Линейная алгебра и геометрия

Лекция 10

Тарас Евгеньевич Панов

Механико-математический факультет МГУ, 1 курс, весна 2021 г.
Текст лекций и другие учебные материалы доступны на странице:
http://higeom.math.msu.su/people/taras/teaching

16 марта 2021 г.

2.8. Нильпотентные операторы. Нормальный вид

Определение

Оператор $\mathcal{A}$ называется **нильпотентным**, если $\mathcal{A}^k = \mathcal{O}$ для некоторого натурального k. Минимальное число k, для которого $\mathcal{A}^k = \mathcal{O}$, называется **степенью нильпотентности** оператора $\mathcal{A}$.

Пример

Рассмотрим оператор $\mathcal{A}$, заданный в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ матрицей

$$\begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

(над диагональю стоят единицы, а на остальных местах — нули). Действие этого оператора на базисные векторы описывается схемой $\boldsymbol{e}_n \mapsto \boldsymbol{e}_{n-1} \mapsto \ldots \mapsto \boldsymbol{e}_2 \mapsto \boldsymbol{e}_1 \mapsto \boldsymbol{0}$. Отсюда видно, что $\mathcal{A}^n = \mathcal{O}$, т. е. оператор $\mathcal{A}$ нильпотентен и имеет степень n.

Сформулируем несколько простых свойств нильпотентных операторов.

Предложение

Пусть $\mathcal{A}\colon V \to V$ *— нильпотентный оператор, причём* $\dim V = n$. *Тогда*

а) *единственным собственным значением оператора* $\mathcal{A}$ *является* 0;

б) *оператор* $\mathcal{A}$ *диагонализируем тогда и только тогда, когда* $\mathcal{A} = \mathcal{O}$;

в) $\mathcal{A}^n = \mathcal{O}$, *т. е. степень нильпотентности* $\mathcal{A}$ *не превосходит* $n = \dim V$.

Доказательство

Докажем а). Пусть $\mathcal{A}^k = \mathcal{O}$ и $\mathcal{A}^{k-1} \neq \mathcal{O}$. Значит существует такой вектор $\boldsymbol{v}$, что $\boldsymbol{u} := \mathcal{A}^{k-1}\boldsymbol{v} \neq \mathbf{0}$. Тогда $\mathcal{A}\boldsymbol{u} = \mathcal{A}^k\boldsymbol{v} = \mathbf{0}$, т. е. $\boldsymbol{u}$ — собственный вектор с собственным значением 0.

Если теперь $\lambda \neq 0$ — другое собственное значение, то по определению найдётся $\boldsymbol{w} \neq \mathbf{0}$, такой, что $\mathcal{A}\boldsymbol{w} = \lambda\boldsymbol{w}$. Тогда $\mathbf{0} = \mathcal{A}^k\boldsymbol{w} = \lambda^k\boldsymbol{w}$. Отсюда $0 = \lambda^k$, т. е. $\lambda = 0$ — противоречие.

Предложение

Пусть $\mathcal{A}\colon V \to V$ *— нильпотентный оператор, причём* $\dim V = n$*.*
Тогда

а) *единственным собственным значением оператора* $\mathcal{A}$ *является* 0*;*

б) *оператор* $\mathcal{A}$ *диагонализируем тогда и только тогда, когда* $\mathcal{A} = \mathcal{O}$*;*

в) $\mathcal{A}^n = \mathcal{O}$*, т. е. степень нильпотентности* $\mathcal{A}$ *не превосходит* $n = \dim V$*.*

Доказательство (продолжение).

Докажем б). Если $\mathcal{A}$ диагонализируем, то на диагонали его диагональной матрицы стоят собственные значения, которые все равны нулю в силу а). Следовательно, матрица нулевая и $\mathcal{A} = \mathcal{O}$.

Докажем в). Из утверждения а) вытекает, что характеристический многочлен оператора $\mathcal{A}$ есть $(-t)^n$. Тогда $\mathcal{A}^n = \mathcal{O}$ по теореме Гамильтона–Кэли. $\square$

Следующая теорема показывает, что любой нильпотентный оператор является прямой суммой операторов из предыдущего примера.

Теорема

Пусть $\mathcal{A}\colon V \to V$ — нильпотентный оператор. Тогда в пространстве V существует базис, в котором матрица оператора $\mathcal{A}$ имеет блочно-диагональный вид с блоками из матриц

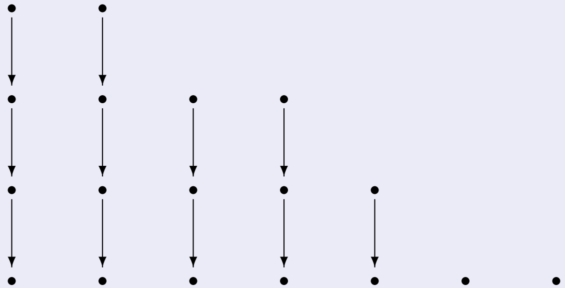
$$\begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

произвольных размеров. Такой вид матрицы оператора единствен с точностью до перестановки блоков.

Базис, существование которого утверждается в этой теореме, называется нормальным, а матрица оператора в таком базисе называется нормальным видом (или нормальной формой) нильпотентного оператора.

Доказательство теоремы

Базис, в котором матрица оператора состоит из блоков указанного вида, удобно изображать в виде диаграммы



В этой диаграмме точки изображают элементы нормального базиса, а стрелки описывают действие оператора $\mathcal{A}$. Элементы нижней строки оператор переводит в нуль, т. е. в ней стоят собственные векторы оператора (с собственным значением 0), входящие в базис. Каждый столбец соответствует одному блоку, причём размер блока равен высоте соответствующего столбца (количеству точек в столбце).

Итак, нам нужно доказать существование базиса, действие оператора $\mathcal{A}$ на элементы которого описывается диаграммой указанного вида.

Проведём индукцию по размерности пространства V. Если $\dim V = 1$, то нильпотентный оператор $\mathcal{A}$ является нулевым, и любой ненулевой вектор в V образует нормальный базис.

Пусть теперь $\dim V = n > 1$, и пусть для размерностей, меньших n, существование нормального базиса уже доказано. Пусть $V_0 = \operatorname{Ker} \mathcal{A}$ — подпространство собственных векторов для $\mathcal{A}$. Так как $\dim V_0 > 0$, имеем $\dim V/V_0 < n$.

Рассмотрим фактор-оператор $\widetilde{\mathcal{A}}\colon V/V_0 \to V/V_0$, $\widetilde{\mathcal{A}}(\boldsymbol{v} + V_0) = \mathcal{A}\boldsymbol{v} + V_0$. По индуктивному предположению $\widetilde{\mathcal{A}}$ имеет нормальный базис. Можно считать его непустым: иначе $V = V_0$ и любой базис в V_0 будет нормальным для $\mathcal{A}$. Построим диаграмму $\widetilde{D}$ для элементов нормального базиса оператора $\widetilde{\mathcal{A}}$, в каждом её столбце возьмём самый верхний вектор $\widetilde{\boldsymbol{e}}_i$, $i = 1, \ldots, m$ (здесь m — количество столбцов в $\widetilde{D}$), и положим $\widetilde{\boldsymbol{e}}_i = \boldsymbol{e}_i + V_0$, $\boldsymbol{e}_i \in V$.

Теперь построим диаграмму D из векторов пространства V следующим образом. Для $i = 1, \ldots, m$ столбец с номером i диаграммы D будет состоять (сверху вниз) из векторов $\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_i, \ldots, \mathcal{A}^{h_i-1}\boldsymbol{e}_i, \mathcal{A}^{h_i}\boldsymbol{e}_i$, где h_i — высота i-го столбца в диаграмме $\widetilde{D}$. Так как $\widetilde{\mathcal{A}}^{h_i}\widetilde{\boldsymbol{e}}_i = \mathbf{0}$, мы имеем $\mathcal{A}^{h_i}\boldsymbol{e}_i \in V_0$ и $\mathcal{A}^{h_i+1}\boldsymbol{e}_i = \mathbf{0}$. Выберем базис в линейной оболочке $\langle \mathcal{A}^{h_1}\boldsymbol{e}_1, \ldots, \mathcal{A}^{h_m}\boldsymbol{e}_m \rangle \subset V_0$, дополним его до базиса V_0 и поставим дополняющие векторы в качестве новых столбцов (высоты один) в нижней строке диаграммы D; оператор $\mathcal{A}$ переводит их нуль.

Таким образом, построенная диаграмма D из векторов пространства V имеет в точности такой вид, как требуется для нормального базиса. Нужно лишь проверить, что векторы, составляющие диаграмму, действительно образуют базис в V.

Сначала покажем, что векторы из D порождают всё V. Пусть $\boldsymbol{v} \in V$. Положим $\widetilde{\boldsymbol{v}} = \boldsymbol{v} + V_0$. По предположению $\widetilde{\boldsymbol{v}} = \sum_{i=1}^{m} \sum_{j=0}^{h_i-1} \lambda_{ij} \widetilde{\mathcal{A}^j} \widetilde{\boldsymbol{e}}_i$. Тогда

$$\boldsymbol{v} - \sum_{i=1}^{m} \sum_{j=0}^{h_i-1} \lambda_{ij} \mathcal{A}^j \boldsymbol{e}_i \in V_0.$$

Но все векторы $\mathcal{A}^j \boldsymbol{e}_i$, $j \leqslant h_i - 1$, лежат в строках диаграммы D, начиная со второй снизу, а подпространство V_0 порождено векторами из нижней строки D по построению. Поэтому $\boldsymbol{v}$ можно представить в виде линейной комбинации векторов из D.

Остаётся проверить линейную независимость векторов из D.

Сначала докажем, что векторы нижней строки линейно независимы. Действительно, если некоторая их нетривиальная линейная комбинация равна нулю, то она должна иметь вид $\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i}\boldsymbol{e}_i = \mathbf{0}$, ибо остальные элементы нижней строки дополняют базис линейной оболочки $\langle \mathcal{A}^{h_1}\boldsymbol{e}_1, \ldots, \mathcal{A}^{h_m}\boldsymbol{e}_m \rangle$ до базиса V_0. Но все $h_i \geqslant 1$, поэтому

$$\mathcal{A}\Bigl(\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i-1}\boldsymbol{e}_i\Bigr) = \mathbf{0},$$

так что

$$\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i-1}\boldsymbol{e}_i \in V_0 \qquad \text{и} \qquad \sum_{i=1}^m \lambda_i \widetilde{\mathcal{A}}^{h_i-1}\widetilde{\boldsymbol{e}}_i = \mathbf{0}.$$

Из последнего соотношения следует, что все $\lambda_i = 0$, так как векторы $\widetilde{\mathcal{A}}^{h_i-1}\widetilde{\boldsymbol{e}}_i$ составляют нижнюю строку диаграммы $\widetilde{D}$ и являются частью базиса пространства V/V_0.

Наконец, покажем, что если имеется любая нетривиальная линейная комбинация векторов D, равная нулю, то из неё можно получить нетривиальную линейную зависимость между векторами нижней строки D. Отметим самую верхнюю строку D, в которой имеются ненулевые коэффициенты этой воображаемой линейной комбинации. Пусть номер этой строки (считая снизу) равен h. Применим к этой комбинации оператор $\mathcal{A}^{h-1}$. При этом её часть, лежащая в h-й строке, перейдёт в нетривиальную линейную комбинацию элементов нижней строки, а остальные слагаемые обратятся в нуль.

Это завершает доказательство существования нормального базиса.

Теперь докажем единственность. Размеры блоков — это высоты столбцов диаграммы. Если расположить столбцы, как на рисунке, в порядке убывания, то их высоты однозначно определяются, если известны длины строк в диаграмме, начиная с нижней, в порядке убывания. Из предыдущего рассуждения следует, что длина нижней строки равна $\dim V_0 = \dim \operatorname{Ker} \mathcal{A}$ и не зависит от выбора базиса. Длина второй снизу строки равна размерности ядра фактор-оператора $\widetilde{\mathcal{A}}$ в пространстве $V / \operatorname{Ker} \mathcal{A}$, т. е. $\dim \operatorname{Ker} \mathcal{A}^2 - \dim \operatorname{Ker} \mathcal{A}$, что также не зависит от выбора базиса. Продолжая далее, мы видим, что длина k-й снизу строки равна размерности ядра фактор-оператора в пространстве $V / \operatorname{Ker} \mathcal{A}^{k-1}$, т. е. $\dim \operatorname{Ker} \mathcal{A}^k - \dim \operatorname{Ker} \mathcal{A}^{k-1}$. Это завершает доказательство единственности. □

Замечание

На практике для нахождения нормального базиса нильпотентного оператора используется следующая модификация процедуры, изложенной в доказательстве теоремы.

Сначала находим степень нильпотентности k оператора $\mathcal{A}$, возводя матрицу в степень, пока не получим 0. Векторы из первой сверху строки диаграммы D соответствуют элементам базиса факторпространства $V/\operatorname{Ker}\mathcal{A}^{k-1}$. Для их нахождения мы выбираем максимальную линейно независимую систему векторов V, не лежащих в $\operatorname{Ker}\mathcal{A}^{k-1}$. Затем мы «спускаем» найденные векторы на одну строку вниз, применяя к ним оператор $\mathcal{A}$. Полученная система векторов лежит в $\operatorname{Ker}\mathcal{A}^{k-1}$, но не лежит в $\operatorname{Ker}\mathcal{A}^{k-2}$, и мы заполняем вторую строку, дополняя эти векторы до базиса в $\operatorname{Ker}\mathcal{A}^{k-1}/\operatorname{Ker}\mathcal{A}^{k-2}$. Затем мы спускаем все векторы из второй строки ещё на одну строку и заполняем третью строку, дополняя векторы, пришедшие из второй строки, до базиса в $\operatorname{Ker}\mathcal{A}^{k-2}/\operatorname{Ker}\mathcal{A}^{k-3}$. И так далее. На последнем шаге мы заполняем нижнюю строку, дополняя векторы, пришедшие сверху, до базиса в $\operatorname{Ker}\mathcal{A}$.