Алгебра 1 курс. 2 модуль. Записки лекций.

И.В.Артамкин

Декабрь 2012

Contents

1 Ядро и образ. **2**

2 Понятие инвариантного подпространства. **3**

3 Собственные векторы и собственные значения. **6**

4 Характеристический многочлен. **8**

5 Минимальный многочлен. **9**

6 Бывает все на свете хорошо: если есть базис, составленный из собственных векторов. **10**

7 Как строить инвариантные подпространства. Матрица Фробениуса. **14**

8 Теорема Гамильтона-Кэли **16**

9 Неприводимые делители минимального многочлена. **18**

10 Корневое разложение. **22**

11 Нильпотентные операторы. **25**

12 Итог: когда характеристический многочлен разлагается на линейные множители. **31**

13 Фактор-пространство **32**

В этих записках мы рассматриваем линейный оператор $f : V \to V$, где V — n-мерное линейное пространство над некоторым полем $\mathbb{K}$.

1 Ядро и образ.

Теорема 1.1. $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$.

Для доказательства выберем базис $e_1, \ldots, e_k$ подпространства $\operatorname{Ker} f$ и дополним его векторами $e_{k+1}, \ldots, e_n$ до базиса $e_1, \ldots, e_n$ всего пространства V. Заметим, что поскольку $f(e_1) = \ldots = f(e_k) = 0$, векторы $f(e_{k+1}), \ldots, f(e_n)$ порождают образ $\operatorname{Im} f$. Тем самым нам достаточно доказать, что они составляют базис в пространстве $\operatorname{Im} f$, а для этого достаточно доказать, что они линейно независимы. Докажем это от противного: пусть имеется соотношение

$$\lambda_{k+1} f(e_{k+1}) + \ldots + \lambda_n f(e_n) = 0,$$

в котором не все коэффициенты равны нулю. Следовательно,

$$f(\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n) = 0,$$

так что

$$\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n \in \operatorname{Ker} f.$$

Но всякий вектор из $\operatorname{Ker} f$ разлагается по его базису $e_1, \ldots, e_k$, так что

$$\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n = \mu_1 e_1 + \ldots + \mu_k e_k,$$

что дает нам линейное между векторами базиса пространства V, в котором, как мы предполагали, не все λ_i равны нулю. Противоречие. □

Отметим, что только что доказанное соотношение интересно, в частности, по следующей причине. Если наше пространство V как-то разложено в прямую сумму двух подпространств $V = U \oplus W$, то размерности прямых слагаемых U и W удовлетворяют точно такому же соотношению: $\dim U + \dim W = \dim V$. Это наблюдение может дать основание надеяться, что пространство V разлагается в прямую сумму ядра и образа любого линейного оператора. Но это, конечно, не так: в случае прямой суммы подпространства U и V пересекаются только по нулю, в то время как ядро и образ оператора вполне могут иметь ненулевое пересечение.

Задача 1.1. *Докажите, что у линейного оператора в n-мерном пространстве, заданного матрицей $a_{i,j} = \delta_{i+1,j}$ имеется ненулевое ядро (найдите его размерность!), которое целиком содержится в образе (найдите его размерность!).*

Задача 1.2. *Пусть в пространстве V имеются два произвольные подпространства U и V, размерности которых удовлетворяют соотношению $\dim U + \dim W = \dim V$. Докажите, что существует такой линейный оператор $f : V \to V$, что $U = \operatorname{Ker} f$ и $W = \operatorname{Im} f$.*

2 Понятие инвариантного подпространства.

Важное средство для изучения линейных операторов — понятие инвариантного подпространства.

Определение 2.1. *Пусть $f : V \to V$ линейный оператор. Линейное подпространство $U \subset V$ называется* **инвариантным подпространством** *(для оператора f), если $f(U) \subset U$.*

Отметим, что совсем не требуется, чтобы образ $f(U)$ пространства U в точности совпадал с U, требуется лишь, чтобы он был подмножеством U.

Задача 2.1. *Докажите, что $\operatorname{Ker} f$ и $\operatorname{Im} f$ являются инвариантными подпространствами для оператора f.*

Задача 2.2. *Предположим, что операторы f и g перестановочны, т.е. $fg = gf$. (Это, конечно, означает, что $\forall\ v \in V\ \ f(g(v)) = g(f(v))$.) Докажите, что $\operatorname{Ker} g$ и $\operatorname{Im} g$ являются инвариантными подпространствами для оператора f.*

Задача 2.3. *Докажите, что $\operatorname{Ker} f^m$ и $\operatorname{Im} f^m$ являются инвариантными подпространствами для оператора f для любого натурального m.*

Задача 2.4. *а) Перечислите все инвариантные подпространства, получаемые по рецепту предыдущей задачи для оператора из задачи 1.1. б) Докажите, у этого оператора нет других нетривиальных инвариантных подпространств.*

Ответ, получившийся в предыдущей задаче, является типичным: у линейных операторов часто бывает совсем немного инвариантных подпространств. Однако, это, все-таки, далеко не всегда бывает именно так: например, у оператора гомотетии с коэффициентом $\lambda \in \mathbb{K}$, $f(v) = \lambda v$, любое подпространство является инвариантным.

Наличие инвариантного подпространства — это геометрическое свойство линейного оператора. Однако эту информацию можно переработать и в алгебраическую, или, точнее, координатную, — в информацию о матрице этого оператора. Для этого выберем в инвариантном подпространстве U какой-нибудь базис $e_1, \ldots, e_k$ и дополним его до базиса всего пространства V векторами $e_{k+1}, \ldots, e_n$. Матрица оператора f в этом базисе составляется из столбцов координат образов базисных векторов. Но в силу инвариантности U образы первых k базисных векторов лежат в U и потому разлагаются только по векторам $e_1, \ldots, e_k$, а все остальные коэффициенты равны нулю. Это значит, что матрица оператора f в этом базисе будет блочной верхнетреугольной:

$$\hat{f} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \tag{2.1}$$

где A и B —квадратные матрицы размером, соответственно, $k \times k$ и $(n-k) \times (n-k)$. При этом нетрудно видеть, что матрица A есть в точности матрица ограничения оператора f на подпространство U, записанная в базисе $e_1, \ldots, e_k$. Очевидно, верно и обратное: если матрица оператора f в базисе $e_1, \ldots, e_n$ оказалась блочной верхнетреугольной вида (2.1), то подпространство $U = \langle e_1, \ldots, e_k \rangle$ является инвариантным подпространством, а матрица A есть в точности матрица ограничения оператора f на подпространство U, записанная в базисе $e_1, \ldots, e_k$. Подведем итог.

Предложение 2.1. *Инвариантность подпространства U относительно оператора f равносильна тому, что матрица оператора f в базисе, первые k векторов которого являются базисом U имеет блочный верхнетреугольный вид (2.1).* $\square$

Задача 2.5. *а) Докажите, что если $P(t) \in \mathbb{K}[t]$, то для блочной верхнетреугольной матрицы (2.1)*

$$P(\hat{f}) = \begin{pmatrix} P(A) & * \\ 0 & P(B) \end{pmatrix}. \tag{2.2}$$

б) Вычислите явно правый верхний угол X в n-ой степени матрицы (2.1):

$$\hat{f}^n = \begin{pmatrix} A^n & X \\ 0 & B^n \end{pmatrix}. \tag{2.3}$$

Отметим одно странное обстоятельство: когда мы благодаря наличию инвариантного подпространства U получили блочно-верхнетреугольную матрицу оператора f, мы сразу легко проинтерпретировали матрицу A в (2.1): она оказалась матрицей ограничения оператора f на подпространство U. Естественно ожидать, что подобным образом можно проинтерпретировать и матрицу B. Однако нетрудно понять, что нельзя ожидать, что матрица тоже B окажется матрицей ограничения оператора f на какое-нибудь другое инвариантное подпространство, например, по той причине, что бывают операторы, у которых всего одно инвариантное подпространство. Простейший пример такого рода — оператор, заданный матрицей $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Легко видеть (проверьте!), что здесь имеется всего одна инвариантная прямая, состоящая из векторов с нулевой второй координатой.

Действительно, для объяснения геометрического смысла матрицу B в (2.1) необходимо привлечь другую геометрическую конструкцию (фактор-пространства), которая здесь еще не обсуждалась и появится у нас позже.

Но если бы все-таки мы стали искать какое-нибудь подпространство, которое "отвечало бы" за матрицу B, то самым естественным кандидатом была бы линейная оболочка векторов $e_{k+1}, \dots, e_n$, которыми мы дополнили базис подпространстве U до базиса всего V. Неприятность, однако, состоит в том, что подпространство $W = \langle e_{k+1}, \dots, e_n \rangle$ не инвариантно относительно оператора f, по крайней мере в том случае, когда матрица C ненулевая. Действительно, если какое-то $c_{p,q} \neq 0$ ($p \leq i$, $q > i$), то это значит, что в разложении по базису $e_1, \dots, e_n$ пространства V вектора $e_q \in W$ коэффициент при векторе $e_p \notin W$, равный как раз $c_{p,q}$, отличен от нуля, поэтому $f(e_q) \notin W$. Зато если матрица $C = 0$, то очевидно, W инвариантно относительно оператора f, и в этом случае, конечно, матрица B есть в точности матрица ограничения оператора f на подпространство W. Заметим еще, что пространство V есть прямая сумма подпространств U и W, т.е. $V = U \oplus W$.

Легко видеть, что и наоборот, если пространство V удалось представить в виде прямой суммы двух инваринатных подпространств, то в базисе пространства V, составленном из базисов этих подпространств, матрица оператора f будет блочной диагональной матрицей

$$\hat{f} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{2.4}$$

Подведем итог.

Предложение 2.2. *Разложение пространства V в прямую сумму инвариантных подпространств равносильна тому, что в базисе пространства V, составленном из базисов этих подпространств, матрица оператора f имеет блочный диагональный вид (2.4).*$\square$

3 Собственные векторы и собственные значения.

Самый важный и часто встречающийся пример инвариантного подпространства — это одномерное инвариантное подпространство. В этом случае подпространство U порождено одним вектором $u \in V$: $U = \langle u \rangle$. Тогда условие инвариантности подпространства U означает, что образ $f(u)$ вектора u снова лежит в U, и, значит, пропорционален вектору u с каким-нибудь коэффициентом пропорциональности $\lambda \in \mathbb{K}$, т.е. $f(u) = \lambda u$.

Определение 3.1. *Вектор $u \in V$ называется* **собственным вектором** *оператора f, если $f(u) = \lambda u$ для некоторого $\lambda \in \mathbb{K}$. Число λ при этом называется* **собственным значением** *оператора f.*

Пусть $u \in V$ ненулевой собственный вектор оператора f с собственным значением λ, т.е. $f(u) = \lambda u$. Это равносильно тому, что вектор u лежит в ядре оператора $f - \lambda \operatorname{Id}_V$. Но оператор $f - \lambda \operatorname{Id}_V$ имеет ненулевое ядро тогда и только тогда, когда

$$\det(f - \lambda \operatorname{Id}_V) = 0. \tag{3.1}$$

Записав оператор матрицей в каком-нибудь базисе, мы видим, что (3.1) представляет собой алгебраическое уравнение степени n на λ; оно называется **характеристическим уравнением** для оператора f. Легко видеть, что, поскольку все переходы в наших предыдущих рассуждениях были равносильными, верно и обратное: если $\lambda \in \mathbb{K}$ является решением характеристического уравнения, то существует ненулевой вектор

$u \in \operatorname{Ker}(f - \lambda \operatorname{Id}_V)$, который, тем самым, является собственным вектором оператора f с собственным значением λ. Тем самым мы доказали очень важную теорему.

Теорема 3.1. *Ненулевой собственный вектор с собственным значением λ существует тогда и только тогда, когда λ является корнем характеристического уравнения (3.1).*$\square$

Следствие 3.1. *У линейного оператора в n-мерном пространстве может быть не более n различных собственных значений.*$\square$

Для каждого корня характеристического уравнения λ можно рассмотреть множество V_λ всех собственных векторов с собственным значением λ.

Задача 3.1. *Докажите, что V_λ является линейным подпространством в V.*

V_λ называется собственным подпространством, отвечающим собственному значению λ.

Задача 3.2. *Пусть $V = \mathbb{R}^2$ евклидова плоскость. Докажите, что поворот на ненулевой угол $\varphi \neq 180^\circ$ не имеет собственных подпространств, зеркальная симметрия относительно прямой l имеет два одномерных собственных подпространства, отвечающие собственным значениям -1 и 1, ортогональное проектирование на прямую l имеет два одномерных собственных подпространства, отвечающие собственным значениям 0 и 1, а оператор, являющийся композицией ортогонального проектирования на прямую l и поворота на 90° имеет ровно одно одномерное собственное подпространство, отвечающие собственному значению 0. (Выпишите матрицы этих четырех операторов и напишите и решите их характеристические уравнения!)*

Отметим, что нас пока что интересовали только корни характеристического уравнения, хотя на самом деле, как мы увидим чуть позже, само уравнение также содержит очень важную информацию об операторе.

4 Характеристический многочлен.

Определение 4.1. *Многочлен*

$$\chi_f(t) = \det(t\,\mathrm{Id}_V - f) \in \mathbb{K}[t] \tag{4.1}$$

называется **характеристическим многочленом** *оператора* f.

Заметим, что, хотя определитель оператора есть понятие инвариантное, не зависящее от выбора базиса, посчитать такой определитель мы умеем только после того, как запишем наш оператор матрицей в каком-нибудь базисе. Корни характеристического многочлена определены независимо от базиса — это собственные значения оператора, старший коэффициент многочлена равен, как легко видеть, единице, поэтому, в случае, когда характеристический многочлен разлагается на линейные множители он, согласно теореме Виета, определен однозначно, т.е. не зависит от выбора базиса, в котором мы записываем его матрицу и считаем определитель. Нетрудно, однако, привести и явное вычисление, показывающее, что характеристический многочлен в любом случае не зависит от выбора базиса.

Предложение 4.1. *Характеристический многочлен оператора не зависит от выбора базиса.*

Предположим, что в пространстве V имеются два базиса $\mathcal{E}$ и $\mathcal{G}$, и T — матрица перехода от $\mathcal{E}$ к $\mathcal{G}$. Тогда, как мы видели, матрицы оператора f в базисах $\mathcal{E}$ и $\mathcal{G}$, обозначаемые у нас, соответственно, $\hat{f}_{\mathcal{E}}$ и $\hat{f}_{\mathcal{G}}$, связаны соотношением $\hat{f}_{\mathcal{G}} = T^{-1}\hat{f}_{\mathcal{E}}T$. Тогда $\det(tE - \hat{f}_{\mathcal{G}}) = \det(tE - T^{-1}\hat{f}_{\mathcal{E}}T) = \det(T^{-1}tET - T^{-1}\hat{f}_{\mathcal{E}}T) = \det\left[T^{-1}(tET - \hat{f}_{\mathcal{E}})T\right] = \det(T^{-1})\det(tE - \hat{f}_{\mathcal{E}})\det T = \det(tE - \hat{f}_{\mathcal{E}})(\det T)^{-1}\det T = \det(tE - \hat{f}_{\mathcal{E}}).\square$

Задача 4.1. *Докажите, что свободный член характеристического многочлена оператора* f *равен* $(-1)^n \det f$, *а коэффициент при* t^{n-1} *равен* $-\operatorname{tr} f$.

Задача 4.2. * *Докажите, что коэффициент при* t^{n-k} *характеристического многочлена равен сумме всех миноров* k*-ого порядка, симметричных относительно главной диагонали (сколько всего таких миноров?) домноженной на* $(-1)^k$.

5 Минимальный многочлен.

Кроме характеристического, существует еще один многочлен, естественно сопоставляемый линейному оператору — это его минимальный многочлен. Пусть $P(t) = t^m + a_1 t^{m-1} + \ldots + a_{m-1} t + a_m \in \mathbb{K}[t]$ — некоторый многочлен. Нетрудно объяснить, как подставить оператор f в многочлен: это будет линейный оператор $P(f) = a_0 f^m + a_1 f^{m-1} + \ldots + a_{m-1} f + a_m \operatorname{Id}_V$. Если A — матрица оператора f в некотором базисе, то матрицей оператора $P(f)$ будет, конечно, матрица $P(A) = a_0 A^m + a_1 A^{m-1} + \ldots + a_{m-1} A + a_m E$, где $E = (\delta_{i,j})$ — единичная матрица.

Докажем, что для любого оператора f существует многочлен, при подстановке в который оператора f получается 0. Про такой многочлен говорят, что он **аннулирует** оператор f. Для того, чтобы убедиться в существовании аннулирующего многочлена для данного оператора f, достаточно заметить, что пространство всех линейных операторов из V в V имеет размерность n^2 (при выборе базиса это просто пространство $n \times n$ матриц), поэтому операторы $\operatorname{Id}_V, f, f^2, f^3, \ldots, f^{n^2}$ линейно зависимы, поэтому между ними существует некоторое линейное соотношение $b_0 \operatorname{Id}_V + b_1 f + b_2 f^2 + b_3 f^3 + \ldots + b_{n^2} f^{n^2} = 0$, в котором не все b_i равны нулю. Но это как раз означает, что многочлен $b_0 + b_1 t + b_2 t^2 + b_3 t^3 + \ldots + b_{n^2} t^{n^2} \in \mathbb{K}[t]$ аннулирует оператор f. На самом деле, как мы скоро увидим, степень этого многочлена (n^2) здесь сильно завышена: мы скоро докажем, что всегда можно найти аннулирующий многочлен степени не выше n.

Определение 5.1. *Многочлен наименьшей степени со старшим коэффициентом* 1*, аннулирующий оператор f, называется* **минимальным многочленом** *оператора f и обозначается $\mu_f(t)$.*

Предложение 5.1. *Пусть многочлен $P(t) \in \mathbb{K}[t]$ аннулирует оператор f, т.е. $P(f) = 0$. Тогда многочлен $P(t)$ нацело делится на минимальный многочлен $\mu_f(t)$ оператора f.*

Доказательство почти очевидно: разделим $P(t)$ с остатком на $\mu_f(t)$; получим

$$P(t) = \mu_f(t) S(t) + Q(t), \quad \text{где } \deg Q < \deg \mu_f.$$

Подставляя в это равенство оператор f, получаем $P(f) = \mu_f(f) S(t) + Q(f)$, откуда в силу $P(f) = 0$ и $\mu_f(f) = 0$ следует $Q(f) = 0$, что противоречит минимальности многочлена μ_f. Следовательно, Q может быть только нулевым многочленом, что и требовалось доказать.$\square$

6 Бывает все на свете хорошо: если есть базис, составленный из собственных векторов.

Предположим, что мы нашли такой базис $e_1, \ldots, e_n$, что каждый базисный вектор e_i является собственным вектором с собственным значением λ_i. Это означает, что матрица оператора f в базисе $e_1, \ldots, e_n$ является диагональной с диагональными элементами $\lambda_1, \ldots, \lambda_n$:

$$\hat{f} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}. \tag{6.1}$$

Очевидно, верно и обратное: если матрица оператора в некотором базисе имеет диагональный вид (6.1), то базисные векторы $e_1, \ldots, e_n$ являются собственными с собственными значениями $\lambda_1, \ldots, \lambda_n$.

Ясно, что в этом случае характеристический многочлен имеет корни и разлагается на линейные множители $\chi_f(t) = (t-\lambda_1)(t-\lambda_2)\ldots(t-\lambda_n)$.

Нетрудно в этом случае вычислить и минимальный многочлен оператора f. Для этого достаточно заметить, что если $P(t) \in \mathbb{K}[t]$, то для диагональной матрицы (6.1)

$$P(\hat{f}) = \begin{pmatrix} P(\lambda_1) & 0 & \ldots & 0 \\ 0 & P(\lambda_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P(\lambda_n) \end{pmatrix}. \tag{6.2}$$

Следовательно, $P(f) = 0$ тогда и только тогда, когда $P(\lambda_i) = 0 \ \forall i$, т.е. когда все λ_i являются корнями многочлена P. Таким образом, минимальный многочлен $\mu_f(t)$ представляет собой произведение линейных множителей $t - \lambda_i$ по всем *различным* собственным значениям λ_i. для удобства записи обозначим через $\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_s}$ все *различные* собственные значения, а через $k_1, k_2, \ldots, k_s$ — их кратности (т.е. число λ_{i_m} встречается на диагонали матрицы (6.1) k_m раз ($m = 1, 2, \ldots, s$), так что $k_1 + k_2 + \ldots + k_s = n$. Тогда наши результаты можно переписать так: $\chi_f(t) = (t-\lambda_{i_1})^{k_1}(t-\lambda_{i_2})^{k_2}\ldots(t-\lambda_{i_s})^{k_s}$, и $\mu_f(t) = (t-\lambda_{i_1})(t-\lambda_{i_2})\ldots(t-\lambda_{i_s})$. Соберем вместе всю полученную информацию о характеристическом и минимальном многочленах в рассматриваемом случае.

Предложение 6.1. *Пусть существует такой базис $e_1, \dots, e_n$, что каждый базисный вектор e_i является собственным вектором с собственным значением λ_i. Тогда матрица оператора f в базисе $e_1, \dots, e_n$ является диагональной матрицей (6.1). Пусть $\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_s}$ все различные собственные значения, а $k_1, k_2, \dots, k_s$ — их кратности. Тогда $\chi_f(t) = (t-\lambda_{i_1})^{k_1}(t-\lambda_{i_2})^{k_2}\dots(t-\lambda_{i_s})^{k_s}$, и $\mu_f(t) = (t-\lambda_{i_1})(t-\lambda_{i_2})\dots(t-\lambda_{i_s})$.*$\square$

Следствие 6.1. *Пусть существует базис, состоящий из собственных векторов оператора f. Тогда*

(1) Минимальный многочлен $\mu_f(t)$ является делителем характеристического многочлена $\chi_f(t)$.

(2) $\mu_f(t)$ и $\chi_f(t)$ имеют один и тот же набор неприводимых делителей.

(3) Если все корни $\chi_f(t)$ различны, то $\mu_f(t) = \chi_f(t)$.$\square$

Мы так подробно сформулировали это очень простое следствие потому, что все эти результаты оказываются верными для любого линейного оператора, независимо от наличия у него базиса, состоящего из собственных векторов. Отметим, что первое утверждение следствия можно переписать в виде $\chi_f(f) = 0$. Это утверждение называется обычно **теоремой Гамильтона-Кэли**. Ее доказательство в общем случае — одна из наших ближайших задач.

Важное замечание: у нас получилось, что если у оператора имеется базис, состоящий из собственных векторов, то характеристический многочлен разлагается на линейные множители. Обратное, вообще говоря, неверно, точнее говоря, верно только в том случае, когда все корни различны это мы докажем чуть позже, см. следствие 6.2. А при наличии кратных корней собственного базиса для оператора f может и не быть.

Задача 6.1. *Докажите, что для линейного оператора в n-мерном пространстве, заданного матрицей $a_{i,j} = \delta_{i+1,j}$ не существует базиса, составленного из собственных векторов. (Найдите все собственные векторы!)*

Отметим еще, что обсуждаемый в этом разделе случай (существование базиса, составленного из собственных векторов) на первый взгляд представляется не самым удобным для проверки: чтобы выяснить, имеет ли место данный случай, нам надо найти все собственные векторы и после этого выяснить, можно ли из них выбрать базис. Однако, как мы сейчас увидим, в большинстве случаев дело обстоит проще: мы можем гарантировать наличие такого базиса, зная только, что характеристический многочлен имеет ровно n различных корней.

Лемма 6.1. *Пусть $u_1, \ldots, u_m$ — собственные векторы оператора f, причем соответствующие им собственные значения $\lambda_1, \ldots, \lambda_m$ все различны. Тогда векторы $u_1, \ldots, u_m$ линейно независимы.*

Будем доказывать лемму от противного: предположим, что эти векторы линейно зависимы. В этом случае они должны удовлетворять некоторому линейному соотношению, или даже, возможно, нескольким линейным соотношениям. Выберем из таких соотношений самое короткое, и для удобства перенумеруем наши векторы таким образом, чтобы в самом коротком линейном соотношении участвовали в точности первые k векторов $u_1, \ldots, u_k$. Пусть это соотношение имеет вид

$$\alpha_1 u_1 + \ldots + \alpha_k u_k = 0. \tag{6.3}$$

Поскольку мы предположили, что это самое короткое соотношение, все $\alpha_i \neq 0$. Применим к обеим частям 6.3 оператор f:

$$f(\alpha_1 u_1 + \ldots + \alpha_k u_k) = f(0) = 0, \tag{6.4}$$

и, пользуясь линейностью оператора и тем, что u_i — собственные векторы, получаем

$$\alpha_1 \lambda_1 u_1 + \ldots + \alpha_k \lambda_k u_= 0. \tag{6.5}$$

Домножая теперь (6.3) на λ_k и вычитая из (6.5), получаем более короткое соотношение

$$\alpha_1(\lambda_1 - \lambda_k)u_1 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)u_{k-1} = 0, \tag{6.6}$$

в котором все коэффициенты отличны от нуля. Противоречие. Следовательно, векторы $u_1, \ldots, u_m$ линейно независимы.□

Следствие 6.2. *Если характеристический многочлен разлагается на линейные множители и не имеет кратных корней (т.е. имеется ровно n различных собственных значений), то существует базис, составленный из собственных векторов.*□

Заметим, что этот, наиболее часто встречающийся случай открывает очень большие возможности для вычислений с матрицами. Продемонстрируем это на одном простом примере. Пусть нам надо вычислить n-ую степень матрицы

$$A = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}. \tag{6.7}$$

Для этого вычислим $\chi_A(t) = t^2 - 2t - 3 = (t-3)(t+1)$. Следовательно, у характеристического многочлена два различных собственных значения $\lambda_1 = 3$ и $\lambda_2 = -1$, поэтому два соответствующих собственных вектора будут образовывать базис. Найдем эти векторы. Согласно определению, собственные векторы, соответствующие собственному значению λ — это ядро оператора $f - \lambda\,\mathrm{Id}$, который записывается матрицей $A - \lambda E$. Поэтому для $\lambda_1 = 3$ нам надо найти ядро для

$$A - 3E = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ 3 & -1 \end{pmatrix}, \tag{6.8}$$

а для $\lambda_2 = -1$ надо найти ядро для

$$A - (-1)E = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix} - (-1)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}. \tag{6.9}$$

Нахождение ядра — это решение системы однородных линейных уравнений. Для $\lambda_1 = 3$ это система

$$\begin{cases} -3x + y = 0 \\ 3x - y = 0 \end{cases}, \tag{6.10}$$

а для $\lambda_2 = -1$ это система

$$\begin{cases} x + y = 0 \\ 3x + 3y = 0 \end{cases}. \tag{6.11}$$

В обеих системах одно уравнение есть следствие второго. Это, конечно, не случайно: мы как раз и искали такие значения λ, при которых определитель этой системы будет нулевым, что означает, что уравнения должны оказаться линейно зависимыми. Решение каждой из систем и будет соответствующим собственным вектором:

$$u_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \text{ для } \lambda_1 = 3 \quad \text{и } u_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ для } \lambda_2 = -1. \tag{6.12}$$

Следовательно, в базисе u_1, u_2 матрица соответствующего оператора диагональна и имеет вид

$$\Lambda = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.13}$$

Две матрицы одного и того же оператора в двух разных базисах связаны соотношением

$$\Lambda = T^{-1}AT, \tag{6.14}$$

где T — матрица перехода от начального базиса к базису u_1, u_2. Напишем эту матрицу перехода

$$T = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}, \tag{6.15}$$

и вычислим обратную матрицу

$$T^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}. \tag{6.16}$$

(Случайно получилось, что T^{-1} оказалась пропорциональна T.)

Осталось выразить A из (6.14) и возвести в степень n:

$$\begin{aligned} A^n = (T\Lambda T^{-1})^n = T\Lambda^n T^{-1} = \\ = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 3^n & 0 \\ 0 & (-1)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} = \\ = \frac{1}{4}\begin{pmatrix} 3^n + 3\cdot(-1)^n & 3^n - (-1)^n \\ 3^{n+1} - 3\cdot(-1)^n & 3^{n+1} + (-1)^n \end{pmatrix}. \end{aligned} \tag{6.17}$$

7 Как строить инвариантные подпространства. Матрица Фробениуса.

Как мы видели в прошлом разделе, проще всего устроены операторы, у которых имеется базис, состоящий из собственных векторов. В частности, к этому классу относятся операторы, у которых характеристический многочлен имеет n различных корней. В случае же, когда собственных векторов мало или нет совсем, для исследования оператора оказываются полезны и инвариантные подпространства большей размерности. (Напомним, что собственные векторы это в точности одномерные инвариантные подпространства.) Сейчас мы опишем более или менее универсальный способ построения инвариантных подпространств.

Возьмем произвольный ненулевой вектор $v \in V$. Если v является собственным вектором, то инвариантное подпространство уже построено. Если же v не является собственным вектором, то векторы v и $f(v)$ линейно независимы. Тогда рассмотрим вектор $f^2(v) = f(f(v))$.

Предположим сначала, что $f^2(v)$ выражается линейно через v и $f(v)$, т.е. $f^2(v) = b_0 v + b_1 f(v)$, то линейная оболочка векторов $u_1 = v$ и $u_2 = f(v)$ является двумерным инвариантным подпространством. Действительно, базисом этой линейной оболочки являются векторы $u_0 = v$ и $u_1 = f(v)$, и каждый из них переводится оператором f в ту же плоскость: $f(u_0) = u_1$ и $f(u_1) = b_0 u_0 + b_1 u_1$. Тем самым мы построили двумерное инвариантное подпространство $U = \langle u_0, u_1 \rangle$, причем матрица ограничения оператора f на это инвариантное подпространство имеет следующий вид:

$$\widehat{f|_U} = \begin{pmatrix} 0 & b_0 \\ 1 & b_1 \end{pmatrix}. \tag{7.1}$$

Однако вполне могло случиться, что вектор $f^2(v)$ не выражается линейно через v и $f(v)$, что, собственно, означает, что три вектора v, $f(v)$ и $f^2(v)$ линейно зависимы. Тогда надо повторить ту же процедуру и выяснить, принадлежит ли вектор $f^3(v)$ линейной оболочке $\langle v, f(v), f^2(v) \rangle$, и повторять так до тех пор, пока не окажется, что векторы $v, f(v), f^2(v) \dots f^{k-1}(v)$ линейно независимы, а вектор $f^k(v)$ через них линейно выражается: $f^k(v) = b_0 v + b_1 f(v) + b_2 f^2(v) + \ldots + b_{k-1} f^{k-1}(v)$. (В n-мерном пространстве V это случится не позже чем через n шагов, поскольку любые $n+1$ векторов в V уже линейно зависимы.) Тогда линейная оболочка U векторов $u_0 = v, u_1 = f(v), u_2 = f^2(v), \ldots, u_{k-1} = f^{k-1}(v)$ будет, очевидно, инвариантным подпространством, причем в базисе $u_0, u_1, u_2 \dots u_{k-1}$ матрица ограничения оператора f на это инвариантное подпространство U имеет следующий вид:

$$\widehat{f|_U} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & b_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & b_1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & b_2 \\ 0 & 0 & 1 & \cdots & 0 & 0 & b_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & b_{k-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & b_{k-1} \end{pmatrix} \tag{7.2}$$

Матрица такого вида называется матрицей Фробениуса. Для нее нетрудно явно посчитать минимальный и характеристический многочлены (это задача из 6 листочка!) и убедиться, что оба эти многочлена совпадают, причем коэффициенты этого многочлена зашифрованы в его последнем столбце. Подведем итоги.

Предложение 7.1. *Пусть векторы* $v, f(v), f^2(v) \dots f^{k-1}(v)$ *линейно независимы, а вектор* $f^k(v)$ *через них линейно выражается:* $f^k(v) = b_0 v + b_1 f(v) + b_2 f^2(v) + \ldots + b_{k-1} f^{k-1}(v)$. *Тогда линейная оболочка* $U = \langle v, f(v), f^2(v), \ldots, f^{k-1}(v) \rangle$ *является инвариантным подпространством, причем в этом базисе матрица ограничения оператора* f *на* U *является матрицей Фробениуса вида (7.2). Минимальный и характеристический многочлены этой матрицы совпадают и равны* $\chi_F(t) = \mu_F(t) = t^k - b_{k-1} t^{k-1} - \ldots - b_1 t - b_0$.$\Box$

Мы видим, что имеется взаимно-однозначное соответствие между многочленами данной степени и матрицами Фробениуса данного размера. Удобно для данного многочлена $P(t) = t^k + a_{k-1} t^{k-1} + \ldots + a_1 t + a_0$ обозначить через F_P матрицу Фробениуса, имеющую P своим минимальным и характеристическим многочленом; тогда, очевидно,

$$F_P = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & 0 & -a_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -a_{k-1} \end{pmatrix}. \tag{7.3}$$

8 Теорема Гамильтона-Кэли

Теперь мы готовы в общем случае доказать теорему Гамильтона-Кэли, утверждающую, что минимальный многочлен всегда является делителем характеристического, или, что в силу Предложения 5.1 равносильно, что характеристический многочлен оператора аннулирует этот оператор.

Теорема 8.1. $\chi_f(f) = 0$.

Будем доказывать эту теорему индукцией по размерности пространства V. В одномерном случае всякий линейный оператор f есть просто умножение на константу λ, поэтому характеристический многочлен этого оператора есть просто $\chi_f(t) = t - \lambda$. Подставляя, получаем $\chi_f(f) = \lambda - \lambda = 0$.

Предположим, что теорема Гамильтона-Кэли доказана для всех размерностей, меньших $\dim V$. Для оператора f на V возможен один из двух случаев: либо у f имеется ненулевое инвариантное подпространство $U \subset V$, $U \neq V$, либо такого инвариантного подпространства нет.

Разберем сначала случай, когда инвариантное подпространство есть. Тогда, согласно предложению 2.1, матрицу оператора f в подходящем базисе имеет блочный верхнетреугольный вид:

$$\hat{f} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}. \tag{8.1}$$

Очевидно, что

$$\chi_f(t) = \chi_A(t)\chi_B(t). \tag{8.2}$$

Поскольку матрицы A и B имеют размер, меньший, чем $\dim V$, по предположению индукции $\chi_A(A) = 0$ и $\chi_B(B) = 0$. Вычислим теперь $\chi_f(f)$:

$$\chi_f(\hat{f}) = \chi_A(\hat{f}) \cdot \chi_B(\hat{f}) = \chi_A \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \cdot \chi_B \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} =$$
$$= \begin{pmatrix} \chi_A(A) & * \\ 0 & \chi_A(B) \end{pmatrix} \begin{pmatrix} \chi_B(A) & * \\ 0 & \chi_B(B) \end{pmatrix} = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix} \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{8.3}$$

(В этом вычислении, как обычно, звездочками обозначены матрицы, конкретный вид которых нам не известен и не важен.) Тем самым в случае наличия инвариантного подпространства шаг индукции сделан.

Предположим теперь, что ненулевого инвариантного подпространства, отличного от всего V, нет. Но ведь мы можем взять любой ненулевой вектор и применить к нему алгоритм, описанный в разделе 7. В результате гарантированно получится ненулевое инвариантное подпространство, которое, следовательно, должно совпасть с V. Следовательно, полученная в этом разделе матрица Фробениуса (7.2) является матрицей оператора f. Но мы уже доказали в предложении 7.1, что у матрицы Фробениуса минимальный и характеристический многочлены совпадают, так что и в этом случае шаг индукции сделан. Теорема Гамильтона-Кэли доказана.□

На самом деле теорема Гамильтона-Кэли дает не всю информацию о взаимоотношениях минимального и характеристического многочленов: она утверждает лишь, что второй делится на первый. Оказывается, их связь теснее: их разложение в произведение неприводимых сомножителей содержит одни и тот же набор неприводимых многочленов, а отличаться могут лишь степени, в которых они входят в разложение.

Доказательство этого факта и выяснение его геометрической природы — наша следующая задача.

9 Неприводимые делители минимального многочлена.

Предположим, что мы разложили минимальный многочлен на множители, и $P(t) \in \mathbb{K}[t]$ — один из неприводимых сомножителей. Тогда, конечно, по теореме Гамильтона-Кэли $P(t)$ является также и делителем характеристического многочлена. Можно попробовать угадать, какие геометрические обстоятельства могли бы к этому привести. Один вариант такой ситуации мы уже рассматривали: так могло бы получиться, если бы нашлось такое инвариантное подпространство U, ограничение на которое оператора f имело бы характеристический (а, следовательно, и минимальный!) многочлен в точности совпадающий с $P(t)$. Чуть позже мы докажем, что именно так всегда и бывает.

Сначала рассмотрим уже хорошо известный нам пример: предположим, что многочлен $P(t)$ оказался самым простым неприводимым многочленом, а именно линейным, т.е. $P(t) = t - \lambda$. Это в точности означает, что λ является корнем характеристического уравнения, т.е. собственным значением, которому, как мы видели, обязательно соответствует хоть один ненулевой собственный вектор $v \in V$. тогда одномерное подпространство, порожденное этим собственным вектором v как раз и будет требуемым инвариантным подпространством U. Действительно, поскольку $f(v) = \lambda v$, ограничение нашего оператора f на U есть в точности умножение на λ, а характеристический многочлен оператора умножения на λ в одномерном пространстве есть в точности $t-\lambda$, т.е. $P(t)$. Следовательно, в этом, самом простом случае, существование

такого инвариантного подпространства нам уже известно. Однако нам будет полезно воспроизвести и схему доказательства существования собственного вектора из теоремы 3.1. Вектор v там получался как элемент ядра оператора
$$f - \lambda \operatorname{Id}_v, \tag{9.1}$$
а число λ специально находили так, чтобы этот оператор $f - \lambda \operatorname{Id}_v$ имел ненулевое ядро. (Для этого, конечно, требовалось, чтобы определитель оператора (9.1) был равен нулю, что и давало характеристическое уравнение для нахождения собственных значений.) Для перенесения этих соображений на общий случай нам осталось только догадаться, почему здесь оказался задействован именно оператор (9.1): это же в точности $P(f)$ для линейного многочлена $P(t) = t - \lambda$.

Попробуем рассуждать таким же образом и в общем случае: пусть $P(t)$ — какой-нибудь неприводимый делитель минимального многочлена $\mu_f(t)$. В рассуждении с собственным вектором нам существенно помогло то, что у оператора $g = P(f)$ имелось ненулевое ядро. Докажем, что в нашем случае это тоже так.

Лемма 9.1. *Пусть $P(t)$ — какой-нибудь неприводимый делитель минимального многочлена $\mu_f(t)$. Тогда у оператора*
$$g = P(f) \tag{9.2}$$
имеется ненулевое ядро.

Для доказательства запишем разложение минимального многочлена на два множителя:
$$\mu_f(t) = P(t)Q(t). \tag{9.3}$$
Предположим противное, что ядро оператора $g = P(f)$ нулевое, тогда оператор $g = P(f)$ обратим. Подставим в (9.3) оператор f вместо t, получим $\mu_f(f) = P(f) \cdot Q(f)$, то есть $0 = g \cdot Q(f)$. Но если оператор g обратим, то мы можем умножить последнее равенство слева на g^{-1}, что даст $Q(f) = 0$. Но это противоречит минимальности многочлена μ_f, поскольку Q является его делителем, и, следовательно, имеет меньшую степень. Полученное противоречие доказывает лемму. □

Итак, $\operatorname{Ker} g$ является ненулевым подпространством, инвариантным относительно f. (Инвариантность следует из задачи 2.2, поскольку операторы f и $g = P(f)$ очевидно, перестановочны.) Каким может быть минимальный многочлен ограничения оператора f на $\operatorname{Ker} g$? По определению $g = P(f)$ является нулевым оператором на $\operatorname{Ker} g$, поэтому многочлен P аннулирует ограничение оператора f на $\operatorname{Ker} g$. Следовательно, согласно предложению 5.1, многочлен $P(t)$ делится на его минимальный многочлен. Следовательно, в силу неприводимости $P(t)$, минимальный многочлен с ним просто совпадает.

Однако эти рассуждения пока не позволяют нам что-то сказать о характеристическом многочлене. Чтобы это сделать, возьмем ненулевой вектор $v \in \operatorname{Ker} g$. Применяя, согласно алгоритму из раздела 7(см. предложение 7.1), последовательные степени оператора f к вектору v до тех пор, пока эти эти векторы остаются линейно независимыми, получим инвариантное подпространство U в $\operatorname{Ker} g$, матрица ограничения на которое нашего оператора f есть матрица Фробениуса, у которой, как мы знаем (см. предложение 7.1) характеристический многочлен совпадает с минимальным. Что же это может быть за многочлен? Понятно, что минимальный многочлен ограничения оператора на меньшее подпространство является делителем минимального многочлена того же оператора на большем подпространстве $\operatorname{Ker} g$, то есть делителем неприводимого многочлена $P(t)$, и, значит, должен совпадать с $P(t)$. Итак, мы доказали следующее.

Теорема 9.1. *Пусть неприводимый многочлен $P(t) \in \mathbb{K}[t]$, $\deg P = k$, является делителем минимального многочлена μ_f оператора f. Тогда существует такое k-мерное инвариантное подпространство U, что $P(t)$ является минимальным и характеристическим многочленом ограничения оператора f на U, а матрица ограничения оператора f на U в подходящем базисе есть матрица Фробениуса F_P многочлена $P(t)$ (см. (7.3)).*$\square$

То же самое можно переформулировать на матричном языке.

Следствие 9.1. *Пусть неприводимый многочлен $P(t) \in \mathbb{K}[t]$, $\deg P = k$, является делителем минимального многочлена μ_f оператора f. Тогда в некотором базисе матрица оператора f имеет блочный верхнетреугольный вид*

$$\hat{f} = \begin{pmatrix} F_P & C \\ 0 & B \end{pmatrix}, \tag{9.4}$$

где F_P есть матрица Фробениуса многочлена $P(t)$ (см. (7.3)).$\square$

Теперь мы можем применить нашу теорему в оператору, заданному матрицей B из (9.4) и выделить в ней еще один блок и действовать так до тех пор, пока матрица нашего оператора не приобретет такой блочный верхнетреугольный вид:

$$\hat{f} = \begin{pmatrix} F_{P_1} & * & \cdots & * \\ 0 & F_{P_2} & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & F_{P_s} \end{pmatrix} \tag{9.5}$$

Здесь, как обычно, звездочками обозначены матрицы, конкретный вид которых нам не известен и не важен, а F_{P_i} есть матрица Фробениуса неприводимого многочлена $P_i(t)$. Отметим, отнюдь не предполагается, что неприводимые многочлены $P_i(t)$ все различны. Теперь мы можем легко найти характеристический многочлен нашего оператора: поскольку $\chi_{F_P}(t) = \mu_{F_P}(t) = P(t)$,

$$\chi_f = P_1 P_2 \dots P_s. \tag{9.6}$$

О минимальном многочлене мы можем сказать только то, что μ_f должен, по крайней мере, аннулировать каждый диагональный блок, и, следовательно, должен делиться на каждый многочлен P_i. Следовательно, каждый неприводимый делитель характеристического многочлена является также и делителем минимального. При этом, конечно, возможно, что один и тот же неприводимый многочлен входит в разложение μ_f с меньшей кратностью, чем в разложение χ_f. (Так например, обстоит дело с тождественным оператором, у которого $\chi_{\mathrm{Id}_V}(t) = (t-1)^n$, но $\mu_{\mathrm{Id}_V}(t) = t-1$.) В результате получилась важная теорема, дополняющая теорему Гамильтона-Кэли.

Теорема 9.2. *Пусть $\chi_f(t) = P_1(t)P_2(t)\dots P_s(t)$ — разложение характеристического многочлена на неприводимые сомножители. Тогда в подходящем базисе матрица оператора f имеет блочный верхнетреугольный вид (10.8), где F_{P_i} есть матрица Фробениуса неприводимого многочлена $P_i(t)$. При этом минимальный многочлен $\mu_f(t)$ делится на все неприводимые многочлены $P_i(t)$.*$\square$

10 Корневое разложение.

В теореме 9.2 мы научились по разложению характеристического многочлена на неприводимые множители приводить матрицу оператора к блочному верхнетреугольному виду. Сейчас мы получим более сильный результат: оказывается, взаимно простые сомножители характеристического многочлена отвечают разложению пространства V в прямую сумму инвариантных подпространств, что позволяет получить не только блочно-верхнетреугльную матрицу, но даже блочно-диагональную.

Сгруппируем одинаковые неприводимые делители характеристического многочлена:

$$\chi_f = P_1^{r_1} P_2^{r_2} \dots P_m^{r_m} \tag{10.1}$$

где $P_1, P_2, \dots P_m$ — различные неприводимые многочлены. Тогда, как следует из той же теоремы 9.2, минимальный многочлен состоит из тех же делителей, только, возможно, в меньших степенях:

$$\mu_f = P_1^{l_1} P_2^{l_2} \dots P_m^{l_m}, \quad 0 < l_i \le r_i. \tag{10.2}$$

Пусть $P(t)$ один из многочленов $P_1, P_2, \dots P_m$, скажем, $P = P_1$. Снова, как и в разделе 9, рассмотрит оператор

$$g = P(f); \tag{10.3}$$

тогда лемма 9.1 утверждает, что $\operatorname{Ker} g$ ненулевое.

Можно ли придумать разложение V в прямую сумму инвариантных подпространств, в котором подпространство $\operatorname{Ker} g$ было бы одним из слагаемых? Естественным кандидатом на второе прямое слагаемое является подпространство $\operatorname{Im} g$, поскольку оно, во-первых, инвариантно (это та же задача 2.2), а, во-вторых, имеет нужную размерность: $\dim \operatorname{Ker} g + \dim \operatorname{Im} g = \dim V$. Однако для разложения V в прямую сумму этого еще недостаточно: необходимо еще (и достаточно) чтобы $\operatorname{Ker} g$ и $\operatorname{Im} g$ имели нулевое пересечение. Но на это надеяться не приходится — мы знаем примеры операторов, у которых ядро и образ, например, совпадают. Однако эту неприятность можно исправить, если вместо g рассмотреть степени этого оператора.

Заметим, что ядра степеней оператора g образуют возрастающую последовательность подпространств: $\operatorname{Ker} g \subset \operatorname{Ker} g^2 \subset \operatorname{Ker} g^3 \subset \dots$, и, аналогично, образы степеней оператора g образуют убывающую последовательность подпространств: $\operatorname{Im} g \supset \operatorname{Im} g^2 \supset \operatorname{Im} g^3 \supset \dots$. Поскольку размерность подпространства может принимать конечное число значений (это натуральное число, меньшее $\dim V$), некоторые включения в этой последовательности на самом деле являются равенствами. Предположим, что это в первый раз произошло на k-ом шаге: оказалось, что $\operatorname{Ker} g^k = \operatorname{Ker} g^{k+1}$, в то время как все предыдущие включения были строгими. Заметим, что поскольку для любой степени q $\dim \operatorname{Ker} g^q + \dim \operatorname{Im} g^q = \dim V$, то же самое верно и относительно последовательности образов: $\operatorname{Im} g^k = \operatorname{Im} g^{k+1}$, а все предыдущие включения также строгие. Заметим, что из этого следует, что ограничение оператора g на подпространство $W = \operatorname{Im} g^k = \operatorname{Im} g^{k+1}$ является обратимым (на этом подпространстве W) оператором. Действительно, $\forall v \in W = \operatorname{Im} g^{k+1}$ $v = g^{k+1}(v')$ при некотором $v' \in V$. Но тогда $v = g^{k+1}(v') = g(g^k(v')) = g(w)$, где $w = g^k(v') \in W = \operatorname{Im} g^k$. Следовательно, оператор g сюръективен на W, но на конечномерном пространстве сюръективность линейного оператора равносильна его обратимости. Заметим, что из этого следует, что дальше каждая последовательность стабилизируется: поскольку g взаимно однозначен на W, все последующие применения оператора g будут иметь образом все тоже подпространство W. Из обратимости g на W также сразу следует, что W не может содержать ненулевых векторов из $\operatorname{Ker} g^k$ (и вообще, векторов из ядра любой степени g). Следовательно, подпространства W и $U = \operatorname{Ker} g^k$ пересекаются только по нулю, сумма их размерностей равна $\dim V$, и при этом оба они инвариантны относительно f.

Следовательно, мы нашли искомое разложение в прямую сумму $V = U \oplus W$, что, согласно предложению 2.2, позволяет записать матрицу оператора f в блочно-диагональном виде

$$\hat{f} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \tag{10.4}$$

где A и B являются матрицами ограничения оператора f соответственно, на U и W.

Ясно, что $\chi_f(t) = \chi_A(t)\chi_B(t)$, и нам осталось только понять, как распределились сомножители (10.1) между $\chi_A(t)$ и $\chi_B(t)$. Но поскольку $U = \operatorname{Ker} g^k = \operatorname{Ker} P(f)^k$, это означает, что многочлен $P(t)^k$ аннулирует ограничение оператора f на U, то есть, аннулирует матрицу A. Следовательно, согласно предложению 5.1, многочлен $\mu_A(t)$ является делителем многочлена $P(t)^k$. Степень неприводимого многочлена может делиться только на меньшую степень того же многочлена, поэтому $\mu_A(t) = P(t)^l$, где $l \leq k$. Следовательно, согласно теореме 9.2 $\chi_A(t)$ также является какой-то степенью многочлена $P(t)$, поскольку минимальный и характеристический многочлены имеют один и тот же набор неприводимых делителей. А может ли $P(t)$ быть делителем $\chi_B(t)$, или, что равносильно, делителем $\mu_B(t)$? Если бы это было так, то оператор $g = p(f)$ согласно лемме 9.1 имел бы ненулевое ядро, в то время как мы доказывали, что оператор $g = p(f)$ обратим на W. Следовательно, $\chi_B(t)$ не делится на $P(t)$, так что (вспоминая, что через P мы для краткости обозначили P_1 из (10.1)), $\chi_A(t) = P_1(t)^{r_1}$, а $\chi_B(t) = P_2^{r_2} \dots P_m^{r_m}$. Подведем итоги.

Лемма 10.1. *Пусть $\chi_f(t) = P(t)^r Q(t)$, где многочлен $P(t)$ неприводим и не делит $Q(t)$. Тогда существует разложение V в прямую сумму инвариантных подпространств $V = U \oplus W$, так что $\chi_{f|_U}(t) = P(t)^r$ и $\chi_{f|_W}(t) = Q(t)$.* $\square$

Применяя эту лемму последовательно к каждому сомножителю разложения (10.1), получаем теорему о корневом разложении.

Теорема 10.1. *Пусть $\chi_f = P_1^{r_1} P_2^{r_2} \dots P_m^{r_m}$, где $P_1, P_2, \dots P_m$ — различные неприводимые многочлены. Тогда существует разложение V в прямую сумму инвариантных подпространств*

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_m, \tag{10.5}$$

называемое корневым разложением, такое что

$$\chi_{f|_{U_i}}(t) = P_i(t)^{r_i}. \tag{10.6}$$

В базисе, полученном последовательным объединением базисов корневых подпространств U_i, матрица оператора f имеет блочно-диагональный вид

$$\hat{f} = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_m \end{pmatrix}, \tag{10.7}$$

где A_i это матрица ограничения оператора f на подпространство U_i.$\square$

Наконец, мы кое-что можем сказать и о виде матриц A_i, если на каждом U_i мы воспользуемся теоремой 9.2.

Следствие 10.1. *В каждом корневом подпространстве U_i можно выбрать базис таким образом, что матрица A_i из (10.7) имеет блочный верхнетреугольный вид*

$$A_i = \begin{pmatrix} F_{P_i} & * & \cdots & * \\ 0 & F_{P_i} & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & F_{P_i} \end{pmatrix}, \tag{10.8}$$

где на диагонали стоят r_i одинаковых матриц Фробениуса F_{P_i}, а звездочками, как обычно, обозначены матрицы, конкретный вид которых нам не известен и не важен.$\square$

Более подробную информацию о виде матриц A_i можно получить в случае, когда многочлен P_i линеен, то есть $P_i(t) = t - \lambda_i$. В этом случае матрица Фробениуса представляет собой просто число λ_i, так что матрица A_i превращается в обычную верхнетреугольную матрицу, причем на диагонали всюду стоит одно и то же число λ_i. Тогда матрица $A_i - \lambda_i E$ является нильпотентной матрицей, что, конечно соответствует уже многократно использованному факту, что на U_i ограничение оператора $g_i = P_i(f)$ нильпотентно. (Здесь $P_i(f) = f - \lambda_i \operatorname{Id}$.) Строение нильпотентных операторов мы обсудим в следующем разделе.

11 Нильпотентные операторы.

Пусть $g : v \to V$ — нильпотентный оператор, т.е. $g^N = 0$ для некоторого натурального N. Это значит, что многочлен t^N аннулирует g, и потому делится на его минимальный многочлен, так что $\mu_g(t) = t^k$. Характеристический многочлен $\chi_g(t)$ — это многочлен степени $n = \dim V$ имеющий тот же набор неприводимых делителей, что и $\mu_g(t)$, поэтому $\chi_g(t) = t^n$. Подведем первые итоги.

Предложение 11.1. *Если g — нильпотентный оператор, то $\chi_g(t) = t^n$, $\mu_g(t) = t^k$ где $n = \dim V$, $k \le n$.*

Напомним, что в разделе 7 был описан совершенно общий алгоритм построения инвариантного подпространства, начиная с любого вектора $v \in V$, в результате применения которого получалось, что ограничение нашего оператора на это подпространство имеет очень простой стандартный вид и записывается матрицей Фробениуса (7.3). Оказывается, что нильпотентный оператор устроен настолько просто, что что пространство V можно разложить в прямую сумму таких подпространств — в этом, собственно, и состоит основная теорема, которую мы докажем в этом разделе. Однако пред тем, как доказывать эту теорему, давайте выясним, какой может получиться матрица Фробениуса для какого-нибудь инвариантного подпространства, построенного по рецепту раздела 7. Поскольку ограничение g на любое инвариантное подпространство также будет нильпотентным многочленом, это будет матрица Фробениуса многочлена t^m, где m — размерность инвариантного подпространства, так что последний столбик матрицы Фробениуса (7.3) будет просто нулевой:

$$F_{t^n} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}. \tag{11.1}$$

Напомним (подробности см. в разделе 7), что это матрица ограничения оператора g на инвариантное подпространство U, полученное многократным применением к какому-нибудь вектору $v \in V$ оператора g:

$$U = \langle v, g(v), g^2(v), \ldots, g^{m-1}(v) \rangle, \tag{11.2}$$

причем векторы

$$v, g(v), g^2(v), \ldots, g^{m-1}(v) \tag{11.3}$$

составляют базис U, а $g^m(v) = 0$. Отметим, что следствием конструкция инвариантного подпространства из раздела 7 оказался следующий не вполне очевидный факт: если оператор g нильпотентен, а векторы $v, g(v), g^2(v), \ldots, g^{m-1}(v)$ ненулевые, то они линейно независимы.

Задача 11.1. *Докажите это утверждение непосредственно, без ссылки на конструкцию из раздела 7.*

Некоторым недостатком матрицы (11.1) является то, что она является нижнетреугольной, в то время как все получавшиеся у нас блочные матрицы оказывались верхнетреугольными. Этой беде легко помочь: надо просто перенумеровать базис (11.3) задом наперед

$$u_1 = g^{m-1}(v), u_2 = g^{m-2}(v), \dots, u_{m-2} = g^2(v), u_{m-1} = g(v), u_m = v. \quad (11.4)$$

Нетрудно проверить, что в базисе $u_1, \dots, u_m$ матрица ограничения g на то же самое подпространство U имеет аналогичный верхнетреугольный вид:

$$\widehat{g|_U} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \quad (11.5)$$

Матрица (11.5) обычно называется матрицей Жордана или жордановой клеткой, хотя и для (11.1) это название тоже иногда используется. Мы будем использовать для жордановой клетки (11.5) обозначение $J_{0,m}$, где второй индекс указывает на размерность, а первый — на то, что на диагонали этой матрицы стоят нули. Дело в том, что общий вид жордановой клетки

$$J_{\lambda,m} = J_{0,m} + \lambda E = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \quad (11.6)$$

нас в этом разделе интересовать не будет, потому что нильпотентной она будет только при $\lambda = 0$.

Теперь мы возвращаемся к обещанной теореме. Пусть g — нильпотентный оператор, $\chi_g(t) = t^n$, $\mu_g(t) = t^k$. Далее мы берем произвольный вектор $v \in V$ и строим инвариантное подпространство (11.2). Как уже было сказано, мы хотим доказать, что V можно разложить в прямую сумму таких подпространств. Однако, оказывается, не любое построенное таким образом инвариантное подпространство U выделяется прямым слагаемым: надо правильно подобрать начальный вектор v. А именно, поскольку $\mu_g(t) = t^k$, это означает, что $g^k = 0$, в то время как оператор g^{k-1} еще ненулевой. Так вот, мы выберем вектор v таким образом, чтобы вектор $g^{k-1}(v)$ также был ненулевым, так что инвариантное подпространство U будет иметь максимальную возможную размерность k.

Лемма 11.1. *Пусть g — нильпотентный оператор, $\chi_g(t) = t^n$, $\mu_g(t) = t^k$, а вектор $v \in V$ таков, что $g^{k-1}(v) \neq 0$. Тогда инвариантное подпространство U (11.2) выделяется прямым слагаемым, т.е. существует такое инвариантное подпространство W, что $V = U \oplus W$.*

Поскольку подпространство U по построению инвариантно, согласно предложению 2.1 матрица оператора g в базисе, первые k элементов которого составляют базис U, может быть записана в виде

$$\hat{g} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \tag{11.7}$$

A — это матрица ограничения оператора g на U, то есть $A = J_{0,k}$ из (11.5). Условие $g^k = 0$ означает тогда, что $A^k = 0$ (что для матрицы (11.5) выполнено автоматически),

$$B^k = 0 \tag{11.8}$$

и

$$A^{k-1}C + A^{k-2}CB + A^{k-3}CB^2 + \ldots + CB^{k-1} = 0. \tag{11.9}$$

(В левой части последнего равенства стоит выражение из правого верхнего угла матрицы $\hat{g}^k$, вычислить которое предлагалось в задаче 2.3. Те, кто еще не решил эту задачу, могут легко доказать выписанный ответ методом математической индукции.)

При этом дополнение базиса подпространства U до базиса всего пространства V выбиралось произвольно. Нам же желательно выбрать эти векторы так, чтобы матрица C оказалась нулевой: согласно предложению 2.2 это в точности равносильно разложению в прямую сумму инвариантных подпространств. Для этого попробуем построить новый базис, в котором первые k векторов (т.е. базис U)мы менять не будем, а остальные вектора изменим, вычитая из них подходящие векторы из U. Нетрудно понять, что это означает, что матрица перехода к новому базису будет иметь следующий вид

$$T = \begin{pmatrix} E & -X \\ 0 & E \end{pmatrix}. \tag{11.10}$$

Действительно, поскольку в столбцах матрицы перехода стоят векторы нового базиса, в (11.11) написано, что первые k векторов нового базиса совпадают с первыми k векторами старого базиса (напомним, что мы обозначили их в (11.4) $u_1, \dots, u_k$, а каждый j-ый вектор нового базиса при $j > k$ получается вычитанием из j-ого вектора старого базиса поправки $\sum_{i=1}^k X_{i,j} u_i \in U$. ($X_{i,j}$ — это элементы матрицы X.)

Таким образом, мы хотим подобрать матрицу X так, чтобы у матрицы оператора в новом базисе, т.е. у матрицы $T^{-1}\hat{g}T$ в правом верхнем углу был нуль. Вычисляем:

$$T^{-1}\hat{g}T = \begin{pmatrix} E & X \\ 0 & E \end{pmatrix} \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \begin{pmatrix} E & -X \\ 0 & E \end{pmatrix} = \begin{pmatrix} A & C + XB - AX \\ 0 & C \end{pmatrix}. \tag{11.11}$$

Таким образом, нам нужно только решить уравнение $C + XB - AX = 0$, или

$$AX = C + XB. \tag{11.12}$$

Нам будет удобно записать эти уравнения отдельно для каждой строчки матрицы X, и для этого давайте обозначим i-ую строчку матрицы X через x_i, а i-ую строчку матрицы C через c_i, так что

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}; \quad C = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}. \tag{11.13}$$

Матрица A — это жорданова клетка (11.5), и легко проверить, что умножение на нее слева сдвигает строки матрицы X на одну позицию вверх, а последнюю строчку заполняет нулями:

$$AX = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_k \\ 0 \end{pmatrix} \tag{11.14}$$

Умножение же матрицы X справа на B приводит просто к умножению каждой строки x_i на B справа:

$$XB = \begin{pmatrix} x_1B \\ x_2B \\ \vdots \\ x_{k-1}B \\ x_kB \end{pmatrix}, \tag{11.15}$$

так что теперь мы можем переписать (11.12) более явно:

$$\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_k \\ 0 \end{pmatrix} = \begin{pmatrix} x_1B \\ x_2B \\ \vdots \\ x_{k-1}B \\ x_kB \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k-1} \\ c_k \end{pmatrix} \tag{11.16}$$

Можно переписать то же самое как систему уравнений (на строки матрицы X):

$$\begin{cases} x_2 = x_1B + c_1 \\ x_3 = x_2B + c_2 \\ \dots \\ x_k = x_{k-1}B + c_{k-1} \\ 0 = x_kB + c_k \end{cases} \tag{11.17}$$

Первые $k-1$ уравнение очень легко решить последовательно, выражая все через x_1:

$$\begin{cases} x_2 = x_1B + c_1 \\ x_3 = x_1B^2 + c_1B + c_2 \\ x_4 = x_1B^3 + c_1B^2 + c_2B + c_3 \\ \dots \\ x_k = x_1B^{k-1} + c_1B^{k-2} + c_2B^{k-3} + \ldots + c_{k-2}B + c_{k-1}, \end{cases} \tag{11.18}$$

а затем необходимо подставить x_k в последнее уравнение (11.17):

$$0 = x_1B^k + c_1B^{k-1} + c_2B^{k-2} + \ldots + c_{k-1}B + c_k. \tag{11.19}$$

Мы видим, что, поскольку $B^k = 0$ (см. (11.8)), это равенство не зависит от X. Тем самым решение системы (11.17) существует если и только если выполняется равенство

$$c_1B^{k-1} + c_2B^{k-2} + \ldots + c_{k-1}B + c_k = 0. \tag{11.20}$$

Нам осталось всего лишь узнать в этом уравнении первую строчку равенства (11.9): в самом деле, c_1 — это первая строка матрицы C, c_2 — первая строка матрицы AC, c_3 — первая строка матрицы A^2C, и так далее вплоть до c_k, которая, конечно, является первой строкой матрицы $A^{k-1}C$. Следовательно, решение нашей системы существует, так что мы можем привести матрицу оператора g к требуемому блочно-диагональному виду, и тем самым лемма доказана. □

Далее мы можем последовательно применить эту же лемму матрицам, получающимся в правом нижнем углу, и в конце концов получить блочно-диагональную матрицу, диагональными блоками которой будут жордановы клетки. Это и есть жорданово разложение нильпотентного оператора.

Теорема 11.1. *Пусть g — нильпотентный оператор на конечномерном пространстве V. Тогда существует разложение V в прямую сумму инвариантных подпространств $V = U_1 \oplus \ldots \oplus U_s$, что ограничение g на каждое U_i записывается в подходящем базисе жордановой клеткой:*

$$\widehat{g|_{U_i}} = J_{0,k_i} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{11.21}$$

Тем самым матрица оператора g будет блочно-диагональной матрицей, диагональными блоками которой будут жордановы клетки:

$$\hat{g} = \begin{pmatrix} J_{0,k_1} & 0 & \cdots & 0 \\ 0 & J_{0,k_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & J_{0,k_s} \end{pmatrix}. \tag{11.22}$$

12 Итог: когда характеристический многочлен разлагается на линейные множители.

Случай, когда характеристический многочлен разлагается на линейные множители, является практически самым важным и потому востребованным; это, например, всегда так, если поле $\mathbb{K}$ алгебраически замкнуто.

В этом случае, если

$$\chi_f(t) = (t-\lambda_1)^{n_1}\dots(t-\lambda_s)^{n_s}, \qquad (12.1)$$

где $\lambda_1,\dots,\lambda_s$ все различны, теорема 10.1 дает корневое разложение $V = V_1\oplus\dots\oplus V_s$, где V_i — такие инвариантные подпространства, что ограничение оператора f на них имеет характеристический многочлен

$$\chi_{f|_{V_i}}(t) = (t-\lambda_i)^{n_i}.$$

Это значит, что оператор $g_i = f|_{V_i} - \lambda_i\,\mathrm{Id}_{V_i}$ нильпотентен на V_i и потому имеет там жорданово разложение (11.22). Следовательно, в подходящем базисе на V_i

$$\widehat{g|_{V_i}} = \widehat{f|_{V_i}} - \lambda_i E = \begin{pmatrix} J_{0,k_1} & 0 & \cdots & 0 \\ 0 & J_{0,k_2} & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & J_{0,k_s} \end{pmatrix}. \qquad (12.2)$$

Выражая отсюда $\widehat{f|_{V_i}}$, (с учетом $J_{0,m} + \lambda E = J_{\lambda,m}$, см. (11.6)) получаем вид матрицы ограничения нашего оператора на корневое подпространство:

$$\widehat{f|_{V_i}} = \begin{pmatrix} J_{\lambda_i,k_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_i,k_2} & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & J_{\lambda_i,k_s} \end{pmatrix}. \qquad (12.3)$$

В итоге получается теорема о жордановом разложении.

Теорема 12.1. *Если характеристический многочлен оператора разлагается на линейные множители то в подходящем базисе матрица этого оператора имеет блочно-диагональный вид, причем на диагонали стоят жордановы клетки с собственными значениями на диагонали, так что сумма размеров клеток, отвечающих одному собственному значению, равна кратности этого собственного значения в характеристическом многочлене.*

13 Фактор-пространство

Пусть U — подпространство пространства V. Введем на V отношение эквивалентности: $v\sim v'$ если $v - v'\in U$.

Задача 13.1. *Докажите, что это, действительно, отношение эквивалентности.*

Легко проверить, что это отношение эквивалентности согласовано со структурой линейного пространства, т.е.

$$\text{если } v \sim v' \text{ и } w \sim w', \text{ то } \lambda v \sim \lambda v' \text{ и } v + w \sim v' + w'. \qquad (13.1)$$

Задача 13.2. *Докажите это.*

Классы эквивалентности в этом случае называются **смежными классами**; если вектор v лежит в каком-то смежном классе, то остальные элементы этого смежного класса имеют вид $v + u$, где u — любой вектор из U. Поэтому смежный класс, содержащий вектор v, обозначается

$$v + U = \{v + u, \quad u \in U\}.$$

В этом обозначении, естественно, заложена существенная неоднозначность, поскольку два разных элемента v и v' могут определять один и тот же смежный класс:

$$v + U = v' + U, \quad \text{если} \quad v - v' \in U.$$

Благодаря согласованности сложения и умножения на число (13.1) мы можем ввести на множестве смежных классов структуру линейного пространства, положив

$$(v + U) + (w + U) = (v + w) + U; \quad \lambda(v + U) = \lambda v + U. \qquad (13.2)$$

Множество всех смежных классов обозначается V/U и называется **фактор-пространством**. Обратите внимание, что нулем фактор-пространства является само U; при желании его можно рассматривать как смежный класс $0 + U$.

Задача 13.3. *Докажите, что это определение фактор-пространства корректно.*

Нетрудно вычислить размерность фактор-пространства. Пусть $e_1, \dots, e_k$ — базис подпространства U, дополним его какими-нибудь векторами $e_{k+1}, \dots, e_n$ до базиса всего пространства V. Проверим, что смежные классы

$$e_{k+1} + U, \dots, e_n + U \qquad (13.3)$$

образуют базис фактор-пространства V/U. Для этого нужно проверить две вещи: во-первых, что через эти смежные классы можно выразить любой смежный класс из V/U, а во-вторых, что они линейно независимы.

Возьмем любой вектор $v \in V$ и докажем, что его смежный класс $v+U$ можно выразить через векторы (13.3). Для этого напишем разложение вектора v по базису $e_1, \ldots, e_n$: $v = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{k} x_i e_i + \sum_{i=k+1}^{n} x_i e_i$. Первое слагаемое — элемент U, поэтому $v + U = (\sum_{i=k+1}^{n} x_i e_i) + U = \sum_{i=k+1}^{n} x_i(e_i + U)$.

Аналогично доказывается независимость: пусть есть соотношение $\sum_{i=k+1}^{n} x_i(e_i + U) = U$. (Напоминаем: в правой части стоит смежный класс, являющийся нулевым элементом в V/U.) Это означает, что $\sum_{i=k+1}^{n} x_i e_i \in U$, но тогда его можно разложить по базису U: $\sum_{i=k+1}^{n} x_i e_i = \sum_{i=1}^{k} y_i e_i$. Получили соотношение между элементами базиса пространства V, следовательно, все коэффициенты в нем нулевые. Итак, мы доказали, что (13.3) является базисом в V/U.

Предложение 13.1.

$$\dim V/U = \dim V - \dim U. \ \square \tag{13.4}$$

Теорема 13.1. *Пусть $f : V \to W$ — линейное отображение линейных пространств, которые не обязательно предполагаются конечномерными. Тогда*

$$\operatorname{Im} f \cong V/\operatorname{Ker} f. \tag{13.5}$$

Для доказательства достаточно заметить, что для любого $w \in \operatorname{Im} f$, $w = f(v)$ прообразом $f^{-1}(w)$ является в точности смежный класс $v + \operatorname{Ker} f$, поэтому имеется взаимно-однозначное соответствие между смежными классами и элементами $\operatorname{Im} f$, так что остается только проверить, что это взаимно-однозначное соответствие согласовано с операциями, что оставляется читателю в качестве упражнения. $\square$

Теперь мы, наконец, можем выполнить старое обещание — дать инвариантную интерпретацию правому нижнему углу матрицы (2.1). Напомним, там речь шла о том, что имеется линейный оператор $f : V \to V$ и подпространство U инвариантно относительно f. Тогда, если выбрать базис подпространства U $e_1, \ldots, e_k$ подпространства U и дополнить его векторами $e_{k+1}, \ldots, e_n$ до базиса всего пространства V, то как мы выяснили в разделе 2 матрица оператора f в этом базисе будет блочной верхнетреугольной:

$$\hat{f} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \tag{13.6}$$

где матрица A есть матрица ограничения оператора f на подпространство U, записанная в базисе $e_1, \ldots, e_k$.

Определим в этой ситуации новый оператор $\bar{f}$, действующий на фактор-пространстве V/U по формуле $\bar{f}(v+U) = f(v)+U$. Поскольку запись смежного класса в виде $v+U$ неоднозначна, нам необходимо проверить корректность этого определения, т.е. доказать, что для другой записи того же смежного класса $v+U = v'+U$ результат применения $\bar{f}$ будет тот же самый. Это просто: $v+U = v'+U \Leftrightarrow v-v' = u \in U$, но тогда $f(v)-f(v') = f(v-v') = f(u) \in U$, поэтому $f(v)+U = f(v')+U$. Итак, линейный оператор $\bar{f}$ определен корректно.

Предложение 13.2. *Матрицей B из (13.6) — это матрица оператора $\bar{f}$ в базисе $e_{k+1}+U, \ldots, e_n+U$.* $\square$