Лекции по аналитической геометрии и высшей алгебре (прикладная информатика) 2 семестр.

1 Геометрия на плоскости и в пространстве.

Целью данного раздела состоит в рассмотрении таких геометрических понятий как расстояние, площадь, объём с последующим обобщением этих понятий и их переносом на произвольные линейные пространства.

1.1 Скалярное произведение.

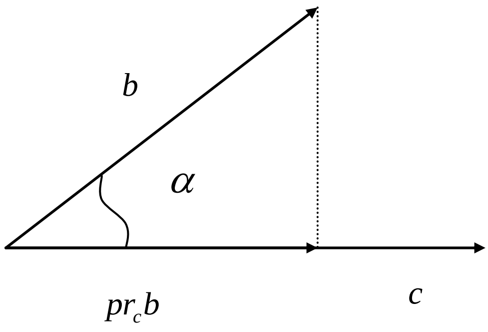
Определение 1.1. Скалярным произведением геометрических векторов a и b называется число, равное произведению длин этих векторов на косинус угла между ними. Скалярное произведение векторов обозначают (a,b).

Из определения следует, что длина вектора равна $|a| = \sqrt{(a,a)}$.

Приведём свойства скалярного произведения.

1. $(a,b) = (b,a)$. **Симметричность**
2. $(\alpha a,b) = \alpha(a,b)$ **Линейность**
3. $(a+b,c) = (a,c)+(b,c)$

В доказательстве нуждается только третье равенство. Если c=0, то равенство очевидно. Пусть $c \neq 0$. Проекция вектора b на c равна $pr_c b = (|\,b\,| \cos\alpha / |\,c\,|)\, c = \left((b,c)/|c|^2\right)c$.



Из равенства $pr_c(a+b) = pr_c a + pr_c b$ и приведённой выше формулы выводим $(a+b,c)/|c|^2\, c = (a,c)/|c|^2\, c + (b,c)/|c|^2\, c$. Приравняем коэффициенты при векторе c в левой и правой частях равенства $(a+b,c)/|c|^2 = (a,c)/|c|^2 + (b,c)/|c|^2$ и умножим на квадрат длины вектора c, получим свойство 3.

Задание длин векторов определяет скалярное произведение. Действительно, из свойств скалярного произведения выводим равенство $|a+b|^2 = (a+b,a+b) = (a,a)+2(a,b)+(b,b) = |a|^2 + 2(a,b)+|b|^2$, которое перепишем в виде $\left(|a+b|^2 - |a|^2 - |b|^2\right)/2 = (a,b)$. Таким образом, задание длин векторов равносильно заданию скалярного произведения и наоборот.

Выразим скалярное произведение через координаты перемножаемых векторов. Пусть f_1, f_2, f_3 - базис пространства векторов, и $a = \alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3$, $b = \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_3$ - разложения векторов a,b по этому базису. Тогда по свойствам скалярного произведения выводим

$$(a,b) = \alpha_1(\beta_1(f_1,f_1)+\beta_2(f_1,f_2)+\beta_3(f_1,f_3)) + +\alpha_2(\beta_1(f_2,f_1)+\beta_2(f_2,f_2)+\beta_3(f_2,f_3)) + \alpha_3(\beta_1(f_3,f_1)+\beta_2(f_3,f_2)+\beta_3(f_3,f_3))$$

. Обозначим через

$$\Gamma(f_1,f_2,f_3) = \begin{pmatrix} (f_1,f_1) & (f_1,f_2) & (f_1,f_3) \\ (f_2,f_1) & (f_2,f_2) & (f_2,f_3) \\ (f_3,f_1) & (f_3,f_2) & (f_3,f_3) \end{pmatrix}$$

матрицу Грамма от векторов f_1, f_2, f_3, составленную из скалярных произведений этих векторов, через $[a]_f = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ - координаты вектора a в базисе f. В этих обозначениях скалярное произведение можно записать с помощью матричных операций следующим образом $(a,b) = [a]_f^{\mathrm{T}} \Gamma(f_1, f_2, f_3)[b]_f$.

Векторы называются ортогональными (перпендикулярными) если угол между ними равен $\pi/2$. Условие ортогональности векторов равносильно равенству нулю их скалярного произведения.

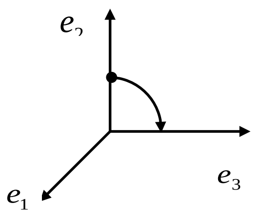
Базис e_1, e_2, e_3 называется ортогональным, если базисные векторы попарно ортогональны. Матрица Грамма ортогональной системы векторов – диагональная. Выражение скалярного произведения через координаты векторов в ортогональном базисе принимает более простой вид, а именно, $(a,b) = \alpha_1\beta_1|e_1|^2 + \alpha_2\beta_2|e_2|^2 + \alpha_3\beta_3|e_3|^2$.

В ортогональном базисе скалярное произведение вектора a на базисный вектор равно $(a,e_i) = (\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3, e_i) = \alpha_i |e_i|^2$, то есть, координаты вектора a находятся по формулам $\alpha_i = (a,e_i)\big/|e_i|^2$.

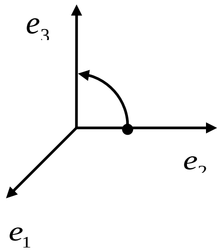
Ортогональный базис e_1, e_2, e_3, в котором длина каждого базисного вектора равна 1, называется ортонормированным. В ортонормированном базисе координаты вектора x определяются по формулам $x_i = (a, e_i)$, а скалярное произведение векторов равно $(a,b) = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3$.

1.2 Векторное и смешанное произведение.

Множество всех ортонормированных троек векторов можно разбить на два класса. Будем говорить, что тройка имеет левую ориентацию, если со стороны первого вектора тройки движение (по кратчайшему пути) от второго к третьему по часовой стрелке, в противном случае тройка имеет правую ориентацию.



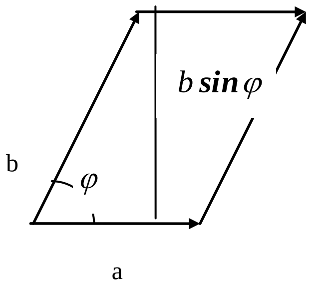
Левый базис



Правый базис

Векторным произведением $[a,b]$ векторов a и b называется вектор, удовлетворяющий следующим трём условиям:

1. Длина вектора $[a,b]$ равна площади параллелограмма натянутого на векторы a,b.
2. Вектор $[a,b]$ ортогонален векторам a и b.
3. Тройка векторов a,b, $[a,b]$ – имеет правую ориентацию.



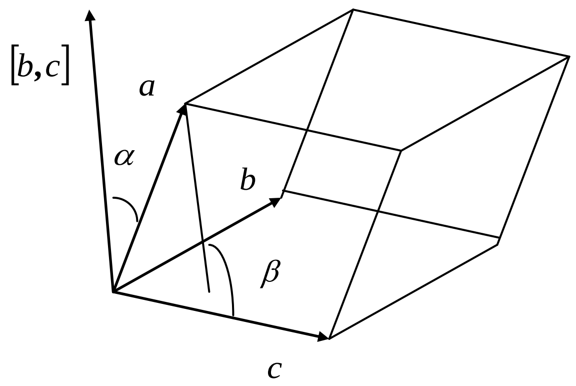
Из определения вытекает, что $|[a,b]| = |a||b|\sin\varphi$. Если векторы a,b коллинеарные, то векторное произведение равно 0. Приведём свойства векторного произведения.

Свойство 1.1 Векторное произведение антикоммутативно, то есть $[a,b] = -[b,a]$.

Действительно, модуль векторного произведения не зависит от порядка сомножителей. Далее, вектор $[a,b]$ коллинеарен вектору $[b,a]$. Однако, переставляя множителей, мы должны изменить направление произведения, чтобы было выполнено условие 3.

Смешанным произведением векторов a,b,c называется число $(a,[b,c])$ и обозначается (a,b,c).

Свойство 1.2 Смешанное произведение векторов (a,b,c) по модулю равно объёму параллелепипеда натянутого на тройку векторов a,b,c. Знак смешанного произведения определяется ориентацией тройки векторов a,b,c, плюс – если тройка правая и минус – если левая.



Доказательство. По определению смешанного произведения $(a,b,c)=|a|\,|[b,c]|\cos\alpha$, где α - угол между вектором a и векторным произведением $[b,c]$, а β - угол между векторами b и c. Произведение $|a|\cos\alpha$ равно высоте параллелепипеда, а $|[b,c]|$ - площади основания параллелепипеда. Произведение этих величин равно объёму параллелепипеда. Знак произведения определяется знаком $\cos\alpha$. Если угол острый, то тройка векторов правая и смешанное произведение положительно. Если угол тупой, то тройка левая и знак смешанного произведения отрицательный.

Свойство 1.3 $(a,b,c)=(c,a,b)=(b,c,a)=-(b,a,c)=-(a,c,b)=-(c,b,a)$.

Для доказательства достаточно заметить, что по модулю все приведённые величины равны и совпадают с объёмом параллелепипеда, натянутого на векторы a,b,c , а знак определяется в зависимости от ориентации тройки векторов.

Свойство 1.4. $[\alpha a+\beta b,c]=\alpha[a,c]+\beta[b,c]$

Доказательство. Рассмотрим смешанное произведение $(x,\alpha a+\beta b,c)$. Выпишем цепочку равенств, используя свойства смешанного и скалярного произведения:

$(x,\alpha a+\beta b,c)=(\alpha a+\beta b,c,x)=(\alpha a+\beta b,[c,x])=\alpha(a,c,x)+\beta(b,c,x)=$
$=\alpha(x,[a,c])+\beta(x,[b,c])=\alpha([a,c],x)+\beta([b,c],x)=(\alpha[a,c]+\beta[b,c],x)=$. Вычтем из левой части равенства правую $0=(x,\alpha a+\beta b,c)-(\alpha[a,c]+\beta[b,c],x)=([\alpha a+\beta b,c],x)-(\alpha[a,c]+\beta[b,c],x)$ и получим равенство $0=([\alpha a+\beta b,c]-\alpha[a,c]-\beta[b,c],x)$ справедливое при любом выборе x. Положим $x=[\alpha a+\beta b,c]-\alpha[a,c]-\beta[b,c]$, тогда $|[\alpha a+\beta b,c]-\alpha[a,c]-\beta[b,c]|^2=0$ и, значит, $[\alpha a+\beta b,c]=\alpha[a,c]+\beta[b,c]$.

Свойство 1.5 $[c,\alpha a+\beta b]=\alpha[c,a]+\beta[c,b]$

Доказательство. $[c,\alpha a+\beta b]=-[\alpha a+\beta b,c]=-\alpha[a,c]-\beta[b,c]=\alpha[c,a]+\beta[c,b]$.

Выразим координаты векторного произведения через координаты исходных векторов в правом ортонормированном базисе. Пусть $a=\alpha_1e_1+\alpha_2e_2+\alpha_3e_3$ и $b=\beta_1e_1+\beta_2e_2+\beta_3e_3$. Используя свойства векторного произведения, найдём $[b,e_1]=\beta_3e_2-\beta_2e_3$, $[b,e_2]=\beta_1e_3-\beta_3e_1$ и $[b,e_3]=\beta_2e_1-\beta_1e_2$. Поскольку базис ортонормированный, то первая координата $[a,b]$ равна $(e_1,[a,b])=(a,[b,e_1])=\alpha_2\beta_3-\alpha_3\beta_2$, вторая координата $(e_2,[a,b])=(a,[b,e_2])=\alpha_3\beta_1-\alpha_1\beta_3$ и третья координата $(e_3,[a,b])=(a,[b,e_3])=\alpha_1\beta_2-\alpha_2\beta_1$.

Таким образом, векторное произведение может быть получено в результате раскрытия по третьему столбцу символического определителя $\begin{vmatrix}\alpha_1 & \beta_1 & e_1\\ \alpha_2 & \beta_2 & e_2\\ \alpha_3 & \beta_3 & e_3\end{vmatrix}$.

Выразим смешанное произведение через координаты исходных векторов в ортонормированном базисе. Разложим векторы a,b,c по базису $a=\alpha_1e_1+\alpha_2e_2+\alpha_3e_3$, $b=\beta_1e_1+\beta_2e_2+\beta_3e_3$, $c=\gamma_1e_1+\gamma_2e_2+\gamma_3e_3$. Раскроем смешенное произведение

$(c,[a,b]) = \gamma_1\alpha_2\beta_3 + \gamma_3\alpha_1\beta_2 + \gamma_2\alpha_3\beta_1 - \gamma_2\alpha_1\beta_3 - \gamma_1\alpha_3\beta_2 - \gamma_3\alpha_2\beta_1$. Выражение в правой части есть определитель матрицы $\begin{vmatrix} \gamma_1 & \alpha_1 & \beta_1 \\ \gamma_2 & \alpha_2 & \beta_2 \\ \gamma_3 & \alpha_3 & \beta_3 \end{vmatrix}$.

Таким образом, определитель матрицы, составленной из координат векторов по абсолютной величине равен объёму параллелепипеда натянутого на эти вектора, а его знак показывает ориентацию этой тройки векторов. Знак положителен, если ориентация совпадает с ориентацией базисных векторов и отрицателен, если ориентации не совпадают.

Матрица Грама от трёх векторов, заданных в ортонормированном базисе равна произведению матриц $\Gamma(a,b,c) = ([a]_e,[b]_e,[c]_e)^{\mathrm{T}}([a]_e,[b]_e,[c]_e)$, следовательно, определитель матрицы Грама равен квадрату объёма параллелепипеда натянутого на эти векторы.

1.3 Уравнение прямой и плоскости в пространстве

Плоскость – линейное многообразие размерности 2. Плоскость в пространстве задаётся одним уравнением $\alpha x + \beta y + \gamma z = \delta$. Подпространство, соответствующее плоскости, задаётся однородным уравнением $\alpha x + \beta y + \gamma z = 0$. В ортонормированном базисе левая часть уравнения является скалярным произведением вектора (α,β,γ) и вектора плоскости (x,y,z). Таким образом, множество векторов плоскости состоит только из тех векторов, которые ортогональны вектору нормали (α,β,γ). Расстояние от точки (x,y,z) до плоскости $\alpha x + \beta y + \gamma z = \delta$ равно $\dfrac{|\alpha x + \beta y + \gamma z - \delta|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}$. Следовательно, коэффициент δ определяет удалённость плоскости от начала координат

Прямая в пространстве задаётся системой из двух уравнений (см. раздел **Error! Reference source not found.Ошибка! Источник ссылки не найден.**) $\begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 z = \delta_1 \\ \alpha_2 x + \beta_2 y + \gamma_2 z = \delta_2 \end{cases}$, причём ранг матрицы, образованной коэффициентами при неизвестных, равен 2. Разберём геометрический смысл коэффициентов. Представив прямую как пересечение двух плоскостей, приходим к выводу, что векторы $(\alpha_1,\beta_1,\gamma_1)$ и $(\alpha_2,\beta_2,\gamma_2)$ образуют базис плоскости перпендикулярной исходной прямой.

2 Евклидово пространство. Скалярное произведение.

Пусть V линейное пространство над полем вещественных чисел. Функция $\varphi(\mathrm{x},\mathrm{y})$, ставящая каждой паре векторов в соответствие число, называется скалярным произведением если выполнены аксиомы

1. Линейность по первому аргументу $\varphi(\alpha x + \beta z, y) = \alpha \cdot \varphi(x,y) + \beta \cdot \varphi(z,y)$.
2. Симметричность: $\varphi(x,y) = \varphi(y,x)$
3. Положительная определенность $\varphi(x,x) > 0$ при $x \neq 0$.

Пространство над полем вещественных чисел в котором введено скалярное произведение называется евклидовым.

Величина $|\mathrm{x}| = \sqrt{\varphi(x,x)}$ называется длиной вектора.

Пусть $e_1,\ldots,e_n$ базис V. Выразим скалярное произведение векторов через координаты векторов. Координаты вектора x в базисе e обозначим через $[\mathrm{x}]_e = \langle \mathrm{x}_1,\ldots,x_n \rangle$. Тогда $\varphi(x,y) = \varphi\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right)$. Пользуясь свойством линейности выводим $\varphi(x,y) = \sum_{i=1}^n x_i \varphi\left(e_i, \sum_{j=1}^n y_j e_j\right)$. Используя симметричность скалярного произведения и линейности по первому аргументу выводим $\varphi(x,y) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j \varphi(e_i,e_j)$. Обозначим через G матрицу Грама базисных векторов, то есть матрицу на пересечении строки i столбца j стоит скалярное произведение i-го и j-го вектора $\varphi(e_i,e_j)$. Используя матричные операции умножения получаем $\varphi(x,y) = [x]_e^{\mathrm{T}} G_e [y]_e$.

2.1 Изменение матрицы Грама при изменении базиса.

Допустим, в евклидовом пространстве V заданы два базиса $e_1,\ldots,e_n$ и $f_1,\ldots,f_n$. Обозначим через P матрицу перехода, связывающие координаты вектора в разных базисах. Пусть для определённости $[x]_e = P[x]_f$. Скалярное произведение не зависит от выбора базиса, поэтому $(x,y)=[x]_f^{\mathrm{T}} G_f [y]_f = [x]_e^{\mathrm{T}} G_e [y]_e$. Подставим в правую часть равенства вместо координат вектора в базисе e их выражение через координаты в базисе f. В результате придём к равенству $[x]_f^{\mathrm{T}} G_f [y]_f = [x]_f^{\mathrm{T}} P^{\mathrm{T}} G_e P[y]_e$. Поскольку полученное равенство справедливо для любых векторов x и y, то выводим $G_{\mathrm{f}} = \mathrm{P}^{\mathrm{T}} G_e P$.

2.2 Ортогональность.

Определение 2.1. Векторы называются ортогональными, если их скалярное произведение равно 0.

Теорема 2.1 (Пифагора). Пусть векторы x и y ортогональны, тогда $|\mathrm{x}+\mathrm{y}|^2 = |x|^2 + |y|^2$.

Доказательство. $|\mathrm{x}+\mathrm{y}|^2 = (x+y,x+y) = (x,x) + 2(x,y) + (y,y) = |x|^2 + |y|^2$, т.к. $(x,y)=0$ в силу ортогональности.

Теорема 2.2 (неравенство Бесселя). Пусть векторы x и y ортогональны, тогда $|\mathrm{x}+\mathrm{y}| \geq |x|$.

Доказательство. По теореме Пифагора $|\mathrm{x}+\mathrm{y}|^2 = |x|^2 + |y|^2$. Поскольку $|y|^2 \geq 0$, то $|\mathrm{x}+\mathrm{y}|^2 \geq |x|^2$, что и требовалось.

Теорема 2.3 (неравенство Коши-Буняковского-Шварца). $|(x,y)| \leq |x||y|$.

Доказательство. Для любого a справедливо неравенство $(x+ay,x+ay) \geq 0$. Раскроем левую часть $|x|^2 + 2a(x,y) + a^2|y|^2 \geq 0$. В левой части неравенства записан квадратный трехчлен. Выделим из него полный квадрат $|y|^2 (a + (x,y)/|y|^2) + |x|^2 - (x,y)^2/|y|^2 \geq 0$. Положив $a = -(x,y)/|y|^2$ получим неравенство $|x|^2 - (x,y)^2/|y|^2 \geq 0$ из которого вытекает $(x,y)^2 \leq |x|^2 |y|^2$. Извлекая квадратный корень, получаем требуемое.

Неравенство Коши-Буняковского-Шварца позволяет ввести угол между векторами, то есть косинус угла равен отношению $(x,y)/|y||\mathrm{x}|$.

Определение 2.2 Система векторов называется ортогональной, если каждая пара векторов из этой системе ортогональна.

Свойство 2.1. Ортогональная система векторов линейно не зависима.

Доказательство. Пусть $f_1,\ldots,f_k$ - ортогональная система векторов и $\sum_{i=1}^k \alpha_i f_i = 0$. Тогда $0 = (0, f_j) = \left(\sum_{i=1}^k \alpha_i f_i, f_j\right) = \alpha_j |f_j|^2$. Таким образом $\alpha_j = 0$ и система векторов линейно независима.

Свойство 2.2. Матрица Грама ортогональной системы векторов – диагональная.

2.3 Процесс ортогонализации.

Пусть $f_1,\ldots,f_k$ линейно не зависимая система векторов. Следующий процесс позволяет строить эквивалентную ей ортогональную систему векторов:

Положим $g_1 = f_1$, $g_2 = f_2 - \frac{(f_2, g_1)}{|g_1|^2} g_1$, …, $g_i = f_i - \sum_{j=1}^{i-1} \frac{(f_i, g_j)}{|g_j|^2} g_j$ … . Процесс не может быть продолжен только в случае, когда $g_i = 0$. Но тогда $0 = f_i - \sum_{j=1}^{i-1} \frac{(f_i, g_j)}{|g_j|^2} g_j$, и, значит, $f_i \in L(f_1, \ldots, f_{i-1})$, что противоречит линейной независимости исходной системы векторов. Ортогональность построенной системы проверяется непосредственно. Допустим, ортогональность системы векторов $g_1, \ldots, g_{i-1}$ установлена. Покажем, что вектор g_i ортогонален всем векторам, построенным ранее него. Действительно, $(g_i, g_k) = \left(f_i - \sum_{j=1}^{i-1} \frac{(f_i, g_j)}{|g_j|^2} g_j, g_k \right) = (f_i, g_k) - \sum_{j=1}^{i-1} \frac{(f_i, g_j)}{|g_j|^2} (g_j, g_k)$, где $k=1,2,\ldots i-1$. В силу ортогональности системы векторов $g_1, \ldots, g_{i-1}$ в сумме из правой части равенства только одно не нулевое слагаемое, получаемое при $j=k$. Следовательно,

$$(g_i, g_k) = (f_i, g_k) - \frac{(f_i, g_k)}{|g_k|^2} (g_k, g_k) = (f_i, g_k) - (f_i, g_k) = 0.$$

Следствие 2.1 В любом подпространстве конечномерного евклидова пространства имеется ортогональный базис.

Доказательство. Возьмем базис подпространства и применим к нему процесс ортогонализации. В результате будет построена ортогональная система векторов (а, значит, и линейно независимая) из этого подпространства. Поскольку количество векторов в построенной системе совпадает с размерностью подпространства, то, следовательно, построенная ортогональная система векторов является базисом подпространства.

Следствие 2.2. Любую ортогональную систему векторов можно дополнить до ортогонального базиса всего пространства.

Доказательство. Пусть $f_1, \ldots, f_k$ - ортогональная система векторов. Дополним ее до базиса всего пространства векторами $f_{k+1}, \ldots, f_n$ и к полученной системе $f_1, \ldots, f_n$ применим процесс ортогонализации. В результате будет построен ортогональный базис $g_1, \ldots, g_n$ всего пространства. Поскольку первые k векторов были ортогональны, то в процессе ортогонализации они не изменились, т.е. $g_1 = f_1, \ldots, g_k = f_k$. Таким образом, векторы $g_{k+1}, \ldots, g_n$ дополняют ортогональную систему $f_1, \ldots, f_k$ до ортогонального базиса всего пространства.

Следствие 2.3. Пусть $f_1, \ldots, f_n$ - базис пространства, а $g_1, \ldots, g_n$ - ортогональный базис пространства, полученный из базиса $f_1, \ldots, f_n$ процессом ортогонализации. Тогда матрица перехода от одного базиса к другому является треугольной, и на ее главной диагонали стоят 1.

Доказательство. Согласно процессу ортогонализации имеем $g_1 = f_1$, $\frac{(f_2, g_1)}{|g_1|^2} g_1 + g_2 = f_2$, …, $\sum_{j=1}^{i-1} \frac{(f_i, g_j)}{|g_j|^2} g_j + g_i = f_i$ …, а, значит, матрица перехода P (ее столбцы – координаты базисных векторов) равна $\begin{pmatrix} 1 & (f_2, g_1)/|g_1|^2 & \cdots & (f_n, g_1)/|g_1|^2 \\ 0 & 1 & \cdots & (f_n, g_2)/|g_2|^2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$.

2.4 Ортогональное дополнение. Ортогональная проекция и составляющая.

Пусть V – евклидово пространство со скалярным произведением (x,y), W – подпространство V.

Множество всех векторов x, ортогональных всем векторам из W, которое обозначим $W^{\perp}$, называется ортогональным дополнением к подпространству W. Опишем свойства ортогонального дополнения.

Свойство 2.3. $W^{\perp}$ **- линейное подпространство** V.

Доказательство. Пусть $x, y \in W^{\perp}$, тогда $\forall z \in W$ справедливы равенства $(x,z)=0$ и $(y,z)=0$. Из этих равенств выводим равенства $(x+y,z)=0$ и $(\alpha x,z)=0$, то есть $x+y, \alpha x \in W^{\perp}$. Тем самым свойство доказано.

Свойство 2.4 $V = W \oplus W^{\perp}$**.**

Доказательство. Построим ортогональный базис $e_1,\dots,e_k$ подпространства W и дополним его до ортогонального базиса $e_1,\dots,e_n$ всего пространства V. Векторы $e_{k+1},\dots,e_n$ ортогональны векторам $e_1,\dots,e_k$, а, значит и любому вектору из W. Следовательно, векторы $e_{k+1},\dots,e_n$ принадлежат ортогональному дополнению к W. Разложим произвольный вектор x по базису $x = x_1e_1+\dots+x_ne_n$ и положим $y = x_1e_1+\dots+x_ke_k$, $z = x_{k+1}e_{k+1}+\dots+x_ne_n$. Поскольку $x=y+z$ и $y \in W$, $z \in W^{\perp}$, то установлено равенство $V = W + W^{\perp}$.

Покажем, что сумма прямая. Пусть $x \in W \cap W^{\perp}$, тогда $(x,x)=0$ как скалярное произведение вектора из W на вектор из ортогонального дополнения к W. Единственный вектор нулевой длины равен 0, и, значит, пересечение содержит только нулевой вектор и сумма прямая.

Следствие 2.4 $\boldsymbol{dim} W^{\perp} = \boldsymbol{dim} V - \boldsymbol{dim} W$ **.**

Доказательство вытекает из свойства прямой суммы подпространств.

Любой вектор x пространства V можно представить в виде суммы вектора y из подпространства W и вектора z из $W^{\perp}$, причем векторы y и z определяются единственным образом. Вектор y называется ортогональной проекцией x на W и обозначается $pr_W x$, а вектор z – ортогональной составляющей вектора x и обозначается $ort_W x$. О способах построения ортогональной проекции и ортогональной составляющей будет разговор в п.2.6.

Свойство 2.5. $W = \left(W^{\perp}\right)^{\perp}$**.**

Доказательство. Применив **Следствие 2.4**, получим $\boldsymbol{dim} W = \boldsymbol{dim}\left(W^{\perp}\right)^{\perp}$. Пусть x – произвольный вектор из W. Поскольку для произвольного вектора $y \in W^{\perp}$ скалярное произведение $(x,y)=0$, то $x \in \left(W^{\perp}\right)^{\perp}$. Тем самым показано включение $W \subseteq \left(W^{\perp}\right)^{\perp}$, из которого, в силу совпадения размерностей, выводим равенство $W = \left(W^{\perp}\right)^{\perp}$.

Пусть $e_1,\dots,e_k$ базис W. Вектор z принадлежит ортогональному дополнению к W тогда и только тогда, когда $(e_1,z)=0$, $(e_2,z)=0$, …, $(e_k,z)=0$. Пусть $h_1,\dots,h_n$ базис пространства V. В координатах, эти равенства превращаются в систему линейных уравнений
$$\begin{cases} (e_1,h_1)z_1 + (e_1,h_2)z_2 + \dots + (e_1,h_n)z_n = 0 \\ (e_2,h_1)z_1 + (e_2,h_2)z_2 + \dots + (e_2,h_n)z_n = 0 \\ \qquad\qquad\ddots \\ (e_k,h_1)z_1 + (e_k,h_2)z_2 + \dots + (e_k,h_n)z_n = 0 \end{cases}.$$

Взяв в качестве W ортогональное дополнение к нему, получим следующее утверждение.

Свойство 2.6. Любое подпространство может быть задано системой линейных однородных уравнений.

В случае, если базис $h_1,\ldots,h_n$ ортонормированный, коэффициентами при неизвестных в системе линейных уравнений являются координаты базисных векторов ортогонального дополнения.

2.5 Геометрический смысл определителя матрицы Грама. Неравенство Адамара.

Свойство 2.7. Определитель матрицы Грама от линейно зависимой системы векторов равен 0.

Доказательство. Пусть система векторов $f_1,\ldots,f_k$ - линейно зависима. Тогда, либо система содержит нулевой вектор, и утверждение в этом случае очевидно, либо найдется вектор $f_i=\alpha_1 f_1+\cdots+\alpha_{i-1}f_{i-1}$, линейно выражающийся через предыдущие векторы системы. В матрице Грама $G(f_1,\ldots,f_k)$ вычтем из *i*-ой строки, предыдущие строки с коэффициентами $\alpha_1,\ldots,\alpha_{i-1}$. Определитель матрицы Грама при этом не изменится, а *i*-ая строка станет равной нулю. Определитель матрицы с нулевой строкой равен нулю, а, значит, и определитель матрицы Грама равен нулю.

Рассмотрим геометрический смысл матрицы Грама от линейной не зависимой системы векторов $f_1,\ldots,f_k$. Если k=1, то $G(f_1)=|f_1|^2$ - квадрат длины вектора. Если k>1, то применим к системе векторов $f_1,\ldots,f_k$ процесс ортогонализации и построим ортогональную систему векторов $g_1,\ldots,g_k$. Обозначим через P матрицу перехода от системы $f_1,\ldots,f_k$ к системе $g_1,\ldots,g_k$. Эта матрица имеет треугольный вид, а на ее главной диагонали стоят 1, и ее определитель равен 1. Кроме того, $P^{\mathrm{T}}G(f_1,\ldots,f_k)P=G(g_1,\ldots,g_k)$ и, следовательно, определители матриц Грама равны. Поскольку система векторов $g_1,\ldots,g_k$ - ортогональна, то матрица Грама от этой системы векторов – диагональная, и ее определитель равен произведению квадратов длин векторов этой системы. Таким образом, установлено равенство $|G(f_1,\ldots,f_k)|=|g_1|^2\cdots|g_k|^2$. Рассмотрим случай k=2. Тогда $|g_2|$ равна длине высоты параллелограмма, опущенного на сторону g_1 (см. **рис. 1**). Следовательно, произведение $|g_1||g_2|$ равно площади параллелограмма натянутого на векторы f_1, f_2, а определитель матрицы Грама $G(f_1,f_2)$ равен квадрату площади этого параллелограмма. Если k=3, то вектор g_3 является ортогональной составляющей вектора f_3 к плоскости, натянутой на векторы f_1, f_2. Следовательно, определитель матрицы Грама от трех векторов равен квадрату объема параллелепипеда, натянутого на векторы f_1, f_2, f_3. Поскольку все рассуждения обобщаются на произвольную размерность, то тем самым установлено свойство.

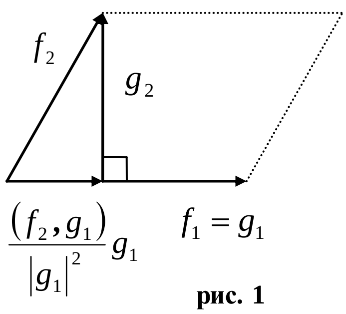


рис. 1

Свойство 2.8 Определитель матрицы Грама от системы векторов равен 0, если система линейно зависима, и квадрату объема *k*-мерного параллелепипеда натянутого на векторы $f_1,\ldots,f_k$ иначе.

Покажем теперь неравенство Адамара.

Теорема 2.4. $\boldsymbol{det}\, G(f_1,\ldots,f_k)\le|f_1|^2\cdots|f_k|^2$

Доказательство. Если система векторов $f_1,\ldots,f_k$ линейно зависимая, то неравенство очевидно. Пусть эта система векторов линейно независимая. Применим к ней процесс ортогонализации и построим ортогональную систему векторов $g_1,\ldots,g_k$. Вектор g_i является ортогональной составляющей вектора f_i на линейную оболочку векторов $f_1,\ldots,f_{i-1}$, и, значит, $|g_i|\le|f_i|$ по неравенству Бесселя (Теорема 2.2). Далее, $|G(f_1,\ldots,f_k)|=|g_1|^2\cdots|g_k|^2\le|f_1|^2\cdots|f_k|^2$, что и требовалось доказать.

Неравенство Адамара обращается в равенство, только если исходная система векторов является ортогональной. В остальных случаях неравенство – строгое.

Следствие 2.5 Справедливы неравенства $|\boldsymbol{det}\, A|\le\prod_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}^2\right)^{1/2}$ **и** $|\boldsymbol{det}\, A|\le\prod_{j=1}^{n}\left(\sum_{i=1}^{n}a_{ij}^2\right)^{1/2}$**.**

Доказательство. В n-мерном арифметическом пространстве определим скалярное произведение по формуле $(x, y) = \sum_{i=1}^{n} x_i y_i$. Рассмотрим систему векторов, образованную столбцами матрицы A. Матрица Грама от этой системы векторов равна $A^{\mathrm{T}} A$ и по неравенству Адамара $\boldsymbol{det}\left(A^{\mathrm{T}} A\right) \le \prod_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij}^2\right)$. Поскольку $\boldsymbol{det}\left(A^{\mathrm{T}} A\right) = (\boldsymbol{det}\, A)^2$, то неравенство $|\boldsymbol{det}\, A| \le \prod_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij}^2\right)^{1/2}$ установлено. Применяя полученное неравенство к транспонированной матрице, выводим $|\boldsymbol{det}\, A| \le \prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$.

Следствие 2.6 Пусть $\alpha = \boldsymbol{max}_{ij} |a_{ij}|$. Тогда $|\boldsymbol{det}\, A| \le \left(\sqrt{n}\alpha\right)^n$.

Доказательство очевидно.

Положим $F_1 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ и, далее, по индукции $F_n = F_1 \times F_{n-1}$. Матрица F_n имеет порядок 2^n, ее определитель равен $2^{n \cdot 2^n}$ и все ее элементы равны ± 1. Легко убедиться, что неравенство (**Следствие 2.6**) обращается на этой матрице в равенство.

2.6 . Расстояния. Псевдорешения. Нормальные решения. Нормальные псевдорешения.

Расстоянием между множествами X и Y называется $\boldsymbol{inf}_{x \in X, y \in Y} |x\text{-}y|$.
Рассмотрим задачу нахождения расстояния от точки x до подпространства W. В начале рассмотрим случай, когда подпространство задано в виде линейной оболочки системы векторов.

Теорема 2.5. Расстояние от точки до подпространства достигается на перпендикуляре, опущенном из точки x на подпространство.

Доказательство. Представим $x = pr_W x + ort_W x$. Расстояние от точки x до подпространства W равно $\boldsymbol{inf}_{y \in W} |x - y| = \boldsymbol{inf}_{y \in W} |ort_W x + (pr_W x - y)|$. Векторы $ort_W x$ и $pr_W x - y$ ортогональны друг другу, и по неравенству Бесселя $|ort_W x + (pr_W x - y)| \ge |ort_W x|$, причем равенство достигается только в случае $y = pr_W x$. Тем самым установлено $\boldsymbol{inf}_{y \in W} |x - y| = |ort_W x|$, что и требовалось.

Пусть $W = L(e_1, \dots, e_k)$ и система векторов $e_1, \dots, e_k$ линейно независимая. Расстояние от точки x до подпространства W можно найти как отношение объема k+1-мерного параллелепипеда натянутого на векторы $e_1, \dots, e_k, x$ к объему k-мерного параллелепипеда натянутого на векторы $e_1, \dots, e_k$. Таким образом, справедлива формула $|ort_W x|^2 = |G(e_1, \dots, e_k, x)| / |G(e_1, \dots, e_k)|$. К сожалению, эта формула не позволяет находить проекцию и ортогональную составляющую вектора. Для нахождения проекции можно поступать следующим образом. Представим $pr_W x = x_1 e_1 + \dots + x_k e_k$ и $x = x_1 e_1 + \dots + x_k e_k + ort_W x$, а затем умножим скалярно на векторы $e_1, \dots, e_k$ вектор x. Получим систему линейных уравнений $\begin{cases} (e_1, e_1) x_1 + \dots + (e_1, e_k) x_k = (e_1, x) \\ \vdots \\ (e_k, e_1) x_1 + \dots + (e_k, e_k) x_k = (e_k, x) \end{cases}$. Коэффициенты при неизвестных образуют матрицу Грама, определитель которой не равен нулю. Следовательно, система имеет единственное решение. Решив эту систему, найдем проекцию вектора x, а затем и ортогональную составляющую.

Рассмотрим случай, когда линейное подпространство задано системой однородных линейных уравнений Ax=0. Для простоты проведения рассуждений будем считать, что строки матрицы A линейно независимы. В ортонормированном базисе, коэффициенты при неизвестных в уравнении являются координатами вектора из ортогонального дополнения (см. п.2.4). Таким образом, по системе линейных уравнений можно найти базис ортогонального дополнения к пространству W. Обозначим базис $W^{\perp}$ через $e_1, \dots, e_m$. Тогда представим $ort_W x = x_1 e_1 + \dots + x_m e_m$ и $x = pr_W x + x_1 e_1 + \dots + x_m e_m$, а затем умножим скалярно на векторы $e_1,\dots,e_m$ вектор x. Получим систему линейных уравнений

$$\begin{cases} (e_1,e_1)x_1+\dots+(e_1,e_m)x_m=(e_1,x) \\ \vdots \\ (e_m,e_1)x_1+\dots+(e_m,e_m)x_m=(e_m,x) \end{cases}$$

. Коэффициенты при неизвестных образуют матрицу Грама, определитель которой не равен нулю. Следовательно, система имеет единственное решение. Решив эту систему, найдем ортогональную составляющую вектора x, а затем и проекцию.

Рассмотрим теперь задачу нахождения расстояния от точки x до линейного многообразия M. Эта задача легко сводится к аналогичной задаче построения расстояния от точки до подпространства. Действительно, пусть $M=z+W$, где z – произвольная точка из M, а W – подпространство. Тогда $\boldsymbol{inf}_{y\in M}|x-y|=\boldsymbol{inf}_{y\in W}|(x-z)-y|$, то есть задача свелась к определению расстояния от точки x-z до подпространства W.

Линейное многообразие, заданное как множество решений одного линейного уравнения $ax=b$ называется гиперплоскостью. Рассмотрим задачу отыскания расстояния от точки y до гиперплоскости $ax=b$. Перпендикуляр, опущенный из y на гиперплоскость равен αa и $a(y-\alpha a)=b$. Отсюда находим неизвестный параметр $\alpha=(ay-b)/|a|^2$, а затем и расстояние $|\alpha a|=|ay-b|/|a|$.

Рассмотрим задачу определения расстояния между двумя линейными многообразиями $M_1=y+W_1$ и $M_2=z+W_2$. Расстояние между ними равно $\boldsymbol{inf}_{a\in M_1,b\in M_2}|a-b|=\boldsymbol{inf}_{a\in W_1,b\in W_2}|(y-z)+a-b|=\boldsymbol{inf}_{a\in W_1+W_2}|(y-z)+a|$, то есть задача свелась к нахождению расстояния от точки y-z до подпространства W_1+W_2. Заметим, что расстояние между линейными многообразиями достигается на общем перпендикуляре.

2.6.1 Псевдорешения. Метод наименьших квадратов.

Рассмотрим несовместную систему линейных уравнений $Ax=b$. Псевдорешением системы линейных уравнений называется вектор x, на котором достигается минимум нормы невязки $|Ax$-$b|$. Задача построения псевдорешения возникает при подборе параметров физических процессов. Левая часть системы уравнений определяется конкретным видом зависимости от параметров, а правая – конкретными измерениями. Поскольку каждое измерение производится с некоторой точностью, то обычно их проводят с избытком. В результате получается несовместная система линейных уравнений, а задача подбора параметров сводится к построению псевдорешения. Сам способ перехода от задачи решения системы линейных уравнений к нахождению минимума длины невязки называется метод наименьших квадратов. Такое название связано с тем, что $|Ax-b|^2=\sum_{i=1}^m\left(\sum_{j=1}^n a_{ij}x_j-b_i\right)^2$.

Обозначим через W линейную оболочку столбцов матрицы A. Задача построения псевдорешения эквивалентна задаче определения расстояния от b до W, а точнее к определению проекции b на W. Коэффициенты разложения проекции по столбцам матрицы A являются решениями системы уравнений $A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$. Тем самым, задача построения псевдорешения свелась к решению системы линейных уравнений.

Если исходная система имела решение, то оно является также псевдорешением. Необходимым и достаточным условием единственности псевдорешения является условие линейной независимости столбцов матрицы A.

2.6.2 Нормальное решение

В ряде случаев, из множества решений, следует выбрать какое то одно. Нормальным решением системы линейных уравнений $Ax=b$ называется решение наименьшей длины.

Задача отыскания нормального решения сводится к задаче определения расстояния от начала координат до линейного многообразия, заданного системой линейных уравнений $Ax=b$.

Перпендикуляр, опущенный из начала координат на это линейное многообразие, представляется в виде $A^{\mathrm{T}}u$ линейной комбинации строк матрицы A. Следовательно, задача построения нормального решения сводится к решению системы линейных уравнений $AA^{\mathrm{T}}u=b$ и вычислению ответа $x=A^{\mathrm{T}}u$.

Нормальное решение всегда единственно, чего нельзя сказать о решении системы $AA^{\mathrm{T}}u=b$.

Необходимым и достаточным условием единственности решения указанной системы является условие линейной независимости строк матрицы A.

2.6.3 Нормальное псевдорешение.

Задача построения нормального псевдорешения сводится к решению системы $A^{\mathrm{T}}AA^{\mathrm{T}}u = A^{\mathrm{T}}b$ и вычисления нормального псевдорешения по формуле $x = A^{\mathrm{T}}u$.

3 Унитарное пространство.

Пусть *V* линейное пространство над полем комплексных чисел. Можно ли обобщить понятие скалярного произведения на такое пространство. Оказывается, да! Для этого достаточно незначительно изменить аксиомы скалярного произведения.

1. $\varphi(\alpha x + \beta z, y) = \alpha \cdot \varphi(x, y) + \beta \cdot \varphi(z, y)$.
2. $\varphi(x, y) = \overline{\varphi(y, x)}$
3. $\varphi(x, x) > 0$ при $x \neq 0$.

Черта в свойстве 2 обозначает знак комплексного сопряжения. Пространство над полем комплексных чисел, в котором введено скалярное произведение называется унитарным.
Обозначим через G матрицу Грама базисных векторов, то есть матрицу на пересечении строки i столбца j стоит скалярное произведение i-го и j-го вектора $\varphi(e_i, e_j)$. Используя матричные операции умножения, получаем $\varphi(x, y) = [x]_e^{\mathrm{T}} G_e \overline{[y]}_e$. Матрицы Грама в разных базисах связаны формулой $G_{\mathrm{f}} = \mathrm{P}^{\mathrm{T}} G_e \overline{P}$, где *P* матрица перехода. Все остальные свойства скалярного произведения полностью сохраняются.

4 Билинейные функции, квадратичные формы.

4.1 Билинейные формы. Квадратичные формы.

Пусть *V* – линейное пространство над полем *P*. Функция, ставящая в соответствие паре векторов вещественное число, и обладающая свойствами линейности называется билинейной формой. Другими словами, функция $f : V \times V \to P$ называется билинейной, если

1. $f(\beta b + \gamma c, a) = \beta f(b, a) + \gamma\ f(c, a)$,
2. $f(a, \beta b + \gamma c) = \beta\ f(a, b) + \gamma\ f(a, c)$,

где $a, b, c \in V$, $\beta, \gamma \in P$.

Примером билинейной функции является скалярное произведение.

Теорема 4.1 Билинейная форма полностью определяется своими значениями на базисных векторах.

Доказательство. Пусть $e_1, \ldots, e_n$ - базис *V*. Разложим векторы *b* и *c* по базису $b = \beta_1 e_1 + \cdots + \beta_n e_n$, $c = \gamma_1 e_1 + \cdots + \gamma_n e_n$. Тогда из линейности выводим $f(b, c) = \sum_{i=1}^n \beta_i \sum_{j=1}^n \gamma_j\ f(e_i, e_j)$. Теорема доказана.

Обозначим через $[b]_e$ столбец, составленный из координат вектора *b*, а через F_e – матрицу, на пересечении *i*-ой строки и *j*-го столбца которой расположено значение билинейной формы от базисных векторов $f(e_i, e_j)$. Легко убедиться в равенстве $f(b, c) = [b]_e^{\mathrm{T}} F_e [c]_e$. Матрица F_e называется матрицей билинейной формы *f* в базисе $e_1, \ldots, e_n$.

Следствие 4.1 Билинейная форма полностью определяется своей матрицей.

Билинейная форма называется *симметричной*, если ее значение не меняется от перестановки аргументов, то есть $f(b, c) = f(c, b)$.

Следствие 4.2 Билинейная форма симметрична тогда и только тогда, когда найдется базис, в котором ее матрица симметрична.

Доказательство. Если билинейная форма симметричная, то в любом базисе ее матрица симметрична. Обратно, пусть в некотором базисе матрица билинейной формы симметричная. Тогда

$$f(x, y) = [x]^{\mathrm{T}} F[y] = \left([x]^{\mathrm{T}} F[y]\right)^{\mathrm{T}} = [y]^{\mathrm{T}} F[x] = f(y, x).$$

Квадратичной формой называется значение билинейной формы от одного аргумента, то есть *f*(*x*,*x*).

Одну и ту же квадратичную форму можно получить из разных билинейных форм. Например, квадратичную форму x_1x_2 можно получить из следующих билинейных форм $\alpha x_1 y_2 + (1-\alpha) y_1 x_2$, где $\alpha \in R$.

Между квадратичными формами и симметричными билинейными формами существует взаимно однозначное соответствие, определяемое формулой 0,25(*f*(*x*+*y*,*x*+*y*)-*f*(*x*-*y*,*x*-*y*)). Матрица симметричной билинейной формы, соответствующей квадратичной форме, называется матрицей квадратичной формы.

4.2 Полуторалинейные формы. Эрмитовы формы.

Пусть *V* – линейное пространство над полем комплексных чисел *C*. Функция, ставящая в соответствие паре векторов комплексное число, и обладающая свойствами линейности по первому аргументами и «почти линейностью» по второму, называется полуторалинейной формой. Точнее, функция $f : V \times V \to C$ называется полуторалинейной, если

1. $f(\beta b + \gamma c, a) = \beta f(b,a) + \gamma f(c,a)$,
2. $f(a, \beta b + \gamma c) = \overline{\beta} f(a,b) + \overline{\gamma} f(a,c)$,

где $a,b,c \in V$, $\beta, \gamma \in C$.

Примером полуторалинейной функции является скалярное произведение в унитарном пространстве.

Теорема 4.2. Полуторалинейная форма полностью определяется своими значениями на базисных векторах.

Доказательство. Пусть $e_1,\dots,e_n$ - базис *V*. Разложим векторы *b* и *c* по базису $b = \beta_1 e_1 + \dots + \beta_n e_n$, $c = \gamma_1 e_1 + \dots + \gamma_n e_n$. Тогда $f(b,c) = \sum_{i=1}^n \beta_i \sum_{j=1}^n \overline{\gamma_j} f(e_i, e_j)$. Теорема доказана.

Обозначим через $[b]_e$ столбец, составленный из координат вектора *b*, а через F_e – матрицу, на пересечении *i*-ой строки и *j*-го столбца которой расположено значение полуторалинейной формы от базисных векторов $f(e_i, e_j)$. Легко убедиться в равенстве $f(b,c) = [b]_e^{\mathrm{T}} F_e \overline{[c]}_e$, где черта обозначает знак комплексного сопряжения. Матрица F_e называется матрицей полуторалинейной формы *f* в базисе $e_1,\dots,e_n$.

Следствие 4.3 Полуторалинейная форма полностью определяется своей матрицей.

Полуторалинейная форма называется эрмитовой, если ее значение меняется от перестановки аргументов на комплексно сопряженное, то есть $f(b,c) = \overline{f(c,b)}$.

Следствие 4.4 Полуторалинейная форма является эрмитовой тогда и только тогда, когда найдется базис *e*, в котором ее матрица удовлетворяет равенству $F_e = \overline{F_e^{\mathrm{T}}}$.

Для эрмитовых форм определен аналог квадратичной формы $f(x,x)$.

Значение квадратичной эрмитовой формы – всегда вещественное число, так как $f(x,x) = \overline{f(x,x)}$.

4.3 Изменение матрицы билинейной (полуторалинейной) формы при изменении базиса.

Пусть *V* – линейное пространство, *f* – билинейная форма, *e* и *g* – базисы *V*. Согласно полученным ранее формулам, имеем равенства $f(x,y) = [x]_e^{\mathrm{T}} F_e [y]_e$, $f(x,y) = [x]_g^{\mathrm{T}} F_g [y]_g$, $[x]_e = P[x]_g$, где *P* – матрица перехода. После элементарных преобразований получим равенство $[x]_g^{\mathrm{T}} F_g [y]_g = [x]_e^{\mathrm{T}} P^{\mathrm{T}} F_e P [y]_e$, из которого выводим формулу $F_g = P^{\mathrm{T}} F_e P$ изменения матрицы билинейной формы при изменении базиса.
Рассматривая матрицу квадратичной формы как матрицу симметричной билинейной формы, получаем, что матрица квадратичной формы изменяется по формуле $F_g = P^{\mathrm{T}} F_e P$.
Аналогичным образом выводим формулу изменения матрицы полуторалинейной формы при изменении базиса $F_g = P^{\mathrm{T}} F_e \overline{P}$.

4.4 *Приведение квадратичных форм (симметричных билинейных форм, эрмитовых форм) к простейшему виду.*

4.4.1 Метод выделения квадратов (Лагранжа).

Базис $e_1, e_2, ..., e_n$ называется каноническим для симметричной (эрмитовой) билинейной функции, если ее матрица в этом базисе диагональная.

Теорема 4.3 Лагранжа. Для любой эрмитовой функции существует канонический базис.

Доказательство проведём индукцией по рангу r матрицы эрмитовой формы F. Если r=0, то матрица нулевая, утверждение очевидно. Допустим, что теорема верна для r-1. Докажем ее истинность для r. Рассмотрим три случая

а) $f_{11} \neq 0$, тогда положим $x'_1 = 1/f_{11} \sum_{j=1}^{n} f_{j1} x_j$ и $x'_k = x_k$, где k>1. В данном случае матрица перехода S будет отличаться от единичной матрицы только первой строкой, равной $\langle 1 \quad f_{21}/f_{11} \quad \cdots \quad f_{n1}/f_{11} \rangle$ и $S[x]=[x']$, $Q[x']=[x]$, где $Q = S^{-1}$. Матрица Q отличается от единичной матрицы только первой строкой, равной $(1, -f_{21}/f_{11}, \ldots, -f_{n1}/f_{11})$. После замены координат, получим матрицу билинейной формы $Q^T F \overline{Q}$, которая имеет следующий блочный вид $\left(\begin{array}{c|c} f_{11} & 0 \\ \hline 0 & F_1 \end{array}\right)$. Поскольку ранг F_1 равен r-1, то по предположению индукции эрмитову матрицу F_1 можно привести к каноническому виду. Пусть $Q'^T F_1 \overline{Q'} = D'$. Тогда $\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q' \end{array}\right)^{T} \left(\begin{array}{c|c} f_{11} & 0 \\ \hline 0 & F_1 \end{array}\right) \overline{\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q' \end{array}\right)} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & D' \end{array}\right)$ и теорема в этом случае доказана.

б) $f_{11} = 0$ и существует k, что $f_{kk} \neq 0$ переставим первый и k базисные вектора, и далее перейдем к пункту а).

в) $f_{kk} = 0$ для всех k и найдётся не нулевой элемент $f_{kl} \neq 0$, где $k \neq l$. Возможны два случая:

1. $\boldsymbol{Re(}\, f_{kl} \boldsymbol{)} \neq 0$ тогда прибавим к k базисному вектору l базисный вектор и получим случай б)
2. $\boldsymbol{Re(}\, f_{kl} \boldsymbol{)} = 0$ тогда прибавим к k базисному вектору l базисный вектор, умноженный на i, и получим случай б). Теорема доказана.

Базис эрмитовой билинейной функции $f(x,y)$ называется нормальным, если матрица билинейной функции в этом базисе имеет диагональный вид, и ее главная диагональ равна (1,..,1,-1,..,-1,0..,0).
Для отыскания матрицы перехода можно поступать следующим образом. Припишем к матрице F единичную матрицу справа. Затем будем производить элементарные преобразования со строками расширенной матрицы и столбцами матрицы F. Причем, если к строке k прибавим строку j, умноженную на число α, то затем к столбцу k прибавим столбец j, умноженный на число $\overline{\alpha}$. После приведения матрицы F к диагональному виду справа будет расположена матрица, все элементы которой комплексно сопряжены к матрице перехода.

Следствие 4.5 Для эрмитовой формы существует нормальный базис если поле R или C.

Доказательство. Построим канонический базис. Далее, если $f_{jj} \neq 0$, то умножим j базисный вектор на число $1/\sqrt{f_{jj}}$. Затем перестановкой базисных векторов приведем матрицу к нормальному виду.

Следствие 4.6 Если все угловые миноры матрицы F отличны от нуля, то существует верхняя треугольная матрица Q, которая приводит F к диагональному виду.

Доказательство проведем индукцией по рангу F. По теореме Лагранжа существует матрица Q, приводящая F к диагональному виду. Докажем, что она верхняя треугольная матрица. Обозначим через Δ_j угловой минор j-го порядка матрицы F. Так как $\Delta_1 = f_{11} \neq 0$, то выполняется пункт а) теоремы Лагранжа. Матрица перехода Q верхняя треугольная. Угловой минор матрицы F_1 порядка k-1, умноженный на f_{11}, равен Δ_k (угловому минору порядка k матрицы F). По предположению индукции, найдется верхняя треугольная матрица Q’, приводящая матрицу F_1 к диагональному виду. Но тогда $T=Q\begin{pmatrix}1 & 0\\ 0 & Q'\end{pmatrix}$ - верхняя треугольная матрица, а $T^{\mathrm{T}}F\overline{T}$ - диагональная матрица.

4.4.2 Приведение квадратичных форм к нормальному виду элементарными преобразованиями

Симметричные матрицы A и F назовем конгруэнтными, если найдется невырожденная матрица P, что $A=P^{\mathrm{T}}FP$. Матрицы билинейной формы в различных базисах конгруэнтны. Из теоремы Лагранжа вытекает, что симметричная квадратная матрица конгруэнтна диагональной матрице $diag$(1,…,1.-1,…,-1,0,…,0). Опишем алгоритм приведения симметричной квадратной матрицы F к диагональному виду элементарными преобразованиями. Отметим, если мы совершаем какие то действия со строками матрицы F, то те же самые действия надо совершить и со столбцами матрицы. Номера столбцов будут указываться в квадратных скобках, а номера строк – в круглых скобках.

1. Положим r=1.
2. Если $f_{rr}=0$, то перейдем на шаг 4, иначе шаг 3.
3. Положим $(i)=(i)-f_{ir}/f_{rr}(r)$, $[i]=[i]-f_{ir}/f_{rr}[r]$, где $i=1,\dots,r-1,r+1,\dots,n$. Затем увеличим r на 1 и вернемся на шаг 2.
4. Если найдется i, что $f_{ir}\neq 0$, то положим $(r)=(r)+(i)$, $[r]=[r]+[i]$ и вернемся на шаг 2. В противном случае перейдем на шаг 5.
5. Если для всех $i,j>r$ справедливо неравенство $f_{ij}=0$, то алгоритм работу закончил. В противном случае найдутся номера i,j, для которых $f_{ij}\neq 0$. Тогда переставим строки $(r)\leftrightarrow(i)$ и столбцы $[r]\leftrightarrow[i]$ и вернемся на шаг 2.

Легко проверить, что предложенный алгоритм построит диагональную матрицу конгруэнтную исходной матрице. Преобразованиями вида $(i)=1\big/\sqrt{f_{ii}}\,(i)$, $[i]=1\big/\sqrt{f_{ii}}\,[i]$ можно добиться, чтобы на главной диагонали стояли только 0,1,-1. Перестановками строк и столбцов элементы матрицы, стоящие по главной диагонали, можно расположить в порядке не возрастания.

Если приписать справа единичную матрицу, то элементарные преобразования можно запомнить в ней.

4.4.3 Закон инерции квадратичных форм.

Теорема 4.4 Закон инерции. Для любой эрмитовой билинейной функции нормальный вид определяется единственным образом.

Доказательство проведем методом от противного. Пусть найдется два базиса $e_1,\dots,e_n$ и $h_1,\dots,h_n$, в которых нормальный вид различный. Пусть $[x]_e=(x_1,\dots,x_n)^{\mathrm{T}}$ и $[x]_h=(y_1,\dots,y_n)^{\mathrm{T}}$ и $f(x,x)=\sum_{i=1}^{s_1}x_i^2-\sum_{i=s+1}^{r_1}x_i^2=\sum_{i=1}^{s_2}y_i^2-\sum_{i=s+1}^{r_1}y_i^2$. Поскольку $r_1=rgF_e=rgF_h=r_2=r$, то нормальный вид различен только если $s_1\neq s_2$. Для определенности положим $s_1>s_2$. Обозначим через W линейную оболочку векторов $e_1,\dots,e_{s_1}$, а через U – линейную оболочку векторов $h_{s_2+1},\dots,h_n$. Для не нулевого вектора $x\in W$ имеем $[x]_e=(x_1,\dots,x_{s_1},0,\dots,0)^{\mathrm{T}}$ и $f(x,x)=\sum_{i=1}^{s_1}x_i^2>0$, а для вектора $x\in U$ выполняются равенства $[x]_h=(0,\dots,0,y_{s_2+1},\dots,y_n)^{\mathrm{T}}$ и $f(x,x)=-\sum_{i=s_2+1}^{r}y_i^2\leq 0$. Пересечение подпространств $W\cap U$ не может содержать векторов отличных от нуля, но $\boldsymbol{dim}W\cap U=\boldsymbol{dim}(W+U)-\boldsymbol{dim}W-\boldsymbol{dim}U\geq n-s_1-(n-s_2)>0$. К полученному противоречию привело допущение $s_1\neq s_2$. Таким образом теорема доказана.

Число положительных слагаемых в нормальном виде эрмитовой формы называется положительным индексом инерции, а число слагаемых с отрицательными коэффициентами – отрицательный индекс инерции. Индексы инерции эрмитовых форм не зависят от выбора базиса.

4.4.4 Теорема Якоби

Обозначим через Δ_j угловой минор *j*-го порядка матрицы *F* (и положим $\Delta_0 = 1$).

Теорема 4.5 Якоби. Пусть $\Delta_k \neq 0$ (*k*=1,2,..,*r*). Существует канонический базис $e_1, e_2, ..., e_n$, для которого $f(e_k, e_k) = \Delta_k / \Delta_{k-1}$, при *k*=1,2,..,*r*, и $f(e_k, e_k) = 0$ при *k*>*r*.

Доказательство очевидным образом повторяет Следствие 4.6.

4.4.5 Критерий Сильвестра.

Эрмитовая форма $f(x, y)$ называется положительно определенной, если для любого $x \neq 0$ справедливо неравенство $f(x, x) > 0$.

Теорема 4.6 Критерий Сильвестра положительной определенности. Эрмитова форма положительна определена тогда и только тогда, когда все ее главные миноры (расположенные по главной диагонали) строго больше нуля.

Доказательство. Если все главные миноры *F* строго больше нуля, то и все угловые миноры матрицы *F* строго больше нуля. По теореме Якоби найдется базис, в котором эрмитова форма имеет вид $\sum_{i=1}^{n} \Delta_i / \Delta_{i-1} |x_i|^2$. Поскольку все коэффициенты строго больше нуля, то эрмитова форма положительно определена.

Покажем обратное. Допустим, найдется главный минор матрицы *F*, не больше нуля. Не нарушая общности можно считать, что это угловой минор порядка *k*, так как в противном случае перенумеруем переменные соответствующим образом. Далее, можно считать, что все угловые миноры до (*k*-1)-го порядка больше нуля. Действительно, иначе можно положить *k* равным меньшему значению. Положим все переменные с номером больше *k* равными нулю. В результате получим эрмитову форму от *k* переменных с матрицей $\begin{pmatrix} f_{11} & \cdots & f_{1k} \\ \vdots & \ddots & \vdots \\ f_{k1} & \cdots & f_{kk} \end{pmatrix}$. Угловые миноры этой матрицы до (*k*-1)-го порядка больше нуля, и, значит можно воспользоваться теоремой Якоби. В некотором базисе эта форма имеет вид $\sum_{j=1}^{k} \Delta_j / \Delta_{j-1} |y_j|^2$. По построению $\Delta_k / \Delta_{k-1} \leq 0$, и, значит, найдется не нулевой вектор, значение эрмитовой формы $f(x, x)$ на котором не больше нуля, что противоречит ее положительной определенности. К полученному противоречию привело допущение о существовании не положительных главных миноров матрицы *F*. Следовательно, все главные миноры больше нуля.

5 Квадрики.

5.1 *Алгебраическая поверхность*

Пусть $f(x_1, \ldots, x_n)$ некоторый полином от *n* переменных. Алгебраической поверхностью называется множество решений уравнения $f(x_1, \ldots, x_n) = 0$, порядком алгебраической поверхности называется максимальная степень полинома. Если *n*=2, то алгебраическая поверхность называется алгебраической кривой.

Под аффинной заменой координат будем понимать замену вида $x = b + Py$, где P – невырожденная матрица.

Теорема 5.1При аффиной заменой координат порядок поверхности не меняется.

Доказательство. Непосредственной подстановкой легко проверить, что при аффинной замене координат порядок поверхности не возрастает. Возврат к исходной системе координат является аффинным преобразованием. Значит, порядок поверхности не может и убывать.

5.2 Уравнение квадрики.

Алгебраическая поверхность второго порядка называется квадрикой. Общее уравнение квадрики можно записать в следующем виде $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha = 0$. Положим $\hat{A} = \begin{pmatrix} A & a \\ a^{\mathrm{T}} & \alpha \end{pmatrix}$ и $\hat{x} = \begin{pmatrix} x \\ 1 \end{pmatrix}$. Уравнение квадрики в новых обозначениях принимает вид $\hat{x}^{\mathrm{T}}\hat{A}\hat{x} = 0$. Матрица $\hat{A}$ называется расширенной матрицей квадрики.

5.3 Изменение квадрики при аффинном преобразовании

Ответим на вопрос об изменении уравнения квадрики при аффинной замене координат $x=h+Ty$. Положим $\hat{T} = \left(\begin{array}{c|c} T & h \\ \hline 0 & 1 \end{array}\right)$, тогда равенство $x=h+Ty$ эквивалентно равенству $\hat{x} = \hat{T}\hat{y}$, и, значит, $\hat{x}^{\mathrm{T}}\hat{A}\hat{x} = \hat{y}^{\mathrm{T}}\hat{T}^{\mathrm{T}}\hat{A}\hat{T}y$. Тем самым установлена теорема.

Теорема 5.2. Пусть из квадрики $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha = 0$ аффинной заменой координат $x=h+Ty$ получается квадрика $y^{\mathrm{T}}By + 2b^{\mathrm{T}}y + \beta = 0$, тогда $\hat{B} = \hat{T}^{\mathrm{T}}\hat{A}\hat{T}$, $B = T^{\mathrm{T}}AT$, $b = T^{\mathrm{T}}a + T^{\mathrm{T}}Ah$, $\beta = h^{\mathrm{T}}Ah + 2a^{\mathrm{T}}h + \alpha$.

Обозначим через $s(A)$ положительный индекс инерции, а через $t(A)$ – отрицательный индекс инерции квадратичной формы. Из приведенных формул вытекает полезное следствие.

Следствие 5.1. Пусть из квадрики $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha = 0$ аффинной заменой координат получена квадрика $y^{\mathrm{T}}By + 2b^{\mathrm{T}}y + \beta = 0$. Тогда $s(A) = s(B)$, $t(A) = t(B)$, $rg(A) = rg(B)$ и $s(\hat{A}) = s(\hat{B})$, $t(\hat{A}) = t(\hat{B})$, $rg(\hat{A}) = rg(\hat{B})$.

Доказательство вытекает из закона инерции квадратичных форм и формул изменения квадрики при аффинной замене системы координат.

Следствие 5.2. Величины $s(A)$, $rg(A)$, $s(\hat{A})$, $rg(\hat{A})$ являются аффинными инвариантами квадрики.

5.4 Приведение уравнения квадрики к простейшему виду

Теорема 5.3 Для любого уравнения квадрики $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha = 0$ найдется аффинная замена системы координат, приводящую квадрику к виду $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = \begin{cases} \alpha \\ 2y_{1+rgA} \end{cases}$.

Доказательство. По теореме Лагранжа существует невырожденная матрица T, приводящая матрицу A к нормальному виду, т.е. $T^T AT$ - диагональная матрица, по главной диагонали которой расположено $s(A)$ единиц и rgA-$s(A)$ минус единиц. После замены координат $x=Ty$ получим уравнение квадрики $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 + 2\sum_{i=1}^{n} b_i y_i + \alpha = 0$. Преобразуем уравнение $\sum_{i=1}^{s(A)} (y_i + b_i)^2 - \sum_{i=1+s(A)}^{rgA} (y_i - b_i)^2 + 2\sum_{i=1+rgA}^{n} b_i y_i + \gamma = 0$, где $\gamma = \alpha - \sum_{i=1}^{s(A)} b_i^2 + \sum_{i=1+s(A)}^{rgA} b_i^2$.

Положим $z_i = \begin{cases} y_i + b_i,\ при\ i = 1,\dots,s(A) \\ y_i - b_i,\ при\ i = 1+s(A),\dots,rgA \\ y_i,\ при\ i = 1+rgA,\dots,n \end{cases}$. В новой системе координат уравнение квадрики имеет вид $\sum_{i=1}^{s(A)} z_i^2 - \sum_{i=1+s(A)}^{rgA} z_i^2 + 2\sum_{i=1+rgA}^{n} b_i z_i + \gamma = 0$.

Если $b_i = 0$, где $i = 1+rgA,\dots,n$, то $\sum_{i=1}^{s(A)} z_i^2 - \sum_{i=1+s(A)}^{rgA} z_i^2 = -\gamma$, и теорема в этом случае доказана.

Пусть найдется i, при котором выполняется неравенство $b_i \neq 0$. Если i=1+rgA, то сделаем аффинную замену координат $u_j = \begin{cases} z_j\ при\ j \neq 1+rgA \\ -\gamma/2 - \sum_{k=1+rgA}^{n} b_k z_k\ при\ j = 1+rgA \end{cases}$, а если i>1+rgA, то замену

$u_j = \begin{cases} z_j \ при \ j \neq 1+rgA, j \neq i \\ -\gamma/2 - \sum_{k=1+rgA}^{n} b_k z_k \ при \ j = 1+rgA \\ z_{1+rgA} \ при \ j = i \end{cases}$. В результате получим уравнение квадрики $\sum_{i=1}^{s(A)} u_i^2 - \sum_{i=1+s(A)}^{rgA} u_i^2 = 2u_{1+rgA}$, что и требовалось доказать.

Если уравнение квадрики умножить на не нулевое число, то множество решений уравнения не изменится. Два уравнения квадрики называются *аффинно-эквивалентными*, если от одного к другому можно перейти в результате аффинного преобразования или умножения уравнения на произвольное ненулевое число.

Теорема 5.4 Уравнение квадрики аффинно эквивалентно одному из следующих уравнений $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = \begin{cases} 0, \pm 1 \\ 2y_{1+rgA} \end{cases}$**, причем** $2s(A) \geq rgA$**. Если** $2s(A) = rgA$**, то правая часть не может равняться -1. Все эти уравнения аффинно не эквивалентны между собой.**

Доказательство. Аффинной заменой координат любое уравнение приводится к виду $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = \begin{cases} \alpha \\ 2y_{1+rgA} \end{cases}$. Если $2s(A) < rgA$, то умножим уравнение на -1. Аналогично, если $2s(A) = rgA$ и $\alpha < 0$, то умножим уравнение на -1. Если $\alpha \neq 0$, то умножим уравнение на $1/|\alpha|$. В результате этих преобразований получим уравнение вида $\sum_{i=1}^{s(A)} \lambda y_i^2 - \sum_{i=1+s(A)}^{rgA} \lambda y_i^2 = \begin{cases} 0, \pm 1 \\ 2\beta y_{1+rgA} \end{cases}$, где $\lambda > 0$ и $2s(A) \geq rgA$. Причем, если $2s(A) = rgA$, то правая часть уравнения не может равняться -1.

Сделаем замену координат $z_i = \begin{cases} \sqrt{\lambda} y_i \ при \ i = 1, \dots, rgA \\ \beta y_i \ при \ i = 1+rgA \\ y_i \ при \ i = 2+rgA, \dots, n \end{cases}$ и получим одно из уравнений квадрики, приведенных в условии теоремы.

Рассмотрим матрицы $\hat{A}$ для уравнений квадрик, приведенных в условии теоремы. Для квадрики $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = \delta$, где $\delta \in \{-1, 0, 1\}$, расширенная матрица

$$\hat{A} = \left(\begin{array}{ccc|c} E_{s(A)} & 0 & 0 & 0 \\ 0 & -E_{rgA-s(A)} & 0 & 0 \\ 0 & 0 & 0_{n-rgA} & 0 \\ \hline 0 & 0 & 0 & \delta \end{array}\right)$$

, а для квадрики $\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = 2y_{1+rgA}$ расширенная матрица

$$\hat{A} = \left(\begin{array}{cccc|c} E_{s(A)} & 0 & 0 & 0 & 0 \\ 0 & -E_{rgA-s(A)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0_{n-rgA-1} & 0 \\ \hline 0 & 0 & -1 & 0 & 0 \end{array}\right)$$

. Приведем таблицу аффинных инвариантов (**Следствие 5.2**).

$rg\hat{A}$	$s(\hat{A})$	квадрика
rgA	$s(A)$	$\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = 0$
$1+rgA$	$s(A)$	$\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = 1$

$1+rgA$	$s(A)+1$	$\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = -1$
$2+rgA$	$s(A)+1$	$\sum_{i=1}^{s(A)} y_i^2 - \sum_{i=1+s(A)}^{rgA} y_i^2 = 2y_{1+rgA}$

Поскольку все наборы инвариантов различны, то теорема доказана.

5.5 Аффинная классификация кривых второго порядка.

Теорема 5.5. Любая кривая второго порядка аффинно эквивалентна одной из 9 кривых, приведенных в таблице. Приведенные кривые аффинно не эквивалентны между собой.

Название кривой	Каноническое уравнение кривой	Расширенная матрица	rgA	$S(A)$	$rg\hat{A}$	$s(\hat{A})$
Эллипс	$x_1^2 + x_2^2 = 1$	diag(1,1,-1)	2	2	3	2
Мнимый эллипс	$x_1^2 + x_2^2 = -1$	diag(1, 1, 1)	2	2	3	3
Гипербола	$x_1^2 - x_2^2 = 1$	diag(–1, 1, –1)	2	1	3	1
Пара пересекающихся мнимых прямых	$x_1^2 + x_2^2 = 0$	diag(0, 1, 1)	2	2	2	2
Пара пересекающихся прямых	$x_1^2 - x_2^2 = 0$, $x_1 = \pm x_2$	diag(0, 1, –1)	2	1	2	1
Парабола	$x_1^2 = 2x_2$	$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$	1	1	3	2
Пара параллельных прямых	$x_1^2 = 1$	diag(–1, 1, 0)	1	1	2	1
Пара параллельных мнимых прямых	$x_1^2 = -1$	diag(1, 1, 0)	1	1	2	2
Пара совпавших параллельных	$x_1^2 = 0$	diag(0, 1, 0)	1	1	1	1

прямых						

Доказательство. Любую кривую 2-го порядка в соответствующих аффинных координатах можно описать одним из перечисленных канонических уравнений. Действительно, $2s(A) \geq rgA$ и rgA может принимать лишь два значения 1 или 2, поэтому матрица A может иметь один из следующих трёх видов: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ или $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Очевидно, приведённая таблица исчерпывает все возможные варианты расширенных матриц, соответствующих каждой из трёх матриц A.

5.6 Аффинная классификация поверхностей второго порядка

Теорема 5.6. Любая поверхность второго порядка аффинно эквивалентна одной из 17 поверхностей, приведенных в таблице. Приведенные поверхности аффинно не эквивалентны между собой.

Название поверхности	Каноническое уравнение поверхности	Расширенная матрица	rgA	$S(A)$	$rg\hat{A}$	$s(\hat{A})$
Поверхности вращения						
Эллипсоид	$x_1^2 + x_2^2 + x_3^2 = 1$	diag(–1, 1, 1, 1)	3	3	4	3
Мнимый эллипсоид	$x_1^2 + x_2^2 + x_3^2 = -1$	diag(1, 1, 1,1)	3	3	4	4
Однополостный гиперболоид	$x_1^2 + x_2^2 - x_3^2 = 1$	diag(–1, 1, 1, –1)	3	2	4	2
Двуполостный гиперболоид	$x_1^2 + x_2^2 - x_3^2 = -1$	diag(1, 1, 1, –1)	3	2	4	3
Мнимый конус	$x_1^2 + x_2^2 + x_3^2 = 0$	diag(0, 1, 1, 1)	3	3	3	3
Конус	$x_1^2 + x_2^2 - x_3^2 = 0$	diag(0, 1, 1, –1)	3	2	3	2
Эллиптический параболоид	$x_1^2 + x_2^2 = 2x_3$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$	2	2	4	3
Цилиндрические поверхности						
Эллиптический цилиндр	$x_1^2 + x_2^2 = 1$	diag(1,1,-1)	2	2	3	2
Мнимый эллиптический цилиндр	$x_1^2 + x_2^2 = -1$	diag(1, 1, 1)	2	2	3	3
гиперболический цилиндр	$x_1^2 - x_2^2 = 1$	diag(–1, 1, –1)	2	1	3	1

Пара пересекающихся мнимых плоскостей	$x_1^2+x_2^2=0$	diag(0, 1, 1)	2	2	2	2
Пара пересекающихся плоскостей	$x_1^2-x_2^2=0$, $x_1=\pm x_2$	diag(0, 1, –1)	2	1	2	1
Параболический цилиндр	$x_1^2=2x_2$	$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$	1	1	3	2
Пара параллельных плоскостей	$x_1^2=1$	diag(–1, 1, 0)	1	1	2	1
Пара параллельных мнимых плоскостей	$x_1^2=-1$	diag(1, 1, 0)	1	1	2	2
Пара совпавших плоскостей	$x_1^2=0$	diag(0, 1, 0)	1	1	1	1
Гиперболический параболоид (седло)	$x_1^2-x_2^2=2x_3$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$	2	1	4	2

6 Линейный оператор

6.1 Линейный оператор. Матрица линейного оператора.

Пусть W и V линейные пространства над числовым полем P. Однозначное отображение φ линейного пространства W в линейное пространство V называется линейным оператором, если для любых векторов x,y из W и чисел α, β из поля P справедливо равенство $\varphi(\alpha x+\beta y)=\alpha\varphi(x)+\beta\varphi(y)$.

6.1.1 Примеры линейных операторов.

1. Линейная функция
2. Дифференцирование функций
3. Проекция вектора
4. Пседообратная матрица

6.1.2 Матрица линейного оператора.

Пусть $e_1, \cdots, e_n$ базис *W*. Разложим вектор *x* из *W* по этому базису $x = x_1 e_1 + \cdots + x_n e_n$ и найдем его образ $\varphi(x) = x_1\varphi(e_1) + \cdots + x_n\varphi(e_n)$. Из полученного равенства видно, что образ вектора определяется координатами вектора и значениями линейного оператора на базисных векторах. Обозначим через $f_1, \ldots, f_m$ базис *V*. Координаты вектора *x* из *W* в базисе $e_1, \cdots, e_n$ обозначим через $[x]_e$, а координаты вектора *y* из *V* в базисе $f_1, \ldots, f_m$ обозначим через $[y]_f$. Перейдем в последнем равенстве от равенства векторов к равенству их координат $[\varphi(x)]_f = x_1[\varphi(e_1)]_f + \cdots + x_n[\varphi(e_n)]_f$, которое можно записать используя матричное умножение следующим образом $[\varphi(x)]_f = ([\varphi(e_1)]_f, \cdots, [\varphi(e_n)]_f)[x]_e$. Матрица $([\varphi(e_1)]_f, \cdots, [\varphi(e_n)]_f)$ называется матрицей линейного оператора и обозначается $[\varphi]_{e\,f}$.

6.1.3 Изменение матрицы линейного оператора при изменении базиса.

Получим формулу изменения матрицы линейного оператора при изменении базиса в пространствах *V* и *W*. Пусть $h_1, \cdots, h_n$ новый базис *W*, а $g_1, \ldots, g_m$ новый базис *V*. Координаты вектора в разных базисах связаны матрицей перехода. Пусть $[x]_e=T[x]_h$ и $[y]_f=Q[y]_g$. Отсюда и равенства $[\varphi(x)]_f = [\varphi]_{e\,f}[x]_e$ выводим $Q[\varphi(x)]_g = [\varphi]_{e\,f}T[x]_h$ или $[\varphi(x)]_g = Q^{-1}[\varphi]_{e\,f}T[x]_h$. Сопоставляя полученное равенство с $[\varphi(x)]_g = [\varphi]_{h\,g}[x]_h$, получаем равенство матриц $[\varphi]_{h\,g} = Q^{-1}[\varphi]_{e\,f}T$.

6.2 Алгебра линейных операторов.

Обозначим через WV множество линейных операторов, действующих из пространства *W* в пространство *V*. На множестве WV определим операции умножения оператора на скаляр $(\alpha\cdot\varphi)(x) = \alpha\cdot\varphi(x)$ и сложение операторов $(\varphi+\psi)(x) = \varphi(x)+\psi(x)$. Оператор назовем нулевым, если все векторы переводятся в ноль. Нулевой оператор обозначим через 0, т.е $0(x) = 0$. Относительно операций умножения на скаляр и сложения множество линейных операторов WV образует линейное пространство. Отметим, что $[\alpha\cdot\varphi] = \alpha[\varphi]$ и $[\varphi+\psi] = [\varphi]+[\psi]$.

Пусть *W,V,U* – линейные пространства над полем *P*, а φ линейный оператор из *W* в *V*, ψ - линейный оператор из *V* в *U*. Отображение $\psi(\varphi(x))$ из *W* в *U* является линейным оператором и обозначается $\psi\varphi$. Пусть $e_1, \cdots, e_n$ - базис *W*, $h_1, \cdots, h_m$ - базис *V*, $g_1, \cdots, g_k$ - базис *U*, тогда $[\psi\varphi]_{eg} = [\psi]_{hg}[\varphi]_{eh}$.

6.3 Простейший вид матрицы линейного оператора.

6.3.1 Эквивалентность матриц

Матрицы *A* и *B* называются эквивалентными, если найдутся невырожденные матрицы *Q* и *T*, что *A*=*QBT*.

Теорема 6.1. Если матрицы эквивалентны, то их ранги равны.

Доказательство. Поскольку ранг произведения не превосходит ранги сомножителей, то $rgA \le rgB$. Так как $B = Q^{-1}AT^{-1}$, то $rgB \le rgA$. Объединяя два неравенства, получаем требуемое утверждение.

Теорема 6.2. Элементарными преобразованиями со строками и столбцами матрицу *A* можно привести к блочному виду $\left(\begin{array}{c|c} E_{rgA} & 0 \\ \hline 0 & 0 \end{array}\right)$, где E_k - единичная матрица порядка *k*, а 0 – нулевая матрица соответствующих размеров.

Доказательство. Приведем алгоритм приведения матрицы *A* к указанному виду. Номера столбцов будут указываться в квадратных скобках, а номера строк – в круглых скобках.

1. Положим *r*=1.
2. Если $a_{rr} = 0$ то перейдем на шаг 4, иначе перейдем на шаг 3.

3. Сделаем преобразования со строками $(i)=(i)-a_{ir}/a_{rr}(r)$, где $i=r+1,\ldots,m$, и со столбцами $[j]=[j]-a_{rj}/a_{rr}[r]$, где $j=r+1,\ldots,n$, и $[r]=1/a_{rr}[r]$. Увеличим r на 1 и вернемся на шаг 2.
4. Если $a_{ij}=0$, при $i=r+1,\ldots,m$, $j=r+1,\ldots,n$, то конец. В противном случае найдем $i,j>r$, что $a_{ij}\neq 0$. Переставим строки $(r)\leftrightarrow(i)$ и столбцы $[r]\leftrightarrow[j]$, вернемся на шаг 2.

Очевидно, что алгоритмом будет строиться последовательность эквивалентных матриц, последняя из которых имеет требуемый вид.

Теорема 6.3. Матрицы A и B одинаковых размеров эквивалентны тогда и только тогда, когда их ранги равны.

Доказательство. Если матрицы эквивалентны, то их ранги равны (**Теорема 6.1**). Пусть ранги матриц равны. Тогда найдутся невырожденные матрицы, что $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}=Q_1AT_1=Q_2BT_2$, где $r=rgA=rgB$ (**Теорема 6.2**).

Следовательно, $A=\left(Q_1^{-1}Q_2\right)B\left(T_2T_2^{-1}\right)$, и матрицы A и B – эквивалентны.

Результаты данного пункта позволяют находить простейший вид матрицы линейного оператора и базисы пространств, в которых матрица линейного оператора имеет этот простейший вид.

6.3.2 Ранг, дефект линейного оператора.

Образ нуля равен нулю. Действительно, $\varphi(0)=\varphi(0+0)=2\varphi(0)$, отсюда $\varphi(0)=0$.
Множество векторов из W, образ которых равен 0, называется *ядром* линейного оператора. Ядро линейного преобразования обозначим $Ker\,\varphi$ ($Ker\,\varphi=\{x|\ \varphi(x)=0\}$). Ядро является подпространством W (докажите) и его размерность называют *дефектом* и обозначают $def\ \varphi$.

Множество всех образов векторов из W обозначают $\boldsymbol{Im}\,\varphi$ ($\boldsymbol{Im}\,\varphi=\{y|\ \exists x\in W, y=\varphi(x)\}$). Множество образов является подпространством V (докажите), его размерность называют рангом линейного оператора и обозначают $rg\,\varphi$.

Теорема 6.4. $\boldsymbol{dim}W=rg\,\varphi+def\ \varphi$.

Доказательство. Пусть $g_1,\ldots,g_k$ – базис $\boldsymbol{Im}\,\varphi$. По определению $\boldsymbol{Im}\,\varphi$ для каждого вектора g_i существует прообраз h_i из W. Система векторов $h_1,\ldots,h_k$ является линейно независимой. Действительно, из равенства $\alpha_1h_1+\ldots+\alpha_kh_k=0$, выводим $\varphi(\alpha_1h_1+\ldots+\alpha_kh_k)=\varphi(0)$, или

$\alpha_1g_1+\ldots+\alpha_kg_k=0$. В силу линейной независимости, все коэффициенты равны 0, и система $h_1,\ldots,h_k$ является линейно независимой. Аналогично показывается, что пересечение линейной оболочки векторов $h_1,\ldots,h_k$ и $Ker\,\varphi$ состоит только из нулевого вектора. Действительно, из включения $\alpha_1h_1+\ldots+\alpha_kh_k\in Ker\,\varphi$, выводим $\varphi(\alpha_1h_1+\ldots+\alpha_kh_k)=0$, и далее, $\alpha_1g_1+\ldots+\alpha_kg_k=0$. Для любого вектора x из W найдутся коэффициенты, что $\varphi(x)=\alpha_1g_1+\cdots+\alpha_kg_k$, и $\varphi(x-\alpha_1h_1-\cdots-\alpha_kh_k)=0$. Таким образом W представляется в виде прямой суммы линейной оболочки векторов $h_1,\ldots,h_k$ и $Ker\,\varphi$. Теорема вытекает из свойства прямой суммы.

Следствие 6.1. Можно выбрать базисы в пространствах W и V так, чтобы матрица линейного оператора имела диагональный вид, причем по диагонали расположены 1 и 0. Количество ненулевых элементов на диагонали равно рангу оператора.

Доказательство. Пусть $g_1,\ldots,g_k$ и $h_1,\ldots,h_k$ имеют тот же смысл, что и в доказательстве предыдущей теоремы. Дополним векторы $g_1,\ldots,g_k$ до базиса V, а векторы $h_1,\ldots,h_k$ до базиса W векторами из $Ker\,\varphi$. Полученные базисы обозначим через $g_1,\ldots,g_m$ и $h_1,\ldots,h_n$, соответственно. Построим матрицу линейного оператора в этих базисах. Заметим, $\varphi(h_i)=\begin{cases} g_i \text{ при } i\le k \\ 0 \text{ при } i>k \end{cases}$, а координаты вектора g_i в базисе

$g_1,\dots,g_m$ равны (0,…,0,1,0,…,0), где 1 стоит на i-ом месте. Таким образом, матрица линейного оператора в этих базисах имеет диагональный вид, причем по диагонали расположены 1 и 0. Количество 1 равно рангу оператора.

7 Линейное преобразование

7.1 Линейное преобразование. Его матрица

Однозначное отображение φ линейного пространства V над числовым полем P в себя называется линейным преобразованием, если оно сохраняет линейность, то есть $\varphi(\alpha x+\beta y)=\alpha\varphi(x)+\beta\varphi(y)$ для любых $x,y\in V$ и $\alpha,\beta\in P$.

Линейное преобразование полностью определяется своими значениями на базисных векторах. Действительно, пусть $e_1,\cdots,e_n$ базис V. Вектор x разложим по базису $x=x_1e_1+\cdots+x_ne_n$, где $(x_1,\cdots,x_n)^{\mathrm{T}}$ - координаты вектора x. По свойству линейного преобразования имеем $\varphi(x)=x_1\varphi(e_1)+\cdots+x_n\varphi(e_n)$. Перейдем в последнем равенстве от равенства векторов к равенству их координат $[\varphi(x)]_e=x_1[\varphi(e_1)]_e+\cdots+x_n[\varphi(e_n)]_e$, которое можно записать используя матричное умножение следующим образом $[\varphi(x)]_e=([\varphi(e_1)]_e,\cdots,[\varphi(e_n)]_e)[x]_e$. Матрица $([\varphi(e_1)]_e,\cdots,[\varphi(e_n)]_e)$ называется *матрицей линейного преобразования* и обозначается $[\varphi]_e$. Матрица линейного преобразования связывает координаты образа с координатами исходного вектора $[\varphi(x)]_e=[\varphi]_e[x]_e$.

7.2 Изменение матрицы линейного преобразования при изменении базиса.

Поскольку линейное преобразование частный случай линейного оператора, то можно воспользоваться полученной ранее формулой $[\varphi]_f=P^{-1}[\varphi]_eP$, где P – матрица перехода. Матрицы A и B называются подобными, если существует невырожденная матрица P, что $A=P^{-1}BP$. Вопрос о подобии матриц сводится к решению системы линейных уравнений $PA=BP$, где в роли неизвестных выступают элементы матрицы P, с дополнительным нелинейным условием $\det P\neq 0$.

7.3 Алгебра линейных преобразований.

На множестве всех линейных преобразований пространства V расмотрим операции:

1. Умножение на число: $(\alpha\varphi)(x)=\alpha\varphi(x)$.
2. Сложение (вычитание) $(\varphi+\phi)(x)=\varphi(x)+\phi(x)$
3. Умножение $(\varphi\phi)(x)=\varphi(\phi(x))$.

Легко проверить линейность всех этих преобразований и вывести следующие формулы, связывающие их матрицы

1. $[\alpha\varphi]=\alpha[\varphi]$
2. $[\varphi+\phi]=[\varphi]+[\phi]$
3. $[\varphi\phi]=[\varphi][\phi]$

Линейное преобразование, переводящее каждый вектор в себя, называется тождественным преобразованием и обозначается ε. В любом базисе матрица тождественного преобразования равна единичной.

Пусть $f(t)=\alpha_0+\alpha_1t+\cdots+\alpha_kt^k$ - некоторый многочлен, φ - линейное преобразование пространства V. Сопоставим многочлену $f(t)$ линейное преобразование $f(\varphi)=\alpha_0\varepsilon+\alpha_1\varphi+\cdots+\alpha_k\varphi^k$. Будем говорить, что преобразование $f(\varphi)$ получено подстановкой φ в многочлен $f(t)$. Матрица $f(\varphi)$ может быть вычислена по формуле $[f(\varphi)]=\alpha_0E+\alpha_1[\varphi]+\cdots+\alpha_k[\varphi]^k$.

Свойство 7.1. Пусть $f(t)=g(t)h(t)$. Тогда $f(\varphi)=g(\varphi)h(\varphi)$.

7.4 Инвариантные пространства

Подпространство W называется *инвариантным* относительно линейного преобразования φ, если для любого x из W его образ $\varphi(x)$ также принадлежит W.

Свойство 7.2. $Ker(\varphi)$ **- инвариантное подпространство.**

Доказательство. Пусть $x \in Ker(\varphi)$. Тогда $\varphi(x) = 0 \in Ker(\varphi)$.

Свойство 7.3. $\boldsymbol{Im}(\varphi)$ **- инвариантное подпространство.**

Доказательство. Пусть $x \in \boldsymbol{Im}(\varphi)$, **тогда** $\varphi(x) \in \boldsymbol{Im}(\varphi)$.

Свойство 7.4. Пусть $f(t)$ **- многочлен, тогда** $Ker(f(\varphi))$ **инвариантное пространство относительно** φ**.**

Доказательство. Пусть $x \in Ker(f(\varphi))$, то есть $f(\varphi)(x) = 0$. Далее, $f(\varphi)(\varphi(x)) = (f(\varphi)\varphi)(x) = \varphi(f(\varphi)(x)) = \varphi(0) = 0$, то есть $\varphi(x) \in Ker(f(\varphi))$.

Свойство 7.5. Пусть $f(t)$ **- многочлен, тогда** $\boldsymbol{Im}(f(\varphi))$ **инвариантное пространство относительно** φ**.**

Доказательство. Пусть $x \in \boldsymbol{Im}(f(\varphi))$, тогда $x = f(\varphi)(y)$. Далее, $\varphi(x) = \varphi(f(\varphi)(y)) = f(\varphi)(\varphi(y))$, то есть $\varphi(x) \in \boldsymbol{Im}(f(\varphi))$.

Знание инвариантных подпространств позволяет найти базис пространства, в котором матрица линейного преобразования имеет простую структуру. Действительно, пусть $e_1, \ldots, e_k$ базис инвариантного подпространства W. Дополним его до базиса всего пространства векторами $e_{k+1}, \ldots, e_n$. Координаты образов первых k векторов могут иметь только k первых ненулевых компонент. Следовательно, в матрице линейного преобразования содержится в левом нижнем углу блок размером $(n-k)*k$, состоящий из одних нулей.

Если пространство V представляется в виде прямой суммы инвариантных подпространств W и U, то построим базис пространства V, объединив базисы W и U. В построенном базисе матрица линейного преобразования будет иметь блочно диагональный вид.

Таким образом, структура матрицы линейного преобразования имеет тем более простой вид, чем меньше размерность инвариантных подпространств, в прямую сумму которых расщепляется исходное пространство V.

7.5 Собственные векторы и собственные числа. Характеристическое уравнение.

Базис одномерного инвариантного подпространства называется собственным вектором. Другими словами, ненулевой вектор x называется собственным, если $\varphi(x) = \lambda x$. Число λ называется собственным. Запишем это равенство в координатах $[\varphi][x] = \lambda[x]$, или $([\varphi] - \lambda E)[x] = 0$. Последнее равенство можно рассматривать как квадратную систему линейных уравнений с n неизвестными. По правилу Крамера, если $\det([\varphi] - \lambda E) \neq 0$, то система имеет единственное нулевое решение. Следовательно, собственные числа являются корнями уравнения $\det([\varphi] - \lambda E) = 0$. Данное уравнение называется характеристическим. Обратно, если λ корень характеристического уравнения, то система $([\varphi] - \lambda E)[x] = 0$ имеет ненулевое решение, и значит, λ является собственным числом. Тем самым доказана теорема.

Теорема 7.1. Корнями характеристического уравнения являются только собственные числа. Все собственные числа являются корнями характеристического уравнения.

Коэффициенты характеристического уравнения не зависят от выбора базиса. Действительно, матрицы линейного преобразования в разных базисах связаны уравнением $[\varphi]_e = P^{-1}[\varphi]_f P$, откуда

$$\det([\varphi]_t - \lambda E) = \det P^{-1}([\varphi]_f - \lambda E)P = \det([\varphi]_f - \lambda E).$$

Собственные векторы для собственного числа λ принадлежат ядру линейного преобразования $\varphi - \lambda\varepsilon$. Подпространство $Ker(\varphi - \lambda\varepsilon)$ называется *корневым* подпространством, соответствующим собственному числу λ.

Приведем простые факты.

Следствие 7.1. Линейное преобразование φ линейного пространства V над полем комплексных чисел имеет собственный вектор.

Доказательство. Над полем комплексных чисел характеристический многочлен имеет хотя бы один корень, а, значит, линейное преобразование имеет собственный вектор.

Следствие 7.2. Линейное преобразование φ линейного пространства V над полем вещественных чисел имеет инвариантное подпространство размерности не выше 2.

Доказательство. Пусть φ - линейное преобразование пространства *V* над полем *R*. Если характеристический многочлен имеет вещественный корень, то утверждение леммы очевидно. На множестве $\widetilde{V} = \{x + iy | \, x, y \in V\}$ определим операцию сложения $(x+iy)+(a+ib)=(x+a)+i(y+b)$ и умножения на комплексное число $(\alpha + i\beta)(x+iy) = (\alpha x - \beta y) + i(\beta x + \alpha y)$. Множество $\widetilde{V}$ относительно введенных операций сложения векторов и умножения на скаляр образует линейное пространство над *C*. Вектор *x* из *V* можно рассматривать как вектор из пространства $\widetilde{V}$, записанный в виде *x*+*i*0. Базис пространства *V* является базисом пространства $\widetilde{V}$, и, значит, размерности пространств *V* и $\widetilde{V}$ совпадают. В пространстве $\widetilde{V}$ рассмотрим линейное преобразование $\widetilde{\varphi}(x+iy) = \varphi(x) + i\varphi(y)$. Пусть $e_1, \ldots, e_n$ - базис *V*. Тогда $e_1, \ldots, e_n$ - базис $\widetilde{V}$ и $[\varphi]_e = [\widetilde{\varphi}]_e$. Пусть $\alpha = \mu + i\nu$ - комплексное собственное число, а $x + iy$ - соответствующий собственный вектор линейного преобразования $\widetilde{\varphi}$. Тогда $\widetilde{\varphi}(x+iy) = \varphi(x) + i\varphi(y) = (\mu + i\nu)(x+iy) = (\mu x - \nu y) + i(\nu x + \mu y)$, и, значит, $\varphi(x) = \mu x - \nu y$, $\varphi(y) = \nu x + \mu y$. Линейная оболочка векторов *x,y* образует двумерное инвариантное подпространство.

7.6 Коэффициенты характеристического уравнения. След матрицы.

Теорема 7.2. Коэффициент характеристического уравнения при $(-\lambda)^{n-k}$ равен

$$\sum_{1 \le i_1 < \cdots < i_k \le n} \boldsymbol{det}\, A\begin{pmatrix} i_1, \ldots, i_k \\ i_1, \ldots, i_k \end{pmatrix}.$$

Доказательство получается раскрытием определителя $\boldsymbol{det}(A - \lambda E)$.

Сумма элементов матрицы *A*, расположенных на главной диагонали, называется следом матрицы. След матрицы является коэффициентом характеристического многочлена и не зависит от выбора базиса..

7.7 Диагонализируемые преобразования

Линейное преобразование называется диагонализируемым, если существует базис, в котором матрица линейного преобразования имеет диагональный вид. Заметим, что базис, в котором матрица линейного преобразования имеет диагональный вид, образован собственными векторами. Верно и обратное. В базисе из собственных векторов матрица линейного преобразования имеет диагональный вид. Не каждое линейное преобразование диагонализируемо. Например, линейное преобразование, заданное матрицей $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ не диагонализируемо.

Теорема 7.3. Собственные векторы, отвечающие различным собственным значениям, линейно независимы.

Доказательство. Пусть $e_{i1}, \ldots, e_{ik_i}$ - линейно независимая система собственных векторов, соответствующих собственному значению α_i, где *i*=1,…,*s*. Покажем линейную независимость системы векторов $e_{11}, \ldots, e_{sk_s}$ индукцией по *s*. При *s*=1 утверждение очевидно. Пусть оно верно для *s*-1. Покажем его справедливость для s. Допустим, система $e_{11},\dots,e_{sk_s}$ - линейно зависима. Тогда найдутся коэффициенты β_{ij} не все равные нулю, что $\beta_{11}e_{11}+\dots+\beta_{sk_s}e_{sk_s}=0$. Из этого равенства выводим $(\varphi-\alpha_s\varepsilon)(\beta_{11}e_{11}+\dots+\beta_{sk_s}e_{sk_s})=0$ или $\beta_{11}(\alpha_1-\alpha_s)e_{11}+\dots+\beta_{s-1k_{s-1}}(\alpha_{s-1}-\alpha_s)e_{s-1k_{s-1}}=0$. По предположению индукции все коэффициенты в этом равенстве равны 0, и, значит $\beta_{ij}=0$ при $i<s$. Но тогда система $e_{s1},\dots,e_{sk_s}$ - линейно зависима, что противоречит условиям теоремы. К полученному противоречию привело допущение о линейной зависимости системы векторов $e_{11},\dots,e_{sk_s}$, значит, эта система линейно независима, что и требовалось доказать.

Рассмотрим вопрос о количестве линейно независимых собственных векторов, соответствующих собственному числу α.

Геометрической кратностью собственного числа α называется дефект преобразования $\varphi-\alpha\varepsilon$, а *алгебраической кратностью* называется кратность корня α в характеристическом многочлене.

Теорема 7.4. Геометрическая кратность α не превосходит его алгебраической кратности.

Доказательство. Пусть геометрическая кратность α равна k. Дополним базис $e_1,\dots,e_k$ ядра преобразования $\varphi-\alpha\varepsilon$ до базиса всего пространства $e_1,\dots,e_n$. Матрица линейного преобразования в этом базисе имеет вид $[\varphi]_e=\begin{pmatrix}\alpha E_k & A\\ 0 & B\end{pmatrix}$ и характеристический многочлен равен $\boldsymbol{det}([\varphi]_e-\lambda E)=(\alpha-\lambda)^k\,\boldsymbol{det}(B-\lambda E)$. Таким образом, алгебраическая кратность α не меньше геометрической кратности, что и требовалось доказать.

Теорема 7.5 Линейное преобразование φ линейного пространства V над числовым полем P диагонализируемо тогда и только тогда, когда характеристический многочлен раскладывается над полем P на линейные множители и алгебраическая кратность каждого корня совпадает с его геометрической кратностью.

Доказательство очевидно.

7.8 Теорема Шура

Пусть φ - линейное преобразование пространства V над полем комплексных чисел C. Линейное преобразование φ имеет хотя бы один собственный вектор (Следствие 7.1). Этот факт можно усилить.

Теорема 7.6. Пусть φ - линейное преобразование пространства V над полем комплексных чисел C. Существует базис V, в котором матрица линейного преобразования φ имеет верхний треугольный вид.

Доказательство проведем индукцией по размерности V. Пусть утверждение верно для линейных преобразований (n-1)-мерных пространств. Покажем его справедливость для линейного преобразования φ n-мерного линейного пространства V. Поскольку линейное пространство над полем C, то существует собственный вектор h этого линейного преобразования. Дополним этот вектор до базиса всего пространства векторами $h_2,\dots,h_n$. Матрица линейного преобразования в этом базисе имеет блочный вид $[\varphi]_h=\left(\begin{array}{c:c}\alpha & a\\ \hdashline 0 & A\end{array}\right)$, где α - собственное число для вектора h. Обозначим через W линейную оболочку векторов $h_2,\dots,h_n$. Векторы $h_2,\dots,h_n$ образуют базис W. Обозначим через ψ линейное преобразование W, матрица которого в базисе $h_2,\dots,h_n$ равна A. По предположению индукции в подпространстве W существует базис $f_2,\dots,f_n$, в котором матрица линейного преобразования имеет верхний треугольный вид. Пусть T – матрица перехода к этому базису. Тогда $T^{-1}AT$ - верхняя треугольная матрица. Матрица перехода от базиса $h, h_2, \ldots, h_n$ к базису $h, f_2, \ldots, f_n$ равна $\begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}$, и, значит, матрица φ в базисе $h, f_2, \ldots, f_n$ равна $\begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}^{-1} \begin{pmatrix} \alpha & a \\ 0 & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix} = \begin{pmatrix} \alpha & aT \\ 0 & T^{-1}AT \end{pmatrix}$, то есть является верхней треугольной.

Аналогом доказанной теоремы над полем вещественных чисел является следующий результат.

Теорема 7.7. Пусть φ - линейное преобразование пространства *V* над полем вещественных чисел *R*. Существует базис *V*, в котором матрица линейного преобразования φ имеет блочный верхний треугольный вид. По главной диагонали стоят блоки первого и второго порядка.

Доказательство проведем индукцией по размерности *n* пространства *V*. Пусть утверждение верно для линейных преобразований пространств размерности меньшей *n*. Покажем его справедливость для линейного преобразования φ *n*-мерного линейного пространства *V*. Линейное преобразование φ имеет либо одномерное, либо двумерное инвариантное подпространство (**Следствие 7.2**). Дополним базис этого инвариантного подпространства до базиса всего пространства векторами $h_k, \ldots, h_n$, где *k* равно либо 2, либо 3. Матрица линейного преобразования в этом базисе имеет блочный вид $[\varphi]_h = \left(\begin{array}{c:c} \alpha & a \\ \hdashline 0 & A \end{array}\right)$, где α - блок либо первого, либо второго порядка. Далее, рассуждения повторяют доказательство теоремы 7.6.

Теорема 7.8. (теорема Шура). Для линейного преобразования φ унитарного пространства *V* существует ортонормированный базис, в котором матрица линейного преобразования φ имеет верхний треугольный вид.

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ - базис *V*, в котором матрица линейного преобразования φ имеет верхний треугольный вид (**Теорема 7.6**). Применим к базису процесс ортогонализации и построим ортогональный базис $h_1, h_2, \ldots, h_n$. Матрица перехода *T* от базиса $e_1, e_2, \ldots, e_n$ к базису $h_1, h_2, \ldots, h_n$ - верхняя треугольная и $[\varphi]_h = T^{-1}[\varphi]_e T$. Поскольку произведение верхних треугольных матриц является верхней треугольной матрицей, то матрица $[\varphi]_h$ - верхняя треугольная. Положим $f_i = 1/|h_i| h_i$, где *i*=1,…,*n*. Базис $f_1, f_2, \ldots, f_n$ - ортонормированный и матрица линейного преобразования в этом базисе – верхняя треугольная, тем самым теорема доказана.

Теорема 7.9. Для линейного преобразования φ евклидова пространства *V* существует ортонормированный базис, в котором матрица линейного преобразования φ имеет блочный верхний треугольный вид. По главной диагонали расположены блоки первого и второго порядков.

Доказательство аналогично доказательству теоремы 7.7.

8 Сопряженные преобразования.

8.1 Линейное преобразование и билинейные функции

Пусть *V* евклидово (унитарное) пространство. Обозначим через LP множество всех линейных преобразований пространства *V*, а через *B* множество билинейных функций, заданных на *V*. Если $\varphi \in LP$, то функция $f(x, y) = (\varphi(x), y)$ является билинейной. Таким образом, определено однозначное отображение множества линейных преобразований *LP* в множество билинейных функций *B*. Исследуем свойства этого отображения.

Свойство 8.1. Разные линейные преобразования отображаются в разные билинейные функции.

Доказательство проведем методом от противного. Пусть найдутся два разных линейных преобразования φ и ψ, которые отображаются в одну и ту же билинейную функцию. Тогда для любых векторов $x, y \in V$ справедливо равенство $(\varphi(x), y) = (\psi(x), y)$ или $(\varphi(x) - \psi(x), y) = 0$. Положим $y = \varphi(x) - \psi(x)$, тогда $|\varphi(x) - \psi(x)|^2 = 0$ и $\varphi(x) = \psi(x)$ для любого вектора $x \in V$. Это означат, что линейные преобразования равны, что противоречит допущению.

Свойство 8.2. Отображение линейных преобразований в билинейные функции взаимно однозначно.

Доказательство. Покажем, что для любой билинейной функции $f(x,y)$ существует линейное преобразование φ, что $f(x,y)=(\varphi(x),y)$. Для каждого вектора *x* определим подпространство $W_x=\{y|\ f(x,y)=0\}$. Ортогональное дополнение к этому подпространству имеет размерность не выше 1. Действительно, если $z\in W_x^{\perp}$ и $z\neq 0$, то $f(x,z)\neq 0$ и для вектора $y\in V$ справедливо включение $y-f(x,y)/f(x,z)z\in W_x$, и, значит $V=L(z)\oplus W_x$. Определим функцию $\varphi(x)=f(x,z)/|z|^2 z$, где z – базис $W_x^{\perp}$. Если $W_x^{\perp}=0$, то положим $\varphi(x)=0$. Легко убедиться, что $(\varphi(x),y)=f(x,y)$, и, значит функция $\varphi(x)$ - линейное преобразование.

Аналогично, можно рассмотреть отображение *LP* на *B*, задаваемое формулой $(x,\varphi(y))=f(x,y)$. Это отображение взаимно однозначно.

Линейное преобразование ψ называется *сопряженным* преобразованием к φ, если для любых векторов *x,y* из *V* справедливо равенство $(\varphi(x),y)=(x,\psi(y))$. Сопряженное преобразование к φ обозначают φ^*.

8.2 Сопряженное преобразование. Свойства.

Пусть $e_1,\ldots,e_n$ базис *V*, $[\varphi]_e$ - матрица линейного преобразования φ, G_e – матрица Грама скалярного произведения. Перейдем от равенства векторов $(\varphi(x),y)=(x,\varphi^*(y))$ к равенству координат $([\varphi]_e[x]_e)^{\mathrm{T}}G_e\overline{[y]_e}=[x]_e^{\mathrm{T}}G_e\overline{([\varphi^*]_e[y]_e)}$. Из этого равенства выводим $[\varphi^*]_e=\overline{G_e^{-1}[\varphi]_e^{\mathrm{T}}G_e}$. В случае ортонормированного базиса формула принимает более простой вид $[\varphi^*]_e=\overline{[\varphi]_e^{\mathrm{T}}}$. Для евклидова пространства, знак комплексного сопряжения можно опустить.

Свойство 8.3. Перечислим свойства сопряженного преобразования

1) $(\varphi\pm\phi)^*=\varphi^*\pm\phi^*$
2) $(\varphi\phi)^*=\phi^*\varphi^*$
3) $(\alpha\varphi)^*=\overline{\alpha}\varphi^*$
4) $(\varphi^*)^*=\varphi$
5) **Если *W* инвариантное подпространство φ, то ортогональное дополнение к *W* инвариантно относительно φ^*.**

Доказательство. Из равенства $(x,(\varphi+\phi)^*y)=(\varphi(x)+\phi(x),y)=(x,\varphi^*(y))+(x,\phi^*(y))$ выводим первое свойство. Второе свойство получается из равенств $(x,(\varphi\phi)^*y)=(\varphi(\phi(x)),y)=(\phi(x),\varphi^*(y))=(x,\phi^*\phi^*y)$. Для доказательства третьего свойства достаточно рассмотреть равенства $(x,(\alpha\varphi)^*y)=\alpha(\varphi(x),y)=(x,\overline{\alpha}\varphi(y))$. Четвертое свойство доказывается равенствами $(x,(\varphi^*)^*y)=(\varphi^*(x),y)=\overline{(y,\varphi^*(x))}=\overline{(\varphi(y),x)}=(x,\varphi(y))$. Докажем пятое свойство. Для произвольного вектора *x* из *W* и произвольного вектора $y\in W^{\perp}$ скалярное произведение $(\varphi(x),y)=0$. По определению сопряженного преобразования $0=(\varphi(x),y)=(x,\varphi^*(y))$, и, значит $\varphi^*(y)\in W^{\perp}$, что и требовалось доказать.

Пятое свойство позволяет дать другое доказательство теоремы Шура.

8.3 Нормальное преобразование и его свойства.

Преобразование называется нормальным, если оно перестановочно с сопряженным преобразованием, то есть $\varphi\varphi^*=\varphi^*\varphi$.

Свойство 8.4. Если *x* собственный вектор нормального преобразования φ с собственным значением α, то *x* собственный вектор φ^* с собственным значением $\overline{\alpha}$.

Доказательство. Пусть $\varphi(x)=\alpha x$. Поскольку $(\varphi(x),\varphi(x))=(x,\varphi^*\varphi(x))=(x,\varphi\varphi^*(x))=(\varphi^*(x),\varphi^*(x))$ и $(x,\varphi(x))=(\varphi^*(x),x)$, то $0=|\varphi(x)-\alpha x|^2=|\varphi(x)|^2-\alpha(x,\varphi(x))-\overline{\alpha}(\varphi(x),x)+|\alpha x|^2=$
$=|\varphi^*(x)|^2-\alpha(\varphi^*(x),x)-\overline{\alpha}(x,\varphi^*(x))+|\overline{\alpha}x|^2=|\varphi^*(x)-\overline{\alpha}x|^2$.

Свойство 8.5. Собственные векторы нормального преобразования, соответствующие разным собственным значениям ортогональны.

Доказательство. Пусть x и y – собственные векторы нормального преобразования φ, соответствующие разным собственным значениям α и β ($\varphi(x)=\alpha x$, $\varphi(y)=\beta y$). Из равенств $\varphi^*\varphi(x)=|\alpha|^2 x$ и $\varphi^*\varphi(y)=|\beta|^2 y$ (**Свойство 8.4**) выводим $(\varphi^*\varphi(x),y)=|\alpha|^2(x,y)$,
$(\varphi^*\varphi(x),y)=(\varphi(x),\varphi(y))=\alpha\overline{\beta}(x,y)$, $(\varphi^*\varphi(x),y)=(\varphi^*(x),\varphi^*(y))=\overline{\alpha}\beta(x,y)$,
$(\varphi^*\varphi(x),y)=(x,\varphi^*\varphi(x))=|\beta|^2(x,y)$. Далее, $0=\left(|\alpha|^2-\overline{\alpha}\beta-\alpha\overline{\beta}+|\beta|^2\right)(x,y)=|\alpha-\beta|^2(x,y)$, откуда $(x,y)=0$.

Теорема 8.1. Для нормального преобразования конечномерного унитарного пространства существует ортонормированный базис из собственных векторов.

Доказательство. Путь $e_1,\ldots,e_n$ - ортонормированный базис унитарного пространства *V*, в котором матрица нормального преобразования φ является верхней треугольной. Пусть $[\varphi]_e=A$, тогда $[\varphi^*]_e=\overline{A^{\mathrm{T}}}$. Из равенства $A\cdot\overline{A^{\mathrm{T}}}=\overline{A^{\mathrm{T}}}\cdot A$ вытекает, что матрица *A* – диагональная, и, значит, базис $e_1,\ldots,e_n$ составлен из собственных векторов.

Построение ортонормированного базиса из собственных векторов, в котором матрица нормального преобразования диагонализируема, можно осуществлять следующим образом. Найти какой ни будь базис из собственных векторов. При этом, собственные векторы, соответствующие разным собственным числам заведомо ортогональны (**Свойство 8.5**). Условие ортогональности может нарушаться только на собственных векторах, соответствующих одному и тому же собственному значению.

Если матрица линейного преобразования диагонализируема, то всегда можно ввести скалярное произведение таким образом, чтобы линейное преобразование стало нормальным.

Теорема 8.2. Для нормального преобразования конечномерного евклидова пространства существует ортонормированный базис, в котором матрица линейного преобразования имеет блочно-диагональный вид. По главной диагонали расположены блоки первого и второго порядка.

Доказательство. Путь $e_1,\ldots,e_n$ - ортонормированный базис евклидова пространства *V*, в котором матрица нормального преобразования φ является блочной верхней треугольной. Пусть $[\varphi]_e=A$, тогда $[\varphi^*]_e=A^{\mathrm{T}}$. Из равенства $AA^{\mathrm{T}}=A^{\mathrm{T}}A$ вытекает, что матрица *A* – блочно диагональная, что и требовалось доказать.

К сожалению, приведенное доказательство не раскрывает структуру блоков второго порядка, расположенных на главной диагонали. Поэтому дадим другое доказательство этой теоремы.

Доказательство 2. Множество $\widetilde{V}=\{x+iy \mid x,y\in V\}$ является линейным пространством над полем комплексных чисел *C*. В этом линейном пространстве введем скалярное произведение $(x+iy,a+ib)=(x,a)+(y,b)+i((y,a)-(x,b))$. Определим линейное преобразование пространства $\widetilde{V}$ как $\widetilde{\varphi}(x+iy)=\varphi(x)+i\varphi(y)$. Пусть $e_1,\ldots,e_n$ - ортонормированный базис V, тогда $e_1,\ldots,e_n$ - ортонормированный базис унитарного пространства $\widetilde{V}$ и $[\widetilde{\varphi}]_e=[\varphi]_e$ - матрица с вещественными элементами. Далее, $[\varphi^*]_e=[\varphi]_e^{\mathrm{T}}$, $[\widetilde{\varphi}^*]_e=\overline{[\varphi]_e}^{\mathrm{T}}=[\varphi]_e^{\mathrm{T}}$ и из равенства матриц $[\varphi]_e[\varphi]_e^{\mathrm{T}}=[\varphi]_e^{\mathrm{T}}[\varphi]_e$ выводим равенство $\widetilde{\varphi}\widetilde{\varphi}^*=\widetilde{\varphi}\widetilde{\varphi}^*$, то есть преобразование $\widetilde{\varphi}$ - нормальное. Следовательно, существует ортонормированный базис $f_1,\ldots,f_n$ унитарного пространства $\widetilde{V}$ из собственных векторов нормального

преобразования $\tilde{\varphi}$. Пусть $\alpha_1,\dots,\alpha_n$ - собственные числа этих векторов. Заметим, что ортонормированный базис $\tilde{V}$ получается объединением ортонормированных базисов подпространств $\boldsymbol{ker}(\tilde{\varphi}-\alpha_j\varepsilon)$. Если собственное число α вещественное, то ортонормированный базис подпространства $\boldsymbol{ker}(\varphi-\alpha\varepsilon)$ является также ортонормированным базисом подпространства $\boldsymbol{ker}(\tilde{\varphi}-\alpha\varepsilon)$. Поэтому, не нарушая общности можно считать, что вещественным собственным числам в базисе $f_1,\dots,f_n$ соответствуют векторы из V. Пусть $f=x+iy$ – собственный вектор $\tilde{\varphi}$ с комплексным собственным числом $\alpha+i\beta$, тогда из равенств $\tilde{\varphi}(x+iy)=\varphi(x)+i\varphi(y)$ и $\tilde{\varphi}(x+iy)=(\alpha+i\beta)(x+iy)=(\alpha x-\beta y)+i(\beta x+\alpha y)$ выводим $\varphi(x)=\alpha x-\beta y$, $\varphi(y)=\beta x+\alpha y$, то есть линейное подпространство, натянутое на векторы x,y – инвариантно. Из полученных равенств вытекает $\tilde{\varphi}(x-iy)=(\alpha-i\beta)(x-iy)$, то есть вектор x-iy – собственный с собственным числом $\alpha-i\beta$. Если $h_1,\dots,h_k$ ортонормированный базис $\boldsymbol{ker}(\varphi-(\alpha+i\beta)\varepsilon)$, то $\bar{h}_1,\dots,\bar{h}_k$ - ортонормированный базис $\boldsymbol{ker}(\varphi-(\alpha-i\beta)\varepsilon)$, поэтому, можно считать, что в базисе $f_1,\dots,f_n$ собственные векторы с комплексными собственными числами разбиты на пары. Рассмотрим пару $f_{2s}=x_s+iy_s$, $f_{2s+1}=x_s-iy_s$ собственных векторов с собственными числами $\alpha+i\beta$ и $\alpha-i\beta$. Эти векторы ортогональны всем остальным векторам из базиса, следовательно, векторы x_s,y_s ортогональны всем остальным векторам. Далее,
$0=(x_s+iy_s,x_s-iy_s)=|x_s|^2-|y_s|^2+2i(x_s,y_s)$, откуда выводим $|x_s|=|y_s|$ и $(x_s,y_s)=0$. Заменим векторы f_{2s} и f_{2s+1} на $\sqrt{2}x_s,\sqrt{2}y_s$ получим ортонормированный базис пространства V, в котором матрица линейного преобразования имеет блочно диагональный вид. По главной диагонали расположены блоки первого порядка, отвечающие вещественным собственным значениям, и блоки второго порядка $\begin{pmatrix}\alpha & \beta\\ -\beta & \alpha\end{pmatrix}$, отвечающие $\alpha+i\beta$ комплексным собственным значениям.

Если матрица линейного преобразования диагонализируема, то всегда можно ввести скалярное произведение таким образом, чтобы линейное преобразование стало нормальным.

8.4 Ортогональные преобразования

Линейное преобразование называется ортогональным (унитарным) если оно сохраняет скалярное произведение, то есть $(x,y)=(\varphi(x),\varphi(y))$. Из определения выводим $(x,\varphi^*\varphi(y))=(\varphi(x),\varphi(y))=(x,y)$ или $\varphi\varphi^*=\varepsilon=\varphi^*\varphi$. Таким образом ортогональное преобразование является нормальным.

Свойство 8.6. Собственные числа ортогонального преобразования по модулю равны 1.

Доказательство. Пусть $\varphi(x)=\alpha x$, тогда $|x|^2=|\varphi(x)|^2=|\alpha x|^2=|\alpha|^2|x|^2$, и, значит, $|\alpha|=1$.

Следствие 8.1. Ортогональное преобразование евклидова пространства, в некотором ортонормированном базисе, сводится к выполнению последовательности тождественных преобразований, симметрий и поворотов в координатных плоскостях.

Доказательство. Ортогональное преобразование нормально, следовательно, существует ортонормированный базис, в котором матрица линейного преобразования имеет блочно диагональный вид. Блоки первого порядка соответствуют вещественным собственным числам, а блоки второго порядка – комплексным числам. Так как собственные числа ортогонального преобразования по модулю равны 1, то по главной диагонали могут стоять либо 1, либо -1, либо блок второго порядка $\begin{pmatrix}\boldsymbol{\cos}\alpha & \boldsymbol{\sin}\alpha\\ -\boldsymbol{\sin}\alpha & \boldsymbol{\cos}\alpha\end{pmatrix}$. Для доказательства осталось заметить, что геометрический смысл указанных преобразований как раз и есть тождественные преобразования, симметрии и повороты в координатных плоскостях.

8.5 Самосопряженное преобразование.

Линейное преобразование называется самосопряженным, если $\varphi=\varphi^*$.

Свойство 8.7. Собственные числа самосопряженного преобразования – вещественны.

Доказательство. Пусть x –собственный вектор самосопряженного преобразования φ (т.е. $\varphi(x)=\alpha x$). Из равенств $\alpha|x|^2=(\varphi(x),x)=(x,\varphi(x))=\overline{\alpha}|x|^2$ выводим $\alpha=\overline{\alpha}$, то есть $\alpha\in R$.

Следствие 8.2. Для самосопряженного линейного преобразования евклидова пространства существует ортонормированный базис из собственных векторов.

Доказательство. Самосопряженное преобразование является нормальным, и значит, существует ортонормированный базис, в котором матрица линейного преобразования имеет блочно диагональный вид. Поскольку все собственные числа вещественные, то все блоки первого порядка.

8.6 Полярное разложение

Самосопряженное преобразование φ называется положительно определенным, если $(\varphi(x),x)\geq 0$.

Следствие 8.3. Все собственные числа положительно определенного самосопряженного линейного преобразования неотрицательны.

Доказательство. Пусть $\varphi(x)=\alpha x$, тогда $(\varphi(x),x)=\alpha|x|^2\geq 0$, и, значит, $\alpha\geq 0$.

Теорема 8.3. (извлечение корня) Для положительно определенного самосопряженного линейного преобразования φ существует единственное положительно определенное самосопряженное преобразование ψ, что $\varphi=\psi^2$.

Доказательство. Пусть $e_1,\ldots,e_n$ - ортонормированный базис линейного пространства, в котором матрица φ - диагональная. Пусть $[\varphi]_e=diag(\alpha_1,\cdots,\alpha_n)$. Все числа стоящие на главной диагонали неотрицательны. Положим $[\psi]_e=diag\left(\sqrt{\alpha_1},\ldots,\sqrt{\alpha_n}\right)$. Легко убедиться, что линейное преобразование ψ является положительно определенным самосопряженным преобразованием и $\varphi=\psi^2$. Единственность очевидна.

Теорема 8.4 (полярное разложение) Любое линейное преобразование φ можно представить в виде произведения самосопряженного положительно определенного линейного преобразования ψ и ортогонального преобразования χ. Если φ - невырожденное, то представление единственно. Разложение $\varphi=\psi\chi$ называется правым, а разложение $\varphi=\chi\psi$ - левым.

Доказательство. Преобразование $\varphi\varphi^*$ является самосопряженным и положительно определенным. Построим ортонормированный базис $e_1,\ldots,e_n$ преобразования $\varphi\varphi^*$, при этом расположим собственные векторы, соответствующие нулевому собственному значению в конце базиса. Пусть $e_1,\ldots,e_k$ - собственные векторы с не нулевыми собственными значениями, а $e_{k+1},\ldots,e_n$ - собственные векторы с нулевым собственным значением. Матрица $[\varphi\varphi^*]_e=diag(\alpha_1,\ldots,\alpha_k,0,\ldots,0)$ - диагональная, поэтому первые k строк матрицы $[\varphi]_e$ образуют ортогональную систему, а остальные равны 0. Длина j строки равна $\sqrt{\alpha_i}$. Обозначим через $a_1,\ldots,a_k$ первые k строк матрицы $[\varphi]_e$ и дополним ортонормированную систему векторов $b_1=1/\sqrt{\alpha_1}\,a_1,\ldots,b_k=1/\sqrt{\alpha_k}\,a_k$ векторами $b_{k+1},\ldots,b_n$ до ортонормированного базиса всего пространства. Обозначим через χ ортогональное преобразование, матрица которого в базисе $e_1,\ldots,e_n$ образована строками $b_1,\ldots,b_n$, а через ψ - положительно определенное самосопряженное преобразование, матрица которого в базисе $e_1,\ldots,e_n$ диагональная и равна $diag\left(\sqrt{\alpha_1},\ldots,\sqrt{\alpha_k},0,\ldots,0\right)$. Легко убедиться, что $\varphi=\psi\chi$.

Для построения левого разложения достаточно найти правое разложение для сопряженного преобразования.

Поскольку $\varphi\varphi^* = \psi^2$, то преобразование ψ определяется единственным образом. Если преобразование φ - невырожденное, то преобразование ψ невырожденное, и, значит, $\chi = \psi^{-1}\varphi$ определяется единственным образом.

9 Приведение квадратичных форм

9.1 Приведение квадратичных форм к главным осям.

Рассмотрим квадратичную форму $x^{\mathrm{T}}Ax$. Матрица *A* является симметричной. Линейное преобразование, заданное матрицей *A*, является самосопряженным и для этого преобразования существует ортонормированный базис из собственных векторов. Другими словами, найдется ортогональная матрица *T* ($T^{-1} = T^{\mathrm{T}}$), что $T^{-1}AT = diag(\alpha_1,\dots,\alpha_n)$, где $\alpha_1,\dots,\alpha_n$ - собственные числа *A*. Поскольку $T^{-1} = T^{\mathrm{T}}$, то квадратичная форма $x^{\mathrm{T}}Ax$ ортогональной заменой $y = Tx$ переходит в форму $\sum_{i=1}^{n}\alpha_i y_i^2$. Приведение квадратичной формы к каноническому виду ортогональным преобразованием называется приведением к главным осям. Полученный факт оформим в виде теоремы.

Теорема 9.1. Квадратичная форма $x^{\mathrm{T}}Ax$ при помощи ортогонального преобразования всегда может быть приведена к канонической форме $\sum_{i=1}^{n}\alpha_i y_i^2$, де $\alpha_1,\dots,\alpha_n$ - собственные числа *A*.

Отметим, что для квадратичной формы выполняется закон инерции. Следовательно, используя теорему Якоби, можно определить число положительных и число отрицательных собственных значений. Собственные значения матриц *A* и *A-tE* отличаются на *t*, поэтому, определяя число положительных и отрицательных собственных значений матрицы *A-tE*, мы, тем самым, определим количество собственных значений матрицы *A* меньших *t*. Выбирая различные *t* можно найти собственные числа с любой точностью.

9.2 Приведение пары квадратичных форм

Рассмотрим задачу выбора базиса в котором пара квадратичных форм имеют диагональный вид. Не все пары квадратичных форм можно одновременно привести к диагональному виду, например, формы $x^2 - y^2$ и *xy* привести нельзя.

9.2.1 Первый способ

Пусть даны квадратичные формы $x^{\mathrm{T}}Ax$ и $x^{\mathrm{T}}Bx$, причем квадратичная форма $x^{\mathrm{T}}Bx$ - положительно определена. Тогда введем скалярное произведение $(x, y) = x^{\mathrm{T}}By$ и найдем ортонормированный базис, а затем приведем первую квадратичную форму к главным осям. Поскольку ортогональное преобразование не меняет скалярное произведение, то обе квадратичные формы будут приведены к каноническому виду.

9.2.2 Пучок матриц

Пусть даны квадратичные формы $x^{\mathrm{T}}Ax$ и $x^{\mathrm{T}}Bx$. Рассмотрим пучок квадратичных форм $x^{\mathrm{T}}(\mu A - \nu B)x$. Если квадратичные формы $x^{\mathrm{T}}Ax$ и $x^{\mathrm{T}}Bx$ заменой координат *x=Py* приводятся к каноническому виду, то все формы из пучка $x^{\mathrm{T}}(\mu A - \nu B)x$ приводятся к каноническому виду этой же заменой координат. Пусть $P^{\mathrm{T}}AP = diag(\alpha_1,\dots,\alpha_n)$ и $P^{\mathrm{T}}BP = diag(\beta_1,\dots,\beta_n)$, тогда $P^{\mathrm{T}}(\mu A - \nu B)P = diag(\mu\alpha_1 - \nu\beta_1,\dots,\mu\alpha_n - \nu\beta_n)$. Из последнего равенства выводим $\boldsymbol{det}(\mu A - \nu B) = c\prod_{i=1}^{n}(\mu\alpha_i - \nu\beta_i)$, то есть многочлен $\boldsymbol{det}(\mu A - \nu B)$ раскладывается на линейные множители над полем вещественных чисел. Из равенства $(\mu A - \nu B)P = (P^{\mathrm{T}})^{-1}diag(\mu\alpha_1 - \nu\beta_1,\dots,\mu\alpha_n - \nu\beta_n)$ выводим, что *i*-ый столбец матрицы *P* удовлетворяет однородной системе уравнений $(\mu_i A - \nu_i B)x = 0$. Таким образом, получается следующий алгоритм приведения пары квадратичных форм к нормальному виду.

1. Раскладываем многочлен $\boldsymbol{det}(\mu A - \nu B)$ на линейные множители. Если разложения не существует, то искомой замены координат не существует.

2. Для каждого линейного множителя $\mu\alpha - \nu\beta$ многочлена $\boldsymbol{det}(\mu A - \nu B)$ находим базис подпространства $(\beta A + \alpha B)x = 0$. Если размерность подпространства меньше кратности множителя, то искомой замены координат не существует. В противном случае, будет построен базис, в котором квадратичные формы имеют нормальный вид.

Для обоснования этого подхода требуется показать, что объединение линейно независимых систем векторов, соответствующих разным линейным множителям, образует линейно независимую систему. Доказательство проводится также как и для собственных векторов.

9.3 Приведение квадрики ортогональным преобразованием. Ортогональные инварианты и полуинварианты.

Рассмотрим задачу упрощения уравнения квадрики с использованием ортогональным преобразованием системы координат. Отметим, что при ортогональной замене координат сохраняются метрические характеристики.

Опишем алгоритм приведения квадрики $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha$ к простейшему виду ортогональным преобразованием.

1. Приводим квадратичную форму $x^{\mathrm{T}}Ax$ к главным осям ортогональным преобразованием $y = Px$. В результате получим уравнение квадрики $\sum_{i=1}^{k}\beta_i y_i^2 + 2by + \alpha$, где $b = a^{\mathrm{T}}P$, k – ранг матрицы A, а β_i - ее ненулевые собственные числа.
2. Сдвигом начала координат $z_i = y_i + b_i/\beta_i$ при $i \le k$ и $z_i = y_i$ при $i>k$ приведем квадрику к виду $\sum_{i=1}^{k}\beta_i z_i^2 + 2\sum_{i=k+1}^{n} b_i y + \alpha'$, где $\alpha' = \alpha - \sum_{i=1}^{k} b_i^2/\beta_i$. Если $b_i = 0$ при $i>k$, то конец, а иначе перейдем на следующий шаг.
3. Положим $f = \left(1\Big/\sqrt{\sum_{i=k+1}^{n} b_i^2}\right)\sum_{i=k+1}^{n} b_i e_i$. Система векторов $e_1,\ldots,e_k, f$ - ортонормированная. Дополним ее до ортонормированного базиса всего пространства. Пусть T – матрица перехода к новому базису. Сделаем замену переменных $z = Tu + (\alpha/2|b|)e_{k+1}$. Очевидно, сделанная замена является ортогональной. В новой системе координат уравнение квадрики $\sum_{i=1}^{k}\beta_i u_i^2 + 2|b|u_{k+1} = 0$.

Оформим доказанное выше в виде теоремы.

Теорема 9.2. Ортогональным преобразованием, сдвигом начала координат и умножением на ненулевое число уравнение квадрики приводится к одному из следующих четырех видов $\sum_{i=1}^{k}\beta_i u_i^2 + 2u_{k+1} = 0$, $\sum_{i=1}^{k}\beta_i u_i^2 + 1 = 0$, $\sum_{i=1}^{k}\beta_i u_i^2 - 1 = 0$, $\sum_{i=1}^{k}\beta_i u_i^2 = 0$.

Обозначим через $i_k(A)$ сумму всех главных миноров k-го порядка матрицы A. Величина $i_k(A)$ является коэффициентом характеристического многочлена $|A - \lambda E|$ при $(-\lambda)^{n-k}$.

Пусть квадрика $x^{\mathrm{T}}Ax + 2a^{\mathrm{T}}x + \alpha$ ортогональным преобразованием $x=h+Ty$ приводится к виду $y^{\mathrm{T}}By + 2b^{\mathrm{T}}y + \beta$, где $B = T^{\mathrm{T}}AT$, $\hat{B} = \hat{T}^{\mathrm{T}}\hat{A}\hat{T}$, $\hat{T} = \left(\begin{array}{c|c} T & h \\ \hline 0 & 1 \end{array}\right)$. Поскольку T ортогональная матрица, то $B = T^{-1}AT$, и, значит, $i_k(A) = i_k(B)$, где $k=1,\ldots,n$. Кроме того, $|\hat{T}|^2 = |T|^2 = 1$, и, следовательно, $i_{n+1}(\hat{A}) = |\hat{A}| = |\hat{B}| = i_{n+1}(\hat{B})$. Тем самым установлен следующий факт.

Свойство 9.1 При ортогональном преобразовании не меняются следующие величины $i_k(A)$, где $k=1,\ldots,n$, и $i_{n+1}(\hat{A})$, которые называются ортогональными инвариантами квадрики.

К сожалению, ортогональные инварианты не всегда позволяют установить простейший тип квадрики.

Свойство 9.2. Пусть $k = rg\hat{A}$ и $k > rgA$, тогда $i_k(\hat{A})$ не меняется при ортогональном преобразовании.

Доказательство. При ортогональном преобразовании (без сдвига) величины $i_k\left(\hat{A}\right)$ не меняются. Пусть квадратичная форма $x^{\mathrm{T}}Ax$ приводится к главным осям ортогональной заменой координат $x = Py$. Пусть $x = Py + h$ - ортогональное преобразование квадрики. Поскольку

$\hat{T} = \left(\begin{array}{c|c} T & h \\ \hline 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} TP^{-1} & 0 \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} E & PT^{-1}h \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} P & 0 \\ \hline 0 & 1 \end{array}\right)$, то для доказательства утверждения достаточно рассмотреть случай, когда A - диагональная матрица и преобразование заключается в сдвиге на вектор h начала координат. Если $rgA = rg\hat{A} - 1$, то

$$\hat{A} = \left(\begin{array}{ccc:c|c} \multicolumn{3}{c:}{diag\left(\alpha_1,\ldots,\alpha_{k-1}\right)} & 0 & \begin{matrix} a_1 \\ \vdots \\ a_{k-1} \end{matrix} \\ \hdashline \multicolumn{3}{c:}{0} & 0 & 0 \\ \hline a_1 & \cdots & a_{k-1} & 0 & \alpha \end{array}\right)$$

. В этой матрице единственный минор k порядка, не содержащий нулевых строк, определитель которого не зависит от сдвига. Следовательно, утверждение в данном случае доказано. Пусть $rgA = rg\hat{A} - 2$, тогда

$$\hat{A} = \left(\begin{array}{ccc:cc|c} \multicolumn{3}{c:}{diag\left(\alpha_1,\ldots,\alpha_{k-2}\right)} & \multicolumn{2}{c|}{0} & \begin{matrix} a_1 \\ \vdots \\ a_{k-2} \end{matrix} \\ \hdashline \multicolumn{3}{c:}{0} & \multicolumn{2}{c|}{0} & \begin{matrix} a_{k-1} \\ 0 \end{matrix} \\ \hline a_1 & \cdots & a_{k-2} & a_{k-1} & 0 & \alpha \end{array}\right)$$

. В этой матрице единственный минор k порядка, не содержащий нулевых строк, определитель которого не зависит от сдвига. Следовательно, утверждение и в данном случае доказано.

Величины $i_k\left(\hat{A}\right)$ называются полуинвариантами ортогонального преобразования.

Набор инвариантов и полуинвариантов квадрики позволяет однозначно установить простейшее уравнение квадрики.

9.4 Ортогональная классификация кривых второго порядка

Теорема 9.3. Любая кривая второго порядка ортогонально эквивалентна одному из 9 классов кривых, приведенных в таблице. Приведенные кривые ортогонально не эквивалентны между собой.

Название кривой	Каноническое уравнение кривой
Эллипс	$\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = 1$
Мнимый эллипс	$\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = -1$

Гипербола	$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1$
Пара пересекающихся мнимых прямых	$x_1^2+\frac{x_2^2}{b^2}=0$
Пара пересекающихся прямых	$x_1^2-a^2x_2^2=0$, $x_1=\pm a x_2$
Парабола	$x_1^2=2ax_2$
Пара параллельных прямых	$x_1^2=a^2$
Пара параллельных мнимых прямых	$x_1^2=-a^2$
Пара совпавших параллельных прямых	$x_1^2=0$

Доказательство. очевидно

9.5 Ортогональная классификация поверхностей второго порядка.

Теорема 9.4 Любая поверхность второго порядка ортогонально эквивалентна одной из поверхностей в одном из 17 классов, приведенных в таблице. Приведенные поверхности ортогонально не эквивалентны между собой.

Название поверхности	Каноническое уравнение поверхности
Эллипсоид	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}+\frac{x_3^2}{c^2}=1$
Мнимый эллипсоид	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}+\frac{x_3^2}{c^2}=-1$
Однополостный гиперболоид	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}-\frac{x_3^2}{c^2}=1$

Двуполостный гиперболоид	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}-\frac{x_3^2}{c^2}=-1$
Мнимый конус	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}+\frac{x_3^2}{c^2}=0$
Конус	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}-\frac{x_3^2}{c^2}=0$
Эллиптический параболоид	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1$
Эллиптический цилиндр	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=-1$
Мнимый эллиптический цилиндр	$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1$
гиперболический цилиндр	$x_1^2+\frac{x_2^2}{b^2}=0$
Пара пересекающихся мнимых плоскостей	$x_1^2-a^2x_2^2=0$, $x_1=\pm ax_2$
Пара пересекающихся плоскостей	$x_1^2=2ax_2$
Параболический цилиндр	$x_1^2=a^2$
Пара параллельных плоскостей	$x_1^2=-a^2$
Пара параллельных мнимых	$x_1^2=0$

плоскостей	
Пара совпавших плоскостей	$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1$
Гиперболический параболоид (седло)	$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=2x_3$

Доказательство очевидно.

10 Аннулирующий многочлен

10.1 Аннулирующий многочлен вектора.

Рассмотрим наименьшее (по включению) инвариантное подпространство, содержащее вектор x. Очевидно, что с вектором x в нем содержится и векторы $\varphi^k(x)$, где k=1,2,... . Обозначим через k наибольшее число, при котором система векторов $x,\varphi(x),\dots,\varphi^k(x)$ линейно независима. Очевидно, что линейная оболочка этих векторов образует наименьшее инвариантное подпространство, содержащее вектор x. Выразим $\varphi^{k+1}(x)=\alpha_0 x+\alpha_1\varphi(x)+\dots+\alpha_k\varphi^k(x)$. Это равенство запишем в виде $\left(\varphi^{k+1}-\alpha_0\varepsilon-\alpha_1\varphi-\dots-\alpha_k\varphi^k\right)x=0$, где ε - тождественное преобразование. Слева стоит линейное преобразование, по виду являющееся многочленом от линейного преобразования φ. Будем говорить, что многочлен $p(t)$ аннулирует вектор x, если $p(\varphi)x=0$. Многочлен наименьшей степени, аннулирующий вектор x, называется минимальным аннулирующим многочленом вектора x.

Минимальный аннулирующий многочлен определен с точностью до числового множителя. Далее, для определенности будем считать коэффициент при старшей степени равным 1.

Свойство 10.1 Аннулирующий многочлен вектора делится без остатка на минимальный аннулирующий многочлен вектора.

Доказательство. Пусть $f(t)$ –аннулирующий многочлен, а $p(t)$ – минимальный аннулирующий многочлен. Разделим $f(t)$ на $p(t)$ с остатком $f(t)=p(t)g(t)+r(t)$. Тогда $r(\varphi)x=f(\varphi)x-g(\varphi)p(\varphi)x=0$. Так как степень $r(t)$ меньше степени $p(t)$, и многочлен $r(t)$ аннулирует вектор x, то единственная возможность $r(t)=0$.

Теорема 10.1 (Метод и академика Крылова). Пусть векторы $x,\varphi(x),\dots,\varphi^k(x)$ линейно независимы и $\varphi^{k+1}(x)=\sum_{i=0}^k\alpha_i\varphi^i(x)$, тогда многочлен $t^{k+1}-\sum_{i=0}^k\alpha_i t^i$ является минимальным аннулирующим многочленом вектора x.

Доказательство. Очевидно, что многочлен $t^{k+1}-\sum_{i=0}^k\alpha_i t^i$ является аннулирующим для вектора x. Допустим, он не является минимальным аннулирующим многочленом. Следовательно, найдется аннулирующий многочлен меньшей степени $\sum_{i=0}^s\beta_i t^i$, что $\sum_{i=0}^s\beta_i\varphi^i(x)=0$. Последнее равенство не возможно в силу линейной независимости системы векторов $x,\varphi(x),\dots,\varphi^k(x)$.

Теорема 10.2 Минимальный аннулирующий многочлен вектора является делителем характеристического многочлена.

Доказательство. Пусть система векторов $x,\varphi(x),\dots,\varphi^k(x)$ - линейно независима и $\varphi^{k+1}(x)=\sum_{i=0}^k\alpha_i\varphi^i(x)$. Дополним систему векторов $x,\varphi(x),\dots,\varphi^k(x)$ до базиса всего пространства и найдем матрицу линейного преобразования φ в этом базисе. Эта матрица имеет блочный вид

$$[\varphi]=\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}, \text{ где } A=\begin{pmatrix} 0 & \cdots & 0 & \alpha_0 \\ 1 & \cdots & 0 & \alpha_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & \alpha_k \end{pmatrix}$$ - блок порядка $k+1$. По теореме Лапласа

$\left|[\varphi]-\lambda E\right| = |A-\lambda E|\cdot|C-\lambda E|$, а $|A-\lambda E| = (-1)^{k+1}\left(\lambda^{k+1}-\sum_{i=0}^{k}\alpha_i\lambda^i\right)$ - минимальный аннулирующий многочлен вектора x.

Следствие 10.1. (теорема Гамильтона – Кэли) Линейное преобразование является корнем своего характеристического многочлена.

Доказательство. Пусть $\chi(\lambda)$ - характеристический многочлен. Тогда для любого x имеем $\chi(\varphi)(x)=0$, и, следовательно, $\chi(\varphi)=0$.

10.2 Аннулирующий многочлен подпространства

Будем говорить, что многочлен $p(t)$ аннулирует подпространство W, если он аннулирует каждый вектор из W. Аннулирующий многочлен подпространства W наименьшей степени называется минимальным аннулирующим многочленом подпространства W. Как и минимальный аннулирующий многочлен вектора, минимальный аннулирующий многочлен подпространства определен с точностью до множителя. Для определенности, будем считать старший коэффициент минимального аннулирующего многочлена подпространства равным 1.

Свойство 10.2. Аннулирующий многочлен подпространства делится без остатка на минимальный аннулирующий многочлен этого же подпространства.

Доказательство. Пусть $f(t)$ –аннулирующий многочлен, а $p(t)$ – минимальный аннулирующий многочлен. Разделим $f(t)$ на $p(t)$ с остатком $f(t)=p(t)g(t)+r(t)$. Тогда для вектора x из W справедливо равенство $r(\varphi)x = f(\varphi)x - g(\varphi)p(\varphi)x = 0$. Так как степень $r(t)$ меньше степени $p(t)$, и многочлен $r(t)$ аннулирует любой вектор x из W, то единственная возможность $r(t)=0$.

Теорема 10.3. Минимальный аннулирующий многочлен подпространства равен наименьшему общему кратному минимальных аннулирующих базисных векторов.

Доказательство. Пусть $e_1,\ldots,e_k$ - базис подпространства W, h - минимальный аннулирующий многочлен подпространства W,а $f_i(t)$ - минимальный аннулирующий многочлен вектора e_i, где $i=1,\ldots,k$. Многочлены $f_i(t)$ являются делителями $h(t)$ (**Свойство 10.1**). С другой стороны, наименьшее общее кратное этих многочленов аннулирует все базисные векторы, а значит и любой вектор из W.

Следствие 10.2. Минимальный аннулирующий многочлен подпространства является делителем характеристического многочлена.

Доказательство. Пусть $e_1,\ldots,e_k$ - базис подпространства W, а $f_i(t)$ - минимальный аннулирующий многочлен вектора e_i, где $i=1,\ldots,k$. Многочлены $f_i(t)$ являются делителями характеристического многочлена (**Теорема 10.2**), следовательно, характеристический многочлен делится и на их наименьшее общее кратное, равное минимальному аннулирующему многочлену подпространства.

Если в качестве подпространства взять все пространство, то минимальный аннулирующий многочлен подпространства называется *минимальным аннулирующим многочленом.*

Следствие 10.3. Минимальный аннулирующий многочлен является делителем характеристического многочлена и имеет то же самое множество корней.

Доказательство очевидно.

10.3 Функции от матриц

Пусть $f(t)$ некоторый многочлен, и требуется вычислить значение матрицы A от этого многочлена. В арифметическом пространстве матрица A задает линейное преобразование. Обозначим через $g(t)$ минимальный аннулирующий многочлен этого преобразования. Разделим многочлен $f(t)$ на $g(t)$ с остатком $f(t)=h(t)g(t)+r(t)$. При подстановке матрицы A получим равенство $f(A)=h(A)g(A)+r(A)=r(A)$. Таким образом, вычисление значения многочлена от матрицы сводится к вычислению значению его остатка. Остаток от

деления $r(t)$ можно вычислить как интерполяционный многочлен Лагранжа - Сильвестра от корней минимального многочлена.
Ничего не изменится в проведенных рассуждениях, если вместо многочлена $f(t)$ использовать произвольную функцию, значения которой, а также значения ее производных соответствующих порядков, определены на множестве корней минимального многочлена.
В некоторых случаях в качестве минимального многочлена берут характеристический многочлен.

10.4 Вычисление линейных рекуррентных последовательностей

Последовательность $\{a_n\}$ называется линейной рекуррентной, если существуют такие коэффициенты $\alpha_1,\cdots,\alpha_k$, что для любого n справедливо равенство $a_{n+k}=\alpha_1 a_n+\cdots+\alpha_k a_{n+k-1}$. Для задания линейной рекуррентной последовательности, кроме ее коэффициентов, необходимо знать первые k членов $a_1,\cdots,a_k$, которые называются начальными условиями. Рассмотрим задачу выражения n-го члена последовательности через его номер и начальные условия.
Обозначим через $\bar{\mathbf{a}}_n$ вектор столбец, состоящий из k компонент $(a_{n+k-1},\ldots,a_n)$, через A — матрицу размерами $k\times k$ вида $\begin{pmatrix}\alpha_k & \alpha_{k-1} & \cdots & \alpha_2 & \alpha_1\\ 1 & 0 & \cdots & 0 & 0\\ 0 & 1 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 0\end{pmatrix}$. По правилу перемножения матриц имеем:
$A\bar{\mathbf{a}}_{\mathbf{n-1}}=\bar{\mathbf{a}}_n$. Многократным применением полученной формулы выводим $A^{n-1}\bar{\mathbf{a}}_1=\bar{\mathbf{a}}_n$. Задача вычисления n-го члена последовательности свелась, тем самым, к вычислению матрицы A^{n-1}.
Характеристический многочлен $\chi(\lambda)$ матрицы A равен $\chi(\lambda)=(-1)^k(\lambda^k-\alpha_k\lambda^{k-1}-\cdots-\alpha_1)$. Разделим многочлен λ^{n-1} на $\chi(\lambda)$ с остатком. Пусть $\lambda^{n-1}=\chi(\lambda)g(\lambda)+r(\lambda)$, где $r(\lambda)$ - остаток от деления.
Подставив вместо λ матрицу A, получим $A^{n-1}=\chi(A)g(A)+r(A)$. По теореме Гамильтона-Кэли каждая матрица является корнем своего характеристического уравнения, то есть $\chi(A)=0$, где 0 - нулевая матрица. Таким образом, $A^{n-1}=r(A)$, и задача вычисления A^{n-1} свелась к вычислению многочлена $r(\lambda)$.
Разложим многочлен $\chi(\lambda)$ на линейные множители $\chi(\lambda)=(-1)^k\prod_{i=1}^{s}(\lambda-\beta_i)^{t_i}$, где $t_1+\cdots+t_s=k$.
Для каждого неотрицательного j строго меньшего t_i справедливо равенство $\chi(\beta_i)^{(j)}=0$, где $\chi(\lambda)^{(j)}$ - j-ая производная характеристического многочлена. Продифференцировав j раз равенство $\lambda^{n-1}=\chi(\lambda)g(\lambda)+r(\lambda)$ и, подставив в него β_i, получим $(n-1)!/(n-1-j)!\,\beta_i^{n-1-j}=r(\beta_i)^{(j)}$. Этими условиями многочлен $r(\lambda)$ степени k-1 определяется однозначно. В литературе задача вычисления многочлена по таким условиям носит название «интерполяционный многочлен Лагранжа-Сильвестра».
В качестве примера вычислим n-ый член линейной рекуррентной последовательности $a_{n+2}=4a_{n+1}-4a_n$, где $a_1=a_2=1$. Положим $A=\begin{pmatrix}4 & -4\\ 1 & 0\end{pmatrix}$. Характеристический многочлен равен $\chi(\lambda)=(\lambda-2)^2$.
Остаток от деления λ^{n-1} на $(\lambda-2)^2$ удовлетворяет соотношениям $r(2)=2^{n-1}$ и $r'(2)=(n-1)2^{n-2}$.
Единственный многочлен первой степени, удовлетворяющий этим условиям, равен $r(\lambda)=2^{n-1}+(\lambda-2)(n-1)2^{n-2}$. Таким образом,
$A^{n-1}=2^{n-1}E+(n-1)2^{n-2}(A-2E)=\begin{pmatrix}n2^{n-1} & (1-n)2^n\\ (n-1)2^{n-2} & (2-n)2^{n-1}\end{pmatrix}$ и $a_n=(3-n)2^{n-2}$.

10.5 Корневые подпространства. Расщепление пространства в прямую сумму корневых подпространств.

Пусть минимальный многочлен $h(\lambda)$ линейного преобразования φ раскладывается в произведение взаимно простых множителей.