# Семинар 4

## Линейные отображения

Пусть $V$ и $U$ – векторные пространства, например, можно считать, что $V = \mathbb{R}^n$, а $U = \mathbb{R}^m$. Напомню, что линейным отображение $\phi\colon V \to U$ – это отображение, удовлетворяющее двум условиям: (1) $\phi(v+u) = \phi(v) + \phi(u)$ для всех $v, u \in V$ и (2) $\phi(\lambda v) = \lambda\phi(v)$ для всех $v \in V$ и $\lambda \in \mathbb{R}$. Если при этом $\phi$ бьет из одного пространства, в то же самое, т.е. $\phi\colon V \to V$, то $\phi$ называется *линейным оператором*. Напомню, что множество всех линейных отображений из $V$ в $U$ обозначается $\operatorname{Hom}(V, U)$.

Правильно думать про линейные операторы как про «линейные деформации пространства $V$». Например, в $\mathbb{R}^n$ мы можем делать растяжения вдоль координатных осей (на самом деле растяжения вдоль любых прямых годятся). Или можем делать повороты вокруг каких-то прямых. Можно «наклонить» одну координатную ось, зеркальная симметрия, симметрия относительно прямой, плоскости, проекция вектора на прямую, плоскость и еще куча других преобразований описывается линейными операторами.

Важный вопрос: а как задавать линейные отображения и операторы? Оказывается для этого достаточно знать куда отправляется базис.

**Утверждение.** *Пусть $e_1, \ldots, e_n$ – некоторый базис векторного пространства $V$ и $u_1, \ldots, u_n$ – произвольный набор векторов другого пространства $U$. Тогда существует единственное линейное отображение $\phi\colon V \to U$ такое, что $\phi(e_i) = u_i$.*

*Доказательство.* Действительно, пусть $v = x_1e_1 + \ldots + x_ne_n$ – произвольный вектор из $V$. Тогда, если $\phi$ существует, то он должен действовать по правилу

$$\phi(v) = \phi(x_1e_1 + \ldots + x_ne_n) = x_1\phi(e_1) + \ldots + x_n\phi(e_n) = x_1u_1 + \ldots + x_nu_n$$

С другой стороны, легко видеть, что данное равенство однозначно задает линейное отображение. □

## Линейные отображения между $\mathbb{R}^n$ и $\mathbb{R}^m$

В случае $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$ мы можем полностью описать линейные отображения в терминах матриц. Действительно, пусть (в обозначениях предыдущего утверждения) $\phi(e_i) = u_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$, где $e_i$ – стандартный базисный вектор $\mathbb{R}^n$. Тогда

$$\phi\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \phi(x_1e_1 + \ldots + x_ne_n) = x_1u_1 + \ldots + x_nu_n = x_1\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n\begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Пристально вглядевшись в то, что мы только что сделали, можно получить следующее.

**Утверждение.** *Отображение $\mathrm{M}_{m\,n}(\mathbb{R}) \to \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, которое каждой матрице $A$ ставит в соответствие линейное отображение $\phi_A$, действующее $\phi_A(x) = Ax$, где $x \in \mathbb{R}^n$, является изоморфизмом векторных пространств, т.е. это правило биективно и $\phi_{A+B} = \phi_A + \phi_B$ и $\phi_{\lambda A} = \lambda\phi_A$.*

Заметим, что под действием биекции из упражнения выше операция композиции линейных отображений соответствует операции умножения матриц: если $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ и $B \in \mathrm{M}_{n\,k}(\mathbb{R})$, то они соответствуют $\phi_A\colon \mathbb{R}^n \to \mathbb{R}^m$ и $\phi_B\colon \mathbb{R}^k \to \mathbb{R}^n$. Тогда $\phi_A\phi_B\colon \mathbb{R}^k \to \mathbb{R}^m$ совпадает с $\phi_{AB}$. Таким образом, как только в пространствах $V$ и $U$ выбраны базисы, нет разницы между изучением линейных отображений и матриц.

## Удобный формализм

**Матрица линейного отображения** Пусть у нас есть линейное отображение $\phi\colon V \to U$ и пусть $e_1, \ldots, e_n$ – некоторый базис $V$ и $f_1, \ldots, f_m$ – некоторый базис $U$. Тогда каждый вектор $\phi(e_i)$ является линейной комбинацией векторов $f_i$, т.е. $\phi(e_i) = a_{1i}f_1 + \ldots + a_{mi}f_m$. Это можно записать в матричном виде так

$$\begin{pmatrix} \phi(e_1) & \ldots & \phi(e_n) \end{pmatrix} = \begin{pmatrix} f_1 & \ldots & f_m \end{pmatrix}\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

или еще короче

$$\phi \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} = \begin{pmatrix} f_1 & \dots & f_m \end{pmatrix} A$$

Здесь $\phi(e_1, \dots, e_n)$ имеется в виду покомпонентное умножение вектора из $e_i$ на $\phi$ слева. Это одна из форм блочного умножения матриц. Матрица $A$ в этом случае называется матрицей линейного отображения $\phi$ в базисах $e_i$ и $f_i$.

**Действие линейного отображения в координатах** Пусть теперь $v \in V$ – некоторый вектор, который раскладывается по базису $v = x_1 e_1 + \ldots + x_n e_n = (e_1, \dots, e_n)x$, где $x \in \mathbb{R}^n$. Тогда

$$\phi(v) = \phi(e_1, \dots, e_n)x = (f_1, \dots, f_m)Ax$$

То есть вектор $\phi(v)$ раскладывается по базису $f_i$ с координатами $Ax$. Значит в координатах, наше линейное отображение задается по правилу $x \mapsto Ax$. На этот факт можно смотреть так. Если есть отображение $\phi\colon V \to U$, то после выбора базиса в $V$ оно превращается в $\mathbb{R}^n$, после выбора базиса в $U$ оно превращается в $\mathbb{R}^m$, а $\phi$ должен превратиться в отображение умножения на некоторую матрицу слева. Так вот матрица линейного оператора для $\phi$ – это в точности та самая матрица, в которую превратился $\phi$ после выбора базиса.

## Смена базиса и линейные отображения

Линейные отображения – это отображения прежде всего и потому они ничего не знают про выбор базиса. С другой стороны, такие отображения задаются разными матрицами в разных базисах. Тут есть пара вещей которые надо понимать: (1) как меняется матрица линейного отображения и (2) смена базиса позволяет упростить вид матрицы.

Начнем с первого вопроса. Тут есть две ситуации: $\phi\colon V \to U$ и $\phi\colon V \to V$, т.е. случай общего линейного отображения и случай линейного оператора. Главная разница в том, что в первом случае мы можем менять одновременно два базиса и в области определения $\phi$ и в области куда $\phi$ бьет. Во втором случае, базисы меняются одновременно.

**Утверждение.** *Пусть $e_1, \dots, e_n$ и $e'_1, \dots, e'_n$ – два базиса $V$, также $f_1, \dots, f_m$ и $f'_1, \dots, f'_m$ – два базиса $U$. Пусть*

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C \text{ и } (f'_1, \dots, f_m) = (f_1, \dots, f_m)D$$

*где $C \in \mathrm{M}_n(\mathbb{R})$ и $D \in \mathrm{M}_m(\mathbb{R})$ – матрицы перехода. Если $\phi$ задается матрицей $A$ в базисах $e_i$ и $f_i$, то в базисах $e'_i$ и $f'_i$ он задается матрицей $D^{-1}AC$.*

*Доказательство.* Для доказательства воспользуемся замечанием из предыдущего раздела. Нам известно, что $\phi(e_1, \dots, e_n) = (f_1, \dots, f_m)A$, а надо найти матрицу $A'$ такую, что $\phi(e'_1, \dots, e'_n) = (f'_1, \dots, f'_m)A'$. Давайте посчитаем:

$$\phi \begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} = \phi \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} C = \begin{pmatrix} f_1 & \dots & f_m \end{pmatrix} AC = \begin{pmatrix} f'_1 & \dots & f'_m \end{pmatrix} D^{-1}AC$$

Значит $A' = D^{-1}AC$, что и требовалось. ☐

**Следствие.** *Если $\phi\colon V \to V$ в базисе $e_1, \dots, e_n$ записывается матрицей $A$, то в базисе $e'_1, \dots, e'_n$ заданном $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$, $\phi$ записывается матрицей $C^{-1}AC$.*

## Смена базиса в координатах

Пусть теперь $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$, также $e_1, \dots, e_n$ обозначает стандартный базис в $\mathbb{R}^n$ и $f_1, \dots, f_m$ – стандартный базис в $\mathbb{R}^m$. Пусть $e'_1, \dots, e'_n$ – другой базис $\mathbb{R}^n$. Это вектор столбцы, из которых я могу соорудить матрицу $C \in \mathrm{M}_n(\mathbb{R})$, поставив $e'_i$ подряд в качестве столбцов.[1] Аналогично, если $f'_1, \dots, f'_m$ – другой базис из $\mathbb{R}^m$ я могу составить из них матрицу $D \in \mathrm{M}_m(\mathbb{R})$. Обе матрицы $C$ и $D$ невырождены.

Любой вектор $v \in \mathbb{R}^n$ можно записать как

$$v = x_1 e_1 + \ldots + x_n e_n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } x_i$$

[1] В этом случае мы также имеем $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$. Это лишь другой способ описать ту же конструкцию, что и в предыдущем пункте. В столбцах матрицы $C$ стоят координаты векторов $e'_i$ относительно стандартного базиса $e_i$.

С другой стороны, мы можем записать $v$ так

$$v = y_1 e'_1 + \ldots + y_n e'_n = C \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } y_i$$

Аналогично в пространстве $\mathbb{R}^m$ любой вектор $u$ может быть записан в двух системах координат:

$$u = w_1 f_1 + \ldots + w_m f_m = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \text{ или } u = z_1 f'_1 + \ldots + z_m f'_m = D \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}$$

Пусть теперь наше отображение $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ задано матрицей $A$, то есть вектор в координатах $x_i$ переходит в вектор в координатах $w_i$ по правилу

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ или кратко } x \mapsto w = Ax$$

Мы хотим переписать $\phi$ в координатах $y_i$ и $z_i$, то есть записать отображение $\phi$ в виде $y \mapsto z = A'y$. Для этого надо пройти по следующей диаграмме

$$\begin{tikzcd} x = Cy \arrow[r, mapsto] & w = Ax = ACy \arrow[d, mapsto] \\ y \arrow[u, mapsto] \arrow[r, mapsto] & z = D^{-1}w = D^{-1}ACy \end{tikzcd}$$

Стартуем с координат $y$ (левый нижний угол). По ним сначала рассчитываем координаты $x$ (вверх по диаграмме). Потом действуем отображением $\phi$ с помощью матрицы $A$ и получаем вектор $\phi(v)$ в координатах $w$ (вправо по стрелке). Потом пересчитываем координаты $w$ в координаты $z$ (вниз по диаграмме). В результате получаем, что $y \mapsto z = D^{-1}ACy$, т.е. $A' = D^{-1}AC$.

## Образ и ядро отображения

Если $\phi\colon V \to U$ – линейное отображение (как и выше $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$), то с ним можно связать два подпространства. Первое из них – *ядро* $\phi$, а именно: $\ker \phi = \{v \in V \mid \phi(v) = 0\}$.[2] Второе – *образ* $\phi$: $\operatorname{Im} \phi = \phi(V) \subseteq U$, то есть все, что можно получить из $V$, применив к нему $\phi$.

**Связь со СЛУ** Пусть $\phi$ задается матрицей $A \in \mathrm{M}_{m\,n}(\mathbb{R})$, то есть наше отображение имеет вид $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ по правилу $x \mapsto y = Ax$, здесь $x \in \mathbb{R}^n$ и $y \in \mathbb{R}^m$.

- **Ядро** – это пространство решений однородной системы линейных уравнений $\{y \in \mathbb{R}^n \mid Ay = 0\}$.
- **Образ**. Введем следующие обозначения для столбцов матрицы $A$: $A = (A_1|\ldots|A_n)$. Тогда по определению в образе $\phi$ лежат все возможные векторы вида $Ax$. Давайте распишем это так:

  $$\operatorname{Im} \phi = \{Ax \mid x \in \mathbb{R}^n\} = \{x_1 A_1 + \ldots + x_n A_n \mid x_i \in \mathbb{R}\} = \langle A_1, \ldots, A_n \rangle$$

  То есть образ – это линейная оболочка столбцов матрицы $A$. Если $e_1, \ldots, e_n$ – это стандартный базис $\mathbb{R}^n$, то есть все координаты $e_i$ кроме $i$-ой равны нулю, а $i$-я равна единице, тогда $i$-ый столбец матрицы $A$ – это образ вектора $e_i$.
- **Прообраз вектор**. Пусть мы зафиксировали вектор $b \in \mathbb{R}^m$ и хотим найти все векторы $x \in \mathbb{R}^n$ такие, что они переходят в $b$ под действием $\phi$. Тогда это означает, что нам надо решить уравнение $Ax = b$, то есть решение неоднородной системы означает, что мы ищем прообраз к некоторому вектору.

[2] В англоязычной технической литературе ядро еще называют nullspace, что можно перевести как нулевой пространство.

- **Связь между ОСЛУ и СЛУ**. Пусть $x_0$ – произвольное решение для $Ax = b$ и $\ker\phi = \{y \in \mathbb{R}^n \mid Ax = 0\}$ – решения однородной системы. Тогда все решения системы $Ax = b$ имеют вид $x_0 + z$, где $z \in \ker\phi$. То есть прообраз любого вектора $b$ является сдвигом ядра отображения $\phi$. Однако, обратите внимание, прообраз вектора $b$ может быть пуст, а ядро всегда не пусто, в нем как минимум всегда найдется нулевой вектор. Таким образом ядро отвечает за единственность решения, если оно есть.

  Полезно понимать, что для любого $b$ найдется прообраз относительно $\phi$, если в системе $Ax = 0$ (или $Ax = b$) количество главных переменных равно количеству строк матрицы $A$, то есть $m$. В терминах ранга это означает, что $\operatorname{rk} A = m$.

## Свойсва ядра и образа

**Утверждение.** *Пусть $V$ и $U$ – векторные пространства и $\varphi\colon V \to U$ – линейное отображение. Тогда*

1. *$\varphi$ сюръективно тогда и только тогда, когда $\operatorname{Im}\varphi = U$.*
2. *$\varphi$ инъективно тогда и только тогда, когда $\ker\varphi = 0$.*
3. *$\dim\ker\varphi + \dim\operatorname{Im}\varphi = \dim V$.*

*Доказательство.* (1) Это просто переформулировка сюръективности на другом языке.

(2) Так как $\ker\varphi = \varphi^{-1}(0)$ и прообраз всегда содержит 0, то из инъективности вытекает, что $\ker\varphi = 0$. Наоборот, пусть $\varphi(v) = \varphi(v')$, тогда $\varphi(v) - \varphi(v') = 0$. А значит, $\varphi(v - v') = 0$. То есть $v - v'$ лежит в ядре, а значит равен 0, что и требовалось.

(3) Этот пункт я пояснять не буду.

□

Еще полезно понимать, что если в пространствах $V$ и $U$ задать пару подпространств $V' \subseteq V$ и $U' \subseteq U$ такую, что $\dim V' + \dim U' = \dim V$, то найдется (и не одно) линейное отображение $\phi\colon V \to U$ такое, что $\ker\phi = V'$, а $\operatorname{Im}\phi = U'$.

**Проекторы** Пусть $P\colon V \to V$ – линейное отображение со свойством $P^2 = P$. Обозначим $U = \operatorname{Im} P$ и $W = \ker P$. Тогда

1. Для любого $u \in U$ верно $Pu = u$.
2. $U \cap W = 0$.
3. Любой вектор $v \in V$ есть сумма $v = u + w$ для некоторых $u \in U$ и $w \in W$.

Действительно, если $u \in U$, это значит, что $u = Pv$. Тогда $Pu = P^2v = Pv = u$. Если вектор $v \in U \cap W$, то с одной стороны $v = Pv$ из предыдущего, с другой стороны $Pv = 0$ по определению $W$. Для любого вектора $v \in V$ верно $v = Pv + (E - P)v$. Тогда вектор $Pv \in U$, а вектор $(E - P)v \in W$, так как $P(E - P)v = (P - P^2)v = 0$.

Линейный оператор $P\colon V \to V$ со свойством $P^2 = P$ называется проектором. Он проектирует любой вектор $v \in V$ на $\operatorname{Im} P$ и на образе действует тождественно. Геометрически все пространство распадается в «сумму»[3] непересекающихся подпространств $U$ и $W$, так что на $W$ оператор $P$ действует нулем, а на $U$ тождественно.

### Формула БАБА

Давайте я в начале разберу задачу нахождения проекции вектора на подпространство вдоль другого подпространства (здесь нам не нужно никакое скалярное произведение). Пусть $V$ – некоторое векторное пространство и пусть $U, W \subseteq V$ – такие подпространства, что $U \cap W = 0$ и при этом любой вектор $v \in V$ представляется в виде суммы вектора из $U$ и вектора $W$. Тогда на пространстве $V$ задан оператор проекции $P\colon V \to V$ такой, что $\ker P = W$ и $\operatorname{Im} P = U$, то есть, если $v \in V$ раскладывается в сумму $v = u + w$, где $u \in U$ и $w \in W$, то $Pv = u$ – оператор вычисления проекции на $U$ вдоль $W$.

Теперь мы хотим научиться эффективно считать $P$. Для этого предположим $V = \mathbb{R}^n$, $U = \langle u_1, \ldots, u_k\rangle$, $W = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in \mathrm{M}_{s\,n}(\mathbb{R})$. В этом случае $P\colon \mathbb{R}^n \to \mathbb{R}^n$ задается некоторой матрицей. Наша задача – найти эту матрицу.

[3] Я не стал определять формально сумму пространств, но это значит, что любой вектор из $V$ представляется в виде суммы векторов из $U$ и $W$, то что сказано во втором пункте.

Предположим для простоты, что векторы $u_1, \ldots, u_k$ образуют базис $U$, а строки матрицы $A$ линейно независимы. Определим матрицу $B = (u_1|\ldots|u_k) \in \mathrm{M}_{n\,k}(\mathbb{R})$. Тогда утверждаются следующие вещи:

1. Количество столбцов $B$ совпадает с количеством строк $A$, то есть $k = s$.

2. Матрица $AB$ обратима.

3. Оператор проекции задается формулой $P = B(AB)^{-1}A$. Мнемоническое правило «БАБА».

*Доказательство.* 1) Это следует из условия $U \cap W = 0$ и условия, что любой вектор из $V$ раскладывается в сумму векторов из $U$ и $W$. Я позволю себе пропустить эту часть.

(2) Теперь рассмотрим матрицу $AB$. Чтобы доказать ее обратимость надо проверить, что $ABy = 0$ влечет $y = 0$. В этом случае положим $z = By$. Тогда $Az = 0$, то есть $z \in W$ по определению. Кроме того, $z = By$, то есть $z$ – линейная комбинация столбцов $B$. То есть $z \in U$ по определению. Но так как $U \cap W = 0$, то $z = 0$. То есть $By = 0$. Но так как столбцы $B$ линейно независимы, от сюда следует, что $y = 0$.

(3) Теперь выведем формулу для $P$. Пусть $v = u + w$, где $v \in \mathbb{R}^n$ – произвольный вектор, $u \in U$ и $w \in W$ – его разложение по подпространствам $U$ и $W$. Тогда $Av = Au + Aw = Au$. С другой стороны, так как $u \in U$, мы имеем $u = Bx$ для некоторого $x \in \mathbb{R}^k$. Тогда $Av = ABx$. Так как $AB$ обратимая квадратная матрица, имеем $x = (AB)^{-1}Av$. Значит $u = Bx = B(AB)^{-1}Av$, что и требовалось. ☐

Обратите внимание, что проектор $P$ на $U$ вдоль $W$ зависит от двух подпространств, а не только от $U$. Если вы измените одно из них, то проектор изменится.

## Характеристики операторов

**След оператора** Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – квадратная матрица. Тогда *след матрицы* $A$ – это сумма ее диагональных элементов, т.е. $\operatorname{tr} A = a_{11} + \ldots + a_{nn} = \sum_{i=1}^n a_{ii}$. Заметим важное свойство следа: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ (это непосредственная проверка влоб). В частности $\operatorname{tr}(C^{-1}AC) = \operatorname{tr}(A)$ для любых $A, C \in \mathrm{M}_n(\mathbb{R})$ с условием, что $C$ обратима.

Пусть теперь $\phi\colon V \to V$ – некоторый линейный оператор. Тогда в некотором базисе $e_1, \ldots, e_n$ он задается матрицей $A$. Определим *след линейного оператора* $\phi$ как след этой матрицы $A$. Это определение не зависит от базиса. Действительно, в другом базисе оператор $\phi$ задается матрицей $A' = C^{-1}AC$, тогда $\operatorname{tr}(A') = \operatorname{tr}(C^{-1}AC) = \operatorname{tr}(A)$. След оператора $\phi$ также обозначается через $\operatorname{tr}\phi$. Важно понимать, что след – это характеристика линейного оператора, а не его матрицы, т.е. эта штука не зависит от матрицы, которой задается оператор. Однако, мы не можем определить эту характеристику не пользуясь базисом. Более того, в принципе невозможно определить след без базиса!

**Определитель оператора** Пусть $\phi\colon V \to V$ – линейный оператор. Тогда в некотором базисе он задается матрицей $A$. Положим $\det\phi = \det A$. Надо лишь проверить, что это определение не зависит от выбора базиса. Действительно, в другом базисе $\phi$ задается $C^{-1}AC$, а значит $\det(C^{-1}AC) = \det A$. Величина $\det\phi$ называется *определителем линейного оператора*. Как и в случае следа, определитель линейного оператора не зависит от базиса, но его нельзя определить не пользуясь базисом.

**Многочлены от операторов** Пусть $\phi\colon V \to V$ – линейный оператор. Тогда определен линейный оператор $\lambda\phi$ для любого $\lambda \in \mathbb{R}$. Уточним на всякий случай, что $(\lambda\phi)(v) = \lambda\phi(v)$ по определению. Более того, определены все натуральные степени оператора $\phi$, как композиция, т.е. $\phi^n(v) = \phi(\ldots\phi(v)\ldots)$ – где композиция берется $n$ раз, например, $\phi^3(v) = \phi(\phi(\phi(v)))$.[4] Кроме того, линейные операторы можно складывать. Напомню, что $(\phi + \psi)(v) = \phi(v) + \psi(v)$ по определению.

Из выше сказанного следует, что можно составлять выражения вида $\sqrt{\pi}\phi - 2\phi^4$ и их результат будет линейный оператор на $V$. Более того, можно определить $\phi^0$ как тождественный оператор $\mathrm{Id}$, т.е. $\mathrm{Id}(v) = v$.[5] Тогда можно писать выражения вроде $2/3 + \phi^2$, где имеется в виду $2/3\,\mathrm{Id} + \phi^2$.

Таким образом для любого многочлена $p(t) = a_0 + a_1 t + \ldots + a_n t^n$, где $a_i \in \mathbb{R}$, и любого линейного оператора $\phi\colon V \to V$ определен линейный оператор $p(\phi)\colon V \to V$. Когда мы перейдем к базисам, оператор $\phi$ будет соответствовать матрице $A$. В этом случае $p(\phi)$ соответствует матрице $p(A)$.[6]

[4] На самом деле умные люди вообще не пишут скобки, ибо они только загромождают обозначения. Ведь куда приятнее смотреть на $\phi\phi\phi v$ вместо $\phi(\phi(\phi(v)))$. Но еще приятнее смотреть на $\phi^3 v$.

[5] В любом базисе тождественный оператор соответствует единичной матрице.

[6] Предлагаю вам самим восстановить детали того, что такое подставить матрицу в многочлен по аналогии с операторами.

Основной бонус от подстановки матриц и линейных операторов в многочлены состоит вот в чем. Пусть многочлен $p(t)$ раскладывается в произведение $(t-\lambda_1)\dots(t-\lambda_n)$. Тогда оператор $p(\phi)$ раскладывается в композицию операторов $(\phi-\lambda_1 \operatorname{Id})\dots(\phi-\lambda_n \operatorname{Id})$. Скоро (очень очень скоро) будет видно зачем все это нужно.

**Характеристический многочлен** Пусть $\phi\colon V\to V$ – некоторый линейный оператор. Опять же для удобства, можно считать, что после выбора базиса $V=\mathbb{R}^n$ и $\phi$ соответствует некоторой матрице $A\in \operatorname{M}_n(\mathbb{R})$. Тогда выражение $\det(\phi-\lambda \operatorname{Id})=\det(A-\lambda E)$ является многочленом от $\lambda$ степени $n$. Действительно,

$$\det(A-\lambda E)=\det\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22}-\lambda & \dots & a_{an}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix}$$

С усилием вспоминая явную формулу для определителя через перестановки, понимаем, что получается многочлен от $\lambda$. Еще чуть внимательнее присмотревшись к нему, можно заметить, что

$$\det(A-\lambda E)=\det(A)+\ldots+(-1)^{n-1}\operatorname{tr} A\lambda^{n-1}+(-1)^n\lambda^n$$

Данный многочлен обозначим через $\chi_\phi(\lambda)$ или $\chi_A(\lambda)$ и будем называть *характеристическим многочленом* оператора $\phi$ или соответствующей матрицы $A$ (в зависимости от того, о чем идет речь).[7] Еще полезно видеть перед глазами следующее равенство

$$(-1)^n\chi_A(\lambda)=\lambda^n-\operatorname{tr}(A)\lambda^{n-1}+\ldots+(-1)^n\det(A)$$

**Утверждение** (Теорема Гамильтона-Кэли)**.** *Пусть $\phi\colon V\to V$ – линейный оператор и $\chi(t)$ – его характеристический многочлен. Тогда $\chi(\phi)=0$.*[8]

**Минимальный многочлен** Если $\phi\colon V\to V$ – линейный оператор на $n$ мерном пространстве, то как мы видели выше, он зануляется своим характеристическим многочленом, т.е. существуют многочлены $p(t)$ такие, что $p(\phi)=0$. На самом деле, факт существования таких многочленов доказывается проще, чем теорема Гамильтона-Кэли. Наш $\phi$ соответствует матрице $A$. Но тогда $E, A, A^2,\dots,A^{n^2}$ не могут быть линейно независимы. А значит $a_0+a_1A+\ldots+a_{n^2}A^{n^2}=0$, то есть $A$ зануляется многочленом $p(t)=a_0+a_1t+\ldots+a_{n^2}t^{n^2}$. Суть теоремы Гамильтона-Кэли в том, чтобы понизить степень многочлена с $n^2$ до $n$.

Мы скажем, что ненулевой многочлен $p(t)$ является минимальным для $\phi$, если $p(\phi)=0$ и степень многочлена $p$ является минимально возможной. Минимальный многочлен делит любой многочлен зануляющий $\phi$, так как остаток тоже должен занулять $\phi$ и степени меньше. Потому, если у минимальный многочлен нормировать так, чтобы его старший коэффициент был единицей, то он определен однозначно.

**Утверждение.** *Пусть $\phi\colon V\to V$ – линейный оператор. Тогда*

1. *Минимальный многочлен $p(t)$ для $\phi$ со старшим коэффициентом определен однозначно.*
2. *Многочлен $p(t)$ делит любой многочлен зануляющий $\phi$.*
3. *Существует такое число $m$, что характеристический многочлен $\chi_\phi(t)$ делит $p^m(t)$.*

## Собственные значения и вектора оператора

Пусть $\phi\colon V\to V$ – линейный оператор на пространстве $V$. Будем говорить, что вектор $v\in V$ является *собственным*, если $\phi v=\lambda v$.[9] То есть на собственный вектор оператор $\phi$ действует растяжением. Если $\phi v=\lambda v$ для $v\neq 0$, число $\lambda$ называется *собственным значением* оператора $\phi$. При фиксированном $\lambda\in\mathbb{R}$ множество

[7] Надо отметить, что часто характеристическим многочленом называют $\det(\lambda E-A)$, так как в этом случае старший коэффициент по $\lambda$ становится 1. Наш многочлен от этого отличается на $(-1)^n$. Для многих вопросов это не принципиально.

[8] Есть два доказательства этого факта: (1) кустарное, методами линейной алгебры и (2) концептуальное методами коммутативной алгебры. Первое доказательство использует теорему о Жордановой нормальной форме (по сути классификацию всех линейных операторов) и очень геморройное. Второе доказательство в одну строчку – формулы Крамера для модулей над коммутативными кольцами. Его беда в том, что надо объяснить все дурацкие слова в доказательстве. Это не сложно, но требует кучу времени и усилий, чтобы их осознать.

[9] Нулевой вектор является собственным для любого $\lambda\in\mathbb{R}$. Обратите внимание, что в некоторых учебниках собственные вектора обязательно считаются ненулевыми, но это идейно не правильно.

всех собственных векторов с собственным значением $\lambda$, т.е. $\{v \in V \mid \phi v = \lambda v\}$, будем обозначать через $V_\lambda$. Все $V_\lambda$ обязательно будут подпространствами.[10]

Если мы выберем базис в пространстве $V$, то оно превратится в $\mathbb{R}^n$. Наш оператор $\phi$ будет задаваться матрицей $A \in \mathrm{M}_n(\mathbb{R})$. В этом случае, собственный вектор задается уравнением $Ax = \lambda x$, где $x \in \mathbb{R}^n$. Беда в том, что мы пока заранее не знаем, какие $\lambda$ нам подходят. Чтобы это выяснить нужно переписать уравнение так: $(A - \lambda E)x = 0$. Оно имеет решение тогда и только тогда, когда $A - \lambda E$ – вырожденная матрица. Это, в свою очередь, происходит тогда и только тогда, когда $\det(A - \lambda E) = 0$. Напомним, что характеристический многочлен $\phi$ (он же характеристический для $A$) это $\chi(\lambda) = \det(A - \lambda E)$, т.е. получаем следующее.

**Утверждение.** *Пусть $\phi\colon V \to V$ – некоторый линейный оператор с матрицей $A \in \mathrm{M}_n(\mathbb{R})$ в некотором базисе. Тогда*

1. *Все собственные значения оператора $\phi$ – это в точности корни характеристического многочлена $\chi(\lambda) = \det(A - \lambda E)$.*

2. *Если $\lambda$ – НЕ корень характеристического многочлена, то $V_\lambda = 0$.*

3. *Если $\lambda$ – корень характеристического многочлена, то $V_\lambda$ – ненулевое подпространство $V$. Кроме того, $\dim V_\lambda$ не превосходит кратности корня $\lambda$ у характеристического многочлена.*[11]

## Привет от комплексных чисел

Заметим, что собственные значения являются корнями многочлена. С действительными числами есть беда: многочлены могут вообще не иметь корней. Например: пусть $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, тогда $\chi_A(\lambda) = \lambda^2 + 1$. У этого многочлена нет вещественных корней. Потому нет собственных значений, а значит и ненулевых собственных векторов. На самом деле, для любого многочлена можно подобрать матрицу так, что он будет ее характеристическим многочленом. Так что это не случайное явление.

Так как собственные значения и вектора хотелось бы иметь, то нам придется в этом вопросе переходить к комплексным числам. и вместо пространства $\mathbb{R}^n$ рассматривать $\mathbb{C}^n$. Тогда над комплексными числами каждый многочлен имеет ровно столько корней (с учетом кратности), какова его степень. Это первое место в линейной алгебре, где появляется разница в том, какие коэффициенты использовать.

## Кто такие комплексные числа

По простому, мы хотим построить множество «чисел», которые бы содержали вещественные числа и на них были определены все нужные операции: сложения, вычитания, умножения и деления на любое ненулевое число. Есть несколько конструкций, я рассмотрю две.

**Классическая конструкция** Рассмотрим множество картинок вида $a+bi$, где $a, b \in \mathbb{R}$, а $i$ – просто символ. Как множество $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$. Теперь на $\mathbb{C}$ определим следующие операции:

1. Сложение: $(a + bi) + (c + di) = (a + c) + (b + d)i$

2. Вычитание: $(a + bi) - (c + di) = (a - c) + (b - d)i$

3. Умножение: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$.

4. Сопряжение: $\overline{a + bi} = a - bi$.

В этом случае нулем будет число вида $0 + 0i$, единицей $1 + 0i$. Если $z = a + bi$, то число $z\bar{z} = a^2 + b^2$ является неотрицательным вещественным числом. Модуль комплексного числа $z$ – это $|z| = \sqrt{z\bar{z}}$. Обратный к числу $z$ имеет вид $\frac{\bar{z}}{|z|^2}$.

Числа вида $a + 0i$ можно отождествить с вещественными числами $a \in \mathbb{R}$. Таким образом $\mathbb{R} \subseteq \mathbb{C}$. Более того, операции определены так, что это вложение с ними согласовано. Обратим внимание на новое число $i = 0 + 1i$. По определению $i^2 = -1$. На самом деле верно следующее.

**Утверждение.** *Для любого многочлена $p(t) = a_0 + a_1 t + \ldots + a_{n-1}t^{n-1} + t^n$, где $a_i \in \mathbb{C}$ существует ровно $n$ комплексных корней с учетом кратности.*

[10]Заметим, что $V_\lambda = \ker(\phi - \lambda\,\mathrm{Id})$.

[11]Для многочлена $p(t)$ число $\lambda$ является корнем тогда и только тогда, когда $p(t) = (t - \lambda)q(t)$. Если $\lambda$ корень для $q(t)$, мы можем еще раз вынести множитель $t - \lambda$ и так далее. В итоге, можно записать $p(t) = (t - \lambda)^k h(t)$, где $h(\lambda) \neq 0$. Такое число $k$ называется кратностью корня $\lambda$.

**Матричная конструкция** Рассмотрим матрицы вида

$$T = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}) \right\}$$

Заметим, что если сложить или перемножить любые две матрицы из $T$, то получим матрицу из $T$. Более того, все матрицы из $T$ кроме нулевой обратимы. Множество $T$ можно отождествить с $\mathbb{C}$, построенным выше, следующим образом: $a+bi \mapsto \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$. То есть $T$ и $\mathbb{C}$ – это одно и тоже. Обратим внимание, что на этом языке сопряжение – это транспонирование, а определитель равен квадрату модуля комплексного числа.[12]

## Собственный базис

**Утверждение.** *Пусть $\phi\colon V \to V$ – линейный оператор и пусть $\lambda_1, \ldots, \lambda_k$ – его разные собственные значения (тут не важно из $\mathbb{R}$ или $\mathbb{C}$) и $v_1, \ldots, v_k \in V$ – соответствующие им ненулевые собственные вектора. Тогда $v_1, \ldots, v_k$ линейно независимы.*

*Доказательство.* Предположим противное, что $a_1 v_1 + \ldots a_k v_k = 0$. Мы можем считать, что все $a_i$ не равны нулю. Это можно записать так

$$a_1 v_1 + \ldots a_k v_k = \begin{pmatrix} a_1 v_1 & \ldots & a_k v_k \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

Применим к этой линейной комбинации $\phi$, получим новую линейную комбинацию

$$a_1 \lambda_1 v_1 + \ldots a_k \lambda_k v_k = \begin{pmatrix} a_1 v_1 & \ldots & a_k v_k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = 0$$

Продолжим применять $\phi$ суммарно $k-1$ раз. В результате имеем

$$\begin{pmatrix} a_1 v_1 & \ldots & a_k v_k \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \ldots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \ldots & \lambda_k^{k-1} \end{pmatrix} = \begin{pmatrix} 0 & \ldots & 0 \end{pmatrix}$$

Но определитель матрицы выше есть определитель вандермонда. Значит, матрица обратима и на нее можно поделить. Значит, все вектора $a_i v_i = 0$. Так как по предположению $v_i \neq 0$ это означает, что $a_i = 0$, противоречие. □

**Утверждение.** *Пусть $\phi\colon V \to V$ – оператор на $n$-мерном пространстве (не важно комплексном или вещественном), при этом его характеристический многочлен имеет $n$ различных корней $\lambda_1, \ldots, \lambda_n$. Тогда соответствующие ненулевые собственные вектора $v_1, \ldots, v_n$ образуют базис $V$ и в этом базисе матрица $\phi$ диагональная с числами $\lambda_i$ на диагонали.*

*Доказательство.* Действительно, для каждого такого $\lambda_i$ обязательно найдется ненулевой собственный вектор. Из предыдущего утверждения все такие собственные вектора линейно независимы, а значит образуют базис. По определению в этом базисе $\phi v_i = \lambda v_i$, т.е. $\phi(v_1, \ldots, v_n) = (v_1, \ldots, v_n) \left(\begin{smallmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{smallmatrix}\right)$. □

На это утверждение можно смотреть так: если есть квадратная матрица $A \in \mathrm{M}_n(\mathbb{R})$ (или $A \in \mathrm{M}_n(\mathbb{C})$) такая, что $\det(A - \lambda E)$ имеет $n$ различных корней, то существует такая невырожденная матрица $C \in \mathrm{M}_n(\mathbb{R})$ (соответственно из $\mathrm{M}_n(\mathbb{C})$), что $C^{-1}AC$ является диагональной и на диагонали стоят корни многочлена $\det(A - \lambda E)$. Комплексный случай хорош лишь тем, что корни обязательно существуют у многочлена, надо лишь чтобы они были различными. В вещественном случае существование корней не гарантировано.

Важно понимать, что если матриц взялась «из жизни» или из «непрерывных случайных данных», то с вероятностью один, характеристический многочлен такой матрицы будет иметь $n$ различных комплексных корней. То есть над комплексными числами любая случайная матрица с вероятностью один превращается в диагональную в некотором базисе.

[12] Первая конструкция обычно рассказывается в школе и потому более привычная. Вторая хороша тем, что нам не надо проверять, что операции ведут себя хорошо, все следует из знаний о матрицах. Плюс это дает некий мостик в правильную линейную алгебру над вещественными числами.

## Поиск собственных значений и векторов

Следующий алгоритм годится как для комплексных так и для вещественных матриц. Разница лишь в том, что в вещественном случае у нас вообще говоря будет меньше собственных значений. Для определенности алгоритм рассказывается для комплексных матриц.

**Дано** Матрица $A \in \mathrm{M}_n(\mathbb{C})$.

**Задача** Найти все собственные значения $\lambda_i$ для $A$ и для каждого $\lambda_i$ найти базис пространства $V_{\lambda_i}$.

**Алгоритм**

1. Посчитать характеристический многочлен $\chi_A(\lambda) = \det(A - \lambda E)$.

2. Найти корни многочлена $\chi(\lambda)$. Корни $\{\lambda_1, \ldots, \lambda_k\}$ будут собственным значениями $A$.

3. Для каждого $\lambda_i$ найти ФСР системы $(A - \lambda_i E)x = 0$. Тогда ФСР будет базисом $V_{\lambda_i}$.

Отметим, что общее количество собственных векторов для всех собственных значений $\lambda_i$ не превосходит $n$ – размерности матрицы, так как $\dim V_{\lambda_i}$ не превосходит кратности корня $\lambda_i$, а сумма кратностей всех корней в точности равна степени многочлена $\chi(\lambda)$, которая есть $n$ – размер матрицы $A$.

Если количество собственных векторов оказалось равно $n$, то матрица $A$ приводится в диагональный вид. Пусть $v_{i1}, \ldots, v_{in_i}$ – собственные вектора с собственным значением $\lambda_i$, при этом $n_i$ будет кратность собственного значения $\lambda_i$. Пусть $C$ – матрица составленная из векторов $v_{ij}$. Пусть $D$ – диагональная матрица с диагональю $(\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2, \ldots, \lambda_k, \ldots, \lambda_k)$, где каждое $\lambda_i$ повторяется $n_i$ раз. Тогда $C^{-1}AC = D$.

## Проверка на диагонализуемость

**Дано** Матрица $A \in \mathrm{M}_n(\mathbb{R})$, задающая линейный оператор $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$.

**Задача** Выяснить существует ли базис, в котором $\varphi$ задается диагональной матрицей и если задается, то какой именно. На матричном языке: существует ли невырожденная матрица $C \in \mathrm{M}_n(\mathbb{R})$ такая, что $C^{-1}AC$ является диагональной и найти эту диагональную матрицу.

**Алгоритм**

1. Найдем характеристический многочлен $\chi(t)$ для $\varphi$, он же для $A$ по формуле $\chi(t) = \det(A - tE)$.

2. Проверим, раскладывается ли $\chi(t)$ на линейные множители, то есть представляется ли он в виде $\chi(t) = (t - \lambda_1)^{d_1} \ldots (t - \lambda_k)^{d_k}$. Если не представляется, то $\varphi$ (или что то же самое $A$) не диагонализируется

3. Если $\chi(t) = (t - \lambda_1)^{d_1} \ldots (t - \lambda_k)^{d_k}$. Найдем для каждого $\lambda_i$ базис $V_{\lambda_i}$ как ФСР системы $(A - \lambda_i E)x = 0$. Если для хотя бы одного $i$ количество элементов в ФСР меньше соответствующей кратности корня $d_i$, то $\varphi$ не диагонализируется.

4. Если для каждого $i$ мы получили, что размер ФСР совпадает с кратностью корня, то есть $\dim V_{\lambda_i} = d_i$. То $\varphi$ диагонализируется и диагональная матрица $C^{-1}AC$ на диагонали содержит числа $\lambda_i$ в количестве $d_i$ штук.

Заметим, что если задача изначально дана для комплексной матрицы $A \in \mathrm{M}_n(\mathbb{C})$, которая задает в этом случае оператор $\varphi\colon \mathbb{C}^n \to \mathbb{C}^n$, то первый шаг алгоритма выполнен автоматически, а именно, над комплексными числами любой многочлен разлагается на линейные множители. Потому над комплексными числами вопрос о диагонализируемости – это лишь проверка всех равенств $\dim V_{\lambda_i} = d_i$.

## Жорданова нормальная форма (ЖНФ)

Самый главный вопрос о линейных операторах: на сколько хорошим можно выбрать базис, чтобы максимально упростить матрицу оператора в этом базисе? В случае «общего положения» как в предыдущем параграфе мы можем диагонализировать матрицу. И это самый популярны в приложениях случай. Но есть и плохие матрицы, которые нельзя диагонализировать. В общем случае ответ будет чуть-чуть сложнее.

Для начала несколько определений. *Жорданова клетка* $J_n(\lambda)$ размера $n$ с числом $\lambda \in \mathbb{C}$ – это матрица вида[13]

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \dots & \lambda \end{pmatrix} \in \mathrm{M}_n(\mathbb{C})$$

Будем говорить, что матрица $A \in \mathrm{M}_n(\mathbb{C})$ имеет *Жорданову нормальную форму*, если она имеет следующий блочный вид

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix} \text{ где } A_i \in \mathrm{M}_{n_i}(\mathbb{R}) \text{ имеет вид } \begin{pmatrix} J_{r_1}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_2}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_s}(\lambda_i) \end{pmatrix}$$

Следующая теорема – это полная классификация линейных операторов на векторном пространстве.

**Утверждение** (Теорема о Жордановой нормальной форме)**.** *Пусть $V$ – комплексное векторное пространство и $\phi\colon V \to V$ – произвольный линейный оператор. Пусть $\lambda_1, \dots, \lambda_k$ – корни его характеристического многочлена с кратностями $n_1, \dots, n_k$. Тогда, существует базис $V$ такой, что матрица $\phi$ имеет следующий вид:*

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix}$$

*где $A_i \in \mathrm{M}_{n_i}(\mathbb{C})$ (размер равен кратности собственного значения). А каждая $A_i$ имеет вид*

$$\begin{pmatrix} J_{r_{i1}}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_{i2}}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_{is_i}}(\lambda_i) \end{pmatrix}$$

*где $\lambda_i$ – соответствующее собственное значение. При этом числа $r_{i1}, \dots, r_{is_i}$ определены однозначно и могут отличаться только порядком.*[14]

**Замечания** Пусть $A \in \mathrm{M}_n(\mathbb{C})$, тогда она обязательно приводится в ЖНФ. Если $\lambda$ – это собственное значение для $A$, тогда оно является корнем характеристического многочлена и пусть его кратность будет $d$.[15]

1. Число $d$ – это суммарный размер жордановых клеток в ЖНФ для $A$ с числом $\lambda$ на диагонали.

2. Число $\dim V_\lambda$ – это количество жордановых клеток в ЖНФ для $A$ с числом $\lambda$ на диагонали.

[13] Можно определить Жорданову клетку и для действительных чисел и для рациональных и вообще каких угодно, но я буду тут обсуждать только комплексный случай.

[14] Существуют алгоритмы нахождения базиса, в котором матрица имеет Жорданову нормальную форму, но мы их изучать не будем.

[15] На самом деле можно дать полный список числовых инвариантов, которые характеризуют ЖНФ для $A$, но это выходит за рамки нашего обсуждения.