M. Докажем включение $M \subset (M^{\perp})^{\perp}$. Пусть $w \in M$, тогда $(w, v) = 0$ для любого вектора $v \in M^{\perp}$. Но это означает, что $w \in (M^{\perp})^{\perp}$

2. *Пересечение любого непустого подмножества $M \subset E$ со своим ортогональным дополнением есть нулевой вектор: $M \cap M^{\perp} = \{o\}$.*

Действительно, только нулевой вектор ортогонален самому себе.

3. *Если L – подпространство E ($L \triangleleft E$), то $E = L \oplus L^{\perp}$.*

Действительно, возьмем в L ортогональный базис $(e) = (e_1, \ldots, e_k)$. Дополним его векторами $(f) = (f_{k+1}, \ldots, f_n)$ до ортогонального базиса $(e), (f)$ всего пространства E. Тогда произвольный вектор $w \in E$ можно представить в виде суммы

$$w = \underbrace{\sum_{i=1}^{k} w_i e_i}_{u} + \underbrace{\sum_{j=k+1}^{n} w_j f_j}_{v} = u + v ,$$

где $u \in L$, а $v \in L^{\perp}$, так как $(v, e_i) = \sum_{j=k+1}^{n} w_j \underbrace{(f_j, e_i)}_{0} = 0$ для $i = 1, \ldots, k$. Следовательно, любой вектор пространства E раскладывается по подпространствам L и $L^{\perp}$, т.е. $E = L + L^{\perp}$. Эта алгебраическая сумма является прямой суммой по свойству 2, поскольку $L \cap L^{\perp} = \{o\}$. Следовательно, $E = L \oplus L^{\perp}$.

4. *Если $L \triangleleft E$, то* $\dim L^{\perp} = \dim E - \dim L$.

5. *Если L – подпространство E, то* $L = (L^{\perp})^{\perp}$.

Из первого свойства следует включение $L \subset (L^{\perp})^{\perp}$. Докажем, что $(L^{\perp})^{\perp} \subset L$. Действительно, пусть $w \in (L^{\perp})^{\perp}$. По свойству 3: $w = u + v$, где $u \in L$, $v \in L^{\perp}$. Найдем скалярное произведение $\underbrace{(w, v)}_{0} = (u + v, v) = \underbrace{(u, v)}_{0} + (v, v) = (v, v)$. Следовательно, $(v, v) = 0$, и согласно аксиоме 4 скалярного произведения (см. разд. 8.8.1) $v = o$, поэтому $w = u + v = u + o = u \in L$. Значит, $(L^{\perp})^{\perp} \subset L$. Из двух включений $L \subset (L^{\perp})^{\perp}$ и $(L^{\perp})^{\perp} \subset L$ следует равенство $L = (L^{\perp})^{\perp}$.

6. *Если $L_1 \triangleleft E$ и $L_2 \triangleleft E$, то*

$$(L_1 + L_2)^{\perp} = L_1^{\perp} \cap L_2^{\perp} \quad \text{и} \quad (L_1 \cap L_2)^{\perp} = L_1^{\perp} + L_2^{\perp} .$$

Последние свойства аналогичны свойствам алгебраических дополнений (см. разд. 8.6.4).

НАХОЖДЕНИЕ ОРТОГОНАЛЬНОГО ДОПОЛНЕНИЯ ПОДПРОСТРАНСТВА

В разд. 8.6.5 для описания подпространств линейных пространств использовались два способа описания (внешний и внутренний). Рассмотрим применение этих способов описания для нахождения ортогональных дополнений подпространств. Учитывая изоморфизм евклидовых пространств, будем рассматривать арифметическое пространство R^n со скалярным произведением (8.27).

Для заданного подпространства $L \triangleleft R^n$ требуется найти его ортогональное дополнение $L^{\perp}$. В зависимости от способа описания подпространства L используем одно из следующих двух утверждений.

1. *Если подпространство $L \triangleleft R^n$ задано как линейная оболочка $L = Lin(a_1, \ldots, a_k)$ столбцов матрицы $A = (a_1 \ \cdots \ a_k)$, то множество решений однородной системы $A^T x = o$ является его ортогональным дополнением $L^{\perp} \triangleleft R^n$*, т.е.

$$L = Lin(a_1, \ldots, a_k) \quad \Rightarrow \quad L^{\perp} = \{ A^T x = o \}. \tag{8.34}$$

2. *Если подпространство $L \triangleleft R^n$ задано как множество решений однородной системы $Ax = o$ m уравнений с n неизвестными, то линейная оболочка столбцов $a_1^T, \ldots, a_m^T$ транспонированной матрицы $A^T = (a_1^T \ \cdots \ a_m^T)$ является его ортогональным дополнением $L^{\perp} \triangleleft R^n$*, т.е.

$$L = \{ Ax = o \} \quad \Rightarrow \quad L^{\perp} = Lin(a_1^T, \ldots, a_m^T), \tag{8.35}$$

где a_i^T – i-й столбец матрицы A^T.

21. *Процесс ортогонализации Грама – Шмидта*

8.8.5. Процесс ортогонализации Грама–Шмидта

Рассмотрим следующую задачу. Дана линейно независимая система $v_1, v_2, \ldots, v_k$ векторов конечномерного евклидова пространства. Требуется построить ортогональную систему $w_1, w_2, \ldots, w_k$ векторов того же пространства так, чтобы совпадали линейные оболочки:

$$Lin(w_1, w_2, \ldots, w_k) = Lin(v_1, v_2, \ldots, v_k).$$

Решение задачи находится при помощи ***процесса ортогонализации*** (Грама – Шмидта), выполняемого за k шагов.

1. Положить $w_1 = v_1$.

2. Найти $w_2 = v_2 - \alpha_{21} \cdot w_1$, где $\alpha_{21} = \dfrac{(v_2, w_1)}{(w_1, w_1)}$.

3. Найти $w_3 = v_3 - \alpha_{31} w_1 - \alpha_{32} w_2$, где $\alpha_{31} = \dfrac{(v_3, w_1)}{(w_1, w_1)}$, $\alpha_{32} = \dfrac{(v_3, w_2)}{(w_2, w_2)}$;

и т.д.

k) Найти $w_k = v_k - \sum\limits_{i=1}^{k-1} \alpha_{ki} w_i$, где $\alpha_{ki} = \dfrac{(v_k, w_i)}{(w_i, w_i)}$, $i = 1, \ldots, k-1$.

Поясним процесс ортогонализации. Искомый на втором шаге вектор w_2 представлен в виде линейной комбинации $w_2 = v_2 - \alpha \cdot w_1$. Коэффициент α подберем так, чтобы обеспечить ортогональность векторов w_2 и w_1. Приравняем нулю скалярное произведение этих векторов $(w_2, w_1) = (v_2, w_1) - \alpha \cdot (w_1, w_1) = 0$. Отсюда получаем, что $\alpha = \alpha_{21}$ (см. п.2 алгоритма). Подбор коэффициентов α_{ji} на j-м шаге алгоритма делается так, чтобы искомый вектор w_j был ортогонален всем ранее найденным векторам $w_1, w_2, \ldots, w_{j-1}$.

З а м е ч а н и я 8.11.

1. Векторы, найденные в процессе ортогонализации, обладают следующими свойствами:

а) $w_j \perp Lin(w_1, \ldots, w_{j-1})$, $j = 2, \ldots, k$;

б) $Lin(w_1) = Lin(v_1)$, $Lin(w_1, \ldots, w_j) = Lin(v_1, \ldots, v_j)$, $j = 2, \ldots, k$.

Первое свойство следует из свойства 4 ортогональных векторов (см. разд. 8.8.4). Второе свойство следует из того, что каждый вектор системы $w_1, \ldots, w_j$ линейно выражается через векторы $v_1, \ldots, v_j$, и наоборот.

2. В процессе ортогонализации любой вектор w_j можно заменить на коллинеарный ему ненулевой вектор $\lambda \cdot w_j$. При этом свойства, перечисленные в п.1, не нарушаются.

3. Если система $v_1, v_2, \ldots, v_k$ векторов линейно зависима, то в процессе ортогонализации будем получать (на некоторых шагах) нулевые векторы. Действительно, если подсистема $v_1, v_2, \ldots, v_j$ линейно зависима, то $v_j \in Lin(w_1, \ldots, w_{j-1})$. Тогда вектор $w_j = v_j - \sum\limits_{i=1}^{j-1} \alpha_{ji} w_i$ одновременно удовлетворяет двум условиям $w_j \perp Lin(w_1, \ldots, w_{j-1})$ и $w_j \in Lin(w_1, \ldots, w_{j-1})$. Значит, это нулевой вектор $w_j = o$.

Поэтому в данном случае формулы вычисления коэффициентов α_{ji} на j-м шаге следует записывать в виде:

$$\alpha_{ji} = \begin{cases} \dfrac{(v_j, w_i)}{(w_i, w_i)}, & w_i \ne o, \\ 0, & w_i = o, \end{cases} \qquad i = 1, \ldots, j-1.$$

В остальном процесс ортогонализации остается неизменным.

4. Процесс ортогонализации можно дополнить ***процессом нормировки***, разделив каждый вектор ортогональной системы $w_1, w_2, \ldots, w_k$ на его длину:

$$e_i = \frac{1}{|w_i|} \cdot w_i, \; i = 1, \ldots, k.$$

В результате получим ортонормированную систему $e_1, e_2, \ldots, e_k$, отвечающую условию $Lin(e_1, \ldots, e_k) = Lin(v_1, \ldots, v_k)$. Если исходная система векторов является линейно зависимой, то среди векторов ортогональной системы $w_1, w_2, \ldots, w_k$ будут нулевые. Чтобы получить ортонормированную систему, нулевые векторы следует исключить, а остальные векторы нормировать.

Пример 8.18. Даны системы векторов евклидовых пространств:

а) $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $y = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ – элементы пространства R^2 со скалярным произведением (8.26): $(x, y) = x^T \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} y = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$;

б) $p_1(x) = 1$, $p_2(x) = x$, $p_3(x) = x^2$ – элементы пространства $C[-1;1]$ со скалярным произведением (8.28): $(f, g) = \int\limits_{-1}^{1} f(x) g(x) dx$.

Провести ортогонализацию данных векторов.

☐ а) Заметим, что система векторов x, y, z линейно зависимая, так как x и y пропорциональны, поэтому используем процесс ортогонализации Грама – Шмидта с учетом п.3 замечаний 8.11.

1. Полагаем $u = x$.

2. Вычисляем $\alpha_{21} = \dfrac{(y, u)}{(u, u)} = \dfrac{2 \cdot 2 \cdot 1 + 1 \cdot 0 + 0 \cdot 2 + 0 \cdot 0}{2 \cdot 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0} = \dfrac{4}{2} = 2$ и находим

$$v = y - \alpha_{21} \cdot u = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Получили нулевой вектор.

3. Вычисляем $\alpha_{31} = \dfrac{(z, u)}{(u, u)} = \dfrac{2 \cdot 0 \cdot 1 + 0 \cdot 0 + 1 \cdot 1 + 1 \cdot 0}{2 \cdot 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0} = \dfrac{1}{2}$; $\alpha_{32} = 0$ согласно п.3 замечаний 8.11, так как $v = o$, и находим

$$w = z - \alpha_{31} \cdot u - \alpha_{32} v = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 0 \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}.$$

Проверим условие ортогональности $(u, w) = 2 \cdot 1 \cdot \left(\frac{-1}{2}\right) + 1 \cdot 1 + 0 \cdot \left(\frac{-1}{2}\right) + 0 \cdot 1 = 0$.

Для получения ортонормированной системы исключаем нулевой вектор $v = o$, а остальные нормируем (см. п.4 замечаний 8.11):

$$|u| = \sqrt{(u, u)} = \sqrt{2};$$

$$|w| = \sqrt{(w, w)} = \sqrt{2(-0{,}5)(-0{,}5) + (-0{,}5) \cdot 1 + 1 \cdot (-0{,}5) + 1 \cdot 1} = \sqrt{0{,}5};$$

$$\hat{u} = \frac{1}{|u|} \cdot u = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \hat{w} = \frac{1}{|w|} \cdot w = \frac{1}{\sqrt{0{,}5}} \cdot \begin{pmatrix} -0{,}5 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \sqrt{2} \end{pmatrix}.$$

Таким образом, для системы трех векторов x, y, z построена ортогональная система из трех векторов u, v, w и ортонормированная система из двух векторов $\hat{u}$, $\hat{w}$. Линейные оболочки этих трех систем совпадают между собой (и со всем пространством R^2).

б) 1. Полагаем $q_1(x) = p_1(x) = 1$.

2. Вычисляем $\alpha_{21} = \dfrac{(p_2, q_1)}{(q_1, q_1)} = \left(\int\limits_{-1}^{1} x \cdot 1\, dx\right) : \left(\int\limits_{-1}^{1} 1 \cdot 1\, dx\right) = \dfrac{1}{2} x^2 \Big|_{-1}^{1} : 2 = 0$ и находим $q_2(x) = x - 0 \cdot 1 = x$.

3. Вычисляем $\alpha_{31} = \dfrac{(p_3, q_1)}{(q_1, q_1)} = \left(\int\limits_{-1}^{1} x^2 \cdot 1\, dx\right) : 2 = \dfrac{1}{3} x^3 \Big|_{-1}^{1} : 2 = \dfrac{2}{3} \cdot \dfrac{1}{2} = \dfrac{1}{3}$,

$\alpha_{32} = \dfrac{(p_3, q_2)}{(q_2, q_2)} = \left(\int\limits_{-1}^{1} x^2 \cdot x\, dx\right) : \left(\int\limits_{-1}^{1} x \cdot x\, dx\right) = \dfrac{1}{4} x^4 \Big|_{-1}^{1} : \dfrac{1}{3} x^3 \Big|_{-1}^{1} = 0$ и находим

$$q_3(x) = x^2 - \alpha_{31} \cdot 1 - \alpha_{32} \cdot x = x^2 - \frac{1}{3}.$$

Получили ортогональные многочлены $q_1(x) = 1$, $q_2(x) = x$, $q_3(x) = x^2 - \frac{1}{3}$. Выполним нормировку: $|q_1(x)| = \sqrt{(q_1(x), q_1(x))} = \sqrt{2}$; $|q_2(x)| = \sqrt{(q_2(x), q_2(x))} = \sqrt{\frac{2}{3}}$; $|q_3(x)| = \sqrt{(q_3(x), q_3(x))} = \frac{2}{3}\sqrt{\frac{2}{5}}$;

$$\hat{q}_1(x) = \frac{1}{|q_1(x)|} \cdot q_1(x) = \frac{1}{\sqrt{2}} \cdot 1 = \frac{1}{\sqrt{2}}; \quad \hat{q}_2(x) = \frac{1}{|q_2(x)|} \cdot q_2(x) = \sqrt{\frac{3}{2}} \cdot x;$$

$$\hat{q}_3(x) = \frac{1}{|q_3(x)|} \cdot q_3(x) = \frac{3}{2}\sqrt{\frac{5}{2}} \cdot \left(x^2 - \frac{1}{3}\right).$$

Получили ортонормированные многочлены (см. многочлены Лежандра [8,25,42]). ■

22. Задача о перпендикуляре и ее решение.

8.8.8. Задача о перпендикуляре

Пусть L – подпространство конечномерного евклидова пространства E. Для любого вектора $v \in E$ (по свойству 3 ортогонального дополнения) существует единственное разложение:

$$v = l + h, \text{ где } l \in L, \; h \in L^{\perp}. \tag{8.36}$$

Вектор l называется ***ортогональной проекцией*** вектора v на подпространство L, а вектор h – ***ортогональной составляющей*** вектора v относительно подпространства L. По аналогии с привычными терминами курса элементарной геометрии ортогональную составляющую h называют ***перпендикуляром, опущенным из конца вектора v на подпространство L***. Из-за ортогональности составляющих l и h разложение (8.36) называют ***ортогональным***.

Задача о перпендикуляре ставится следующим образом. В n-мерном евклидовом пространстве заданы вектор $v \in E$ и подпространство $L \triangleleft E$. Требуется найти ортогональную проекцию $l \in L$ вектора v и его ортогональную составляющую (перпендикуляр) $h \in L^{\perp}$, т.е. представить заданный вектор v в виде (8.36).

Для решения задачи о перпендикуляре нужно выполнить следующие действия.

1. Взять любой базис $e_1, \ldots, e_r$ подпространства L (полагаем, что $\dim L = r \le n$).

2. Составить неоднородную систему

$$\begin{cases} (e_1, e_1) \cdot l_1 + (e_1, e_2) \cdot l_2 + \ldots + (e_1, e_r) \cdot l_r = (e_1, v), \\ \quad \vdots \\ (e_r, e_1) \cdot l_1 + (e_r, e_2) \cdot l_2 + \ldots + (e_r, e_r) \cdot l_r = (e_r, v) \end{cases}$$

r уравнений с r неизвестными $l_1, \ldots, l_r$.

3. Решить систему, составленную в п.2.

4. Найти ортогональную проекцию $l = l_1 \cdot e_1 + \ldots + l_r \cdot e_r$, а затем – ортогональную составляющую (перпендикуляр) $h = v - l$.

Поясним алгоритм решения. Разложив ортогональную проекцию $l = l_1 \cdot e_1 + \ldots + l_r \cdot e_r$ по базису подпространства, запишем ортогональную составляющую (перпендикуляр): $h = v - l = v - l_1 \cdot e_1 - \ldots - l_r \cdot e_r$. Затем найдем скалярные произведения (h, e_i), $i = 1, \ldots, r$, умножая последнее равенство последовательно на $e_1, \ldots, e_r$. Учитывая, что $(h, e_i) = 0$, $i = 1, \ldots, r$, получаем систему из п.2 алгоритма. Заметим, что матрица полученной системы – это матрица Грама $G(e_1, \ldots, e_r)$ линейно независимой системы векторов (базиса) $e_1, \ldots, e_r$. По свойству 1 определителя Грама $\det G(e_1, \ldots, e_r) \ne 0$, значит, рассматриваемая система имеет единственное решение.

Пример 8.20. В пространстве R^4 со стандартным скалярным произведением (8.27) заданы: вектор $v = (-3 \;\; 2 \;\; 0 \;\; 0)^T$ и подпространство L – множество решений однородной системы

$$\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 0, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 0. \end{cases}$$

Требуется найти ортогональную проекцию $l \in L$ и ортогональную составляющую $h \in L^{\perp}$ вектора v относительно подпространства L.

☐ 1. Базис подпространства был найден в примере 8.9: $L = Lin(\varphi_1, \varphi_2)$ где $\varphi_1 = (-6 \;\; 4 \;\; 1 \;\; 0)^T$, $\varphi_2 = (-2 \;\; 1 \;\; 0 \;\; 1)^T$.

2. Вычисляем скалярные произведения

$(\varphi_1, \varphi_1) = (-6)^2 + 4^2 + 1^2 + 0^2 = 53$; $(\varphi_1, \varphi_2) = (-6) \cdot (-2) + 4 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 = 16$;
$(\varphi_2, \varphi_2) = (-2)^2 + 1^2 + 0^2 + 1^2 = 6$; $(\varphi_1, v) = (-6) \cdot (-3) + 4 \cdot 2 + 1 \cdot 0 + 0 \cdot 0 = 26$;
$(\varphi_2, v) = (-2) \cdot (-3) + 1 \cdot 2 + 0 \cdot 0 + 1 \cdot 0 = 8$, $(\varphi_2, \varphi_1) = (\varphi_1, \varphi_2) = 16$

и составляем неоднородную систему $\begin{cases} 53 \cdot l_1 + 16 \cdot l_2 = 26, \\ 16 \cdot l_1 + 6 \cdot l_2 = 8. \end{cases}$

3. Решаем систему по правилу Крамера (см. разд.5.2):

$$l_1 = \frac{26 \cdot 6 - 16 \cdot 8}{53 \cdot 6 - 16 \cdot 16} = \frac{28}{62} = \frac{14}{31}; \qquad l_2 = \frac{53 \cdot 8 - 16 \cdot 26}{53 \cdot 6 - 16 \cdot 16} = \frac{8}{62} = \frac{4}{31}.$$

4. Находим ортогональную проекцию и ортогональную составляющую

$$l = \frac{14}{31}\begin{pmatrix} -6 \\ 4 \\ 1 \\ 0 \end{pmatrix} + \frac{4}{31}\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{31}\begin{pmatrix} -92 \\ 60 \\ 14 \\ 4 \end{pmatrix}; \quad h = \begin{pmatrix} -3 \\ 2 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{31}\begin{pmatrix} -92 \\ 60 \\ 14 \\ 4 \end{pmatrix} = \frac{1}{31}\begin{pmatrix} -1 \\ 2 \\ -14 \\ -4 \end{pmatrix}.$$

Проверим ортогональность составляющих:

$$(l, h) = \frac{1}{31^2}\left((-92) \cdot (-1) + 60 \cdot 2 + 14 \cdot (-14) + 4 \cdot (-4)\right) = 0. \; ■$$

З а м е ч а н и я 8.14.

1. Из теоремы Пифагора $|v|^2 = |l|^2 + |h|^2$ следуют неравенства: $|l| \le |v|$, $|h| \le |v|$. Равенства возможны только тогда, когда $v \in L$ или $v \perp L$ соответственно. В остальных случаях неравенства строгие, т.е. получаем утверждения, знакомые читателю из курса геометрии: $|l| < |v|$ – проекция меньше наклонной, $|h| < |v|$ – перпендикуляр есть кратчайшее расстояние от конца вектора v до подпространства L.

2. Для одномерного подпространства $L = Lin(e)$ составляющую $l \in L$ в разложении (8.36) называют ***ортогональной проекцией на ось, задаваемую вектором e*** (или ***на направление, задаваемое вектором e***). Если ось задается единичным вектором ($|e| = 1$), то длина ортогональной проекции равна $l = |l| = (l, e)$.

3. Если в подпространстве L взять ортонормированный базис $e_1, \ldots, e_r$, то квадрат длины вектора l можно вычислить по формуле $|l|^2 = l_1^2 + l_2^2 + \ldots + l_r^2$, где $l_i = (l, e_i)$, $i = 1, \ldots, r$. Тогда из неравенства (см. п.1) $|l| \le |v|$ следует:

$$\sum_{i=1}^{r} l_i^2 \le |v|^2 \qquad \text{(\textit{неравенство Бесселя})},$$

т.е. квадрат длины вектора не меньше суммы квадратов длин его проекций на любые r взаимно ортогональных направлений.

4. В процессе ортогонализации системы векторов $v_1, v_2, \ldots, v_k$ на каждом шаге фактически решается задача о перпендикуляре. Например, на j-м шаге находится ортогональная составляющая w_j вектора v_j относительно подпространства $Lin(w_1, \ldots, w_{j-1})$ (см. п.1 замечаний 8.11).

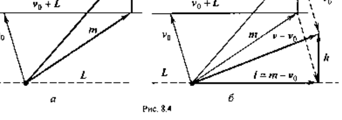

Рис. 8.4

5. В приложениях приходится также рассматривать задачу о перпендикуляре не для подпространства, а для многообразия. Пусть в n-мерном евклидовом пространстве заданы вектор $v \in E$ и многообразие $v_0 + L$ (рис. 8.4, *а*). Требуется найти разложение $v = m + h$, где $m \in v_0 + L$, $h \in L^{\perp}$.

Здесь h – перпендикуляр, опущенный из конца вектора v на многообразие $v_0 + L$. Заметим, что составляющие m и h в общем случае не ортогональны.

Поставленная задача сводится к задаче нахождения ортогональной проекции $l = m - v_0$ и ортогональной составляющей h вектора $v - v_0$ относительно подпространства L (см. рис. 8.4, *б*). Найдя ортогональное разложение $v - v_0 = l + h$, можно получить и искомое разложение $v = m + h$, где $m = l + v_0$.

23. Определитель Грама, его свойства

$$G(e_1, \ldots, e_n) = \begin{pmatrix} (e_1, e_1) & \cdots & (e_1, e_n) \\ \vdots & \ddots & \vdots \\ (e_n, e_1) & \cdots & (e_n, e_n) \end{pmatrix}, \tag{8.33}$$

которая называется ***матрицей Грама*** системы векторов $e_1, \ldots, e_n$.

Определитель матрицы (8.33) называется ***определителем Грама***. Рассмотрим свойства этого определителя

1. Критерий Грама линейной зависимости векторов: *система векторов $v_1, \ldots, v_k$ линейно зависима тогда и только тогда, когда определитель Грама этой системы равен нулю.*

Действительно, если система $v_1, v_2, \ldots, v_k$ линейно зависима, то существуют такие числа $x_1, \ldots, x_k$, не равные нулю одновременно, что

$$x_1 v_1 + \ldots + x_k v_k = o.$$

Умножая это равенство скалярно на v_1, затем на v_2 и т.д. на v_k, получаем однородную систему уравнений $G(v_1, v_2, \ldots, v_k) \cdot x = o$, которая имеет нетривиальное решение $x = (x_1 \;\; \cdots \;\; x_k)^T$. Следовательно, ее определитель равен нулю (см. разд.5.5). Необходимость доказана. Достаточность доказывается, проводя рассуждения в обратном порядке.

Следствие. *Если какой-либо главный минор матрицы Грама равен нулю, то и определитель Грама равен нулю.*

Главный минор матрицы Грама системы $v_1, v_2, \ldots, v_k$ представляет собой определитель Грама подсистемы векторов. Если подсистема линейно зависима, то и вся система линейно зависима (см. разд.8.2.2).

2. *Определитель Грама* $\det G(v_1, \ldots, v_k)$ *не изменяется в процессе ортогонализации системы векторов $v_1, \ldots, v_k$.* Другими словами, *если в процессе ортогонализации векторов $v_1, \ldots, v_k$ получены векторы $w_1, w_2, \ldots, w_k$, то*

$$\det G(v_1, v_2, \ldots, v_k) = \det G(w_1, w_2, \ldots, w_k) = (w_1, w_1) \cdot (w_2, w_2) \cdot \ldots \cdot (w_k, w_k).$$

Действительно, в процессе ортогонализации (см. разд. 8.8.5) по векторам $v_1, v_2, \ldots, v_k$ последовательно строятся векторы

$$w_1 = v_1, \quad w_2 = v_2 - \alpha_{21} w_1, \ldots, \quad w_k = v_k - \sum_{j=1}^{k-1} \alpha_{kj} w_j.$$

После первого шага определитель Грама не изменяется

$$\det G(v_1, v_2, \ldots, v_k) = \det G(w_1, v_2, \ldots, v_k).$$

Выполним с определителем $\det G(w_1, v_2, \ldots, v_k)$ следующие преобразования. Прибавим ко второй строке первую, умноженную на число $(-\alpha_{21})$, а затем ко второму столбцу прибавим первый, умноженный на $(-\alpha_{21})$. Получим определитель $\det G(w_1, v_2 - \alpha_{21} w_1, \ldots, v_k) = \det G(w_1, w_2, v_3, \ldots, v_k)$. Так как при этих преобразованиях определитель не изменяется (см. разд. 2.3), то

$$\det G(v_1, v_2, \ldots, v_k) = \det G(w_1, v_2, \ldots, v_k) = \det G(w_1, w_2, v_3, \ldots, v_k).$$

Значит, после второго шага в процессе ортогонализации определитель не изменяется. Продолжая аналогично, получаем после k шагов:

$$\det G(v_1, v_2, \ldots, v_k) = \det G(w_1, w_2, \ldots, w_k).$$

Вычислим правую часть этого равенства. Матрица $G(w_1, w_2, \ldots, w_k)$ Грама ортогональной системы $w_1, w_2, \ldots, w_k$ векторов является диагональной, так как $(w_i, w_j) = 0$ при $i \ne j$. Поэтому ее определитель равен произведению элементов, стоящих на главной диагонали:

$$\det G(w_1, w_2, \ldots, w_k) = (w_1, w_1) \cdot (w_2, w_2) \cdot \ldots \cdot (w_k, w_k).$$

3. *Определитель Грама любой системы $v_1, v_2, \ldots, v_k$ векторов удовлетворяет двойному неравенству*

$$0 \le \det G(v_1, v_2, \ldots, v_k) \le (v_1, v_1) \cdot (v_2, v_2) \cdot \ldots \cdot (v_k, v_k).$$

Докажем неотрицательность определителя Грама. Если система $v_1, v_2, \ldots, v_k$ линейно зависима, то определитель равен нулю (по свойству 1). Если же система $v_1, v_2, \ldots, v_k$ линейно независима, то, выполнив процесс ортогонализации, получим ненулевые векторы $w_1, w_2, \ldots, w_k$, для которых по свойству 2:

$$\det G(v_1, v_2, \ldots, v_k) = \det G(w_1, w_2, \ldots, w_k) = |w_1|^2 \cdot |w_2|^2 \cdot \ldots \cdot |w_k|^2 > 0.$$

Оценим теперь скалярный квадрат (v_j, v_j). Выполняя процесс ортогонализации, имеем $v_j = w_j + \alpha_{j1} w_1 + \ldots + \alpha_{j\,j-1} w_{j-1}$. Отсюда

$$(v_j, v_j) = (w_j, w_j) + \sum_{i=1}^{j-1} \alpha_{ji}^2 (w_i, w_i) \ge (w_j, w_j).$$

Следовательно, по свойству 2:

$$(v_1, v_1) \cdot (v_2, v_2) \cdot \ldots \cdot (v_k, v_k) \ge (w_1, w_1) \cdot (w_2, w_2) \cdot \ldots \cdot (w_k, w_k) = \det G(w_1, w_2, \ldots, w_k).$$

З а м е ч а н и я 8.12.

1. Матрица Грама любой системы векторов является неотрицательно определенной (см. разд. 6.6.3), так как все ее главные миноры также являются определителями Грама соответствующих подсистем векторов и неотрицательны в силу свойства 3.

2. Матрица Грама любой линейно независимой системы векторов является положительно определенной, так как все ее угловые миноры положительны (в силу свойств 1,3), поскольку являются определителями Грама линейно независимых подсистем векторов.

24. Метрические приложения определителя Грама.

МЕТРИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПРЕДЕЛИТЕЛЯ ГРАМА

Пусть $v_1, v_2, \ldots, v_k$ – линейно независимая система векторов n-мерного евклидова пространства ($k \le n$). Определим по индукции понятие ***многомерного объема***. Обозначим через h_j – перпендикуляр, опущенный из конца вектора v_j на подпространство $Lin(v_1, \ldots, v_{j-1})$, $j = 2, \ldots, k$.

Обозначим

$V_{v_1} = |v_1|$ – одномерный объем – длина вектора v_1;

$V_{v_1, v_2} = V_{v_1} \cdot |h_2| = |v_1| \cdot |h_2|$ – двумерный объем – площадь параллелограмма, построенного на векторах v_1, v_2;

$V_{v_1, v_2, v_3} = V_{v_1, v_2} \cdot |h_3| = |v_1| \cdot |h_2| \cdot |h_3|$ – трехмерный объем – объем параллелепипеда, построенного на векторах v_1, v_2, v_3;

. . .

$V_{v_1, \ldots, v_k} = V_{v_1, \ldots, v_{k-1}} \cdot |h_k| = |v_1| \cdot |h_2| \cdot \ldots \cdot |h_k|$ – k-мерный объем – объем параллелепипеда, построенного на векторах $v_1, v_2, \ldots, v_k$.

Проводя ортогонализацию системы векторов $v_1, v_2, \ldots, v_k$, получаем, согласно п.4 замечаний 8.14, перпендикуляры $h_1 = v_1, h_2, \ldots, h_k$. Тогда по свойству 2 определителя Грама имеем

$$V^2_{v_1, \ldots, v_k} = |h_1|^2 \cdot |h_2|^2 \cdot \ldots \cdot |h_k|^2 = \det G(v_1, \ldots, v_k), \tag{8.37}$$

т.е. *определитель Грама векторов $v_1, v_2, \ldots, v_k$ равен квадрату k-мерного объема параллелепипеда, построенного на этих векторах.* В этом заключается геометрический смысл определителя Грама.

Расстоянием от конца вектора v до подпространства L называется наименьшее значение длин векторов $(v - l)$, где $l \in L$, т.е.

$$d = \min_{l \in L} |v - l|.$$

Аналогично определяется расстояние от конца вектора до многообразия.

Углом между ненулевым вектором v и подпространством L называется наименьший угол между вектором v и ненулевыми векторами подпространства, т.е.

$$\varphi = \min_{l \in L}\left(\arccos \frac{(v, l)}{|v| \cdot |l|}\right).$$

Аналогично определяется угол между вектором и многообразием, как угол между вектором и однородной частью многообразия.

Из неравенств п.1 замечаний 8.14 следует, что

1) расстояние от конца вектора v до подпространства L равно длине перпендикуляра, опущенного из конца вектора v на подпространство L, т.е. $d = |h|$;

2) угол между ненулевым вектором v и подпространством L равен углу между вектором v и его ортогональной проекцией на подпространство L.

Для нахождения расстояний и углов можно использовать формулу (8.37).

Пусть задан вектор v и подпространство $L = Lin(e_1, \ldots, e_r)$, причем векторы $e_1, \ldots, e_r$ линейно независимы. Тогда $V_{e_1, \ldots, e_r, v} = V_{e_1, \ldots, e_r} \cdot |h|$, где h – ортогональная составляющая вектора v относительно подпространства L. Отсюда, $|h| = \dfrac{V_{e_1, \ldots, e_r, v}}{V_{e_1, \ldots, e_r}}$. Используя (8.37) для вычисления объемов, получаем, что длина $|h|$ ортогональной составляющей (расстояние от конца вектора v до подпространства $L = Lin(e_1, \ldots, e_r)$) находится по формуле

$$|h| = \sqrt{\frac{\det G(e_1, \ldots, e_r, v)}{\det G(e_1, \ldots, e_r)}}, \tag{8.38}$$

а угол φ между ненулевым вектором v и подпространством находится по формуле

$$\varphi = \arcsin \frac{|h|}{|v|}. \tag{8.39}$$

Пример 8.22. В пространстве R^4 со стандартным скалярным произведением (8.27) заданы: вектор $v = (-3 \;\; 2 \;\; 0 \;\; 0)^T$ и подпространство L – множество решений однородной системы:

$$\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 0, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 0. \end{cases}$$

Требуется найти расстояние $|h|$ от конца вектора v до подпространства L и угол между вектором v и подпространством L.

☐ Базис подпространства был найден в примере 8.9: $L = Lin(\varphi_1, \varphi_2)$, где $\varphi_1 = (-6 \;\; 4 \;\; 1 \;\; 0)^T$, $\varphi_1 = (-2 \;\; 1 \;\; 0 \;\; 1)^T$. Составляем определители Грама ($(v, v) = |v|^2 = (-3)^2 + 2^2 + 0^2 + 0^2 = 13$, остальные скалярные произведения векторов найдены в примере 8.20):

$$\det G(\varphi_1, \varphi_2, v) = \begin{vmatrix} 53 & 16 & 26 \\ 16 & 6 & 8 \\ 26 & 8 & 13 \end{vmatrix} = 14, \qquad \det G(\varphi_1, \varphi_2) = \begin{vmatrix} 53 & 16 \\ 16 & 6 \end{vmatrix} = 62.$$

Тогда $|h| = \sqrt{\dfrac{14}{62}} = \sqrt{\dfrac{7}{31}}$, а $\varphi = \arcsin \dfrac{\sqrt{\frac{7}{31}}}{\sqrt{13}} = \arcsin \sqrt{\dfrac{7}{403}}$. В примере 8.20 найдены ортогональная проекция $l = \left(\frac{-92}{31} \;\; \frac{60}{31} \;\; \frac{14}{31} \;\; \frac{4}{31}\right)^T$ и ортогональная составляющая $h = \left(\frac{-1}{31} \;\; \frac{2}{31} \;\; \frac{-14}{31} \;\; \frac{-4}{31}\right)^T$. Вычисляя длину вектора h, получаем $|h| = \sqrt{\dfrac{7}{31}}$. Результаты совпадают. ■

25. Неравенства Адамара, Бесселя.

3. Определитель квадратной матрицы A (n-го порядка) удовлетворяет *неравенству Адамара*:

$$(\det A)^2 \le \prod_{i=1}^{n}\left(a_{i1}^2 + a_{i2}^2 + \ldots + a_{in}^2\right).$$

Действительно, обозначив $a_1,\ldots,a_n$ столбцы матрицы A, элементы матрицы A^TA можно представить как скалярные произведения (8.27): $(a_i, a_j) = (a_i)^T a_j$. Тогда $A^TA = G(a_1,\ldots,a_n)$ – матрица Грама системы $a_1,\ldots,a_n$ векторов пространства R^n. По свойству 3, теореме 2.2 и свойству 1 определителя (см. разд.2.3.1) получаем доказываемое неравенство:

$$(\det A)^2 = \det A\cdot\det A = \det A^T\cdot\det A = \det\left(A^TA\right) =$$

$$= \det G(a_1,\ldots,a_n) \le |\,a_1\,|^2\cdot\ldots\cdot|\,a_n\,|^2 = \prod_{i=1}^{n}\left(a_{i1}^2 + a_{i2}^2 + \ldots + a_{in}^2\right).$$

3. Если в подпространстве L взять ортонормированный базис $e_1,\ldots,e_r$, то квадрат длины вектора l можно вычислить по формуле $|\,l\,|^2 = l_1^2 + l_2^2 + \ldots + l_r^2$, где $l_i = (l, e_i)$, $i = 1,\ldots,r$. Тогда из неравенства (см. п.1) $|\,l\,| \le |\,v\,|$ следует:

$$\sum_{i=1}^{r} l_i^2 \le |\,v\,|^2 \qquad (\textit{неравенство Бесселя}),$$

т.е. квадрат длины вектора не меньше суммы квадратов длин его проекций на любые r взаимно ортогональных направлений.

26. Отображения: определение, образ, полный прообраз. Сюръективные, инъективные, биективные, тождественные и обратимые отображения.

9.1. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

9.1.1. Определение линейных отображений

Напомним основные определения [19,25,43], связанные с понятием отображения (функции, оператора).

Пусть V и W – заданные множества. Говорят, что на множестве V ***определено отображение (функция)*** f, если каждому элементу $v \in V$ поставлен в соответствие единственный элемент $f(v)$ множества W. Такое соответствие называют также ***отображением множества*** V ***в множество*** W и обозначают $f: V \to W$, или $V \xrightarrow{f} W$. Если отображение f элементу $v \in V$ ставит в соответствие элемент $w \in W$, т.е. $w = f(v)$, то элемент w называется ***образом*** v, а элемент v – ***прообразом*** w.

Два отображения $f: V \to W$ и $g: V \to W$ называются ***равными***, если $f(v) = g(v)\ \ \forall v \in V$.

Отображение $f: V \to W$ называется:

инъективным, если разным элементам множества V соответствуют разные образы: $v_1 \ne v_2 \Rightarrow f(v_1) \ne f(v_2)$;

сюръективным, если для каждого элемента из множества W имеется хотя бы один прообраз: $\forall w \in W\ \exists v \in V: w = f(v)$;

биективным (взаимно однозначным), если оно инъективно и сюръективно одновременно.

Сюръективное отображение называется также ***отображением множества*** V ***на множество*** W.

Композицией отображений $g: U \to V$ и $f: V \to W$ называется отображение $f \circ g: U \to W$, определяемое равенством $(f \circ g)(u) = f(g(u))$.

Отображение $\mathcal{E}_V: V \to V$ называется ***тождественным***, если каждому элементу множества V ставится в соответствие этот же элемент: $\mathcal{E}_V(v) = v$ $\forall v \in V$.

Отображение $f^{-1}: W \to V$ называется ***обратным*** для отображения $f: V \to W$, если $f^{-1} \circ f = \mathcal{E}_V: V \to V$ и $f \circ f^{-1} = \mathcal{E}_W: W \to W$. Отображение f называется ***обратимым***, если для него существует обратное отображение. Необходимым и достаточным условием обратимости является условие биективности (взаимной однозначности) отображения.

Пусть V и W – линейные пространства (над одним и тем же числовым полем). Отображение $\mathcal{A}: V \to W$ называется ***линейным***, если

1. $\mathcal{A}(v_1 + v_2) = \mathcal{A}(v_1) + \mathcal{A}(v_2)\ \ \forall v_1 \in V,\ \forall v_2 \in V$;
2. $\mathcal{A}(\lambda\cdot v) = \lambda\cdot\mathcal{A}(v)\ \ \forall v \in V$ и любого числа λ (из данного числового поля).

Условие 1 называется ***аддитивностью*** отображения, а условие 2 – ***однородностью***. Пространство V называется ***пространством прообразов***, а пространство W – ***пространством образов***.

Замечания 9.1.

1. Линейное отображение $\mathcal{A}: V \to W$ нулевому элементу o_V пространства V ставит в соответствие нулевой элемент o_W пространства W.

2. Условия аддитивности и однородности можно заменить одним условием ***линейности*** отображения:

$$\mathcal{A}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1\mathcal{A}(v_1) + \lambda_2\mathcal{A}(v_2)\quad \forall v_1 \in V,\ \forall v_2 \in V$$

и любых чисел λ_1 и λ_2 из данного числового поля.

3. При линейном отображении образ линейной комбинации является линейной комбинацией образов:

$$\mathcal{A}\left(\sum_{i=1}^{k}\lambda_i v_i\right) = \sum_{i=1}^{k}\lambda_i\,\mathcal{A}(v_i).$$

4. Взаимно однозначное линейное отображение является изоморфизмом (см. разд. 8.5).

9.1.2. Примеры линейных отображений

1. Обозначим $\mathcal{O}: V \to W$ – нулевое отображение, которое ставит в соответствие любому вектору $v \in V$ нулевой элемент o_W пространства W. Условия аддитивности и однородности такого отображения, разумеется, выполняются. Это отображение не является инъективным (разным прообразам v_1 и v_2 соответствует один и тот же образ o_W), не является сюръективным (из всех векторов пространства W только у нулевого имеется прообраз). Поэтому нулевое отображение не является биективным и, следовательно, обратимым.

27. Композиция отображений. Теорема об обратном отображении.

Композицией отображений $g: U \to V$ и $f: V \to W$ называется отображение $f \circ g: U \to W$, определяемое равенством $(f \circ g)(u) = f(g(u))$.

Отображение $\mathcal{E}_V: V \to V$ называется ***тождественным***, если каждому элементу множества V ставится в соответствие этот же элемент: $\mathcal{E}_V(v) = v$ $\forall v \in V$.

Отображение имеет обратное тогда и только тогда, когда является взаимно-однозначным (биективным) отображением.

Доказательство теоремы

> Предположим сначала, что отображение $f: X \to Y$ имеет обратное $f^{-1}: Y \to X$. Тогда из равенства $f^{-1} \circ f = e_X$, согласно лемме, следует, что f инъективно, а равенство $f \circ f^{-1} = e_Y$ означает, что f сюръективно, поэтому f является биекцией.

Обратно, предположим, что отображение f биективно. Тогда для любого элемента $y \in Y$ найдется единственный элемент x такой, что $f(x) = y$. Определим теперь отображение $g(y)$ соотношением:

$$\forall y \in Y,\quad g(y) = x, \qquad (2)$$

где x удовлетворяет равенству $f(x) = y$.

Покажем, что отображение $g(y)$ функционально. Предположим противное: пусть $g(y)$ определено равенством (1.3), но не является функциональным, т.е существует элемент $y_0 \in Y$, имеющий два различных образа:

$$g(y_0) = x_1 \text{ и } \quad g(y_0) = x_2 \text{, причем,} \quad x_1 \ne x_2.$$

Тогда, в силу определения $g(y)$, имеем: $f(x_1) = y_0 = f(x_2)$, и, следовательно, отображение f не является инъективным (последнее, очевидно, противоречит предположению о биективности f).

Наконец, докажем, что отображение g является обратным к отображению f. Для этого достаточно показать справедливость двух равенств: $g \circ f = e_X$ и $f \circ g = e_Y$. В самом деле: $\forall x \in X$, $g \circ f(x) = g(f(x)) = g(y) = x$, и значит $g \circ f = e_X$. Точно так же $\forall y \in Y$, $f \circ g(y) = f(g(y)) = f(x) = y$, поэтому $f \circ g = e_Y$.

28. Линейные отображения: определение, примеры, свойства.

9.1.2. Примеры линейных отображений

1. Обозначим $\mathcal{O}: V \to W$ – нулевое отображение, которое ставит в соответствие любому вектору $v \in V$ нулевой элемент o_W пространства W. Условия аддитивности и однородности такого отображения, разумеется, выполняются. Это отображение не является инъективным (разным прообразам v_1 и v_2 соответствует один и тот же образ o_W), не является сюръективным (из всех векторов пространства W только у нулевого имеется прообраз). Поэтому нулевое отображение не является биективным и, следовательно, обратимым.

2. Пусть в n-мерном линейном пространстве V задан базис $e_1,\ldots,e_n$. Обозначим $\ae: V \to R^n$ отображение, которое ставит в соответствие каждому вектору v его координатный столбец $v = (v_1\ \cdots\ v_n)^T$ относительно заданного базиса. Такое отображение является линейным, так как при сложении векторов в одном и том же базисе их координаты складываются, а при умножении вектора на число – координаты вектора умножаются на это число (см. разд. 8.4.2). Это отображение является инъективным (разные векторы имеют разные координаты (в одном и том же базисе)), является сюръективным (для любого столбца $v = (v_1\ \cdots\ v_n)^T \in R^n$ существует прообраз $v = v_1e_1 + \ldots + v_ne_n$). Поэтому отображение $\ae$ биективное и, следовательно, обратимое. Напротив, отображение, которое каждому вектору $v \in V$ ставит в соответствие столбец $v = (v_1 + 1\ \cdots\ v_n + 1)^T \in R^n$ не является линейным, так как образом нулевого вектора $o_V \in V$ служит столбец $(1\ \cdots\ 1)^T \ne o$, отличный от нулевого.

3. Пусть в n-мерном евклидовом пространстве E задан ненулевой вектор $e \in E$. Обозначим $\text{пр}_e(v) = \dfrac{(e,v)}{|e|}$ – алгебраическое значение проекции вектора $v \in E$ на направление, задаваемое вектором e. Тогда отображение $\text{пр}_e: E \to R$ будет линейным, так как скалярное произведение линейно по второму сомножителю (см. разд. 8.8.1). Это отображение не является инъективным (разные векторы могут иметь одну и ту же проекцию), является сюръективным (для любого действительного числа λ, задающего величину проекции, найдется прообраз, например вектор $\dfrac{\lambda\cdot e}{|e|}$). Поэтому отображение не является биективным и, следовательно, обратимым. Отображение $E \to R$, которое каждому вектору $v \in E$ ставит в соответствие его длину $|\,v\,| \in R$ не является линейным, поскольку не выполняется, например, условие однородности: $|\,\lambda v\,| \ne \lambda|\,v\,|$ для отрицательных λ.

4. Пусть $P_n(R)$ и $P_{n-1}(R)$ – пространства многочленов с действительными коэффициентами степени не выше n или $(n-1)$ соответственно. Обозначим через $\mathcal{D}(p(x)) = \dfrac{dp(x)}{dx}$ производную многочлена $p(x) \in P_n(R)$. Тогда отображение (оператор дифференцирования) $\mathcal{D}: P_n(R) \to P_{n-1}(R)$ ставит в соответствие каждому многочлену $p(x) \in P_n(R)$ его производную, т.е. многочлен из пространства $P_{n-1}(R)$. Этот оператор линейный, так как производная суммы равна сумме производных, а производная произведения функции на число равна произведению производной на это число [19,25,43]. Оператор дифференцирования не является инъективным (два многочлена, отличающиеся свободными членами имеют одну и ту же производную), является сюръективным (для любого многочлена $p_{n-1}(x)$ имеется прообраз – многочлен из множества первообразных $\int p_{n-1}(x)\,dx + C$, где C – произвольная постоянная). Поэтому оператор дифференцирования не является биективным и, следовательно, обратимым. Оператор интегрирования $\mathcal{I}: P_{n-1}(R) \to P_n(R)$, который многочлену $p_{n-1}(x) \in P_{n-1}(R)$ ставит в соответствие многочлен $p_n(x) = \int_0^x p_{n-1}(t)\,dt$, также является линейным (см. свойства интеграла в [19,25,43]). Этот оператор является инъективным (из равенства образов, дифференцируя по верхнему пределу интегрирования, получаем равенство прообразов), не является сюръективным (многочлен с отличным от нуля свободным членом не имеет прообраза). Поэтому оператор интегрирования не является биективным и, следовательно, обратимым.

9.1.3. Свойства линейных отображений

Пусть $\mathcal{A}: V \to W$ – линейное отображение.

1. *Если векторы $v_1,\ldots,v_k$ линейно зависимы, то их образы также линейно зависимы.*

Действительно, если нетривиальная линейная комбинация равна нулевому вектору: $\lambda_1v_1 + \ldots + \lambda_kv_k = o_V$, то, применяя к обеим частям отображение $\mathcal{A}$, в силу его линейности с учетом п.1,3 замечаний 9.1, получаем

$$\lambda_1\mathcal{A}(v_1) + \ldots + \lambda_k\mathcal{A}(v_k) = o_W,$$

т.е. равную нулевому вектору нетривиальную линейную комбинацию образов заданных векторов. Значит, образы $\mathcal{A}(v_1),\ldots,\mathcal{A}(v_k)$ заданных векторов линейно зависимы.

2. *Пусть $\mathcal{A}: V \to W$ – сюръективное отображение пространства V на пространство W и векторы $w_1,\ldots,w_k$ пространства W образуют линейно независимую систему. Тогда в пространстве V существует такая линейно независимая система векторов $v_1,\ldots,v_k$, что $\mathcal{A}(v_i) = w_i$, $i = 1,\ldots,k$.*

Действительно, в силу сюръективности отображения у векторов $w_1,\ldots,w_k$ найдутся прообразы $v_1,\ldots,v_k$. Если система $v_1,\ldots,v_k$ линейно зависима, то и система $w_1,\ldots,w_k$ была бы линейно зависимой (по свойству 1). Поэтому найденная система прообразов линейно независимая.

3. *При линейном сюръективном отображении $\mathcal{A}: V \to W$ конечномерного пространства размерность пространства образов не превосходит размерности пространства прообразов*, т.е. $\dim W \le \dim V$.

В самом деле, в пространстве образов W нет линейно независимой системы из большего, чем $\dim V$, количества векторов. Если бы такая система векторов была, то прообразы этих векторов были бы линейно независимы (по свойству 2). Но в пространстве V не может быть линейно независимой системы из большего, чем $\dim V$, количества векторов.

4. *Композиция линейных отображений является линейным отображением.*

Действительно, пусть $\mathcal{C} = \mathcal{B} \circ \mathcal{A}$ – композиция линейных отображений $\mathcal{A}: U \to V$ и $\mathcal{B}: V \to W$. Тогда отображение $\mathcal{C}$ аддитивно:

$$\mathcal{C}(u_1 + u_2) = (\mathcal{B} \circ \mathcal{A})(u_1 + u_2) = \mathcal{B}(\mathcal{A}(u_1 + u_2)) = \mathcal{B}(\mathcal{A}(u_1) + \mathcal{A}(u_2)) =$$

$$= \mathcal{B}(\mathcal{A}(u_1)) + \mathcal{B}(\mathcal{A}(u_2)) = \mathcal{C}(u_1) + \mathcal{C}(u_2).$$

Однородность отображения $\mathcal{C}$ доказывается аналогично.

5. *Если линейное отображение $\mathcal{A}: V \to W$ обратимое (взаимно однозначное), то обратное отображение $\mathcal{A}^{-1}: W \to V$ – линейное.*

Докажем, например, аддитивность обратного отображения

$$\mathcal{A}^{-1}(w_1 + w_2) = \mathcal{A}^{-1}(w_1) + \mathcal{A}^{-1}(w_2).$$

Обозначим $v = \mathcal{A}^{-1}(w_1) + \mathcal{A}^{-1}(w_2)$. Тогда в силу линейности $\mathcal{A}$, получаем

$$\mathcal{A}(v) = \mathcal{A}\left(\mathcal{A}^{-1}(w_1) + \mathcal{A}^{-1}(w_2)\right) = \mathcal{A}\left(\mathcal{A}^{-1}(w_1)\right) + \mathcal{A}\left(\mathcal{A}^{-1}(w_2)\right) =$$

$$= \left(\mathcal{A} \circ \mathcal{A}^{-1}\right)(w_1) + \left(\mathcal{A} \circ \mathcal{A}^{-1}\right)(w_2) = \mathcal{E}_W(w_1) + \mathcal{E}_W(w_2) = w_1 + w_2.$$

Следовательно, $\mathcal{A}^{-1}(w_1 + w_2) = v = \mathcal{A}^{-1}(w_1) + \mathcal{A}^{-1}(w_2)$, что и требовалось доказать. Однородность обратного отображения доказывается аналогично.

6. *Линейное отображение конечномерного пространства однозначно задается образами базисных векторов.*

В самом деле, пусть $e_1,\ldots,e_n$ – базис пространства V, а $f_1,\ldots,f_n$ – произвольная система векторов пространства W. Докажем, что существует единственное линейное отображение $\mathcal{A}: V \to W$, удовлетворяющее условиям $\mathcal{A}(e_i) = f_i$, $i = 1,\ldots,n$.

Рассмотрим отображение $\mathcal{A}(v) = \sum_{i=1}^{n} v_i f_i$, где $v_1,\ldots,v_n$ – координаты вектора v в заданном базисе: $v = v_1e_1 + \ldots + v_ne_n$. Это отображение удовлетворяет заданным условиям, так как $\mathcal{A}(e_i) = f_i$. Покажем, что оно аддитивное и однородное:

$$\mathcal{A}(u + v) = \mathcal{A}\left(\sum_{i=1}^{n}(u_i + v_i)e_i\right) = \sum_{i=1}^{n}(u_i + v_i)f_i = \sum_{i=1}^{n}u_i f_i + \sum_{i=1}^{n}v_i f_i = \mathcal{A}(u) + \mathcal{A}(v);$$

$$\mathcal{A}(\lambda v) = \mathcal{A}\left(\sum_{i=1}^{n}(\lambda v_i)e_i\right) = \sum_{i=1}^{n}(\lambda v_i)f_i = \lambda\sum_{i=1}^{n}v_i f_i = \lambda\,\mathcal{A}(v).$$

Существование доказано. Единственность докажем от противного. Пусть $\mathcal{B}$ – еще одно линейное отображение, удовлетворяющее условиям $\mathcal{B}(e_i) = f_i$, $i = 1,\ldots,n$. Для любого вектора $v = v_1e_1 + \ldots + v_ne_n$ имеем

$$\mathcal{B}(v) = \mathcal{B}\left(\sum_{i=1}^{n}v_ie_i\right) = \sum_{i=1}^{n}v_i\mathcal{B}(e_i) = \sum_{i=1}^{n}v_if_i = \mathcal{A}(v).$$

Следовательно, $\mathcal{B} = \mathcal{A}$.

29. Матрицы линейных отображений и их свойства.

9.1.4. Матрица линейного отображения

Пусть $\mathcal{A}: V \to W$ – линейное отображение n-мерного пространства V в m-мерное пространство W. Зафиксируем в пространстве V произвольный базис $(e)=(e_1,\ldots,e_n)$, а в пространстве W базис $(f)=(f_1,\ldots,f_m)$. Линейное отображение однозначно задается образами базисных векторов (см. свойство 6). Разложим образы $\mathcal{A}(e_i)$, $i=1,\ldots,n$, базисных векторов (e) по базису (f):

$$\mathcal{A}(e_i)=\sum_{j=1}^{m} a_{ji} f_j, \quad i=1,\ldots,n.$$

Из координатных столбцов векторов $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$ относительно базиса (f) составим матрицу размеров $m\times n$:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \qquad (9.1)$$

Она называется ***матрицей линейного отображения*** $\mathcal{A}$ ***в базисах*** (e) и (f). Матрицу отображения обозначают также $\underset{(e)(f)}{A}$, чтобы подчеркнуть ее зависимость от выбранных базисов.

При помощи матрицы отображения найдем координаты образа $w=\mathcal{A}(v)$ по координатам прообраза v. Пусть $v=(v_1 \;\cdots\; v_n)^T$ – координатный столбец вектора v, а $w=(w_1 \;\cdots\; w_m)^T$ – координатный столбец вектора w, т.е. $v=v_1e_1+\ldots+v_ne_n$ и $w=w_1f_1+\ldots+w_mf_m$. Тогда

$$w=\mathcal{A}(v)=\sum_{i=1}^{n} v_i\mathcal{A}(e_i)=\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji}v_i f_j.$$

В силу единственности разложения вектора w по базису (f) получаем

$$w_j=\sum_{i=1}^{n} a_{ji}v_i, \quad j=1,\ldots,m.$$

Используя матричные операции, связь координат можно записать в виде

$$\begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \Leftrightarrow \quad w=Av, \qquad (9.2)$$

где A – матрица (9.1) отображения $\mathcal{A}$.

Таким образом, *для каждого линейного отображения n-мерного пространства V в m-мерное пространство W (с фиксированными базисами (e) и (f) соответственно) определена единственная матрица* (9.1) *этого отображения*, и наоборот, *любая числовая матрица размеров $m\times n$ является матрицей некоторого линейного отображения n-мерного пространства V в m-мерное пространство W*.

Для нахождения матрицы отображения $\mathcal{A}: V \to W$ нужно выполнить следующие действия:

1) зафиксировать базисы $(e)=(e_1,\ldots,e_n)$ и $(f)=(f_1,\ldots,f_m)$ пространств V и W;

2) найти образ $\mathcal{A}(e_1)$ первого базисного вектора и разложить его по базису (f). Полученные координаты записать в первый столбец матрицы (9.1) отображения $\mathcal{A}$;

3) найти образ $\mathcal{A}(e_2)$ второго базисного вектора и разложить его по базису (f). Полученные координаты записать во второй столбец матрицы (9.1) отображения и т.д. В последний столбец матрицы (9.1) записать координаты образа $\mathcal{A}(e_n)$ последнего базисного вектора.

Найдем матрицы отображений, рассмотренных в разд.9.1.2.

1. Матрица нулевого отображения $\mathcal{O}: V\to W$ нулевая относительно любых базисов пространств V и W, так как образ любого базисного вектора равен нулевому вектору o_W, координаты которого равны нулю (относительно любого базиса пространства W).

2. Пусть в n-мерном линейном пространстве V задан базис $e_1,\ldots,e_n$. Рассмотрим отображение $æ: V\to R^n$, которое ставит в соответствие каждому вектору $v=v_1e_1+\ldots+v_ne_n$ его координатный столбец $v=(v_1 \;\cdots\; v_n)^T$ относительно заданного базиса. В пространстве R^n выберем стандартный базис $e_1,\ldots,e_n$ (см. п.3 в разд. 8.3.2). Напомним, что в стандартном базисе координатный столбец вектора $x=(x_1 \;\cdots\; x_n)^T$ совпадает с самим столбцом x, так как

$$x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=x_1\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}+x_2\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}+\ldots+x_n\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}=x_1e_1+x_2e_2+\ldots+x_ne_n.$$

Поэтому образ $æ(e_1)$ первого базисного вектора e_1 имеет координатный столбец $e_1=(1 \;\; 0 \;\cdots\; 0)^T$, совпадающий с первым базисным вектором $e_1\in R^n$. Образ $æ(e_2)=e_2$ и т.д. Составляя из этих столбцов матрицу отображения $æ: V\to R^n$, получаем единичную матрицу E n-го порядка.

3. В n-мерном евклидовом пространстве E возьмем ортонормированный базис $e_1,\ldots,e_n$. В качестве базиса одномерного линейного пространства R возьмем единицу. Рассмотрим отображение $\text{пр}_{e_1}: E\to R$, где $\text{пр}_{e_1}(v)=(e_1,v)$ – алгебраическое значение проекции вектора v на направление, задаваемое вектором e_1. Тогда матрица отображения пр_{e_1} имеет вид $(1 \;\; 0 \;\cdots\; 0)$, так как $\text{пр}_{e_1}(e_1)=(e_1,e_1)=1$, а $\text{пр}_{e_1}(e_i)=(e_1,e_i)=0$ для $i\ne 1$.

4. Взяв в пространствах $P_n(R)$ и $P_{n-1}(R)$ стандартные базисы (см. п.6 в разд.8.3.2), находим образы базисных векторов (первые производные многочленов):

$$\mathcal{D}(1)=0=0\cdot 1+0\cdot x+0\cdot x^2+\ldots+0\cdot x^{n-1};$$
$$\mathcal{D}(x)=1=1\cdot 1+0\cdot x+0\cdot x^2+\ldots+0\cdot x^{n-1};$$
$$\mathcal{D}(x^2)=2x=0\cdot 1+2\cdot x+0\cdot x^2+\ldots+0\cdot x^{n-1};$$
$$\ldots$$
$$\mathcal{D}(x^n)=n\,x^{n-1}=0\cdot 1+0\cdot x+0\cdot x^2+\ldots+n\cdot x^{n-1}.$$

Записывая найденные координаты по столбцам матрицы отображения, получаем матрицу размеров $n\times(n+1)$:

$$D=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \end{pmatrix}.$$

СВОЙСТВА МАТРИЦ ЛИНЕЙНЫХ ОТОБРАЖЕНИЙ

При фиксированных базисах линейных пространств:

1) *матрица суммы линейных отображений равна сумме их матриц;*

2) *матрица произведения линейного отображения на число равна произведению матрицы отображения на то же самое число;*

3) *матрица обратного отображения является обратной для матрицы отображения;*

4) *матрица композиции $C=\mathcal{B}\circ\mathcal{A}$ отображений равна произведению матриц отображений:* $C=BA$

Докажем, например, последнее свойство. Пусть в линейных пространствах V, W, U фиксированы базисы $(e),(f),(g)$ соответственно. Отображения $\mathcal{A}: V\to W$, $\mathcal{B}: W\to U$, а также их композиция $\mathcal{C}=\mathcal{B}\circ\mathcal{A}$, имеют матрицы A, B, C относительно соответствующих базисов. Для координатных столбцов v, w, u векторов $v\in V$, $w=\mathcal{A}(v)$, $u=\mathcal{B}(w)=\mathcal{C}(v)$ запишем связи (9.2): $w=Av$, $u=Bw$, $u=Cv$. Тогда $Cv=Bw=BAv$ для координатного столбца v произвольного вектора $v\in V$. Отсюда следует, что $C=BA$.

30. Ядро и образ линейного отображения: определение, примеры, свойства.

9.1.5. Ядро и образ линейного отображения

Ядром линейного отображения $\mathcal{A}: V\to W$ называется множество таких векторов $v\in V$, что $\mathcal{A}(v)=o_W$, т.е. множество векторов из V, которые отображаются в нулевой вектор пространства W. Ядро отображения $\mathcal{A}: V\to W$ обозначается:

$$\mathbf{Ker}\,\mathcal{A}=\{v: v\in V, \mathcal{A}(v)=o_W\}.$$

Образом линейного отображения $\mathcal{A}: V\to W$ называется множество образов $\mathcal{A}(v)$ всех векторов v из V. Образ отображения $\mathcal{A}: V\to W$ обозначается $\mathbf{Im}\,\mathcal{A}$ или $\mathcal{A}(V)$:

$$\mathbf{Im}\,\mathcal{A}=\mathcal{A}(V)=\{w: w=\mathcal{A}(v),\ \forall v\in V\}.$$

Заметим, что символ $\mathbf{Im}\,\mathcal{A}$ следует отличать от $\mathrm{Im}\,z$ – мнимой части комплексного числа.

ПРИМЕРЫ ЯДЕР И ОБРАЗОВ ЛИНЕЙНЫХ ОТОБРАЖЕНИЙ

1. Ядром нулевого отображения $\mathcal{O}: V\to W$ является все пространство V, а образом служит один нулевой вектор, т.е. $\mathbf{Ker}\,\mathcal{O}=V$, $\mathbf{Im}\,\mathcal{O}=\{o_W\}$.

2. Рассмотрим отображение $æ: V\to R^n$, которое ставит в соответствие каждому вектору v n-мерного линейного пространства V его координатный столбец $v=(v_1 \;\cdots\; v_n)^T$ относительно заданного базиса $e_1,\ldots,e_n$. Ядром этого отображения является нулевой вектор o_V пространства V, поскольку только этот вектор имеет нулевой координатный столбец $æ(o_V)=o\in R^n$. Образ преобразования $æ$ совпадает со всем пространством R^n, так как это преобразование сюръективно (любой столбец из R^n является координатным столбцом некоторого вектора пространства V).

3. Рассмотрим отображение $\text{пр}_e: E\to R$, которое каждому вектору v n-мерного евклидова пространства E ставит в соответствие алгебраическое значение $\text{пр}_e(v)=(e,v)$ его проекции на направление, задаваемое единичным вектором e. Ядром этого преобразования является ортогональное дополнение $\{e\}^{\perp}$ – множество векторов, ортогональных e. Образом является все множество действительных чисел R.

4. Рассмотрим отображение $\mathcal{D}: P_n(R)\to P_{n-1}(R)$, которое каждому многочлену степени не выше n ставит в соответствие его производную. Ядром этого отображения является множество $P_0(R)$ многочленов нулевой степени, а образом – все пространство $P_{n-1}(R)$.

СВОЙСТВА ЯДРА И ОБРАЗА ЛИНЕЙНОГО ОТОБРАЖЕНИЯ

1. *Ядро любого линейного отображения $\mathcal{A}: V\to W$ является подпространством:* $\{o_V\}\triangleleft \mathbf{Ker}\,\mathcal{A}\triangleleft V$.

В соответствии с определением (см. разд. 8.6.1) требуется доказать, что множество $\mathbf{Ker}\,\mathcal{A}$ является непустым и замкнутым относительно операций сложения векторов и умножения вектора на число. В самом деле, из однородности отображения следует, что

$$\mathcal{A}(o_V)=\mathcal{A}(0\cdot v)=0\cdot\mathcal{A}(v)=o_W,$$

т.е. нулевой вектор o_V отображается в нулевой вектор o_W. Следовательно, ядро любого линейного отображения не является пустым и содержит, по крайней мере, нулевой элемент: $o_V\in\mathbf{Ker}\,\mathcal{A}$. Покажем, что множество $\mathbf{Ker}\,\mathcal{A}$ замкнуто по отношению к операциям сложения векторов и умножения вектора на число. Действительно:

$$\left.\begin{array}{l} v_1\in\mathbf{Ker}\,\mathcal{A}\Rightarrow\mathcal{A}(v_1)=o_W \\ v_2\in\mathbf{Ker}\,\mathcal{A}\Rightarrow\mathcal{A}(v_2)=o_W \end{array}\right\}\Rightarrow\ \mathcal{A}(v_1+v_2)=\mathcal{A}(v_1)+\mathcal{A}(v_2)=o_W \ \Rightarrow$$

$$\Rightarrow\quad v_1+v_2\in\mathbf{Ker}\,\mathcal{A};$$

$$v\in\mathbf{Ker}\,\mathcal{A}\Rightarrow\mathcal{A}(v)=o_W\Rightarrow\mathcal{A}(\lambda\cdot v)=\lambda\cdot\mathcal{A}(v)=\lambda\cdot o_W=o_W\Rightarrow\lambda\cdot v\in\mathbf{Ker}\,\mathcal{A}.$$

Следовательно, множество $\mathbf{Ker}\,\mathcal{A}$ является линейным подпространством пространства V.

2. *Образ любого линейного отображения $\mathcal{A}: V\to W$ является подпространством:* $\mathbf{Im}\,\mathcal{A}\triangleleft W$.

В самом деле, докажем, например, замкнутость множества $\mathbf{Im}\,\mathcal{A}$ по отношению к операции умножения вектора на число. Если $w\in\mathbf{Im}\,\mathcal{A}$, то существует вектор $v\in V$ такой, что $w=\mathcal{A}(v)$. Тогда $\mathcal{A}(\lambda\cdot v)=\lambda\cdot\mathcal{A}(v)=\lambda\cdot w$, т.е. $\lambda\cdot w\in\mathbf{Im}\,\mathcal{A}$.

Поскольку ядро и образ линейного отображения являются линейными подпространствами (свойства 1 и 2), можно говорить об их размерностях.

Дефектом линейного отображения называется размерность его ядра: $d=\dim(\mathbf{Ker}\,\mathcal{A})$, а ***рангом линейного отображения*** – размерность его образа: $\mathrm{rg}\,\mathcal{A}=r=\dim(\mathbf{Im}\,\mathcal{A})$.

3. *Ранг линейного отображения равен рангу его матрицы (определенной относительно любых базисов).*

В самом деле, если $e_1,\ldots,e_n$ любой базис пространства V, то $\mathbf{Im}\,\mathcal{A}=Lin(\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n))$. Поэтому максимальное число линейно независимых векторов системы $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$ (ранг системы векторов) равно максимальному числу линейно независимых столбцов матрицы A отображения, т.е. рангу матрицы: $\dim(\mathbf{Im}\,\mathcal{A})=\mathrm{rg}\,A=\mathrm{rg}\,\mathcal{A}$.

4. *Линейное отображение $\mathcal{A}: V\to W$ инъективно тогда и только тогда, когда* $\mathbf{Ker}\,\mathcal{A}=\{o_V\}$, другими словами, *когда дефект отображения равен нулю:* $d=\dim(\mathbf{Ker}\,\mathcal{A})=0$.

Действительно, образом нулевого вектора o_V служит нулевой вектор o_W. Поэтому, если отображение инъективно, то ядро содержит только нулевой вектор o_V, иначе два разных вектора имели бы один и тот же образ o_W. Обратно, при условии $\mathbf{Ker}\,\mathcal{A}=\{o_V\}$ разные векторы $v_1\ne v_2$ не могут иметь одинаковые образы $\mathcal{A}(v_1)=\mathcal{A}(v_2)$, так как в этом случае из равенства $\mathcal{A}(v_1)-\mathcal{A}(v_2)=\mathcal{A}(v_1-v_2)=o_W$ следует, что ненулевой вектор $(v_1-v_2)\in\mathbf{Ker}\,\mathcal{A}$ (приходим к противоречию).

5. *Линейное отображение $\mathcal{A}: V\to W$ сюръективно тогда и только тогда, когда* $\mathbf{Im}\,\mathcal{A}=W$, другими словами, *когда ранг отображения равен размерности пространства образов:* $r=\dim(\mathbf{Im}\,\mathcal{A})=\dim W$.

6. *Линейное отображение $\mathcal{A}: V\to W$ биективно* (значит, *обратимо*) *тогда и только тогда, когда* $\mathbf{Ker}\,\mathcal{A}=\{o_V\}$ и $\mathbf{Im}\,\mathcal{A}=W$ *одновременно.*

31. Теорема о размерностях ядра и образа.

Теорема 9.1 (о размерностях ядра и образа). *Сумма размерностей ядра и образа любого линейного отображения $\mathcal{A}: V\to W$ равна размерности пространства прообразов:*

$$\dim(\mathbf{Ker}\,\mathcal{A})+\dim(\mathbf{Im}\,\mathcal{A})=\dim V. \qquad (9.3)$$

Действительно, пусть $d=\dim(\mathbf{Ker}\,\mathcal{A})$, $\dim V=n$. Выберем в подпространстве $\mathbf{Ker}\,\mathcal{A}\triangleleft V$ базис $e_1,\ldots,e_d$ и дополним его векторами $e_{d+1},\ldots,e_n$ до базиса $e_1,\ldots,e_n$ всего пространства V. Покажем, что векторы $\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n)$ образуют базис подпространства $\mathbf{Im}\,\mathcal{A}\triangleleft W$.

Во-первых, $\mathbf{Im}\,\mathcal{A}=Lin(\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n))$, так как образ любого вектора $v=v_1e_1+\ldots+v_de_d+v_{d+1}e_{d+1}+\ldots+v_ne_n$ линейно выражается через векторы $\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n)$:

$$\mathcal{A}(v)=\mathcal{A}\left(\sum_{i=1}^{n}v_i e_i\right)=\sum_{i=1}^{n}v_i\,\mathcal{A}(e_i)=\sum_{i=1}^{d}v_i\underbrace{\mathcal{A}(e_i)}_{o_W}+\sum_{i=d+1}^{n}v_i\,\mathcal{A}(e_i)=\sum_{i=d+1}^{n}v_i\,\mathcal{A}(e_i).$$

Во-вторых, образующие $\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n)$ линейно независимы. Если их линейная комбинация равна нулевому вектору: $o_W=\sum_{i=d+1}^{n}\lambda_i\,\mathcal{A}(e_i)=\mathcal{A}\left(\sum_{i=d+1}^{n}\lambda_i e_i\right)$, то вектор $\sum_{i=d+1}^{n}\lambda_i e_i$ принадлежит ядру (его образ – нулевой вектор). Однако, по построению этот вектор принадлежит алгебраическому дополнению $(\mathbf{Ker}\,\mathcal{A})^{+}$. Учитывая, что $\mathbf{Ker}\,\mathcal{A}\cap(\mathbf{Ker}\,\mathcal{A})^{+}=\{o_V\}$, заключаем: $\sum_{i=d+1}^{n}\lambda_i e_i=o_V$. Получили разложение нулевого вектора по линейно независимой системе $e_{d+1},\ldots,e_n$ векторов, значит, все коэффициенты $\lambda_i=0$. Поэтому равенство $\sum_{i=d+1}^{n}\lambda_i\,\mathcal{A}(e_i)=o_W$ справедливо только для тривиальной линейной комбинации, т.е. система векторов $\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n)$ линейно независимая.

Таким образом, векторы $\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n)$ образуют базис подпространства $\mathbf{Im}\,\mathcal{A}=Lin(\mathcal{A}(e_{d+1}),\ldots,\mathcal{A}(e_n))$, а его размерность определяется количеством базисных векторов, т.е. $\dim(\mathbf{Im}\,\mathcal{A})=n-d$, что равносильно (9.3).

Следствие. *Линейное отображение $\mathcal{A}: V\to W$ биективно* (значит, *обратимо*) *тогда и только тогда, когда обратима его матрица* (*определенная относительно любых базисов*).

Действительно, для обратимости преобразования $\mathcal{A}: V\to W$ (см. свойство 6) его матрица A (размеров $m\times n$) должна удовлетворять условиям (см. свойства 3, 4, 5):

$$\mathrm{rg}\,A=\mathrm{rg}\,\mathcal{A}=\dim(\mathbf{Im}\,\mathcal{A})=\dim W=m, \qquad d=\dim(\mathbf{Ker}\,\mathcal{A})=0.$$

Тогда по теореме 9.1 заключаем, что $m=n-d=n$, т.е. матрица A – квадратная n-го порядка и невырожденная ($\mathrm{rg}\,A=n$), что и требовалось доказать.

Обратимые линейные отображения называются также ***невырожденными*** (имея в виду невырожденность их матрицы).

32. Линейные преобразования: определение, примеры.

Линейным преобразованием (***линейным оператором***) линейного пространства V называется линейное отображение $\mathcal{A}: V\to V$ пространства V в себя.

Поскольку линейное преобразование является частным случаем линейного отображения, к нему применимы все понятия и свойства, рассмотренные для отображений: инъективность, сюръективность, биективность, обратимость, ядро, образ, дефект, ранг и т.д.

Матрицей линейного преобразования $\mathcal{A}: V\to V$ ***в базисе*** $e_1,\ldots,e_n$ ***пространства*** V называется квадратная матрица A, составленная из координатных столбцов образов базисных векторов $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$, найденных относительно базиса $e_1,\ldots,e_n$.

Матрица биективного линейного преобразования обратима, т.е. невырождена. Поэтому биективное (обратимое) преобразование называют также ***невырожденным***.

ПРИМЕРЫ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЙ

1. Обозначим $\mathcal{O}: V\to V$ – нулевое преобразование n-мерного пространства V, которое ставит в соответствие любому вектору $v\in V$ нулевой элемент o пространства V. Это преобразование не является инъективным, сюръективным, биективным, обратимым. Матрица нулевого преобразования (в любом базисе) нулевая, ядро преобразования $\mathbf{Ker}\,\mathcal{O}=V$, образ преобразования $\mathbf{Im}\,\mathcal{O}=\{o\}$, дефект $d=n$, ранг $r=0$.

2. Обозначим $\mathcal{E}: V\to V$ – тождественное преобразование n-мерного пространства V, которое ставит в соответствие каждому вектору $v\in V$ этот же вектор $\mathcal{E}(v)=v$. Это преобразование является инъективным, сюръективным, биективным, обратимым. Матрица тождественного преобразования (в любом базисе) единичная n-го порядка, ядро преобразования $\mathbf{Ker}\,\mathcal{E}=\{o\}$, образ преобразования $\mathbf{Im}\,\mathcal{E}=V$, дефект $d=0$, ранг $r=n$.

3. Обозначим $\mathcal{Z}_o: V\to V$ – ***центральную симметрию*** n-мерного пространства V (относительно нулевого вектора o), т.е. преобразование, которое каждому вектору ставит в соответствие противоположный ему вектор: $\mathcal{Z}_o(v)=-v$. Это преобразование линейное, инъективное, сюръективное, биективное, обратимое. Матрица преобразования противоположна единичной (в любом базисе): $Z_o=-E$; ядро преобразования $\mathbf{Ker}\,\mathcal{Z}_o=\{o\}$, образ преобразования $\mathbf{Im}\,\mathcal{Z}_o=V$, дефект $d=0$, ранг $r=n$.

4. Обозначим $\mathcal{H}_\lambda: V\to V$ – ***гомотетию*** n-мерного пространства V (с коэффициентом λ), т.е. преобразование, которое каждому вектору ставит в соответствие коллинеарный ему вектор: $\mathcal{H}_\lambda(v)=\lambda\cdot v$. Это преобразование линейное. При $\lambda\ne 0$ оно инъективное, сюръективное, биективное, обратимое. Матрица преобразования пропорциональна единичной (в любом базисе): $H_\lambda=\lambda\cdot E$, ядро преобразования $\mathbf{Ker}\,\mathcal{H}_\lambda=\{o\}$, образ преобразования $\mathbf{Im}\,\mathcal{H}_\lambda=V$, дефект $d=0$, ранг $r=n$. При $\lambda=0$: $\mathcal{H}_0=\mathcal{O}$ (см. п.1); при $\lambda=1$: $\mathcal{H}_1=\mathcal{E}$ (см. п.2); при $\lambda=-1$: $\mathcal{H}_{(-1)}=\mathcal{Z}_o$ (см. п.3).

5. Рассмотрим линейное пространство V_2 радиус-векторов (с общим началом в точке O), принадлежащих одной плоскости (рис. 9.1). Обозначим $\mathcal{R}_\varphi: V_2\to V_2$ – поворот вокруг точки O (на угол φ в положительном направлении (против часовой стрелки)). Это преобразование линейное, инъективное, сюръективное, биективное, обратимое. Найдем матрицу поворота в стандартном ортонормированном базисе $\vec{i}, \vec{j}$. Раскладывая образы $\vec{i}\,'=\mathcal{R}_\varphi(\vec{i})$, $\vec{j}\,'=\mathcal{R}_\varphi(\vec{j})$ базисных векторов по базису, получаем

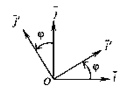

Рис. 9.1

$$\vec{i}\,'=\cos\varphi\cdot\vec{i}+\sin\varphi\cdot\vec{j},$$
$$\vec{j}\,'=-\sin\varphi\cdot\vec{i}+\cos\varphi\cdot\vec{j}.$$

Составляем матрицу (9.1) преобразования, записывая найденные координаты образов по столбцам:

$$R_\varphi=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Ядро преобразования $\mathbf{Ker}\,\mathcal{R}_\varphi=\{\vec{o}\}$, образ преобразования $\mathbf{Im}\,\mathcal{R}_\varphi=V_2$, дефект $d=0$, ранг $r=2$. При $\varphi=2\pi k$, $k\in Z$: $\mathcal{R}_{2\pi k}=\mathcal{E}$ (см. п.2); при $\varphi=\pi+2\pi k$, $k\in Z$: $\mathcal{R}_{\pi+2\pi k}=\mathcal{Z}_o$ (см. п.3).

6. Обозначим $\mathcal{D}: P_n(R)\to P_n(R)$ – оператор дифференцирования, который каждому многочлену степени не выше n ставит в соответствие его производную, рассматриваемую как многочлен степени не выше n: $\mathcal{D}(p(x))=p'(x)$. Это преобразование линейное, неинъективное, несюръективное, небиективное, необратимое. Квадратная матрица ($(n+1)$-го порядка) преобразования в стандартном базисе имеет вид

$$D=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Ядро преобразования $\mathbf{Ker}\,\mathcal{D}=P_0(R)$ – пространство многочленов нулевой степени, образ $\mathbf{Im}\,\mathcal{D}=P_{n-1}(R)$ – пространство многочленов степени не выше $(n-1)$, дефект $d=1$, ранг $r=n$, $\dim P_n(R)=n+1$.

Рассмотрим преобразование $\mathcal{D}: T_\omega(R)\to T_\omega(R)$ линейного пространства тригонометрических многочленов (частоты $\omega\ne 0$) с действительными коэффициентами: $T_\omega(R)=Lin(\sin\omega t,\cos\omega t)$, т.е. $T_\omega(R)$ – множество функций вида $f(t)=a\sin\omega t+b\cos\omega t$, где $a\in R$, $b\in R$ (см. п.7 в разд.8.1.3). Заметим, что это множество является двумерным вещественным линейным пространством (см. п.7 в разд.8.3.2). Стандартный базис пространства $T_\omega(R)$ образуют функции $e_1(t)=\sin\omega t$, $e_2(t)=\cos\omega t$, поскольку они линейно независимы (тождественное равенство нулю $a\sin\omega t+b\cos\omega t\equiv 0$ возможно только в тривиальном случае $a=b=0$). При дифференцировании функции $f(t)$ получаем функцию $f'(t)=-b\,\omega\sin\omega t+a\,\omega\cos\omega t$ того же вида. Следовательно, преобразование $\mathcal{D}: T_\omega(R)\to T_\omega(R)$ определено. Это преобразование линейное, инъективное, сюръективное, биективное, обратимое. Найдем матрицу преобразования в стандартном базисе $e_1(t)=\sin\omega t$, $e_2(t)=\cos\omega t$. Раскладывая образы базисных векторов, получаем

$$\mathcal{D}(e_1)=\omega\cdot\cos\omega t=0\cdot\sin\omega t+\omega\cdot\cos\omega t,$$
$$\mathcal{D}(e_2)=-\omega\cdot\sin\omega t=-\omega\cdot\sin\omega t+0\cdot\cos\omega t.$$

Составляем матрицу (9.1) преобразования, записывая найденные координаты образов по столбцам: $D=\begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$. Ядро преобразования $\mathbf{Ker}\,\mathcal{D}=\{o(t)\}$ – нулевое подпространство, образ $\mathbf{Im}\,\mathcal{D}=T_\omega(R)$, дефект $d=0$, ранг $r=2$, $o(t)=0\cdot\sin\omega t+0\cdot\cos\omega t$.

Аналогичными свойствами обладает преобразование $\mathcal{D}: T_\omega(C)\to T_\omega(C)$, где $T_\omega(C)=Lin(\sin\omega t,\cos\omega t)$ – множество функций вида $a\sin\omega t+b\cos\omega t$ с комплексными коэффициентами $a\in C$ и $b\in C$. Множество $T_\omega(C)$ является двумерным комплексным линейным пространством.

7. Пусть линейное пространство разлагается в прямую сумму подпространств $V=L_1\oplus L_2$. Обозначим $\Pi_{L_1}: V\to V$ – ***оператор проектирования на подпространство L_1 параллельно подпространству L_2***, который каждому вектору $v=v_1+v_2$, где $v_1\in L_1$, $v_2\in L_2$, ставит в соответствие его составляющую (проекцию) $v_1\in L_1$, т.е. $\Pi_{L_1}(v_1+v_2)=v_1$ (рис.9.2). Это преобразование линейное. При $L_1\ne V$ оно неинъективное, несюръективное, небиективное, необратимое. Ядро преобразования $\mathbf{Ker}\,\Pi_{L_1}=L_2$, образ преобразования $\mathbf{Im}\,\Pi_{L_1}=L_1$, дефект $d=\dim L_2$, ранг $r=\dim L_1$. При $L_1=V$: $\Pi_V=\mathcal{E}$; при $L_2=V$: $\Pi_{\{o\}}=\mathcal{O}$.

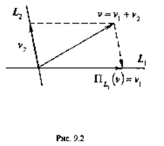

Рис. 9.2

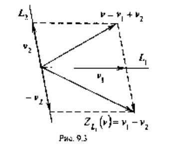

Рис. 9.3

8. Пусть линейное пространство разлагается в прямую сумму подпространств $V=L_1\oplus L_2$. Обозначим $Z_{L_1}: V\to V$ – ***оператор отражения в подпространстве L_1 параллельно подпространству L_2*** (или ***преобразование симметрии относительно подпространства L_1 параллельно подпространству L_2***), который каждому вектору $v=v_1+v_2$, где $v_1\in L_1$, $v_2\in L_2$, ставит в соответствие вектор (v_1-v_2), т.е. $Z_{L_1}(v_1+v_2)=v_1-v_2$ (рис. 9.3). Это преобразование линейное, инъективное, сюръективное, биективное, обратимое. Ядро преобразования $\mathbf{Ker}\,Z_{L_1}=\{o\}$, образ преобразования $\mathbf{Im}\,Z_{L_1}=V$, дефект $d=0$, ранг $r=\dim V$. При $L_1=V$: $Z_V=\mathcal{E}$.

33. Матрицы линейного преобразования в разных базисах.

9.2.2. Матрицы линейного преобразования в разных базисах

Найдем связь матриц одного и того же линейного преобразования в разных базисах.

Пусть в базисе $(e)=(e_1,\ldots,e_n)$ преобразование $\mathcal{A}:V\to V$ имеет матрицу $\underset{(e)}{A}$, а в базисе $(f)=(f_1,\ldots,f_n)$ – матрицу $\underset{(f)}{A}$. Если S – матрица перехода от базиса (e) к базису (f), то

$$\underset{(f)}{A}=S^{-1}\underset{(e)}{A}S. \qquad (9.4)$$

Докажем формулу (9.4). Пусть векторы v и w в базисах (e) и (f) имеют координатные столбцы $\underset{(e)}{v},\underset{(f)}{v}$ и $\underset{(e)}{w},\underset{(f)}{w}$ соответственно. Если $w=\mathcal{A}(v)$, то по формуле (9.2) имеем

$$\underset{(e)}{w}=\underset{(e)}{A}\,\underset{(e)}{v},\quad \underset{(f)}{w}=\underset{(f)}{A}\cdot\underset{(f)}{v}.$$

Подставляя в первое равенство связи координат векторов в разных базисах (см. разд. 8.4.3) $\underset{(e)}{v}=S\cdot\underset{(f)}{v}$, $\underset{(e)}{w}=S\cdot\underset{(f)}{w}$, получаем $S\cdot\underset{(f)}{w}=\underset{(e)}{A}\cdot S\cdot\underset{(f)}{v}$ или, учитывая обратимость матрицы S: $\underset{(f)}{w}=S^{-1}\cdot\underset{(e)}{A}\cdot S\cdot\underset{(f)}{v}$. Сравнивая последнее равенство с $\underset{(f)}{w}=\underset{(f)}{A}\cdot\underset{(f)}{v}$, убеждаемся в справедливости (9.4).

Замечания 9.2.

1. Матрицы линейного преобразования в разных базисах оказываются подобными (см. разд. 7.2.2). И наоборот, любые две подобные матрицы являются матрицами некоторого линейного преобразования, найденными относительно разных базисов.

2. Для матриц преобразований справедливы свойства, рассмотренные в разд. 9.1.4. В частности, при фиксированном базисе матрица суммы преобразований равна сумме их матриц, матрица произведения преобразования на число равна произведению матрицы преобразования на это же число, матрица композиции преобразований равна произведению матриц преобразований, матрица обратного преобразования является обратной для матрицы обратимого преобразования

34. Алгебра линейных преобразований: сложение, умножение на число, произведение и степень линейных операторов.

9.2.3. Алгебра линейных операторов

Рассмотрим множество $\mathcal{L}(V)$ – линейных преобразований (операторов) n-мерного линейного пространства V. Напомним, что два преобразования $\mathcal{A}:V\to V$ и $\mathcal{B}:V\to V$ называются равными, если $\mathcal{A}(v)=\mathcal{B}(v)$ $\forall v\in V$.

На множестве $\mathcal{L}(V)$ определены две линейные операции: сложение еобразований и умножение преобразования на число, поскольку в результате этих операций получается линейное преобразование (см. разд. 9.1.3).

Нетрудно показать, что эти операции удовлетворяют условиям:

1. $\mathcal{A}+\mathcal{B}=\mathcal{B}+\mathcal{A}\quad \forall\mathcal{A},\mathcal{B}\in\mathcal{L}(V)$;
2. $\mathcal{A}+(\mathcal{B}+\mathcal{C})=(\mathcal{A}+\mathcal{B})+\mathcal{C}\quad\forall\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{L}(V)$;
3. существует нулевое преобразование $\mathcal{O}\in\mathcal{L}(V)$ такое, что $\mathcal{A}+\mathcal{O}=\mathcal{A}$ $\forall\mathcal{A}\in\mathcal{L}(V)$;
4. для каждого преобразования $\mathcal{A}$ существует ***противоположное преобразование*** $(-\mathcal{A})=(-1)\cdot\mathcal{A}$ такое, что $\mathcal{A}+(-\mathcal{A})=\mathcal{O}$;
5. $\lambda\cdot(\mathcal{A}+\mathcal{B})=\lambda\cdot\mathcal{A}+\lambda\cdot\mathcal{B}\quad\forall\mathcal{A},\mathcal{B}\in\mathcal{L}(V)$ и любого числа λ;
6. $(\lambda+\mu)\cdot\mathcal{A}=\lambda\cdot\mathcal{A}+\mu\cdot\mathcal{A}\quad\forall\mathcal{A}\in\mathcal{L}(V)$ и любых чисел λ, μ;
7. $\lambda(\mu\cdot\mathcal{A})=(\lambda\mu)\cdot\mathcal{A}\quad\forall\mathcal{A}\in\mathcal{L}(V)$ и любых чисел λ, μ;
8. $1\cdot\mathcal{A}=\mathcal{A}\quad\forall\mathcal{A}\in\mathcal{L}(V)$.

В условиях 5–7 говорится о числах из того же числового поля, над которым определено линейное пространство V.

Условия 1–8 повторяют аксиомы линейного пространства (см. разд. 8.1.1). Поэтому множество $\mathcal{L}(V)$ с линейными операциями является линейным пространством. Если пространство V вещественное (комплексное), то и пространство $\mathcal{L}(V)$ вещественное (комплексное).

Найдем размерность пространства $\mathcal{L}(V)$. При фиксированном базисе имеется взаимно однозначное соответствие между линейными преобразованиями и их матрицами, причем это соответствие сохраняет линейные операции (см. п.2 замечаний 9.2). Следовательно, пространство $\mathcal{L}(V)$ изоморфно пространству $M_{n\times n}$ – квадратных матриц n-го порядка (см. п.5 в разд.8.3.2). Размерность пространства $M_{n\times n}$ равна n^2. По теореме 8.3 (см. разд.8.5): $\dim\mathcal{L}(V)=\dim M_{n\times n}=n^2$, т.е. $\dim\mathcal{L}(V)=(\dim V)^2$.

Кроме линейных операций в множестве $\mathcal{L}(V)$ определена операция умножения элементов. ***Произведением преобразований*** $\mathcal{A}$ и $\mathcal{B}$ назовем их композицию, т.е. $\mathcal{A}\mathcal{B}=\mathcal{A}\circ\mathcal{B}$. В результате композиции линейных преобразований получается линейное преобразование. Операция умножения удовлетворяет следующим условиям:

1. $\mathcal{A}(\mathcal{B}\mathcal{C})=(\mathcal{A}\mathcal{B})\mathcal{C}\quad\forall\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{L}(V)$;
2. $\mathcal{A}(\mathcal{B}+\mathcal{C})=\mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}\quad\forall\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{L}(V)$;
3. $(\mathcal{A}+\mathcal{B})\mathcal{C}=\mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C}\quad\forall\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{L}(V)$;
4. существует тождественное преобразование $\mathcal{E}$ такое, что $\mathcal{A}\mathcal{E}=\mathcal{E}\mathcal{A}=\mathcal{A}\quad\forall\mathcal{A}\in\mathcal{L}(V)$.

Первое условие выражает ассоциативность операции умножения, условия 2 и 3 – законы дистрибутивности, условие 4 – существование нейтрального элемента (см. разд. В.2.2, В.2.3). Множество $\mathcal{L}(V)$ с операциями сложения и умножения элементов является кольцом с единицей (вообще говоря, некоммутативное, так как в общем случае $\mathcal{A}\mathcal{B}\ne\mathcal{B}\mathcal{A}$).

Операции умножения преобразований и произведения преобразования на число (из заданного числового поля) удовлетворяют условию:

5. $(\lambda\cdot\mathcal{A})\mathcal{B}=\mathcal{A}(\lambda\cdot\mathcal{B})=\lambda\cdot(\mathcal{A}\mathcal{B})$.

Линейное пространство, которое является кольцом, удовлетворяющим условию 5, называется ***алгеброй***. Поэтому множество $\mathcal{L}(V)$ называют ***алгеброй линейных операторов (преобразований)***.

МНОГОЧЛЕНЫ ОТ ЛИНЕЙНОГО ПРЕОБРАЗОВАНИЯ

В алгебре $\mathcal{L}(V)$ можно определить целую неотрицательную степень оператора $\mathcal{A}:V\to V$, полагая по определению

$$\mathcal{A}^0=\mathcal{E},\quad \mathcal{A}^1=\mathcal{A},\quad \mathcal{A}^2=\mathcal{A}\mathcal{A},\ldots,\ \mathcal{A}^n=\mathcal{A}^{n-1}\mathcal{A}.$$

Пусть $p(\lambda)=a_m\lambda^m+\ldots+a_1\lambda+a_0$ – многочлен переменной λ. ***Многочленом $p(\mathcal{A})$ от линейного преобразования*** $\mathcal{A}$ называется преобразование $p(\mathcal{A})=a_m\mathcal{A}^m+\ldots+a_1\mathcal{A}+a_0\mathcal{E}$.

Многочлен $p(\lambda)=a_m\lambda^m+\ldots+a_1\lambda+a_0$ называется ***аннулирующим для линейного преобразования*** $\mathcal{A}$, если $p(\mathcal{A})=\mathcal{O}$ – нулевое преобразование. Заметим, что у каждого линейного преобразования $\mathcal{A}:V\to V$ n-мерного линейного пространства V существует аннулирующий многочлен степени не выше n^2. Действительно, система из (n^2+1) элементов $\mathcal{E},\mathcal{A},\ldots,\mathcal{A}^{n^2}$ линейного пространства $\mathcal{L}(V)$ линейно зависима (так как $\dim\mathcal{L}(V)=n^2$). Поэтому существуют такие числа $a_0,a_1,\ldots,a_{n^2}$, не все равные нулю одновременно, что $a_{n^2}\mathcal{A}^{n^2}+\ldots+a_1\mathcal{A}+a_0\mathcal{E}=\mathcal{O}$.

Следовательно, многочлен $p(\lambda)=a_{n^2}\lambda^{n^2}+\ldots+a_1\lambda+a_0$ – аннулирующий для преобразования $\mathcal{A}$.

35. Инвариантные подпространства: определение, примеры. Сужение (ограничение) оператора на подпространство. Свойства инвариантных подпространств.

Пусть $\mathcal{A}:V\to V$ – линейное преобразование линейного пространства V. Линейное подпространство $L\triangleleft V$ называется ***инвариантным*** относительно преобразования $\mathcal{A}$, если образ любого вектора из L принадлежит подпространству L, т.е. $\mathcal{A}(v)\in L\quad\forall v\in L$. Другими словами, инвариантное подпространство L включает свой образ $\mathcal{A}(L)$: $\mathcal{A}(L)\subset L$. Нулевое подпространство $\{o\}$ и все пространство V являются инвариантными подпространствами для любого линейного преобразования $\mathcal{A}:V\to V$.

Пусть L – инвариантное подпространство относительно преобразования $\mathcal{A}:V\to V$. Линейный оператор $\mathcal{A}:L\to L$, рассматриваемый как линейное преобразование пространства L в себя, называется ***сужением (ограничением) линейного преобразования*** $\mathcal{A}:V\to V$ ***на инвариантное подпространство*** $L\triangleleft V$ и обозначается $\mathcal{A}_L:L\to L$, или $\mathcal{A}|_L:L\to L$.

Для всех векторов $v\in L$ выполняется равенство $\mathcal{A}_L(v)=\mathcal{A}(v)$, т.е. $\forall v\in L$ образы, порождаемые оператором $\mathcal{A}$ и его сужением $\mathcal{A}_L$, совпадают.

ПРИМЕРЫ ИНВАРИАНТНЫХ ПОДПРОСТРАНСТВ

Рассмотрим инвариантные подпространства линейных преобразований, определенных в разд. 9.2.1.

1. Для нулевого преобразования $\mathcal{O}:V\to V$ любое подпространство $L\triangleleft V$ является инвариантным, так как $\mathcal{O}(L)=\{o\}\subset L$. Сужение нулевого преобразования $\mathcal{O}_L:L\to L$ является нулевым преобразованием.

2. Для тождественного преобразования $\mathcal{E}:V\to V$ любое подпространство $L\triangleleft V$ является инвариантным, так как $\mathcal{E}(L)=L$. Сужение тождественного преобразования $\mathcal{E}_L:L\to L$ является тождественным преобразованием.

3. Для центральной симметрии $\mathcal{Z}_o:V\to V$ любое подпространство $L\triangleleft V$ является инвариантным, так как $\mathcal{Z}_o(L)=L$. Сужение центральной симметрии $\mathcal{Z}_o|_L:L\to L$ является центральной симметрией.

4. Для гомотетии $\mathcal{H}_\lambda:V\to V$ любое подпространство $L\triangleleft V$ является инвариантным, так как $\mathcal{H}_\lambda(L)=L$ (при $\lambda\ne0$). Сужение гомотетии $\mathcal{H}_\lambda|_L:L\to L$ является гомотетией.

5. Для поворота $\mathcal{R}_\varphi:V_2\to V_2$ плоскости (при $\varphi\ne\pi k$, $k\in Z$) имеются два инвариантных подпространства: нулевое $\{\bar{o}\}$ и вся плоскость V_2. Других инвариантных подпространств нет.

6. Для оператора дифференцирования $\mathcal{D}:P_n(R)\to P_n(R)$ каждое из подпространств $\{o(x)\}\triangleleft P_0(R)\triangleleft P_1(R)\triangleleft\ldots\triangleleft P_n(R)$ является инвариантным, так как при дифференцировании степень многочлена уменьшается.

7. Рассмотрим оператор $\Pi_{L_1}:V\to V$ проектирования на подпространство L_1 параллельно подпространству L_2. Здесь $V=L_1\oplus L_2$, $\Pi_{L_1}(v_1+v_2)=v_1$ для $v=v_1+v_2$, $v_1\in L_1$, $v_2\in L_2$. Для этого оператора подпространства L_1 и L_2 инвариантные, так как $\Pi_{L_1}(L_1)=L_1$ и $\Pi_{L_1}(L_2)=\{o\}\subset L_2$. Сужение оператора проектирования на подпространство L_1 является тождественным преобразованием $\Pi_{L_1}|_{L_1}=\mathcal{E}$, а сужение на подпространство L_2 – нулевым $\Pi_{L_1}|_{L_2}=\mathcal{O}$.

9.3.2. Свойства инвариантных подпространств

1. *Если L – инвариантное подпространство относительно обратимого линейного преобразования $\mathcal{A}:V\to V$, то его сужение $\mathcal{A}_L:L\to L$ также обратимое линейное преобразование.*

2. *Для любого линейного преобразования $\mathcal{A}:V\to V$ ядро $Ker\,\mathcal{A}$ и образ $Im\,\mathcal{A}$ являются инвариантными подпространствами*, так как $\mathcal{A}(Ker\,\mathcal{A})=\{o\}\triangleleft Ker\,\mathcal{A}$ и $\mathcal{A}(Im\,\mathcal{A})\triangleleft Im\,\mathcal{A}$.

3. *Если L – инвариантное подпространство относительно линейного преобразования $\mathcal{A}:V\to V$, то L – инвариантно относительно любой натуральной степени этого преобразования, причем*

$$\mathcal{A}^m(L)\triangleleft\ldots\triangleleft\mathcal{A}(L)\triangleleft\mathcal{E}(L)=L.$$

В самом деле, каждое из указанных множеств является линейным подпространством, так как это образы сужений линейных операторов, например, $\mathcal{A}^m(L)=Im(\mathcal{A}_L)^m$. Докажем, например, включение $\mathcal{A}^2(L)\triangleleft\mathcal{A}(L)$. Для любого $w\in\mathcal{A}^2(L)$ существует вектор $v\in\mathcal{A}(L)\triangleleft L$, что $w=\mathcal{A}(v)$. Следовательно, $w\in\mathcal{A}(L)$.

4. *Если L – инвариантное подпространство относительно линейного преобразования $\mathcal{A}:V\to V$, то L – инвариантно относительно любого многочлена от этого преобразования.*

36. Теорема о матрицах оператора и его сужения на инвариантное подпространство, следствие.

Теорема 9.2 (о матрицах оператора и его сужения на инвариантное подпространство). *Пусть $\mathcal{A}:V\to V$ – линейное преобразование n-мерного пространства V, а L – подпространство, инвариантное относительно преобразования $\mathcal{A}$. Тогда существует базис $(e)=(e_1,\ldots,e_n)$ пространства V, в котором матрица A преобразования $\mathcal{A}$ имеет нулевой угол:*

$$A=\left(\begin{array}{c|c}B&C\\ O&D\end{array}\right),$$

где B – матрица сужения $\mathcal{A}_L$ преобразования $\mathcal{A}$ на подпространство L, O – нулевая матрица размеров $(n-l)\times l$, $l=\dim L$. И наоборот, *если в некотором базисе (e) матрица A преобразования $\mathcal{A}$ имеет нулевой угол (нулевую матрицу O размеров $(n-l)\times l$), то преобразование $\mathcal{A}$ имеет l-мерное инвариантное подпространство.*

В самом деле, возьмем базис $e_1,\ldots,e_l$ подпространства L и дополним его векторами $e_{l+1},\ldots,e_n$ до базиса $e_1,\ldots,e_n$ всего пространства V. Раскладывая образы первых l базисных векторов по этому базису, получаем

$$\mathcal{A}(e_i)=a_{1i}e_1+\ldots+a_{li}e_l+0\cdot e_{l+1}+\ldots+0\cdot e_n,$$

так как $\mathcal{A}(e_i)\in L$, $i=1,\ldots,l$. Следовательно, последние $(n-l)$ элементов первых l столбцов матрицы A преобразования $\mathcal{A}$ равны нулю. Обратное утверждение доказывается, проводя аналогичные рассуждения в обратном порядке.

Следствие. *Если n-мерное пространство V представлено в виде прямой суммы ненулевых инвариантных относительно преобразования $\mathcal{A}$ подпространств*

$$V=L_1\oplus\ldots\oplus L_k,$$

то существует базис, в котором матрица преобразования имеет блочно-диагональный вид

$$A=diag(A_1,\ldots,A_k)=\begin{pmatrix}A_1&&O\\&\ddots&\\O&&A_k\end{pmatrix},$$

где A_i – матрица сужения $\mathcal{A}_{L_i}$ преобразования $\mathcal{A}$ на подпространство L_i, $i=1,\ldots,k$.

Например, рассмотрим операторы проектирования $\Pi_{L_1}:V\to V$ и отражения $Z_{L_1}:V\to V$ (см. п.7, 8 в разд. 9.3.1). Объединяя базисы подпространств L_1 и L_2, получаем базис пространства $V=L_1\oplus L_2$, в котором матрицы преобразований имеют блочно-диагональный вид

$$\Pi_{L_1}=\left(\begin{array}{c|c}E&O\\ \hline O&O\end{array}\right),\quad Z_{L_1}=\left(\begin{array}{c|c}E&O\\ \hline O&-E\end{array}\right).$$

37. Собственные векторы и собственные значения линейного преобразования. Геометрический смысл собственных векторов. Примеры . Свойства .

9.4.1. Собственные векторы и собственные значения

Пусть $\mathcal{A}:V\to V$ – линейное преобразование n-мерного линейного пространства V. Ненулевой вектор s линейного пространства V, удовлетворяющий условию

$$\mathcal{A}(s)=\lambda s, \qquad (9.5)$$

называется ***собственным вектором линейного преобразования*** $\mathcal{A}$. Число λ в равенстве (9.5) называется ***собственным значением преобразования*** $\mathcal{A}$. Говорят, что собственный вектор ***соответствует (принадлежит) собственному значению*** λ. Если пространство V вещественное (комплексное), то собственное значение λ – действительное (комплексное) число.

Множество всех собственных значений линейного преобразования называется его ***спектром***.

Поясним геометрический смысл собственных векторов. Ненулевой вектор s является собственным для преобразования $\mathcal{A}$, если его образ $\mathcal{A}(s)$ коллинеарен прообразу s. Другими словами, *если s – собственный вектор, то преобразование $\mathcal{A}$ имеет одномерное инвариантное подпространство $Lin(s)$*. Справедливо и обратное утверждение.

В самом деле, пусть собственный вектор s соответствует некоторому собственному значению λ. Любой вектор v из $Lin(s)$ имеет вид $v=\alpha s$, где α – любое число из заданного поля. Найдем образ этого вектора

$$\mathcal{A}(v)=\mathcal{A}(\alpha s)=\alpha\mathcal{A}(s)=\alpha\lambda s\in Lin(s).$$

Следовательно, $\mathcal{A}(v)\in Lin(s)$ для любого вектора $v\in Lin(s)$, т.е. подпространство $Lin(s)$ инвариантно относительно преобразования $\mathcal{A}$. Размерность подпространства $Lin(s)$ равна единице, так как $s\ne o$ по определению.

Обратное утверждение доказывается, проводя рассуждения в обратном порядке.

9.4.2. Примеры собственных векторов

1. Для нулевого преобразования $\mathcal{O}:V\to V$ любой ненулевой вектор $s\in V$ является собственным, соответствующим нулевому собственному значению $\lambda=0$, так как $\mathcal{O}(s)=0\cdot s\quad\forall s\in V$.

2. Для тождественного преобразования $\mathcal{E}:V\to V$ любой ненулевой вектор $s\in V$ является собственным, соответствующим единичному собственному значению $\lambda=1$, так как $\mathcal{E}(s)=1\cdot s\quad\forall s\in V$.

3. Для центральной симметрии $\mathcal{Z}_o:V\to V$ любой ненулевой вектор $s\in V$ является собственным, соответствующим собственному значению $\lambda=-1$, так как $\mathcal{Z}_o(s)=(-1)\cdot s\quad\forall s\in V$.

4. Для гомотетии $\mathcal{H}_\lambda:V\to V$ любой ненулевой вектор $s\in V$ является собственным, соответствующим собственному значению λ (коэффициенту гомотетии), так как $\mathcal{H}_\lambda(s)=\lambda\cdot s\quad\forall s\in V$.

5. Для поворота $\mathcal{R}_\varphi:V_2\to V_2$ плоскости (при $\varphi\ne\pi k$, $k\in Z$) собственных векторов нет, так как при повороте на угол, не кратный π, образ каждого ненулевого вектора неколлинеарен прообразу. Здесь рассматривается поворот вещественной плоскости, т.е. двумерного векторного пространства над полем действительных чисел.

6. Для оператора дифференцирования $\mathcal{D}:P_n(R)\to P_n(R)$ любой ненулевой многочлен нулевой степени (не равный тождественно нулю) является собственным вектором, соответствующим нулевому собственному значению $\lambda=0$, так как $\mathcal{D}(s(x))=0\cdot s(x)\quad\forall s(x)\equiv\text{const}$. Любой многочлен ненулевой степени не является собственным вектором, так как многочлен не пропорционален своей производной: $\mathcal{D}(s(x))=s'(x)\ne\lambda\cdot s(x)$, поскольку они имеют разные степени.

7. Рассмотрим оператор $\Pi_{L_1}:V\to V$ проектирования на подпространство L_1 параллельно подпространству L_2. Здесь $V=L_1\oplus L_2$, $\Pi_{L_1}(v_1+v_2)=v_1$ для $v=v_1+v_2$, $v_1\in L_1$, $v_2\in L_2$. Для этого оператора любой ненулевой вектор $v_1\in L_1$ является собственным, соответствующим собственному значению $\lambda=1$, так как $\Pi_{L_1}(v_1)=1\cdot v_1$, а любой ненулевой вектор $v_2\in L_2$ является собственным, соответствующим собственному значению $\lambda=0$, так как $\Pi_{L_1}(v_2)=0\cdot v_2$. Другие векторы не являются собственными.

38. Теорема о собственных векторах линейного преобразования и его матрицы

Теорема 9.3 (о собственных векторах линейного преобразования и его матрицы). *Пусть $\mathcal{A}:V\to V$ – линейное преобразование n-мерного линейного пространства V с базисом $e_1,\ldots,e_n$. Тогда собственное значение λ и координатный столбец s собственного вектора s преобразования $\mathcal{A}$ являются собственным значением и собственным вектором матрицы A* (см. разд. 7.2.1) *этого преобразования, определенной относительно базиса $e_1,\ldots,e_n$*, т.е.

$$\mathcal{A}(s)=\lambda\cdot s\quad\Rightarrow\quad As=\lambda\cdot s,$$

где $s=s_1e_1+\ldots+s_ne_n$, $s=(s_1\ \cdots\ s_n)^T$. Обратное утверждение справедливо при дополнительных условиях: *если столбец $s=(s_1\ \cdots\ s_n)^T$ и число λ являются собственным вектором и собственным значением матрицы A, причем числа $s_1,\ldots,s_n,\lambda$ принадлежат тому же числовому полю, над которым определено линейное пространство V, то вектор $s=s_1e_1+\ldots+s_ne_n$ и число λ являются собственным вектором и собственным значением линейного преобразования $\mathcal{A}:V\to V$ с матрицей A в базисе $e_1,\ldots,e_n$.*

В самом деле, условие (9.5) в координатной форме имеет вид (9.6), что совпадает с определением (7.13) собственного вектора матрицы. Наоборот, из равенства (9.6) следует равенство (9.5) при условии, что векторы $s=s_1e_1+\ldots+s_ne_n$ и $\lambda\cdot s$ определены, т.е. числа $s_1,\ldots,s_n,\lambda$ принадлежат тому же числовому полю, над которым определено линейное пространство.

Напомним, что нахождение собственных значений матрицы сводится к решению ее характеристического уравнения $\Delta_A(\lambda)=0$, где $\Delta_A(\lambda)=\det(A-\lambda E)$ – характеристический многочлен матрицы A. Для линейного преобразования введем аналогичные понятия.

Характеристическим многочленом линейного преобразования $\mathcal{A}:V\to V$ n-мерного линейного пространства называется характеристический многочлен $\Delta_{\mathcal{A}}(\lambda)=\det(A-\lambda E)$ матрицы A этого преобразования, найденной относительно любого базиса пространства V.

Уравнение $\Delta_{\mathcal{A}}(\lambda)=0$ называется ***характеристическим уравнением линейного преобразования***.

Преобразование $\mathcal{A}-\lambda\mathcal{E}$ называется ***характеристическим*** для линейного преобразования $\mathcal{A}:V\to V$.

39. *Алгоритм нахождения собственных значений и собственных векторов линейного преобразования.*

Для нахождения собственных векторов и собственных значений линейного преобразования $\mathcal{A}:V\to V$ вещественного линейного пространства V следует выполнить следующие действия.

1. Выбрать произвольный базис $e_1,\ldots,e_n$ линейного пространства V и найти в этом базисе матрицу A преобразования $\mathcal{A}$.

2. Составить характеристический многочлен преобразования $\mathcal{A}$:
$$\Delta_{\mathcal{A}}(\lambda)=\det(A-\lambda E).$$

3. Найти все различные действительные корни $\lambda_1,\ldots,\lambda_k$ характеристического уравнения $\Delta_{\mathcal{A}}(\lambda)=0$. Комплексные (но не действительные) корни характеристического уравнения следует отбросить (см. п.2. замечаний 9.4).

4. Для корня $\lambda=\lambda_1$ найти фундаментальную систему $\varphi_1,\varphi_2,\ldots,\varphi_{n-r}$ решений однородной системы уравнений
$$(A-\lambda_1E)x=o,$$
где $r=rg(A-\lambda_1E)$. Для этого можно использовать либо алгоритм решения однородной системы (см. разд. 5.5), либо один из способов нахождения фундаментальной матрицы (см. п.3 замечаний 5.3, п.1 замечаний 5.5).

5. Записать линейно независимые собственные векторы преобразования $\mathcal{A}$, отвечающие собственному значению λ_1:
$$s_1=\varphi_{11}e_1+\ldots+\varphi_{n1}e_n,$$
$$s_2=\varphi_{12}e_1+\ldots+\varphi_{n2}e_n,$$
$$\vdots$$
$$s_{n-r}=\varphi_{1\,n-r}\,e_1+\ldots+\varphi_{n\,n-r}\,e_n.$$

Для нахождения совокупности всех собственных векторов, отвечающих собственному значению λ_1, образовать ненулевые линейные комбинации

$$s=C_1s_1+C_2s_2+\ldots+C_{n-r}s_{n-r},$$

где $C_1,\ldots,C_{n-r}$ – произвольные постоянные, не равные нулю одновременно.

Повторить п.4, 5 для остальных собственных значений $\lambda_2,\ldots,\lambda_k$.

Для нахождения собственных векторов линейного преобразования комплексного линейного пространства нужно в п.3 определить все корни характеристического уравнения и, не отбрасывая комплексные корни, выполнить для них п.4, 5.

40. Характеристический многочлен линейного преобразования и его свойства.

Характеристическим многочленом линейного преобразования $\mathcal{A}:V\to V$ n-мерного линейного пространства называется характеристический многочлен $\Delta_{\mathcal{A}}(\lambda)=\det(A-\lambda E)$ матрицы A этого преобразования, найденной относительно любого базиса пространства V.

Уравнение $\Delta_{\mathcal{A}}(\lambda)=0$ называется ***характеристическим уравнением линейного преобразования***.

Преобразование $\mathcal{A}-\lambda\mathcal{E}$ называется ***характеристическим*** для линейного преобразования $\mathcal{A}:V\to V$.

Замечания 9.4.

1. *Характеристический многочлен линейного преобразования не зависит от базиса, в котором найдена матрица преобразования.*

В самом деле, матрицы $\underset{(e)}{A}$ и $\underset{(f)}{A}$ линейного преобразования $\mathcal{A}$ в базисах $(e)=(e_1,\ldots,e_n)$ и $(f)=(f_1,\ldots,f_n)$ являются, согласно (9.4), подобными: $\underset{(f)}{A}=S^{-1}\underset{(e)}{A}S$, где S – матрица перехода от базиса (e) к базису (f). Как показано в разд. 7.2.3, характеристические многочлены подобных матриц совпадают (см. свойство 3). Поэтому для характеристического многочлена преобразования $\mathcal{A}$ можно использовать обозначение $\Delta_{\mathcal{A}}(\lambda)$, не указывая матрицу этого преобразования.

2. Из теоремы 9.3 следует, что любой комплексный (действительный, рациональный) корень характеристического уравнения является собственным значением линейного преобразования $\mathcal{A}:V\to V$ линейного пространства V, определенного над полем комплексных (действительных, рациональных) чисел.

3. Из теоремы 9.3 следует, что любое линейное преобразование комплексного линейного пространства имеет одномерное инвариантное подпространство, так как это преобразование имеет собственное значение (см. п.2), а следовательно, и собственные векторы. Таким подпространством является, например, линейная оболочка любого собственного вектора. У преобразования вещественного линейного пространства одномерных инвариантных подпространств может и не быть, если все корни характеристического уравнения комплексные (но не действительные).

41. Теорема об инвариантных подпространствах линейного преобразования вещественного линейного пространства.

Теорема 9.4 (об инвариантных подпространствах линейного преобразования вещественного пространства). *У всякого линейного преобразования вещественного линейного пространства существует одномерное или двумерное инвариантное подпространство.*

Действительно, составим матрицу A линейного преобразования $\mathcal{A}: V \to V$ n-мерного вещественного линейного пространства V в произвольном базисе $e_1, \ldots, e_n$. Элементы этой матрицы – действительные числа. Следовательно, характеристический многочлен $\Delta_{\mathcal{A}}(\lambda) = \det(A - \lambda E)$ – это многочлен степени n с действительными коэффициентами. Согласно следствиям 3, 4 основной теоремы алгебры (см. разд. В.4), такой многочлен может иметь действительные корни и пары комплексных сопряженных корней.

Если $\lambda = \lambda$ – действительный корень характеристического уравнения, то и соответствующий собственный вектор $s = (s_1 \ \cdots \ s_n)^T$ матрицы A также действительный. Поэтому он определяет собственный вектор $s = s_1 e_1 + \ldots + s_n e_n$ линейного преобразования (см. теорему 9.3). В этом случае существует одномерное инвариантное относительно $\mathcal{A}$ подпространство $Lin(s)$ (см. геометрический смысл собственных векторов).

Если $\lambda = \alpha \pm \beta i$ – пара комплексных сопряженных корней ($\beta \ne 0$), то собственный вектор $s \ne o$ матрицы A также с комплексными элементами: $s = (x_1 + y_1 i \ \cdots \ x_n + y_n i)^T$. Его можно представить в виде $s = x + yi$, где x, y – действительные столбцы. Равенство (9.6) при этом будет иметь вид

$$A \cdot (x + yi) = (\alpha + \beta i) \cdot (x + yi).$$

Выделяя действительную и мнимую части, получаем систему

$$\begin{cases} Ax = \alpha x - \beta y, \\ Ay = \beta x + \alpha y. \end{cases} \tag{9.7}$$

Покажем, что столбцы x и y линейно независимы. Рассмотрим два случая. Если $x = o$, то из первого уравнения (9.7) следует, что $y = o$, так как $\beta \ne 0$. Тогда $s = o$, что противоречит условию $s \ne o$. Предположим, что $x \ne o$ и столбцы x и y пропорциональны, т.е. существует такое действительное число γ, что $y = \gamma x$. Тогда из системы (9.7) получаем

$$\begin{cases} Ax = (\alpha - \beta\gamma)\, x, \\ \gamma Ax = (\beta + \alpha\gamma) x. \end{cases}$$

Прибавляя ко второму уравнению первое, умноженное на $(-\gamma)$, приходим к равенству $[(\beta + \alpha\gamma) - \gamma(\alpha - \beta\gamma)] x = o$. Так как $x \ne o$, то выражение в квадратных скобках равно нулю, т.е. $(\beta + \alpha\gamma) - \gamma(\alpha - \beta\gamma) = \beta(1 + \gamma^2) = 0$. Поскольку $\beta \ne 0$, то $\gamma^2 = -1$. Этого не может быть, так как γ – действительное число. Получили противоречие. Таким образом, столбцы x и y линейно независимы.

Рассмотрим подпространство $Lin(x, y)$, где $x = x_1 e_1 + \ldots + x_n e_n$, $y = y_1 e_1 + \ldots + y_n e_n$. Это подпространство двумерное, так как векторы x, y линейно независимы (как показано выше, их координатные столбцы x, y линейно независимы). Из (9.7) следует, что

$$\begin{cases} \mathcal{A}(x) = \alpha x - \beta y, \\ \mathcal{A}(y) = \beta x + \alpha y, \end{cases}$$

т.е. образ любого вектора, принадлежащего $Lin(x, y)$, также принадлежит $Lin(x, y)$. Следовательно, $Lin(x, y)$ – двумерное подпространство, инвариантное относительно преобразования $\mathcal{A}$, что и требовалось доказать.

42. Собственные и корневые подпространства линейного преобразования, цепочка инвариантных подпространств.

3. *Для собственного значения λ линейного преобразования $\mathcal{A}: V \to V$ существует цепочка инвариантных подпространств*

$$\{o\} \triangleleft K_\lambda^1 \triangleleft K_\lambda^2 \triangleleft \ldots \triangleleft K_\lambda^m \triangleleft V, \tag{9.8}$$

где $K_\lambda^1 = Ker(\mathcal{A} - \lambda \cdot \mathcal{E})$, $K_\lambda^2 = Ker(\mathcal{A} - \lambda \cdot \mathcal{E})^2, \ldots,$ $K_\lambda^m = Ker(\mathcal{A} - \lambda \cdot \mathcal{E})^m$; m *– некоторое натуральное число* ($m \le n = \dim V$).

Все перечисленные в цепочке (9.8) множества K_λ^k, $k = 1, \ldots, m$, являются линейными подпространствами по свойству ядра линейного преобразования. Каждое из подпространств K_λ^k инвариантно относительно преобразования $\mathcal{A}$, поскольку для любого вектора $v \in K_\lambda^k$ его образ $w = \mathcal{A}(v) \in K_\lambda^k$, так как в силу перестановочности многочленов от одного и того же линейного преобразования (см. п.2 замечаний 9.3)

$$(\mathcal{A} - \lambda \cdot \mathcal{E})^k (w) = (\mathcal{A} - \lambda \cdot \mathcal{E})^k \mathcal{A}(v) = \mathcal{A}(\mathcal{A} - \lambda \cdot \mathcal{E})^k (v) = \mathcal{A}(o) = o,$$

так как $(\mathcal{A} - \lambda \cdot \mathcal{E})^k (v) = o$ $\forall v \in K_\lambda^k$ согласно определения ядра.

Докажем включение $K_\lambda^1 \triangleleft K_\lambda^2$. Если $v \in K_\lambda^1$, то $(\mathcal{A} - \lambda \cdot \mathcal{E})(v) = o$, при этом очевидно, что $(\mathcal{A} - \lambda \cdot \mathcal{E}) \cdot (\mathcal{A} - \lambda \cdot \mathcal{E})(v) = (\mathcal{A} - \lambda \cdot \mathcal{E})(o) = o$, т.е. $v \in K_\lambda^2$. Остальные включения доказываются аналогично.

Из цепочки (9.8) "расширяющихся" подпространств следует, что их размерности не убывают

$$0 \le \dim K_\lambda^1 \le \dim K_\lambda^2 \le \ldots \le \dim K_\lambda^m \le \dim V,$$

поэтому в силу конечномерности пространства V существует такое m, что $\dim K_\lambda^m = \dim K_\lambda^{m+1}$, т.е. $K_\lambda^m = K_\lambda^{m+1}$. Покажем, что дальнейшего "увеличения" подпространств нет, т.е. $K_\lambda^m = K_\lambda^{m+1} = \ldots = K_\lambda^{m+k}$ для любого натурального k. Предположим противное. Пусть $K_\lambda^m = K_\lambda^{m+1}$ и для некоторого $k > 1$ пространства не совпадают: $K_\lambda^{m+k} \ne K_\lambda^{m+k+1}$, т.е. существует вектор $v \in K_\lambda^{m+k+1}$, который не принадлежит пространству K_λ^{m+k}. Обозначим $w = (\mathcal{A} - \lambda \cdot \mathcal{E})^k (v)$. Тогда, с одной стороны, $w \in K_\lambda^{m+1}$, так как $(\mathcal{A} - \lambda \cdot \mathcal{E})^{m+1}(w) = (\mathcal{A} - \lambda \cdot \mathcal{E})^{m+k+1}(v) = o$, поскольку $v \in K_\lambda^{m+k+1}$. С другой стороны, $w \notin K_\lambda^m$, так как $(\mathcal{A} - \lambda \cdot \mathcal{E})^m (w) = (\mathcal{A} - \lambda \cdot \mathcal{E})^{m+k}(v) \ne o$, поскольку $v \notin K_\lambda^{m+k}$. Следовательно, $w \in K_\lambda^{m+1}$ и $w \notin K_\lambda^m$ одновременно, что противоречит предположению $K_\lambda^m = K_\lambda^{m+1}$.

Таким образом, в цепочке (9.8) размерности пространств K_λ^k, $k = 1, \ldots, m$, возрастают. Поэтому $m \le n = \dim V$.

Корневым подпространством линейного преобразования $\mathcal{A}$ для собственного значения λ называется линейное подпространство $K_\lambda^m = Ker(\mathcal{A} - \lambda \cdot \mathcal{E})^m$ с наименьшим натуральным показателем m, для которого $K_\lambda^m = K_\lambda^{m+1}$.

4. *Если λ – собственное значение линейного преобразования $\mathcal{A}: V \to V$, то пространство V можно представить в виде прямой суммы $V = K_\lambda^m \oplus L$, где K_λ^m – корневое подпространство, а $L = Im(\mathcal{A} - \lambda \cdot \mathcal{E})^m$ – инвариантное относительно $\mathcal{A}$ подпространство, в котором нет собственных векторов, принадлежащих собственному значению λ.*

В самом деле, покажем, что пересечение этих подпространств есть нулевой вектор: $K_\lambda^m \cap L = \{o\}$. Выберем вектор $w \in K_\lambda^m \cap L$. Так как вектор $w \in L$, то существует такой вектор $v \in V$, что $w = (\mathcal{A} - \lambda\mathcal{E})^m (v)$. Поскольку $w \in K_\lambda^m$, то $(\mathcal{A} - \lambda\mathcal{E})^m (w) = o$. Тогда $(\mathcal{A} - \lambda\mathcal{E})^{2m}(v) = (\mathcal{A} - \lambda\mathcal{E})^m (w) = o$. Следовательно, вектор $v \in K_\lambda^{2m}$, но $K_\lambda^{2m} = K_\lambda^m$, так как K_λ^m – корневое подпространство. Значит, $v \in K_\lambda^m \Rightarrow w = (\mathcal{A} - \lambda\mathcal{E})^m (v) = o$, т.е. $K_\lambda^m \cap L = \{o\}$. По теореме 9.1 о размерности ядра и образа получаем, что $\dim K_\lambda^m + \dim L = \dim V$. Следовательно, пространство V можно представить в виде прямой суммы подпространств $V = K_\lambda^m \oplus L$ (см. признаки прямых сумм подпространств в разд.8.6.4).

Докажем, что в L нет собственных векторов, принадлежащих собственному значению λ. Действительно, пусть s – собственный вектор, соответствующий собственному значению λ. Тогда $s \in K_\lambda^1$ и в силу (9.8) $s \in K_\lambda^m$. Подпространство L имеет с K_λ^m только один общий вектор (нулевой). Поэтому $s \notin L$, так как $s \ne o$. Инвариантность подпространства L следует из перестановочности операторов $\mathcal{A}$ и $(\mathcal{A} - \lambda \cdot \mathcal{E})^m$ (см. п.2 замечаний 9.3). В самом деле, для любого вектора $w \in L$ существует прообраз $v \in V$: $w = (\mathcal{A} - \lambda \cdot \mathcal{E})^m (v)$. Поэтому в силу перестановочности операторов

$$\mathcal{A}(w) = \mathcal{A}(\mathcal{A} - \lambda \cdot \mathcal{E})^m (v) = (\mathcal{A} - \lambda \cdot \mathcal{E})^m \mathcal{A}(v) \in L,$$

поскольку $\mathcal{A}(v) \in V$ и $L = Im(\mathcal{A} - \lambda\mathcal{E})^m$. Таким образом, инвариантность подпространства L доказана, так как $\mathcal{A}(w) \in L$ $\forall w \in L$.

43. Теорема о разложении пространства в прямую сумму корневых подпространств.

Теорема 9.5 (о разложении пространства в сумму корневых подпространств). *Если все различные корни $\lambda_1, \ldots, \lambda_k$ характеристического уравнения линейного преобразования $\mathcal{A}: V \to V$ являются его собственными значениями, то пространство V можно разложить в прямую сумму инвариантных (корневых) подпространств:*

$$V = K_{\lambda_1}^{m_1} \oplus \ldots \oplus K_{\lambda_k}^{m_k}, \tag{9.9}$$

где $K_{\lambda_i}^{m_i} = Ker(\mathcal{A} - \lambda_i \cdot \mathcal{E})^{m_i}$ – корневое подпространство, соответствующее собственному значению λ_i, $i = 1, \ldots, k$.

В самом деле, по свойству 4 можно "отщепить" корневое подпространство $K_{\lambda_1}^{m_1}$, т.е. представить пространство V в виде прямой суммы инвариантных подпространств $V = K_{\lambda_1}^{m_1} \oplus L_1$, причем в L_1 нет собственных векторов, принадлежащих собственному значению λ_1. В пространстве L_1 определено сужение $\mathcal{A}_{L_1}: L_1 \to L_1$ преобразования $\mathcal{A}$. Применяя свойство 4 к сужению $\mathcal{A}_{L_1}: L_1 \to L_1$, аналогичным образом можно "отщепить" корневое подпространство $K_{\lambda_2}^{m_2}$, т.е. представить пространство L_1 в виде прямой суммы инвариантных подпространств: $L_1 = K_{\lambda_2}^{m_2} \oplus L_2$. Этот процесс следует продолжить до тех пор, пока не исчерпаются все корни характеристического уравнения.

Следствие. *Если все различные корни $\lambda_1, \ldots, \lambda_k$ характеристического уравнения линейного преобразования $\mathcal{A}: V \to V$ являются его собственными значениями, то существует базис пространства V, в котором матрица A линейного преобразования имеет блочно-диагональный вид*

$$A = diag(A_1, \ldots, A_k),$$

где $A_1, \ldots, A_k$ – матрицы сужений $\mathcal{A}_{K_{\lambda_i}^{m_i}}: K_{\lambda_i}^{m_i} \to K_{\lambda_i}^{m_i}$, $i = 1, \ldots, k$, преобразования $\mathcal{A}$ на корневые подпространства.

Согласно следствию из теоремы 9.2, такой базис можно получить, записывая последовательно базисы корневых подпространств (9.9).

44. Теорема об алгебраической и геометрической кратностях собственных значений линейного преобразования.

Алгебраической кратностью собственного значения λ_1 линейного преобразования $\mathcal{A}: V \to V$ называется кратность корня $\lambda = \lambda_1$ характеристического многочлена $\Delta_{\mathcal{A}}(\lambda)$ (или, что то же самое, кратность корня характеристического уравнения $\Delta_{\mathcal{A}}(\lambda) = 0$).

Геометрической кратностью собственного значения λ_1 линейного преобразования $\mathcal{A}: V \to V$ называется размерность собственного подпространства $K_{\lambda_1}^1 = Ker(\mathcal{A} - \lambda_1 \cdot \mathcal{E})$, соответствующего этому собственному значению.

Теорема 9.6 (о кратностях собственных значений). *Геометрическая кратность собственного значения не превосходит его алгебраической кратности.*

Представим пространство V в виде прямой суммы $V = K_{\lambda_1}^{m_1} \oplus L$ (см. свойство 4) и обозначим $r = \dim K_{\lambda_1}^{m_1}$. Выбрав базис пространства $K_{\lambda_1}^{m_1}$, дополним его до базиса всего пространства. В этом базисе, согласно следствию теоремы 9.5, матрица A преобразования $\mathcal{A}$ будет иметь блочно-диагональный вид $A = diag(A_1, A_2)$, где квадратная матрица A_1 порядка r является матрицей сужения $\mathcal{A}_{K_{\lambda_1}^{m_1}}$ преобразования $\mathcal{A}$ на подпространство $K_{\lambda_1}^{m_1}$, а матрица A_2 является матрицей сужения $\mathcal{A}_L$. Характеристический многочлен матрицы A имеет вид (см. определитель блочно-диагональной матрицы в разд. 2.3.4)

$$\det(A - \lambda E) = \det(A_1 - \lambda E) \cdot \det(A_2 - \lambda E) = p_1(\lambda) \cdot p_2(\lambda),$$

где $p_1(\lambda)$, $p_2(\lambda)$ – многочлены степеней r и $(n - r)$ соответственно. Так как сужение $\mathcal{A}_{K_{\lambda_1}^{m_1}}$ не имеет собственных значений, отличных от λ_1, то $p_1(\lambda) = (-1)^r (\lambda - \lambda_1)^r$, в силу того, что $p_1(\lambda_1) = 0$ и основной теоремы алгебры. Поскольку сужение $\mathcal{A}_L$ не имеет собственных векторов, принадлежащих собственному значению λ_1, то $p_2(\lambda_1) \ne 0$. Следовательно, r – алгебраическая кратность собственного значения λ_1. Тогда утверждение теоремы следует из включения (9.8): $\dim K_{\lambda_1}^1 \le \dim K_{\lambda_1}^{m_1}$, так как $K_{\lambda_1}^1 \triangleleft K_{\lambda_1}^{m_1}$.

45. Теорема о диагонализируемости матрицы линейного преобразования и ее следствия.

Говорят, что линейное преобразование $\mathcal{A}: V \to V$ n-мерного линейного пространства V ***приводится к диагональному виду***, если существует базис, в котором матрица A преобразования диагональная, т.е. $A = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$, где $\lambda_1, \ldots, \lambda_n$ – некоторые числа, среди которых могут быть равные. Если преобразование $\mathcal{A}$ приводится к диагональному виду, то оно называется ***диагонализируемым***.

В разделе 7.2.2 было сформулировано необходимое и достаточное условие приводимости матрицы к диагональному виду. Переформулируем это условие для линейного преобразования: *линейное преобразование $\mathcal{A}: V \to V$ приводится к диагональному виду тогда и только тогда, когда в пространстве V существует базис из собственных векторов.*

Действительно, предположим, что в базисе $s_1, \ldots, s_n$ матрица преобразования имеет диагональный вид $A = diag(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_n \end{pmatrix}$. Найдем образ $\mathcal{A}(s_1)$. Умножая матрицу A на координатный столбец $s_1 = (1 \ 0 \ \cdots \ 0)^T$ базисного вектора $s_1 = 1 \cdot s_1 + 0 \cdot s_2 + \ldots + 0 \cdot s_n$, получаем $As_1 = \lambda_1 s_1$. Значит, $\mathcal{A}(s_1) = \lambda_1 s_1$, т.е. вектор s_1 является собственным, а преобразование $\mathcal{A}$ действует в подпространстве $Lin(s_1)$ как гомотетия (с коэффициентом λ_1). Аналогичный вывод можно сделать и про другие базисные векторы. Следовательно, базис пространства состоит из собственных векторов $s_1, \ldots, s_n$. Необходимость доказана. Достаточность доказывается путем приведения тех же рассуждений, но в обратном порядке.

Критерий диагонализируемости линейного преобразования можно сформулировать иначе.

Теорема 9.7 (о диагонализируемости линейного преобразования). *Для того чтобы линейное преобразование приводилось к диагональному виду, необходимо и достаточно, чтобы все корни характеристического многочлена являлись собственными значениями преобразования и геометрическая кратность каждого собственного значения была равна его алгебраической кратности.*

Достаточность следует из теорем 9.5 и 9.6. Действительно, если для каждого из различных собственных значений $\lambda_1, \ldots, \lambda_k$ геометрическая кратность равна алгебраической кратности, то $\dim K_{\lambda_i}^1 = \dim K_{\lambda_i}^{m_i}$, т.е. $K_{\lambda_i}^1 = K_{\lambda_i}^{m_i}$ для всех $i = 1, \ldots, k$. Поэтому в равенстве (9.9) корневые подпространства можно заменить собственными:

$$V = K_{\lambda_1}^1 \oplus \ldots \oplus K_{\lambda_k}^1.$$

Выбрав в каждом собственном подпространстве базис и объединив все эти базисы в единую систему, получим базис всего пространства, составленный из собственных векторов. Достаточность доказана. Необходимость доказывается путем приведения тех же рассуждений, но в обратном порядке.

В разд. 7.2.2 было доказано, что существование n линейно независимых собственных векторов необходимо и достаточно для диагонализируемости матрицы при помощи преобразования подобия (см. теорему 5.7). Теорема 9.7 дает более тонкое условие диагонализируемости матрицы линейного преобразования. В разд. 7.2.2 отмечалось также, что не всякую матрицу можно привести к диагональному виду при помощи преобразования подобия (см. пример 7.9). Аналогичный вывод справедлив и для линейного преобразования.

Следствие 1. *Если характеристическое уравнение линейного преобразования комплексного (вещественного) пространства имеет n попарно различных комплексных (действительных) корней (короче говоря, преобразование имеет* ***простой спектр****), то это преобразование приводится к диагональному виду.*

Действительно, в этом случае алгебраическая и геометрическая кратности каждого собственного значения равны единице.

46. Жорданова форма матрицы. Собственные и присоединенные векторы.

9.5.2. Приведение линейного преобразования к каноническому виду

Говорят, что линейное преобразование $\mathcal{A}: V \to V$ n-мерного линейного пространства V ***приводится к каноническому виду***, если существует базис, в котором матрица A преобразования имеет нормальную жорданову форму (см. разд.7.3). Такой базис называется ***жордановым***.

Напомним, что жордановой клеткой r-го порядка, соответствующей собственному значению λ, называют квадратную матрицу r-го порядка (7.35):

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}. \tag{9.10}$$

Жордановой матрицей называют блочно-диагональную матрицу вида (7.38):

$$J_A = diag(J_{r_1}(\lambda_1), \ldots, J_{r_k}(\lambda_k)) = \left(\begin{array}{c|c|c|c} J_{r_1}(\lambda_1) & O & \cdots & O \\ \hline O & J_{r_2}(\lambda_2) & \cdots & O \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O & O & \cdots & J_{r_k}(\lambda_k) \end{array}\right), \tag{9.11}$$

на диагонали которой стоят жордановы клетки (9.10), причем среди собственных значений $\lambda_1, \ldots, \lambda_k$ могут быть равные, порядки $r_1, \ldots, r_k$ жордановых клеток (всех или некоторых) могут совпадать.

СУЩЕСТВОВАНИЕ И СТРУКТУРА ЖОРДАНОВА БАЗИСА

Пусть преобразование $\mathcal{A}: V \to V$ имеет собственный вектор s, соответствующий собственному значению λ. Вектор $s^{(1)}$, удовлетворяющий условию $\mathcal{A}(s^{(1)}) = \lambda s^{(1)} + s$, называется ***присоединенным вектором 1-го порядка***. Вектор $s^{(2)}$, удовлетворяющий условию $\mathcal{A}(s^{(2)}) = \lambda s^{(2)} + s^{(1)}$, называется ***присоединенным вектором 2-го порядка*** и т.д. Присоединенный вектор $s^{(p)}$ p-го порядка определяется соотношением $\mathcal{A}(s^{(p)}) = \lambda s^{(p)} + s^{(p-1)}$, где $s^{(p-1)}$ – присоединенный вектор $(p-1)$-го порядка.

Определение присоединенных векторов можно записать эквивалентным образом, используя преобразование $\mathcal{B} = \mathcal{A} - \lambda\mathcal{E}$ (характеристическое преобразование для преобразования $\mathcal{A}$): вектор $s^{(p)}$ является присоединенным p-го порядка, если $\mathcal{B}^p(s^{(p)}) = s$, где s – собственный вектор преобразования $\mathcal{A}$.

Действительно, при $p = 1$ из условия $\mathcal{B}(s^{(1)}) = s$ получаем $(\mathcal{A} - \lambda\mathcal{E})(s^{(1)}) = s$, т.е. $\mathcal{A}(s^{(1)}) = \lambda s^{(1)} + s$. При $p = 2$ из условия $\mathcal{B}^2(s^{(2)}) = s$ имеем $(\mathcal{A} - \lambda\mathcal{E})^2(s^{(2)}) = (\mathcal{A} - \lambda\mathcal{E})[(\mathcal{A} - \lambda\mathcal{E})(s^{(2)})] = s$. Отсюда $(\mathcal{A} - \lambda\mathcal{E})(s^{(2)}) = s^{(1)}$, т.е. $\mathcal{A}(s^{(2)}) = \lambda s^{(2)} + s^{(1)}$ и т.д. Заметим, что присоединенный вектор $s^{(p)}$ p-го порядка по определению удовлетворяет одновременно двум условиям:

$$\mathcal{B}^p(s^{(p)}) \ne o \quad \text{и} \quad \mathcal{B}^{p+1}(s^{(p)}) = o,$$

так как собственный вектор $s = \mathcal{B}^p(s^{(p)})$ ненулевой, а $\mathcal{B}^{p+1}(s^{(p)}) = \mathcal{B}(s) = (\mathcal{A} - \lambda\mathcal{E})(s) = o$, поскольку $\mathcal{A}(s) = \lambda s$.

Необходимость рассмотрения присоединенных векторов объясняется следующим свойством: *жорданов базис состоит из собственных и присое-*

47. Теорема о приведении линейного преобразования к каноническому виду (без доказательства).

Теорема 9.8 (о приведении линейного преобразования к каноническому виду). *Если все корни характеристического уравнения являются собственными значениями преобразования, то это преобразование приводится к каноническому виду, т.е. существует базис пространства, в котором матрица преобразования имеет жорданову форму.*

Замечания 9.5.

1. По таблице (9.13) нетрудно выразить количество k_p жордановых клеток порядка p через ранги преобразований $r_p = \operatorname{rg}(\mathcal{A} - \lambda\mathcal{E})^p$, учитывая, что $\dim K_\lambda^p = n - r_p$. Количество k_m жордановых клеток порядка m равно количеству векторов $s_1^{(m-1)}, \ldots, s_{k_m}^{(m-1)}$ в первой строке таблицы (9.13): $k_m = \dim K_\lambda^m - \dim K_\lambda^{m-1} = (n - r_m) - (n - r_{m-1}) = r_{m-1} - r_m$. Количество k_{m-1} жордановых клеток порядка $(m-1)$ равно количеству векторов $s_1^{(m-2)}, \ldots, s_{k_{m-1}}^{(m-2)}$ во второй строке таблицы (9.13): $k_{m-1} = \dim K_\lambda^{m-1} - \dim K_\lambda^{m-2} - k_m = (n - r_{m-1}) - (n - r_{m-2}) - (r_{m-1} - r_m) = r_{m-2} - 2r_{m-1} + r_m$.

Продолжая вычисления аналогичным образом, получаем следующий результат: *количество k_p жордановых клеток порядка p находится по формуле*

$$k_p = r_{p-1} - 2r_p + r_{p+1}, \; p = 1, \ldots, m, \tag{9.14}$$

где $r_p = \operatorname{rg}(\mathcal{A} - \lambda\mathcal{E})^p$, $r_0 = n$; m *– наименьшее натуральное число, при котором* $r_{m+1} = r_m$.

2. Число m в (9.14) равно кратности корня λ минимального многочлена матрицы A линейного преобразования [10].

3. Из п.1 следует единственность жордановой формы матрицы линейного преобразования (с точностью до перестановок жордановых клеток), так как состав жордановых клеток (количество и порядки) полностью определяется по размерностям инвариантных подпространств (9.8), (9.9). От выбора базиса зависит расположение жордановых клеток на главной диагонали матрицы (9.11). Поэтому жордановы формы, отличающиеся перестановкой жордановых клеток, считаются одинаковыми.

48. *Алгоритм приведения линейного преобразования к каноническому виду.*

Задача приведения линейного преобразования к каноническому виду формулируется следующим образом. Требуется найти базис n-мерного линейного пространства V, в котором матрица линейного преобразования $\mathcal{A}: V \to V$ имеет жорданову форму J_A, т.е

– найти жорданову форму J_A матрицы A преобразования (*первый этап*);

– найти жорданов базис (*второй этап*).

НАХОЖДЕНИЕ ЖОРДАНОВОЙ ФОРМЫ МАТРИЦЫ ЛИНЕЙНОГО ПРЕОБРАЗОВАНИЯ

Для нахождения жордановой формы J_A матрицы линейного преобразования нужно выполнить следующие действия.

1. Выбрать произвольный базис $e_1, \ldots, e_n$ линейного пространства V и найти в этом базисе матрицу A преобразования $\mathcal{A}$.

2. Составить характеристический многочлен преобразования $\mathcal{A}$:

$$\Delta_{\mathcal{A}}(\lambda) = \det(A - \lambda E).$$

3. Найти все различные корни $\lambda_1, \ldots, \lambda_k$ характеристического уравнения $\Delta_{\mathcal{A}}(\lambda) = 0$ и их алгебраические кратности $n_1, \ldots, n_k$.

4. Для корня $\lambda = \lambda_1$ кратности n_1 найти ранги матриц

$$r_p = \operatorname{rg} B^p, \quad p = 1, \ldots, m_1,$$

где $B = A - \lambda_1 E$, а m_1 – наименьшее натуральное число ($m_1 \le n_1$), при котором $r_{m_1+1} = r_{m_1}$. Ранг матрицы B^p можно найти одним из способов, рассмотренных в разд 3.3. Если в дальнейшем предполагается искать жорданов базис, то для нахождения ранга лучше использовать метод Гаусса (разд. 3.3.2), приводя матрицу B^p элементарными преобразованиями строк к модифицированному ступенчатому виду $(B^p)_{\text{ст}}$. Он получается в результате вычеркивания нулевых строк из матрицы ступенчатого вида (см. разд. 1.6.1).

5. Определить количество k_p жордановых клеток $J_p(\lambda_1)$ порядка p:

$$k_p = r_{p-1} - 2r_p + r_{p+1}, \; p = 1, \ldots, m_1,$$

где $r_0 = n$ (см. п.1 замечаний 9.5).

Повторить п.4, 5 для остальных собственных значений $\lambda_2, \ldots, \lambda_k$.

6. Составить искомую матрицу J_A блочно-диагонального вида (9.11), располагая найденные жордановы клетки на главной диагонали.

НАХОЖДЕНИЕ ЖОРДАНОВА БАЗИСА

Пусть в базисе $(e) = (e_1, \ldots, e_n)$ линейного пространства V преобразование $\mathcal{A}: V \to V$ имеет матрицу A. Требуется найти матрицу перехода S от *диненных векторов линейного преобразования, взятых в определенном порядке.*

Предположим, что в базисе $e_1, e_2, \ldots, e_n$ пространства V матрица A преобразования $\mathcal{A}: V \to V$ имеет жорданову форму (9.11). Рассмотрим преобразование первых r_1 базисных векторов (r_1 – порядок первой жордановой клетки $J_{r_1}(\lambda_1)$ в (9.11)). Умножая координатный столбец $e_1 = (1 \ \ 0 \ \cdots \ 0)^T$ базисного вектора e_1 на матрицу A с учетом (9.10), получаем $Ae_1 = \lambda_1 e_1$, т.е. $\mathcal{A}(e_1) = \lambda_1 e_1$. Следовательно, вектор e_1 – собственный. Умножая матрицу A на координатный столбец $e_2 = (0 \ \ 1 \ \cdots \ 0)^T$ базисного вектора e_2, получаем $Ae_2 = \lambda_1 e_2 + e_1$, т.е. $\mathcal{A}(e_2) = \lambda_1 e_2 + e_1$. Следовательно, e_2 – присоединенный вектор (1-го порядка). Аналогично заключаем, что векторы $e_3, \ldots, e_{r_1}$ также присоединенные (от второго до $(r_1 - 1)$-го порядков соответственно). Для других жордановых клеток выводы аналогичные.

Таким образом, жорданов базис составляют собственные и присоединенные векторы, взятые в следующем порядке (базисные векторы $e_1, e_2, \ldots, e_n$ удобно обозначить иначе):

$$\underbrace{s_1, s_1^{(1)}, s_1^{(2)}, \ldots, s_1^{(r_1-1)}}_{J_{r_1}(\lambda_1)}, \underbrace{s_2, s_2^{(1)}, s_2^{(2)}, \ldots, s_2^{(r_2-1)}}_{J_{r_2}(\lambda_2)}, \ldots, \underbrace{s_k, s_k^{(1)}, s_k^{(2)}, \ldots, s_k^{(r_k-1)}}_{J_{r_k}(\lambda_k)}, \quad (9.12)$$

где $s_1, s_2, \ldots, s_k$ – собственные векторы, а остальные векторы – соответственно присоединенные к ним, $r_1 + r_2 + \ldots + r_k = n$. Каждой из k жордановых клеток в (9.11) отвечает одна группа векторов в (9.12). Количество собственных векторов в (9.12) равно количеству жордановых клеток в (9.11). Перестановке групп векторов в базисе (9.12) соответствует перестановка жордановых клеток в (9.11), при этом форма матрицы остается жордановой.

Докажем существование жорданова базиса для преобразования $\mathcal{A}: V \to V$ n-мерного линейного пространства V при условии, что все корни характеристического уравнения являются собственными значениями преобразования $\mathcal{A}$.

Сначала рассмотрим случай, когда преобразование $\mathcal{A}$ имеет единственное собственное значение λ. В этом случае в прямой сумме (9.9) имеется одно слагаемое $V = K_\lambda^m$ (см. теорему 9.5), а цепочка (9.8) инвариантных подпространств имеет вид

$$\{o\} \triangleleft K_\lambda^1 \triangleleft K_\lambda^2 \triangleleft \ldots \triangleleft K_\lambda^m = V,$$

базиса (e) к жорданову базису $(s) = (s_1, \ldots, s_n)$: $(s) = (e) \cdot S$ (см. разд.8.4.3). Предполагаем, что жорданова форма J_A матрицы A известна.

1. Для собственного значения λ_1 (алгебраической кратности n_1) найти характеристическую матрицу $B = A - \lambda_1 E$ и по жордановой форме J_A определить наибольший порядок m_1 жордановых клеток, соответствующих собственному значению λ_1.

2. Привести матрицы B^p, $p = 1, \ldots, m_1$, к модифицированному ступенчатому виду $(B^p)_{\text{ст}}$ (он получается в результате удаления нулевых строк из матрицы ступенчатого вида (см. разд.1.6.1)).

3. Найти фундаментальную матрицу Φ_{m_1} однородной системы уравнений

$$(B^{m_1})_{\text{ст}} (B^{m_1-1})^*_{\text{ст}} x = o$$

и составить матрицу $S^{(m_1-1)} = (B^{m_1-1})^*_{\text{ст}} \Phi_{m_1}$. Если B^{m_1} – нулевая матрица, то $S^{(m_1-1)} = (B^{m_1-1})^*_{\text{ст}}$, так как в этом случае Φ_{m_1} – единичная матрица.

Вычислить матрицу $BS^{(m_1-1)}$, найти фундаментальную матрицу Φ_{m_1-1} однородной системы уравнений

$$\left(\frac{(B^{m_1-1})_{\text{ст}}}{(BS^{(m_1-1)})^*} \right) (B^{m_1-2})^*_{\text{ст}} x = o$$

и составить матрицу $S^{(m_1-2)} = \left(BS^{(m_1-1)} \middle| (B^{m_1-2})^*_{\text{ст}} \Phi_{m_1-1} \right)$. Если однородная система не имеет фундаментальной матрицы (система имеет только тривиальное решение), то $S^{(m_1-2)} = BS^{(m_1-1)}$.

Вычислить матрицу $BS^{(m_1-2)}$, найти фундаментальную матрицу Φ_{m_1-2} однородной системы уравнений

$$\left(\frac{(B^{m_1-2})_{\text{ст}}}{(BS^{(m_1-2)})^*} \right) (B^{m_1-3})^*_{\text{ст}} x = o$$

и составить матрицу $S^{(m_1-3)} = \left(BS^{(m_1-2)} \middle| (B^{m_1-3})^*_{\text{ст}} \Phi_{m_1-2} \right)$. Если фундаментальная матрица Φ_{m_1-2} не существует, то $S^{(m_1-3)} = BS^{(m_1-2)}$.

Продолжить аналогичным образом построение матриц $S^{(m_1-4)}, \ldots, S^{(3)}, S^{(2)}$.

Вычислить матрицу $BS^{(2)}$, найти фундаментальную матрицу Φ_2 однородной системы уравнений

$$\left(\frac{(B^2)_{\text{ст}}}{(BS^{(2)})^*} \right) (B)^*_{\text{ст}} x = o$$

и составить матрицу $S^{(1)} = \left(BS^{(2)} \middle| (B)^*_{\text{ст}} \Phi_2 \right)$. Если фундаментальная матрица Φ_2 не существует, то $S^{(1)} = BS^{(2)}$.

Вычислить матрицу $BS^{(1)}$, найти фундаментальную матрицу Φ_1 однородной системы уравнений

$$\left(\frac{(B)_{\text{ст}}}{(BS^{(1)})^*} \right) x = o$$

и составить матрицу $S^{(0)} = \left(BS^{(1)} \middle| \Phi_1 \right)$. Если фундаментальная матрица Φ_1 не существует, то $S^{(0)} = BS^{(1)}$

Если $m_1 = 1$, то $S^{(0)} = \Phi_1$, где Φ_1 – фундаментальная матрица однородной системы уравнений $(B)_{\text{ст}} x = o$.

4. Из столбцов полученных матриц

$$\begin{gathered}
S^{(0)} = \left(B^{m_1-1} s_1^{(m_1-1)} \ldots B^{m_1-1} s_{k_{m_1}}^{(m_1-1)} \middle| \ \ldots \ \middle| B^2 s_1^{(2)} \ldots B^2 s_{k_3}^{(2)} \middle| Bs_1^{(1)} \ldots Bs_{k_2}^{(1)} \middle| s_1^{(0)} \ldots s_{k_1}^{(0)} \right), \\
S^{(1)} = \left(B^{m_1-2} s_1^{(m_1-1)} \ldots B^{m_1-2} s_{k_{m_1}}^{(m_1-1)} \middle| \ \ldots \ \middle| Bs_1^{(2)} \ldots Bs_{k_3}^{(2)} \middle| s_1^{(1)} \ldots s_{k_2}^{(1)} \right), \\
\ldots \qquad (9.15) \\
S^{(m_1-2)} = \left(Bs_1^{(m_1-1)} \ldots Bs_{k_{m_1}}^{(m_1-1)} \middle| s_1^{(m_1-2)} \ldots s_{k_{m_1-1}}^{(m_1-2)} \right), \\
S^{(m_1-1)} = \left(s_1^{(m_1-1)}, \ldots, s_{k_{m_1}}^{(m_1-1)} \right),
\end{gathered}$$

составить первые n_1 столбцов искомой матрицы S, записывая первые столбцы матриц $S^{(0)}, S^{(1)}, \ldots, S^{(m_1-2)}, S^{(m_1-1)}$, затем вторые столбцы этих матриц и т.д.

Выполнить п.1- 4 для остальных собственных значений $\lambda_2, \ldots, \lambda_k$, получая следующие $n_2, \ldots, n_k$ столбцов искомой матрицы S соответственно (при этом m_1 заменяется на $m_2, \ldots, m_k$).

Данный алгоритм использует метод нахождения относительных алгебраических дополнений, рассмотренный в разд. 8.6.5.

49 *Алгоритм нахождения многочлена от матрицы*.

7.3.4. Многочлены от матриц

Напомним определение многочлена от матрицы (см. разд.1.3.4). Пусть заданы многочлен (степени m) переменной λ:

$$f(\lambda)=a_m\lambda^m+a_{m-1}\lambda^{m-1}+\ldots+a_1\lambda+a_0, \tag{7.40}$$

A – квадратная матрица n-го порядка. Выражение вида

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\ldots+a_1A+a_0E \tag{7.41}$$

называется многочленом от матрицы A.

При больших значениях m и n вычисление выражения (7.41) затруднительно из-за операции возведения матрицы в натуральную степень. Поэтому требуется найти другие, эквивалентные определению (7.41), формы записи и алгоритмы эффективного вычисления многочлена от матрицы. Для упрощения (7.41) имеются две возможности. Во-первых, можно упростить матрицу A так, чтобы многочлен (7.40) от упрощенной матрицы уже вычислялся сравнительно просто. Например, выражение (7.41) легко вычисляется, если матрица A диагональная. Во-вторых, можно понизить степень m многочлена, тогда самая трудоемкая операция – возведение матрицы в степень – упрощается.

ИСПОЛЬЗОВАНИЕ ЖОРДАНОВОЙ ФОРМЫ МАТРИЦЫ

Использование жордановой формы для нахождения многочлена от матрицы основано на трех свойствах.

1. *Многочлены от подобных матриц подобны.*

Действительно, пусть при помощи преобразования подобия матрица A приведена к жордановой форме J_A: $J_A=S^{-1}AS$. Подставим $A=S\,J_AS^{-1}$ в правую часть (7.41):

$$f(A)=a_m\left(SJ_AS^{-1}\right)^m+a_{m-1}\left(SJ_AS^{-1}\right)^{m-1}+\ldots+a_1\left(SJ_AS^{-1}\right)+a_0E.$$

Учитывая, что $\left(SJ_AS^{-1}\right)^k=\underbrace{SJ_AS^{-1}\cdot SJ_AS^{-1}\cdot\ldots\cdot SJ_AS^{-1}}_{k\ \text{множителей}\ J_A}=S\,J_A^k\,S^{-1}$ для любого натурального k, получаем

$$f(A)=S\left(a_mJ_A^m+a_{m-1}J_A^{m-1}+\ldots+a_1J_A+a_0E\right)S^{-1}=S\,f(J_A)\,S^{-1}.$$

Таким образом, многочлены $f(A)$ и $f(J_A)$ подобны (с той же самой преобразующей матрицей S):

$$A=S\,J_AS^{-1}\quad\Rightarrow\quad f(A)=S\,f(J_A)\,S^{-1}.$$

2. *Многочлен от блочно-диагональной матрицы является блочно-диагональной матрицей.*

Пусть $A=\left(\begin{array}{c|c}A_1&O\\\hline O&A_2\end{array}\right)$, где A_1 и A_2 – квадратные матрицы, а O – нулевые матрицы соответствующих размеров. Для блочно-диагональных матриц справедливы равенства (они следуют из операций над блочными матрицами (см. разд.1.5.1)):

$$A^k=\left(\begin{array}{c|c}A_1&O\\\hline O&A_2\end{array}\right)^k=\left(\begin{array}{c|c}A_1^k&O\\\hline O&A_2^k\end{array}\right),\quad a_kA^k=\left(\begin{array}{c|c}a_kA_1^k&O\\\hline O&a_kA_2^k\end{array}\right),\ \text{где}\ k\in N.$$

Поэтому $f(A)=\left(\begin{array}{c|c}f(A_1)&O\\\hline O&f(A_2)\end{array}\right)$. Для большего числа блоков доказательство аналогичное.

3. *Многочлен (7.41) от жордановой клетки $J_r(\lambda_0)$ имеет вид*

$$f(J_r(\lambda_0))=\begin{pmatrix}f(\lambda_0)&\frac{1}{1!}f'(\lambda_0)&\cdots&\frac{1}{(r-1)!}f^{(r-1)}(\lambda_0)\\0&f(\lambda_0)&\cdots&\frac{1}{(r-2)!}f^{(r-2)}(\lambda_0)\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&f(\lambda_0)\end{pmatrix}. \tag{7.42}$$

Это верхняя треугольная матрица r-го порядка, на главной диагонали которой стоят значения функции $f(\lambda)$ в точке λ_0, над диагональю – значения первой производной в этой же точке и т.д., т.е. коэффициенты ряда Тейлора [19,25,44] для функции $f(\lambda)$.

Действительно, разложим многочлен (7.40) по формуле Тейлора в окрестности точки $\lambda=\lambda_0$:

$$f(\lambda)=f(\lambda_0)+\frac{1}{1!}f'(\lambda_0)(\lambda-\lambda_0)+\frac{1}{2!}f''(\lambda_0)(\lambda-\lambda_0)^2+\ldots+\frac{1}{m!}f^{(m)}(\lambda_0)(\lambda-\lambda_0)^m.$$

Остаточный член в данном случае равен нулю, так как все производные более высокого порядка, чем m, тождественно равны нулю. При вычислении $f(J_r(\lambda_0))$ линейный двучлен $(\lambda-\lambda_0)$ заменяется матрицей

$$I=(J_r(\lambda_0)-\lambda_0E)=\begin{pmatrix}\lambda_0&1&\cdots&0&0\\0&\lambda_0&\ddots&0&0\\\vdots&\vdots&\ddots&\ddots&\vdots\\0&0&\cdots&\lambda_0&1\\0&0&\cdots&0&\lambda_0\end{pmatrix}-\begin{pmatrix}\lambda_0&0&\cdots&0&0\\0&\lambda_0&\ddots&0&0\\\vdots&\vdots&\ddots&\ddots&\vdots\\0&0&\cdots&\lambda_0&0\\0&0&\cdots&0&\lambda_0\end{pmatrix}=$$

$$=\begin{pmatrix}0&1&\cdots&0&0\\0&0&\ddots&0&0\\\vdots&\vdots&\ddots&\ddots&\vdots\\0&0&\cdots&0&1\\0&0&\cdots&0&0\end{pmatrix},$$

у которой элементы над главной диагональю равны единице, а остальные элементы равны нулю, т.е. $I=(0\ \ e_1\ \ e_2\ \ \cdots\ \ e_{r-1})$, где e_i – i-й столбец единичной матрицы r-го порядка.

Можно показать, что при возведении в степень единичные элементы матрицы I смещаются вверх:

$$I^2=(0\ \ 0\ \ e_1\ \ e_2\ \ \cdots\ \ e_{r-2}),\ I^3=(0\ \ 0\ \ 0\ \ e_1\ \ \cdots\ \ e_{r-3})\quad\text{и т.д.,}$$

причем I^k – нулевая матрица при $k\ge r$. Подставляя эти матрицы в формулу Тейлора, получаем

$$f(A)=f(\lambda_0)\cdot E+\frac{1}{1!}f'(\lambda_0)\cdot I+\frac{1}{2!}f''(\lambda_0)\cdot I^2+\ldots+\frac{1}{m!}f^{(m)}(\lambda_0)\cdot I^m.$$

Складывая матрицы в правой части, получаем квадратную матрицу r-го порядка, у которой элементы главной диагонали равны $f(\lambda_0)$, элементы над главной диагональю равны $\frac{1}{1!}\cdot f'(\lambda_0)$ и т.д., т.е. матрицу вида (7.42).

Пример 7.16. Найти многочлен $f(\lambda)=\lambda^2+\lambda+1$ от матриц

а) $A=\begin{pmatrix}2&1&0\\0&2&1\\0&0&2\end{pmatrix}$, б) $B=\begin{pmatrix}2&1&0\\0&2&0\\0&0&2\end{pmatrix}$, в) $C=\begin{pmatrix}2&1&0\\0&2&0\\0&0&3\end{pmatrix}$.

☐ а) Матрица A – это жорданова клетка 3-го порядка, соответствующая собственному значению 2: $A=J_3(2)$. Находим значения функции и ее производных в точке $\lambda=2$: $f(2)=7$, $f'(2)=5$, $f''(2)=2$. Составляем матрицу вида (7.42), учитывая, что $r=3$:

$$f(A)=f(J_3(2))=\begin{pmatrix}f(2)&\frac{f'(2)}{1!}&\frac{f''(2)}{2!}\\0&f(2)&\frac{f'(2)}{1!}\\0&0&f(2)\end{pmatrix}=\begin{pmatrix}7&5&1\\0&7&5\\0&0&7\end{pmatrix}.$$

б) Матрица B имеет жорданову форму $B=diag(J_2(2),J_1(2))$, т.е. является блочно-диагональной. По свойству 2 многочлен от матрицы B является блочно-диагональной матрицей. Записываем многочлен $f(\lambda)$ от каждой жордановой клетки по формуле (7.42):

$$f(J_2(2))=\begin{pmatrix}7&5\\0&7\end{pmatrix},\quad f(J_1(2))=f(2)=7.$$

Здесь число 7 рассматривается как квадратная матрица 1-го порядка. Составляем из этих квадратных матриц искомую блочно-диагональную матрицу

$$f(B)=diag\left(\begin{pmatrix}7&5\\0&7\end{pmatrix},(7)\right)=\left(\begin{array}{cc|c}7&5&0\\0&7&0\\\hline 0&0&7\end{array}\right).$$

в) Матрица C имеет жорданову форму $C=diag(J_2(2),J_1(3))$, т.е. является блочно-диагональной. Записываем многочлен $f(\lambda)$ от каждой жордановой клетки по формуле (7.42):

$$f(J_2(2))=\begin{pmatrix}7&5\\0&7\end{pmatrix},\quad f(J_1(3))=f(3)=13.$$

Составляем из этих квадратных матриц искомую блочно-диагональную матрицу

$$f(C)=diag\left(\begin{pmatrix}7&5\\0&7\end{pmatrix},(13)\right)=\left(\begin{array}{cc|c}7&5&0\\0&7&0\\\hline 0&0&13\end{array}\right).\ ■$$

ПЕРВЫЙ СПОСОБ НАХОЖДЕНИЯ МНОГОЧЛЕНА ОТ МАТРИЦЫ

1. Привести матрицу A к жордановой форме $J_A=S^{-1}AS$, т.е. определить жорданову форму J_A и преобразующую матрицу S.

2. Составить блочно-диагональную матрицу $f(J_A)$, размещая на ее диагонали многочлены от жордановых клеток (7.42).

3. Найти многочлен от матрицы A по формуле $f(A)=S\,f(J_A)S^{-1}$.

50. Теорема Гамильтона – Кэли .

Теорема 7.7 (теорема Гамильтона–Кэли) *Характеристический многочлен матрицы является аннулирующим для нее, т.е. $\Delta_A(A)=O$.*

В самом деле, обозначим через $(A-\lambda E)^+$ матрицу, присоединенную к характеристической матрице $(A-\lambda E)$. Тогда из (7.7) следует:

$$(A-\lambda E)\cdot(A-\lambda E)^+=\Delta_A(\lambda)E\ \ \text{и}\ \ (A-\lambda E)^+\cdot(A-\lambda E)=\Delta_A(\lambda)E. \tag{7.27}$$

Правые части этих равенств можно рассматривать как многочлены с матричными коэффициентами (каждый коэффициент характеристического многочлена умножается на единичную матрицу). Из (7.27) следует, что λ-матрица $\Delta_A(\lambda)E$ делится на $(A-\lambda E)$ слева и справа без остатка, т.е. остаток равен нулевой матрице. По обобщенной теореме Безу (теорема 7.2) остаток равен левому и правому значениям многочлена $\Delta_A(\lambda)E$ при подстановке матрицы A вместо λ. Отсюда получаем $\Delta_A(A)E=O$, т.е. $\Delta_A(A)=O$, что и требовалось доказать.

Пример 7.11. Показать, что характеристический многочлен матрицы $A=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$ является для нее аннулирующим.

☐ Находим характеристический многочлен матрицы (см. пример 7.8)

$$\Delta_A(\lambda)=\det(A-\lambda E)=\begin{vmatrix}1-\lambda&1&1\\1&1-\lambda&1\\1&1&1-\lambda\end{vmatrix}=(1-\lambda)^3+2-3(1-\lambda)=3\lambda^2-\lambda^3.$$

Подставляя вместо переменной λ матрицу A, получаем

$$\Delta_A(A)=3\cdot A^2-A^3=3\cdot\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}^2-\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}^3=3\cdot\begin{pmatrix}3&3&3\\3&3&3\\3&3&3\end{pmatrix}-\begin{pmatrix}9&9&9\\9&9&9\\9&9&9\end{pmatrix}=O,$$

что и требовалось показать. ■

Теорема Гамильтона–Кэли говорит о том, что для квадратной матрицы A n-го порядка всегда найдется аннулирующий многочлен n-й степени (характеристический многочлен имеет n-ю степень). Возникает вопрос о существовании аннулирующего многочлена меньшей степени.

Минимальным многочленом матрицы A называется ее аннулирующий многочлен наименьшей степени (со старшим коэффициентом, равным единице). Минимальный многочлен будем обозначать $\mu_A(\lambda)$

51. Ортогональные преобразования: определение, примеры, свойства.

9.6.1. Ортогональные преобразования

Преобразование $\mathcal{A}:E\to E$ n-мерного евклидова пространства E называется ***ортогональным***, если оно сохраняет скалярное произведение векторов, т.е.

$$(\mathcal{A}(v),\mathcal{A}(w))=(v,w)\quad\forall\,v,w\in E. \tag{9.16}$$

Из определения следуют ***простейшие свойства***: *при ортогональном преобразовании не изменяются длины векторов, а также углы между векторами*, поскольку $|\mathcal{A}(v)|^2=(\mathcal{A}(v),\mathcal{A}(v))=(v,v)=|v|^2$ и для ненулевых векторов $\cos\varphi=\dfrac{(v,w)}{|v|\cdot|w|}=\dfrac{(\mathcal{A}(v),\mathcal{A}(w))}{|\mathcal{A}(v)|\cdot|\mathcal{A}(w)|}$.

Перейдем к изучению других свойств ортогональных преобразований.

СВОЙСТВА ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ

1. *Ортогональное преобразование – линейное.*

В определении намеренно не говорится о линейности преобразования $\mathcal{A}$, поскольку она следует из (9.16). Действительно, выберем в E ортонормированный базис $(e)=(e_1,\ldots,e_n)$. Тогда векторы $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$ также образуют ортонормированный базис пространства E, так как по определению

$$(\mathcal{A}(e_i),\mathcal{A}(e_j))=(e_i,e_j)=\begin{cases}1,&i=j,\\0,&i\ne j.\end{cases} \tag{9.17}$$

Найдем координаты образа $\mathcal{A}(v)$ произвольного вектора $v=v_1e_1+\ldots+v_ne_n$ в базисе $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$. Так как $(\mathcal{A}(v),\mathcal{A}(e_i))=(v,e_i)=v_i$, $i=1,\ldots,n$, получаем

$$\mathcal{A}(v)=v_1\mathcal{A}(e_1)+\ldots+v_n\mathcal{A}(e_n). \tag{9.18}$$

Найдем образ произведения вектора v на число λ

$$\mathcal{A}(\lambda\cdot v)=(\lambda v_1)\mathcal{A}(e_1)+\ldots+(\lambda v_n)\mathcal{A}(e_n)=\lambda\left(v_1\mathcal{A}(e_1)+\ldots+v_n\mathcal{A}(e_n)\right)=\lambda\cdot\mathcal{A}(v).$$

Следовательно, преобразование $\mathcal{A}$ – однородное. Аддитивность доказывается аналогично.

2. *Линейное преобразование ортогонально тогда и только тогда, когда оно отображает ортонормированный базис в ортонормированный.*

Необходимость следует из (9.17). Докажем достаточность. Пусть $e_1,\ldots,e_n$ и $\mathcal{A}(e_1),\ldots,\mathcal{A}(e_n)$ – ортонормированные базисы пространства E. В силу линейности преобразования для любого вектора $v=v_1e_1+\ldots+v_ne_n$ справедливо (9.18). Поэтому

$$(\mathcal{A}(v),\mathcal{A}(w))=\left(\sum_{i=1}^n v_i\mathcal{A}(e_i),\ \sum_{j=1}^n w_j\mathcal{A}(e_j)\right)=\sum_{i=1}^n\sum_{j=1}^n v_iw_j(\mathcal{A}(e_i),\mathcal{A}(e_j))=$$

$$=\sum_{i=1}^n v_iw_i=\sum_{i=1}^n\sum_{j=1}^n v_iw_j(e_i,e_j)=\left(\sum_{i=1}^n v_ie_i,\sum_{j=1}^n w_je_j\right)=(v,w),$$

что и требовалось доказать.

3. *Линейное преобразование $\mathcal{A}:E\to E$ ортогонально тогда и только тогда, когда его матрица A в любом ортонормированном базисе является ортогональной, т.е. $A^T=A^{-1}$.*

В самом деле, пусть в ортонормированном базисе $(e)=(e_1,\ldots,e_n)$ ортогональное преобразование $\mathcal{A}$ имеет матрицу A. Найдем произведение образов $\mathcal{A}(e_i)=\sum_{k=1}^n a_{ki}e_k$ и $\mathcal{A}(e_j)=\sum_{l=1}^n a_{lj}e_l$ базисных векторов (см. разд. 9.3.2). Согласно (9.17) имеем

$$(\mathcal{A}(e_i),\mathcal{A}(e_j))=\left(\sum_{k=1}^n a_{ki}e_k,\sum_{l=1}^n a_{lj}e_l\right)=\sum_{k=1}^n\sum_{l=1}^n a_{ki}a_{lj}(e_k,e_l)=\sum_{k=1}^n a_{ki}a_{kj}=\begin{cases}1,&i=j,\\0,&i\ne j.\end{cases}$$

Последнюю сумму можно рассматривать как произведение элементов i-й строки транспонированной матрицы A^T на соответствующие элементы j-го столбца матрицы A. Поэтому $A^TA=E$, тогда и $AA^T=E$. Следовательно, $A^T=A^{-1}$. Для доказательства достаточности проводим рассуждения в обратном порядке и приходим к заключению, что преобразование $\mathcal{A}$ (с ортогональной матрицей A) отображает один ортонормированный базис в другой. По свойству 2 такое преобразование ортогональное.

4. *Ортогональное преобразование обратимо, т.е. инъективно и сюръективно: $Ker\mathcal{A}=\{o\}$, $Im\,\mathcal{A}=E$.* Это следует из свойств ядра и образа линейного отображения (см. разд. 9.1.5).

5. *Корни характеристического уравнения ортогонального преобразования по модулю равны единице (собственные значения равны +1 или −1).*

Действительно, пусть A – матрица ортогонального преобразования в ортонормированном базисе. Характеристическое уравнение имеет действительные коэффициенты, так как матрица A действительная. Если $\lambda_{1,2}=\alpha\pm\beta i$ – пара комплексных сопряженных корней характеристического уравнения, то найдется ненулевой столбец z (с комплексными элементами), для которого $Az=\lambda z$. Выполняя комплексное сопряжение и транспонирование обеих частей равенства, получаем $z^*A^*=\overline{\lambda}z^*$ (см. разд. 1.4.2). Перемножим оба равенства: $z^*A^*Az=\lambda\overline{\lambda}z^*z$. Так как $A^*=A^T=A^{-1}$, то $z^*z=\lambda\overline{\lambda}z^*z$. Для ненулевого столбца ($z\ne o$) $z^*z\ne0$ и, следовательно, $\lambda\overline{\lambda}=|\lambda|^2=1$. Поэтому для собственных значений, которые являются действительными числами, получаем $\lambda=\pm1$.

6. *Определитель матрицы ортогонального преобразования равен +1 или −1.* Это свойство следует из равенства $AA^T=E$, так как

$$(\det A)^2=\det A\cdot\det A^T=\det(AA^T)=\det E=1\quad\Rightarrow\quad\det A=\pm1.$$

Ортогональное преобразование $\mathcal{A}$ называется ***собственным***, если $\det A=1$ и ***несобственным***, если $\det A=-1$.

7. *Пусть L – инвариантное относительно ортогонального преобразования $\mathcal{A}:E\to E$ подпространство E. Тогда его ортогональное дополнение $L^\perp$ также инвариантно по отношению к преобразованию $\mathcal{A}$.*

По свойству 4 ортогональное преобразование $\mathcal{A}_L:L\to L$ (сужение преобразования $\mathcal{A}:E\to E$ на инвариантное подпространство $L\triangleleft E$) обратимо. Поэтому для любого $w\in L$ найдется прообраз $v\in L$: $w=\mathcal{A}(v)$. Тогда для любого $u\in L^\perp$ имеем $(\mathcal{A}(u),w)=(\mathcal{A}(u),\mathcal{A}(v))=(u,v)=0$, т.е. $\mathcal{A}(u)\in L^\perp$. Следовательно, подпространство $L^\perp\triangleleft E$ инвариантно относительно преобразования $\mathcal{A}$, т.е. $\mathcal{A}(L^\perp)\subset L^\perp$ и даже $\mathcal{A}(L^\perp)=L^\perp$ в силу обратимости $\mathcal{A}$.

8. *Пусть $\lambda=\alpha+\beta i$ – пара комплексных сопряженных корней ($\beta\ne0$) характеристического многочлена ортогонального преобразования $\mathcal{A}:E\to E$. Тогда существует такая пара равных по длине ортогональных векторов x и y, что*

$$\begin{cases}\mathcal{A}(x)=\alpha x-\beta y,\\\mathcal{A}(y)=\beta x+\alpha y.\end{cases} \tag{9.19}$$

В самом деле, в теореме 9.4 доказано существование линейно независимых векторов x и y, удовлетворяющих системе (9.19), которая в координатной форме имеет вид (9.7). Матрица A_L сужения $\mathcal{A}_L:L\to L$ преобразования $\mathcal{A}$ на двумерное инвариантное подпространство $L=Lin(x,y)$ относительно базиса x, y имеет вид $A_L=\begin{pmatrix}\alpha&\beta\\-\beta&\alpha\end{pmatrix}$. Она составлена из координатных столбцов векторов $\mathcal{A}(x)$, $\mathcal{A}(y)$ в базисе x, y, т.е. из коэффициентов разложений (9.19) этих векторов по базису.

Докажем ортогональность векторов x и y. Находим скалярные произведения, учитывая (9.19) и ортогональность преобразования:

$$|x|^2=(\mathcal{A}(x),\mathcal{A}(x))=\alpha^2|x|^2-2\alpha\beta(x,y)+\beta^2|y|^2,$$

$$(x,y)=(\mathcal{A}(x),\mathcal{A}(y))=\alpha\beta|x|^2+(\alpha^2-\beta^2)(x,y)-\alpha\beta|y|^2.$$

Подставляя $\alpha^2=1-\beta^2$ (см. свойство 5) и сокращая на $\beta\ne0$, получаем систему уравнений

$$\begin{cases}\beta\left(|y|^2-|x|^2\right)-2\alpha(x,y)=0,\\-\alpha\left(|y|^2-|x|^2\right)-2\beta(x,y)=0\end{cases}$$

относительно двух неизвестных $\left(|y|^2-|x|^2\right)$ и (x,y). Определитель матрицы системы $\begin{vmatrix}\beta&-2\alpha\\-\alpha&-2\beta\end{vmatrix}=-2(\alpha^2+\beta^2)=-2\ne0$, следовательно, система имеет только тривиальное решение: $|y|^2-|x|^2=0$, $(x,y)=0$, т.е. векторы x и y имеют равные длины и перпендикулярны.

52. Канонический вид ортогонального преобразования и его геометрический смысл.

КАНОНИЧЕСКИЙ ВИД ОРТОГОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ

Рассмотрим инвариантные подпространства ортогонального преобразования. По теореме 9.4 линейное преобразование вещественного пространства имеет одномерное или двумерное инвариантное подпространство. Выясним геометрический смысл сужения ортогонального преобразования на инвариантное подпространство.

1. Пусть L – одномерное инвариантное подпространство с базисом e_1. Тогда, e_1 – собственный вектор преобразования: $\mathcal{A}(e_1)=\lambda e_1$. По свойству 5 $\lambda=\pm1$. Следовательно, ортогональное преобразование $\mathcal{A}:L\to L$ одномерного пространства – это либо тождественное преобразование $\mathcal{A}(e_1)=e_1$, либо отражение (симметрия) $\mathcal{A}(e_1)=-e_1$.

2. Пусть L – двумерное инвариантное подпространство с ортонормированным базисом e_1,e_2. Запишем для матрицы $A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ ортогонального преобразования $\mathcal{A}:L\to L$ равенство $A^T=A^{-1}$:

$$\begin{pmatrix}a&c\\b&d\end{pmatrix}=\frac{1}{\det A}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}.$$

По свойству 6 для собственного преобразования $\det A=1$, поэтому $d=a$, $c=-b$, $\det A=a^2+b^2=1$, т.е. $a=\cos\varphi$ и $b=\sin\varphi$, где φ – некоторый угол. Следовательно, матрица собственного ортогонального преобразования двумерного пространства совпадает с матрицей поворота на угол $(-\varphi)$ (см. п.5 в разд.9.2.1): $A=R_\varphi=\begin{pmatrix}\cos\varphi&\sin\varphi\\-\sin\varphi&\cos\varphi\end{pmatrix}$. Для несобственного преобразования $\det A=-1$ (см. свойство 6), поэтому $d=-a$, $c=b$, $\det A=-a^2-b^2=-1$. Матрица $A=\begin{pmatrix}a&b\\b&-a\end{pmatrix}$ имеет два действительных собственных значения ($\lambda_1=1$, $\lambda_2=-1$), так как $\det(A-\lambda E)=\begin{vmatrix}a-\lambda&b\\b&-a-\lambda\end{vmatrix}=\lambda^2-a^2-b^2=\lambda^2-1$. Поэтому несобственное ортогональное преобразование имеет два одномерных инвариантных подпространства (см. п.1).

Получим канонический вид преобразования. Пусть L – одномерное или двумерное инвариантное подпространство для ортогонального преобразования $\mathcal{A}:E\to E$. Представим пространство в виде прямой суммы $E=L\oplus L^\perp$. Выберем в L ортонормированный базис и дополним его до ортонормированного базиса всего пространства. В этом базисе матрица A преобразования будет иметь блочно-диагональный вид $A=diag(A_L,A_{L^\perp})$, где A_L – матрица сужения $\mathcal{A}_L$ преобразования $\mathcal{A}$ на L, а $A_{L^\perp}$ – матрица сужения $\mathcal{A}_{L^\perp}$ преобразования $\mathcal{A}$ на $L^\perp$. Согласно п.1, 2: $A_L=(1)$ или $A_L=(-1)$ при $\dim L=1$, либо $A_L=R_\varphi=\begin{pmatrix}\cos\varphi&\sin\varphi\\-\sin\varphi&\cos\varphi\end{pmatrix}$ при $\dim L=2$. Следовательно, в матрице $A=diag(A_L,A_{L^\perp})$ ортогонального преобразования блок A_L имеет один из указанных трех видов. Поскольку подпространство $L^\perp$ инвариантно относительно ортогонального преобразования $\mathcal{A}$ (см. свойство 7), то к матрице $A_{L^\perp}$ применимы те же выводы, что и к матрице A. Таким образом, справедливо следующее утверждение.

Теорема 9.9 (о каноническом виде ортогонального преобразования). *Для каждого ортогонального преобразования $\mathcal{A}:E\to E$ n-мерного евклидова пространства E существует ортонормированный базис, в котором матрица преобразования имеет канонический вид:*

$$\underset{(s)}{A}=\begin{pmatrix}\pm1&&&&&\\&\ddots&&&O&\\&&+1&&&\\&&&R_{\varphi_1}&&\\&O&&&\ddots&\\&&&&&R_{\varphi_k}\end{pmatrix}. \tag{9.20}$$

На главной диагонали матрицы стоят либо числа 1 или (-1), либо блоки вида $R_\varphi=\begin{pmatrix}\cos\varphi&\sin\varphi\\-\sin\varphi&\cos\varphi\end{pmatrix}$, а остальные элементы матрицы равны нулю.

Базис $(s)=(s_1,\ldots,s_n)$, в котором матрица преобразования имеет вид (9.20), называется ***каноническим***. Заметим, что канонический базис определяется неоднозначно.

53 *Алгоритм приведения ортогонального преобразования к каноническому виду.*

ПРИВЕДЕНИЕ ОРТОГОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ К КАНОНИЧЕСКОМУ ВИДУ

Задача приведения ортогонального преобразования к каноническому виду формулируется следующим образом: требуется найти базис (*канонический*), в котором матрица ортогонального преобразования имеет канонический вид (9.20). Для приведения ортогонального преобразования к каноническому виду нужно выполнить следующие действия.

Нахождение канонического вида ортогонального преобразования (*первый этап*).

1. Выбрать базис $e_1,\ldots,e_n$ пространства E и найти матрицу A преобразования в этом базисе.

2. Составить характеристическое уравнение $\det(A-\lambda E)=0$ и найти различные его корни $\lambda_1,\ldots,\lambda_k$ (а также их алгебраические кратности).

3. Записать блочно-диагональную матрицу (9.20) канонического вида ортогонального преобразования:

каждый действительный корень λ_1 кратности n_1 поместить на главной диагонали n_1 раз;

для каждой пары $\lambda=\alpha\pm\beta i$ комплексных сопряженных корней кратности m записать m блоков вида $R_\varphi=\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix}$ (см. доказательство свойства 8 ортогональных преобразований).

Нахождение канонического базиса (*второй этап*).

4. Для действительного корня λ_1 кратности n_1 найти фундаментальную систему $x_1,\ldots,x_{n_1}$ решений однородной системы $(A-\lambda_1E)x=o$. Линейно независимую систему $x_1,\ldots,x_{n_1}$ векторов (пространства R^n) ортогонализировать и нормировать (см. разд. 8.8.5). Получим векторы $s_1,\ldots,s_{n_1}$.

5. Для пары $\lambda=\alpha\pm\beta i$ комплексных сопряженных корней кратности m найти фундаментальную систему $z_1,\ldots,z_m$ решений однородной системы $(A-(\alpha+\beta i)E)z=o$. Выделяя действительные $x_j=\operatorname{Re} z_j$ и мнимые части $y_j=\operatorname{Im} z_j$, $j=1,\ldots,m$, комплексных столбцов $z_1,\ldots,z_m$, получить m пар ортогональных векторов $x_1,y_1;x_2,y_2;\ldots;x_m,y_m$ (пространства R^n). Эту систему векторов ортогонализировать и нормировать. Получим $2m$ векторов $s_1,\ldots,s_{2m}$.

6. Выполнить п.4 или п.5 для всех различных корней характеристического уравнения. Получаемые в результате группы столбцов последовательно записать в матрицу S перехода от базиса $e_1,\ldots,e_n$ к искомому каноническому базису $s_1,\ldots,s_n$: $(s)=(e)S$. Матрица $S^{-1}AS$ преобразования $\mathcal{A}$ будет иметь канонический вид (9.20), полученный в п.3.

54. Сопряженные преобразования: определение, примеры, свойства .

Пусть $\mathcal{A}:E\to E$ – линейное преобразование n-мерного евклидова пространства E. Преобразование $\mathcal{A}^*:E\to E$ называется ***сопряженным*** преобразованию $\mathcal{A}$, если для любых векторов x и y из пространства E выполняется равенство

$$(\mathcal{A}(x),y)=(x,\mathcal{A}^*(y)). \tag{9.21}$$

СВОЙСТВА СОПРЯЖЕННОГО ПРЕОБРАЗОВАНИЯ

1. *Сопряженное преобразование – линейное.*

Докажем, например, однородность: $\mathcal{A}^*(\lambda y)=\lambda\mathcal{A}^*(y)$ $\forall\lambda\in R$. Пусть $(e)=(e_1,\ldots,e_n)$ – ортонормированный базис пространства E. Тогда

$$(e_1,\mathcal{A}^*(\lambda y))=(\mathcal{A}(e_1),\lambda y)=\lambda(\mathcal{A}(e_1),y)=\lambda(e_1,\mathcal{A}^*(y))=(e_1,\lambda\mathcal{A}^*(y)),$$

т.е. первые координаты векторов $\mathcal{A}^*(\lambda y)$ и $\lambda\mathcal{A}^*(y)$ равны. Аналогично показывается, что равны и остальные координаты этих векторов. Значит, это равные векторы. Аддитивность сопряженного преобразования доказывается аналогично.

2. *Для каждого линейного преобразования существует единственное сопряженное преобразование, причем матрица сопряженного преобразования (в любом ортонормированном базисе) является транспонированной по отношению к матрице данного преобразования (в том же базисе).*

Пусть в ортонормированном базисе $(e)=(e_1,\ldots,e_n)$ преобразование $\mathcal{A}$ имеет матрицу A. Рассмотрим преобразование $\mathcal{A}'$, которое в данном базисе имеет матрицу A^T. Для координатных столбцов x, y любых векторов x, y имеем равенство $(Ax,y)=x^TA^Ty=(x,A^Ty)$ (см. (8.27) в разд.8.8.2). Следовательно, $(\mathcal{A}(x),y)=(x,\mathcal{A}'(y))$, т.е. согласно (9.21) преобразование $\mathcal{A}'$ – сопряженное: $\mathcal{A}'=\mathcal{A}^*$. Итак, сопряженное преобразование существует и его матрица в любом ортонормированном базисе является транспонированной A^T по отношению к матрице данного преобразования. Отсюда также следует единственность, так как транспонированная матрица находится однозначно.

3. *Если L – подпространство, инвариантное относительно линейного преобразования $\mathcal{A}:E\to E$, то его ортогональное дополнение $L^\perp$ является инвариантным подпространством относительно сопряженного преобразования $\mathcal{A}^*$.* Действительно, покажем, что образ $\mathcal{A}^*(y)$ любого вектора $y\in L^\perp$ ортогонален любому вектору $x\in L$, т.е $\mathcal{A}^*(y)\in L^\perp$. Учитывая, что $\mathcal{A}(x)\in L$, по определению (9.21) получаем $(x,\mathcal{A}^*(y))=(\mathcal{A}(x),y)=0$, что и требовалось доказать.

Замечания 9.8.

1. Из второго свойства следует, что на сопряженные преобразования переносятся свойства транспонированных матриц (см. разд.1.4.1). В частности: $(\mathcal{A}^*)^*=\mathcal{A}$, $(\mathcal{A}\mathcal{B})^*=\mathcal{B}^*\mathcal{A}^*$, а также $(\mathcal{A}^{-1})^*=(\mathcal{A}^*)^{-1}$ для обратимого преобразования.

2. Матрица A^* сопряженного преобразования $\mathcal{A}^*$ в произвольном (неортонормированном) базисе связана с матрицей A преобразования $\mathcal{A}$ следующей формулой

$$A^*=G^{-1}A^TG,$$

где G – матрица Грама данного базиса.

3. Условие ортогональности преобразования $\mathcal{A}$ (см. свойство 2) можно представить в виде $\mathcal{A}^*=\mathcal{A}^{-1}$.

55. Самосопряженные преобразования: определение, примеры, свойства.

9.6.3. Самосопряженные преобразования

Линейное преобразование $\mathcal{A}:E\to E$ n мерного евклидова пространства E называется ***самосопряженным***, если оно является сопряженным самому себе, а именно $\mathcal{A}^*=\mathcal{A}$, т.е. $(\mathcal{A}(x),y)=(x,\mathcal{A}(y))$ для любых векторов x и y из пространства E.

Например, самосопряженными преобразованиями являются нулевое преобразование $\mathcal{O}$ и тождественное $\mathcal{E}$.

СВОЙСТВА САМОСОПРЯЖЕННОГО ПРЕОБРАЗОВАНИЯ

1. *Матрица A самосопряженного преобразования в любом ортонормированном базисе является симметрической* ($A^T=A$), и наоборот, *если в каком-либо ортонормированном базисе матрица преобразования симметрическая, то это преобразование самосопряженное.*

2. *Все корни характеристического уравнения самосопряженного преобразования действительные.*

В самом деле, предположим противное, а именно существование пары комплексных сопряженных корней $\lambda=\alpha\pm\beta i$, $\beta\ne0$. По теореме 9.4 преобразование имеет двумерное инвариантное подпространство с линейно независимыми образующими x и y, удовлетворяющими системе (9.19), которая следует из (9.7):

$$\begin{cases}\mathcal{A}(x)=\alpha x-\beta y,\\ \mathcal{A}(y)=\beta x+\alpha y.\end{cases}$$

Найдем скалярные произведения:

$$(\mathcal{A}(x),y)=\alpha(x,y)-\beta|y|^2,\quad(x,\mathcal{A}(y))=\beta|x|^2+\alpha(x,y).$$

Левые части равенств совпадают из-за самосопряженности преобразования $\mathcal{A}$. Значит, равны и правые части: $\alpha(x,y)-\beta|y|^2=\beta|x|^2+\alpha(x,y)$. Отсюда $\beta(|x|^2+|y|^2)=0$. Поскольку $\beta\ne0$, то $x=y=o$, что противоречит линейной независимости x и y.

3. *Собственные векторы, принадлежащие различным собственным значениям самосопряженного преобразования, ортогональны.*

Действительно, пусть $\mathcal{A}(x)=\lambda_1x$ и $\mathcal{A}(y)=\lambda_2y$, $\lambda_1\ne\lambda_2$. Тогда

$$(\mathcal{A}(x),y)=\lambda_1(x,y)\ \text{и}\ (x,\mathcal{A}(y))=\lambda_2(x,y).$$

Так как $(\mathcal{A}(x),y)=(x,\mathcal{A}(y))$, то $\lambda_1(x,y)=\lambda_2(x,y)\ \Leftrightarrow\ (\lambda_1-\lambda_2)(x,y)=0$. Отсюда $(x,y)=0$, так как $\lambda_1\ne\lambda_2$. Значит, собственные векторы x и y ортогональны.

4. *Если L – подпространство, инвариантное относительно самосопряженного преобразования $\mathcal{A}:E\to E$, то его ортогональное дополнение $L^\perp$ также инвариантно относительно преобразования $\mathcal{A}$.*

Это следует из свойства 3 сопряженных преобразований (см. разд. 9.6.2).

56. Теорема о диагонализируемости матрицы самосопряженного преобразования

Теорема 9.10 (о диагонализируемости самосопряженного преобразования). *Для всякого самосопряженного преобразования $\mathcal{A}:E\to E$ n-мерного евклидова пространства E существует ортонормированный базис (из собственных векторов), в котором матрица преобразования имеет диагональный вид*

$$\Lambda=diag(\lambda_1,\ldots,\lambda_n), \tag{9.22}$$

где $\lambda_1,\ldots,\lambda_n$ – собственные значения преобразования $\mathcal{A}$, повторенные в соответствии с их кратностью.

Диагональный вид (9.22) называется также ***каноническим видом*** самосопряженного преобразования, а базис, в котором матрица имеет вид (9.22), – ***каноническим***.

Для доказательства теоремы 9.10 нужно показать, что если существует ортонормированный базис пространства E, состоящий из собственных векторов преобразования, тогда оно приводится к диагональному виду (см. разд. 9.5.1). Действительно, для собственного значения λ_1 найдем единичный собственный вектор s_1. Представим пространство в виде прямой суммы $E=L_1\oplus L_1^\perp$, где $L_1=Lin(s_1)$ – одномерное инвариантное подпространство. Сужение преобразования $\mathcal{A}$ (по свойству 4) на инвариантное подпространство $L_1^\perp$ является самосопряженным. Поэтому в $L_1^\perp$ можно найти одномерное инвариантное подпространство $L_2=Lin(s_2)$, где s_2 – собственный вектор, перпендикулярный s_1. Продолжая аналогичным образом, получим $E=L_1\oplus\ldots\oplus L_n$, где $L_i=Lin(s_i)$ – одномерное инвариантное подпространство, причем базис $s_1,\ldots,s_n$ из собственных векторов ортогональный, а после нормировки – ортонормированный.

57 *Алгоритм приведения самосопряженного преобразования к каноническому виду.*

ПРИВЕДЕНИЕ САМОСОПРЯЖЕННОГО ПРЕОБРАЗОВАНИЯ К ДИАГОНАЛЬНОМУ ВИДУ

Пусть в некотором ортонормированном базисе $(e)=(e_1,\ldots,e_n)$ самосопряженное преобразование $\mathcal{A}:E\to E$ имеет матрицу A. Требуется найти базис $(s)=(s_1,\ldots,s_n)$, в котором матрица преобразования имеет диагональный вид (9.22). Для решения задачи нужно выполнить следующие действия.

Нахождение диагонального вида матрицы самосопряженного преобразования (*первый этап*).

1. Составить характеристическое уравнение $\det(A-\lambda E)=0$, найти его корни $\lambda_1,\ldots,\lambda_k$ и их алгебраические кратности $n_1,\ldots,n_k$, $n_1+\ldots+n_k=n$.

2. Составить искомую диагональную матрицу (9.22):

$$\Lambda=diag(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2},\ldots,\underbrace{\lambda_k,\ldots,\lambda_k}_{n_k}).$$

Нахождение матрицы S перехода от данного базиса (e) к каноническому базису (s) (*второй этап*).

3. Для корня λ_1 кратности n_1 найти фундаментальную систему $\varphi_1,\ldots,\varphi_{n_1}$ решений однородной системы $(A-\lambda_1E)x=o$. Столбцы $\varphi_1,\ldots,\varphi_{n_1}$ ортогонализировать и нормировать. Получим n_1 столбцов $s_1,\ldots,s_{n_1}$.

4. Записать полученные столбцы $s_1,\ldots,s_{n_1}$ в первые n_1 столбцов матрицы S.

Выполнить п.3, 4 для остальных собственных значений $\lambda_2,\ldots,\lambda_k$, добавляя полученные столбцы в матрицу S. В результате получим искомую матрицу перехода: $(s)=(e)S$.

Пример 9.6. Самосопряженное преобразование $\mathcal{A}:E\to E$ в ортонормированном базисе e_1,e_2,e_3 имеет матрицу $A=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$. Привести это преобразование к диагональному виду, т.е найти ортонормированный базис s_1,s_2,s_3, в котором матрица преобразования имеет диагональный вид (9.22), и найти эту диагональную матрицу.

□ *Первый этап.* Находим диагональный вид матрицы преобразования.

1. При решении примера 9.2 были найдены корни характеристического уравнения $\lambda_1=0$ (кратности $n_1=2$) и $\lambda_2=3$ (кратности $n_2=1$).

2. Составляем искомую диагональную матрицу $\Lambda=diag(0,0,3)$.

Нахождение матрицы S перехода к каноническому базису (*второй этап*).

3^1. Для собственного значения $\lambda_1=0$ в примере 9.2 была найдена фундаментальная система решений $\varphi_1=(1\ \ 0\ \ -1)^T$, $\varphi_2=(1\ \ -1\ \ 0)^T$. Ортогонализируем их, используя метод Грама–Шмидта (см. разд.8.8.5). Положим $\psi_1=\varphi_1=(1\ \ 0\ \ -1)^T$, $\psi_2=\varphi_2-\alpha\psi_1$. Коэффициент α выбираем из условия ортогональности $(\psi_1,\psi_2)=0$:

$$(1\ \ 0\ \ -1)\left[\begin{pmatrix}1\\-1\\0\end{pmatrix}-\alpha\begin{pmatrix}1\\0\\-1\end{pmatrix}\right]=1\cdot(1-\alpha)+0\cdot(-1)+(-1)\cdot\alpha=0\ \Rightarrow\ 1-\alpha\cdot2=0.$$

Следовательно, $\alpha=0{,}5$ и $\psi_2=(0{,}5\ \ -1\ \ 0{,}5)^T$. Нормируем столбцы ($|\psi_1|=\sqrt2$, $|\psi_2|=\sqrt{\tfrac32}$):

$$s_1=\frac{1}{|\psi_1|}\psi_1=\left(\frac{\sqrt2}{2}\ \ 0\ \ -\frac{\sqrt2}{2}\right)^T,\quad s_2=\frac{1}{|\psi_2|}\psi_2=\left(\frac{\sqrt6}{6}\ \ -\frac{\sqrt6}{3}\ \ \frac{\sqrt6}{6}\right)^T.$$

4^1. Полученные столбцы записываем в искомую матрицу $S=\begin{pmatrix}\frac{\sqrt2}{2}&\frac{\sqrt6}{6}&*\\0&-\frac{\sqrt6}{3}&*\\-\frac{\sqrt2}{2}&\frac{\sqrt6}{6}&*\end{pmatrix}$. Звездочкой обозначены неизвестные пока элементы матрицы.

3^2. Для собственного значения $\lambda_2=3$ фундаментальная система решений содержит одно решение $\varphi_3=(1\ \ 1\ \ 1)^T$ (см. пример 9.2). Нормируя этот столбец, получаем $s_3=\left(\frac{\sqrt3}{3}\ \ \frac{\sqrt3}{3}\ \ \frac{\sqrt3}{3}\right)^T$.

4^2. Полученный столбец дописываем в матрицу, полученную в п. 4^1,

$$S=\begin{pmatrix}\frac{\sqrt2}{2}&\frac{\sqrt6}{6}&\frac{\sqrt3}{3}\\0&-\frac{\sqrt6}{3}&\frac{\sqrt3}{3}\\-\frac{\sqrt2}{2}&\frac{\sqrt6}{6}&\frac{\sqrt3}{3}\end{pmatrix}.$$

Матрица перехода к каноническому базису найдена. ■

58. Теорема о структуре невырожденного линейного преобразования евклидова пространства и ее геометрический смысл .

Самосопряженное преобразование называется ***положительным (неотрицательным)***, если $(\mathcal{A}(x),x)>0$ для любого ненулевого вектора $x\in E$ (соответственно $(\mathcal{A}(x),x)\ge0$ для любого вектора $x\in E$).

Эти понятия связаны с положительностью (неотрицательностью) симметрических матриц и квадратичных форм (см. разд. 6.5.4). Действительно, запишем неравенство $(\mathcal{A}(x),x)\ge0$ в координатной форме (в ортонормированном базисе). Учитывая, что $(\mathcal{A}(x),x)=(Ax,x)=x^TA^Tx=x^TAx$, получаем $x^TAx\ge0$ для любого столбца $x\in R^n$, что совпадает с определением неотрицательности квадратичной формы x^TAx.

Отметим следующие свойства положительных и неотрицательных преобразований.

1. *Преобразование $\mathcal{A}$ положительно (неотрицательно) тогда и только тогда, когда все его собственные значения положительные (неотрицательные).*

2. *Для любого неотрицательного (положительного) преобразования $\mathcal{A}$ существует такое единственное неотрицательное (положительное) преобразование $\mathcal{B}$*, что $\mathcal{B}^2=\mathcal{A}$.

Действительно, в каноническом базисе $(s)=(s_1,\ldots,s_n)$ матрица преобразования $\mathcal{A}$ имеет диагональный вид (9.22). Преобразование $\mathcal{B}$ определяем его матрицей в базисе (s), полагая $B=diag(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n})$. Тогда $B^2=A$.

3. *Преобразования $\mathcal{A}^*\mathcal{A}$ и $\mathcal{A}\mathcal{A}^*$ являются самосопряженными неотрицательными (положительными) для любого (невырожденного) преобразования $\mathcal{A}$.*

Теорема 9.11 (о разложении невырожденного линейного преобразования). *Любое невырожденное линейное преобразование $\mathcal{A}:E\to E$ n-мерного евклидова пространства E можно представить в виде композиции положительного самосопряженного преобразования и ортогонального преобразования.*

Действительно, рассмотрим самосопряженное положительное преобразование $\mathcal{D}=\mathcal{A}^*\mathcal{A}$ (см. свойство 3). Для него существует такое положительное самосопряженное преобразование $\mathcal{S}$, что $\mathcal{D}=\mathcal{S}\mathcal{S}=\mathcal{S}^*\mathcal{S}$ (свойство 2). Рассмотрим преобразование $\mathcal{B}=\mathcal{A}\mathcal{S}^{-1}$. Это преобразование ортогональное (см. п.3 замечаний 9.8), так как

$$\mathcal{B}^*\mathcal{B}=(\mathcal{S}^{-1})^*\underbrace{\mathcal{A}^*\mathcal{A}}_{\mathcal{D}}\mathcal{S}^{-1}=(\mathcal{S}^{-1})^*\mathcal{D}\mathcal{S}^{-1}=(\mathcal{S}^{-1})^*\underbrace{\mathcal{S}^*\mathcal{S}}_{\mathcal{D}}\mathcal{S}^{-1}=\underbrace{(\mathcal{S}\mathcal{S}^{-1})^*}_{\mathcal{E}}\underbrace{\mathcal{S}\mathcal{S}^{-1}}_{\mathcal{E}}=\mathcal{E}.$$

Следовательно, $\mathcal{A}=\mathcal{S}\mathcal{B}$ – композиция положительного самосопряженного и ортогонального преобразований.

Замечания 9.9.

1. Из теоремы 9.10 следует, что *для любой действительной симметрической матрицы A существует диагональная матрица $\Lambda=diag(\lambda_1,\ldots,\lambda_n)$ (с собственными числами матрицы A на главной диагонали) и ортогональная матрица S ($S^T=S^{-1}$), что $\Lambda=S^TAS$.*

2. Всякое обратимое самосопряженное преобразование можно представить как композицию растяжений (с коэффициентами, равными собственным числам $\lambda_1,\ldots,\lambda_n$) вдоль взаимно перпендикулярных направлений (задаваемых ортонормированным базисом $s_1,\ldots,s_n$ из собственных векторов). Растяжение с отрицательным коэффициентом $\lambda_1<0$ понимается как композиция зеркального отражения и растяжения с коэффициентом $|\lambda_1|$.

3. Теорема 9.11 справедлива для любого линейного преобразования, если условие положительности самосопряженного преобразования заменить условием его неотрицательности.

4. Геометрический смысл теоремы 9.11 следующий: любое невырожденное линейное преобразование можно представить как композицию преобразований, каждое из которых есть либо простое отражение (относительно гиперплоскости), либо простой поворот (двумерной плоскости), либо растяжение вдоль взаимно перпендикулярных направлений.

59. Приведение квадратичной формы к главным осям.

9.6.4. Приведение квадратичной формы к главным осям

В разд. 6.5.3 была рассмотрена задача приведения вещественной квадратичной формы

$$q(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=x^TAx \tag{9.23}$$

n переменных $x_1,\ldots,x_n$ к каноническому виду (6.18)

$$\tilde q(y)=\lambda_1y_1^2+\lambda_2y_2^2+\ldots+\lambda_ny_n^2 \tag{9.24}$$

при помощи невырожденной линейной замены переменных $x=Sy$. Для решения этой задачи использовался метод Лагранжа (см. разд.6.5.2).

Рассмотрим другой подход к решению. Линейную невырожденную замену переменных $x=Sy$ с ортогональной матрицей S ($S^{-1}=S^T$) будем называть ***ортогональной заменой переменных*** (или ***ортогональным преобразованием переменных***).

Сформулируем задачу ***приведения квадратичной формы к главным осям***: требуется найти ортогональную замену переменных $x=Sy$ ($S^{-1}=S^T$), приводящую квадратичную форму (9.23) к каноническому виду (9.24).

Для решения используем следующий геометрический смысл задачи. Будем считать переменные $x_1,\ldots,x_n$ координатами вектора x n-мерного евклидова пространства E в ортонормированном базисе $(e)=(e_1,\ldots,e_n)$, а матрицу A квадратичной формы (9.23) – матрицей некоторого линейного преобразования $\mathcal{A}:E\to E$ в том же базисе. Причем это преобразование самосопряженное, так как его матрица симметрическая: $A^T=A$. Квадратичную форму (9.23) можно представить в виде скалярного произведения

$$q(x)=(\mathcal{A}(x),x)=(x,\mathcal{A}(x)).$$

Ортогональной замене переменных $x=Sy$ соответствует переход от одного ортонормированного базиса к другому. Действительно, пусть S – матрица перехода от ортонормированного базиса (e) к ортонормированному базису $(s)=(s_1,\ldots,s_n)$, т.е. $(s)=(e)S$ и $S^{-1}=S^T$. Тогда координаты x вектора x в базисе (e) и координаты y того же вектора в базисе (s) связаны формулой (8.11): $x=Sy$.

Таким образом, задача приведения квадратичной формы к главным осям может быть сформулирована так: требуется найти в пространстве E такой базис, в котором матрица самосопряженного преобразования $\mathcal{A}$ имеет диагональный вид. Эта задача была решена в разд. 9.6.3. По теореме 9.10 нужно выбрать ортонормированный базис из собственных векторов самосопряженного преобразования. При этом матрица перехода S к каноническому базису оказывается ортогональной: $S^T=S^{-1}$.

Сформулируем этот результат для квадратичной формы.

Теорема 9.12 (о приведении квадратичной формы к главным осям). *Вещественная квадратичная форма (9.23) при помощи ортогонального преобразования переменных $x=Sy$ может быть приведена к каноническому виду (9.24), где $\lambda_1,\ldots,\lambda_n$ – собственные значения матрицы A.*

Следствие. *Квадратичная форма (9.23) является положительно определенной (неотрицательно определенной) тогда и только тогда, когда все собственные значения ее матрицы положительны (неотрицательны).*

Замечания 9.10.

1. При линейной невырожденной замене переменных матрица квадратичной формы изменяется по формуле (6.10): $A'=S^TAS$. Для ортогональной матрицы S эта формула принимает вид $A'=S^{-1}AS$, который совпадает с формулой (9.4) изменения матрицы линейного преобразования при замене базиса.

2. Для нахождения канонического вида (9.24) достаточно определить все корни $\lambda_1,\ldots,\lambda_n$ (среди которых могут быть равные) характеристического уравнения $\det(A-\lambda E)=0$.

3. Следствие теоремы 9.12 можно использовать для анализа знакоопределенности квадратичной формы:

если все собственные значения положительные (отрицательные), то квадратичная форма положительно (отрицательно) определенная;

если все собственные значения неотрицательные (неположительные), то квадратичная форма неотрицательно (неположительно) определенная;

если имеются собственные значения разных знаков, то квадратичная форма неопределенная (знакопеременная).

4. Результаты, сформулированные в п.3 замечаний, могут быть использованы для проверки достаточных и необходимых условий второго порядка в задаче поиска безусловного экстремума функций (см. разд. 6.5.5). Для этого следует найти собственные значения $\lambda_1,\ldots,\lambda_n$ матрицы Гессе $\frac{d^2f(x)}{dx^Tdx}$ в каждой из стационарных точек x^* функции $f(x)=f(x_1,\ldots,x_n)$.

Если все собственные значения положительные: $\lambda_i>0$, $i=1,\ldots,n$, то в точке x^* локальный минимум;

если все собственные значения отрицательные: $\lambda_i<0$, $i=1,\ldots,n$, то в точке x^* локальный максимум;

если все собственные значения неотрицательные: $\lambda_i\ge0$, $i=1,\ldots,n$, то в точке x^* может быть локальный минимум;

если все собственные значения неположительные: $\lambda_i\le0$, $i=1,\ldots,n$, то в точке x^* может быть локальный максимум;

если собственные значения λ_i, $i=1,\ldots,n$, разных знаков, то в точке x^* нет экстремума;

если все собственные значения нулевые: $\lambda_i=0$, $i=1,\ldots,n$, то требуется дополнительное исследование.

5. Задача приведения квадратичной формы к главным осям решается при помощи алгоритма, рассмотренного в разд. 9.6.3. При этом находится диагональный вид матрицы квадратичной формы и ортогональная матрица S замены переменных $x=Sy$, приводящей квадратичную форму к каноническому виду (к главным осям).

60. Приведение пары квадратичных форм к каноническому виду (конспект лекций).