Линейные операторы

Содержание

Литература:

В. Винберг Э. Б. Курс алгебры. М.: МЦНМО, 2013.

К. Кострикин А. И. Введение в алгебру. Часть II. Линейная алгебра. М.: Физико-математическая литература, 2000.

Ф. Фаддеев Д. К. Лекции по алгебре. М.: «Наука», 1984.

Г. Гайфуллин А. А., Пенской А. В., Смирнов С. В. Задачи по линейной алгебре и геометрии. М.: МЦНМО, 2014. (Содержит подробные решения)

Лекция I

В., с. 234-240, 200-205; **К.**, с. 60-77; **Ф.**, с. 314-319; **Г.**, с. 40-44, 59-61;

Данный модуль посвящён теории линейных операторов в векторных пространствах, которая является основным источником приложений линейной алгебры в различных областях математики.

§1. Линейные операторы. Ядро и образ

Определение 1.1. ***Линейным оператором*** в векторном пространстве V (***эндоморфизмом*** пространства V) называется отображение $\mathcal{A}\colon V \to V$, удовлетворяющее условиям:

1. $\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y$ для любых $x, y \in V$;
2. $\mathcal{A}(\lambda x) = \lambda \mathcal{A}x$ для любых $x \in V$, $\lambda \in F$.

Множество всех линейных операторов в пространстве V будем обозначать $\mathrm{End}(V)$.

Пример 1.1. а) Нулевой оператор $\mathcal{O}$: переводит любой вектор любого пространства в нулевой;

б) Тождественный оператор $\mathcal{E}$: переводит любой вектор любого пространства в себя;

в) «Растяжение» $\lambda\mathcal{E}$, $\lambda \neq 0$: переводит любой вектор x в вектор λx;

г) Поворот на угол α — линейный оператор в плоскости E^2;

д) Пусть $V = U \oplus W$, тогда проектор на U параллельно W является линейным оператором в V;

е) Дифференцирование $\mathcal{D}\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ — линейный оператор в пространстве многочленов с вещественными коэффициентами степени не выше n;

ж) $\mathcal{T}\colon M_n(F) \to M_n(F)$ — линейный оператор транспонирования в пространстве квадратных матриц размера $n \times n$ с элементами из поля F.

Линейные операторы в одном векторном пространстве можно складывать и умножать на скаляры как обычные функции: $(\mathcal{A}+\mathcal{B})x = \mathcal{A}x + \mathcal{B}x, (\lambda\mathcal{A})x = \lambda(\mathcal{A}x)$. Относительно этих операций они образуют векторное пространство. Далее, если $\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$, то их произведение (композиция) $\mathcal{A}\mathcal{B}$ также является линейным оператором. Умножение линейных операторов ассоциативно. Легко понять, что $\mathcal{E}\mathcal{A} = \mathcal{A}\mathcal{E} = \mathcal{A}$ для любого $\mathcal{A} \in \mathrm{End}(V)$. Столь же легко понять, что в общем случае $\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}$ (приведите пример). То есть операторы со сложением и умножением являются ассоциативным кольцом с единицей. Векторное пространство + кольцо + $((\lambda a)b = a(\lambda b) = \lambda(ab))$ = ***алгебра*** (см. **В.**, с. 38-41).

Пример 1.2. Оператор транспонирования $\mathcal{T}$ удовлетворяет условию $\mathcal{T}^2 = \mathcal{E}$ (пример ***инволюции***), а проектор $\mathcal{P}$ — условию $\mathcal{P}^2 = \mathcal{P}$ (пример ***идемотента***).

Определение 1.2. Для линейного оператора $\mathcal{A}$ определяется его ***образ*** $\mathrm{Im}\,(\mathcal{A}) = \{\mathcal{A}x \,|\, x \in V\}$ и ***ядро*** $\mathrm{Ker}(\mathcal{A}) = \{x \in V \,|\, \mathcal{A}x = 0\}$.

Замечание 1.1. Образ и ядро являются подпространствами в соответствующем векторном пространстве, то есть замкнуты относительно сложения векторов и умножения на скаляры.

Пример 1.3. а) $\mathrm{Im}\,(\mathcal{O}) = \{0\}$, $\mathrm{Ker}(\mathcal{O}) = V$ в любом пространстве V;

б) $\mathrm{Im}\,(\mathcal{E}) = V$, $\mathrm{Ker}(\mathcal{E}) = \{0\}$ в любом пространстве V;

в) $\mathrm{Im}\,(\lambda\mathcal{E}) = V$, $\mathrm{Ker}(\lambda\mathcal{E}) = \{0\}$ в любом пространстве V, $\lambda \neq 0$;

г) Образ оператора поворота в E^2 — вся плоскость E^2, его ядро — только нулевой вектор.

д) Если $V = U \oplus W$ и тогда $\mathcal{P}$ — проектор на U параллельно W, то $\mathrm{Im}\,\mathcal{P} = U$, $\mathrm{Ker}\,\mathcal{P} = W$;

е) $\mathcal{D}\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_n$, тогда $\mathrm{Im}\,(\mathcal{D}) = \mathbb{R}[x]_{n-1}$, $\mathrm{Ker}(\mathcal{D}) = \mathbb{R}[x]_0$ (константы);

ж) $\mathcal{T}\colon M_n(F) \to M_n(F)$, тогда $\mathrm{Im}\,(\mathcal{D}) = M_n(F)$, $\mathrm{Ker}(\mathcal{D}) = \{O\}$.

Теорема 1.1. $\dim \mathrm{Im}\,\mathcal{A} + \dim \mathrm{Ker}\,\mathcal{A} = \dim V$.

Доказательство. Выберем базис $e_1, e_2, \dots, e_k$ подпространства $\mathrm{Ker}\,\mathcal{A}$ и дополним его векторами $e_{k+1}, \dots, e_n$ до базиса всего пространства V. Достаточно показать, что векторы $\mathcal{A}(e_{k+1}), \dots, \mathcal{A}(e_n)$ составляют базис $\mathrm{Im}\,\mathcal{A}$. Они порождают образ, так как для любого $x = \sum\limits_{i=1}^{k} \lambda_i e_i + \sum\limits_{i=k+1}^{n} \lambda_i e_i$ имеем $\mathcal{A}x = \underbrace{\sum\limits_{i=1}^{k} \lambda_i(\mathcal{A}e_i)}_{0} + \sum\limits_{i=k+1}^{n} \lambda_i(\mathcal{A}e_i) = \sum\limits_{i=k+1}^{n} \lambda_i(\mathcal{A}e_i)$. Векторы $\mathcal{A}e_{k+1}, \dots, \mathcal{A}e_n$ линейно независимы, так как из равенства $0 = \sum\limits_{i=k+1}^{n} \lambda_i(\mathcal{A}e_i) = \mathcal{A}\left(\sum\limits_{i=k+1}^{n} \lambda_i e_i\right)$ следует, что $\sum\limits_{i=k+1}^{n} \lambda_i e_i \in \mathrm{Ker}\,\mathcal{A}$ и является линейной комбинацией векторов $e_1, e_2, \dots, e_k$, что возможно только если все λ_i равны 0.

Следствие 1.1.1. *Следующие свойства линейного оператора $\mathcal{A}$ эквивалентны:* 1) $\mathcal{A}$ *— изоморфизм;* 2) $\mathrm{Ker}\,\mathcal{A} = \{0\}$*;* 3) $\mathrm{Im}\,\mathcal{A} = V$.

§2. Матрица линейного оператора

Если в пространстве V выбран базис $e_1, e_2, \dots, e_n$, то линейный оператор можно задать матрицей.

Определение 2.1. ***Матрицей линейного оператора*** $\mathcal{A}$ в базисе $e_1, e_2, \dots, e_n$ называется матрица $A = (a_{ij})$, определяемая из равенств $\mathcal{A}e_j = \sum\limits_{i=1}^{n} a_{ij} e_i$.

То есть в j-том столбце матрицы A стоят координаты образа j-того базисного вектора в базисе $e_1, e_2, \ldots, e_n$. Можно записать определение матрицы линейного оператора как $(\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n) = (e_1, e_2, \ldots, e_n)A$. Для любых векторов $x_1, x_2, \ldots, x_n \in V$ существует единственный линейный оператор, переводящий $e_1, e_2, \ldots, e_n$ в $x_1, x_2, \ldots, x_n$ — это оператор, переводящий каждый вектор $u = \sum\limits_{i=1}^{n} u_i e_i$ в вектор $\sum\limits_{i=1}^{n} u_i x_i$. Таким образом, линейный оператор однозначно определяется своей матрицей и наоборот, любая квадратная матрица n-ного порядка является матрицей некоторого линейного оператора в данном базисе. Операциям над линейными оператора соответствуют такие же операции над их матрицами. Для линейных операций это очевидно, поэтому проверим для умножения.

Пусть $\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$, A, B — соответственно их матрицы в базисе $e_1, e_2, \ldots, e_n$. Тогда $(\mathcal{A}\mathcal{B})e_k = \mathcal{A}(\mathcal{B}e_k) = \mathcal{A}\left(\sum\limits_{j=1}^{n} b_{jk} e_j\right) = \sum\limits_{j=1}^{n} (\mathcal{A}e_j) = \sum\limits_{i,j=1}^{n} a_{ij} b_{jk} e_i$, следовательно, матрица оператора $\mathcal{A}\mathcal{B}$ есть $C = (c_{ik})$, где $c_{ik} = \sum\limits_{j=1}^{n} a_{ij} b_{jk}$, то есть $C = AB$. Таким образом, кольцо $\mathrm{End}(V)$ изоморфно кольцу $M_n(F)$.[1]

Замечание 2.1. Если $\dim V = n$, то размерность End(V) как векторного пространства равна n^2.

Пример 2.1. а) Матрицей нулевого оператора $\mathcal{O}$ является нулевая матрица O;

б) Матрицей тождественного оператора $\mathcal{E}$ является единичная матрица E;

в) Матрицей оператора «растяжения» $\lambda\mathcal{E}$ является скалярная матрица λE;

г) Матрица оператора поворота на угол α в плоскости E^2 в ортонормированном базисе: $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$;

д) Матрица проектора на U параллельно W в объединении базисов U и W имеет блочный вид $\begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix}$, $k = \dim U$;

е) Матрица оператора дифференцирования в базисе $x^n, x^{n-1}, \ldots, x, 1$ имеет вид $\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ n & 0 & 0 & \ldots & 0 & 0 \\ 0 & n-1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & n-2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$;

[1] «Преподавание математики все еще страдает от энтузиазма, вызванного открытием этого изоморфизма. Следствием было то, что геометрия фактически исключалась и заменялась вычислениями <...> Мой опыт показывает, что доказательства, включающие в себя матрицы, могут быть сокращены на 50%, если выбросить матрицы» (Э. Артин. Геометрическая алгебра. М.: Изд-во «Наука», 1969 г., с. 28 и далее).

ж) Матрица оператора транспонирования в $M_2(F)$ в базисе из стандартных матричных единиц $E_{11}, E_{12}, E_{21}, E_{22}$ имеет вид $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$;

Найдём выражение для координат образа y вектора x при действии оператора $\mathcal{A}$: $y = \mathcal{A}x$. Пусть $x = \sum\limits_{j=1}^{n} x_j e_j$, тогда $y = \mathcal{A}\left(\sum\limits_{j=1}^{n} x_j e_j\right) = \sum\limits_{j=1}^{n} x_j(\mathcal{A}e_j) = \sum\limits_{i,j=1}^{n} a_{ij}x_j e_i = \sum\limits_{i=1}^{n} y_i e_i$, где $y_i = \sum\limits_{j=1}^{n} a_{ij}x_j$. Если X и Y — столбцы координат векторов x и y соответственно, то полученные равенства можно записать как $Y = AX$.

Пусть $e = \{e_1, e_2, \ldots, e_n\}$ и $\widetilde{e} = \{\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n\}$ — два базиса векторного пространства V. Выясним, как преобразуется матрица линейного оператора при переходе от одного базиса к другому. Пусть $C = (e \rightsquigarrow \widetilde{e})$, тогда $(\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n) = (e_1, e_2, \ldots, e_n)C$. $(\mathcal{A}\widetilde{e}_1, \mathcal{A}\widetilde{e}_2, \ldots, \mathcal{A}\widetilde{e}_n) = (\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n)C = (e_1, e_2, \ldots, e_n)AC = (\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_n)C^{-1}AC$. Следовательно, если $\widetilde{A}$ — матрица оператора $\mathcal{A}$ в базисе $\widetilde{e}$, то

$$\widetilde{A} = C^{-1}AC.$$

Одна из основных задач теории линейных операторов состоит в нахождении базиса, в котором матрица имеет «наиболее простой» вид.

Замечание 2.2. Определитель матрицы оператора зависит только от самого оператора, но не от базиса, в котором записана эта матрица. Действительно, $\det \widetilde{A} = \det(C^{-1}AC) = (\det C)^{-1} \det A \det C = \det A$. Это позволяет говорить об ***определителе оператора*** $\det \mathcal{A}$ и рассмотреть ***невырожденные операторы*** в пространстве V, у которых $\det \mathcal{A} \neq 0$. Невырожденные линейные операторы в пространстве V образуют группу $\mathrm{GL}(V)$, называемую ***полной линейной группой*** пространства V.

Кроме того, верна следующая

Теорема 2.1. $\dim \operatorname{Im} \mathcal{A} = \operatorname{rank} A$.

Доказательство. Поскольку $\operatorname{Im} \mathcal{A}$ является линейной оболочкой образов базисных векторов $e_1, e_2, \ldots, e_n$, то $\dim \operatorname{Im} \mathcal{A}$ является рангом системы векторов $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n$. Однако столбцы матрицы A и есть координаты этих векторов в базисе $e_1, e_2, \ldots, e_n$.

Замечание 2.3. В силу следствия 1.1.1 и теоремы 2.1 оператор обратим тогда и только тогда, когда он невырожденный.

§3. Инвариантные подпространства

Определение 3.1. Подпространство $U \leqslant V$ называется ***инвариантным*** относительно оператора $\mathcal{A}$ ($\mathcal{A}$-инвариантным), если $\mathcal{A}U \leqslant U$, то есть для любого $x \in U$ его образ $\mathcal{A}x \in U$.

Замечание 3.1. Нулевое подпространство и всё пространство V инвариантны для любого оператора. Любое подпространство, содержащееся в ядре оператора $\mathcal{A}$, и любое подпространство, содержащее его образ, $\mathcal{A}$-инвариантны. Сумма и пересечение инвариантных подпространств являются инвариантными подпространствами.

Пример 3.1. Пусть оператор — осевая симметрия относительно оси абсцисс в декартовой системе координат на плоскости. Тогда $U_1 = \langle \mathbf{i} \rangle$ и $U_2 = \langle \mathbf{j} \rangle$ инвариантны, а $U_3 = \langle \mathbf{i} + \mathbf{j} \rangle$ — нет.

Пример 3.2. Инвариантные подпространства оператора дифференцирования в $\mathbb{R}[x]_n$ имеют вид $\mathbb{R}[x]_k$, $k \leqslant n$ (см. **Г**, с. 61, задача 48).

Ограничение (***сужение***) $\mathcal{A}|_U$ линейного оператора $\mathcal{A}$ на инвариантное подпространство U является линейным оператором в U.

Если выбрать базис $e_1, e_2, \ldots, e_n$ пространства V так, чтобы инвариантное подпространство U было линейной облочкой первых k базисных векторов, то матрица оператора в этом базисе будет иметь вид $\begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$, где B — матрица оператора $\mathcal{A}|_U$ в базисе $e_1, e_2, \ldots, e_k$. Обратно, если матрица оператора $\mathcal{A}$ имеет такой блочный вид (где B — квадратная матрица размера $k \times k$, а под ней матрица из нулей), то $U = \langle e_1, e_2, \ldots, e_k \rangle$ — инвариантное подпространство.

Если удаётся разложить V в прямую сумму $V = V_1 \oplus V_2 \oplus \ldots \oplus V_k$ инвариантных подпространств V_i, то в базисе пространства V, составленном из базисов этих подпространств, матрица оператора $\mathcal{A}$ имеет блочно-диагональный вид

$$\begin{pmatrix} A_1 & & & \mathbf{O} \\ & A_2 & & \\ & & \ddots & \\ \mathbf{O} & & & A_k \end{pmatrix},$$

где A_i — матрица оператора $\mathcal{A}|_{V_i}$. Из этого ясно, что поиск инвариантных подпространств является важным шагом в решении задачи поиска «наиболее простого» вида матрицы линейного оператора.

Пример 3.3. Для оператора дифференцирования в $\mathbb{R}[x]_n$ инвариантные подпространства вложены друг в друга, поэтому ни у какого нетривиального инвариантного подпространства нет инвариантного прямого дополнения. Поэтому ни в каком базисе матрица этого оператора не может иметь блочно-диагональный вид.

Пример 3.4. Рассмотрим поворот на угол α вокруг какой-либо оси в пространстве E^3. В ортонормированном базисе e_1, e_2, e_3, если вектор e_3 направлен по оси поворота, матрица оператора поворота имеет вид

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

который согласуется с разложением E^3 в прямую сумму $E^3 = \langle e_1, e_2 \rangle \oplus \langle e_3 \rangle$

Особую роль играют одномерные инвариантные подпространства, которые приводят к понятию собственного вектора.

Лекция II

В., с. 240-246; **К.**, с. 77-82; **Г.**, с. 44-46;

§4. Собственные векторы и собственные значения

Определение 4.1. Ненулевой вектор $x \in V$ называется ***собственным вектором*** оператора $\mathcal{A}$, если $\mathcal{A}x = \lambda x$. Число $\lambda \in F$ называется при этом ***собственным значением*** (***собственным числом***) оператора $\mathcal{A}$, отвечающим собственному вектору x.

Собственный вектор порождает одномерное инвариантное подпространство. В базисе, составленном из собственных векторов (если он существует) матрица оператора имеет диагональный вид, что является «наиболее простым» видом.

Пример 4.1. а) Для нулевой оператора $\mathcal{O}$ каждый вектор является собственным с собственным значением 0;

б) Для тождественного оператора $\mathcal{E}$ каждый вектор является собственным с собственным значением 1;

в) Для оператора «растяжения» $\lambda\mathcal{E}$ каждый вектор является собственным с собственным значением λ;

г) Собственные векторы оператора поворот на угол α в пространстве E^3: если $\alpha \neq \pi k$, то это векторы, лежащие на оси поворота, собственное значение 1; если же $\alpha = \pi k$, то к ним добавляются векторы, перпендикулярные оси поворота, собственные значения $(-1)^k$;

д) $V = U \oplus W$, $\mathcal{P}$ — проектор на U параллельно W является линейным оператором в V. Известно, что $\mathcal{P}^2 = \mathcal{P}$. Тогда если x — собственный вектор, то $\lambda x = \mathcal{P}x = \mathcal{P}^2 x = \mathcal{P}(\mathcal{P}x) = \mathcal{P}(\lambda x) = \lambda\mathcal{P}x = \lambda^2 x$, значит λ равно 0 или 1. Любой вектор из U — собственный с собственным значением 1, из W — собственный с собственным значением 0.

е) Для оператора дифференцирования $\mathcal{D}\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ собственными являются константы, соответствующие собственному значению 0.

ж) Оператор транспонирования удовлетворяет условию $\mathcal{P}^2 = \mathcal{E}$. Аналогично д) можно показать, что собственные значения инволюции только ± 1. В данном случае, значение 1 соответствует симметрическим матрицам, -1 — кососимметрическим.

Замечание 4.1. Число $\lambda \in F$ является собственным для оператора $\mathcal{A}$ тогда и только тогда, когда подпространство $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) \leqslant V$ ненулевое, то есть когда оператор $\mathcal{A} - \lambda\mathcal{E}$ вырожден, то есть $\det(\mathcal{A} - \lambda\mathcal{E}) = 0$.

Определение 4.2. Подпространство $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})$ называется ***собственным подпространством*** оператора $\mathcal{A}$, соответствующим собственному значению λ и обозначается V_λ. Помимо собственных векторов, оно содержит нулевой.

Замечание 4.2. Любое собственное подпространство оператора $\mathcal{A}$ является $\mathcal{A}$-инвариантным.

Определение 4.3. ***Геометрической кратностью*** $g(\lambda)$ собственного значения λ называется размерность соответствующего ему собственного подпространства: $g(\lambda) = \dim V_\lambda$.

§5. Характеристический многочлен

Из замечания 4.1 следует, что для нахождения собственных подпространств удобнее сначала найти собственные значения из условия $\det(\mathcal{A} - \lambda\mathcal{E}) = 0$. Пусть A — матрица оператора $\mathcal{A}$ в каком-либо базисе, тогда

$$\det(\mathcal{A} - t\mathcal{E}) = \begin{vmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{vmatrix}.$$

Определение 5.1. Многочлен $\chi_{\mathcal{A}}(t) = (-1)^t \det(\mathcal{A} - t\mathcal{E}) = \det(t\mathcal{E} - \mathcal{A})$ называется ***характеристическим многочленом*** оператора $\mathcal{A}$.

Корни характеристического многочлена называются ***спектром оператора*** $\mathcal{A}$.

Если из контекста ясно, о каком операторе идёт речь, индекс $\mathcal{A}$ будем опускать.

Замечание 5.1. Определение корректно, так как характеристический многочлен не зависит от выбора базиса. Действительно, пусть C — матрица перехода от старого базиса к новому, $\widetilde{A} = C^{-1}AC$, тогда $\det(tE - \widetilde{A}) = \det(C^{-1}(tE)C - C^{-1}AC) = \det(C^{-1}(tE - A)C) = (\det C)^{-1}\det(tE - A)\det C = \det(tE - A)$.

Замечание 5.2. Нетрудно заметить, что старший коэффициент характеристического многочлена равен 1, коэффициент при t^n равен $-\operatorname{tr} A$, а свободный коэффициент равен $\chi_{\mathcal{A}}(0) = (-1)^n \det A$. Так как параметрический многочлен не зависит от выбора базиса, то и след матрицы оператора не зависит от базиса, поэтому можно говорить о ***следе оператора*** $\operatorname{tr}\mathcal{A}$. Собственные значения оператора — это в точности корни его характеристического многочлена, их количество не превосходит $n = \dim V$. В силу основной теоремы алгебры, любой линейный оператор в ненулевом конечномерном векторном пространстве над полем $\mathbb{C}$ имеет собственный вектор.

Определение 5.2. ***Алгебраической кратностью*** $m(\lambda)$ собственного значения λ называется его кратность как корня характеристического многочлена.

Лемма 5.1. *Геометрическая кратность собственного значения не превосходит его алгебраической кратности.*

Доказательство. Рассмотрим собственное подпространство $V_\lambda \leqslant V$, оно инвариантно, следовательно, в согласованном базисе матрица оператора имеет вид

$$A = \begin{pmatrix} \lambda E & D \\ \mathbf{0} & C \end{pmatrix},$$

матрица λE квадратная размера $g(\lambda)$. Тогда характеристический многочлен имеет вид $\chi(t) = (t-\lambda)^{g(\lambda)} \cdot f(t)$, $f(t) = \det(tE - C)$. Так как у многочлена f может быть корень λ, то $m(\lambda) \geqslant g(\lambda)$.

§6. Собственный базис и диагонализируемость

Определение 6.1. Подпространства $V_1, \ldots, V_k$ называются ***линейной независимыми***, если равенства $v_1 + \ldots + v_k = 0$, $v_k \in V_k$ следует, что $v_1 = \ldots = v_k = 0$.

Замечание 6.1. Можно сказать, что разложение пространства V в прямую сумму подпространств $V_1, \ldots, V_k$ — это разложение в сумму линейной независимых подпространств $V_1, \ldots, V_k$.

Теорема 6.1. *Собственные подпространства, отвечающие попарно различным собственным значениями $\lambda_1, \lambda_2, \ldots, \lambda_n$ оператора $\mathcal{A}$ линейно независимы.*

Доказательство. Индукция по k. При $k = 1$ доказывать нечего. Пусть $k > 1$ и $x_1 + \ldots + x_k = 0$. Применим к обеим частям оператор $\mathcal{A}$:

$$\lambda_1 x_1 + \ldots \lambda_k x_k = 0.$$

Вычтем из этого равенства равенство $x_1 + \ldots + x_k = 0$, умноженное на λ_k, получим

$$(\lambda_1 - \lambda_k)x_1 + \ldots + (\lambda_{k-1} - \lambda_k)x_{k-1} = 0.$$

Каждое из слагаемых лежит в соответствующем подпространстве. Так как по предположению индукции они линейно независимы и рассматриваются разные собственные значения, то $x_1 = \ldots = x_{k-1} = 0$. Но тогда и $x_k = 0$.

Следствие 6.1.1. *Если характеристический многочлен оператора имеет $n = \dim V$ различных корней (оператор с* ***простым спектром****), то существует базис из собственных векторов этого оператора.*

Замечание 6.2. Данное условие не является необходимым для существования собственного базиса. Например, любой базис состоит из собственных векторов тождественного оператора, однако его характеристический многочлен имеет единственный корень 1 (кратности n).

Определение 6.2. Линейный оператор в конечномерном векторном пространстве называется ***диагонализируемым***, если существует базис, в котором матрицам этого оператора имеет диагональный вид.

Другими словами, диагонализируемость оператора эквивалентна существованию собственного базиса.

Пусть оператор $\mathcal{A}$ диагонализируем и $V = \bigoplus_{i=1}^{n} V_{\lambda_i}$. Рассмотрим проектор $\mathcal{P}_i$ на подпространство V_{λ_i} параллельно прямой сумме оставшихся подпространств. Тогда $\mathcal{P}_i^2 = \mathcal{P}_i$, $\mathcal{P}_i\mathcal{P}_j = \mathcal{O}$ при $i \neq j$ и $\sum\limits_{i=1}^{k} \mathcal{P}_i = \mathcal{E}$. Легко проверяется, что оператор $\mathcal{A}$ действует на любой вектор так же, как оператор $\sum\limits_{i=1}^{k} \lambda_i \mathcal{P}_i$. Выражение $\mathcal{A} = \sum\limits_{i=1}^{k} \lambda_i \mathcal{P}_i$ называется ***спектральным разложением*** оператора $\mathcal{A}$.

Теорема 6.2. (***критерий диагонализируемости***) *Оператор диагонализируем тогда и только тогда, когда выполняются следующие два условия:*
1) *Характеристический многочлен раскладывается на линейные сомножители (то есть все его корни лежат в поле F);*
2) *Геометрическая кратность каждого собственного значения равна его алгебраической кратности.*

Доказательство. Диагонализируемость эквивалента наличию собственного базиса, откуда следует, что $V = \bigoplus_{i=1}^{n} V_{\lambda_i}$ (объединение базисов собственных подпространств — собственный базис V), но тогда $n = \dim V = g(\lambda_1) + \ldots + g(\lambda_k)$. Но $m(\lambda_i) \geqslant g(\lambda_i)$, а $m(\lambda_1) + \ldots + m(\lambda_k) \leqslant \deg \chi_{\mathcal{A}} = n$. Отсюда следует, что $n = g(\lambda_1) + \ldots + g(\lambda_k)$ тогда и только тогда, когда $m(\lambda_i) = g(\lambda_i)$.

Пример 6.1. Пусть $\mathcal{A}$ — поворот на угол $\frac{\pi}{2}$ относительно начала координат декартовой плоскости. В ортонормированном базисе $\mathbf{i}, \mathbf{j}$ матрица оператора имеет вид

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Следовательно, $\chi_{\mathcal{A}}(t) = t^2 + 1$. Его корни не лежат в поле $\mathbb{R}$. Поэтому, оператор $\mathcal{A}$ не диагонализируем над $\mathbb{R}$. Если же рассмотреть ту же матрицу над полем $\mathbb{C}$, то она оказывается подобной матрице

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Пример 6.2. Пусть $\mathcal{D}$ — оператор дифференцирования в $\mathbb{R}[x]_1$. В базисе $x, 1$ его матрица оператора вид

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$\chi_{\mathcal{D}}(t) = t^2$. Его корень 0 кратности 2 лежит в поле $\mathbb{R}$. $\operatorname{Ker}(\mathcal{D} - 0\mathcal{E}) = \mathbb{R}[x]_0$, $\dim \operatorname{Ker}(\mathcal{D} - 0\mathcal{E}) = 1$. Оператор $\mathcal{D}$ не диагонализируем.

Лекция III

В., с. 258-264; **К.**, с. 90-96; **Ф.**, с. 335-339.

Диагональная матрица как раз и является «наиболее простым» видом матрицы линейного оператора. Если мы смогли привести матрицу оператора к диагональному виду, то легко можем найти любою степень оператора. Это открывает путь к вычислению, например, многочленов от операторов. Но как показывает предыдущий пример, линейный оператор может не иметь собственного базиса из даже в том случае, когда все корни его характеристического многочлена лежат в поле, над которым определено векторное пространство. Далее мы введём понятие корневого вектора, обобщающее понятие собственного вектора, и покажем, что в случае, когда спектр линейного оператора содержится в поле, над которым определено векторное пространство, пространство раскладывается в прямую сумму так называемых корневых подпространств.

§7. Корневые векторы и корневые подпространства

Определение 7.1. Вектор $x \in V$ называется ***корневым вектором*** оператора $\mathcal{A}$, отвечающим собственному значению $\lambda \in F$, если существует такое целое неотрицательное число k, что $(\mathcal{A} - \lambda\mathcal{E})^k x = 0$. Наименьшее такое k называется ***высотой*** корневого вектора x.

Замечание 7.1. Если x — корневой вектор высоты k, то $\widetilde{x} = (\mathcal{A} - \lambda\mathcal{E})x$ является корневым вектором высоты $k-1$.

Пример 7.1. а) Корневые векторы высоты 0 — нулевые векторы;

б) Корневые векторы высоты 1 — собственные векторы;

в) Каждый многочлен есть корневой вектор с собственным числом 0 оператора дифференцирования пространства многочленов, причём высота многочлена как корневого вектора равна $n+1$, где n — степень этого многочлена;

Пример 7.2. Пусть V — пространство бесконечно дифференцируемых функций на вещественной прямой с вещественными значениями, $\mathcal{D}$ — оператора дифференцирования. Тогда:

1) $f \in V$ — собственный с собственным значением λ: $\mathcal{D}f = f' = \lambda f$. Следовательно $\dfrac{f'}{f} = \lambda \Leftrightarrow \ln|f| = \lambda x + C \Leftrightarrow |f| = e^C e^{\lambda t} \Leftrightarrow f = Ce^{\lambda t}$, $C \in R \setminus \{0\}$.

2) Корневые векторы: положим $f(t) = e^{\lambda t} g(t)$, $g \in V$, тогда $(\mathcal{D} - \lambda\mathcal{E})f = f'(t) - \lambda f(t) = \lambda e^{\lambda t} g(t) + e^{\lambda t} g'(t) - \lambda f(t) = e^{\lambda t} g'(t)$. То есть f — корневой тогда и только тогда, когда существует такое k, что $g^{(k)}(t) = 0$, то есть $g \in \mathbb{R}[t]$. Таким образом, корневые векторы для оператора дифференцирования — это функции вида $e^{\lambda t} g(t)$, где $g(t)$ — многочлен. Высота такого корневого вектора равна $\deg g + 1$. Они называются квазимногочленами. Вспомните о них, когда будете изучать линейные дифференциальные уравнения ☺

Корневые векторы, отвечающие собственному значению λ, высоты $\leqslant k$ — это $\mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E})^k \leqslant V$. Возникает цепочка подпространств

$$V_\lambda = \mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E}) \leqslant \mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E})^2 \leqslant \ldots \mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E})^k \leqslant \ldots \leqslant V^\lambda,$$

где V^λ=$\{$все корневые векторы с собственным значение $\lambda\}$ — ***корневое подпространство*** с собственным значением λ:

$$V^\lambda = \bigcup_{i=1}^{\infty} \mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E})^i.$$

В конечномерном случае эта цепочка с некоторого момента стабилизируются — размерности подпространств растут до тех пор, пока мы не дойдём до размерности корневого подпространства. Будем считать, что $\dim V < \infty$.

Теорема 7.1. (***свойства корневых подпространств***)
1) V^λ *$\mathcal{A}$-инвариантно;*
2) $(\mathcal{A}-\lambda\mathcal{E})|_{V^\lambda} = \mathcal{N}$ — ***нильпотентный*** *оператор, то есть существует такое неотрицательное целое m, то $\mathcal{N}^m = \mathcal{O}$;*
3) $(\mathcal{A}-\mu\mathcal{E})|_{V^\lambda}$ *невырожден при $\mu \neq \lambda$;*
4) $\dim V^\lambda = m(\lambda)$ (*геометрический смысл алгебраической кратности*).

Доказательство.

1) Пусть $V_k^\lambda = \mathrm{Ker}(\mathcal{A}-\lambda\mathcal{E})^k$, тогда $V_\lambda \leqslant V_1^\lambda \leqslant V_2^\lambda \leqslant \ldots \leqslant V_m^\lambda = V^\lambda$. Заметим, что $(\mathcal{A}-\lambda\mathcal{E})V_k^\lambda \leqslant V_{k-1}^\lambda \leqslant V_k^\lambda$. То есть, V_k^λ инвариантно относительно $\mathcal{A}-\lambda\mathcal{E}$, следовательно, V_k^λ и $\mathcal{A}$-инвариантно. Это верно и для $V_m^\lambda = V^\lambda$.

2) Выберем в V^λ базис, согласованный с цепочкой подпространств V_i^k: $e_1, \ldots, e_{l_1}$ — базис V_1^λ, $e_1, \ldots, e_{l_1}, \ldots, e_{l_2}$ — базис V_2^λ и т.д., $e_1, \ldots, e_{l_m}$ — базис $V_m^\lambda = V^\lambda$. $\mathcal{N}(V_k^\lambda) \leqslant V_{k-1}^\lambda$ (положим $V_0^\lambda = \{0\}$). Из этого следует, что $\mathcal{N}e_j \in \langle e_1, \ldots, e_{j-1}\rangle$, следовательно $\mathcal{N}^{l_m} = \mathcal{O}$, где $l_m = \dim V^\lambda$.

3) В базисе $e_1, \ldots, e_{l_1}, \ldots, e_{l_m}$ матрица оператора $\mathcal{N}$ верхнетреугольная с нулями на главной диагонали (***верхненильтреугольная***), тогда матрица оператора $\mathcal{A}|_{V^\lambda} = \mathcal{N}+\lambda\mathcal{E}$ верхнетреугольная с λ-ми на главной диагонали, а матрица оператора $(\mathcal{A}-\mu\mathcal{E})|_{V^\lambda}$ верхнетреугольная с $\lambda-\mu$ на главной диагонали. Следовательно (так как $\lambda \neq \mu$) она невырожденная и, значит, оператор $(\mathcal{A}-\mu\mathcal{E})|_{V^\lambda}$ тоже невырожден.

4) Дополним базис V^λ до базиса всего пространства V. В этом базисе матрица оператора $\mathcal{A}$ имеет блочно-треугольный вид:

$$\begin{pmatrix} A|_{V^\lambda} & D \\ 0 & C \end{pmatrix}.$$

Тогда $\chi_\mathcal{A}(t) = \chi_{A|_{V^\lambda}}(t)\det(tE-C) = (t-\lambda)^{l_m}\det(tE-C)$. Нужно показать, что λ не является собственным значением оператора $\mathcal{C}$ в пространстве $\langle e_{l_m+1}, \ldots, e_n\rangle$ с матрицей C. Пусть существует такой вектор $0 \neq x \in \langle e_{l_m+1}, \ldots, e_n\rangle$, что

$\mathcal{C}x = \lambda x$. Это означает, что $\mathcal{A}x = \lambda x + y$, $y \in V^\lambda$. Следовательно, $(\mathcal{A} - \lambda\mathcal{E})x = y$ — корневой вектор, но тогда и x — корневой вектор, что противоречит определению V^λ.

Теорема 7.2. *Корневые подпространства, отвечающие различным собственным значениями, линейно независимы.*

Доказательство. Аналогично доказательству теоремы 6.1. См. **В.**, с. 260.

§8. Структура нильпотентных операторов

Пусть $\mathcal{N}$ — нильпотентный оператор, то есть существует такое неотрицательное целое m, что $\mathcal{N}^m = \mathcal{O}$. Наименьшее из таких m называют ***высотой*** нильпотентного оператора. Для него все векторы V — корневые с собственным значением 0, высоты не больше m.

Пример 8.1. Оператор дифференцирования в пространстве $\mathbb{R}[x]_n$ — нильпотентный высоты $n+1$.

В силу пункта 3 теоремы 7.1 изучение произвольного оператора сводится к изучению нильпотентного оператора на соответствующем корневом подпространстве.

Лемма 8.1. *Пусть $x \in V$ — вектор высоты $k > 0$. Тогда векторы $x, \mathcal{N}x, \ldots, \mathcal{N}^{k-1}x$ линейно независимы.*

Доказательство. Индукция по k. Если $k = 1$, в этом случае $x \neq 0$ и доказывать нечего. Пусть $\alpha_0 x + \alpha_1 \mathcal{N}x + \ldots + \alpha_{k-1}\mathcal{N}^{k-1}x = 0$. Применим оператора N к обеим частям равенства, получим $\alpha_0\mathcal{N}x + \alpha_1\mathcal{N}^2x + \ldots + \alpha_{k-2}\mathcal{N}^{k-1}x = 0$, так как $\mathcal{N}^k x = 0$. Пусть $\mathcal{N}x = y$, его высота $k-1$ и $\alpha_0 y + \alpha_1\mathcal{N}y + \ldots + \alpha_{k-2}\mathcal{N}^{k-2}y = 0$. Поскольку по предположению индукции векторы $y, \mathcal{N}y, \ldots, \mathcal{N}^{k-1}y$ линейно независимы, то $\alpha_0 = \ldots = \alpha_{k-2} = 0$. Но тогда и $\alpha_{k-1}\mathcal{N}x = 0$. Так как высота x равна k, то $\mathcal{N}^{k-1}x \neq 0$, значит, $\alpha_{k-1} = 0$, следовательно, $x, \mathcal{N}x, \ldots, \mathcal{N}^{k-1}x$ линейно независимы.

Определение 8.1. Подпространство $U = \langle x, \mathcal{N}x, \mathcal{N}^2x, \ldots, \rangle$ называется ***циклическим подпространством*** нильпотентного оператора $\mathcal{N}$, порождённым вектором x.

Циклическое подпространство U — наименьшее $\mathcal{N}$-инвариантное подпространство, содержащее x, $\dim U = k$, где k — высота вектора x.

Базис $U\colon x_1, x_2, \ldots x_k$, где $x_i = N^{k-i}x$. Такой базис называется ***жордановой цепочкой***: $0 \xleftarrow{\mathcal{N}} x_1 \xleftarrow{\mathcal{N}} x_2 \xleftarrow{\mathcal{N}} \ldots \xleftarrow{\mathcal{N}} x_{k-1} \xleftarrow{\mathcal{N}} x_k$, то есть первый вектор переходит при действии $\mathcal{N}$ в нулевой, второй — в первый и т.д, последний —

в предпоследний. Следовательно, матрица оператора $\mathcal{N}$ в базисе $x_1, x_2, \dots x_k$ имеет вид

$$J_k(0) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

называемый ***нильпотентной жордановой клеткой*** порядка k.

Пример 8.2. Пусть $\mathcal{N} = \mathcal{D}^2 = \dfrac{d^2}{dx^2} : \mathbb{R}[x]_7 \to \mathbb{R}[x]_7$, $f_i = \dfrac{x^k}{k!}$, $i = \overline{0,7}$ — базис $\mathbb{R}[x]_7$. Действие $\mathcal{N}$ на базисных векторах даёт следующие жордановы цепочки

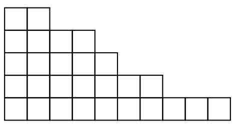
$$0 \xleftarrow{\mathcal{N}} f_1 \xleftarrow{\mathcal{N}} f_3 \xleftarrow{\mathcal{N}} f_5 \xleftarrow{\mathcal{N}} f_7,$$

$$0 \xleftarrow{\mathcal{N}} f_0 \xleftarrow{\mathcal{N}} f_2 \xleftarrow{\mathcal{N}} f_4 \xleftarrow{\mathcal{N}} f_6.$$

Теорема 8.1. (***основная теорема о структуре нильпотентного оператора***) *Пусть $\mathcal{N}$ — нильпотентный оператор на V. Тогда существует разложение пространства V в прямую сумму циклических подпространств этого оператора $V = \bigoplus U_i$. Количество слагаемых в таком разложении равно* $\dim \operatorname{Ker} \mathcal{N}$.

Пример 8.3. В предыдущем примере $\mathbb{R}[x]_7 = \langle f_1, f_3, f_5, f_7 \rangle \oplus \langle f_0, f_4, f_2, f_6 \rangle$. $\dim \operatorname{Ker} \mathcal{N} = \dim \langle f_0, f_1 \rangle = \dim \mathbb{R}[x]_1 = 2$. Нетрудно заметить, что в данном случае разложение в прямую сумму циклических — это разложение в прямую сумму подпространств чётных и нечётных многочленов степени не выше 7.

Наглядно можно изображать структуру нильпотентного оператора с помощью так называемой ***диаграммы Юнга***, которая в данном случае схематически показывает, как действует нильпотентный оператор на базисных векторах жорданова базиса:



С помощью такой диаграммы нильпотентный оператор задаётся однозначно. Квадратики — векторы жорданова базиса, нильпотентный оператор действует на них сверху вниз.

- Высота строки соответствует высоте базисного вектора, а высота произвольного вектора (линейной комбинации базисных) определяется как наибольшая высота базисного вектора, входящего в эту линейную комбинацию с ненулевым коэффициентом

- i-тый столбец соответствует жордановой цепочке — базису циклического пространства U_i
- Ядро оператора $\mathcal{N}^k$ — линейная оболочка векторов, стоящих в строках высоты не больше k
- Векторы, лежащие в нижней строке, при действии оператора переходят в нулевые

Пример 8.4. Для примеров 8.2 и 8.3 диаграмма Юнга имеет вид:

$$\begin{array}{|c|c|}\hline f_7 & f_6 \\ \hline f_5 & f_4 \\ \hline f_3 & f_2 \\ \hline f_1 & f_0 \\ \hline \end{array}$$

Высота вектора $2f_4 - 8f_1$ равна $\max\{3, 1\} = 3$, $\operatorname{Ker} \mathcal{N}^2 = \langle f_0, f_1, f_2, f_3 \rangle$.

Лекция IV

В., с. 262-264; **К.**, с. 97-99; **Ф.**, с. 337-340; **Г.**, с. 46-57.

§9. Жорданова нормальная форма оператора

Если вернуться к произвольному линейному оператору $\mathcal{A}$, то можно заметить, что на циклическом подпространстве нильпотентного оператора $\mathcal{N} = \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})|_{V^\lambda}$ оператор $\mathcal{A}$ задаётся матрицей

$$J(\lambda) = J(0) + \lambda E = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix},$$

называемой ***жордановой клеткой*** с собственным значением λ

Определение 9.1. ***Жордановой матрицей*** называется блочно-диагональная матрица

$$J = \begin{pmatrix} J_1 & & & \mathbf{O} \\ & J_2 & & \\ & & \ddots & \\ \mathbf{O} & & & J_k \end{pmatrix},$$

где $J_1, J_2, \dots J_k$ — какие-то жордановы клетки.

Жорданова матрица также называется ***жордановой нормальной формой*** (ЖНФ) для оператора $\mathcal{A}$. Верна следующая

Теорема 9.1. (***основная теорема о структуре оператора***) *Если характеристический многочлен $\chi_{\mathcal{A}}$ раскладывается на линейные сомножители, то существует базис, в котором матрица оператора $\mathcal{A}$ жорданова. Она определена однозначно, с точностью до перестановки жордановых клеток.*

Следствие 9.1.1. *Матрица любого оператора над полем комплексных чисел приводится к ЖНФ.*

Другими словами, любая комплексная матрица подобна жордановой.

В выборе жорданова базиса есть значительная свобода. Поймём, что количество жордановых клеток не зависит от выбора жорданова базиса. Достаточно сделать это для нильпотентного оператора.

Пусть $n(\lambda)$ — количество жордановых клеток в $J(N)$ с собственным значением λ. Оно равно количеству столбцов в диаграмме Юнга, то есть количеству векторов в нижней строке диаграммы. То есть $n(\lambda) = \dim \operatorname{Ker} \mathcal{N}$. С другой стороны $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) = V_\lambda \leqslant V^\lambda$, следовательно, $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) = \operatorname{Ker} \mathcal{N}$, поэтому

$$n(\lambda) = \dim \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}).$$

Пусть $n_k(\lambda)$ — число жордановых клеток размера $k \times k$ с собственным значением λ, они соответствуют столбцам высоты k. Так как ядро оператора $\mathcal{N}^k$ — это линейная оболочка векторов, стоящих в строках высоты не больше k, то их количество — это разница между количество вектор высоты не меньше k ($\dim \operatorname{Ker} \mathcal{N}^k$) и количество вектора высоты меньше k ($\dim \operatorname{Ker} \mathcal{N}^{k-1}$). А чтобы найти количество столбцов высоты ровно k, нужно из этого числа вычесть количество столбцов высоты не больше $k+1$, которое вычисляется аналогично ($\dim \operatorname{Ker} \mathcal{N}^{k+1} - \dim \operatorname{Ker} \mathcal{N}^k$). Итого получаем

$$n_k(\lambda) = \dim \operatorname{Ker} \mathcal{N}^k - \dim \operatorname{Ker} \mathcal{N}^{k-1} - (\dim \operatorname{Ker} \mathcal{N}^{k+1} - \dim \operatorname{Ker} \mathcal{N}^k),$$

то есть

$$n_k(\lambda) = 2 \dim \operatorname{Ker} \mathcal{N}^k - \dim \operatorname{Ker} \mathcal{N}^{k-1} - \dim \operatorname{Ker} \mathcal{N}^{k+1}.$$

Так как $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) \leqslant V^\lambda$, то $\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^k = \operatorname{Ker} \mathcal{N}^k$, и

$$n_k(\lambda) = 2 \dim \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^k - \dim \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^{k-1} - \dim \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^{k+1}.$$

Полученные формулы показывают, что число жордановых клеток не зависит от выбора жорданова базиса, следовательно, ЖНФ единственная с точностью до перестановки клеток.

Для удобства при ручном нахождении ЖНФ, запишем формулы в виде, содержащем ранги матриц. Пусть $r_k = \operatorname{rank}(\mathcal{A} - \lambda\mathcal{E})^k$, $n = \dim V$ = размер матрицы оператора. Тогда $\dim \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^k = n - r_k(\lambda)$. Следовательно,

$$n(\lambda) = n - r_1(\lambda),$$

$$n_k(\lambda) = r_{k-1}(\lambda) - 2r_k(\lambda) + r_{k+1}(\lambda).$$

Итак, если нам нужно найти ЖНФ оператора $\mathcal{A} \in \operatorname{End}(V)$, то:

1) Находим собственные значения $\lambda_1, \ldots \lambda_s$. Тогда $V = \bigoplus\limits_{i=1}^{s} V^{\lambda_i}$ и в согласованном базисе оператор имеет блочно-диагональную матрицу

$$\begin{pmatrix} A_1 & & & \mathbf{O} \\ & A_2 & & \\ & & \ddots & \\ \mathbf{O} & & & A_s \end{pmatrix};$$

2) Выясняем, как устроен каждый блок A_i для $\lambda_i = \lambda$. Для этого надо понять устройство оператора $\mathcal{A}|_{V^\lambda}$. Нам известно, что $\mathcal{A}|_{V^\lambda} = \lambda\mathcal{E} + \mathcal{N}$, $\mathcal{N}$ — нильпотентный оператор. Тогда $V^\lambda = \bigoplus\limits_{i=1}^{m} U_i$, U_i — циклические подпространства $\mathcal{N}$. Схема действия $\mathcal{N}$ на V^λ однозначно задаётся диаграммой Юнга;

3) В базисе V^λ, составленном из базисов U_i, матрица нильпотентного оператора N имеет блочно-диагональный вид, где на диагонали стоят нильпотентные жордановы клетки — получаем ***нильпотентную жорданову матрицу***. Матрица оператора $\mathcal{A}|_{V^\lambda} = \lambda\mathcal{E} + \mathcal{N}$ получается из неё прибавлением λ по главной диагонали;

4) Собираем из всех блоков A_i ЖНФ оператора $\mathcal{A}$.

Пример 9.1. Найдём ЖНФ оператора $\mathcal{A} \in \operatorname{End}\left(\mathbb{R}^4\right)$, заданного матрицей

$$A = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix},$$

Давайте одно из собственных значений угадаем. Заметим, что $\lambda_1 = 1$ является собственным значением, так как матрица $A - E$ вырожденная, поскольку имеет нулевой столбец:

$$A = \begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix}.$$

Найдём её ранг:

$$\operatorname{rank}\begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 3 & 0 & -3 \\ 0 & 3 & 0 & -1 \end{pmatrix} = 2,$$

$r_1(1) = \operatorname{rank}(A - E) = 2$. Тогда геометрическая кратность $\lambda = 1$ как минимум $\dim\operatorname{Ker}(\mathcal{A} - \mathcal{E}) = 2 = 4 - 2$, то есть ещё как минимум ещё одно собственное значение равно 1. Найдём оставшиеся собственные значения по теореме Виета:

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= \operatorname{tr} A, \\ \lambda_1\lambda_2\lambda_3\lambda_4 &= \det A, \end{cases} \quad \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 4, \\ \lambda_1\lambda_2\lambda_3\lambda_4 &= 1, \end{cases}$$

откуда $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$. Следовательно, всё пространство является корневым с собственным значением 1 и $n(1) = n - r_1(1) = 4 - 2 = 2$, то есть всего две жордановы клетки. Вычислим нужные ранги:

$$r_0(1) = \operatorname{rank}(A - E)^0 = \operatorname{rank} E = 4,$$

$$r_1(1) = \operatorname{rank}(A - E)^1 = 2,$$

$$r_2(1) = \operatorname{rank}(A - E)^2 = 1,$$

$$r_3(1) = \operatorname{rank}(A - E)^3 = 0.$$

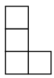
Тогда

$$n_1(1) = r_0(1) - 2r_1(1) + r_2(1) = 1,$$
$$n_2(1) = r_1(1) - 2r_2(1) + r_3(1) = 0,$$
$$n_3(1) = r_2(1) - 2r_3(1) + r_4(1) = 1.$$

То есть у нас одна жорданова клетка размером 1×1 и одна жорданова клетка размером 3×3, то есть оператор $\mathcal{A}$ имеет следующую ЖНФ:

$$J(\mathcal{A}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Диаграмма Юнга нильпотентного оператора $\mathcal{A} - \mathcal{E}$ имеет вид



Пример 9.2. Найдём ЖНФ оператора из примера 8.2. Оператор нильпотентен. Достаточно одного взгляда на диаграмму Юнга, чтобы написать ЖНФ:

$$\begin{pmatrix} J_4(0) & \mathbf{O} \\ \mathbf{O} & J_4(0) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

§10. Построение жорданова базиса

Как найти сам жорданов базис? Есть два разных подхода:

- Начать с собственных векторов и подниматься вверх
- Начать с корневых векторов наибольшей высоты и спускаться вниз

Далее будет показать второй подход. Одной из причин этого является то, что так действуют системы компьютерной алгебры. Дело в том, что подходящих «самых высоких» векторов оказывается столь много, что с большой вероятностью подойдут какие-то векторы стандартного базиса (ниже мы это видим) или их нехитрые линейные комбинации.

Пример 10.1. Продолжим предыдущий пример. Отметим, что нумерация векторов жорданова базиса должна соответствовать выбранной перестановке клеток в ЖНФ! Так, если рассмотреть матрицу нильпотентного оператора

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

то первый вектор при его действии должен переходит в нулевой, второй — тоже в нулевой, третий — во второй, четвёртый — в третий:

$$\begin{array}{|c|c|} \cline{1-1} e_4 & \multicolumn{1}{c}{} \\ \cline{1-1} e_3 & \multicolumn{1}{c}{} \\ \hline e_2 & e_1 \\ \hline \end{array}$$

Итак, самый высокий столбец в нашей диаграмме Юнга высоты 3. «Самый высокий» вектор должен обнулять матрицу $(A-E)^3$ (в данном случае это выполнится автоматически, так как всё пространство — корневое для нильпотентного оператора высоты 3) и не должен обнулять матрицу $(A-E)^2$. Так как

$$(A-E)^2 = \begin{pmatrix} 3 & 9 & 0 & -18 \\ 1 & 3 & 0 & -6 \\ 3 & 9 & 0 & -18 \\ 1 & 3 & 0 & 6 \end{pmatrix},$$

то в качестве «самого высокого» вектора e_4 жорданова базиса можно взять первый, второй или четвёртый векторы стандартного базиса, так как их образы не являются нулевыми (именно они записаны в первом, втором и четвёртом столбцах матрицы). Пусть $e_4 = (1,0,0,0)^T$. Тогда $(A-E)e_4 = e_3$, $(A-E)^2 e_4 = (A-E)e_3 = e_2$:

$$e_3 = \begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 0 \\ -1 \end{pmatrix}.$$

Ровно это мы и ожидали получить ☺

$$e_2 = \begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 1 \end{pmatrix}.$$

Осталось найти e_1. Он дополняет вектор e_2 до базиса $\mathrm{Ker}(\mathcal{A}-\mathcal{E})$, поэтому можно взять любой вектор из этого ядра, линейно независимый с e_2, например, $e_1 = (0,0,1,0)^T$.

Итак, у нас следующий жорданов базис: $e_1 = (0,0,1,0)^T$, $e_2 = (3,1,3,1)^T$, $e_3 = (0,-2,0,-1)^T$, $e_4 = (1,0,0,0)$. Проверим себя. Матрица перехода от стандартного базиса к жорданову имеет вид

$$C = \begin{pmatrix} 0 & 3 & 0 & 1 \\ 0 & 1 & -2 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix},$$

тогда

$$C^{-1}AC = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Подумайте, как изменится нумерация векторов при перестановке жордановых клеток в ЖНФ и проверьте себя.

Замечание. Вообще, если ЖНФ J найдена, то найти матрицу перехода можно с помощью матричного уравнения $XJ - AX = 0$, а столбцы матрицы перехода и есть векторы жорданова базиса. Решение данного матричного уравнения сводится к решению однородной СЛАУ. Пространство её решений, вообще говоря, многомерно. Из решений можно взять любое, которое даёт невырожденную матрицу.

Лекция V

В., с. 265-276; **К.**, с. 67-69, 301-304; **Ф.**, с. 142, 321-323; **Г.**, с. 57-59.

§11. Многочлены от линейного оператора

Определение 11.1. Пусть $f = a_m t^m + a_{m-1} t^{m-1} + \ldots + a_1 t + a_0 \in F[t]$, $\mathcal{A} \in \operatorname{End}(V)$. Тогда

$$\operatorname{End}(V) \ni f(\mathcal{A}) = a_m \mathcal{A}^m + a_{m-1} \mathcal{A}^{m-1} + \ldots + a_1 \mathcal{A} + a_0 \mathcal{E}.$$

Если у оператора $\mathcal{A}$ в некотором базисе матрица $A \in M_n(F)$, то $f(A)$ в том же базисе имеет матрицу $a_m A^m + a_{m-1} A^{m-1} + \ldots + a_1 A + a_0 E \in M_n(F)$.

Говорят, что многочлен f аннулирует оператор $\mathcal{A}$ (матрицу A), если $f(\mathcal{A}) = 0$ ($f(A) = 0$).

Определение 11.2. ***Минимальный многочлен*** $\mu_{\mathcal{A}}(t)$ оператора/матрицы — аннулирующий многочлен наименьшей степени со старшим коэффициентом единица.

Если из контекста ясно, о каком операторе идёт речь, нижний индекс будем опускать.

Лемма 11.1. 1) *Для любого линейного оператора $\mathcal{A}$ существует и единственен минимальный многочлен;*
2) *Для любого многочлена $f \in F[t]$ верно следующее: $f(\mathcal{A})$ равно остатку от деления f на $\mu(t)$;*
3) *Многочлен f аннулирует оператора $\mathcal{A}$ тогда и только тогда, когда f делится на $\mu(t)$.*

Доказательство.
1) Существование: Пусть $\dim V = n$, тогда $\operatorname{End}(V) \simeq M_n(F)$ и $\dim \operatorname{End}(V) = n^2$. Тогда система $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots \mathcal{A}^{n^2}$ линейно зависима, то есть существует нетривиальная линейная комбинация, равная нулю: $a_0\mathcal{E} + a_1\mathcal{A} + a_2\mathcal{A}^2 + \ldots + a_{n^2}\mathcal{A}^{n^2} = 0$. Следовательно, $a_0 + a_1 t + a_2 t^2 + \ldots + a_{n^2} t^{n^2}$ как раз ненулевой аннулирующий многочлен. Выберем из всех таких многочленов многочлен наименьшей степени. Поделив его на старший коэффициент, можно считать, что старший коэффициент равен единице.

Единственность: Пусть есть два различных минимальных многочлена $\mu(t) = b_0 + b_1 t + \ldots + b_{m-1} t^{m-1} + t^m$ и $\widetilde{\mu}(t) = \widetilde{b}_0 + \widetilde{b}_1 t + \ldots + \widetilde{b}_{m-1} t^{m-1} + t^m$. Тогда $\mu(t) - \widetilde{\mu}(t)$ — ненулевой (поскольку есть несовпадающие коэффициенты) аннулирующий многочлен меньшей степени. Пришли к противоречию.

2) Поделим f на μ с остатком: $f = \mu g + r$, $\deg r < \deg \mu$. Тогда $f(\mathcal{A}) = \underbrace{\mu(\mathcal{A})}_{0} g(\mathcal{A}) + r(\mathcal{A}) = r(\mathcal{A})$.

3) $f(\mathcal{A}) = 0$ тогда и только тогда, когда $f(\mathcal{A}) = 0$ (предыдущий пункт), следовательно, $r \equiv 0$, так как $\deg r < \deg \mu$, а ненулевым r быть не может, так как это противоречило бы минимальности μ.

Если матрица оператора $\mathcal{A}$ имеет блочно-диагональный вид

$$\begin{pmatrix} A_1 & & & \mathbf{O} \\ & A_2 & & \\ & & \ddots & \\ \mathbf{O} & & & A_s \end{pmatrix},$$

то всё пространство раскладывается в прямую сумму инвариантных подпространств, на каждом из которых ограничение оператора имеет матрицу A_i, поэтому на всём пространстве оператор $f(\mathcal{A})$ имеет матрицу

$$\begin{pmatrix} f(A_1) & & & \mathbf{O} \\ & f(A_2) & & \\ & & \ddots & \\ \mathbf{O} & & & f(A_s) \end{pmatrix}.$$

Условие $f(\mathcal{A}) = 0$ эквивалентно тому, что $f(A_i) = 0$. А это эквивалентно тому, что f делится на $\mu_{\mathcal{A}_i}$. Многочлен наименьшей степени, удовлетворяющий этому условию, это НОК$(\mu_{\mathcal{A}_1}, \mu_{\mathcal{A}_2}, \dots, \mu_{\mathcal{A}_s}) = \mu$ — минимальный многочлен оператора $\mathcal{A}$.

Лемма 11.2. 1) *$f \in F[t]$, $f(t) = c_0 + c_1(t-\lambda) + c_2(t-\lambda)^2 + ...c_s(t-\lambda)^s$ — разложение по степеням $t-\lambda$. $J_k(\lambda)$ — жорданова клетка порядка k с собственным значением λ. Тогда*

$$f(J_k(\lambda)) = \begin{pmatrix} c_0 & c_1 & c_2 & \dots & \dots & c_{k-1} \\ 0 & \ddots & \ddots & \ddots & \ddots & \dots \\ \ddots & \ddots & c_0 & c_1 & \ddots & \dots \\ & & 0 & c_0 & \ddots & c_2 \\ & & & \ddots & \ddots & c_1 \\ \mathbf{O} & & & \ddots & 0 & c_0 \end{pmatrix}$$

— нильверхнетреугольная матрица.
2) *Минимальный многочлен жордановой клетки $J_k(\lambda)$ равен $(t-\lambda)^k$.*

Доказательство. 1) $J_k(\lambda) = \lambda E + J_k(0)$. Тогда $f(J_k(\lambda)) = c_0 E + c_1 J_k + c_2 J_k^2 + \ldots + c_s J_k^s$. $J_k(0)$ — матрица нильпотентного оператора $\mathcal{N}$, который k-тый базисный вектор переводит в $(k-1)$-ый. Следовательно, при $i > j$ вектор e_i переводится оператором N^j в e_{j-i}, а при $i \leqslant j$ — в нулевой. Отсюда получаем доказываемый вид матрицы.

2) f аннулирует $J_k(\lambda)$ тогда и только тогда, когда $c_0 = c_1 = \ldots = c_{k-1} = 0$, что равносильно тому, что f делится на $(t-\lambda)^k$. Следовательно, искомый минимальный многочлен имеет вид $(t-\lambda)^k$.

Если мы рассматриваем оператор $\mathcal{A}$ векторное пространство над полем $\mathbb{C}$ (в нём каждый многочлен раскладывается на линейные сомножители), то непосредственно из полученного выше следует

Теорема 11.1.
$$\mu_{\mathcal{A}}(t) = (t-\lambda_1)^{k_1}(t-\lambda_2)^{k_2}\ldots(t-\lambda_l)^{\lambda_l},$$
где λ_i — все различные собственные значения оператора $\mathcal{A}$, k_i — максимальный размер жордановой клетки с собственным значением λ_i.

Следствие 11.1.1. *ЖНФ оператора диагональна тогда и только тогда, когда его минимальный многочлен не имеет кратных корней.*

Теорема 11.2. (***теорема Гамильтона-Кэли***)

$$\chi_{\mathcal{A}}(\mathcal{A}) = 0,$$

то есть характеристический многочлен оператора является аннулирующим для него.

Доказательство. $\chi_A(t) = (t-\lambda_1)^{m_1}\ldots(t-\lambda_l)^{m_l}$, $l_i = dim V^{\lambda_i}$. Заметим, что $m_i \geqslant k_i$, поскольку k_i — это наибольшая высота корневого вектора собственным значением λ_i, f m_i — сумма размеров жордановых клеток с тем же собственным значением. Следовательно, характеристический многочлен делится на минимальный и поэтому является аннулирующим.

Замечание. В доказательстве мы воспользовались разложимость характеристического многочлена на линейные сомножители, однако теорема Гамильтона-Кэли верна над любым полем.

В частности, для линейного оператора $\mathcal{A}$ в двумерном пространстве получается
$$\mathcal{A}^2 - (\operatorname{tr}\mathcal{A})\mathcal{A} + (\det\mathcal{A})\mathcal{E} = \mathcal{O}.$$

Пример 11.1. Найдём все линейные операторы $\mathcal{A}$ в двумерном пространстве такие, что $\mathcal{A}^3 = \mathcal{A}^2$. $\mathcal{A}^3 - \mathcal{A}^2 = 0$ означает, что многочлен $t^3 - t^2$ является для оператора A аннулирующим. Значит, минимальный многочлен делит $t^3 - t^2 = t^2(t-1)$ и его степень не больше двух (поскольку пространство двумерно). Тогда в ЖНФ оператора могут быть только клетки вида $J_2(0)$, $J_1(0)$, $J_1(1)$, сумма их порядков равна двум. То есть подходят операторы с ЖНФ

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Пусть $f(t) = \mu(t)g(t) + r(t)$, $\deg f = s$, тогда $f(t) - r(t)$ делится на $\mu(t)$, то есть многочлен $f(t) - r(t)$ имеет корни λ_i, кратности которых не ниже, чем k_i соответственно. Вспомним, что корень многочлена кратности k — это и корень всех его производных вплоть до $(k-1)$-го порядка. Тогда $f^{(k)}(\lambda_i) = r^{(k)}(\lambda_i)$ для всех k от 0 до $k_i - 1$. А это СЛАУ размера $s \times s$ на коэффициенты многочлена $r(t)$.

Вместо минимального многочлена можно использоваться характеристический, однако его степень, естественно, может быть больше степени минимального и нужно решать СЛАУ большего размера.

Пример 11.2. Найдём A^{2024}, где $A = \begin{pmatrix} 5 & 3 & -8 \\ 6 & 2 & -8 \\ 6 & 3 & -9 \end{pmatrix}$. Найдём собственные значения: $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = 0$. $n(-1) = 3 - \operatorname{rank}(A + E) = 3 - 1 = 2$. Тогда ЖНФ имеет вид

$$J = \begin{pmatrix} J_1(-1) & & \mathbf{O} \\ & J_1(-1) & \\ \mathbf{O} & & J_1(0) \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Тогда $\mu(t) = (t+1)t$ (в данном случае степень минимального многочлена меньше степени характеристического). Поделим $f(t) = t^{50}$ на $\mu(t)$ с остатком: $f(t) = \mu(t)g(t) + \underbrace{At + B}_{r(t)}$. Тогда

$$\begin{cases} f(-1) &= r(-1), \\ f(0) &= r(0); \end{cases} \Rightarrow \begin{cases} -A + B &= 1, \\ B &= 0; \end{cases}$$

Следовательно. $r(t) = -t$, тогда

$$r(A) = -A \Rightarrow A^{2024} = \begin{pmatrix} -5 & -3 & 8 \\ -6 & -2 & 8 \\ -6 & -3 & 9 \end{pmatrix}.$$

§12. Аналитические функции от линейного оператора

Данный способ работает и для нахождения произвольных аналитических функций от операторов/матриц. ***Аналитическая функция*** f задаётся сходящимся степенным рядом $f(t) = c_0 + c_1(t-t_0) + c_2(t-t_0)^2 + \ldots + c_s(t-t_0)^s + \ldots$, $t_0 \in \mathbb{C}$. В данный момент нас не интересуют вопросы сходимости, ими вы займётесь на математическом анализе ☺

Если считать, что со сходимостью для функции f всё в порядке (собственные значения оператора лежат в так называемом круге сходимости), то $f(A) = \lim\limits_{s\to\infty} f_s(A) = \lim\limits_{s\to\infty} r_s(A)$, где $f_s(t) = c_0 + c_1(t-t_0) + c_2(t-t_0)^2 + \ldots + c_s(t-t_0)^s$, $r_s(t)$ — остаток от деления f_s на минимальный многочлен. Коэффициенты остатков r_s стремятся к некоторым предельным значениям при $s \to \infty$. Это следует из формул Крамера для решения квадратных СЛАУ и возможности почленного дифференцирования аналитических функций. То есть $r_s(t)$ стремится к некоторому многочлену $r(t)$, удовлетворяющего СЛАУ, аналогичной СЛАУ в случае, когда f — многочлен.

Пример 12.1. Найдём e^A, где $A = \begin{pmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -1 \end{pmatrix}$. Найдём собственные значения: $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 1$. $n(0) = 3 - r_1(0) = 3 - \operatorname{rank} A = 3 - 2 = 1$. Тогда ЖНФ имеет вид

$$J = \begin{pmatrix} J_2(0) & 0 \\ 0 & J_1(1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Тогда $\mu(t) = t^2(t-1)$ (в данном случае минимальный многочлен совпадает с характеристическим). Тогда $e^A = r(A)$ и $\deg r(t) < 3$, то есть $r(t) = At^2 + Bt + C$.

$$\left\{ \begin{array}{rcl} r(0) & = & f(0), \\ r'(0) & = & f'(0), \\ r(1) & = & f(1); \end{array} \right. \Rightarrow \left\{ \begin{array}{rcccl} C & = & e^0 & = & 1, \\ B & = & e^0 & = & 1, \\ A + B + C & = & e^1 & = & e; \end{array} \right.$$

Следовательно. $r(t) = (e-1)t^2 + t + 1$, тогда

$$e^A = r(A) = (e-1)A^2 + A + E = \begin{pmatrix} 3e-1 & e & -3e+1 \\ 3e & e+3 & -3e-3 \\ 3e-1 & e+1 & -3e \end{pmatrix}.$$