

Содержание

1	Линейные пространства. Пространства со скалярным произведением.	3
1.1	Неравенство Коши-Бунковского-Шварца.	3
1.2	Процесс ортогонализации Грама-Шмидта и QR-разложение матрицы. . .	4
1.3	Матрица Грама и критерий линейной зависимости.	5
1.4	Общий вид скалярного произведения в конечномерном пространстве. . .	6
1.5	Задача о наилучшем приближении вектора на конечномерном подпространстве в пространстве со скалярным произведением.	7
2	Линейные операторы.	8
2.1	Матрица линейного оператора в паре базисов. Изменение матрицы оператора при изменении пары базисов.	8
2.2	Эквивалентность матриц, подобие матриц и инварианты подобия.	9
2.3	Ядро и образ линейного оператора. Соотношение между рангом и дефектом линейного оператора.	11
2.4	Обратимый оператор. Критерий обратимости. Линейность обратного оператора.	12
2.5	Оператор проектирования.	12
2.6	Собственные значения и собственные векторы. Характеристический многочлен линейного оператора (матрицы).	13
2.7	Геометрическая и алгебраическая кратность собственного значения. . . .	14
2.8	Операторы простой структуры и диагонализуемые матрицы. Критерий диагонализуемости.	15
2.9	Верхняя треугольная форма матрицы линейного оператора в комплексном пространстве.	15
2.10	Многочлен от линейного оператора (матрицы). Теорема Гамильтона-Кэли. .	17
2.11	Нильпотентные и квазискалярные операторы (матрицы). Критерий нильпотентности.	17
2.12	Прямая сумма линейных операторов (матриц). Теорема о расщеплении вырожденного оператора.	18
2.13	Корневое расщепление линейного оператора.	19
2.14	Нерасщепляемые операторы и подпространства Крылова.	20
2.15	Условие линейной независимости составной системы векторов Крылова нильпотентного оператора.	21
2.16	Максимальное расщепление и жорданова форма нильпотентного оператора. .	22
2.17	Теорема Жордана о структуре линейного оператора.	24
2.18	Единственность жордановой формы линейного оператора (матрицы). . .	25
2.19	Критерий подобия комплексных матриц.	26
2.20	Блочно-диагональная жорданова форма вещественной матрицы.	26
2.21	Минимальный многочлен матрицы (оператора).	28
2.22	Условие совпадения минимального и характеристического многочленов. .	29

3	Линейные операторы в пространствах со скалярным произведением.	29
3.1	Существование, линейность, единственность сопряжённого оператора. . .	29
3.2	Матрицы оператора и сопряжённого к нему в паре биортогональных базисов.	31
3.3	Критерии нормальности оператора (матрицы).	31
3.4	Критерии унитарности и эрмитовости оператора (матрицы).	32
3.5	Эрмитово разложение оператора (матрицы) и эрмитовость знакоопределённого оператора в унитарном пространстве.	33
3.6	Существование и единственность неотрицательно определённого квадратного корня из неотрицательно определённой матрицы.	34
3.7	Блочно-диагональная форма вещественной нормальной матрицы.	35
3.8	Блочно-диагональная форма ортогональной матрицы. Матрицы вращения и отражения.	35
3.9	Полярное разложение оператора (матрицы).	36
3.10	Сингулярное разложение линейного оператора (матрицы). Три формы записи сингулярного разложения.	37
3.11	Вариационные свойства собственных значений эрмитовых матриц, теорема Куранта-Фишера. Вариационные свойства сингулярных чисел.	38
3.12	Соотношения разделения собственных значений эрмитовой матрицы и её главной подматрицы. Соотношения разделения для сингулярных чисел. .	39
4	Билинейные и квадратичные формы.	39
4.1	Закон сохранения инерции эрмитовой матрицы при переходе к эрмитово конгруэнтной матрице.	39
4.2	Квадратичные и эрмитово квадратичные формы и их приведение к алгебраической сумме квадратов.	40
4.3	Общий вид линейного функционала и эрмитовой билинейной формы в конечномерном пространстве со скалярным произведением.	41

1 Линейные пространства. Пространства со скалярным произведением.

1.1 Неравенство Коши-Буняковского-Шварца.

Опр. Пусть V - вещественное линейное пространство, на котором каждой упорядоченной паре векторов $x, y \in V$ поставлено в соответствие вещественное число (x, y) таким образом, что:

- $(x, x) \geq 0 \forall x \in V; (x, x) = 0 \Leftrightarrow x = 0;$
- $(x, y) = (y, x) \forall x, y \in V;$
- $(x + y, z) = (x, z) + (y, z) \forall x, y, z \in V;$
- $(\alpha x, y) = \alpha(x, y) \forall \alpha \in \mathbb{R} \forall x, y \in V.$

Число (x, y) называется скалярным произведением векторов x, y . Вещественное линейное пространство со скалярным произведением называется евклидовым.

Опр. Пусть V - комплексное линейное пространство, на котором каждой упорядоченной паре векторов $x, y \in V$ поставлено в соответствие комплексное число (x, y) таким образом, что:

- $(x, x) \geq 0 \forall x \in V; (x, x) = 0 \Leftrightarrow x = 0;$
- $(x, y) = \overline{(y, x)} \forall x, y \in V;$
- $(x + y, z) = (x, z) + (y, z) \forall x, y, z \in V;$
- $(\alpha x, y) = \alpha(x, y) \forall \alpha \in \mathbb{C} \forall x, y \in V.$

Число (x, y) называется скалярным произведением векторов x, y . Комплексное линейное пространство со скалярным произведением называется унитарным.

Опр. В произвольном евклидовом или унитарном пространстве величина $|x| := \sqrt{(x, x)}$ называется длиной вектора.

Теорема (Неравенство Коши-Буняковского-Шварца). Скалярное произведение векторов и их длины связано неравенством $|(x, y)| \leq |x||y|$. Равенство достигается в том и только в том случае, когда векторы x и y линейно зависимы.

Д-во. Случай $(x, y) = 0$ очевиден. В противном случае запишем $(x, y) = |(x, y)|\xi$, где $\xi = e^{i\phi}$, и рассмотрим функцию вещественного аргумента $F(t) = (x + t\xi y, x + t\xi y) = (x, x) + t\xi(x, y) + t\bar{\xi}(x, y) + t^2\xi\bar{\xi}(y, y) = t^2|y|^2 + 2t|(x, y)| + |x|^2$. В силу свойств скалярного произведения $F(t) \geq 0$ при всех вещественных t . Значит $D \leq 0$, $D = |(x, y)|^2 - |x|^2|y|^2 \leq 0 \Rightarrow |(x, y)| \leq |x||y|$. Равенство означает, что $D = 0 \Rightarrow (x + t\xi y, x + t\xi y) = 0 \Rightarrow x + t\xi y = 0$. \square

1.2 Процесс ортогонализации Грама-Шмидта и QR-разложение матрицы.

Теорема. Для любой линейно независимой системы векторов a_1, \dots, a_m существует ортогональная система p_1, \dots, p_m такая, что $L(p_1, \dots, p_k) = L(a_1, \dots, a_k)$, $1 \leq k \leq m$.

Д-во. Положим, что $p_1 = a_1 \implies L(p_1) = L(a_1)$. Предположим, что уже построена ортогональная система p_1, \dots, p_{k-1} такая, что $L(p_1, \dots, p_i) = L(a_1, \dots, a_i)$ при $1 \leq i \leq k-1$. Тогда вектор

$$p_k = a_k - \sum_{i=1}^{k-1} \frac{(a_k, p_i)}{(p_i, p_i)} p_i.$$

будет ортогонален каждому из векторов p_1, \dots, p_{k-1} :

$$(p_k, p_j) = (a_k, p_j) - \left(\sum_{i=1}^{k-1} \frac{(a_k, p_i)}{(p_i, p_i)} p_i, p_j \right) = (a_k, p_j) - \frac{(a_k, p_j)}{(p_j, p_j)} (p_j, p_j) = 0.$$

Кроме того, $p_k \in L(p_1, \dots, p_{k-1}, a_k) = L(a_1, \dots, a_{k-1}, a_k)$ и $a_k \in L(p_1, \dots, p_{k-1}, p_k) \implies L(p_1, \dots, p_{k-1}, p_k) = L(a_1, \dots, a_{k-1}, a_k)$. \square

Теорема об ортогонализации содержит, по существу, следующий алгоритм построения ортонормированной системы q_1, \dots, q_m в линейной оболочке заданной линейно независимой системы a_1, \dots, a_m :

$$p_k := a_k - \sum_{i=1}^{k-1} (a_k, q_i) q_i, \quad q_k := \frac{p_k}{|p_k|}, \quad k = 1, 2, \dots, m.$$

Этот алгоритм называется процессом ортогонализации Грама-Шмидта.

Пусть матрица A имеет линейно независимые столбцы a_1, \dots, a_m , а процесс ортогонализации ее столбцов относительно естественного скалярного произведения дает ортонормированные столбцы q_1, \dots, q_m . Процесс ортогонализации устроен таким образом, что a_k есть линейная комбинация столбцов q_1, \dots, q_k :

$$a_k = \sum_{i=1}^k r_{ik} q_i \Leftrightarrow A = QR, \quad Q = [q_1, \dots, q_m], \quad R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ & r_{22} & \dots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}.$$

Опр. Разложение $A = QR$, где Q имеет ортонормированные столбцы, а R - верхняя треугольная матрица, называется QR-разложением матрицы A . Таким образом, для любой прямоугольной матрицы с линейно независимыми столбцами существует QR-разложение.

Теорема (Теорема о QR-разложении). Любая квадратная комплексная матрица представима в виде произведения унитарной и верхней треугольной матрицы.

Д-во. Любая квадратная матрица A является пределом последовательности невырожденных матриц $A_k = A - \alpha_k I$, так как заведомо имеется последовательность чисел $\alpha_k \rightarrow 0$, отличных от собственных значений матрицы A . Для каждой невырожденной матрицы A_k , как мы уже знаем, существует QR -разложение: $A_k = Q_k R_k$. Последовательность Q_k принадлежит компактному множеству матриц, поэтому из нее можно выделить сходящуюся подпоследовательность $Q_{k_l} \rightarrow Q$. Матрица Q будет, конечно, унитарной, а предел последовательности $R_{k_l} = Q_{k_l}^* A_{k_l} \rightarrow Q^* A$ является, очевидно, верхней треугольной матрицей. \square

1.3 Матрица Грама и критерий линейной зависимости.

Теорема (теорема о перпендикуляре). *Для любого вектора x в произвольном пространстве со скалярным произведением и любого конечномерного подпространства $L \subset V$ существуют и единственны перпендикуляр h и проекция z такие, что*

$$x = z + h, \quad z \in L, \quad h \perp L, \quad |x - z| = |h| \leq |x - y| \quad \forall y \in L.$$

Д-во. Если $x \in L$, то полагаем $z = x$ и $h = 0$. Пусть v_1, \dots, v_k - базис подпространства L . В случае $x \notin L$ система v_1, \dots, v_k, x будет линейно независимой. Применив к ней процесс ортогонализации Грама-Шмидта, получим ортонормированную системы q_1, \dots, q_k, q_{k+1} такую, что $L = L(q_1, \dots, q_k)$ и $x \in L(q_1, \dots, q_k, q_{k+1})$, а искомые проекция и перпендикуляр получаются из разложения $x = \alpha_1 q_1 + \dots + \alpha_k q_k + \alpha_{k+1} q_{k+1}$ очевидным образом: $z = \alpha_1 q_1 + \dots + \alpha_k q_k$, $h = \alpha_{k+1} q_{k+1}$.

Единственность: если $x = z + h = z' + h'$, где $z, z' \in L$ и $h, h' \perp L$, то $c := z - z' = h' - h \in L \cap L^\perp \implies c = 0$.

Наконец, для любого $y \in L$ находим $x - y = (z - y) + h$, и, согласно теореме Пифагора, $|x - y|^2 = |z - y|^2 + |h|^2 \geq |h|^2$. Равенство, очевидно, имеет место в том и только в том случае, когда $y = z$. \square

Если v_1, \dots, v_k - произвольный базис подпространства L , то ортогональная проекция $z = x_1 v_1 + \dots + x_k v_k$ вектора x на L однозначно определяется уравнением $x - z \perp L$. Для этого необходимо и достаточно, чтобы вектор $x - z$ был ортогонален каждому из векторов v_1, \dots, v_k :

$$\begin{cases} (v_1, v_1)x_1 + \dots + (v_k, v_1)x_k = (x, v_1) \Leftrightarrow (x - z, v_1) = 0 \\ (v_1, v_2)x_1 + \dots + (v_k, v_2)x_k = (x, v_2) \Leftrightarrow (x - z, v_2) = 0 \\ \dots \\ (v_1, v_k)x_1 + \dots + (v_k, v_k)x_k = (x, v_k) \Leftrightarrow (x - z, v_k) = 0 \end{cases}$$

Из теоремы о перпендикуляре следует, что эта система линейных алгебраических уравнений имеет и притом единственное решение, определяющее коэффициенты x_1, \dots, x_k .

Опр. Матрицы $A = [a_{ij}]$ полученной нами системы линейных алгебраических уравнений имеет элементы $a_{ij} = (v_i, v_j)$. Матрица такого вида называется матрицей Грама системы векторов v_1, \dots, v_k .

Теорема. Для линейно независимой системы матрица Грама невырождена.

Д-во. Сразу следует из теоремы о перпендикуляре, так как система должна иметь единственное решение. \square

Теорема. Матрица Грама неотрицательно определена для любой системы векторов и положительно определена в том и только в том случае, когда система линейно независима.

Д-во. Пусть A - матрица Грама системы v_1, \dots, v_k и x - вектор столбец с элементами x_1, \dots, x_k . Тогда $x^*Ax = \sum_{i,j=1}^k \bar{x}_i a_{ij} x_j = \sum_{i,j=1}^k \bar{x}_i (v_i, v_j) x_j = \sum_{i=1}^k \bar{x}_i \left(v_i, \sum_{j=1}^k \bar{x}_j v_j \right) = \sum_{i=1}^k \bar{x}_i (v_i, v) = \left(\sum_{i=1}^k \bar{x}_i v_i, v \right) = (v, v) \geq 0, v = \bar{x}_1 v_1 + \dots + \bar{x}_k v_k.$ \square

1.4 Общий вид скалярного произведения в конечномерном пространстве.

Теорема. Пусть V - вещественное скалярное или комплексное пространство размерности n и e_1, \dots, e_n - произвольный фиксированный базис V . Тогда для произвольной положительно определенной матрицы A порядка n формула

$$(x, y) = [y]_e^* A [x]_e = [\bar{y}_1, \dots, \bar{y}_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \text{ где } x = \sum_{j=1}^n x_j e_j, y = \sum_{j=1}^n y_j e_j,$$

задает некоторое скалярное произведение на V и для произвольного скалярного произведения является тождеством, в котором A является матрица Грама базиса e_1, \dots, e_n .

Д-во. Пусть A — эрмитова положительно определенная матрица и $f(u, v) = v^* A u$ — функция от векторов-столбцов $u, v \in \mathbb{C}^n$. Проверка свойств скалярного произведения для данной функции выполняется непосредственно: линейность по первому аргументу очевидна, а положительная определенность и симметричность вытекает их положительной определенности и эрмитовости матрицы.

В тоже время, произвольное скалярное произведение векторов $x = \sum_{i=1}^n x_i e_i$ и $y = \sum_{i=1}^n y_i e_i$ имеет вид

$$(x, y) = \left(\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i \right) = \sum_{i,j=1}^n \bar{y}_i (e_j, e_i) x_j = [\bar{y}_1 \dots \bar{y}_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A - матрица с элементами $a_{ij} = (e_j, e_i)$. \square

1.5 Задача о наилучшем приближении вектора на конечномерном подпространстве в пространстве со скалярным произведением.

Опр. Пусть V - нормированное пространство и M - непустое подмножество векторов из V . Вектор $z \in M$ называется элементом наилучшего приближения вектора $x \in V$ на множестве M , если $\|x - z\| \leq \|x - y\| \forall y \in M$.

Теорема. Для любого $x \in V$ и любого конечномерного подпространства $M \in V$ существует единственное наилучшее приближение.

Д-во. Если M состоит из одного вектора, то он и является наилучшим приближением. Далее полагаем, что в M больше одного вектора. Пусть $y, z \in M$. Представим z в виде $z = y + h$, $h \in M$. Тогда

$$(x - z, x - z) = (x - y - h, x - y - h) = (x - y, x - y) - (x - y, h) - (h, x - y) + (h, h)$$

$$\|x - z\|^2 = \|x - y\|^2 - (x - y, h) - (h, x - y) + \|h\|^2.$$

Если $(x - y, h) = 0 \forall h \in M$, то $\|x - y\| \leq \|x - z\| \forall z \in M$.

Если $\|x - y\| \leq \|x - z\| \forall z \in M$, то $-(x - y, h) - (h, x - y) + (h, h) \geq 0 \forall h \in M$. Заменим что вектор h на $h_1 = \frac{(x - y, h)}{\|h\|^2} h$. Получим

$$\begin{aligned} & - \left(x - y, \frac{(x - y, h)}{\|h\|^2} h \right) - \left(\frac{(x - y, h)}{\|h\|^2} h, x - y \right) + \left(\frac{(x - y, h)}{\|h\|^2} h, \frac{(x - y, h)}{\|h\|^2} h \right) = \\ & = - \frac{\overline{(x - y, h)}}{\|h\|^2} (x - y, h) - \frac{(x - y, h)}{\|h\|^2} \overline{(x - y, h)} + \frac{|(x - y, h)|^2}{\|h\|^4} (h, h) = \\ & = -2 \frac{|(x - y, h)|^2}{\|h\|^2} + \frac{|(x - y, h)|^2}{\|h\|^2} = - \frac{|(x - y, h)|^2}{\|h\|^2} \geq 0 \end{aligned}$$

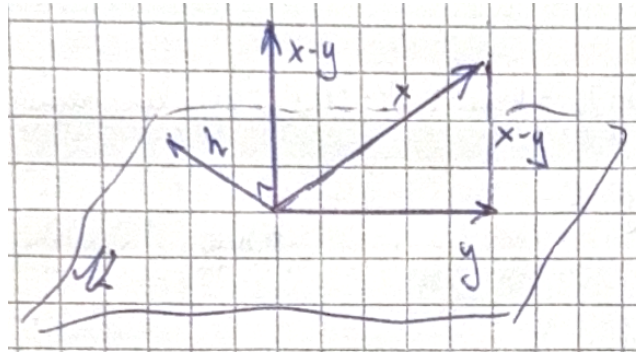
Полученное неравенство верно только при $(x - y, h) = 0$.

Итак, чтобы вектор $y \in M$ был наилучшим приближением к вектору $x \in V$ необходимо и достаточно, чтобы $(x - y, h) = 0 \forall h \in M$ (вектор $x - y$ должен быть ортогонален подпространству M).

Докажем, что вектор y , удовлетворяющий условию $(x - y, h) = 0 \forall h \in M$ однозначно определяется вектором x .

Пусть $(x - y, h) = 0 \forall h \in M$ и существует вектор еще один вектор $\tilde{y} \in M$ такой, что

$(x - \tilde{y}, h) = 0 \forall h \in M$. Тогда $(y - \tilde{y}, h) = 0 \forall h \in M$. Пологая $h = y - \tilde{y}$, получим, что $(y - \tilde{y}, y - \tilde{y}) = 0 \implies y = \tilde{y}$.



Докажем теперь, что существует вектор $y \in M$, удовлетворяющий условию $(x - y, h) = 0 \forall h \in M$.

Пусть e_1, \dots, e_m - базис M . Условие $(x - y, h) = 0 \forall h \in M$ эквивалентно тому, что $(x - y, e_k) = 0, k = \overline{1, m}$. Будем искать y в виде разложения по базису: $y = \sum_{i=1}^m y_i e_i$. Тогда

$$\left(\sum_{i=1}^m y_i e_i, e_k \right) = (x, e_k), \quad k = \overline{1, m}.$$

$$\sum_{i=1}^m y_i (e_i, e_k) = (x, e_k), \quad k = \overline{1, m}.$$

— СЛАУ относительно y_1, \dots, y_m , в которой матрица коэффициентов A — матрица Грама векторов e_1, \dots, e_m . A невырождена \implies система имеет единственное решение. \square

2 Линейные операторы.

2.1 Матрица линейного оператора в паре базисов. Изменение матрицы оператора при изменении пары базисов.

Пусть $e = (e_1, \dots, e_n)$ и $f = (f_1, \dots, f_n)$ - базисы пространств V и W . Линейный оператор $A \in L(V, W)$ однозначно определяется заданием векторов Ae_1, \dots, Ae_n . В свою очередь векторы $Ae_i, i = 1, \dots, n$, однозначно определяются своими координатами в базисе f , т.е. коэффициентами разложений

$$\begin{cases} Ae_1 = a_{11}f_1 + a_{21}f_2 + \dots + a_{m1}f_m, \\ Ae_2 = a_{12}f_1 + a_{22}f_2 + \dots + a_{m2}f_m, \\ \dots \\ Ae_n = a_{1n}f_1 + a_{2n}f_2 + \dots + a_{mn}f_m. \end{cases}$$

Опр. Матрица

$$A_{fe} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

называется матрицей оператора A в паре базисов e и f .

Пусть e и $t = C_{et}^{-1}e$ - два базиса пространства V с матрицей перехода C_{et} , а f и $s = D_{fs}^{-1}f$ - два базиса пространства W с матрицей перехода D_{fs} . Одному и тому же линейному оператору $A \in L(V, W)$ в паре базисов e и f соответствует матрица A_{ef} , а в паре базисов t и s - матрица A_{st} .

Теорема. Матрицы A_{fe} и A_{st} линейного оператора в различных парах базисов связаны соотношением

$$A_{st} = D_{fs}^{-1} A_{fe} C_{et}.$$

Д-во. Для произвольного вектора $x \in V$ и его образа $y = Ax$ имеем

$$y_f = A_{fe} x_e, \quad y_s = A_{st} x_t.$$

В свою очередь, $x_e = C_{et} x_t$, $y_f = D_{fs} y_s$. Подставив эти соотношения, получим, что $D_{fs} y_s = A_{fe} C_{et} x_t$ или $D_{fs} A_{st} x_t = A_{fe} C_{et} x_t$. Так как это соотношение имеет место при любых x_t , то $D_{fs} A_{st} = A_{fe} C_{et}$. В силу невырожденности матрицы перехода получаем, что $A_{st} = D_{fs}^{-1} A_{fe} C_{et}$. \square

2.2 Эквивалентность матриц, подобие матриц и инварианты подобия.

Опр. Две матрицы $A, B \in \mathbb{P}^{m \times n}$ называются эквивалентными, если существуют невырожденные матрицы P и Q такие, что $A = PBQ$.

Утверждение. Эквивалентность матриц является соотношением эквивалентности.

Д-во. (рефлексивность) $A \sim A$, т.к. $A = IAI$.

(симметричность) $A \sim B \implies \exists P, Q$, т.ч. $|P| \neq 0$, $|Q| \neq 0$, $A = PBQ \implies B = P^{-1}AQ^{-1} \implies B \sim A$.

(транзитивность) $A \sim B, B \sim C \implies \exists$ невырожденные P_1, P_2, Q_1, Q_2 , т.ч. $A = P_1 B Q_1$, $B = P_2 C Q_2 \implies A = (P_1 P_2) B (Q_1 Q_2) \implies A \sim C$. \square

Теорема. Две матрицы A и B над полем \mathbb{P} одинакового размера эквивалентны тогда и только тогда, когда они являются матрицами одного линейного оператора $A \in L(V, W)$, где V и W - линейные пространства над полем \mathbb{P} размерностей n и m соответственно.

Д-во. (\implies) Пусть $A, B \in \mathbb{P}^{m \times n}$ и $B = D^{-1}AC$. Рассмотрим любые линейные пространства V и W над полем \mathbb{P} такие, что $\dim V = n$, $\dim W = m$. Возьмем в пространстве V произвольный базис e , а в пространстве W - базис f . В силу взаимной однозначности соответствия между $\mathbb{P}^{m \times n}$ и $L(V, W)$ существует единственный оператор $A \in L(V, W)$, который в паре базисов e и f имеет матрицу A . Тогда матрица B будет матрицей этого же оператора в паре базисов $t = Ce$ и $s = Df$.

(\impliedby) Пусть A и B - матрицы линейного оператора $A \in L(V, W)$ в парах базисов e, f и t, s соответственно. Причем $t = C^{-1}e$, $s = D^{-1}f$. Тогда $B = D^{-1}AC \implies$ матрицы A и B эквивалентны. \square

Теорема. Любая невырожденная матрица $A \in \mathbb{F}^{m \times n}$ ранга r эквивалентна матрице $I_r \in \mathbb{F}^{m \times n}$ вида

$$I_r = \left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & \\ 0 & 1 & \dots & 0 & \emptyset \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & 1 & \\ \hline & \emptyset & & & \emptyset \end{array} \right].$$

Д-во. Любую матрицу можно привести к диагональному виду элементарными преобразованиями. Если привести матрицу A к диагональному виду, а затем поделить каждую ненулевую строку на ненулевой элемент в ней, то получится матрица вида I_r . Это означает, что существу, матрицы элементарных преобразований Q_1, \dots, Q_k и P_1, \dots, P_s , такие, что $I_r = Q_1 \dots Q_k A P_1 \dots P_s$. Значит $A \sim I_r$. \square

Теорема. Две матрицы $A, B \in \mathbb{F}^{m \times n}$ эквивалентны тогда и только тогда, когда их ранги совпадают.

Д-во. (\implies) Вытекает из того, что умножение на невырожденную матрицу не меняет ранга матрицы.

(\impliedby) Следует из предыдущей теоремы и транзитивности эквивалентности матриц. \square

Опр. Матрицы $A, B \in \mathbb{F}^{n \times n}$ называются подобными, если существует невырожденная матрица $C \in \mathbb{F}^{n \times n}$, т.ч. $A = C^{-1}BC$.

Теорема. Инварианты подобия:

1. Ранг матрицы;
2. Определитель матрицы;
3. След матрицы.

Д-во. 1) Сразу следует из предыдущей теоремы.

2) $|A| = |P^{-1}BP| = |P^{-1}||B||P| = |P^{-1}P||B| = |B|$.

3)

$$\begin{aligned} \text{tr}(A) &= \text{tr}(P^{-1}BP) = \sum_{i=1}^n (P^{-1}BP)_{ii} = \sum_{i=1}^n \sum_{j=1}^n (P^{-1})_{ij} (BP)_{ji} = \\ &= \sum_{i=1}^n \sum_{j=1}^n (P^{-1})_{ij} \sum_{k=1}^n b_{jk} (P)_{ki} = \sum_{j=1}^n \sum_{k=1}^n b_{jk} \sum_{i=1}^n (P)_{ki} (P)_{ij} = \\ &= \sum_{j=1}^n \sum_{k=1}^n b_{jk} (I)_{kj} = \sum_{j=1}^n b_{jj} = \text{tr}(B) \end{aligned}$$

\square

2.3 Ядро и образ линейного оператора. Соотношение между рангом и дефектом линейного оператора.

Опр. *Образом линейного оператора называется множество $\text{im } A = \{y \in W \mid \exists x \in V : Ax = y\}$. Ядром линейного оператора называется множество $\ker A = \{x \in V \mid Ax = 0\}$. Рангом линейного оператора называется размерность его образа, а дефектом - размерность его ядра.*

Теорема. *Если $A \in L(V, W)$, то $\ker A$ - линейное подпространство пространства V , $\text{im } A$ - линейное подпространство пространства W .*

Д-во. Для того чтобы подмножество было подпространством достаточно, чтобы применение операций сложения векторов и умножения вектора на число давало результат в подмножестве. Для данных подмножеств данные условия легко проверяются. \square

Теорема. *Если e_1, \dots, e_n - базис пространства V , то $\text{im } A = L(Ae_1, \dots, Ae_n)$.*

Д-во. Достаточно показать для множеств $\text{im } A$ и $L(Ae_1, \dots, Ae_n)$ имеет место двухстороннее вложение. С одной стороны, если $y \in \text{im } A$, то $y = Ax = A \sum_{i=1}^n x_i e_i = \sum_{i=1}^n x_i Ae_i \in L(Ae_1, \dots, Ae_n)$. С другой стороны, если $y \in L(Ae_1, \dots, Ae_n)$, то $y = \sum_{i=1}^n x_i Ae_i = A \sum_{i=1}^n x_i e_i = Ax \in \text{im } A$. \square

Теорема. *Если $A \in L(V, W)$, то $\text{rank } A + \text{def } A = \dim V$.*

Д-во. Пусть $\ker A \neq \{\theta\}$ и e_1, \dots, e_k - базис $\ker A$. Дополним его до базиса $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ пространства V . $\text{im } A = L(Ae_1, \dots, Ae_k, Ae_{k+1}, \dots, Ae_n) = L(Ae_{k+1}, \dots, Ae_n)$. Покажем, что векторы Ae_{k+1}, \dots, Ae_n линейно независимы. Пусть это не так. Тогда для нетривиальной линейной комбинации этих векторов имеет место соотношение $\alpha_{k+1} Ae_{k+1} + \dots + \alpha_n Ae_n = A(\alpha_{k+1} e_{k+1} + \dots + \alpha_n e_n) = \theta$. Следовательно, $\alpha_{k+1} e_{k+1} + \dots + \alpha_n e_n \in \ker A$. Это означает, что вектор $\alpha_{k+1} e_{k+1} + \dots + \alpha_n e_n$ линейно выражается через e_1, \dots, e_k , что невозможно в силу линейной независимости e_1, \dots, e_n . Таким образом, $\dim \ker A = k$, $\dim \text{im } A = n - k$. \square

Теорема. *Пусть M - конечномерное линейное пространство над полем \mathbb{P} . Тогда для любых его линейных подпространств V_1 и V_2 , т.ч. $\dim V_1 + \dim V_2 = \dim V$, существует линейный оператор $A \in L(V, V) : \text{im } A = V_1, \ker A = V_2$.*

Д-во. Пусть $\dim V_1 = p$, $\dim V_2 = q$, $\dim V = n$, $n = p + q$ и e_{p+1}, \dots, e_n - базис V_2 . Дополним его до базиса $V : e_1, \dots, e_p, e_{p+1}, \dots, e_n$. Выберем произвольный базис $V_1 : g_1, \dots, g_p$ и зададим линейный оператор $A \in L(V, V)$:

$$\begin{cases} Ae_1 = g_1, \dots, Ae_p = g_p \\ Ae_{p+1} = Ae_{p+2} = \dots = Ae_n = 0 \end{cases}$$

Тогда $\text{im } A = L(Ae_1, \dots, Ae_p) = A(g_1, \dots, g_p) = V_1$ и $\ker A = L(e_{p+1}, \dots, e_n) = V_2$. \square

2.4 Обратимый оператор. Критерий обратимости. Линейность обратного оператора.

Опр. Оператор $A : V \rightarrow W$ называется обратимым или невырожденным, если существует оператор $B : W \rightarrow V$ такой, что $AB = I_W$ и $BA = I_V$.

Утверждение. Если линейный оператор обратим, то обратный оператор определен однозначно и является линейным.

Д-во. 1) Пусть $A \in L(V, W)$ и $B_1, B_2 \in L(W, V)$ обратные к A . Тогда

$$\left. \begin{aligned} B_1AB_2 &= (B_1A)B_2 = I_V B_2 = B_2 \\ B_1AB_2 &= B_1(AB_2) = B_1I_W = B_2 \end{aligned} \right| \implies B_1 = B_2.$$

2) Пусть $A \in L(V, W)$ и $B \in L(W, V)$ — обратный к A . Тогда $\forall y_1, y_2 \in W \exists x_1, x_2 \in V : y_1 = Ax_1, y_2 = Ax_2$, значит $By_1 = x_1, By_2 = x_2$. $\forall \alpha_1, \alpha_2 \in \mathbb{P}$:

$$\begin{aligned} B(\alpha_1 y_1 + \alpha_2 y_2) &= B(\alpha_1 Ax_1 + \alpha_2 Ax_2) = B(A(\alpha_1 x_1 + \alpha_2 x_2)) = \\ &= (BA)(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 x_1 + \alpha_2 x_2 = \\ &= \alpha_1 By_1 + \alpha_2 By_2. \end{aligned}$$

□

Теорема. Пусть V и W — конечномерные пространства над общим полем. Тогда для обратимости линейного оператора $A \in L(V, W)$ необходимо и достаточно, чтобы $\dim V = \dim W$ и $\ker A = \{\theta\}$

Д-во. (\implies) Если $x_0 \in \ker A$, то $\forall x \in V : A(x + x_0) = Ax + Ax_0 = Ax + \theta = Ax = y \in W$. Значит $A^{-1}y = x = x + x_0$, т.е. $x_0 = \theta$. $\operatorname{im} A = L(Ae_1, \dots, Ae_{\dim V}) \subseteq W \implies \dim W \geq \dim V$ и $\operatorname{im} A^{-1} = L(A^{-1}f_1, \dots, A^{-1}f_{\dim W}) \subseteq V \implies \dim V \geq \dim W$. Значит $\dim V = \dim W$.

(\impliedby) Пусть $\dim V = \dim W$ и $\ker A = \{\theta\}$. Согласно теореме о размерности ядра и образа: $\operatorname{rank} A = n \implies$ оператор сюръективен и инъективен, а значит для каждого $y \exists! x = x(y) \in V : Ax = y$. Пусть оператор $B : W \rightarrow V$ определяется правилом $By = x(y)$. Тогда $(AB)y = y, (BA)x = x \implies$ выполнены условия обратимости оператора A . □

2.5 Оператор проектирования.

Опр. Пусть $V = L \oplus M$. Тогда любой вектор $x \in V$ однозначно представляется в виде суммы $x = u + v$, где $u \in L, v \in M$. Оператор P , переводящий x в u называется оператором проектирования на подпространство L параллельно подпространству M .

Утверждение. P является линейным оператором.

Д-во. $y_1, y_2 \in L$, $z_1, z_2 \in M$, $x_1 = z_1 + y_1$, $x_2 = y_2 + z_2$, $\lambda x_1 = \lambda y_1 + \lambda z_1$:

$$\begin{aligned} P(x_1 + x_2) &= y_1 + y_2 = Px_1 + Px_2 \\ P(\lambda x_1) &= \lambda y_1 = \lambda Px_1 \end{aligned}$$

□

Теорема. Для того чтобы линейный оператор $P \in L(V, V)$ был оператором проектирования, необходимо и достаточно, чтобы $P^2 = P$.

Д-во. (\Rightarrow) $V = L \oplus M \forall x \in V \exists! u \in L, v \in M : x = u + v$ и $Px = u$. Значит $Pu = u$ ($u = u + \theta$) и $P^2x = P(Px) = Pu = u = Px$, т.е. $P^2 = P$.

(\Leftarrow) Пусть $P^2 = P$. Положим $L = \text{im } P$, $M = \ker P$. Тогда $\dim L + \dim M = \dim V$. Если $w \in L \cap M$, то $w = Px$ и $Pw = \theta$. Поэтому $Pw = P^2x = Px = w = \theta$. Значит $L \cap M = \{0\}$. Значит $L \oplus M = V$. □

2.6 Собственные значения и собственные векторы. Характеристический многочлен линейного оператора (матрицы).

Опр. Пусть V - линейное пространство над полем \mathbb{P} . $A \in L(V, V)$. Число $\lambda \in \mathbb{P}$ и вектор $\theta \neq 0 \in V$ называются собственным значением и собственным вектором оператора A , если $A\theta = \lambda\theta$.

Теорема. Собственные векторы x_1, \dots, x_k , отвечающие различным собственным значениям $\lambda_1, \dots, \lambda_k$ линейно независимы.

Д-во. Применим индукцию по k . Для $k = 1$ утверждение очевидно. Пусть оно верно для любой системы из $k - 1$ векторов. Докажем его для k векторов x_1, \dots, x_k . Приравняем нулевому вектору линейную комбинацию этих векторов: $\alpha_1 x_1 + \dots + \alpha_k x_k = \theta$. Под действием оператора A это равенство перейдет в равенство $\alpha_1 \lambda_1 x_1 + \dots + \alpha_k \lambda_k x_k = \theta$ (*). $(*) - \lambda_k (*) = \alpha_1 (\lambda_1 - \lambda_k) x_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) x_{k-1} = \theta$. В силу индуктивного предположения отсюда следует, что $\alpha_1 = \dots = \alpha_{k-1} = 0$. Значит и $\alpha_k = 0$. Значит x_1, \dots, x_k линейно независимы. □

Следствие. Линейный оператор, действующий в n -м-ном пространстве, не может иметь более чем n различных собственных векторов.

Опр. Характеристическим многочленом матрицы $A \in \mathbb{P}^{m \times n}$ называется функция $f(\lambda) = |A - \lambda I|$.

Теорема. Характеристический многочлен матрицы является инвариантом подобия.

Д-во. Пусть $B = P^{-1}AP$. Тогда

$$\begin{aligned} |B - \lambda I| &= |(P^{-1}AP) - \lambda P^{-1}P| = |P^{-1}(A - \lambda I)P| = |P^{-1}||A - \lambda I||P| = \\ &= |P^{-1}||P||A - \lambda I| = |P^{-1}P||A - \lambda I| = |A - \lambda I|. \end{aligned}$$

□

Свойства характеристического многочлена.

- Характеристический многочлен индуцированного оператора является делителем характеристического многочлена порождающей его матрицы.
- Если $V = L_1 \oplus \dots \oplus L_k$, где L_1, \dots, L_k - инвариантные подпространства относительно оператора $A \in L(V, V)$, то характеристический многочлен $f(\lambda)$. Равен произведению характеристических многочленов $f_1(\lambda), \dots, f_k(\lambda)$ индуцированных операторов $A|_{L_1}, \dots, A|_{L_k}$.

Теорема. Пусть V - линейное пространство над полем \mathbb{P} . Число $\lambda \in \mathbb{P}$ является собственным значением оператора $A \in L(V, V)$ тогда и только тогда, когда λ - корень его характеристического многочлена.

Д-во. Число λ является собственным значением оператора A тогда и только тогда, когда существует вектор x , удовлетворяющий условиям

$$\begin{cases} Ax = \lambda x, \\ x \neq 0, \\ \lambda \in \mathbb{P}. \end{cases} \Leftrightarrow \begin{cases} (A - \lambda I)x = 0, \\ x \neq 0, \\ \lambda \in \mathbb{P}. \end{cases}$$

Это равносильно вырожденности оператора $A - \lambda I$ при некотором λ , т.е. $|A - \lambda I| = 0$. \square

2.7 Геометрическая и алгебраическая кратность собственного значения.

Опр. Пусть λ_0 - собственное значение оператора A . Множество $W_{\lambda_0} = \{x \in V : Ax = \lambda_0 x\}$ называется собственным подпространством оператора A , отвечающим собственному значению λ_0 .

Очевидно, что $W_{\lambda_0} = \ker(A - \lambda_0 I)$, поэтому собственное подпространство является линейным подпространством пространства V .

Опр. Размерность собственного подпространства W_{λ_0} называется геометрической кратностью собственного значения λ_0 , а кратность λ_0 как корня характеристического многочлена - его алгебраической кратностью.

Теорема. Геометрическая кратность собственного значения не превосходит его алгебраической кратности.

Д-во. Пусть m и s - алгебраическая и геометрическая кратность собственного значения λ_0 оператора $A \in L(V, V)$. Собственное подпространство W_{λ_0} инвариантно относительно оператора A , следовательно, можно рассматривать индуцированный оператор $A|_{W_{\lambda_0}}$. Найдем его характеристический многочлен $f_1(\lambda)$. Пусть e_1, \dots, e_s - базис W_{λ_0} . Тогда матрица оператора $A|_{W_{\lambda_0}}$ в этом базисе будет диагональной матрицей s -го порядка с элементами λ_0 на главной диагонали. Следовательно, $f_1(\lambda) = (\lambda_0 - \lambda)^s$. $(\lambda_0 - \lambda)^s$ является делителем характеристического многочлена $f(\lambda)$ оператора A , но $(\lambda_0 - \lambda)$ входит в характеристический многочлен $f(\lambda)$ ровно m раз. Значит, $s \leq m$. \square

2.8 Операторы простой структуры и диагонализуемые матрицы. Критерий диагонализуемости.

Опр. Линейный оператор $A \in L(V, V)$ называется оператором простой структуры, если в пространстве V существует базис из собственных векторов оператора A .

Теорема. Линейный оператор $A \in L(V, V)$ имеет простую структуру тогда и только тогда, когда в пространстве V существует базис, в котором он имеет диагональную матрицу.

Д-во. Пусть $\dim V = n$. Согласно определению оператор A имеет простую структуру тогда и только тогда, когда он имеет n линейно независимых собственных векторов e_1, \dots, e_n , в котором матрица A_e оператора A имеет вид

$$A_e = \begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{bmatrix},$$

где $\lambda_1, \dots, \lambda_n$ — собственные значения, соответствующие собственным векторам e_1, \dots, e_n . □

Следствие. В n -мерном пространстве линейный оператор, имеющий n различных значений, является оператором простой структуры.

Следствие. Если матрица порядка n имеет n попарно различных собственных значений, то она диагонализуема.

Теорема. Линейный оператор $A \in L(V, V)$ имеет простую структуру тогда и только тогда, когда $W_{\lambda_1} \oplus \dots \oplus W_{\lambda_p} = V$.

Д-во. (\implies) Пусть A имеет простую структуру. Тогда в пространстве V существует базис e_1, \dots, e_n , состоящий из собственных векторов оператора A . Рассмотрим подпространство $W_{\lambda_1} + \dots + W_{\lambda_p}$, оно содержится в V . С другой стороны, каждый вектор базиса e_1, \dots, e_n принадлежит одному из собственных подпространств, поэтому $P \subset \sum_{i=1}^n W_{\lambda_i} \implies W_{\lambda_1} + W_{\lambda_p} = V$. Эта сумма прямая, т.к. собственные подпространства $W_{\lambda_1}, \dots, W_{\lambda_p}$ имеют тривиальное пересечение. □

2.9 Верхняя треугольная форма матрицы линейного оператора в комплексном пространстве.

Вопрос о существовании собственных векторов сводится к вопросу о существовании корней характеристического многочлена, принадлежащих основному полю. В алгебраическом поле \mathbb{C} любой многочлен степени $n \geq 1$ имеет n корней. Отсюда вытекает следующее утверждение.

Теорема. Произвольный линейный оператор, действующий в n -мерном комплексном пространстве, имеет:

1. n собственных значений, если каждое собственное значение считать столько раз, какова его кратность как корня характеристического многочлена;
2. Хотя бы один собственный вектор;
3. На любом своем инвариантном подпространстве хотя бы один собственный вектор.

Лемма. Линейный оператор, действующий в n -мерном комплексном пространстве, обладает инвариантным пространством размерности $n - 1$.

Д-во. Линейный оператор A действующий в комплексном пространстве V , имеет собственное значение λ . Значит, $|A - \lambda I| = 0$ и $\text{rank}(A - \lambda I) \leq n - 1$. Следовательно, $\dim \text{im}(A - \lambda I) \leq n - 1$ и в пространстве V существует подпространство L размерности $n - 1$, которое содержит $\text{im}(A - \lambda I)$. Очевидно, что L инвариантно относительно оператора $A - \lambda I$. Покажем, что оно инвариантно и относительно A . Пусть $x \in L$, тогда $(A - \lambda I)x = y \in L \implies Ax = \lambda x + y \in L$. \square

Теорема. В n -мерном комплексном пространстве V для любого линейного оператора $A \in L(V, V)$ существует система n вложенных друг в друга инвариантных подпространств L_1, \dots, L_n всех размерностей от 1 до n , т.е. таких, что $L_1 \subset L_2 \subset \dots \subset L_n = V$, где $\dim L_k = k$, $k = 1, \dots, n$.

Д-во. Используем индукцию по n . Для $n = 1$ утверждение теоремы очевидно. Пусть теорема верна для всех линейных операторов размерности $n - 1$. Тогда, согласно лемме оператор A , действующий в n -мерном комплексном пространстве V , имеет инвариантное пространство L_{n-1} размерности $n - 1$. Тогда для индуцированного оператора $A|_{L_{n-1}}$ существует система вложенных инвариантных подпространств $L_1 \subset L_2 \subset \dots \subset L_{n-1}$. Так как действия операторов A и $A|_{L_{n-1}}$ совпадают, то подпространства L_1, \dots, L_{n-1} инвариантны относительно оператора A . Остается добавить, что $L_{n-1} \subset L_n = V$. \square

Теорема. Для любого комплексного оператора A , действующего в комплексном пространстве, существует базис, в котором матрица линейного оператора имеет треугольную форму.

Д-во. Для оператора A найдется система инвариантных подпространств L_1, \dots, L_n таких, что $\dim L_k = k$ и $L_1 \subset L_2 \subset \dots \subset L_n = V$. Искомый базис e_1, \dots, e_n строим так: в качестве вектора e_1 берем любой базис L_1 , в качестве e_k , $k > 1$ - вектор, дополняющий базис L_{k-1} до базиса L_k . В силу инвариантности подпространств L_1, \dots, L_n матрица A_e имеет верхнюю треугольную форму. \square

2.10 Многочлен от линейного оператора (матрицы). Теорема Гамильтона-Кэли.

Опр. Зафиксируем квадратную матрицу $A \in \mathbb{P}^{n \times n}$. Рассмотрим произвольный многочлен $f(\lambda) = \sum_{i=0}^k f_i \lambda^i$ и поставим ему в соответствие матрицу $\sum_{i=0}^n f_i A^i = f(A)$. $f(A)$ называется многочленом от матрицы A , соответствующий многочлену $f(\lambda)$ с коэффициентами из поля \mathbb{P} . Если $f(A) = 0$ $f(\lambda) \neq 0$, то говорят, что многочлен f аннулирует матрицу A .

Утверждение. Для любой матрицы можно найти многочлен, который ее аннулирует.

Д-во. Рассмотрим матрицы $I = A^0$, $A^1 = A$, A^2, \dots, A^{n^2} . Их $n^2 + 1$ штука \implies они ЛЗ $\implies \exists$ нетривиальный набор a_0, \dots, a_{n^2} , т.ч. $a_0 I + a_1 A + \dots + a_{n^2} A^{n^2} = O$ — искомым многочлен, т.к. $f(\lambda) \neq 0$ в силу нетривиальности набора a_0, \dots, a_{n^2} . \square

Опр. Многочлен, аннулирующий матрицу A и имеющий минимальную степень среди всех аннулирующих ее многочленов, называется минимальным многочленом матрицы A .

Теорема. Линейный оператор, действующий в комплексном (или в вещественном) пространстве, является корнем своего характеристического многочлена.

Д-во. 1. Докажем сначала для комплексного пространства V . Пусть $A \in L(V, V)$ и его характеристический многочлен имеет вид $f(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_j - \lambda)^{m_j} \dots (\lambda_p - \lambda)^{m_p}$. $V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_p}$ и, следовательно, для любого вектора $x \in V$ имеет место разложение $x = x_1 + \dots + x_p$, где $x_j \in K_{\lambda_j}$, $k = 1, \dots, p$. Тогда

$$f(A)x = f(A)x_1 + \dots + f(A)x_j + \dots + f(A)x_p.$$

Каждое слагаемое в этом разложении равно нулевому вектору, так как $f(A)x_j = (\lambda_1 I - A)^{m_1} \dots (\lambda_j I - A)^{m_j} \dots (\lambda_p I - A)^{m_p} x_j = \theta$, ибо операторы в этом произведении перестановочны, а $(A - \lambda_j I)^{m_j} x_j = \theta$. Следовательно, $f(A)x = \theta \forall x \in V$, т.е. $f(A) = O$.

2. Пусть V - вещественное линейное пространство. Возьмем какой-либо базис e пространства V , и пусть A_e - матрица оператора A в этом базисе. Рассмотрим любое комплексное пространство V_1 той же размерности. Пусть f - произвольный базис V_1 , тогда матрица A_e является матрицей оператора $B \in L(V_1, V_1)$ в базисе f , т.е. $A_e = B_f$. Значит характеристические многочлены операторов A и B совпадают, и согласно п. 1, $f(A_e) = O$. \square

2.11 Нильпотентные и квазискалярные операторы (матрицы). Критерий нильпотентности.

Опр. Пусть линейный оператор A действует в n -мерном пространстве. Если он имеет только одно собственное значение λ кратности n , то будем называть его квазискалярным.

Опр. Линейный оператор $A \in L(V, V)$ называется нильпотентным, если существует число $q \in \mathbb{N}$ такое, что $A^q = O$. Наименьшее число q , обладающее таким свойством, называется индексом нильпотентности (высотой) оператора A .

Теорема. В комплексном пространстве линейный оператор нильпотентен тогда и только тогда, когда он является квазискалярный с единственным собственным значением равным нулю.

Д-во. (\Rightarrow) Если λ - собственное значение нильпотентного оператора $A \in L(V, V)$ индекса q и x - собственное значение соответствующее ему, то $Ax = \lambda x \Rightarrow A^2x = \lambda^2x \Rightarrow \dots \Rightarrow A^qx = \lambda^qx$. Отсюда следует, что $\lambda^qx = 0$. Так как $x \neq 0$, то $\lambda = 0$.

(\Leftarrow) Рассмотрим базис e комплексного пространства V , в котором оператор A имеет верхнюю треугольную матрицу с нулями на главной диагонали. Итак,

$$A_e = \begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Нетрудно проверить, что при последовательном возведении этой матрицы в степени $q = 2, 3, \dots, n$ нетривиальный треугольник расположенный над главной диагональю, перемещается каждый раз на одну диагональ выше, так что $(A_e)^n = O$. Значит, $A^n = O$. \square

2.12 Прямая сумма линейных операторов (матриц). Теорема о расщеплении вырожденного оператора.

Опр. Если $V = L_1 \oplus L_2 \oplus \dots \oplus L_p$ - прямая сумма подпространств L_1, \dots, L_p инвариантных относительно линейного оператора $A \in L(V, V)$, то оператор A называется прямой суммой индуцированных операторов $A|L_1, \dots, A|L_p$.

Теорема. Вырожденный и не нильпотентный оператор $A \in L(V, V)$ является прямой суммой нильпотентного и обратимого операторов, причем это разложение единственно.

Д-во. Для доказательства теоремы необходимо показать, что существует единственная пара подпространств L_1, L_2 , инвариантных относительно линейного оператора A и таких, что $V = L_1 \oplus L_2$, $A|L_1$ нильпотентен, $A|L_2$ обратим.

Существование. Обозначим для $k \in \mathbb{N}$: $N_k = \ker A^k$, $T_k = \operatorname{im} A^k$.

1. Покажем, что подпространства N_k строго вложены друг в друга до некоторого момента q , начиная с которого все N_k совпадают, т.е. $N_1 \subset N_2 \subset \dots \subset N_q = N_{q+1} = \dots$.

а) Вложение $N_k \subseteq N_{k+1}$ очевидно, так как если $A^k x = \theta$, то $A^{k+1}x = A(A^k x) = A\theta = \theta$.

б) Пусть $N_k = N_{k+1}$. Тогда $N_{k+1} = N_{k+2}$, так как $N_{k+1} \subseteq N_{k+2}$, $N_{k+2} \subseteq N_{k+1}$. Второе

из этих вложений следует из того, что если $x \in N_{k+2}$, то $A^{k+2}x = \theta$, т.е. $A^{k+1}(Ax) = \theta$. Значит, $Ax \in N_{k+1} = N_k$, откуда $A^k(Ax) = \theta$, т.е. $A^{k+1}x = \theta$.

Из а и б следует, что подпространство N_k либо строго вложено в N_{k+1} , либо совпадает со всеми последующими ядрами. Так как в конечномерном пространстве размерности подпространств N_k не могут бесконечно возрастать, то наступит момент q , начиная с которого все ядра N_k будут совпадать с N_q .

2. Зафиксируем этот момент q и покажем, что $V = N_q \oplus T_q$.

Действительно, $\dim V = \dim N_q + \dim T_q$ в силу теоремы о ранге и дефекте, при этом $N_q \cap T_q = \{\theta\}$, так как если $y \in N_q \cap T_q$, то $A^q y = \theta$, $y = A^q x$, т.е. $A^{2q} x = \theta$. Значит, $x \in N_{2q} = N_q$ и $A^q x = \theta = y$.

3. Подпространства N_q и T_q инвариантны относительно A , т.к.:

а) если $x \in N_q$, то $x \in N_{q+1} = N_q \implies A^{q+1}x = \theta$, т.е. $A^q(Ax) = \theta \implies Ax \in N_q$.

б) если $y \in T_q$, то $y = A^q x$ и $Ay = A^{q+1}x = A^q(Ax) = A^q x_1$, где $x_1 = Ax$, следовательно, $Ay \in T_q$.

4. Оператор $A|_{N_q}$ - нильпотентный оператор индекса q , т.к.:

а) $A^q x = \theta \forall x \in N_q$;

б) $\exists x_0 \in N_q$ такой, что $A^{q-1}x_0 \neq \theta$, ибо $N_{q-1} \neq N_q$.

5. Оператор $A|_{T_q}$ обратим, так как его ядро состоит только из нулевого вектора. Действительно, если $y \in \ker A|_{T_q}$, то $y \in T_q$, $Ay = \theta$, т.е. $y = A^q x$ и $A^{q+1}x = \theta$. Отсюда следует, что $x \in N_{q+1} = N_q$, т.е. $A^q x = \theta$ и $y = \theta$.

Утверждения 2-5 доказывают существование искомого разложения: $L_q = N_q$, $L_2 = T_q$.

Единственность. Пусть существует другое разложение $V = N \oplus T$, обладающее всеми свойствами первого.

1. Нильпотентность оператора $A|_N$ означает, что $A^k x = \theta \forall x \in N$, при некотором $k \in \mathbb{N}$. Следовательно, $N \subseteq N_k \subseteq N_q$ и $\dim N \leq \dim N_q$.

2. Обратимость оператора $A|_T$ означает, что $\text{im } A|_T = T$. Следовательно, для любого вектора $y \in T$ имеет место представление $y = Ay_1$, где $y_1 \in T$. Используя такое же представление для вектора y_1 и всех последующих, получаем, что $y = Ay_1 = A^2 y_2 = \dots = A^q y_q$. Таким образом, $T \subseteq T_q$ и $\dim T \leq \dim T_q$.

Так как $\dim V = \dim N + \dim T = \dim N_q + \dim T_q$ и $\dim N \leq \dim N_q$, $\dim T \leq \dim T_q$, то $N = N_q$ и $T = T_q$. \square

2.13 Корневое расщепление линейного оператора.

Опр. Пусть λ_j - собственное значение оператора A . Вектор $x \in V$ называется *корневым вектором* оператора A , отвечающим собственному значению λ_j , если $(A - \lambda_j I)^k x = \theta$ при некотором $k \in \mathbb{N} \cup \{0\}$. *Высотой* корневого вектора называется наименьшее k , обладающее указанным свойством.

Опр. Множество $K_{\lambda_j} = \{x \in V \mid \exists k \in \mathbb{N} \cup \{0\} : (A - \lambda_j I)^k x = \theta\}$ называется *корневым подпространством* оператора A , отвечающим собственному значению λ_j .

Утверждение. Корневое подпространство K_{λ_j} инвариантно относительно A .

Д-во. $v \in K_{\lambda_j} \implies \exists q_j : (A - \lambda_j I)^{q_j} v = \theta \implies (A - \lambda_j I)^{q_j} (Av) = A(A - \lambda_j I)^{q_j} v = A \cdot \theta = \theta \implies Av \in K_{\lambda_j}$. \square

Оператор $B = A - \lambda_j I$ - вырожденный, но не нильпотентный. Следовательно, к оператору B применима теорема о прямой сумме нильпотентного и обратимого оператора. Согласно этой теореме, если $N_k = \ker B^k$, $T_k = \operatorname{im} B^k$, то $N_1 \subset N_2 \subset \dots \subset N_q = N_{q+1} = \dots$. $V = N_q \oplus T_q$, где N_q и T_q - инвариантны относительно B .

Вернемся к оператору A .

N_1 состоит из корневых векторов оператора A высоты не превосходящей 1, т.е. совпадающим собственному значению λ_j . Таким образом $N_1 = W_{\lambda_1}$ и, следовательно, $\dim N_1 = s_j$, где s_j - геометрическая кратность собственного значения λ_j .

N_2 состоит из корневых векторов оператора A высоты, не превосходящей 2, а N_q состоит из векторов всех высот, т.е. q - максимальная высота коневого вектора, отвечающего собственному вектору λ_j , и N_q совпадает со всем корневым подпространством K_{λ_j} . Таким образом, $K_{\lambda_j} = N_q$.

Из свойств подпространства N_q вытекают важные свойства корневых подпространств: если характеристический многочлен оператора A имеет вид $f(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_j - \lambda)^{m_j} \dots (\lambda_p - \lambda)^{m_p}$, то

а) подпространство K_{λ_j} инвариантно относительно оператора A (в силу инвариантности относительно оператора $A - \lambda_j I$).

б) характеристический многочлен оператора $A|_{K_{\lambda_j}}$ имеет вид $f_j(\lambda) = (\lambda_j - \lambda)^{m_j}$ (т.к. $f_{A|N_q}(\lambda) = (-\lambda)^{m_1}$, $f_{A|T_q} = (\lambda_2 - \lambda)^{m_2} \dots (\lambda_p - \lambda)^{m_p}$)

в) $\dim K_{\lambda_j} = m_j$.

Теорема. Если A - линейный оператор, действующий в комплексном пространстве V и $f(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_p - \lambda)^{m_p}$, $\lambda_i \neq \lambda_k$, при $i \neq k$ - его характеристический многочлен, то пространство V разлагается в прямую сумму его корневых подпространств: $V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_p}$.

Д-во. Воспользуемся индукцией по p . Для $p = 1$, понятно, что $V = K_{\lambda_1}$. Пусть теорема верна для оператора, имеющего $p - 1$ различных собственных значений. Докажем ее для оператора A . Выделим корневое подпространство $K_{\lambda_p} = N_q = \ker(A - \lambda_p I)^{m_p}$. Тогда $V = K_{\lambda_p} \oplus T_q$, $T_q = \operatorname{im} (A - \lambda_p I)^{m_p}$. Обозначим $V_1 = T_q$. Пространство V_1 инвариантно относительно оператора $A - \lambda_p I$, а, следовательно, оно инвариантно и относительно A , при этом характеристический многочлен оператора $A_1 = A|_{V_1}$ имеет вид $f_1(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_{p-1} - \lambda)^{m_{p-1}}$. Оператор A_1 имеет $p - 1$ различных собственных значений, и для него теорема верна. Если учесть, что корневые пространства оператора A_1 совпадают с корневыми подпространствами $K_{\lambda_1}, \dots, K_{\lambda_{p-1}}$ оператора A , то $V_1 = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_{p-1}}$ и $V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_{p-1}} \oplus K_{\lambda_p}$. \square

2.14 Нерасщепляемые операторы и подпространства Крылова.

В максимальном расщеплении линейного оператора каждое инвариантное подпространство не может быть прямой суммой ненулевых инвариантных подпространств. Такие

подпространства и сужение оператора на них естественно называть нерасщепляемыми. Согласно теореме о корневом расщеплении, нерасщепляемый оператор обязан быть квазискалярным.

Опр. Инвариантное подпространство $M = M(A, x)$ оператора A , содержащее заданный ненулевой вектор x , называется минимальным, если данное подпространство содержится в любом инвариантном подпространстве, которому принадлежит вектор x .

Минимальное инвариантное подпространство $M(A, x)$ должно содержать последовательность векторов x, Ax, A^2x, \dots . Векторы такого вида принято называть векторами Крылова, а линейные оболочки $L_k(A, x) = L(x, Ax, A^2x, \dots, A^{k-1}x)$ — пространствами Крылова.

Лемма. Минимальное инвариантное подпространство $M(A, x)$ совпадает с пространством Крылова $L_x(A, x)$, содержащим вектор $A^k x$. Его размерность равна минимальному значению k , при котором $A^k x \in L_k(A, x)$.

Д-во. Пусть $x, Ax, \dots, A^{k-1}x$ — ЛНЗ, а вектор $A^k x$ выражается в виде их линейной комбинации. Ясно, что $\dim L_k(A, x) = k$. Условие $A^k x \in L_k(A, x)$ обеспечивает инвариантность подпространства $L_k(A, x)$. В то же время, любое инвариантное подпространство, содержащее вектор x , обязано содержать все пространство Крылова $\implies M(A, x) = L_k(A, x)$. \square

Лемма. Минимальное инвариантное подпространство $M(A, x)$ нерасщепляемо в том и только в том случае, когда сужение оператора A на нем квазискалярно.

Д-во. Квазискалярность является необходимым условием нерасщепляемости. Докажем его достаточность в случае подпространства $M(A, x)$. $M = M(A, x) = L_k(A, x)$, где $k = \dim L_k(A, x)$ и $A^k x \in L_k(A, x)$. Пусть единственное собственное значение оператора A на M равно λ . Тогда $B = A - \lambda I$ — нильпотентный на M , $M = L_k(B, x)$, система $x, Bx, \dots, B^{k-1}x$ — ЛНЗ и индекс нильпотентности $B|_M$ не больше k . Значит $B^k x = \theta$. Пусть $L \subseteq M$ — произвольное ненулевое инвариантное подпространство B . Возьмем ненулевой вектор $\theta \neq z \in L$, $z = \sum_{j=0}^{k-1} \alpha_j B^j x$, пусть i — минимальное число такое, что $\alpha_i \neq 0$. Тогда $B^{k-1-i} z = \alpha_i B^{k-1} x \in L \implies B^{k-1} x \in L$. Таким образом, любое инвариантное подпространство $L \subseteq M$ оператора B содержит общий вектор $B^{k-1} x$. Значит M нельзя представить в виде прямой суммы двух ненулевых инвариантных подпространств оператора B . Каждое инвариантное пространство оператора A является инвариантным и для оператора $B = A - \lambda I$. \square

2.15 Условие линейной независимости составной системы векторов Крылова нильпотентного оператора.

Лемма. Пусть A — линейный оператор и k_1, \dots, k_t — его индексы нильпотентности на ненулевых векторах x_1, \dots, x_t . Тогда для линейной независимости составной си-

стемы векторов Крылова: $x_1, Ax_1, \dots, A^{k_1-1}x_1, \dots, x_t, Ax_t, \dots, A^{k_t-1}x_t$ (1) необходима и достаточна линейная независимость векторов $A^{k_1-1}x_1, \dots, A^{k_t-1}x_t$ (2).

Д-во. (\Rightarrow) Из системы (1) очевидно следует линейная независимость системы (2).

(\Leftarrow) Пусть (2) линейно независима и $k = \max_{1 \leq i \leq t} k_i$. Индукция по k . При $k = 1$ системы (1) и (2) совпадают. Пусть $k \geq 2$ и $I_k = \{i : k_i = k, i = \overline{1, t}\}$. Пусть $\theta = y =$

$\sum_{i=1}^t \sum_{j=1}^{k_i-1} \alpha_{ij} A^j x_i (*)$. Тогда $\theta = A^{k-1}y = \sum_{i \in I_k} \alpha_{i0} A^{k-1}x_i \Leftrightarrow \alpha_{i0} = 0 \forall i \in I_k$. Из системы

(1) удалим все векторы $x_i, i \in I_k$. Оставшаяся система — составная система векторов Крылова, но индексы нильпотентности A на векторах $x_i, i \notin I_k$ и векторах $Ax_i, i \in I_k$ меньше k . По предположению индукции для векторов, оставшихся в $(*)$ (в силу ЛНЗ) $\alpha_{ij} = 0, \forall i, j : 1 \leq i \leq t, 0 \leq j \leq k_i - 1$. \square

Следствие. Пусть A - нильпотентный оператор и $L_{k_1}(A, x_1), \dots, L_{k_t}(A, x_t)$ - максимальные пространства Крылова. Для того чтобы их сумма была прямой необходимо и достаточно, чтобы $A^{k_1-1}x_1, \dots, A^{k_t-1}x_t$ были линейно независимыми.

2.16 Максимальное расщепление и жорданова форма нильпотентного оператора.

Теорема. Любой нильпотентный оператор A на конечномерном пространстве V расщепляется в прямую сумму операторов на максимальных пространствах Крылова и в любом таком расщеплении число n_k пространств размерности k не зависит от метода, которым было получено расщепление и равно $n_k = 2\text{def } A^k - \text{def } A^{k+1} - \text{def } A^{k-1}$.

Д-во. Пусть нильпотентный оператор A действует в V . Рассмотрим какие-либо максимальные пространства Крылова: $L_1 = L_{k_1}(A, x_1), \dots, L_s = L_{k_s}(A, x_s)$ дающие в сумме (необязательно прямой) пространство V . Чтобы получить их, возьмем в качестве начальных векторов x_1, \dots, x_s , например, базис пространства V .

Если $V = L_1 + \dots + L_s$ - прямая сумма, то нужное расщепление получено. Если нет, то

$A^{k_1-1}x_1, \dots, A^{k_s-1}x_s$ — ЛЗ. Рассмотрим нетривиальную линейную комбинацию $\sum_{i=1}^s \alpha_i A^{k_i-1}x_i =$

θ . Среди векторов линейной комбинации, перед которыми стоит ненулевой коэффициент, выберем тот, который входит в пространство Крылова с наименьшей размерностью. Пусть это будет $A^{j-1}x_j$.

$$A^{k_j-1}x_j + \sum_{i=1, i \neq j}^s \frac{\alpha_i}{\alpha_j} A^{k_i-1}x_i = \theta.$$

$$\begin{aligned}
y_0 &= x_j + \sum_{i=1, i \neq j}^s \frac{\alpha_i}{\alpha_j} A^{k_i - k_j} x_i, \\
y_1 &= Ax_j + \sum_{i=1, i \neq j}^s \frac{\alpha_i}{\alpha_j} A^{k_i - k_j + 1} x_i, \\
&\dots \\
y_{k_j - 1} &= A^{k_j - 1} x_j + \sum_{i=1, i \neq j}^s \frac{\alpha_i}{\alpha_j} A^{k_i - 1} x_i = \theta.
\end{aligned}$$

$y_0, \dots, y_{k_j - 1}$ образуют последовательность векторов Крылова. Если $y_0 = \theta$, то все они нулевые и пространство V - сумма меньшего числа максимальных пространств Крылова. Если $y_0 \neq \theta$, то $V = L_1 + \dots + L_{j-1} + L'_j + L_{j+1} + \dots + L_s$, где $L'_j = L_k(A, y_0)$ - максимальное пространство Крылова размерности $d = \dim L(y_0, \dots, y_{k_j - 2}) < k_j$. Мы получаем новое расщепление, в котором сумма размерностей максимальных пространств Крылова уменьшена. Всякий раз, когда V представляется суммой максимальных пространств Крылова с суммой размерностей больше, чем $\dim V$, мы можем найти аналогичной расщепление с меньшей суммой размерностей. В какой-то момент сумма размерностей будет $\dim V$ и сумма подпространств будет прямой.

Пусть $V = L_{k_1}(A, x_1) \oplus \dots \oplus L_{k_s}(A, x_s)$ - прямая сумма максимальных пространств Крылова и среди них ровно m_k пространств размерности $\geq k$. Докажем, что $m_k = \text{def } A^k - \text{def } A^{k-1}$.

Множество $\{A^d x_i : k \leq d \leq k_i - 1, i = \overline{1, s}\}$ состоит из ненулевых векторов, образующих базис пространства $\text{im } A^k : \text{rank } A^k = \sum_{i=1}^s l_i, l_i = \max(0, k_i - k) \implies \text{def } A^k = \sum_{i=1}^s (k_i - l_i) \implies \ker A^k = L(A^{l_1} x_1, A^{l_1 + 1} x_1, \dots, A^{k_1 - 1} x_1, \dots, A^{l_s} x_s, A^{l_s + 1} x_s, \dots, A^{k_s - 1} x_s)$.

Число образующих векторов равно размерности ядра и все они принадлежат ядру оператора A^k . Если $l_j = k_j - k$, то $A^k A^{l_j} x_j = A^{k + k_j - k} x_j = A^{k_j} x_j = \theta$, а если $l_j = \theta$, то $d_j < d \implies A^d A^{l_j} x_j = A^d x_j = \theta$. Среди образующих векторов все, кроме, возможно, первых векторов $A^{l_1} x_1, \dots, A^{l_s} x_s$ в последовательностях Крылова, принадлежат пространству $\ker A^{k-1}$. А число тех из них, которые ему не принадлежат, как раз и равно числу пространств Крылова размерности, которых $\geq k$. Пусть M_k - их линейная оболочка. Тогда $\dim M_k = m_k, \ker A^k = \ker A^{k-1} \oplus M_k \implies m_k = \text{def } A^k - \text{def } A^{k-1}$.

$n_k = m_k - m_{k+1} = (\text{def } A^k - \text{def } A^{k-1}) - (\text{def } A^{k+1} - \text{def } A^k) = -\text{def } A^{k-1} + 2\text{def } A^k - \text{def } A^{k+1}$. \square

Следствие. Для нильпотентного оператора на конечномерном пространстве расщепление в прямую сумму операторов является максимальным тогда и только тогда, когда операторы расщепления определены на максимальных пространствах Крылова.

Д-во. Если подпространство в прямой сумме не является максимальным пространством

Крылова, то оно может быть разложено в нетривиальную прямую сумму таких подпространств, а значит исходное разложение не максимальное. \square

Пусть λ - единственное собственное значение квазискалярного оператора A и пусть $L_k(A, x)$ — его инвариантное пространство Крылова. В данном пространстве имеется линейно независимая системы векторов $x, Bx, \dots, B^{k-1}x$, где $B = A - \lambda I$. Составим из них базис $e_1 = B^{k-1}x, e_2 = B^{k-2}x, \dots, e_{k-1} = Bx, e_k = x \implies Ae_1 = \lambda e_1, Ae_2 = \lambda e_2 + e_1, \dots, Ae_k = \lambda e_k + e_1$. Матрица оператора $A|_{L_k(A, x)}$ в этом базисе:

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}_{k \times k}$$

— жорданова клетка собственного значения λ .

Жордановой матрицей называется прямая сумма жордановых клеток. Для заданной матрица A подобная ей жорданова матрица $J = P^{-1}AP$ и столбцы матрицы P называются жордановой формой и жордановым базисом матрицы A . Ясно, что $J_n^m(0) \neq O$ при $m = \overline{1, n-1}$ и $J_n^n(0) = O$, т.е. $J_n(0)$ - нильпотентная матрица с индексом нильпотентности равным n .

2.17 Теорема Жордана о структуре линейного оператора.

Теорема. Пусть V — конечномерное линейное пространство, $n = \dim V$, $A \in L(V, V)$ — квазискалярный оператор и λ — его единственное собственное значение. Тогда существует базис h_1, \dots, h_n , т.ч. $A_h = J_{k_1}(\lambda) \oplus \dots \oplus J_{k_r}(\lambda)$, где $k_1 + \dots + k_r = n$.

Д-во. Оператор $B = A - \lambda I$ — нильпотентный. Значит существует базис h_1, \dots, h_n , т.ч. $B_r = J = J_{k_1}(0) \oplus \dots \oplus J_{k_r}(0)$, $k_1 + \dots + k_r = n$. Далее, $A = B + \lambda I \implies A_h = B_h + \lambda I = J + \lambda I = J_{k_1}(\lambda) \oplus \dots \oplus J_{k_r}(\lambda)$. \square

Теорема (Жордана о структуре линейного оператора). Пусть V — конечномерное линейное пространство, $\dim V = n$, $A \in L(V, V)$ и $f(\lambda) = (-1)^n(\lambda - \lambda_1)^{l_1} \dots (\lambda - \lambda_k)^{l_k}$, $\lambda_p \neq \lambda_q$ — характеристический многочлен A . Тогда существует базис h_1, \dots, h_n , т.ч. $A_h = A_1 \oplus \dots \oplus A_k$, где $A_j = J_{i_1}(\lambda_j) \oplus \dots \oplus J_{i_m}(\lambda_j)$, $m = m(j)$, $i_1 + \dots + i_m = l_j$, $j = \overline{1, k}$.

Д-во. Положим $f_j(\lambda) = (\lambda - \lambda_j)^{l_j}$, $j = \overline{1, k}$. Из теоремы о корневом расщеплении $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$, где $W_{\lambda_j} = \ker(A - \lambda_j I)^{l_j}$, $\dim W_{\lambda_j} = l_j$, $AW_{\lambda_j} \subset W_{\lambda_j}$, $j = \overline{1, k}$ и $A_j = A|_{W_{\lambda_j}}$ — сужение оператора A на подпространство W_{λ_j} . Оператор $A_j \in L(W_{\lambda_j}, W_{\lambda_j})$ имеет единственное собственное значение λ_j , алгебраической кратности l_j . Применим предыдущую теорему к каждому оператору A_j , получим искомое утверждение. \square

2.18 Единственность жордановой формы линейного оператора (матрицы).

Пусть K_{λ_j} - корневое пространство оператора A , отвечающее собственному значению λ_j . Положим $B = A - \lambda_j I$, $N_k = \ker B^k$, $n_k = \dim N_k$, $r_k = \text{rank } B^k$.

Построим сначала само корневое подпространство K_{λ_j} . Для этого необходимо найти момент q , начиная с которого все ядра N_q будут совпадать с $N_q = K_{\lambda_j}$, при этом имеем $n_1 = s_j < n_2 < \dots < n_q = m_j$, где s_j и m_j - геометрическая и алгебраическая кратности λ_j . Теперь будем строить базис K_{λ_j} , последовательно просматривая подпространства N_q, N_{q-1}, \dots, N_1 .

N_q) Пусть f_1, \dots, f_{t_q} - векторы, дополняющие произвольный базис N_{q-1} до базиса N_q . Ясно, что:

- 1) они будут корневыми векторами высоты q ;
- 2) их количество равно $n_q - n_{q-1}$;
- 3) $t_q = n_q - n_{q-1} = (n_q - n_{q-1}) - (n_{q+1} - n_q) = -n_{q+1} + 2n_1 - n_{q-1}$, так как $n_{q+1} = n_q$.
- 4) никакая нетривиальная линейная комбинация этих векторов не принадлежит N_{q-1} (такие векторы будем называть линейно независимыми над N_{q-1}).

N_{q-1}) Построим векторы Bf_1, \dots, Bf_{t_q} . Эти векторы являются корневыми векторами высоты $q - 1$, и они линейно независимы над N_{q-2} , так как в противном случае для

нетривиального набора чисел $\alpha_1, \dots, \alpha_{t_q}$ имеем $B^{q-2} \sum_{k=1}^{t_q} \alpha_k B f_k$, т.е. $B^{q-1} \sum_{k=1}^{t_q} \alpha_k f_k = \theta$, и

$\sum_{k=1}^{t_q} \alpha_k f_k \in N_{q-1}$, что противоречит линейной независимости f_1, \dots, f_{t_q} над N_{q-1} .

Дополним эти векторы векторами $g_1, \dots, g_{t_{q-1}} \in N_{q-1}$ так, что векторы $Bf_1, \dots, Bf_{t_q}, g_1, \dots, g_{t_{q-1}}$ дополняли произвольный базис N_{q-2} до базиса N_{q-1} . Ясно, что:

- 1) они будут корневыми векторами высоты $q - 1$;
- 2) их количество равно $n_{q-1} - n_{q-2}$;
- 3) $t_{q-1} = (n_{q-1} - n_{q-2}) - (n_q - n_{q-1}) = -n_q + 2n_{q-1} - n_{q-2}$;
- 4) они линейно независимы над N_{q-2} .

Выполняя далее такие же построения в подпространствах N_{q-2}, N_{q-3}, \dots , придем к подпространству N_1 .

N_1) Здесь строятся векторы

$$B^{q-1}f_1, \dots, B^{q-1}f_{t_q}, B^{q-2}g_1, \dots, B^{q-2}g_{t_{q-1}}, \dots, Bv_1, \dots, Bv_{t_2},$$

которые дополняются векторами u_1, \dots, u_{t_1} до базиса N_1 . Таким образом векторы

$$B^{q-1}f_1, \dots, B^{q-1}f_{t_q}, B^{q-2}g_1, \dots, B^{q-2}g_{t_{q-1}}, \dots, Bv_1, \dots, Bv_{t_2}, u_1, \dots, u_{t_1},$$

- 1) являются собственными векторами;
- 2) их количество равно $n_1 = n_1 - n_0$ (очевидно, $n_0 = \text{def } B^0 = 0$);
- 3) $t_1 = (n_1 - n_0) - (n_2 - n_1) = -n_2 + 2n_1 - n_0$;
- 4) они линейно независимы.

Полученную за q шагов систему вектором удобно объединить в таблицу, которую будем называть жордановой лестницей.

N_1	$\underbrace{f_1, \dots, f_{t_{q-1}}}_{t_q = -n_{q-1} + 2n_q - n_{q+1}}$			
N_{q-1}	Bf_1, \dots, Bf_{t_q}	$\underbrace{g_1, \dots, g_{t_{q-1}}}_{t_{q-1} = -n_{q-2} + 2n_{q-1} - n_q}$		
\vdots	\vdots	\vdots	\ddots	
N_1	$B^{q-1}f_1, \dots, B^{q-1}f_{t_q}$	$B^{q-2}g_1, \dots, B^{q-2}g_{t_{q-1}}$	\dots	$\underbrace{u_1, \dots, u_{t_1}}_{t_1 = -n_0 + 2n_1 - n_2}$

Получили, что число и размер клеток жордана однозначно определяется размерностями ядер операторов $(A - \lambda_j I)^i$.

Таким образом доказали теорему:

Теорема. Любая матрица $A \in \mathbb{C}^{n \times n}$ подобна прямой сумме жордановых клеток $P^{-1}AP = J_1 \oplus \dots \oplus J_n$, причем число и размеры жордановых клеток определяются однозначно по матрице A .

2.19 Критерий подобия комплексных матриц.

Опр. Матрицы A и B называются подобными, если существует невырожденная матрица X , такая, что $A = X^{-1}BX$.

Теорема. Две матрицы $A, B \in \mathbb{C}^{n \times n}$ подобны тогда и только тогда, когда их жордановы формы совпадают.

Д-во. В доказательстве нуждается только подобие жордановых матриц с одинаковым набором жордановых клеток. Это утверждение следует из того, что перемещение столбцов матрицы реализуется умножением матрицы на матрицу элементарных преобразований, которая не вырождена. \square

2.20 Блочно-диагональная жорданова форма вещественной матрицы.

Пусть $A \in \mathbb{R}^{n \times n}$, $f_A(\lambda) \in \mathbb{R}_n[\lambda]$.

Если $\lambda \in \mathbb{R}$, $f_A(\lambda) = 0$, то корневое пространство, порожденное этим собственным значением, расщепляется в прямую сумму жордановых клеток $J_k(\lambda)$.

Пусть $\lambda = a + bi$ ($a, b \in \mathbb{R}$ и $b \neq 0$), $f_A(\lambda) = 0$. Тогда $\bar{\lambda} = a - bi$ тоже собственное значение A и алгебраические кратности λ и $\bar{\lambda}$ совпадают. Пусть h_1, \dots, h_k — жорданова цепочка (система векторов Крылова) клетки $J_k(\lambda)$, тогда $\bar{h}_1, \dots, \bar{h}_k$ — жорданова цепочка, отвечающая жордановой клетке $J_k(\bar{\lambda})$.

Пусть $h_j = x_j + iy_j$ ($x_j, y_j \in \mathbb{R}^n$), $j = \bar{1}, \bar{k}$, тогда

$$\begin{aligned} Ah_1 &= \lambda h_1, & Ah_j &= \lambda h_j + h_{j-1} \quad (2 \leq j \leq k) \\ A\bar{h}_1 &= \bar{\lambda} \bar{h}_1, & A\bar{h}_j &= \bar{\lambda} \bar{h}_j + \bar{h}_{j-1} \quad (2 \leq j \leq k) \end{aligned} \quad (*)$$

Система $h_1, \dots, h_k, \bar{h}_1, \dots, \bar{h}_k$ — ЛНЗ над \mathbb{C} и $L_{\mathbb{C}}(h_1, \bar{h}_1, \dots, h_k, \bar{h}_k) = L_{\mathbb{C}}(x_1, y_1, \dots, x_k, y_k)$.
 Значит $2k = \dim L_{\mathbb{C}}(h_1, \bar{h}_1, \dots, h_k, \bar{h}_k) = \dim L_{\mathbb{C}}(x_1, y_1, \dots, x_k, y_k)$. Поэтому
 $\dim L_{\mathbb{R}}(x_1, y_1, \dots, x_k, y_k) = 2k$.

Перепишем жорданову цепочку (*):

$$\begin{cases} A(x_1 + iy_1) = (a + ib)(x_1 + iy_1), \\ A(x_2 + iy_2) = (x_1 + iy_1) + (a + ib)(x_2 + iy_2), \\ \dots \\ A(x_k + iy_k) = (x_{k-1} + iy_{k-1}) + (a + ib)(x_k + iy_k). \end{cases}$$

Выделим действительные части и мнимые части в этих равенствах, получим:

$$\begin{cases} Ax_1 = ax_1 - by_1, & Ay_1 = bx_1 + ay_1, \\ Ax_2 = 1x_1 + 0y_1 + ax_2 - by_2, & Ay_2 = 0x_1 + 1y_1 + bx_2 + ay_2, \\ \dots & \\ Ax_k = 1x_{k-1} + 0y_{k-1} + ax_k - by_k, & Ay_k = 0x_{k-1} + 1y_k + bx_k + ay_k. \end{cases} \quad (**)$$

Значит вещественное подпространство $L_{\mathbb{R}}(x_1, y_1, \dots, x_k, y_k) = L$ — это инвариантное подпространство размерности $2k$ и A_L — сужение A на L выглядит так:

$$J = \begin{bmatrix} a & b & 1 & 0 & & & \\ -b & a & 0 & 1 & & & \\ & a & b & 1 & 0 & & \\ & -b & a & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & a & b & 1 & 0 \\ & & & -b & a & 0 & 1 \\ & & & & a & b & \\ & & & & -b & a & \end{bmatrix} \in \mathbb{R}^{2k \times 2k}$$

Положим $\Lambda = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ и $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ можем записать:

$$J = \begin{bmatrix} \Lambda & E & & & \\ & \Lambda & E & & \\ & & \ddots & \ddots & \\ & & & \Lambda & E \\ & & & & \Lambda \end{bmatrix}$$

— блочная жорданова клетка порядка $2k$. Таким образом, доказали теорему:

Теорема. Любая вещественная матрица подобна прямой сумме вещественных жордановых клеток и вещественных блочных жордановых клеток.

2.21 Минимальный многочлен матрицы (оператора).

Опр. Многочлен минимальной степени аннулирующий матрицу A называется ее минимальным многочленом.

Лемма. Минимальный многочлен является делителем характеристического многочлена.

Д-во. Пусть $p(t)$ - минимальный многочлен для матрицы A , тогда $f_A(t) = q(t)p(t) + r(t)$, где $p(t), r(t) \in \mathbb{C}[t]$, $r(t) = 0$ или $\deg r(t) < \deg p(t)$. Далее, $f_A(t) = 0$ и $p(A) = 0$, значит $r(A) = 0$, т.е. $r(t)$ — аннулирующий многочлен и неравенство $\deg r(t) < \deg p(t)$ противоречит минимальности $p(t)$. Значит $r(t) = 0$ и $p(t)$ является делителем $f_A(t)$. \square

Лемма. Минимальные многочлены подобных матриц совпадают.

Д-во. Сразу следует из того, что если $B = C^{-1}AC$, то $p(B) = C^{-1}p(A)C$. \square

Следствие. Минимальный многочлен A совпадает с минимальным многочленом ее жордановой формы.

Теорема. Пусть матрица A имеет попарно различные собственные значения $\lambda_1, \dots, \lambda_m$. Тогда степень ее минимального многочлена равна сумме $n_1 + \dots + n_m$, где n_i — максимальный порядок жордановых клеток для собственного значения λ_i .

Д-во. Достаточно рассмотреть разложение произвольного вектора x по инвариантным подпространствам L_j максимального расщепления. Пусть подпространства L_{j_1}, \dots, L_{j_m} отвечают собственным значениям $\lambda_1, \dots, \lambda_m$ и имеют размерности n_1, \dots, n_m . Тогда подпространства $\ker(A - \lambda I)^{n_i}$ является корневым подпространством собственного значения $\lambda_i \implies (A - \lambda_i I)^{n_1} \dots (A - \lambda_m I)^{n_m} = 0$. Степень минимального многочлена не выше $n_1 + \dots + n_m$. В то же время, степень минимального многочлена не может быть меньше. Жорданова клетка порядка n_i для λ_i не может быть аннулирована многочленом степень меньше n_i , при этом ее минимальный многочлен есть в точности $(\lambda_i - \lambda)^{n_i}$ и этот многочлен не может аннулировать ни одну из жордановых клеток, отвечающих другому собственному значению. \square

Лемма. Пусть A — линейный оператор на пространстве над произвольным полем, многочлены $f_1(\lambda)$ и $f_2(\lambda)$ над этим полем взаимно просты и являются минимальными для сужений $A_1 = A|_{L_1}$ и $A_2 = A|_{L_2}$, где $L_1 = \ker f_1(A)$, $L_2 = \ker f_2(A)$. Тогда минимальный многочлен для прямой суммы $A_1 \oplus A_2$ есть произведение $f_1(\lambda)f_2(\lambda)$.

Д-во. Пусть A действует на пространстве $L = L_1 \oplus L_2$. Пусть $L \ni x = x_1 + x_2$ — разложение на подпространства L_1 и L_2 . Тогда $f_1(A)f_2(A)x_1 = f_1(A)f_2(A)x_2 = 0 \implies f_1(A)f_2(A) = O$. Если $\varphi(\lambda)$ — многочлен, аннулирующий A , то он аннулирует также A_1 и $A_2 \implies$ делится на $f_1(\lambda)$ и $f_2(\lambda) \implies$ на их произведение $f_1(\lambda)f_2(\lambda)$. \square

2.22 Условие совпадения минимального и характеристического многочленов.

Опр. V — линейное пространство над \mathbb{C} , $x \in V$, $p(t) \in \mathbb{C}[t]$. $p(t)$ называется аннулирующим многочленом вектора x относительно оператора A , если $p(A)x = \theta$. Аннулирующий многочлен минимальной степени со старшим коэффициентом 1 называется минимальным многочленом вектора x .

Лемма. Минимальный многочлен вектора является делителем минимального многочлена оператора.

Д-во. Аналогично доказательству леммы из вопроса 2.21. □

Исходя из теоремы Гамильтона-Кели, характеристический многочлен матрицы A порядка n совпадает с минимальным многочленом тогда и только тогда, когда степень минимального многочлена равна n . Поэтому, если существует вектор x такой, что система векторов Крылова $x, Ax, \dots, A^{n-1}x$ — ЛНЗ, то минимальный многочлен совпадает с характеристическим. Действительно, пусть $p(t) = a_0 + a_1t + \dots + a_{n-1}t^{n-1}$. Тогда $p(A)x = a_0Ix + a_1Ax + \dots + a_{n-1}A^{n-1}x$. Если $p(A)x = 0$, то в силу ЛНЗ: $a_0 = a_1 = \dots = a_{n-1} = 0$, т.е. никакой многочлен степени ниже n не может быть аннулирующим для A .

3 Линейные операторы в пространствах со скалярным произведением.

3.1 Существование, линейность, единственность сопряжённого оператора.

Опр. Пусть V и W — комплексные линейные пространства со скалярными произведениями $(\cdot, \cdot)_V$ и $(\cdot, \cdot)_W$ и $A : V \rightarrow W$ произвольный оператор. Оператор $A^* : W \rightarrow V$ называется сопряженным оператору A , если $(Ax, y)_W = (x, A^*y)_V \quad \forall x \in V \quad \forall y \in W$.

Утверждение. Если оператор обладает сопряженным, то он и его сопряженный оператор являются линейными.

Д-во. Пусть $u, v \in V$, $\alpha, \beta \in \mathbb{C}$. Тогда $(A(\alpha u + \beta v), y)_W = (\alpha u + \beta v, A^*y)_V = \alpha(u, A^*y)_V + \beta(v, A^*y)_V = \alpha(Au, y)_W + \beta(Av, y)_W = (\alpha Au + \beta Av, y)_W = 0 \implies (A(\alpha u + \beta v) - \alpha Au - \beta Av, y)_W = 0 \quad \forall y \in W$. Для вектора $y = A(\alpha u + \beta v) - \alpha Au - \beta Av : (y, y)_W = 0 \implies y = 0$. Аналогично показывается линейность A^* . □

Теорема. Если линейный оператор существует, то он единственный.

Д-во. Пусть $A : V \rightarrow W$, $A_1^*, A_2^* : W \rightarrow V$ и $(Ax, y)_W = (x, A_1^*y)_V = (x, A_2^*y)_V \quad \forall x \in V, y \in W$. Значит $(x, A_1^*y - A_2^*y)_V = 0$. Взяв $x = A_1^*y - A_2^*y$, получим: $(A_1^*y - A_2^*y, A_1^*y - A_2^*y)_V = 0 \quad \forall y \in W \implies A_1^*y = A_2^*y \quad \forall y \in W \implies A_1^* = A_2^* = A^*$. □

Если пространства V и W конечномерные, то в них можно выбрать ортонормированные базисы. Пусть $[u]$ и $[v]$ — векторы-столбцы из координат векторов $u, v \in V$. Аналогично, $[p], [q]$ — векторы-столбцы из координат векторов $p, q \in W$. Тогда скалярные произведения принимают вид:

$$\begin{aligned}(u, v)_V &= [v]^*[u] \\ (p, q)_W &= [q]^*[p]\end{aligned}$$

Теорема. Если пространства V и W конечномерны, то для любого линейного оператора $A : V \rightarrow W$ сопряженный оператор существует и единственен.

Д-во. Учитывая ортонормированность базисов, находим

$$\begin{aligned}(Ax, y)_W &= (x, A^*y)_V \Leftrightarrow [y]^*[Ax] = [A^*y][x] \Leftrightarrow [y]^*[A][x] = [y]^*[A^*]^*[x] \Leftrightarrow \\ &\Leftrightarrow [A] = [A^*]^* \Leftrightarrow [A^*] = [A]^*\end{aligned}$$

При использовании ортонормированных базисов линейный оператор A^* удовлетворяет условиям сопряженного оператора в том и только в том случае, когда он задается матрицей $[A^*] = [A]^*$. \square

Свойства сопряженного оператора.

1. $(A^*)^* = A$;
2. $(A + B)^* = A^* + B^*$;
3. $(\alpha A)^* = \bar{\alpha} A^*$;
4. $(AB)^* = B^* A^*$;
5. $(A^{-1})^* = (A^*)^{-1}$.

Д-во. 1) $A \in L(V, W) \implies A^* \in L(W, V) \implies (A^*)^* \in L(V, W)$ и $\forall x \in V, y \in W$: $(y, (A^*)^*x)_W = (A^*y, x)_V = \overline{(x, A^*y)_V} = \overline{(Ax, y)_W} = (y, Ax)_W$. Возьмем $y = Ax - (A^*)^*x$, тогда $\forall x \in V : (Ax - (A^*)^*x, Ax - (A^*)^*x) = 0 \implies A = (A^*)^*$.

2) $\forall x \in V, y \in W$:

$$\begin{aligned}(x, (A + B)^*y)_V &= ((A + B)x, y)_W = (Ax, y)_W + (Bx, y)_W = \\ &= (x, A^*y)_V + (x, B^*y)_V = \\ &= (x, A^*y + B^*y)_V\end{aligned}$$

Значит, $\forall y \in W : (A + B)^*y = A^*y + B^*y$.

3) $\forall x \in V, y \in W, \alpha \in \mathbb{C}$:

$$\begin{aligned}(x, (\alpha A)^*y)_V &= (\alpha Ax, y)_W = \alpha(Ax, y)_W = \alpha(x, A^*y)_V = \\ &= (x, \bar{\alpha} A^*y)_W\end{aligned}$$

Значит, $\forall y \in W : (\alpha A)^* y = \bar{\alpha} A^* y$.

4) $\forall B \in L(V, W), A \in L(W, K), x \in V, y \in W$:

$$(x, B^* A^* y)_V = (Bx, A^* y)_W = (ABx, y)_K = (x, (AB)^* y)_V$$

Значит, $\forall y \in K : B^* A^* y = (AB)^* y$.

5) $A \in L(V, W), A^{-1} \in L(W, V), (A^{-1})^* \in L(V, W)$. Если $\exists (A^*)^{-1}$, то $(A^*)^{-1} \in L(V, W)$, т.к. $A^* \in L(W, V)$.

Поскольку $\exists A^{-1}$, то A — изоморфизм между V и W и $\forall y \in W \exists! x \in V : y = Ax$. Пусть $A^* y = \theta$, тогда $\forall x \in V : 0 = (x, A^* y)_V = (Ax, y)_W$. Положим $x : Ax = y$, получаем, $(y, y)_W = 0$, т.е. $y = \theta$, а значит $\ker A^* = \{\theta\}$, т.е. существует $(A^*)^{-1}$. Далее $\forall x \in W, y \in V$:

$$\begin{aligned} (x, (A^{-1})^* y)_W &= (A^{-1} x, y)_V = (A^{-1} x, A^* (A^*)^{-1} y)_V = (AA^{-1} x, (A^*)^{-1} y)_W = \\ &= (x, (A^*)^{-1} y)_W \end{aligned}$$

Значит, $\forall y \in V : (A^{-1})^* y = (A^*)^{-1} y$. □

3.2 Матрицы оператора и сопряжённого к нему в паре биортогональных базисов.

Опр. Два ортогональных базиса e_1, \dots, e_n и f_1, \dots, f_n пространства V называются парой биортогональных базисов, если $(e_i, f_j) = \delta_{ij}$.

Теорема. В паре биортогональных базисов e и f унитарного пространства V матрицы операторов A и A^* связаны соотношением $(A^*)_f = (A_e)^H$.

Д-во. Пусть $A_e = [a_{ij}]$, $(A^*)_f = [b_{ij}]$. Тогда $Ae_j = \sum_{k=1}^n a_{kj} e_k$, $A^* f_i = \sum_{k=1}^n b_{ki} f_k$. Умножив первое их равенств скалярно на f_i получим, что $(Ae_j, f_i) = \sum_{k=1}^n a_{kj} (e_k, f_i) = a_{ij}$. С другой стороны, $(Ae_j, f_i) = (e_j, A^* f_i) = \sum_{k=1}^n \bar{b}_{ki} (e_j, f_k) = \bar{b}_{ji}$. Следовательно, $a_{ij} = \bar{b}_{ji}$. □

3.3 Критерии нормальности оператора (матрицы).

Опр. Линейный оператор $A \in L(V, V)$ называется нормальным, если $A^* A = A A^*$.

При выборе ортонормированного базиса сопряженный оператор определяется сопряженной матрицей, поэтому условие нормальности оператора A равносильно условию нормальности его матрицы в ортонормированном базисе: $[A]^* [A] = [A] [A]^*$.

Теорема (Критерий нормальности). Оператор A нормален тогда и только тогда, когда существует ортонормированный базис, состоящий из собственных векторов A .

Д-во. (\implies) Пусть A — нормальный оператор и e — ее базис Шура. Тогда A_e — верхняя треугольная матрица и A_e^* — нижняя треугольная матрица. Сравнив диагональные элементы матриц расположенных в левой и правой части равенства: $A_e A_e^* = A_e^* A_e$. Получим, равенства:

$$\begin{aligned} |a_{11}|^2 + |a_{12}|^2 + \dots + |a_{1n}|^2 &= |a_{11}|^2, \\ |a_{22}|^2 + \dots + |a_{2n}|^2 &= |a_{22}|^2, \\ &\dots \\ |a_{n-1,n-1}|^2 + |a_{n-1,n}|^2 &= |a_{n-1,n-1}|^2, \\ |a_{nn}|^2 &= |a_{nn}|^2, \end{aligned}$$

из которых следует, что $a_{12} = a_{13} = \dots = a_{1n} = a_{24} = a_{25} = \dots = a_{2n} = \dots = a_{n-1,n} = 0$. Следовательно, матрица A_e имеет диагональную форму. Таким образом, базис Шура является ортонормированным базисом из собственных векторов оператора A .

(\Leftarrow) Пусть e — базис из собственных векторов оператора A , тогда $A_e = \text{diag}(\alpha_1, \dots, \alpha_n)$, $A_e^* = \text{diag}(\bar{\alpha}_1, \dots, \bar{\alpha}_n)$. Из перестановочности диагональных матриц следует, что A_e — нормальна матрица. Это означает, что A — нормальный оператор. \square

3.4 Критерии унитарности и эрмитовости оператора (матрицы).

Опр. *Линейный оператор A называется унитарным, если он обратим $A^{-1} = A^*$, т.е. $AA^* = A^*A = I$.*

Теорема. *Нормальный оператор унитарен тогда и только тогда, когда все его собственные значения по модулю равны 1.*

Д-во. (\implies) Пусть U — унитарен, $Ux = \lambda x$, $||x|| = 1$. Тогда $1 = (x, x) = (x, U^*Ux) = (Ux, Ux) = (\lambda x, \lambda x) = \lambda \bar{\lambda} (x, x) = |\lambda|^2$.

(\Leftarrow) U — нормален $\implies \exists e$ — ортонормированный базис в V , в котором $U_e = \text{diag}(\lambda_1, \dots, \lambda_n)$ и $U_e^* = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n)$. Тогда $U_e U_e^* = U_e^* U_e = \text{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2) = I$. \square

Теорема. U — унитарен $\Leftrightarrow \forall x, y \in V : (x, y) = (Ux, Uy)$.

Д-во. (\implies) $\forall x, y \in V : (x, y) = (x, Iy) = (x, U^*Uy) = (Ux, Uy)$.

(\Leftarrow) $\forall x, y \in V : (x, Iy) = (x, y) = (Ux, Uy) = (x, U^*Uy) \implies U^*U = I$. Значит U — невырожденный и $U^* = U^{-1}$. \square

Теорема. U — унитарен \Leftrightarrow переводит любой ортонормированный базис в ортонормированный базис.

Д-во. (\implies) Пусть U — унитарен, e_1, \dots, e_n — ортонормированная система векторов. Значит, $(e_i, e_j) = \delta_{ij}$. Тогда $(Ue_i, Ue_j) = (e_i, e_j) = \delta_{ij}$.

(\Leftarrow) Пусть e_1, \dots, e_n — ОНБ в V , тогда Ue_1, \dots, Ue_n — ОНБ в V . $\forall x, y \in V$:

$$\begin{aligned} x &= x_1 e_1 + \dots + x_n e_n & Ux &= x_1 Ue_1 + \dots + x_n Ue_n \\ y &= y_1 e_1 + \dots + y_n e_n & Uy &= y_1 Ue_1 + \dots + y_n Ue_n \end{aligned}$$

Значит, $(x, y) = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n = (Ux, Uy)$. \square

Теорема. U — унитарен $\Leftrightarrow \forall x \in V : \|Ux\| = \|x\|$.

Д-во. (\Rightarrow) $\|x\| = \sqrt{(x, x)} = \sqrt{(Ux, Ux)} = \|Ux\|$.

(\Leftarrow) $(x, x) = \|x\|^2 = \|Ux\|^2 = (Ux, Ux)$. \square

Опр. Линейный оператор A называется эрмитовым, если $A = A^*$.

Теорема. Нормальный оператор эрмитов \Leftrightarrow все его собственные значения вещественные.

Д-во. (\Rightarrow) Пусть $A = A^*$, $Ax = \lambda x$, $\|x\| = 1$. Тогда

$$\lambda = \lambda \cdot 1 = \lambda(x, x) = (\lambda x, x) = (Ax, x) = (x, A^* x) = (x, Ax) = (x, \lambda x) = \bar{\lambda}(x, x) = \bar{\lambda}$$

(\Leftarrow) Пусть e_1, \dots, e_n — ОНБ из собственных векторов A . Тогда $A_e = \text{diag}(\lambda_1, \dots, \lambda_n)$, $A_e^* = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \Rightarrow A_e = A_e^* \Rightarrow A = A^*$. \square

3.5 Эрмитово разложение оператора (матрицы) и эрмитовость знакоопределенного оператора в унитарном пространстве.

Теорема. Любой линейный оператор $A \in L(V, V)$ допускает однозначно определенное эрмитово разложение: $A = H_1 + iH_2$, H_1, H_2 — эрмитовы операторы.

Д-во. Если $A = H_1 + iH_2$ и $H_1 = H_1^*$, $H_2 = H_2^*$, то $A^* = H_1^* - iH_2^* = H_1 - iH_2$. Значит $H_1 = \frac{1}{2}(A + A^*)$ и $H_2 = \frac{i}{2}(A^* - A)$. \square

Теорема. Если в унитарном пространстве $V : \forall x \in V : (Bx, x) = 0$, то $B = O$.

Д-во. Для любых векторов $y, z \in V$:

$$\begin{aligned} 0 &= (B(y + z), y + z) = (By, y) + (Bz, y) + (By, z) + (Bz, z) = (Bz, y) + (By, z) \\ 0 &= (B(iy + z), iy + z) = (By, y) - i(Bz, y) + i(By, z) + (Bz, z) = -i(Bz, y) + i(By, z) \end{aligned}$$

Прибавим к первому равенству второе, умноженное на $-i$, получим, что $(By, z) = 0 \forall y, z \in V \Rightarrow B = O$. \square

Теорема. Линейный оператор A в унитарном пространстве V эрмитов $\Leftrightarrow (Ax, x) \in \mathbb{R}$.

Д-во. (\implies) Пусть A — эрмитов оператор. Тогда $(Ax, x) = (x, A^*x) = (x, Ax) = \overline{(Ax, x)} \in \mathbb{R}$.

(\impliedby) Пусть $(Ax, x) \in \mathbb{R}$. Тогда $(x, Ax) = \overline{(x, Ax)} = (Ax, x) = (x, A^*x) \implies (x, (A - A^*)x) = 0 \forall x \in V \implies A = A^*$. \square

Опр. Эрмитов линейный оператор $A : V \rightarrow V$ называется неотрицательно определенным, если $\forall x \in V : (Ax, x) \geq 0$, и положительно определенным, если $\forall \theta \neq x \in V : (Ax, x) > 0$.

Теорема. Оператор A — неотрицателен (положителен) $\Leftrightarrow A$ эрмитов с неотрицательными (неположительными) собственными значениями.

Д-во. (\implies) Пусть $Ax = \lambda x$, $x \neq \theta$. Тогда $0 \leq (Ax, x) = \lambda(x, x)$. Значит $\lambda \geq 0$.

(\impliedby) Пусть e_1, \dots, e_n — ОНБ, $Ae_k = \lambda e_k$, $k = \overline{1, n}$ и $x = \sum_{i=1}^n x_i e_i$. Тогда $(Ax, x) = \left(A \sum_{k=1}^n x_k e_k, \sum_{m=1}^n x_m e_m \right) = \sum_{k,m=1}^n \lambda_k x_k \overline{x_m} (e_k, e_m) = \sum_{k=1}^n \lambda_k |x_k|^2$. Если $\lambda_k \geq 0$ (> 0), то $(Ax, x) \geq 0$ (> 0). \square

3.6 Существование и единственность неотрицательно определенного квадратного корня из неотрицательно определенной матрицы.

Теорема. Для любого неотрицательного определенного оператора A существует, и притом единственный, неотрицательно определенный оператор B такой, что $B^2 = A$.

Д-во. *Существование.* Пусть e_1, \dots, e_n — ОНБ в V , тогда $Ae_k = \lambda e_k$, $\lambda_k \geq 0$, $k = \overline{1, n}$. Пусть $B \in L(V, V) : Be_k = \sqrt{\lambda_k} e_k$, $k = \overline{1, n}$. Тогда:

- 1) B — нормален, т.к. e_1, \dots, e_n — ОНБ из собственных векторов;
- 2) B — эрмитов, т.к. его собственные значения $\sqrt{\lambda_k} \in \mathbb{R}$, $k = \overline{1, n}$;
- 3) $B \geq 0$, т.к. его собственные значения $\sqrt{\lambda_k} \geq 0$, $k = \overline{1, n}$;
- 4) $B^2 e_k = B(Be_k) = B(\sqrt{\lambda_k} e_k) = \sqrt{\lambda_k} Be_k = \lambda_k e_k = Ae_k$, $k = \overline{1, n} \implies B^2 = A$.

Единственность. Пусть существует другой оператор $C \geq 0$, т.ч. $C^2 = A$. Тогда существует ортонормированный базис f_1, \dots, f_n пространства из собственных векторов оператора C . Если $Cf_i = \mu_i f_i$, $i = \overline{1, n}$, то $Af_i = C^2 f_i = \mu_i^2 f_i$, $i = \overline{1, n}$. Значит, числа μ_1^2, \dots, μ_n^2 являются собственными значениями оператора A и совпадают с числами $\lambda_1, \dots, \lambda_n$.

Разложим вектор e_i по базису f : $e_i = \sum_{k=1}^n \alpha_k f_k$. Из линейной независимости собственных векторов отвечающих различным собственным значениям следует, что отличными от нуля будут только коэффициенты α_k лишь при тем f_k , который отвечают собственному значению $\mu_k^2 = \lambda_i$. Поэтому $Ce_i = \sum_{k=1}^n \alpha_k C f_k = \sum_{k=1}^n \alpha_k \sqrt{\lambda_i} f_k = \sqrt{\lambda_i} \sum_{k=1}^n \alpha_k f_k = \sqrt{\lambda_i} e_i$, $i = \overline{1, n} \implies C = B$. \square

3.7 Блочно-диагональная форма вещественной нормальной матрицы.

Теорема. Для любого вещественного нормального оператора существует e — ОНБ, в котором его матрица является прямой суммой вещественных блоков порядка 1: $[\lambda]$ и вещественных блоков порядка 2: $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

Д-во. A — нормален \implies все его жордановы клетки имеют порядок 1 (т.к. существует базис из собственных векторов). Пусть $\lambda = a + ib$, $a, b \in \mathbb{R}$, $b \neq 0$ и $f_A(\lambda) = 0$. Тогда $f_A(\bar{\lambda}) = 0$ и $A(x + iy) = (a + ib)(x + iy) = (ax - by) + i(bx + ay)$, т.е. $Ax = ax - by$, $Ay = bx + ay$, где x, y — ненулевые вещественные векторы: система x, y — ЛНЗ (т.к. система $x + iy, x - iy$ — ЛНЗ). Значит $A(L(x, y)) \subset L(x, y)$ и сужение A на $L(x, y)$ в базисе x, y имеет матрицу $A_{x,y} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

Заметим, что собственные векторы нормального оператора, отвечающие различным собственным значениям, ортогональны. Действительно, если A нормальный то собственные векторы A и A^* совпадают. Далее, пусть $Ae = \lambda e$, $Af = \mu f$, $\lambda \neq \mu$, $e \neq \theta$, $f \neq \theta$. Тогда $A^*f = \bar{\mu}f$ и $(Ae, f) = (\lambda e, f) = \lambda(e, f) = (e, A^*f) = (e, \bar{\mu}f) = \mu(e, f) \implies (\lambda - \mu)(e, f) = 0 \implies (e, f) = 0$. Значит векторы $x + iy$ и $x - iy$ ортогональны, т.е.

$0 = (x + iy, x - iy) = (x, x) - (y, y) + 2i(x, y) \Leftrightarrow \begin{cases} (x, x) = (y, y) \\ (x, y) = 0 \end{cases}$. Получили, что если $t = \|x\| = \|y\| > 0$, то $e = \frac{x}{t}$, $f = \frac{y}{t}$: e, f — ОНБ в $L(x, y)$ и $A_{e,f} = A_{x,y}$. \square

3.8 Блочно-диагональная форма ортогональной матрицы. Матрицы вращения и отражения.

Опр. Пусть V — вещественное конечномерное линейное пространство, $A \in L(V, V)$. Оператор A называется ортогональным, если $AA^* = A^*A = I$.

Теорема. Пусть $A \in L(V, V)$ — ортогональный оператор. Тогда существует e_1, \dots, e_n — ОНБ, в котором A_e является прямой суммой вещественных матриц порядка 1 с элементами ± 1 и вещественных матриц вращения порядка 2.

Д-во. $AA^* = I = A^*A \implies A$ — нормален. Значит $\exists e$ — ОНБ, в котором матрица A_e — есть прямая сумма вещественных матриц порядка 1 : $[\lambda]$ и блоков порядка 2 : $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. Далее, A — ортогонален \implies по модулю собственные значения A равны 1. Значит блоки порядка 1 имеют вид $[1]$ или $[-1]$. Блоки порядка 2 отвечают парам собственных значений $\lambda = a + ib$, $\bar{\lambda} = a - ib$, $a, b \in \mathbb{R}$, $b \neq 0$. Так как $|\lambda| = |\bar{\lambda}| = 1$, то $a^2 + b^2 = 1$. Значит существует $\varphi \in (-\pi, \pi]$, т.ч. $a = \cos \varphi$, $b = -\sin \varphi$ и блок второго порядка имеет вид: $\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$. \square

Матрица оператора вращения.

$$R_{k,m}(\varphi) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & \cos \varphi & \dots & -\sin \varphi & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sin \varphi & \dots & \cos \varphi & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{matrix} \leftarrow k \\ \leftarrow m \end{matrix}$$

$\begin{matrix} \uparrow & \uparrow \\ k & m \end{matrix}$

$R_{k,m}(\varphi)$ — матрица Гивенса.

Пусть в ОНБ e матрица оператора $A_e = P_{k,k+1}(\varphi)$. $x = \sum_{i=1}^n x_i e_i$. Пусть $x' = x_k e_k + x_{k+1} e_{k+1}$, $z = x_1 e_1 + \dots + x_{k-1} e_{k-1} + x_{k+2} e_{k+2} + \dots + x_n e_n$, $x = z + x'$, $x' \in L(e_k, e_{k+1})$, $z \in L^\perp(e_k, e_{k+1})$. $y' \in L(e_k, e_{k+1})$ и y' получен из x' поворотом на угол φ ; $y = Ax = z + y'$. Такой оператор называют оператором простого поворота.

Матрица оператора отражения.

$$Q_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \leftarrow k$$

$\begin{matrix} \uparrow \\ k \end{matrix}$

Q_k — матрица отражения.

3.9 Полярное разложение оператора (матрицы).

Опр. Представление комплексной квадратной матрицы в виде $A = HQ$, где H — неотрицательно определенная, а матрица Q — унитарная, называется её полярным разложением.

Теорема. Любая квадратная комплексная матрица A имеет полярное разложение $A = HQ$ и при этом матрица H определена однозначно. Если A невырождена, то и матрица Q тоже определяется однозначно.

Д-во. Из сингулярного разложения $A = V\Sigma U^*$, получаем $A + (V\Sigma V^*)(VU^*)$. Положим $H = V\Sigma V^*$, $Q = VU^*$. Тогда $H^* = (V^*)^*\Sigma^*V^* = V\Sigma V^* = H$ ($\Sigma^* = \Sigma^T = \Sigma$), т.е. H — эрмитова и т.к. её собственные значения — это сингулярные числа матрицы A : $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0 \implies H \geq 0$. Матрица Q , как произведение унитарных, тоже унитарная: $Q^* = (U^*)^*V^* = (U^{-1})^*V^* = (U^*)^{-1}V^{-1} = Q^{-1}$. Единственность H следует из теоремы о квадратном корне: $A = HQ$, $A^* = Q^*H^* = Q^{-1}H$. Значит $AA^* = H^2$. Если A невырожденная, то A^* и AA^* — невырожденные. Тогда H — невырожденная. Значит матрица $Q = H^{-1}A$ определена однозначно. \square

3.10 Сингулярное разложение линейного оператора (матрицы). Три формы записи сингулярного разложения.

Теорема. Пусть V и W — унитарные пространства размерностей n и m , и $A : V \rightarrow W$ — произвольный линейный оператор. Тогда существуют ортонормированные базисы $u_1, \dots, u_n \in V$ и $v_1, \dots, v_m \in W$, в которых матрица линейного оператора имеет вид

$$\Sigma = \begin{bmatrix} \Sigma_r & O_{r \times (n-r)} \\ O_{(m-r) \times r} & O_{(m-r) \times (n-r)} \end{bmatrix} \quad \Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, \quad \sigma_1, \dots, \sigma_r > 0.$$

Д-во. Оператор A^*A эрмитов и, как любой нормальный оператор, обладает ортонормированным базисом из собственных векторов. Обозначим его через u_1, \dots, u_n . Пусть λ — произвольное собственное значение оператора A^*A и x — его собственный вектор. Тогда $A^*Ax = \lambda x \implies \lambda(x, x) = (A^*Ax, x) = (Ax, Ax) \implies \lambda \geq 0$. Все собственные значения оператора A^*A неотрицательные и с учетом кратностей образуется набор $\sigma_1^2, \dots, \sigma_n^2$, где $\sigma_1, \dots, \sigma_n > 0$. Будем считать, что $A \neq O$ и $\sigma_i > 0$ при $1 \leq i \leq r$ и $\sigma_i = 0$ при $r+1 \leq i \leq n$. Положим $v_i = \frac{1}{\sigma_i}Au_i$, $1 \leq i \leq r$, а в случае $r+1 \leq i \leq n$ заметим, что если $A^*Au_i = 0$, то $(A^*Au_i, u_i) = (Au_i, Au_i) = 0 \implies Au_i = 0$. Получаем: $Au_i = \begin{cases} \sigma_i u_i & , 1 \leq i \leq r \\ 0 & , r+1 \leq i \leq n \end{cases}$. Заметим, что векторы v_1, \dots, v_r образуют ортонормированную систему: $(v_i, v_j) = (\frac{1}{\sigma_i}Au_i, \frac{1}{\sigma_j}Au_j) = \frac{1}{\sigma_i\sigma_j}(A^*Au_i, u_j) = \frac{\sigma_i}{\sigma_j}(u_i, u_j) = \delta_{ij}$. Дополнив систему v_1, \dots, v_r до ортонормированного базиса v_1, \dots, v_m получаем:

$$Au_i = \begin{cases} \sum_{j=1}^{i-1} 0 \cdot v_j + \sigma_i v_i + \sum_{j=i+1}^m 0 \cdot v_j & , 1 \leq i \leq r \\ \sum_{j=1}^m 0 \cdot v_j & , r+1 \leq i \leq n \end{cases}$$

\square

Следствие. Для любой матрицы $A \in \mathbb{C}^{m \times n}$ существуют унитарные матрицы $U = [u_1, \dots, u_n] \in \mathbb{C}^{n \times n}$ и $V = [v_1, \dots, v_m] \in \mathbb{C}^{m \times m}$ такие, что $AU = V\Sigma$, $A^*V = U\Sigma^T$, $A = V\Sigma U^*$.

Получили три различные по форме представления комплексной $m \times n$ матрицы A :

$$A = V\Sigma U^* = V_r \Sigma_r U_r^* = \sum_{i=1}^r \sigma_i u_i^* v_i,$$

где $U_r = [u_1, \dots, u_r] \in \mathbb{C}^{n \times r}$, $V_r = [v_1, \dots, v_r] \in \mathbb{C}^{m \times r}$.

3.11 Вариационные свойства собственных значений эрмитовых матриц, теорема Куранта-Фишера. Вариационные свойства сингулярных чисел.

Вариационные свойства собственных значений эрмитовых матриц — свойства, связанные с минимизацией и максимизацией некоторой функции. В качестве такой функции используется отношение Гэлей: $f(x) = \frac{x^* Ax}{x^* x}$. Пусть $\lambda_1 \geq \dots \geq \lambda_n$ — собственные значения эрмитовой матрицы A . Тогда $\lambda_1 = \max_{x \neq \theta} \frac{x^* Ax}{x^* x}$, $\lambda_n = \min_{x \neq \theta} \frac{x^* Ax}{x^* x}$.

Теорема (Теорема Куранта-Фишера).

$$\lambda_k = \max_{\dim L=k} \min_{\theta \neq x \in L} \frac{x^* Ax}{x^* x} = \min_{\dim L=n-k+1} \max_{\theta \neq x \in L} \frac{x^* Ax}{x^* x}.$$

До-во. Пусть $L \subseteq \mathbb{C}^n$ — произвольное подпространство размерности k и M — линейная оболочка, натянутая на ортонормированные собственные векторы, отвечающие собственным значениям $\lambda_k \geq \lambda_{k+1} \geq \dots \geq \lambda_n$ матрицы A . Согласно теореме Грассмана, $\dim(L \cap M) = \dim L + \dim M - \dim(L \cup M) \geq k + (n - k + 1) - n = 1$. Поэтому $\exists \theta \neq z \in L \cap M$. Далее,

$$\min_{\theta \neq x \in L} \frac{x^* Ax}{x^* x} \leq \frac{z^* Az}{z^* z} \leq \max_{\theta \neq x \in M} \frac{x^* Ax}{x^* x} = \lambda_k \implies \max_{\dim L=k} \min_{\theta \neq x \in L} \frac{x^* Ax}{x^* x} \leq \lambda_k.$$

Равенство достигается на подпространстве, натянутом на ортонормированные векторы, отвечающие собственным значениям $\lambda_1 \geq \dots \geq \lambda_k$ \square

Вариационные свойства сингулярных чисел.

Пусть $A : L \rightarrow W$ — эрмитов оператор и $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ — его собственные значения. Тогда оператор $A^* A$ так же эрмитов: $(A^* A)^* = A^* (A^*)^* = A^* A$ и его собственные значения: $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2$, $\sigma_i^2 = \lambda_i \bar{\lambda}_i = |\lambda_i|^2 \in \mathbb{R}$. Тогда по теореме Куранта-Фишера:

$$\sigma_k^2 = \max_{\dim L=k} \min_{\theta \neq x \in L} \frac{x^* A^* A x}{x^* x} = \min_{\dim L=n-k+1} \max_{\theta \neq x \in L} \frac{x^* A^* A x}{x^* x},$$

или, аналогично:

$$\sigma_k = \max_{\dim L=k} \min_{\theta \neq x \in L} \frac{\|Ax\|}{\|x\|} = \min_{\dim L=n-k+1} \max_{\theta \neq x \in L} \frac{\|Ax\|}{\|x\|}.$$

3.12 Соотношения разделения собственных значений эрмитовой матрицы и её главной подматрицы. Соотношения разделения для сингулярных чисел.

Теорема. Собственные значения эрмитовой матрицы A порядка n разделяются собственными значениями её главной подматрицы порядка $n - 1$ B :

$$\lambda_i \geq \mu_i \geq \lambda_{i+1}, \quad 1 \leq i \leq n - 1,$$

где $\lambda_1 \geq \dots \geq \lambda_n$ — собственные значения A , $\mu_1 \geq \dots \geq \mu_{n-1}$ — собственные значения B .

Д-во. Обозначим через M подпространство векторов $x = [x_1, \dots, x_n]^T$, определяемое уравнением $x_n = 0$, и рассмотрим отображение $p : [x_1, \dots, x_{n-1}, x_n]^T \rightarrow [x_1, \dots, x_{n-1}, 0]^T$. Тогда $\forall x \in M : \frac{(Ax, x)}{(x, x)} = \frac{(Bp(x), p(x))}{(p(x), p(x))}$. Пусть $1 \leq k \leq n - 1$. Согласно теореме Куранта-Фишера:

$$\begin{aligned} \lambda_k &= \max_{\dim L=k} \min_{\theta \neq x \in L} \frac{(Ax, x)}{(x, x)} \geq \max_{\dim L=k, L \subset M} \min_{\theta \neq x \in L} \frac{(Ax, x)}{(x, x)} = \\ &= \max_{\dim L=k, L \subset M} \min_{\theta \neq x \in L} \frac{(Bp(x), p(x))}{(p(x), p(x))} = \max_{\dim L=k, L \subset \mathbb{C}^{n-1}} \min_{\theta \neq y \in L} \frac{(By, y)}{(y, y)} = \mu_k. \\ \lambda_k &= \min_{\dim L=k} \max_{\theta \neq x \in L} \frac{(Ax, x)}{(x, x)} \leq \min_{\dim L=k, L \subset M} \max_{\theta \neq x \in L} \frac{(Ax, x)}{(x, x)} = \\ &= \min_{\dim L=k, L \subset M} \max_{\theta \neq x \in L} \frac{(Bp(x), p(x))}{(p(x), p(x))} = \min_{\dim L=k, L \subset \mathbb{C}^{n-1}} \max_{\theta \neq y \in L} \frac{(By, y)}{(y, y)} = \mu_k \end{aligned}$$

□

Сингулярные числа матрицы и её подматрицы.

Пусть $A \in \mathbb{C}^{m \times n}$, $A = [Bv]$, $B \in \mathbb{C}^{m \times (n-1)}$, $v \in \mathbb{C}^m$. Тогда $A^* = \begin{bmatrix} B^* \\ v^* \end{bmatrix} \in \mathbb{C}^{n \times m}$, $A^*A = \begin{bmatrix} B^* \\ v^* \end{bmatrix} [Bv] = \begin{bmatrix} B^*B & B^*v \\ v^*B & v^*v \end{bmatrix} \in \mathbb{C}^{n \times n}$. Пусть $\sigma_1(A) \geq \dots \geq \sigma_n(A)$ — сингулярные числа A , $\sigma_1(B) \geq \dots \geq \sigma_{n-1}(B)$ — сингулярные числа B . Применим только что доказанную теорему, получим:

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \dots \sigma_{n-1}(A) \geq \sigma_{n-1}(B) \geq \sigma_n(A).$$

4 Билинейные и квадратичные формы.

4.1 Закон сохранения инерции эрмитовой матрицы при переходе к эрмитово конгруэнтной матрице.

Опр. Пусть $\pi_+(A), \pi_-(A), \pi_0(A)$ обозначают число положительных, отрицательных и нулевых собственных значений матрицы A . Тройку (π_+, π_-, π_0) принято называть

инерцией матрицы A . Числа π_+ и π_- называют положительным и отрицательным индексами инерции.

Теорема (закон инерции). Эрмитовы матрицы A и B эрмитово конгруэнтны тогда и только тогда, когда они имеют одну и ту же инерцию:

$$\pi_+(A) = \pi_+(B), \quad \pi_-(A) = \pi_-(B), \quad \pi_0(A) = \pi_0(B).$$

Д-во. Пусть Λ и M — диагональные матрицы, составленные из собственных значений $\lambda_1, \dots, \lambda_n$ и μ_1, \dots, μ_n матриц A и B . Ясно, что A и B эрмитово конгруэнтны в том и только в том случае, когда Λ и M эрмитово конгруэнтны. Таким образом, теорему достаточно доказать для вещественных диагональных матриц. Обозначим через k и l положительные индексы инерции для Λ и M , и будем считать, что $\lambda_1, \dots, \lambda_k > 0$ и $\mu_1, \dots, \mu_l > 0$. Пусть $\Lambda = P^*MP$ и $x = Py$. Тогда

$$x^*\Lambda x = \sum_{i=1}^k \lambda_i |x_i|^2 + \varphi(x) = y^*My = \sum_{i=1}^l \mu_i |y_i|^2 + \psi(y),$$

$$\varphi(x) \leq 0, \quad \psi(y) \leq 0 \quad \forall x = Py.$$

Допустим, что $k < l$. Рассмотрим два подпространства

$$L_1 = \{x \in \mathbb{C}^n : x_1 = \dots = x_k = 0\}, \quad L_2 = \{x = Py \in \mathbb{C}^n : y_{l+1} = \dots = y_n = 0\}.$$

Тогда, $\dim(L_1 \cap L_2) = \dim L_1 + \dim L_2 - \dim(L_1 + L_2) \geq (n-k) + l - n \geq 1$. Следовательно, существует ненулевой вектор $x \in L_1 \cap L_2$. Тогда $0 < y^*My = x^*\Lambda x \leq 0$, что, конечно, невозможно, так как $y = P^{-1}x \neq 0$. Поэтому $k \geq l$. Меняя k и l ролями находим $l \leq k$. Значит, $k = l$. Далее, равенство $\pi_0(\Lambda) = \pi_0(M)$ — это равенство дефектов этих матриц. Равенство $\pi_-(\Lambda) = \pi_-(M)$ — очевидное следствие уже доказанного.

Эрмитова конгруэнтность матриц с одинаковой инерцией следует из того, что каждая из них эрмитово конгруэнтна одной и той же диагональной матрице с элементами ± 1 и 0 . \square

4.2 Квадратичные и эрмитово квадратичные формы и их приведение к алгебраической сумме квадратов.

Под вещественной квадратичной формой от переменных x_1, \dots, x_n понимается однородный квадратичный многочлен с вещественными коэффициентами

$$f(x) = f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n f_{ij} x_i x_j = x^T F x.$$

В данном случае $f(x) = x^T A x$, где $A = \frac{F+F^T}{2}$ — вещественная симметрическая матрица. Под эрмитовой квадратичной формой понимается функция вида

$$f(x) = f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n f_{ij} \bar{x}_i x_j = x^* F x.$$

В данном случае $f(x) = x^*Ax$, где $A = \frac{F+F^*}{2}$ — эрмитова матрица. Таким образом, как вещественная, так и эрмитова квадратичная форма определяется некоторой эрмитовой матрицей, которая и называется матрицей квадратичной формы.

При замене переменных $x = Py$ матрица вещественной квадратичной формы $f(Py)$ имеет вид $B = P^TAP$, т.е. является конгруэнтной матрице A . Если матрица P ортогональна, то B называется ортогонально конгруэнтной матрице A . Матрица эрмитовой квадратичной формы $f(Py)$ имеет вид $B = P^*AP$ матрица такого вида называется эрмитово конгруэнтной матрице A . Если матрица P унитарна, то B называется унитарно конгруэнтной матрице A .

Задача о поиске диагональной матрицы, эрмитово конгруэнтной заданной эрмитовой матрице, решается за конечное число арифметических операций с помощью выделения полных квадратов (метод Лагранжа). Пусть $f(x) = x^*Ax$, $A = A^*$. Возможны два случая:

(а) На главной диагонали матрицы A есть ненулевой элемент, тогда симметрической перестановкой строк и столбцов его можно перевести в позицию $(1, 1)$. Если $a_{11} \neq 0$, то выделяем полный квадрат таким образом:

$$a_{11}x_1\bar{x}_1 + a_{1i}\bar{x}_1x_i + \bar{a}_{1i}\bar{x}_ix_1 + a_{ii}x_i\bar{x}_i = a_{11} \left| x_1 + \frac{a_{1i}}{a_{11}}x_i \right|^2 + \left(a_{ii} - \frac{|a_{1i}|^2}{a_{11}} \right) |x_i|^2.$$

(б) Главная диагональ матрицы A нулевая. Если $a_{1i} \neq 0$, то действуем так:

$$a_{1i}\bar{x}_1x_i + \bar{a}_{1i}\bar{x}_ix_1 = \frac{1}{2}(|x_1 + a_{1i}x_i|^2 - |x_1 - a_{1i}x_i|^2).$$

Таким образом, от A можно перейти к эрмитово конгруэнтной матрице, в которой первый столбец и первая строка имеют нули в недиагональных позициях. Дальше то же самое делаем с эрмитовой подматрицей, которая остается после вычеркивания первой строки и первого столбца.

4.3 Общий вид линейного функционала и эрмитовой билинейной формы в конечномерном пространстве со скалярным произведением.

Опр. *Линейный оператор $f : V \rightarrow W$ в случае одномерного пространства W называется линейным функционалом.*

Теорема. *Пусть V — конечномерное унитарное пространство, $f \in L(V, \mathbb{C})$, $f \neq O$. Тогда $\dim(\ker f)^\perp = 1$.*

Д-во. Если $f \neq O$, то $\ker f \neq V$ и $V = \ker f \oplus (\ker f)^\perp$, $(\ker f)^\perp \neq \{\theta\}$. Значит $\forall x \in V \exists! y \in \ker f, z \in (\ker f)^\perp : x = y + z$. Пусть $x_1, x_2 \in (\ker f)^\perp$ и $x = f(x_2)x_1 - f(x_1)x_2 \in (\ker f)^\perp$, но $f(x) = f(x_2)f(x_1) - f(x_1)f(x_2) = 0 \implies x \in \ker f$, т.е. $x \in \ker f \cap (\ker f)^\perp \Leftrightarrow x = \theta$. Поэтому $\dim(\ker f)^\perp = 1$. \square

Теорема. Пусть V — конечномерное унитарное пространство, $f \in L(V, \mathbb{C})$. Тогда существует и однозначно определен вектор $h \in V : \forall x \in V f(x) = (x, h)_V$.

Д-во. 1) (существование) Если $f = O$, то $h = \theta$. Пусть $f \neq O$ и $e \in (\ker f)^\perp : \|e\| = \sqrt{(e, e)_V} = 1$, $(\ker f)^\perp = L(e)$. Далее, поскольку $V = \ker f \oplus L(e)$, то $\forall x \in V \exists! x' \in \ker f$, $\lambda \in \mathbb{C} : x = x' + \lambda e$, где $\lambda = (x, e)_V$. Значит, $f(x) = f(x') + f(\lambda e) = \lambda f(e) = (x, e)_V f(e) = (x, \overline{f(e)}e)_v$, т.е. $h = \overline{f(e)}e$.

2) (единственность) Пусть $h_1, h_2 \in V : \forall x \in V f(x) = (x, h_1)_V = (x, h_2)_V \implies (x, h_1 - h_2)_v = 0 \forall x \in V \implies h_1 = h_2 = h$. \square

Опр. Пусть V — комплексное линейное пространство. Отображение $A : V \times V \rightarrow \mathbb{C}$ называется эрмитовой билинейной формой в V , если $\forall x, y, z \in V, \alpha \in \mathbb{C} :$

1. $A(x + y, z) = A(x, z) + A(y, z);$
2. $A(\alpha x, y) = \alpha A(x, y);$
3. $A(x, y + z) = A(x, y) + A(x, z);$
4. $A(x, \alpha y) = \alpha A(x, y);$
5. $f(x, y) = \overline{f(y, z)}.$

Теорема. Пусть V — конечномерное унитарное пространство, $B(x, y)$ — эрмитова билинейная форма в V . Тогда существует и единственен линейный оператор $A \in L(V, V)$ такой, что $\forall x, y \in V : B(x, y) = (x, Ay)_V$.

Д-во. 1) (существование) Фиксируем $y \in V$, тогда $B(x, y) \in L(v, \mathbb{C})$. Значит $\exists! h = h(y) \in V : \forall x \in V B(x, y) = (x, h(y))_V$. Отображение $A : V \rightarrow V : Ay = h(y)$ является линейным оператором, т.е. $A \in L(V, V)$. Действительно, $\forall y_1, y_2 \in V, \alpha_1, \alpha_2 \in \mathbb{C} :$

$$\begin{aligned} (x, A(\alpha_1 y_1 + \alpha_2 y_2)) &= B(x, \alpha_1 y_1 + \alpha_2 y_2) = \bar{\alpha}_1 B(x, y_1) + \bar{\alpha}_2 B(x, y_2) = \\ &= \bar{\alpha}_1 (x, Ay_1) + \bar{\alpha}_2 (x, Ay_2) = (x, \alpha_1 Ay_1 + \alpha_2 Ay_2). \end{aligned}$$

Значит $A(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 Ay_1 + \alpha_2 Ay_2 \implies A \in L(V, V)$. \square

2) (единственность) Пусть $A_1, A_2 \in L(V, V) : \forall x, y \in V B(x, y) = (x, A_1 y) = (x, A_2 y) \implies (x, (A_1 - A_2)y) = 0 \implies A_1 = A_2 = A$.