Содержание

1 Линейные пространства. Пространства со скалярным произведением. 3

1.1 Неравенство Коши-Бунековского-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

1 Линейные пространства. Пространства со скалярным произведением.

1.1 Неравенство Коши-Бунековского-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 произведение 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 произведение 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 $F(t) = (x + t\xi y, x + t\xi y) = (x, x) + t\xi\overline{(x, y)} + t\overline{\xi}(x, y) + t^2\xi\overline{\xi}(y, y) = t^2|y|^2 + 2t|(x, y)| + |x|^2$. В силу свойств скалярного произведения $F(t) \geq 0$ при всех вещественных t. Значит $D \leq 0$, $D = |(x, y)|^2 - |x|^2|y|^2 \leq 0 \implies |(x, y)| \leq |x||y|$. Равенство означает, что $D = 0 \implies (x + t\xi y, x + t\xi y) = 0 \implies x + t\xi y = 0$. □

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

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

Д-во. Положим, что $p_1 = a_1 \implies L(p_1) = L(a_1)$. Предположим, что уже постоена ортогональная система $p_1, \ldots, p_{k-1}$ такая, что $L(p_1, \ldots, p_i) = L(a_1, \ldots, a_i)$ при $1 \le i \le k-1$. 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 $p_1, \ldots, p_{k-1}$:

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

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

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

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

Этот алгоритм называется процессом ортогонализации Грама-Шмидта.

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

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

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 $\alpha_k \to 0$, 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 $Q_{k_l} \to Q$. Матрица Q будет, конечно, унитарной, а предел последовательности $R_{k_l} = Q^*_{k_l} A_{k_l} \to Q^* A$ является, очевидно, верхней треугольной матрицей. □

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*

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

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 $c := z - z' = h' - h \in L \cap L^{\perp} \implies v = 0$.
Наконец, для любого $y \in L$ находим $x - y = (z - y) + h$, и, согласно теореме Пифагора, $|x - y|^2 = |z - h|^2 + |h|^2 \ge |h|^2$. Равенство, очевидно, имеет место в том и только в том случае, когда $y = z$. □

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 линейны 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. □

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 $v_1, \ldots, v_k$ и x - вектор столбец с элементами $x_1, \ldots, x_k$. Тогда $x^*Ax = \sum\limits_{i,j=1}^{k} \overline{x}_i a_{ij} x_j = \sum\limits_{i,j=1}^{k} \overline{x}_i (v_i, v_j) x_j = \sum\limits_{i=1}^{k} \overline{x}_i \left(v_i, \sum\limits_{j=1}^{k} \overline{x}_j v_j \right) = \sum\limits_{i=1}^{k} \overline{x}_i (v_i, v) = \left(\sum\limits_{i=1}^{k} \overline{x}_i v_i, v \right) = (v, v) \geq 0$, $v = \overline{x}_1 v_1 + \cdots + \overline{x}_k v_k$. □

1.4 Общий вид скалярного произведения в конечномерном пространстве.

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

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

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

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, проивольное скалярное произведение векторов $x = \sum\limits_{i=1}^{n} x_i e_i$ и $y = \sum\limits_{i=1}^{n} y_1 e_i$ имеет вид

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

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

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 $z = y + h,\ h \in$. Тогда

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

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

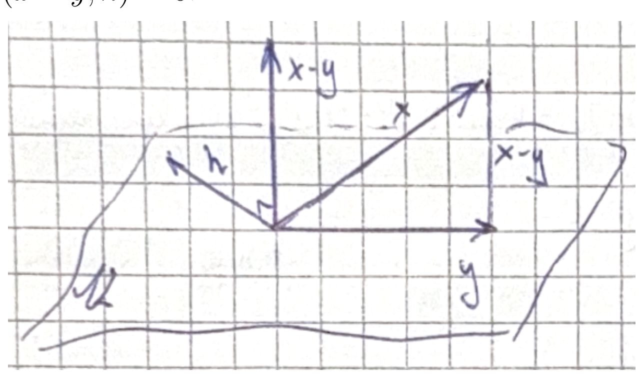
Если $(x - y, h) = 0\ \forall h \in M$, то $||x - y|| \leq ||x - z|| \forall z \in M$.
Если $||x - y|| \leq ||x - z||\ \forall z \in M$, то $-(x - y, h) - (h, x - y) + (h, h) \geq 0\ \forall h \in M$. Заменим что вектор h на $h_1 = \frac{(x-y,h)}{||h||^2} h$. Получим

$$-\left(x - y, \frac{(x - y, h)}{||h||^2} h\right) - \left(\frac{(x - y, h)}{||h||^2} h, x - y\right) + \left(\frac{(x - y, h)}{||h||^2} h, \frac{(x - y, h)}{||h||^2} h\right) =$$

$$= -\frac{\overline{(x - y, h)}}{||h^2||}(x - y, h) - \frac{(x - y, h)}{||h||^2}\overline{(x - y, h)} + \frac{|(x - y, h)|^2}{||h||^4}(h, h) =$$

$$= -2\frac{|(x - y, h)|^2}{||h^2||} + \frac{|(x - y, h)|^2}{||h||^2} = -\frac{|(x - y, h)|^2}{||h||^2} \geq 0$$

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



Докажем теперь, что существует вектор $y \in M$, удовлетворяющий условию $(x - y, h) = 0\ \forall h \in M$.
Пусть $e_1, \ldots, e_m$ - 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: $y = \sum\limits_{i=1}^{m} y_i e_i$. Тогда

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

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

— СЛАУ относительно $y_1, \ldots, y_m$, в которой матрица коэффициентов A — матрица Грама векторов $e_1, \ldots, e_m$. 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 $e = (e_1, \ldots, e_n)$ и $f = (f_1, \ldots, f_n)$ - базисы пространств V и W. Линейный оператор $A \in L(V, W)$ однозначно определяется заданием векторов $Ae_1, \ldots, Ae_n$. В свою очередь векторы Ae_i, $i = 1, \ldots, n$, однозначно определяются своими координатами в базисе f, т.е. коэффициентами разложений

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

Опр. *Матрица*

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

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

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. $|P| \neq 0$, $|Q| \neq 0$, $A = PBQ \implies B = P^{-1} A Q^{-1} \implies B \sim A$.
(транзитивность) $A \sim B, B \sim C \implies$, $\exists$ невырожденные P_1, P_2, Q_1, Q_2, т.ч. $A = P_1 B Q_1$, $B = P_2 C Q_2 \implies A = (P_1 P_2) B (Q_1 Q_2) \implies A \sim C$. □

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 $B = D^{-1} A C$. 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 $B = D^{-1} A C \implies$ матрицы A и B эквивалентны. □

Теорема. *Любая невырожденная матрица $A \in \mathbb{P}^{m\times n}$ ранга r эквивалентна матрице $I_r \in \mathbb{P}^{m\times n}$ 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 $A, B \in \mathbb{P}^{m\times n}$ эквивалентны тогда и только тогда, когда их ранги совпадают.*

Д-во. ($\Longrightarrow$) Вытекает из того, что умножение на невырожденную матрицу не меняет ранга матрицы.
($\Longleftarrow$) Следует из предыдущей теоремы и транзитивности эквивалентности матриц. $\square$

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

Теорема. *Инварианты подобия:*

1. *Ранг матрицы;*
2. *Опрделитель 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)

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

$\square$

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

Опр. *Образом линейного оператора называется множество* $\operatorname{im} A = \{y \in W \mid \exists x \in V : Ax = y\}$. *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, $\operatorname{im} A$ *- 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. □

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

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

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

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. $\operatorname{im} A = L(Ae_1, \dots, Ae_k, Ae_{k+1}, \dots, Ae_n) = L(Ae_{k+1}, \dots, Ae_n)$. 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 $\alpha_{k+1}Ae_{k+1} + \dots + \alpha_n A e_n = A(\alpha_{k+1}e_{k+1} + \dots + \alpha_n e_n) = \theta$. 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, $\dim\ker A = k$, $\dim\operatorname{im} A = n - k$. □

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* $A \in L(V, V) : \operatorname{im} A = V_1,\ \ker A = V_2$.

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

2.4 Обратимый оператор. Критерий обратимости. Линейность обратного оператора.

Опр. *Оператор $A : V \to W$ называется обратимым ил невырожденный, если существует оператор $B : W \to V$ 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

$$\left.\begin{array}{r} B_1AB_2 = (B_1A)B_2 = I_VB_2 = B_2 \\ B_1AB_2 = B_1(AB_2) = B_1I_W = B_2 \end{array}\right| \implies B_1 = B_2.$$

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 $By_1 = x_1,\ By_2 = x_2$. $\forall \alpha_1, \alpha_2 \in \mathbb{P}$:

$$\begin{aligned} B(\alpha_1y_1 + \alpha_2y_2) &= B(\alpha_1Ax_1 + \alpha_2Ax_2) = B(A(\alpha_1x_1 + \alpha_2x_2)) = \\ &= (BA)(\alpha_1x_1 + \alpha_2x_2) = \alpha_1x_1 + \alpha_2x_2 = \\ &= \alpha_1By_1 + \alpha_2By_2. \end{aligned}$$

□

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. $x_0 = \theta$. $\operatorname{im} A = L(Ae_1, \ldots, Ae_{\dim V}) \subseteq W \implies \dim W \geq \dim V$ и $\operatorname{im} A^{-1} = L(A^{-1}f_1, \ldots, A^{-1}f_{\dim W}) \subseteq V \implies \dim V \geq \dim W$. 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 $B : W \to V$ 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. $y_1, y_2 \in L,\ z_1, z_2 \in M,\ x_1 = z_1 + y_1,\ x_2 = y_2 + z_2,\ \lambda x_1 = \lambda y_1 + \lambda z_1$:

$$P(x_1 + x_2) = y_1 + y_2 = Px_1 + Px_2$$
$$P(\lambda x_1) = \lambda y_1 = \lambda P x_1$$

□

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

Д-во. ($\Longrightarrow$) $V = L \oplus M$ $\forall x \in V$ $\exists! u \in L, \mathrm{vim}\, M : x = u + v$ и $Px = u$. Значит $Pu = u$ $(u = u + \theta)$ и $P^2x = P(Px) = Pu = u = Px$, т.е. $P^2 = P$.
($\Longleftarrow$) Пусть $P^2 = P$. Положим $L = \mathrm{im}\, P,\ M = \ker P$. Тогда $\dim L + \dim M = \dim V$. Если $w \in L \cap M$, то $w = Px$ и $Pw = \theta$. Поэтому $Pw = P^2x = Px = 2 = \theta$. 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 $\theta \neq v \in V$ называются собственным значением и собственным вектором оператора A, если $Av = \lambda v$.*

Теорема. *Собственные вектора $x_1, \ldots, x_k$, отвечающие различным собственным значениям $\lambda_1, \ldots, \lambda_k$ 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 $x_1, \ldots, x_k$. Приравняем нулевому вектору линейную комбинацию этих векторов: $\alpha_1 x_1 + \cdots + \alpha_k x_k = \theta$. Под действием оператора A это равенство перейдет в равенство $\alpha_1 \lambda_1 x_1 + \cdots + \alpha_k \lambda_k x_k = \theta$ $(*)$. $(*) - \lambda_k (*) = \alpha_1(\lambda_1 - \lambda_k) + \cdots + \alpha_k(\lambda_{k-1} - \lambda_k)x_{k-1} = \theta$. В силу индуктивного предположения отсюда следует, что $\alpha_1 = \cdots = \alpha_{k-1} = 0$. Значит и $\alpha_k = 0$. Значит $x_1, \ldots, x_k$ линейно независимы. □

Следствие. *Линейный оператор, действующий в n-марном 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

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

□

Свойства характеристического многочлена.

- Характеристический многочлен индуцированного оператора является делителем характеристического многочлена порождающей его матрицы.

- Если $V = L_1 \oplus \cdots \oplus L_k$, где $L_1, \ldots, L_k$ - инвариантные подпространства относительно оператора $A \in L(V, V)$, то характеристический многочлен $f(\lambda)$. Равен произведению характеристических многочленов $f_1(\lambda), \ldots, f_k(\lambda)$ индуцированных операторов $A|L_1, \ldots, A|L_k$.

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

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

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

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 $e_1, \ldots, e_s$ - 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, $s \leq m$. □

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 $e_1, \dots, e_n$. $\square$

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

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

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 $P \subset \sum\limits_{i=1}^{n} W_{\lambda_i} \implies W_{\lambda_1} + W_{\lambda_p} = V$. Эта сумма прямая, т.к. собственные подпространства $W_{\lambda_1}, \dots, W_{\lambda_p}$ имеют тривиальное пересечение. $\square$

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 $\operatorname{rank}(A-\lambda I) \leq n-1$. Следовательно, $\dim \operatorname{im}(A-\lambda I) \leq n-1$ и в пространстве V существует подпространство L размерности $n-1$, которое содержит $\operatorname{im}(A-\lambda I)$. Очевидно, что L инвариантно относительно оператора $A-\lambda I$. Покажем, что оно инвариантно и относительно A. Пусть $x \in L$, тогда $(A-\lambda I)x = y \in L \implies Ax = \lambda x + y \in L$. □

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

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 $L_1 \subset L_2 \subset \cdots \subset L_{n-1}$. Так как действия операторов A и $A|L_{n-1}$ совпадают, то подпространства $L_1, \ldots, L_{n-1}$ инвариантны относительно оператора A. Остается добавить, что $L_{n-1} \subset L_n = V$. □

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 $L_1, \ldots, L_n$ таких, что $\dim L_k = k$ и $L_1 \subset L_2 \subset \cdots \subset L_n = V$. Искомый базис $e_1, \ldots, e_n$ 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 $L_1, \ldots, L_n$ матрица A_e имеет верхнюю треугольную форму. □

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

Опр. *Зафиксируем квадратную матрицу $A \in \mathbb{P}^{n\times n}$. Рассмотрим произвольный многочлен $f(\lambda) = \sum\limits_{i=0}^{k} f_i\lambda^i$ и поставим ему в соответствие матрицу $\sum\limits_{i=0}^{n} f_iA^i = f(A)$. $f(A)$ называется многочленом от матрицы A, соответствующий многочлену $f(\lambda)$ с коэффициентами из поля $\mathbb{P}$. Если $f(A) = 0$ $f(\lambda) \not\equiv 0$, 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. $a_0I + a_1A + \dots + a_{n^2}A^{n^2} = O$ — искомый многочлен, т.к. $f(\lambda) \not\equiv 0$ в силу нетривиальности набора $a_0, \dots, a_{n^2}$. □

Опр. *Многочлен, аннулирующий матрицу A и имеющий минимальную степень среди всех аннулирующий 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 $f(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_j - \lambda)^{m_j}$. $V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_p}$ 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 нулевову 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, $f(A_e) = O$. □

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 $A^n = O$. 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. ($\Longrightarrow$) Если λ - собственное значение нильпотентного оператора $A \in L(V,V)$ индекса q и x - собственное значение соответствующее ему, то $Ax = \lambda x \implies A^2x = \lambda^2 x \implies \dots \implies A^q x = \lambda^q x$. Отсюда следует, что $\lambda^q x = 0$. Так как $x \neq 0$, то $\lambda = 0$.
($\Longleftarrow$) 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 $Ay = A^{q+1}y = A^q(Ax) = A^q x_1$, 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 свойствам 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 $\operatorname{im} A|T = T$. 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 $y = Ay_1 = A^2 y_2 = \cdots = A^q y_q$. 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. $v \in K_{\lambda_j} \implies \exists q_j : (A-\lambda_j I)^{q_j} v = \theta \implies (A-\lambda_j I)^{q_j}(Av) = A(A-\lambda_j I)^{q_j} v = A \cdot \theta = \theta \implies Av \in K_{\lambda_j}$. □

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 $N_k = \ker B^k$, $T_k \mathrm{im}\, B^k$, 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 $V = K_{\lambda_p} \oplus T_q$, $T_q = \mathrm{im}\,(A - \lambda_p I)^{m_p}$. 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 $V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_{p-1}} \oplus K_{\lambda_p}$. □

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, содержащие заданные 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^kx. Его размерность равна минимальному значению k, при котором $A^kx \in L_k(A, x)$.*

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

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 $A^kx \in L_k(A, x)$. 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 $B^kx = \theta$. Пусть $L \subseteq M$ — произвольное ненулевое инвариантное подпространство B. Возьмем ненулевой вектор $\theta \neq z \in L$, $z = \sum\limits_{j=0}^{k-1} \alpha_j B^j x$, 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 $B = A - \lambda I$. □

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

Д-во. ($\Longrightarrow$) Из системы (1) очевидно следует линейная независимость системы (2).
($\Longleftarrow$) Пусть (2) линейно независима и $k = \max\limits_{1\le i\le t} k_i$. Индукция по k. При $k = 1$ системы (1) и (2) совпадают. Пусть $k \ge 2$ и $I_k = \{i : k_i = k, i = \overline{1,t}\}$. Пусть $\theta = y = \sum\limits_{i=1}^{t}\sum\limits_{j=1}^{k_i-1} \alpha_{ij}A^j x_i\,(*)$. Тогда $\theta = A^{k-1}y = \sum\limits_{i\in I_k} \alpha_{i0}A^{k-1}x_i \Leftrightarrow \alpha_{i0} = 0\,\forall i \in I_k$. 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) $\alpha_{ij} = 0$, $\forall i, j : 1 \le i \le t$, $0 \le j \le k_i - 1$. $\square$

Следствие. *Пусть A - нильпотентный оператор и $L_{k_1}(A, x_1), \ldots, L_{k_t}(A, x_t)$ - максимальные пространства Крылова. Для того чтобы их сумма была прямой необходимо и достаточно, чтобы $A^{k_1-1}x_1, \ldots, A^{k_t-1}x_t$ 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 $n_k = 2\operatorname{def} A^k - \operatorname{def} A^{k+1} - \operatorname{def} A^{k-1}$.*

Д-во. Пусть нильпотентный оператор A действует в V. Рассмотрим какие-либо максимальные пространства Крылова:$L_1 = L_{k_1}(A, x_1), \ldots, L_s = L_{k_s}(A, x_s)$ дающие в сумме (необязательно прямой) пространство V. Чтобы получить их, возьмем в качестве начальных векторов $x_1, \ldots, x_s$, например, базис пространства V.
Если $V = L_1 + \cdots + L_s$ - прямая сумма, то нужное расщепление получено. Если нет, то $A^{k_1-1}x_1, \ldots, A^{k_s-1}x_s$ — ЛЗ. Рассмотрим нетривиальную линейную комбинацию $\sum\limits_{i=1}^{s} \alpha_i A^{k_i-1}x_i = \theta$. Среди векторов линейно комбинации, перед которыми стоит ненулевой коэффициент, выберем тот, который входит в пространство Крылова с наименьшей размерностью. Пусть это будет $A^{j_j-1}x_j$.

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

$$y_0 = x_j + \sum_{i=1, i\neq j}^{s} \frac{\alpha_i}{\alpha_j} A^{k_i - k_j} x_i,$$

$$y_1 = Ax_j + \sum_{i=1, i\neq j}^{s} \frac{\alpha_i}{\alpha_j} A^{k_i - k_j + 1} x_i,$$

$$\dots$$

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

$y_0, \dots, y_{k_j-1}$ 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 размерность $d = \dim L(y_0, \dots, y_{k_j-2}) < k_j$. 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 $m_k = \operatorname{def} A^k - \operatorname{def} A^{k-1}$.
Множество $\{A^d x_i : k \leq d \leq k_i - 1, i = \overline{1, s}\}$ состоит из ненулевых векторов, образующих базис пространства $\operatorname{im} A^k$: $\operatorname{rank} A^k = \sum_{i=1}^{s} l_i,\ l_i = \max(0, k_i - k) \implies \operatorname{def} A^k = \sum_{i=1}^{s} (k_i - l_i) \implies \ker A^k = L(A^{l_1} x_1, A^{l_1+1} x_1, \dots, A^{k_1 - 1} x_1, \dots, A^{l_s} x_s, A^{l_s+1} x_s, \dots, A^{k_s-1} x_s)$. Число образующих векторов равно размерности ярда и все они принадлежат ядру оператора A^k. Если $l_j = k_j - k$, то $A^k A^{l_j} x_j = A^{k + k_j - k} x = A^{k_j} x_j = \theta$, 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 размерность, которых $\geq k$. Пусть M_k — их линейная оболочка. Тогда $\dim M_k = m_k$, $\ker A^k = \ker A^{k-1} \oplus M_k \implies m_k = \operatorname{def} A^k - \operatorname{def} A^{k-1}$.
$n_k = m_k - m_{k+1} = (\operatorname{def} A^k - \operatorname{def} A^{k-1}) - (\operatorname{def} A^{k+1} - \operatorname{def} A^k) = -\operatorname{def} A^{k-1} + 2\operatorname{def} A^k - \operatorname{def} A^{k+1}$. □

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 $x, Bx, \ldots, B^{k-1}x$, где $B = A - \lambda I$. Составим из них базис $e_1 = B^{k-1}x$, $e_2 = B^{k-2}x, \ldots, e_{k-1} = Bx$, $e_k = x \implies Ae_1 = \lambda e_1, Ae_2 = \lambda e_2 + e_1, \ldots, Ae_k = \lambda e_k + e_1$. 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 $h_1, \ldots, h_n$, т.ч. $A_h = J_{k_1}(\lambda) \oplus \cdots \oplus J_{k_r}(\lambda)$, где $k_1 + \cdots + k_r = n$.*

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

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

Д-во. Положим $f_j(\lambda) = (\lambda - \lambda_j)^{l_j}$, $j = \overline{1, k}$. Из теоремы о корневом расщеплении $V = W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_k}$, 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 $B = A - \lambda_j I$, $N_k = \ker B^k$, $n_k = \dim N_k$, $r_k = \operatorname{rank} B^k$.
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 случе для нетривиального набора чисел $\alpha_1, \dots, \alpha_{t_q}$ имеем $B^{q-2}\sum\limits_{k=1}^{t_q} \alpha_k B d_f$, т.е. $B^{q-1}\sum\limits_{k=1}^{t_q} \alpha_k f_k = \theta$, и $\sum\limits_{k=1}^{t_p} \alpha_k f_k \in N_{q-1}$, что противоречит линейной независимости $f_1, \dots, f_{t_q}$ над N_{q-1}.
Дополним эти векторы векторами $g_1, .., g_{t_{p-1}} \in N_{q-1}$ так, что векторы $Bf_1, .., Bf_{t_q}, g_1, \dots, g_{t_{p-1}}$ дополняли произвольный базис N_{q-2} до базиса N_{q-1}. Ясно, что:
1) они будут корневыми векторам 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

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

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

$$B^{q-1}f_1, \dots, B^{q-1}f_{t_q}, B^{q-2}g_1, \dots, B^{q-2}g_{t_{q-1}}, \dots, Bv_1, ..., B_{vt_2}, u_1, ..., u_{t-1},$$

1) являются собственными векторами;
2) их количество равно $n_1 = n_1 - n_0$ (очевидно, $n_0 = \operatorname{def} B^0 = 0$);
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	$\underbrace{f_1, \ldots, f_{t_{q-1}}}_{t_q=-n_{q-1}+2n_q-n_{q+1}}$			
N_{q-1}	$Bf_1, ..., Bf_{t_q}$	$\underbrace{g_1, \ldots, g_{t_{q-1}}}_{t_{q-1}=-n_{q-2}+2n_{q-1}-n_q}$		
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
N_1	$B^{q-1}f_1, ..., B^{q-1}f_{t_q}$	$B^{q-2}g_1, \ldots, B^{q-2}g_{t_{q-1}}$	$\cdots$	$\underbrace{u_1, \ldots, u_{t_1}}_{t_1=-n_0+2n_1-n_2}$

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 жоржановых клеток $P^{-1}AP = J_1 \oplus \cdots \oplus J_n$, причем число и размеры жоржановых 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 жоржановых клеток. Это утверждение следует из того, что перемещение столбцов матрицы реализуется умножением матрицы на матрицу элементарных преобразований, которая не вырождена. □

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 $\overline{\lambda} = a - bi$ тоже собственное значение A и алгебраические кратности λ и $\overline{\lambda}$ совпадают. Пусть $h_1, \ldots, h_k$ — жарданова цепочка (система векторов Крылова) клетки $J_k(\lambda)$, тогда $\overline{h}_1, \ldots, \overline{h}_k$ — жорданова цепочка, отвечающая жардановой клетке $J_k(\overline{\lambda})$.
Пусть $h_j = x_j + iy_j$ $(x_j, y_j \in \mathbb{R}^n)$, $j = \overline{1, k}$, тогда

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

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

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

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 называет я 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 $p(t)$ являетя 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 $\lambda_1, \ldots, \lambda_m$. Тогда степень ее минимального многочлена равна сумме $n_1 + \cdots + n_m$, 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 $L_{j_1}, \ldots, L_{j_m}$ отвечают собственным значениям $\lambda_1, \ldots, \lambda_m$ и имеют размерности $n_1, \ldots, n_m$. Тогда подпространства $\ker(A-\lambda I)^{n_i}$ является корневым подпространством собственного значения $\lambda_i \implies (A - \lambda_i I)^{n_1} \ldots (A - \lambda_m I)^{n_m} = 0$. Степень минимального многочлена не выше $n_1 + \cdots + n_m$. 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 является 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 степень 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 в характеристическим. Действительно, пусть $p(t) = a_0 + a_1 t + \dots + a_{n-1}t^{n-1}$. Тогда $p(A)x = a_0 Ix + a_1 Ax + \dots + a_{n-1}A^{n-1}x$. 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.