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.

Theorem 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.

Definition 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.

Proposition 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).*

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):

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

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. $f(u) = \lambda u$.

Definition 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.

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

Corollary 3.1. *У линейного оператора в n-мерном пространстве может быть не более n различных собственных значений.*

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.

Definition 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.

Proposition 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 $\det(tE - \hat{f}_{\mathcal{G}}) = \det(tE - T^{-1}\hat{f}_{\mathcal{E}}T) = \det(T^{-1}tET - T^{-1}\hat{f}_{\mathcal{E}}T) = \det\left[T^{-1}(tET - \hat{f}_{\mathcal{E}})T\right] = \det(T^{-1})\det(tE - \hat{f}_{\mathcal{E}})\det T = \det(tE - \hat{f}_{\mathcal{E}})(\det T)^{-1}\det T = \det(tE - \hat{f}_{\mathcal{E}})$.

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 $P(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE$, 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 $\mathrm{Id}_V, f, f^2, f^3, \ldots, f^{n^2}$ линейно зависимы, поэтому между ними существует некоторое линейное соотношение $b_0\,\mathrm{Id}_V + b_1f + b_2f^2 + b_3f^3 + \ldots + b_{n^2}f^{n^2} = 0$, в котором не все b_i равны нулю. Но это как раз означает, что многочлен $b_0 + b_1t + b_2t^2 + b_3t^3 + \ldots + b_{n^2}t^{n^2} \in \mathbb{K}[t]$ 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.

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

Proposition 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.

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.

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

Corollary 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 $\mu_f(t) = \chi_f(t)$.

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.

Lemma 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.

Corollary 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: $f^k(v) = b_0 v + b_1 f(v) + b_2 f^2(v) + \dots + b_{k-1} f^{k-1}(v)$. (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 $u_0 = v, u_1 = f(v), u_2 = f^2(v), \dots, u_{k-1} = f^{k-1}(v)$ 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.

Proposition 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: $f^k(v) = b_0 v + b_1 f(v) + b_2 f^2(v) + \dots + b_{k-1} f^{k-1}(v)$. Тогда линейная оболочка $U = \langle v, f(v), f^2(v), \dots, f^{k-1}(v) \rangle$ 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 $\chi_F(t) = \mu_F(t) = t^k - b_{k-1}t^{k-1} - \dots - b_1 t - b_0$.*

Мы видим, что имеется взаимно-однозначное соответствие между многочленами данной степени и матрицами Фробениуса данного размера. Удобно для данного многочлена $P(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1 t + a_0$ 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.

Theorem 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 $\mu_f(t)$, так что

$$\mu_f(t) = P(t)Q(t). \tag{9.2}$$

Рассмотрим, как и раньше оператор $g = P(f)$ и докажем, что у него имеется ненулевое ядро. Предположим противное, что ядро нулевое, тогда оператор $g = P(f)$ обратим. Подставим в (9.2) 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, что оператор

$$g = P(f) \tag{9.3}$$

имеет ненулевое ядро. Следовательно, $\operatorname{Ker} g$ 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 ненуле-

вой 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.

Theorem 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)).*

То же самое можно переформулировать на матричном языке.

Corollary 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)).

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.

Theorem 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 (9.5), где F_{P_i} есть матрица Фробениуса неприводимого многочлена $P_i(t)$. При этом минимальный многочлен $\mu_f(t)$ делится на все неприводимые многочлены $P_i(t)$.*

10 Корневое разложение.