Лекция 6
Разложение пространства с оператором

▷Факторизация оператора ▷Минимальный многочлен ▷Разложение пространства по разложению аннулирующего многочлена ▷Корневые подпространства ▷Разложение Жордана ▷Структура нильпотентного оператора ▷Функции от оператора и их вычисление приведением по модулю характеристического многочлена ▷Экспонента и логарифм ▷Пространство с оператором как $k[t]$-модуль ▷Элементарные делители

В этой лекции мы обсудим действие произвольного линейного оператора $F: V \longrightarrow V$ на произвольном векторном пространстве V, не предполагая наличия каких-либо дополнительных структур на V. Иногда мы будем требовать алгебраическую замкнутость основного поля k (при исследовании разложения Жордана и вычислении алгебраических функций от оператора), но значительная часть наших рассмотрений будет сохранять силу над любым полем.

§1. Геометрия пространства с оператором

1.1. Факторизация оператора. Я напомню (см. лекцию 5 прошлого семестра), что по любому подпространству W векторного пространства V можно построить *фактор пространство* V/W. Его векторами, по определению, являются всевозможные параллельные сдвиги подпространства $\mathrm{Aff}(W)$ в $\mathrm{Aff}(V)$. Они обозначаются через

$$[v]_W = v + W = \{v + w \text{ по всем } w \in W\}$$

или просто через $[v]$, если ясно, что за W имеется в виду.

Подмножество $[v] \in V$ может быть описано также *как класс смежности* вектора v по модулю подпространства W. Отмечу, что равенство $[v] = [u]$ равносильно условию $u - v \in W$ (такие u и v называются *сравнимыми по модулю* W). Операции над векторами фактор-пространства определяются правилами

$$[v] + [u] \stackrel{\text{def}}{=} [v + u] \qquad \lambda[v] \stackrel{\text{def}}{=} [\lambda v] \qquad \forall u, v \in V \ \forall \lambda \in k.$$

Задача 6.1. Убедитесь, что эти определения корректны, и что V/W является векторным пространством.

Задача 6.2. Пусть векторы $e_1, e_2, \dots, e_k$ составляют базис в подпространстве W, а классы векторов $[e_{k+1}]$, $[e_{k+2}]$, …, $[e_n]$ составляют базис в фактор-пространстве V/W. Докажите, что векторы $e_1, e_2, \dots, e_n$ составляют базис в V.

Если на пространстве V действует линейный оператор $F: V \longrightarrow V$, то он будет индуцировать линейный оператор $\tilde F: V/W \longrightarrow V/W$ на факторе по подпространству W при условии, что это подпространство является *инвариантным для* F (т.е. будет переходить под действием F в себя). Оператор $\tilde F$ называется *фактор-оператором* F *по модулю инвариантного подпространства* W и определяется формулой $\tilde F([v]) = [F(v)]$.

Задача 6.3. Убедитесь, что это определение корректно (при условии, что W инвариантно).

Задача 6.4. Пусть $\{e_1, e_2, \dots, e_n\}$ — базис в V, такой что $\{e_1, e_2, \dots, e_k\}$ будет базисом в W. Убедитесь, что F-инвариантность W равносильна тому, что матрица F_e имеет блочный вид $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, где 0 — нулевая матрица размера $(n-k)\times k$. Убедитесь также, что

а) блок A будет матрицей ограничения оператора F на W в базисе $\{e_1, e_2, \dots, e_k\}$;

б) классы $[e_\nu]$ с $\nu \geq k+1$ составят базис в V/W и блок C будет матрицей $\tilde F$ в этом базисе.

1.2. Аннулирущие многочлены. Для произвольного многочлена $g(t) \in k[t]$:

$$g(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_m t^m$$

определим результат подстановки в него оператора $F \in \mathrm{Hom}(V, V)$ как оператор

$$g(F) = a_0 \cdot E + a_1 \cdot F + a_2 \cdot F^2 + \dots + a_m \cdot F^m \in \mathrm{Hom}(V, V).$$

Отображение $\mathrm{ev}_F: k[t] \longrightarrow \mathrm{Hom}(V, V): g(t) \mapsto g(F)$ является гомоморфизмом колец. Его образ, состоящий из всевозможных операторов вида $g(F)$, обозначается через $k[F]$. Его ядро $\ker \mathrm{ev}_F \subset k[t]$ состоит из всех многочленов $q(t)$, для которых $q(F)$ является нулевым оператором (т.е. отображает всё V в нулевой вектор). Такие многочлены называются *аннулирующими* для F.

Задача 6.5. Убедитесь, что аннулирующие многочлены составляют идеал в $k[t]$.

Этот идеал не совпадает со всем кольцом многочленов, поскольку не содержит константы, и отличен от нуля, поскольку в конечномерном векторном пространстве $\mathrm{Hom}(V, V)$ бесконечное семейство векторов $F^0 = E$, F^1, F^2, … должно быть линейно зависимо:

$$a_0 \cdot E + a_1 \cdot F + a_2 \cdot F^2 + \dots + a_d \cdot F^d = 0$$

для некоторого $d \in \mathbb{N}$ и подходящих a_ν (не всех нулевых).

Так как $k[t]$ — кольцо главных идеалов, множество многочленов, которые аннулируют данный оператор F, состоит из всех многочленов, делящихся на многочлен $\mu_F(t)$, однозначно определямый как приведённый многочлен наименьшей степени, аннулирующий F. Этот многочлен $\mu_F(t)$ называется *минимальным многочленом* оператора F.

1.2.1. Утверждение. *Степень минимального многочлена не превышает размерности пространства, на котором действует оператор:*

$$\deg \mu_F \leq \dim V.$$

Замечание. Достаточно убедиться, что всякий оператор на V аннулируется некоторым многочленом степени $\dim V = n$. Это немедленно следует из доказанного в конце последней лекции прошлого семестра *тождества Гамильтона-Кэли*, утверждающего, что при подстановке матрицы в её характеристический многочлен получается нулевая матрица (и это верно над любым коммутативным кольцом!). Иначе говоря, $\chi_F(F) = 0$, т.е. *оператор аннулируется своим характеристическим многочленом*. Тем не менее, мы укажем здесь независимое геометрическое построение многочлена степени $\dim V = n$, аннулирующего F.

Доказательство. Индукция по $\dim V$. Если $\dim V = 1$, то $F = \lambda E$ — гомотетия и аннулируется линейным двучленом $t - \lambda$. Если $\dim V > 1$, рассмотрим произвольный вектор $v \neq 0$ и наименьшую натуральную степень F^m, для которой вектор $F^m v$ линейно выражается через предыдущие степени:

$$F^m v = a_0 v + a_1 F v + \cdots + a_{m-1} F^{m-1} v$$

(здесь допускается, чтобы все a_ν были нулевыми!).

Векторы v, Fv, …, $F^{m-1}v$ будут линейно-независимы, так что их линейная оболочка W является F-инвариантным m-мерным подпространством. Ограничение F на это подпространство аннулируется ненулевым многочленом степени m

$$f(t) = t^m - a_{m-1} t^{m-1} - \cdots - a_0.$$

Действительно, $f(F)v = 0$, а значит и $f(F)F^r v = F^r f(F) v = 0$ для каждого r.

По предположению индукции фактор-оператор $\tilde{F}$ по модулю W аннулируется неким многочленом $g(t)$ степени $\dim(V/W) = \dim V - m$. Равенство $g(\tilde{F})\big[V/W\big] = [0]$ означает, что $g(F)V \subset W$. Следовательно, многочлен $f(t)g(t)$ аннулирует F и имеет нужную степень. □

1.3. Разложение пространства по разложению аннулирующего многочлена. Соотношение вида $g(F) = 0$ с $g \in k[t]$ всегда доставляет геометрическую информацию о действии F. Мы уже видели это на примере оператора с квадратом единица $F: \ F^2 = E$, аннулируемого многочленом $t^2 - 1 = (t+1)(t-1)$. Действительно, из соотношения

$$(F+E)(F-E) = 0$$

следуют включения

$$\begin{aligned} \operatorname{im}(F+E) \quad &\subset \quad V_+ = \ker(F-E) = \{v \in V \mid Fv = v\} \\ \operatorname{im}(F-E) \quad &\subset \quad V_- = \ker(F+E) = \{v \in V \mid Fv = -v\}. \end{aligned}$$

Разложение[1] $E = -\frac{1}{2}(F-E) + \frac{1}{2}(F+E)$ позволяет представить любой вектор $v \in V$ в виде

$$v = \frac{1}{2}(F+E)v - \frac{1}{2}(F-E)v = v_+ + v_-$$

[1] Мы здесь молчаливо предполагаем, что $\operatorname{char}(k) \neq 2$, т.к. иначе $t^2 - 1 = (t-1)^2$ и $t - 1 = t + 1$.

с $v_+ \in V_+$ и $v_- \in V_-$. Вместе с очевидным равенством $V_+ \cap V_- = 0$ это даёт разложение $V = V_+ \oplus V_-$ в прямую сумму собственных подпространств оператора F с собственными значениями ± 1 (правда, не гарантирует, что оба они ненулевые). В частности, F диагонализуем.

Непосредственным обобщением этого примера является следующее

1.3.1. Утверждение. *Пусть оператор $F \in \mathrm{Hom}(V, V)$ аннулируется некоторым многочленом $q \in k[t]$, который можно разложить в $k[t]$ в произведение s попарно взаимно-простых множителей*

$$q(t) = q_1(t) \cdot q_2(t) \cdot \cdots \cdot q_s(t),$$

и пусть $Q_j(t) = q(t)/q_j(t)$, $(j = 1, 2, \ldots, s)$ — всевозможные произведения по $(s-1)$ из них. Тогда

(i) $\ker q_j(F) = \mathrm{im}\, Q_j(F)$ и является F-инвариантным подпространством (возможно, нулевым);

(ii) сумма всех ненулевых подпространств из (I) является прямой и совпадает с V.

Доказательство. Прежде всего, мы имеем $\forall j$ включение

$$\mathrm{im}\, Q_j(F) \subset \ker q_j(F),$$

которое вытекает из равенства $0 = q(F)V = q_j(F)Q_j(F)V$.

Далее, подпространства $\ker q_j(F)$ попарно трансверсальны. Это вытекает из взаимной простоты q_i и q_j при $i \neq j$: подберём такие многочлены $a, b \in k[t]$, чтобы выполнялось тождество

$$a(t)q_i(t) + b(t)q_j(t) = 1,$$

подставим в обе части оператор F и применим к вектору $v \in \ker q_i \cap \ker q_j$. Получим

$$v = a(F)q_i(F)v + b(F)q_j(F)v = a(F)0 + b(F)0 = 0.$$

С другой стороны, сумма подпространств $\mathrm{im}\, Q_j(F)$ составляет всё V. Действительно, $\text{НОД}(Q_1, Q_2, \ldots, Q_s) = 1$ и можно подобрать такие $g_\nu \in k[t]$, что $\sum\limits_{\nu=1}^{s} Q_\nu(t)g_\nu(t) = 1$. Подставляя в это равенство F и применяя его к произвольному $v \in V$ получаем разложение

$$v = \sum v_\nu, \quad \text{где } v_\nu = Q_\nu(F)g_\nu(F)v \in \mathrm{im}\, Q_\nu(F).$$

Из этих трёх фактов и вытекает предложение.

1.3.2. Пример: проекторы. Оператор $F:\ V \longrightarrow V$ называется *проектором*, если существует разложение $V = U \oplus W$, такое что F переводит любой вектор $v = (u, w) \in U \oplus W$ в $(0, w)$. В этом случае говорят также, что F проектирует V на W вдоль U.

Из доказанного выше предложения вытекает, что проекторами являются те и только те операторы F, которые аннулируются многочленом $t^2 - t$, т.е. удовлетворяют соотношению $F^2 = F$. Действительно, в этом случае

$$V = \ker(F) \oplus \ker(F - E) \quad \text{и} \quad \ker(F - E) = \mathrm{im}\,(F),$$

т.е. всё пространство распадается в прямую сумму ядра и образа, и действие оператора на образе тождественно.

Другими следствиями из доказанного предложения являются более мощные, чем из лекции 4, критерии диагонализуемости оператора

1.3.3. Следствие. *Если оператор $F \in \operatorname{Hom}_k(V,V)$ аннулируется многочленом без кратных корней, полностью разлагающимся над k на линейные множители, то F диагонализуем.*

Доказательство. Равенство $(F-\lambda_1 E)(F-\lambda_2 E)\cdots(F-\lambda_s E)=0$ при условии, что все λ_ν попарно различны даёт разложение $V=\bigoplus \ker(F-\lambda_\nu E)$ (где суммирование идёт по всем ненулевым из написанных подпространств), что и требуется.

> Задача 6.6. Убедитесь, что верно и обратное: диагонализуемый оператор аннулируется произведением $\prod(t-\lambda_\nu)$, в котором присутствуют без повторений все его различные собственные значения.

1.3.4. Следствие. *Для диагонализуемости оператора над некоторым полем необходимо и достаточно, чтобы его минимальный многочлен полностью разлагался над этим полем на линейные множители и не имел кратных корней.*

> Задача 6.7. Докажите, что ограничение диагонализуемого оператора на любое инвариантное подпространство будет диагонализуемо на этом подпространстве.

> Задача 6.8. Докажите, что произвольная совокупность попарно перестановочных диагонализуемых операторов может быть приведена к диагональному виду одновременно в одном общем базисе
>
> Указание: Воспользуйтесь индукцией по размерности пространства, леммой о коммутирующих операторах и предыдущей задачей.

В качестве последнего следствия отметим, что над алгебраически замкнутым полем, характеристика которого не делит n, будет диагонализуем всякий оператор, удовлетворяющий уравнению $F^n=E$. Поскольку в конечной группе G всякий элемент удовлетворяет уравнению $F^{|G|}=E$, мы получаем

1.3.5. Следствие. *Пусть поле k алгебраически замкнуто, G — конечная подгруппа в $GL_k(V)$, и её порядок $|G|$ не делится на характеристику поля k. Тогда:*

(i) все операторы в G диагонализуемы;

(ii) если G — абелева, то они приводятся к диагональному виду одновременно в одном общем базисе.

§2. Разложение Жордана.

Всюду в этом параграфе мы будем предполагать, что основное поле алгебраически замкнуто. Наша цель состоит в том, чтобы выделить из произвольного оператора его «диагонализуемую составляющую», и выяснить как устроено то, что после этого останется.

2.1. Определение. *Пусть $F \in \operatorname{Hom}_k(V,V)$ и $\lambda \in k$ — любое число. Подпространство*

$$K_\lambda=\bigcup_{n\in\mathbb{N}}\ker(F-\lambda E)^n=\{v\in V \mid (F-\lambda E)^n v=0 \text{ для некоторого } n\in\mathbb{N}\}$$

*незывается **корневым подпространством** со значением λ.*

Задача 6.9. Убедитесь, что это и в самом деле подпространство!

2.1.1. Утверждение. *Всякое корневое подпространство K_λ F-инвариантно, отлично от нуля тогда и только тогда, когда λ является собственным значением оператора F, и трансверсально любому другому корневому подпространству K_μ с $\mu \neq \lambda$.*

Доказательство. F-инвариантнность K_λ проверяется выкладкой:

$$(F - \lambda E)^n v = 0 \Longrightarrow (F - \lambda E)^n F v = F(F - \lambda E)^n v = 0,$$

Если K_λ отлично от нуля, то оно содержит собственный вектор $v \in K_\lambda$, такой что $v \neq 0$ и $Fv = \mu v$. Равенство

$$0 = (F - \lambda E)^n v = (\mu - \lambda)^n v$$

показывает, что его собственное значение μ должно совпадать с λ. Обратно, всякое собственное подпространство $V_\lambda = \ker(F - \lambda E)$ содержится в корневом подпространстве K_λ.

Предыдущие рассуждения показывают заодно, что при $\lambda \neq \mu$ оператор $(F - \mu E)$ (для которого K_λ является инвариантным подпространством, как и для F) не имеет ядра на K_λ, а стало быть, и он, и все его степени действуют на K_λ взаимно однозначно, т.е. $K_\lambda \cap K_\mu = 0$. $\square$

2.2. Корневое разложение. Разложим минимальный многочлен μ_F оператора F на линейные множители

$$\mu_F(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_s)^{m_s}$$

(где все λ_ν предполагаются попарно-разными). Полагая в предложении Утв.1.3.1 $q_j(t) = (t - \lambda_j)^{m_j}$, $Q_j(t) = \prod\limits_{\nu \neq j}(t - \lambda_\nu)^{m_\nu}$, получаем прямое F- инвариантное разложение

$$V = \bigoplus W_{\lambda_j}, \text{ где } W_{\lambda_j} = \ker(F - \lambda E)^{m_j} = \operatorname{im} Q_j(F).$$

Из минимальности μ_F следует, что все подпространства W_{λ_j} ненулевые, поскольку $0 = W_{\lambda_j} = \operatorname{im} Q_j(F)$ означало бы, что F можно аннулировать собственным делителем μ_F. Из включения $0 \neq W_{\lambda_j} \subset K_\lambda$ и трансверсальности корневых подпространств следует очень важный результат

2.3. Теорема. (Теорема о корневом разложении) *Если на векторном пространстве V над алгебраически замкнутым полем действует линейный оператор $F: V \longrightarrow V$, то V распадается в прямую сумму корневых подпространств*

$$V = \bigoplus K_{\lambda_j},$$

где суммирование производится по всем различным собственным значениям λ_j оператора F. Каждое корневое подпространство K_{λ_j} F-инвариантно и совпадает с $\ker(F - \lambda_j E)^{m_j}$, где

$$\prod_j (t - \lambda_j)^{m_j} = \mu_F(t)$$

есть минимальный многочлен оператора F.

Задача 6.10. Убедитесь, что множитель $(t-\lambda_j)^{m_j}$ из разложения минимального многочлена оператора F является минимальным многочленом для ограничения F на корневое подпространство K_{λ_j}.

2.3.1. Следствие. *Характеристический многочлен делится на минимальный и делит некоторую его степень.*

Доказательство. Из предыдущего следует, что $\chi_F(t)$ и $\mu_F(t)$ обладают одинаковым набором корней, поэтому достаточно проверить, что кратности этих корней у характеристического полинома не меньше, чем у минимального. Множитель $(t-\lambda)^m$ из разложения минимального многочлена является минимальным многочленом для ограничения оператора F на корневое подпространство K_λ, а значитб $m \leq \dim K_\lambda$. Характеристический многочлен оператора F равен произведению характеристических многочленов его ограничений на корневые подпространства (это очевидно, если воспользоваться для вычисления $\det(F-tE)$ базисом, согласованным с корневым разложением). Поэтому кратность вхождения $(t-\lambda)$ в $\chi_F(t)$ в точности равна $\dim K_\lambda$, что и требовалось.

Отметим, что из этого следствия вытекает тождество Гамильтона-Кэли для матриц над алгебраически замкнутым полем (наши геометрические рассмотрения проводились, напомню, независимо от этого тождества).

2.3.2. Следствие. $\chi_F(F)=0$

2.4. Определение. *Оператор $F:\ V \longrightarrow V$ называется нильпотентным, если $F^k=0$ для некоторого $k\in\mathbb{N}$ (т.е. если он аннулируется многочленом t^k).*

Задача 6.11. Докажите, что оператор тогда и только тогда нильпотентен, когда все его собственные значения нулевые.

2.5. Теорема. (Аддитивное разложение Жордана.) *Для любого оператора F, действующего в пространстве над алгебраически замкнутым полем существуют диагонализуемый оператор F_{ss} и нильпотентный оператор F_n, такие что:*

(i) $F=F_{ss}+F_n$

(ii) $F_{ss}F_n=F_nF_{ss}$.

Этими двумя свойствами F_{ss} и F_n определяются по F однозначно.

Замечание о терминологии и обозначениях. Индекс ss является сокращенгием от английского semisimple. Дело в том, что на языке теории представлений (а он в данном вопросе является наиболее употребительным) диагонализуемые операторы называются *полупростыми* (а операторы гомотетий — *простыми*), что очень удобно и естественно. Оператор F_{ss} называется *полупростой частью* оператора F, а F_n — *нильпотентной частью* F.

Доказательство. Мы будем проверять единственность — существование при этом получится в качестве «бесплатного приложения».

Пусть F_{ss} и F_n обладают нужными свойствами. Тогда V разлагается в прямую сумму собственных подпространств оператора F_{ss}: $V=\oplus V_{\lambda_j}$. Поскольку F_n перестановочен с F_{ss}, каждое из подпространств V_{λ_j} будет инвариантно для F_n.

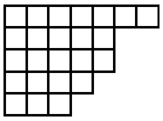
Ограничение F_n на V_{λ_j} совпадает с ограничением $F-\lambda_jE$. Из нильпотентности F_n следует, что V_{λ_j} содержится в корневом подпространстве K_{λ_j} оператора F. Из теоремы о корневом разложении для F вытекает теперь, что $V_{\lambda_j}=K_{\lambda_j}$. Поэтому F_{ss} однозначно определяется по F как оператор, действующий на каждом корневом подпространстве K_{λ_j} оператора F гомотетией с коэффициентом λ_j. Но тогда $F_n = F - F_{ss}$ также однозначно восстанавливается по F.

То, что и наоборот, операторы F_{ss} и F_n, построенные по только что описанным правилам, будут удовлетворять условиям теоремы — очевидно. □

§3. Строение нильпотентного оператора (над любым полем).

Действие нильпотентного опрератора $V \xrightarrow{F} V$ над произвольным полем k очень удобно описывать при помощи *диаграммы Юнга* (мы уже встречались с ними в первом семестре, когда занимались симметрическими многочленами).

3.1. Определение. *Базис пространства V называется жордановым базисом для оператора F, если его базисные векторы можно так расположить в клетках некоторой диаграммы Юнга*

$\sigma_F =$  ,

чтобы при действии F каждый из базисных векторов переходил в базисный вектор, стоящий в соседней слева клетке, а векторы крайнего левого столбца переходили в нуль.

Наглядно действие F на векторы жорданова базиса состоит в сдвиге на один шаг влево вдоль строк диаграммы Юнга. При этом группы векторов, стоящие в строках, называются *жордановыми цепочками*, а набор длин таких цепочек (т.е. длин строк диаграммы), составляющий *разбиение* числа $\dim V$, называется *цикловым типом* оператора F.

3.2. Свойства жорданова базиса. Если оператор F имеет жорданов базис с диаграммой σ_F ширины[2] $l(\sigma_F)$, то $F^{l(\sigma_F)} = 0$. Поэтому жорданов базис бывает только у нильпотентных операторов.

Далее, диаграмма Юнга σ_F однозначно определяется по F и не зависит от выбора жорданого базиса. В самом деле, длина k-того слева столбца диаграммы σ_F равна

$$\dim\Big(\ker F^k / \ker F^{k+1}\Big) .$$

Это оправдывает все введённые выше обозначения и названия.

Задача 6.12. Докажите, что число жордановых цепочек оператора F, состоящих не менее, чем из m векторов, равно $\dim\Big(\operatorname{im} F^{m-1} / \operatorname{im} F^m\Big)$.

Задача 6.13. Докажите, что два нильпотентных оператора F и G тогда и только тогда имеют одинаковый цикловой тип, когда существует изоморфизм $C : V \xrightarrow{\sim} V$, вписывающийся в коммутативную диаграмму

$$\begin{tikzcd} V \arrow[r, "\widetilde{C}"] & V \\ V \arrow[u, "F"] \arrow[r, "\widetilde{C}"] & V \arrow[u, "G"'] \end{tikzcd}$$

(другая формулировка: два нильпотентных оператора тогда и только тогда имеют одинаковый

[2]Напомню, что ширина диаграммы Юнга — это длина её верхней (самой длинной) строчки.

цикловой тип, когда их матрицы подобны — одна получается сопряжением другой при помощи невырожденной матрицы).

3.3. Теорема. *В конечномерном векторном пространстве V над произвольным полем для любого нильпотентного оператора существует жорданов базис.*

Доказательство. Индукция по размерности. Если $\dim V = 1$, или если $F = 0$, то любой базис будет жордановым с диаграммой Юнга, состоящей из одного столбца высоты $\dim V$.

Пусть $\dim V > 1$ и $\ker F \neq V$. Поскольку F нильпотентен, имеем также $\ker F \neq 0$, а значит, $0 < \dim(V/\ker F) < \dim V$, и в фактор-пространстве $V/\ker F$ существует (по индуктивному предположению) жорданов базис для фактор-оператора $\tilde{F}$.

В каждой жордановой цепочке этого базиса

$$[e_1] \leftarrow [e_2] \leftarrow [e_3] \leftarrow \cdots \leftarrow [e_m]$$

класс $[e_j]$ совпадает с классом $[F^{m-j}e_m]$ при всех $j = 1, 2, \ldots$. Действительно, равенство $\tilde{F}[e_\nu] = [e_{\nu-1}]$ в $V/\ker F$ означает, что в V $Fe_\nu = e_{\nu-1} - w_\nu$ с $w_\nu \in \ker F$. Последовательно применяя эти соотношения, получим

$$\begin{aligned}
e_{m-1} &= Fe_m + w_m \\
e_{m-2} &= Fe_{m-1} + w_{m-1} = F(Fe_m + w_m) + w_{m-1} = F^2 e_m + w_{m-1} \\
e_{m-3} &= Fe_{m-2} + w_{m-2} = F(F^2 e_m + w_{m-1}) + w_{m-2} = F^3 e_m + w_{m-2} \\
&\ldots \\
e_1 &= Fe_2 + w_2 = F(F^{m-2}e_m + w_3) + w_2 = F^{m-1}e_m + w_2,
\end{aligned}$$

что и даёт сравнения $e_j \equiv F^{m-j}e_m \pmod{\ker F}$.

Таким образом, любая жорданова цепочка $\tilde{F}$ является приведением по модулю $\ker F$ куска от жордановой цепочки оператора F, причём этот кусок допускает продолжение ещё ровно на один шаг влево — он может быть дополнен ненулевым вектором $e_0 = Fe_1 = F^m e_m \in \ker F$. Эти дополняющие векторы составляют линейно-независимую систему в $\ker F$, поскольку всякая линейная зависимость между ними означала бы точно такую же линейную зависимость по модулю $\ker F$ между последними классами $[e_1]$ жордановых цепочек для $\tilde{F}$ в $V/\ker F$.

Дополним систему векторов $[e_0]$ ещё несколькими векторами до базиса в $\ker F$. Объединение всех добавленных нами векторов с исходными $F^k e_m$ составит по зад.6.2 искомый жорданов базис для F в V. □

Отметим, что из этой теоремы и единственности циклового типа оператора получается такое

3.3.1. Следствие. *Пространство, на котором действует нильпотентный линейный оператор, тогда и только тогда не разлагается в прямую сумму F-инвариантных подпространств меньшей размерности, когда жорданов базис для F состоит из одной жордановой цепочки.*

Поэтому линейные оболочки жордановых цепочек аналогичны простым числам (в том смысле, что и те и те «далее неразложимы»), а теорема о существовании жорданова базиса и единственности циклового типа похожа на теорему о существовании и единственности разложения на простые множители. В конце лекции мы немного остановимся на причинах этой похожести.

3.4. Жорданова нормальная форма матрицы. Предположим теперь опять, что основное поле k алгебраически замкнуто. Тогда произвольный оператор $F: V \longrightarrow V$ можно канонически разложить в сумму $F = F_{ss} + F_n$ полупростой и нильпотентной составляющих, коммутирующих между собой. Соответственно, пространство V распадается в прямую сумму собственных подпространств V_λ оператора F_{ss}, и каждое такое подпространство переводится в себя оператором F_n. Выбирая в каждом V_λ жорданов базис для F_n, мы получаем для оператора F матрицу очень простого вида, называемого *жордановой нормальной формой матрицы оператора* F.

3.4.1. Следствие. *Для любого оператора $F: V \longrightarrow V$, действующего в пространстве над алгебраически замкнутым полем, существует базис, в котором его матрица блочно диагональна с блоками вида*

$$J_m(\lambda) \stackrel{\text{def}}{=} \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & 0 \\ & & \ddots & \ddots & \\ 0 & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \qquad (\text{размер } m \times m).$$

При этом вся совокупность блоков однозначно с точностью до перестановки определяется по оператору и не зависит от выбора базиса, числа λ — это в точности собственные значения, а размеры блоков не превышают кратности вхождения соответствующих $(t - \lambda)$ в разложение минимального многочлена.

Блоки $J_m(\lambda)$ называются *жордановыми клетками*. Количество и размеры жордановых клеток, имеющих одно и то же λ на диагонали, целиком определяется цикловым типом ограничения $F - \lambda E$ на корневое подпространство K_λ.

Задача 6.14. Убедитесь, что наибольший из размеров клеток $J_m(\lambda)$ равен кратности вхождения $(t - \lambda)$ в разложение минимального многочлена, а сумма всех размеров равна кратности вхождения $(t - \lambda)$ в характеристический многочлен.

§4. Функции от оператора.

Среди приложений теории линейных операторов важное место занимают задачи об определении и вычислении различных функций от операторов (а на более приземлённом языке — функций от матриц). Мы обсудим здесь коротко ряд примеров.

4.1. Интерполяционный многочлен. Операторы составляют векторное пространство и кольцо, поэтому при определении функции от оператора мы можем использовать алгебраические действия сложения и умножения, а при наличии нормы на пространстве операторов, — ещё и предельные переходы. Мы будем рассматривать только такие функции из $\mathrm{Hom}(V, V)$ в $\mathrm{Hom}(V, V)$, которые определяются при помощи этих операций.

Пусть $f: \mathrm{Hom}(V, V) \longrightarrow \mathrm{Hom}(V, V)$ — такая функция. Тогда значение $f(F)$ для конкретного $F \in \mathrm{Hom}(V, V)$ будет лежать в линейной оболочке неотрицательных целых степеней F^ν, т.е. в $k[F]$. Это означает, что $f(F)$ всегда совпадает с результатом подстановки F в некоторый многочлен. Всякий такой многочлен назывется *интерполяционным многочленом для вычисления* $f(F)$.

Поскольку $k[F] \simeq k[t]/(\mu_f)$ интерполяционный многочлен всегда можно подобрать таким, чтобы его степень не превосходила степени минимального многочлена оператора F (и этим условием он будет определяться однозначно!). Однако указать степень минимального многочлена не всегда просто, и при определении интерполяционного многочлена удобнее исходить из того, что его степень *строго меньше размерности пространства*. Далее применяется метод неопределённых коэффициентов, а ключом к их вычислению служит тождество Гамильтона-Кэли. Перейдём к примерам.

4.2. Степени оператора. В качестве интерполяционного многочлена для вычисления значения $f(F)$, где $f \in k[t]$ — это многочлен (например, $f(t) = t^N$ — степень), можно взять остаток $r(t)$ от деления $f(t)$ на характеристический многочлен $\chi_F(t)$. Действительно, подставляя F в многочленное равенство

$$f(t) = \chi_F(t) \cdot q(t) + r(t)\,, \quad \deg r \le n-1\,,\ \deg \chi_F = n\,,$$

мы в силу тождества Гамильтона-Кэли получим $f(F) = r(F)$.

Более того, подставляя в это же равенство и в равенства, полученные из него дифференцированиями, собственные значения λ_j оператора F, мы получим для определения коэффициентов $r(t)$ n линейных уравнений:

$$r^{(\nu)}(\lambda_j) = f^{(\nu)}(\lambda_j)$$

(в которых ν меняется от нуля до кратности корня λ_j в χ_F). Таким образом, r можно искать как решение задачи интерполяции с кратными узлами, которую мы обсуждали в лекции 1.

4.2.1. Пример. Вычислим $\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}^k$. Характеристический многочлен

$$\chi(t) = t^2 - t \cdot \vartriangle + \det = t^2 - 4t + 4 = (t-2)^2.$$

Тем самым, интерполяционный многочлен для вычисления t^k имеет вид $r(t) = at + b$, причём

$$\begin{aligned} r(2) &= 2a + b = 2^k \\ r'(2) &= a = k2^{k-1}. \end{aligned}$$

Находим $r(t) = k2^{k-1}t + (1-k)2^k = 2^{k-1}(kt + (2-2k))$, откуда

$$\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}^k = 2^{k-1}\left(k\,\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} + (2-2k)\,\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = 2^{k-1}\begin{pmatrix} 2-k & k \\ -k & 2+k \end{pmatrix}.$$

4.3. Алгебраические функции. Под алгебраической функцией оператора F мы будем понимать оператор G, заданный неявно, как решение уравнения

$$F_0 + F_1G + F_2G^2 + \cdots + F_kG^k = 0,$$

коэффициенты которого $F_\nu = f_\nu(F)$ являются многочленами от оператора F. Примером алгебраической функции может служить корень из оператора $G = \sqrt[k]{F}$, являющийся решением уравнения $G^k - F = 0$. Для определения интерполяционного многочлена, вычисляющего алгебраическую функцию от оператора, достаточно решить по модулю характеристического многочлена $\chi_F(t)$ уравнение

$$f_0(t) + f_1(t)x + f_2(t)x^2 + \cdots + f_k(t)x^k = 0$$

над кольцом многочленов $k[t]$, т.е. подобрать такой многочлен $x = x(t)$, чтобы

$$f_0(t) + f_1(t)x(t) + f_2(t)x(t)^2 + \cdots + f_k(t)x(t)^k$$

делился на $\chi(t)$. Тогда $G = x(F)$ будет искомым оператором.

Мы ограничимся лишь одним примером.

4.3.1. Утверждение. *Если оператор $F:\ V \longrightarrow V$, действующий в пространстве над алгебраически замкнутым полем, не имеет ядра, то $\forall k \in \mathbb{N}$ существует оператор $G:\ V \longrightarrow V$, который является многочленом от F и удовлетворяет равенству $G^k = F$.*

Доказательство. Пусть $\chi_F(t) = \prod (t-\lambda_j)^{m_j}$. Мы должны показать, что если все $\lambda_j \neq 0$, то сравнение

$$x(t)^k \equiv t\,(\operatorname{mod} \chi_F(t))$$

разрешимо в $k[t]$ относительно x. По китайской теореме об остатках достаточно доказать разрешимость этого сравнения по модулю каждого из взаимно-простых множителей $(t-\lambda_j)^{m_j}$ из разложения χ_F. Следовательно, без ограничения общности можно считать, что $\chi_F(t) = (t-\lambda)^m$.

Сделаем замену $t = s + \lambda$, которая сводит задачу к отысканию такого многочлена $x(s)$, что $x(s)^k \equiv s + \lambda\,(\operatorname{mod} s^m)$.

Задача 6.15. Убедитесь, что этим свойством одладает отрезок из k начальных членов формального степенного рада Ньютона

$$(s+\lambda)^{1/m} = \lambda^{1/n}(1+s/\lambda)^{1/m} = \lambda^{1/n}\sum_{\nu\geq 0}\binom{1/m}{\nu}\left(\frac{s}{\lambda}\right)^{\nu}$$

Указание: можно сказать, что ряд Ньютона исходя из этого свойства и выписывался. Более конкретная идея: многочлен делится на s^m тогда и только тогда, когда он обращается в нуль при $s = 0$ вместе со своими производными до $(m-1)$-го порядка включительно.

□

В более сложных случаях вопрос о существовании той или иной алгебраической функции от оператора может быть исследован аналогичным образом при помощи описанной в конце лекции 7 прошлого семестра техники разложения произвольной алгебраической функции в формальный дробно-степенной ряд.

4.3.2. Пример. Вычислим квадратный корень из матрицы $\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$. Характеристический многочлен

$$\chi(t) = t^2 - 4t - 5 = (t+1)(t-5).$$

Интерполяционный многочлен $r(t) = at + b$ определяется тем, что $r(t)^2 - t$ делится на $\chi(t)$, а стало быть, имеет корни -1 и 5. Поэтому

$$\begin{aligned} (-a+b)^2 = -1 &\implies b - a = \pm i \\ (5a+b)^2 = 5 &\implies 5a + b = \pm\sqrt{5} \end{aligned}$$

Заметим что все четыре комбинации знаков приводят к (различным) матрицам, квадрат которых равен заданной в условии, так что алгебраические фукции являются *многозначными*. Выбирая, к примеру, знаки плюс, получим $a = (\sqrt{5} - i)/6$, $b = (\sqrt{5} + 5i)/2$, откуда

$$\begin{aligned}\sqrt{\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}} &= r\left(\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}\right) = \frac{\sqrt{5}-i}{6}\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{5}+5i}{2} & 0 \\ 0 & \frac{\sqrt{5}+5i}{2} \end{pmatrix} = \\ &= \begin{pmatrix} \frac{6\sqrt{5}+13i}{6} & \frac{\sqrt{5}-i}{2} \\ \frac{\sqrt{5}-i}{2} & \frac{6\sqrt{5}+13i}{6} \end{pmatrix}\end{aligned}$$

4.4. Аналитические функции. Пусть теперь V — векторное пространство над полем $\mathbb{R}$ или $\mathbb{C}$, и на V задана норма $||\cdot|| : V \longrightarrow \mathbb{R}$. Тогда по этой норме можно построить норму на пространстве операторов $\mathrm{Hom}(V, V)$, полагая:

$$||F|| \stackrel{\mathrm{def}}{=} \sup_{||v||=1} (||Fv||) = \sup_{v \neq 0} \left(\frac{||Fv||}{||v||}\right).$$

Задача 6.16. Проверьте выполнение свойств нормы.

Интуитивно-геометрически, эта норма есть максимальное растяжение длины, осуществляемое оператором F. Среди всех прочих норм на пространстве $\mathrm{Hom}(V, V)$ (задающих, как мы знаем, одну и ту же сходимость) эта норма выделяется тем, что она хорошо взаймодействует с композицией операторов. А именно, прямо из определения следует неравенство

$$||FG|| \leq ||F||\,||G||.$$

Это неравенство позволяет оценивать норму суммы куска степенного ряда от оператора суммой соответствующего куска числового степенного ряда:

$$\left|\left|\sum_{\nu=n}^{N} a_\nu F^\nu\right|\right| \leq \sum_{\nu=n}^{N} |a_\nu|\,||F||^\nu.$$

При помощи критерия Коши, дословно тем же рассуждением, что и в курсе анализа, решается

Задача 6.17. Пусть степенной ряд $\sum\limits_{\nu\geq 0} a_\nu A^\nu$ с вещественными (или комплексными) коэффициентами абсолютно сходится при некотором вещественном положительном A. Докажите, что ряд из операторов $\sum\limits_{\nu\geq 0} a_\nu F^\nu$ абсолютно и равномерно сходится в любом замкнутом шаре $||F|| \leq R$ радиуса $R < A$.

Таким образом, всякая аналитическая вещественная (или комплексная) функция $\psi(x)$ определяет некоторую функцию на пространстве операторов со значениями в пространстве операторов.

Например, абсолютно сходящийся всюду ряд

$$e^F = \sum_{\nu\geq 0} \frac{1}{\nu!} F^\nu$$

определяет *экспоненту* от оператора.

Важно иметь в виду, что многие тождества, которые имеют место в кольце формальных степенных рядов, перестают выполняться для рядов от операторов, поскольку кольцо операторв *некоммутативно*. К примеру, как правило, $e^{F+G} \neq e^F e^G$, если F и G не коммутируют (придумать два некоммутирующих оператора, для которых бы выполнялось равенство — это довольно известная «завльная » задача из студенческого олимпиадного фольклора).

Разумеется, к коммутирующим операторам весь арсенал формальных тождеств применим вполне, и единственным ограничением здесь является проверка сходимости рядов (т.е. само существование рассматриваемых функций).

4.4.1. Пример. Всякая унитарная матрица является экспонентой от эрмитово кососимметричной матрицы. Действительно, такая матрица является матрицей эрмитовой изометрии, а значит диагонализуема, и по её диагонали будут стоять числа с модулем 1, т.е. числа вида $e^{i\varphi}$. Тем самым этот оператор является экспонентой от оператора, имеющего в том же базисе диагональный вид с чисто мнимыми числами на диагонали, что и требовалось. Ясно, что верно и обратное: экспоненты косоэрмитовых матриц будут унитарными.

Задача 6.18. Убедитесь, что экспонента от вещественной кососимметричной матрицы будет ортогональной матрицей, и покажите, что не всякая ортогональная матрица является экспонентой от кососимметричной.

Задача 6.19. Докажите, что отображение $[0,1] \longrightarrow \mathrm{Mat}_n(\mathbb{R}): \ t \mapsto e^{tA}$, где A — фиксированная вещественная матрица, задаёт бесконечно-гладкую кривую, лежащую в $GL_n(\mathbb{R})$ и соединяющую E с e^A.

Поскольку значение аналитической функции ψ от оператора F лежит в $k[F]$, для для её практического вычисления опять-таки удобнее всего воспользоваться интерполяционным многочленом.

Задача 6.20. Докажите, что интерполяционный многочлен $r(t)$ для вычисления $\psi(F)$ можно выбрать как (единственное на пространстве многочленов степени $\leq \dim V - 1$) решение интерполяционной задачи с узлами в корнях характеристического многочлена, взятых с надлежащими кратностями:

$$r^{(\nu)}(\lambda_j) = \psi^{(\nu)}(\lambda_j).$$

В заключение этой темы ещё одна задача на технику интерполирования

Задача 6.21. Докажите, что над алгебраически замкнутым полем полупростая и нильпотентная жордановы составляющие произвольного оператора могут быть вычислены в виде многочленов без свободного члена от этого оператора.

Из этой задачи вытекает, в частности, что полупростая и нильпотентная составляющие оператора коммутируют с любым оператором, с которым коммутирует исходный оператор.

§5. Пространство с оператором как $k[t]$-модуль.

Пусть $F: \ V \longrightarrow V$ — линейный оператор в векторном пространстве над произвольным полем k.

5.1. Разложение V как модуля над $k[t]$. Зададим на V структуру модуля над кольцом многочленов $k[t]$, определив умножение векторов на многочлены формулой

$$g(t)\cdot v \stackrel{\text{def}}{=} g(F)v.$$

Иными словами, умножение векторов на константы будет происходить по тем же законам что и раньше, но ещё добавится дополнительная операция умножения векторов на t, состоящая, по определению, в применении оператора F.

Задача 6.22. Убедитесь, что все аксиомы модуля при этом будут выполнены.

Фиксируем в V базис $e_1, e_2, \dots, e_n$ над k. Векторы e_ν будут порождать V и как модуль над $k[t]$, но над $k[t]$ они уже будут линейно зависимы. Для описания соотношений между ними рассмотрим свободный $k[t]$-модуль M ранга n, т.е. прямую сумму

$$M = \underbrace{k[t]\oplus k[t]\oplus\cdots\oplus k[t]}_{n}$$

(напомню, что этот модуль состоит из столбцов высоты n из многочленов, а операции над ними определяются покоординатно). Зададим гомоморфизм модулей

$$\pi_F:\ M \longrightarrow V$$

$$\pi_F:\ \begin{pmatrix} g_1\\ g_2\\ \vdots\\ g_n\end{pmatrix} \longrightarrow g_1(F)e_1 + g_2(F)e_2 + \cdots + g_n(F)e_n$$

Задача 6.23. Убедитесь, что это действительно гомоморфизм $k[t]$-модулей.

Гомоморфизм π_F, очевидно, сюрьективен, и из теоремы о строении гомоморфизма модулей мы получаем, что V как $k[T]$-модуль изоморфен $M/\ker\pi_F$.

В прошлом семестре мы видели, что в свободном модуле M любой подмодуль N обладает базисом. Более того, существует такой базис $\{\zeta_1, \zeta_2, \dots, \zeta_n\}$ в M, что возрастающие кратности

$$\{d_1(t)\zeta_1,\ d_2(t)\zeta_2,\ \dots,\ d_r(t)\zeta_r\}\,, \quad (\text{где } d_i \text{ делит } d_j \text{ при } i<j)$$

первых нескольких его базисных векторов составляют базис в подмодуле N, причём набор кратностей $\{d_1, d_2, \dots, d_r\}$ является инвариантом подмодуля и от выбора базиса не зависит.

Применяя эту теорему к подмодулю $\ker\pi_F \subset M$, мы получаем

5.1.1. Утверждение. *Существует такой набор многочленов $\{d_1, d_2, \dots, d_r\}\subset k[t]$, что d_i делят d_j при $i<j$ и имеется изоморфизм $k[t]$-модулей*

$$\frac{k[t]}{(d_1(t))}\oplus\frac{k[t]}{(d_2(t))}\oplus\cdots\oplus\frac{k[t]}{(d_r(t))}\xrightarrow{\sim}V \tag{5-1}$$

Этими условиями набор $\{d_1, d_2, \dots, d_r\}\subset k[t]$ однозначно определяется по оператору F.

Слова «изоморфизм $k[t]$-модулей» означают, что мы можем отождествить векторное пространство V с векторным пространством

$$\frac{k[t]}{(d_1(t))}\oplus\frac{k[t]}{(d_2(t))}\oplus\cdots\oplus\frac{k[t]}{(d_r(t))}\ ,$$

так что оператор F будет действовать как оператор *покоординатного умножения на t.*

Многочлены $d_j(t)$ называются *инвариантными множителями* оператора F.

5.1.2. Следствие. *$d_r(t) = \text{НОК}(d_1, d_2, \ldots, d_r)$ совпадает с минимальным многочленом $\mu_F(t)$ оператора F.*

5.2. Элементарные делители. Если разложить каждый из d_j в произведение неприводимых многочленов

$$d_j(t) = \prod_k p_{jk}(t)^{m_{jk}},$$

то полная совокупность всех участвующих в этих разложениях степеней $p_{jk}^{m_{jk}}$ также будет однозначно определяться оператором F. Эти степени $p_{jk}^{m_{jk}}$ называются *элементарными делителями* оператора F.

По китайской теореме об остатках каждое слагаемое разложения (5-1) допускает дальнейшее разложение в виде

$$\frac{k[t]}{(d_j(t))} = \bigoplus_k \frac{k[t]}{(p_{jk}(t)^{m_{jk}})}$$

откуда получается

5.2.1. Следствие. *Конечномерное векторное пространство V над произвольным полем k, на котором действует линейный оператор $F: V \longrightarrow V$, можно отождествить с прямой суммой фактор-колец*

$$\bigoplus_\nu k[t]/(p_\nu(t)^{m_\nu}),$$

в которой все p_ν неприводимы над k. При этом оператор F перейдёт в оператор умножения на t.

Над алгебраически замкнутым полем это даёт ещё одно доказательство теоремы о приведении матрицы к жордановой нормальной форме:

> Задача 6.24. В пространстве $k[t]/((t-\lambda)^m)$ постройте базис, в котором оператор умножения на t имеет матрицу $J_m(\lambda)$.

5.2.2. Следствие. *Жордановы клетки оператора F, отвечающие собственному значению λ, находятся во взаимно-однозначном соответствии с элементарными делителями оператора F вида $(t-\lambda)^m$, а их размеры равны соответствующим показателям m.*

> Задача 6.25. (необязательная) Как выглядит «вещественная жорданова нормальная форма» матрицы над полем $\mathbb{R}$?

Подчекнём, что набор элементарных делителей является полным набором инвариантов линейного оператора относительно сопряжения: два оператора F и G могут быть в некоторых базисах записаны одинаковыми матрицами[3] тогда и только тогда, когда у них один и тот же набор элементарных делителей. Это вытекает из задачи

[3]Т.е. существует изоморфизм $C: V \xrightarrow{\sim} V$, вписывающийся в коммутативную диаграмму

$$\left(\begin{array}{ccc} V & \overset{C}{\xrightarrow{\sim}} & V \\ {\scriptstyle F}\uparrow & & \uparrow{\scriptstyle G} \\ V & \overset{C}{\xrightarrow{\sim}} & V \end{array}\right),$$

Задача 6.26. Докажите, что два $k[t]$-модуля вида

$$\bigoplus_{\nu} k[t]/(p_\nu(t)^{m_\nu})$$

тогда и только тогда изоморфны, когда они отличаются лишь перестановкой слагаемых.

5.3. Отыскание инвариантных множителей по матрице оператора. Чтобы описать инвариантные множители оператора F в терминах его матрицы F_e в базисе $\{e_1, e_2, \dots, e_n\}$ мы построим систему образующих в $\ker \pi_F$, связанную с этой матрицей.

5.3.1. Лемма. *Ядро $\ker \pi_F$ как $k[t]$-подмодуль в M порождается столбцами матрицы $tE - F_e$, где F_e — это матрица оператора F в некотором базисе $e_1, e_2, \dots, e_n$ пространства V (столбцы матрицы $tE - F_e$ рассматриваются как столбцы из n многочленов, из которых один линеен, а остальные $n-1$ — константы).*

Доказательство. Пусть $F_e = (a_k^j)$, так что

$$Fe_j = \sum_{\nu} a_j^\nu e_\nu .$$

При гомоморфизме π_f j-тый столбец матрицы $tE - F_e$ перейдёт в

$$Fe_j - \sum_{\nu} a_j^\nu e_\nu = 0,$$

так что все столбцы действительно лежат в ядре. Докажем, что они его порождают.

Пусть столбец $(g_\nu) = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix} \in \ker \pi_F$. Разделим каждый g_ν с остатком на $(t - a_\nu^\nu)$:

$$g_\nu(t) = (t - a_\nu^\nu) q_\nu(t) + c_\nu \,, \quad \text{где} \quad c_\nu \in k \,.$$

Вычтем теперь из столбца (g_ν) линейную комбинацию столбцов матрицы $(tE - F_e)$ с коэффициентами q_ν. При этом мы получим столбец, который по-прежнему лежит в ядре π_F и состоит из многочленов

$$\tilde g_n u = a_1^\nu q_1(t) + \dots + a_{(\nu-1)}^\nu q_{\nu-1}(t) + c_\nu + a_{(\nu+1)}^\nu q_{\nu+1}(t) + \dots + a_n^\nu q_n(t) \,.$$

Заметим, что если исходный столбец (g_ν) не был столбцом констант, то максимальная степень многочленов нового столбца

$$\max_\nu(\tilde g_\nu) \leqslant \max_\nu(q_\nu) < \max_\nu(g_\nu)$$

строго уменьшится. Следовательно, итерируя эту процедуру, мы сможем вычесть из исходного столбца (g_ν) такую линейную комбинацию столбцов матрицы $(tE - F)$ с коэффициентами из $k[t]$, что разность будет столбцом констант. Но в ядре π_F имеется, очевидно, только один столбец констант — нулевой. Тем самым, исходный столбец (g_ν) совпадает с построенной линейной комбинацией столбцов матрицы $(tE - F)$, что и требовалось усмотреть. $\square$

Из леммы следует, что для отыскания d_j мы должны элементарными преобразованиями строк и столбцов над кольцом многочленов $k[t]$ привести матрицу $(tE - F_e)$ к диагональному виду

$$\begin{pmatrix} d_1(t) & & 0 \\ & \ddots & \\ 0 & & d_n(t) \end{pmatrix} \tag{5-2}$$

в котором диагональные элементы d_i делят все d_j с $j > i$. Тогда эти диагональные элементы и будут искомыми инвариантными множителями. В частности в правом нижнем углу будет стоять минимальный многочлен оператора F.

В некоторых случаях полезно иметь более явное описание многочленов $d_\nu(t)$ в терминах матрицы $tE - F_e$.

5.3.2. Лемма. *Многочлен $d_j(t)$ равен отношению наибольшего общего делителя всех миноров j-того порядка у матрицы $tE - F_e$ к наибольшему общему делителю всех её миноров $(j-1)$-го порядка.*

Доказательство. У результирующей диагональной матрицы наибольший общий делитель миноров j-того порядка равен произведению первых j диагональных элементов (все последующие на них делятся!). Поэтому для доказательства леммы надо проверить, что наибольший общий делитель всех миноров данного порядка не меняется при элементарных преобразованиях строк и столбцов.

Поскольку обратным к элементарному снова будет элементарное преобразование, инвариантность наибольшего общего делителя вытекает из задачи

> Задача 6.27. Пусть матрица A получена из матрицы B одним элементарным преобразованием строк (или столбцов), и пусть все миноры j-того порядка у B делятся на f. Докажите, что тогда и у A все миноры j-того порядка делятся на f.
>
> Указание: В единственном нетривиальном случае, когда к строке прибавляется строка, разложите минор A, содержащий сумму двух строк, в сумму двух миноров, пришедших из B.

Лемма доказана. $\square$