Глава 12. Линейные преобразования линейных пространств

С одной стороны, мы уже рассматривали линейные отображения одного линейного пространства в другое. С другой стороны, напомню, что преобразованием множества называется отображение этого множества в себя.

? Что такое линейное преобразование пространства?

Достаточно часто линейное преобразование называют ***линейным оператором***.

Итак, пусть L – конечномерное линейное пространство над полем F, f – его линейное преобразование. В пространстве L мы будем фиксировать некоторый базис $a_1, a_2, \ldots, a_n$ и рассматривать матрицу преобразования f в этом базисе.

? Какие размеры имеет матрица линейного преобразования f ?

Естественная цель всей главы – научиться так выбрать базис, чтобы матрица в этом базисе имела как можно более простой вид.

К сожалению, идеал, когда матрица линейного преобразования в некотором базисе диагональна, не всегда достижим. По разным причинам. Скажем так, объективным и субъективным: некоторые связаны со свойствами поля F, а некоторые исключительно со свойствами самого преобразования (например, когда поле алгебраически замкнуто), и начнём мы с изучения тех условий, когда идеал всё-таки достижим.

§ 1. Линейные преобразования простой структуры

Определение 1. Линейное преобразование f линейного пространства L над полем F называется ***преобразованием простой структуры***, если в L существует базис, в котором матрица преобразования f диагональна.

! Приведите примеры линейных преобразований простой структуры.

Теорема 1. Линейное преобразование f является преобразованием простой структуры тогда и только тогда, когда в пространстве L есть базис, состоящий из собственных векторов преобразования f.

? Что такое собственный вектор линейного преобразования?

? Что такое собственное число линейного преобразования?

Докажите теорему самостоятельно.

? Что стоит на диагонали матрицы линейного преобразования простой структуры?

? Как находить собственные числа?

Если линейное преобразование является преобразованием простой структуры, то его характеристический многочлен разлагается в произведение линейных множителей. Здесь-то и возникает первая «объективная» трудность: характеристический многочлен может не разлагаться на линейные множители над полем F. Конечно, если поле F алгебраически замкнуто, то такой неприятности произойти не может.

Но пусть у нас линейное преобразование и поле F не «конфликтуют», т.е. характеристический многочлен разлагается на линейные множители. Тогда, конечно, найдутся и собственные векторы для каждого из корней характеристического многочлена. Но всегда ли из них можно собрать базис пространства L? Некоторую помощь в этом оказывает следующая

Теорема 2. Собственные векторы, соответствующие различным собственным числам линейного преобразования f, линейно независимы.

Доказательство. Занумеруем различные собственные числа в произвольном порядке $\alpha_1, \alpha_2, \ldots, \alpha_m$ и докажем утверждение индукцией по m.

Б.И. $m = 1$. Собственный вектор ненулевой, так что он линейно независим.

Ш. И. Пусть для $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ уже доказано, что их собственные векторы $a_1, a_2, \ldots, a_{m-1}$ линейно независимы. Добавим к ним вектор a_m – собственный для числа α_m. Запишем линейную комбинацию этих векторов и приравняем её к 0:

$$\gamma_1 a_1 + \gamma_2 a_2 + \ldots + \gamma_{m-1} a_{m-1} + \gamma_m a_m = 0.$$

Применим к обеим частям равенства преобразование f:

$$\gamma_1 f(a_1) + \gamma_2 f(a_2) + \ldots + \gamma_{m-1} f(a_{m-1}) + \gamma_m f(a_m) = f(0) = 0.$$

Вспоминая, что $a_1, a_2, \ldots, a_m$ – собственные векторы, получаем:

$$\gamma_1 \alpha_1 a_1 + \gamma_2 \alpha_2 a_2 + \ldots + \gamma_{m-1} \alpha_{m-1} a_{m-1} + \gamma_m \alpha_m a_m = 0.$$

Умножив первое равенство на α_m, вычтем его из второго. Получается

$$\gamma_1(\alpha_1 - \alpha_m) a_1 + \gamma_2(\alpha_2 - \alpha_m) a_2 + \ldots + \gamma_{m-1}(\alpha_{m-1} - \alpha_m) a_{m-1} = 0.$$

Это линейная комбинация векторов a_1, a_2, …, a_{m-1}, а они линейно независимы. Значит,

$$\gamma_1(\alpha_1 - \alpha_m) = \gamma_2(\alpha_2 - \alpha_m) = \ldots = \gamma_{m-1}(\alpha_{m-1} - \alpha_m) = 0.$$

Тогда $\gamma_1 = \gamma_2 = \ldots = \gamma_{m-1} = 0$. Исходное равенство $\gamma_1 a_1 + \gamma_2 a_2 + \ldots + \gamma_{m-1} a_{m-1} + \gamma_m a_m = 0$ превращается в равенство $\gamma_m a_m = 0$, откуда следует, что $\gamma_m = 0$. Тем самым, линейная независимость системы векторов a_1, a_2, …, a_m линейно независима. □

Следствие. Если характеристический многочлен линейного преобразования f линейного n-мерного пространства имеет различных корней, то f является преобразованием простой структуры.

Давайте обсудим наши рассмотрения на матричном языке.

? Что означает для матрицы линейного преобразования переход к новому базису?

? Является ли это преобразование преобразованием простой структуры?

§ 2. Инвариантные подпространства линейного преобразования

Мы будем постепенно приближаться к заветному ответу о матрице наиболее простого вида, создавая нужные инструменты и вводя для них соответствующие понятия. Я хочу на примере познакомить вас с кухней решения проблемы. Первым таким понятием будет понятие инвариантного подпространства.

Определение 1. Пусть f – линейное преобразование линейного пространства L. Подпространство M называется ***инвариантным относительно f***, если $f(x) \in M$ для любого x из M.

У каждого пространства и любого линейного преобразования этого пространства всегда есть, по крайней мере, два инвариантных подпространства: все пространство и нулевое подпространство. Подпространства, отличные от этих двух, называются ***собственными***.

Если у преобразования имеется инвариантное подпространство M, то мы можем рассмотреть на нём преобразование g, которое определяется правилом $g(x) = f(x)$. Ясно, что g является линейным преобразованием пространства M. Оно называется ***ограничением*** преобразования f на подпространство M. Обозначают ограничение на подпространство так: $f|_M$.

Влияние наличия собственного инвариантного подпространства у преобразования f демонстрирует следующая

Теорема 1. Линейное преобразование имеет собственное инвариантное подпространство тогда и только тогда, его матрица в некотором базисе является полураспавшейся.

? Какая матрица называется полураспашейся?

Доказательство. Пусть M – собственное инвариантное подпространство преобразования f пространства L. Выберем в M базис $a_1, a_2, \ldots, a_m$ и дополним его, до базиса $a_1, a_2, \ldots, a_m, \ldots, a_n$ всего пространства L. Запишем матрицу преобразования f в этом базисе. Поскольку $f(a_1), f(a_2), \ldots, f(a_m)$ принадлежат M, ненулевыми могут быть только первые m координат этих векторов, а остальные координаты равны 0. Поэтому в матрице преобразования f в левом нижнем углу стоит нулевая матрица, а в левом верхнем – квадратная матрица преобразования f, ограниченного на подпространство M.

Обратно. Пусть в некотором базисе $a_1, a_2, \ldots, a_m, \ldots, a_n$ пространства L матрица преобразования f имеет вид $\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$, где $a_1, a_2, \ldots, a_m$, та часть базиса, которая соответствует левой части матрицы. Пусть M – подпространство, порождённое векторами $a_1, a_2, \ldots, a_m$.

? Что значит «M – подпространство, порождённое векторами a_1, a_2, …, a_m»?

Каждый вектор x из M записывается как $\gamma_1 a_1 + \gamma_2 a_2 + \ldots + \gamma_{m-1} a_{m-1} + \gamma_m a_m$. Поскольку $f(a_1)$, $f(a_2)$, …, $f(a_m)$ принадлежат M, $f(x)$ также принадлежит M, т.е. M – инвариантное относительно f подпространство. □

Напомним, что матрица называется ***распавшейся***, если она представима в виде $\left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right)$, где A и B – квадратные матрицы.

Теорема 2. Линейное преобразование f имеет в некотором базисе распавшуюся матрицу тогда и только тогда, когда пространство L является прямой суммой двух собственных инвариантных подпространств.

Доказательство проведите самостоятельно.

Следствие. Линейное преобразование f имеет в некотором базисе квазидиагональную матрицу тогда и только тогда, когда пространство L является прямой суммой нескольких собственных инвариантных подпространств.

Доказательство проведите самостоятельно.

§ 3. Корневые подпространства линейного преобразования

Пусть α – корень характеристического многочлена линейного преобразования f линейного пространства L.

Определение 1. Линейное преобразование $f-\alpha\varepsilon$ называется ***характеристическим*** для преобразования f.

Лемма 1. Подпространство M инвариантно относительно преобразования f тогда и только тогда, когда оно инвариантно относительно характеристического преобразования.

Доказательство. Пусть M инвариантно относительно f и $x \in M$. Тогда $(f-\alpha\varepsilon)(x) = f(x) - \alpha\varepsilon(x) = f(x) - \alpha x \in M$, т.е. M инвариантно относительно $f - \alpha\varepsilon$.

Обратно. Пусть M инвариантно относительно $f - \alpha\varepsilon$ и $x \in M$. Тогда $f(x) = (f-\alpha\varepsilon)(x) + \alpha x \in M$, т.е. M инвариантно относительно f. □

Для натурального числа n обозначим через $L_n(\alpha) = \mathrm{Ker}\,(f-\alpha\varepsilon)^n$. Как и ядро любого линейного отображения, $L_n(\alpha)$ является подпространством пространства L.

Лемма 2. 1) $L_1(\alpha) \neq 0$.

2) $L_1(\alpha) \subseteq L_2(\alpha) \subseteq \ldots \subseteq L_n(\alpha) \subseteq \ldots$.

3) $L_n(\alpha)$ инвариантно относительно $f - \alpha\varepsilon$ при любом натуральном n.

4) $\mathrm{Im}\,(f - \alpha\varepsilon)|_{L_1(\alpha)} = 0$ и $\mathrm{Im}\,(f - \alpha\varepsilon)|_{L_n(\alpha)} \subseteq L_{n-1}(\alpha)$ при $n > 1$.

Доказательство. 1) Пусть a – собственный вектор для числа α. Тогда $f(a) = \alpha a$, откуда $(f - \alpha\varepsilon)(a) = 0$, т.е. $a \in L_1(\alpha)$.

2) Пусть $x \in L_{n-1}(\alpha)$, т.е. $(f - \alpha\varepsilon)^{n-1}(x) = 0$. Тогда

$$(f - \alpha\varepsilon)^n(x) = (f - \alpha\varepsilon)\,((f - \alpha\varepsilon)^{n-1}(x)) = 0$$

т.е. $x \in L_n(\alpha)$.

3) Для $n = 1$ утверждение очевидно. Пусть $n > 1$ и $x \in L_n(\alpha)$, а $y = (f - \alpha\varepsilon)(x)$. Тогда $(f - \alpha\varepsilon)^{n-1}(y) = (f - \alpha\varepsilon)^n(x) = 0$, т.е. $y \in L_{n-1}(\alpha) \subseteq L_n(\alpha)$.

4) Утверждение $\mathrm{Im}\,(f - \alpha\varepsilon)|_{L_1(\alpha)} = 0$ следует из доказательства п. 1). Утверждение $\mathrm{Im}\,(f - \alpha\varepsilon)|_{L_n(\alpha)} \subseteq L_{n-1}(\alpha)$ при $n > 1$ следует из доказательства п. 3). □

Следствие. $L_n(\alpha)$ инвариантно относительно f для любого α и при любом натуральном n.

Положим $L(\alpha) = \cup\ \{ L_n(\alpha) \mid n \geq 1\}$.

! Докажите, что $L(\alpha)$ при любом α является подпространством, инвариантным относительно преобразования f.

Определение 1. Подпространство $L(\alpha)$ называется ***α-корневым подпространством*** преобразования f.

Теорема. Сумма корневых подпространств, соответствующих различным корням, является прямой.

Доказательство проведём индукцией по числу слагаемых в сумме. Если слагаемое одно, утверждение очевидно. Пусть $\alpha_1, \alpha_2, \ldots, \alpha_m$ – различные собственные числа и уже доказано, что сумма $L(\alpha_1) + L(\alpha_2) + \ldots + L(\alpha_{m-1})$ прямая, т.е.

$$L(\alpha_1) + L(\alpha_2) + \ldots + L(\alpha_{m-1}) = L(\alpha_1) \oplus L(\alpha_2) \oplus \ldots \oplus L(\alpha_{m-1}).$$

Добавим к этой сумме слагаемое $L(\alpha_m)$. Надо проверить, что

$$(L(\alpha_1) \oplus L(\alpha_2) \oplus \ldots \oplus L(\alpha_{m-1})) + L(\alpha_m) = (L(\alpha_1) \oplus L(\alpha_2) \oplus \ldots \oplus L(\alpha_{m-1})) \oplus L(\alpha_m).$$

? Как проверить, что сумма двух подпространств является прямой?

Надо доказать, что их пересечение равно 0.

Пусть $a \in (L(\alpha_1) \oplus L(\alpha_2) \oplus \ldots \oplus L(\alpha_{m-1})) \cap L(\alpha_m)$. С одной стороны, $a \in L(\alpha_m)$, т.е. существует такое натуральное n, для которого $(f - \alpha_m\varepsilon)^n(a) = 0$. С другой стороны, $a = a_1 + a_2 + \ldots + a_{m-1}$, где $a_i \in L(\alpha_i)$. Для каждого a_i есть своё число $n(i)$, для которого $(f - \alpha_i\varepsilon)^{n(i)}(a_i) = 0$. Поэтому

$\prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}(a) =$

$= \prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}(a_1) + \prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}(a_2) + \ldots + \prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}(a_{m-1}) = 0$,

ибо

$$\prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}(a_j) = \left(\prod_{i=1}^{j-1}(f-\alpha_i\varepsilon)^{n(i)}\right)(f-\alpha_j\varepsilon)^{n(j)}\left(\prod_{i=1}^{j-1}(f-\alpha_i\varepsilon)^{n(i)}\right)(a_j) = 0$$

(здесь мы пользуемся леммой 1).

Рассмотрим многочлены $h(x) = (x - \alpha_1)^{n(1)} (x - \alpha_2)^{n(2)} \ldots (x - \alpha_{m-1})^{n(m-1)}$ и $g(x) = (x - \alpha_m)^n$. Эти два многочлена взаимно просты, поэтому существуют такие многочлены $u(x)$ и $v(x)$, для которых $u(x)\, h(x) + v(x)\, g(x) = 1$.

В каждый многочлен подставим вместо переменной x преобразование f.

? Как в этом случае быть с константой??

$h(f) = \prod_{i=1}^{m-1}(f-\alpha_i\varepsilon)^{n(i)}$;

$g(x) = (f - \alpha_m\varepsilon)^n$;

равенство $u(x)\, h(x) + v(x)\, g(x) = 1$ превратится в $u(f)\, h(f) + v(f)\, g(f) = \varepsilon$.

Тогда

$a = \varepsilon(a) = (u(f)\, h(f) + v(f)\, g(f))\, (a) = (u(f)\, h(f))\, (a) + (v(f)\, g(f))\, (a) =$

$= u(f)\, (h(f)\, (a)) + v(f)\, (g(f)\, (a)) = 0 + 0 = 0.$

А это и требовалось доказать. □