Семинар 4

Линейные отображения

Сами по себе векторные пространства не интересны. Нам бы хотелось уметь их сравнивать между собой. Для этого нам нужны линейные отображения. Кроме того, многие вопросы, возникающие в линейной алгебре формулируются именно в терминах линейных отображений.

Определение. Пусть V и U – векторные пространства над $\mathbb{R}$. Отображение $\phi\colon V \to U$ называется линейным, если

1. $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ для любых $v_1, v_2 \in V$.
2. $\phi(rv) = r\phi(v)$ для любых $r \in \mathbb{R}$ и $v \in V$.

Множество всех линейных отображений из V в U обозначается $\operatorname{Hom}(V, U)$. Если надо подчеркнуть какие скаляры имеются в виду, можно написать $\operatorname{Hom}_{\mathbb{R}}(V, U)$.

Примеры

1. Любое линейное отображение $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ задается в виде $\phi(x) = Ax$, где $A \in \mathrm{M}_{m\,n}(\mathbb{R})$. Таким образом множество $\operatorname{Hom}_{\mathbb{R}}(V, U)$ отождествляется с множеством матриц $\mathrm{M}_{m\,n}(\mathbb{R})$.
2. Любое линейное отображение $\phi\colon \mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_m(\mathbb{R})$ задается единственным образом в виде $\phi(X) = \sum_{i,k=1}^{m}\sum_{j,l=1}^{n} a_{ijkl}E_{ij}XE_{lk}$, где E_{ij} – матричная единица, т.е. матрица такая, что на i-ой строке j-ом столбце стоит 1, а все остальные элементы равны 0.

Определение. Если V – векторное пространство и $\phi\colon V \to V$ – линейное отображение, то ϕ называется *линейным оператором.*[1]

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.

Определение. Пусть V и U – векторные пространства и $\phi\colon V \to U$ – линейное отображение. Мы говорим, что оно является *изоморфизмом* (а пространства V и U *изоморфными*), если существует $\psi\colon U \to V$ – линейное отображение такое, что $\phi\psi = \mathrm{Id}$ и $\psi\phi = \mathrm{Id}$.[2]

Про изоморфные пространства надо думать как про одинаковые. На множество V можно смотреть как на «имена векторов», соответственно, U – множество «новых имен», а ϕ – это переименование наших векторов. А раз это всего лишь переименование, то от него ничего не должно зависеть. Потому изучать V – это все равно, что изучать U.

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 ϕ 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. $\square$

[1] Технически линейный оператор ничем не отличается от линейного отображения, просто мы требуем чтобы мы действовали на одном пространстве, а не между двумя. Однако, это порождает огромную разницу в поведении этих объектов и чтобы не путать их между собой люди даже специально ввели для линейных отображений на одном пространстве отдельное название – «оператор».

[2] Другими словами ϕ обратимо, где обратный $\phi^{-1} = \psi$.

Примеры

- В частности этот критерий позволяет отвечать на вопросы следующего вида: существует ли отображение $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$, со следующим свойством

$$\phi\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}, \quad \phi\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix}, \quad \phi\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$$

  В данном случае векторы

$$v_1 = \begin{pmatrix}1\\1\end{pmatrix}, \quad v_2 = \begin{pmatrix}1\\-1\end{pmatrix}$$

  являются базисом, а

$$v_3 = \begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{2}(v_1 + v_2)$$

  По утверждению, векторы v_1 и v_2 можно отправить куда угодно и тогда найдется единственное $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ со свойствами

$$\phi\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}, \quad \phi\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix}$$

  Теперь осталось лишь проверить, удовлетворяет ли наше ϕ последнему свойству. С одной стороны мы хотим, чтобы

$$\phi\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$$

  С другой стороны, как мы выяснили $v_3 = \frac{1}{2}(v_1 + v_2)$. Значит

$$\phi(v_3) = \frac{1}{2}(\phi(v_1) + \phi(v_2)) = \frac{1}{2}\left(\begin{pmatrix}-1\\1\end{pmatrix} + \begin{pmatrix}2\\0\end{pmatrix}\right) = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}$$

  Не сходится. Значит, не существует. Если бы сошлось, то существовал бы. Отметим, что наивный подход заключается в том, чтобы задать отображение ϕ в виде $x \mapsto Ax$, где $A = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$. Тогда условия на ϕ можно переписать как систему линейных уравнений на a, b, c, d. Три вектора, по две координаты, будет всего 6 условий и 4 неизвестные. Это намного неприятнее, чем предложенный выше метод.

- А что если нам даны векторы $v_1, \ldots, v_k$ в $\mathbb{R}^n$ и векторы $u_1, \ldots, u_k$ в $\mathbb{R}^m$ и нас спрашивают существует линейное отображение $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ такое, что $\phi(v_i) = u_i$. Выше я рассказал, как решать задачу, если $\langle v_1, \ldots, v_k\rangle = \mathbb{R}^n$. На самом деле был изложен способ понять, существует ли линейное отображение $\psi\colon \langle v_1, \ldots, v_k\rangle \to \mathbb{R}^m$ такое, что $\psi(v_i) = u_i$. Если такое отображение не существует на подпространстве $\langle v_1, \ldots, v_k\rangle$, то очевидно, что оно не существует на всем пространстве $\mathbb{R}^n$. То есть в негативном случае задача решается проще. Однако, если же отображение $\psi\colon \langle v_1, \ldots, v_k\rangle \to \mathbb{R}^m$ найдется. То искомое ϕ можно построить так. Пусть $v_1, \ldots, v_s$ – базис в $\langle v_1, \ldots, v_k\rangle$. Так как это линейно независимое множество в $\mathbb{R}^n$ его можно дополнить до базиса в $\mathbb{R}^n$. То есть мы можем найти векторы $w_{s+1}, \ldots, w_n \in \mathbb{R}^n$ такие, что $v_1, \ldots, v_s, w_{s+1}, \ldots, w_n$ являются базисом $\mathbb{R}^n$. Отображение ψ отображает $v_1, \ldots, v_s$ в $u_1, \ldots, u_s$. Отправим векторы $w_{s+1}, \ldots, w_n$ в 0. Это нам даст отображение уже на всем пространстве $\mathbb{R}^n$, которое продолжает желаемое отображение и обладает свойством, что v_i идут в u_i. Тут не важно, во что отправить w_j. Как мы видим таких отображений будет много.

- Другой разумный пример использования утверждения. Если пространство V имеет размерность не меньше, чем пространство U, то всегда можно найти сюръективное линейное отображение $\phi\colon V \to U$. Действительно, пусть $e_1, \ldots, e_n$ – базис V и $f_1, \ldots, f_m$ – базис U и $n \geqslant m$. Тогда существует единственное линейное отображение со свойством $e_1 \mapsto f_1, \ldots, e_m \mapsto f_m, e_{m+1} \mapsto 0, \ldots, e_n \mapsto 0$. Как легко видеть такое отображение получается сюръективным.

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 $\mathrm{M}_{m\,n}(\mathbb{R}) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, которое каждой матрице A ставит в соответствие линейное отображение ϕ_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 ϕ_{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

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

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

Действие линейного отображения в координатах Пусть теперь $v \in V$ – некоторый вектор, который раскладывается по базису $v = x_1e_1 + \ldots + x_ne_n = (e_1, \ldots, e_n)x$, где $x \in \mathbb{R}^n$. Тогда

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

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 ϕ должен превратиться в отображение умножения на некоторую матрицу слева. Так вот матрица линейного оператора для ϕ – это в точности та самая матрица, в которую превратился ϕ 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 ϕ и в области куда ϕ бьет. Во втором случае, базисы меняются одновременно.

Утверждение. *Пусть $e_1, \ldots, e_n$ и $e_1', \ldots, e_n'$ – два базиса V, также $f_1, \ldots, f_m$ и $f_1', \ldots, f_m'$ – два базиса U. Пусть*

$$(e_1', \ldots, e_n') = (e_1, \ldots, e_n)C \text{ и } (f_1', \ldots, f_m) = (f_1, \ldots, f_m)D$$

где $C \in \mathrm{M}_n(\mathbb{R})$ и $D \in \mathrm{M}_m(\mathbb{R})$ – матрицы перехода. Если ϕ 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 $\phi(e_1, \ldots, e_n) = (f_1, \ldots, f_m)A$, а надо найти матрицу A' такую, что $\phi(e'_1, \ldots, e'_n) = (f'_1, \ldots, f'_m)A'$. Давайте посчитаем:

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

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

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

Смена базиса в координатах

Пусть теперь $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$, также $e_1, \ldots, e_n$ обозначает стандартный базис в $\mathbb{R}^n$ и $f_1, \ldots, f_m$ – стандартный базис в $\mathbb{R}^m$. Пусть $e'_1, \ldots, e'_n$ – 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.[3] Аналогично, если $f'_1, \ldots, f'_m$ – 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

$$v = x_1e_1 + \ldots + x_ne_n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } x_i$$

С другой стороны, мы можем записать v так

$$v = y_1e'_1 + \ldots + y_ne'_n = C\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } y_i$$

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

$$u = w_1f_1 + \ldots + w_mf_m = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \text{ или } u = z_1f'_1 + \ldots + z_mf'_m = D\begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}$$

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 ϕ в координатах y_i и z_i, то есть записать отображение ϕ 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 ϕ 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. $A' = D^{-1}AC$.

[3] В этом случае мы также имеем $(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)C$. 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.

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* ϕ, а именно: $\ker\phi = \{v \in V \mid \phi(v) = 0\}$.[4] Второе – *образ* ϕ: $\operatorname{Im}\phi = \phi(V) \subseteq U$, то есть все, что можно получить из V, применив к нему ϕ.

Связь со СЛУ Пусть ϕ 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 ϕ лежат все возможные векторы вида Ax. Давайте распишем это так:

  $$\operatorname{Im}\phi = \{Ax \mid x \in \mathbb{R}^n\} = \{x_1A_1 + \ldots + x_nA_n \mid x_i \in \mathbb{R}\} = \langle A_1, \ldots, A_n\rangle$$

  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 ϕ. Тогда это означает, что нам надо решить уравнение $Ax = b$, то есть решение неоднородной системы означает, что мы ищем прообраз к некоторому вектору.
- **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 ϕ. 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 ϕ, 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. *φ сюръективно тогда и только тогда, когда $\operatorname{Im}\varphi = U$.*
2. *φ 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.

$\square$

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 $\operatorname{Im}\phi = U'$.

Линейные операторы

Напоминание В этом разделе я наконец-то вам начну рассказывать о самых важных объектах в линейной алгебре – линейных операторах. Пусть V – векторное пространство, тогда линейным оператором на V называется линейное отображение $\varphi\colon V \to V$, то есть такое отображение, что $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$ и $\varphi(\lambda v) = \lambda\varphi(v)$. Так как линейный оператор – это частный случай линейного отображения, то для него применимо все, о чем мы уже говорили в случае отображений. Про линейный оператор надо думать как про линейную деформацию пространства V.

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

Примеры

1. $\mathrm{Id}\colon V \to V$, $v \mapsto v$. Тождественный линейный оператор, ничего не деформирует.

2. $0\colon V \to V$, $v \mapsto 0$. Нулевой линейный оператор, который все отправляет в ноль.

3. Растяжения вдоль осей: $D\colon \mathbb{R}^n \to \mathbb{R}^n$, заданный по правилу $x \mapsto Dx$, где

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

   Тогда отображение D растягивает i-ю координату в d_i раз. Если $d_i > 1$, то это растяжение, если $0 < d_i < 1$, то это сжатие, если $-1 < d_i < 0$, то это сжатие и отражение вдоль оси, если $d_i < -1$, то это растяжение и отражение вдоль оси.

4. Поворот на плоскости: $\rho_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$, где вектор x поворачивается на угол α против часовой стрелки. Такое отображение в матричном виде задается так[5]

$$\rho_\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5. Поворот в пространстве вокруг оси OX: $\rho_{1,\alpha}\colon \mathbb{R}^3 \to \mathbb{R}^3$, где вектор x поворачивается вокруг оси OX на угол α против часовой стрелки, если смотреть со стороны вектора $e_1 = (1,0,0)^t$ на начало координат. В матричном виде эта штука имеет вид

$$\rho_{1,\alpha} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Теперь давайте объясним некоторую специфику.

Матрица линейного оператора Пусть в векторном пространстве V задан некоторый базис $e_1, \ldots, e_n$ и пусть $\varphi\colon V \to V$ – линейный оператор. Так как у оператора пространство из которого он бьет и то в которое он бьет совпадают, то мы фиксируем всего лишь один базис (пространство-то у нас одно). Тогда по определению матрица линейного оператора φ – это такая матрица $A_\varphi \in \mathrm{M}_n(\mathbb{R})$, что выполнено $\varphi e = eA_\varphi$, где $e = (e_1, \ldots, e_n)$.

Пусть теперь у нас задан другой базис $e_1', \ldots, e_n'$ в пространстве V с матрицей перехода $C \in \mathrm{M}_n(\mathbb{R})$, то есть $(e_1', \ldots, e_n') = (e_1, \ldots, e_n)C$. Пусть так же $e' = (e_1', \ldots, e_n')$. Тогда матрица φ в базисе e' пусть будет A_φ', то есть $\varphi e' = e'A_\varphi'$. В этом случае связь между матрицами следующая $A_\varphi' = C^{-1}A_\varphi C$. То есть матрица A_φ сопряжена матрице A_φ'.[6]

Замечания

- Отметим, что матрица линейного оператора обязательно квадратная. Таким образом, изучение линейного отображения – это изучение прямоугольной матрицы, а изучение линейного оператора – это всегда изучение только квадратной матрицы.

- Если линейное отображение $\psi\colon V \to U$ бьет между двумя разными пространствами одинаковой размерности, то ему тоже соответствует квадратная матрица. Но принципиальная разница с линейным оператором заключается в том, что для линейного отображения мы можем независимо менять базисы в V и U, что соответствует замене $A_\psi' = C^{-1}A_\psi D$, а для линейного оператора, так как пространство одно и то же, базисы меняются одновременно, что соответствует $A_\varphi' = C^{-1}A_\varphi C$.

[5] Строго говоря я еще не рассказывал про то, что такое движение, но этот и следующий пример можно понять и без общей науки, которая будет чуть позже.

[6] Напомню, что квадратные матрицы B и D называются сопряженными, если найдется обратимая матрица C такая, что $D = C^{-1}BC$.

- Так как линейные операторы – это линейные отображения, то задавать их можно так же как и линейные отображения, например: либо с помощью образа базисных векторов, либо с помощью матрицы в фиксированном базисе.

  Например, можно фиксировать векторы $e_1, \ldots, e_n \in V$, которые являются базисом и зафиксировать любой набор векторов $u_1, \ldots, u_n \in V$, тогда существует единственный линейный оператор $\varphi\colon V \to V$ такой, что $\varphi(e_i) = u_i$. Это самый простой и эффективный способ строить новые линейные операторы на пространстве. Но может быть не всегда самый удобный, чтобы потом с такими операторами работать.

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* ϕ как след этой матрицы A. Это определение не зависит от базиса. Действительно, в другом базисе оператор ϕ задается матрицей $A' = C^{-1}AC$, тогда $\operatorname{tr}(A') = \operatorname{tr}(C^{-1}AC) = \operatorname{tr}(A)$. След оператора ϕ 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 ϕ 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 ϕ, как композиция, т.е. $\phi^n(v) = \phi(\ldots\phi(v)\ldots)$ – где композиция берется n раз, например, $\phi^3(v) = \phi(\phi(\phi(v)))$.[7] 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 ϕ^0 как тождественный оператор Id, т.е. $\mathrm{Id}(v) = v$.[8] 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 ϕ будет соответствовать матрице A. В этом случае $p(\phi)$ соответствует матрице $p(A)$.[9]

Основной бонус от подстановки матриц и линейных операторов в многочлены состоит вот в чем. Пусть многочлен $p(t)$ раскладывается в произведение $(t - \lambda_1)\ldots(t - \lambda_n)$. Тогда оператор $p(\phi)$ раскладывается в композицию операторов $(\phi - \lambda_1\,\mathrm{Id})\ldots(\phi - \lambda_n\,\mathrm{Id})$. 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 ϕ соответствует некоторой матрице $A \in \mathrm{M}_n(\mathbb{R})$. Тогда выражение $\det(\phi - \lambda\,\mathrm{Id}) = \det(A - \lambda E)$ является многочленом от λ степени n. Действительно,

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

[7] 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 $\phi^3 v$.

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

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

С усилием вспоминая явную формулу для определителя через перестановки, понимаем, что получается многочлен от λ. 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 ϕ или соответствующей матрицы A (в зависимости от того, о чем идет речь).[10] 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 $\chi(\phi) = 0$.*[11]

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 ϕ соответствует матрице A. Но тогда $E, A, A^2, \ldots, A^{n^2}$ не могут быть линейно независимы. А значит $a_0 + a_1 A + \ldots + a_{n^2} A^{n^2} = 0$, то есть A зануляется многочленом $p(t) = a_0 + a_1 t + \ldots + a_{n^2} t^{n^2}$. 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 ϕ, если $p(\phi) = 0$ и степень многочлена p является минимально возможной. Минимальный многочлен делит любой многочлен зануляющий ϕ, так как остаток тоже должен занулять ϕ 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 ϕ со старшим коэффициентом определен однозначно.*
2. *Многочлен $p(t)$ делит любой многочлен зануляющий ϕ.*
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 $\phi v = \lambda v$.[12] То есть на собственный вектор оператор ϕ действует растяжением. Если $\phi v = \lambda v$ для $v \neq 0$, число λ называется *собственным значением* оператора ϕ. При фиксированном $\lambda \in \mathbb{R}$ множество всех собственных векторов с собственным значением λ, т.е. $\{v \in V \mid \phi v = \lambda v\}$, будем обозначать через V_λ. Все V_λ обязательно будут подпространствами.[13]

Если мы выберем базис в пространстве V, то оно превратится в $\mathbb{R}^n$. Наш оператор ϕ 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 λ 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 ϕ (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 ϕ – это в точности корни характеристического многочлена $\chi(\lambda) = \det(A - \lambda E)$.*

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

[11] 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.

[12] 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.

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

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

3. *Если λ – корень характеристического многочлена, то V_λ – ненулевое подпространство V. Кроме того, $\dim V_\lambda$ не превосходит кратности корня λ у характеристического многочлена.*[14]

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.*

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.[15]

[14] Для многочлена $p(t)$ число λ является корнем тогда и только тогда, когда $p(t) = (t - \lambda)q(t)$. Если λ 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 λ.

[15] 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.

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 $a_1v_1 + \ldots a_kv_k = 0$. Мы можем считать, что все a_i не равны нулю. Это можно записать так

$$a_1v_1 + \ldots a_kv_k = \begin{pmatrix} a_1v_1 & \ldots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

Применим к этой линейной комбинации ϕ, получим новую линейную комбинацию

$$a_1\lambda_1v_1 + \ldots a_k\lambda_kv_k = \begin{pmatrix} a_1v_1 & \ldots & a_kv_k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = 0$$

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

$$\begin{pmatrix} a_1v_1 & \ldots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \ldots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \ldots & \lambda_k^{k-1} \end{pmatrix} = \begin{pmatrix} 0 & \ldots & 0 \end{pmatrix}$$

Но определитель матрицы выше есть определитель вандермонда. Значит, матрица обратима и на нее можно поделить. Значит, все вектора $a_iv_i = 0$. 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 ϕ диагональная с числами λ_i на диагонали.*

Доказательство. Действительно, для каждого такого λ_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.

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 λ_i для A и для каждого λ_i найти базис пространства V_{λ_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 λ_i найти ФСР системы $(A - \lambda_i E)x = 0$. Тогда ФСР будет базисом V_{λ_i}.

Отметим, что общее количество собственных векторов для всех собственных значений λ_i не превосходит n – размерности матрицы, так как $\dim V_{\lambda_i}$ не превосходит кратности корня λ_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 λ_i, при этом n_i будет кратность собственного значения λ_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 λ_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 φ 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 φ, 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 φ (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 λ_i базис V_{λ_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 φ 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 φ диагонализируется и диагональная матрица $C^{-1}AC$ на диагонали содержит числа λ_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 $\dim V_{\lambda_i} = d_i$.

Обратите внимание, что если $A \in \mathrm{M}_n(\mathbb{R})$ – матрица диагонализуемого оператора, то $\operatorname{rk} A$ равен количеству ненулевых диагональных элементов в ее диагональном виде. То есть количеству ненулевых собственных значений с учетом их кратности в характеристическом многочлене.

Признак диагонализуемости Есть очень удобный признак диагонализуемости, который часто помогает. Если у вас есть линейные оператор $\varphi\colon V \to V$, то обычно в начале фиксируется какой-то базис и после этого V превращается в $\mathbb{R}^n$, а оператор φ в оператор умножения на матрицу $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ задан $x \mapsto Ax$, где $A \in \mathrm{M}_n(\mathbb{R})$. В итоге нам надо понять, можно ли выбрать другой базис для φ, чтобы матрицу A заменить на диагональную.

Утверждение. *Пусть $A \in \mathrm{M}_n(\mathbb{R})$ и $g \in \mathbb{R}[x]$ – зануляющий многочлен для A, то есть $g(A) = 0$. Если он раскладывается на линейные множители $g(x) = (x - \lambda_1) \ldots (x - \lambda_k)$ и все λ_i различны, то существует обратимая матрица $C \in \mathrm{M}_n(\mathbb{R})$ такая, что $C^{-1}AC$ будет диагональной и на диагонали будут числа из множества $\{\lambda_1, \ldots, \lambda_k\}$ (но может быть не все из них и часть может повторяться).*

Обратите внимание, что здесь очень важно, чтобы корни λ_i были различными! Если это условие не выполнено, то утверждение не верно. Например, возьмем матрицу $J_n(\lambda)$. У нее минимальный многочлен будет $(x-\lambda)^n$, но про нее можно доказать, что нельзя сопряжением ее сделать диагональной.

Кроме того, в силу теоремы Гамильтона-Кэли, вам достаточно проверить, что характеристический или минимальный многочлены раскладываются на линейные множители и не имеют кратных корней.

Квадратные корни из единицы Это соображение полезно, если вы хотите решать различные матричные уравнения. Например, пусть мы хотим найти все возможные $A \in \mathrm{M}_n(\mathbb{R})$ такие, что $A^2 = E$, то есть хотим найти все квадратные корни из единицы в матрицах. Тогда это означает, что $g(x) = x^2 - 1$ зануляет матрицу A. При этом $g(x) = (x-1)(x+1)$. Это значит, что найдется обратимая матрица C такая, что $C^{-1}AC$ будет диагональной с числами 1 и -1 на диагонали. Кроме того, если у диагональной матрицы 1 и -1 не идут подряд, то мы ее можем сопрячь некоторой матрицей так, что 1 и -1 пойдут подряд. То есть мы показали, что если A является решением уравнения $A^2 = E$, то найдется такая невырожденная матрица C, что

$$C^{-1}AC = \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}, \text{ следовательно } A = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1}$$

Здесь подразумевается, что блоки с единичной и минус единичной матрицей могут быть пустыми (то есть только одни единицы или минус единицы допустимы).

С другой стороны. Легко видеть, что матрицы полученного вида являются решениями данного уравнения

$$A^2 = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}^2 C^{-1} = CEC^{-1} = E$$

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[16]

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

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 ϕ имеет следующий вид:*

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

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

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 λ_i *– соответствующее собственное значение. При этом числа* $r_{i1}, \dots, r_{is_i}$ *определены однозначно и могут отличаться только порядком.*[17]

Замечания Пусть $A \in \mathrm{M}_n(\mathbb{C})$, тогда она обязательно приводится в ЖНФ. Если λ – это собственное значение для A, тогда оно является корнем характеристического многочлена и пусть его кратность будет d.[18]

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

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

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