§10. Линейные отображения евклидовых пространств

В этом параграфе всюду за исключением алгебраических добавлений из n° 10.3 и n° 10.4 речь идёт про конечномерные евклидовы пространства V над полем $\mathbb{R}$. Евклидово скалярное произведение векторов $u, w \in V$ обозначается через (u, w) или $u \cdot w$.

10.1. Сингулярные числа и сингулярные направления. Напомню, что с каждым линейным отображением $f : U \to W$ между евклидовыми пространствами U, W связано *евклидово сопряжённое* отображение $f^\times : W \to U$, однозначно характеризуемое тем, что $(fu, w) = (u, f^\times w)$ для всех $u \in U$ и $w \in W$. Матрицы $F_{\boldsymbol{wu}}$ и $F^\times_{\boldsymbol{uw}}$ этих отображений в базисах $\boldsymbol{u} = (u_1, \dots, u_n)$ и $\boldsymbol{w} = (w_1, \dots, w_n)$ пространств U и W связаны с матрицами Грама $G_{\boldsymbol{u}} = \boldsymbol{u}^t \cdot \boldsymbol{u}$ и $G_{\boldsymbol{w}} = \boldsymbol{w}^t \cdot \boldsymbol{w}$ этих базисов соотношением

$$F^t_{\boldsymbol{wu}} G_{\boldsymbol{w}} = G_{\boldsymbol{u}} F^\times_{\boldsymbol{uw}}, \tag{10-1}$$

которое получается подстановкой $(fu_1, \dots, fu_n) = \boldsymbol{w} F_{\boldsymbol{wu}}$ и $(f^\times w_1, \dots, f^\times w_n) = \boldsymbol{u} F^\times_{\boldsymbol{uw}}$ в равенство $(fu_1, \dots, fu_n)^t \cdot \boldsymbol{w} = \boldsymbol{u}^t \cdot (f^\times w_1, \dots, f^\times w_n)$, имеющее место по определению сопряжённого оператора. В частности, матрицы евклидово сопряжённых операторов, записанные в произвольных ортонормальных базисах пространств U, W, транспонированы друг другу: $F^\times = F^t$.

Упражнение 10.1. Убедитесь, что сопряжённое к $f : U \to W$ отображение $f^\times : W \to U$ раскладывается в композицию $f^\times = \varepsilon_U^{-1} \circ f^* \circ \varepsilon_W$, где $\varepsilon_U : U \xrightarrow{\sim} U^*$ и $\varepsilon_W : W \xrightarrow{\sim} W^*$ — корреляции[1], задаваемые евклидовыми структурами на U и W, а $f^* : W^* \to U^*$, $\xi \mapsto \xi \circ f$, — двойственное к f отображение.

Поскольку для любой пары линейных отображений $f : U \to V$, $g : V \to W$ и всех $u \in U$ и $w \in W$ выполняются равенства $(gfu, w) = (fu, g^\times w) = (u, f^\times g^\times w)$, сопряжение является антигомоморфизмом по отношению к композиции, т. е. $(gf)^\times = f^\times g^\times$, что согласуется с равенством $(GF)^t = F^t G^t$ между матрицами этих отображений в ортонормальных базисах.

Упражнение 10.2. Покажите, что $\ker f^\times = (\operatorname{im} f)^\perp$.

Лемма 10.1

Для любого линейного отображением $f : U \to W$ между евклидовыми пространствами U, W обе композиции $ff^\times \in \operatorname{End}(W)$, $f^\times f \in \operatorname{End}(U)$ являются самосопряжёнными линейными операторами с неотрицательными собственными числами. Отображение f сюрьективно (соотв. инъективно) если и только если все собственные числа оператора $ff^\times$ (соотв. $f^\times f$) строго положительны.

Доказательство. Самосопряжённость операторов $ff^\times$ и $f^\times f$ очевидна. По лем. 2.5 на стр. 25 все их собственные числа вещественны. Если для некоторого ненулевого вектора $w \in W$ выполняется равенство $ff^\times w = \lambda w$, то $(f^\times w, f^\times w) = (ff^\times w, w) = \lambda \cdot (w, w)$ и либо $w \in \ker f^\times$ и $\lambda = 0$, либо $\lambda = (f^\times w, f^\times w)/(w, w) > 0$. Аналогично, если $f^\times fu = \mu u$ для ненулевого $u \in U$, то либо $\mu = 0$ и $u \in \ker f$, либо $\mu = (fu, fu)/(u, u) > 0$. Поэтому все ненулевые собственные числа обоих операторов положительны. Если отображение f сюрьективно, то по упр. 10.2 $\ker f^\times = (\operatorname{im} f)^\perp = 0$ и, стало быть, все собственные числа оператора $ff^\times$ положительны. Наоборот, если $\operatorname{im} f \neq W$, то $\ker ff^\times \supset \ker f^\times = (\operatorname{im} f)^\perp \neq 0$. Аналогично, если $\ker f = 0$, то все собственные числа оператора $f^\times f$ строго положительны, и наоборот, если $\ker f \neq 0$, то и $\ker f^\times f \supset \ker f \neq 0$. □

[1] См. n° 1.2.2 на стр. 6.

Теорема 10.1

Каждое линейное отображение $f : U \to W$ между евклидовыми пространствами U, W единственным образом раскладывается в композицию $f = \iota_f \circ \delta_f \circ \pi_f$ ортогональной проекции $\pi_f : U \twoheadrightarrow V$ на ортогональное дополнение $V \overset{\text{def}}{=} \ker^{\perp} f$ к ядру $\ker f \subset U$, невырожденного самосопряжённого оператора $\delta_f : V \xrightarrow{\sim} V$ с положительными собственными значениями $\alpha_1, \dots \alpha_r$, где $r = \operatorname{rk} f = \dim \operatorname{im} f$, и изометрического вложения $\iota_f : V \hookrightarrow W$. При этом набор $\alpha_1^2, \dots \alpha_r^2$ квадратов собственных чисел оператора δ_f является набором всех (с учётом кратностей) ненулевых собственных чисел оператора $f^{\times} f : U \to U$.

Доказательство. Согласно предл. 2.2 на стр. 24 в евклидовом пространстве U имеется ортонормальный базис, состоящий из собственных векторов $u_1, \dots, u_n$ самосопряжённого линейного оператора $f^{\times} f : U \to U$, причём все собственные значения этого оператора неотрицательны по лем. 10.1, т. е. $f^{\times} f u_i = \alpha_i^2 u_i$ для некоторых вещественных $\alpha_i \geqslant 0$. Перенумеруем базис так, чтобы $\alpha_i \neq 0$ при $1 \leqslant i \leqslant r$ и $\alpha_i = 0$ при $i > r$. Тогда, как мы видели в доказательстве лем. 10.1, все векторы u_i с $i > r$ лежат в ядре отображения f. Напротив, при $1 \leqslant i, j \leqslant r$ равенства

$$(f u_i, f u_j) = (f^{\times} f u_i, u_j) = \alpha_i^2 (u_i, u_j) = \begin{cases} \alpha_i^2 > 0 & \text{при } i = j \\ 0 & \text{при } i \neq j \end{cases}$$

показывают, что векторы $w_i = f u_i / \alpha_i$ образуют в пространстве W ортонормальную систему. В частности, они линейно независимы. Так как $f(u_j) = 0$ при $j > r$, для любого $u = \sum x_i u_i \in U$ выполняется равенство $f(u) = \alpha_1 x_1 w_1 + \dots + \alpha_r x_r w_r$, т. е. векторы w_i с $1 \leqslant i \leqslant r$ составляют ортонормальный базис в $\operatorname{im} f$, а векторы u_i с $1 \leqslant i \leqslant r$ — ортонормальный базис в ортогональном дополнении V к ядру $\ker f$. Оператор f является композицией изометрического изоморфизма $\iota_f : V \xrightarrow{\sim} \operatorname{im} f$, $u_i \mapsto w_i$, диагонального оператора $\delta_f : V \to V$, $u_i \mapsto \alpha_i u_i$, и ортогональной проекции $\pi_f : U \twoheadrightarrow V$ вдоль $\ker f$.

Если имеется какое-либо ещё разложение $f = \iota \circ \delta \circ \pi_f$, где $\pi_f : U \twoheadrightarrow V$ — ортогональная проекция вдоль $\ker f$, то из предыдущего рассуждения вытекает, что пространство V является прямой ортогональной суммой всех собственных подпространств V_i оператора $f^{\times} f$, отвечающих ненулевым собственным значениям α_i^2 этого оператора, а композиция $\iota\delta : V \xrightarrow{\sim} \operatorname{im} f$ совпадает с ограничением $f|_V$. Поскольку $\delta^{\times} = \delta$ как операторы $V \to V$, а $\iota^{\times} = \iota^{-1}$ как изометрческие операторы $\operatorname{im} f \xrightarrow{\sim} V$, мы заключаем, что $f^{\times} f|_V = \delta^2$. Так как оператор δ^2 диагонализуется в том же самом базисе, что и δ, мы заключаем, что самосопряжённый оператор δ действует на каждом подпространстве V_i умножением на α_i и, тем самым, определяется по f однозначно. А тогда и $\iota = \delta^{-1} \circ f|_V : V \to W$ определяется однозначно. □

Упражнение 10.3. Убедитесь, что оператор $f^{\times} : W \to V$ действует на построенные в доказательстве теор. 10.1 векторы $w_1, \dots, w_r \in W$ по правилу $w_i \mapsto \alpha_i u_i$ и аннулирует ортогональное дополнение к их линейной оболочке. Выведите отсюда, что множества всех (с учётом кратностей) ненулевых собственных чисел у операторов $f^{\times} f$ и $f f^{\times}$ одинаковы.

Определение 10.1 (сингулярные числа и сингулярные направления)

В условиях теор. 10.1 набор из $\dim U$ неотрицательных квадратных корней α_i из собственных значений самосопряжённого оператора $f^{\times} f : U \to U$ называется *набором сингулярных чисел* линейного отображения $f : U \to W$ между евклидовыми пространствами U, W. Ровно $\operatorname{rk} f$ из них строго положительны. Одномерные инвариантные подпространства[1] оператора $f^{\times} f$ называются *сингулярными направлениями* отображения f.

[1]Т. е. одномерные подпространства, порождённые ненулевыми собственными векторами.

Пример 10.1 (этимология эпитета «сингулярный»)
Свяжем с отображением $f : U \to W$ функцию $\varphi : U \smallsetminus 0 \to \mathbb{R}$, $u \mapsto (fu, fu)/(u, u)$. Покажем, что её производная зануляется ровно на собственных направлениях оператора $f^{\times} f$.

Упражнение 10.4. Покажите, что $(u, u)'(v) = 2(u, v)$ и $(fu, fu)'(v) = 2(fu, fv) = 2(f^{\times} fu, v)$.

Согласно правилу дифференцирования дробей, условие $\varphi'(u) = 0$ равносильно тому, что для любого $v \in V$ выполняется равенство $2(f^{\times} fu, v)(u, u) - 2(fu, fu)(u, v) = 0$, означающее, что $f^{\times} fu = u \cdot (fu, fu)/(u, u)$, т. е. что вектор u является собственным для оператора $f^{\times} f$ с собственным значением $(fu, fu)/(u, u) = (f^{\times} fu, u)/(u, u)$.

Следствие 10.1 (полярное разложение)
Каждое биективное линейное преобразование $f \in \mathrm{GL}(V)$ евклидова пространства V допускает единственное разложение $f = gh$, в котором оператор $g \in \mathrm{O}(V)$ ортогонален, а $h \in \mathrm{GL}(V)$ самосопряжён и имеет положительные собственные значения. Квадраты этих собственных значений являются собственными числами оператора $f^{\times} f$.

Доказательство. Поскольку оператор f биективен, правый член его канонического разложения $f = \iota_f \circ \delta_f \circ \pi_f$ из теор. 10.1 является тождественным отображением, а самосопряжённый оператор $h = \delta_f$ не имеет ядра. Следовательно его собственные числа строго положительны. □

Замечание 10.1. (явные формулы для g и h) Компоненты $g \in \mathrm{O}(V)$ и h полярного разложения $f = gh$ однозначно находятся из условий $g^{\times} g = \mathrm{Id}_V$ и $h^{\times} = h$. А именно, $f^{\times} f = h^{\times} g^{\times} gh = h^2$, откуда $h = \sqrt{f^{\times} f}$ и $g = fh^{-1}$. Отметим, что так как нуль не является собственным числом оператора $f^{\times} f$, аналитическая вне нуля функция $\sqrt{t}$ алгебраически вычислима на операторе $f^{\times} f$ при помощи стандартной интерполяционной процедуры, известной из курса алгебры[1].

Упражнение 10.5. Покажите, что каждый невырожденный линейный оператор $f \in \mathrm{GL}(V)$ на евклидовом пространства V также допускает единственное разложение $f = sr$, в котором $r \in \mathrm{O}(V)$, а s самосопряжён и имеет положительные собственные значения, квадраты которых равны собственным числам оператора $ff^{\times}$.

Пример 10.2
Найдём полярное разложение $f = gh$ для оператора $f : \mathbb{R}^3 \to \mathbb{R}^3$ с матрицей

$$F = \begin{pmatrix} 22/15 & -4/3 & 4/15 \\ 4/15 & 2/3 & 28/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix}$$

в стандартном базисе. Поскольку $\det F = -4$, оператор f невырожден. Самосопряжённый оператор $f^{\times} f$ имеет матрицу

$$C = F^t F = \begin{pmatrix} 22/15 & 4/15 & 2/3 \\ -4/3 & 2/3 & 2/3 \\ 4/15 & 28/15 & -1/3 \end{pmatrix} \begin{pmatrix} 22/15 & -4/3 & 4/15 \\ 4/15 & 2/3 & 28/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix} = \begin{pmatrix} 8/3 & -4/3 & 2/3 \\ -4/3 & 8/3 & 2/3 \\ 2/3 & 2/3 & 11/3 \end{pmatrix}$$

у которой след $\mathrm{tr}(C) = 9$, сумма главных 2×2-миноров

$$\det \begin{pmatrix} 8/3 & -4/3 \\ -4/3 & 8/3 \end{pmatrix} = 16/3\,, \quad \det \begin{pmatrix} 8/3 & 2/3 \\ 2/3 & 11/3 \end{pmatrix} = 28/3\,, \quad \det \begin{pmatrix} 8/3 & 2/3 \\ 2/3 & 11/3 \end{pmatrix} = 28/3$$

[1] На всякий случай я напомню её в алгебраическом добавлении из n° 10.4 ниже.

равна 24, определитель $\det(C) = \det^2 F = 16$ и характеристический многочлен

$$\det(tE - C) = t^3 - 9t^2 + 24t - 16 = (t-1)(t-4)^2 .$$

Так как оператор $f^{\times}f$ диагонализуем, он аннулируется многочленом $(t-1)(t-4)$. Следовательно, матрица $H = \sqrt{C}$ самосопряжённого сомножителя h полярного разложения $f = gh$ имеет вид[1] $aE + bC$, где интерполяционный многочлен $p(t) = a + bt$ для вычисления функции $\sqrt{t}$ на матрице C однозначно определяется тем, что $p(1) = \sqrt{1} = 1$ и $p(4) = \sqrt{4} = 2$, т. е. $a + b = 1$ и $a + 4b = 2$, откуда $a = 2/3$, $b = 1/3$. Таким образом, полярное разложение имеет вид $F = GH$, где самосопряжённая матрица $H = \sqrt{C}$ равна

$$\begin{pmatrix} 2/3 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} + \begin{pmatrix} 8/9 & -4/9 & 2/9 \\ -4/9 & 8/9 & 2/9 \\ 2/9 & 2/9 & 11/9 \end{pmatrix} = \begin{pmatrix} 14/9 & -4/9 & 2/9 \\ -4/9 & 14/9 & 2/9 \\ 2/9 & 2/9 & 17/9 \end{pmatrix}$$

а ортогональная матрица $G = FH^{-1}$ равна

$$\begin{pmatrix} 22/15 & -4/3 & 4/15 \\ 4/15 & 2/3 & 28/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 13/18 & 2/9 & -1/9 \\ 2/9 & 13/18 & -1/9 \\ -1/9 & -1/9 & 5/9 \end{pmatrix} = \begin{pmatrix} 11/15 & -2/3 & 2/15 \\ 2/15 & 1/3 & 14/15 \\ 2/3 & 2/3 & -1/3 \end{pmatrix}$$

Упражнение 10.6. Убедитесь, что $G^tG = E$.

Следствие 10.2 (SVD-разложение[2])
Каждая вещественная прямоугольная матрица $F \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ раскладывается в произведение $F = T_m D T_n$, в котором матрицы $T_m \in \mathrm{O}_m$ и $T_n \in \mathrm{O}_n$ ортогональны, а $m\times n$-матрица $D = (d_{ij})$ диагональна и неотрицательна в том смысле, что $d_{ij} = 0$ при $i \neq j$, а все $d_{ii} \geqslant 0$. При этом ровно $\operatorname{rk} F$ диагональных элементов матрицы D отлично от нуля, и они с точностью до перестановки диагональных элементов не зависят от выбора указанного разложения.

Доказательство. Будем воспринимать $F = F_{\boldsymbol{mn}}$ как матрицу оператора $f\colon \mathbb{R}^n \to \mathbb{R}^m$, записанную в стандартных базисах $\boldsymbol{n}$ и $\boldsymbol{m}$ пространств $U = \mathbb{R}^n$ и $W = \mathbb{R}^m$ соответственно. Обозначим через $\boldsymbol{u} = (u_1, \dots, u_n)$ ортонормальный базис пространства U, построенный в доказательстве теор. 10.1, а через $\boldsymbol{w} = (w_1, \dots, w_m)$ — любой ортонормальный базис пространства W, содержащий ортонормальный набор векторов $w_i = fu_i/\alpha_i$, $1 \leqslant i \leqslant r$, из доказательства теор. 10.1. Оператор $f\colon u_i \mapsto \alpha_i w_i$ задаётся в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ диагональной матрицей $D = F_{\boldsymbol{wu}}$, ненулевые диагональные элементы которой суть сингулярные числа $\alpha_1, \dots, \alpha_n$ оператора f. Поэтому $F = F_{\boldsymbol{mn}} = C_{\boldsymbol{mw}} F_{\boldsymbol{wu}} C_{\boldsymbol{un}}$, где $C_{\boldsymbol{mw}}$ — ортогональная матрица перехода от базиса $\boldsymbol{w}$ к стандартному базису $\boldsymbol{m}$ в $\mathbb{R}^m$, а $C_{\boldsymbol{un}} = C_{\boldsymbol{nu}}^{-1} = C_{\boldsymbol{nu}}^{t}$ — ортогональная матрица перехода от стандартного базиса $\boldsymbol{n}$ в $\mathbb{R}^n$ к базису $\boldsymbol{u}$. Для любого другого разложения $F = T_m A T_n$ с ортогональными T_n, T_m и диагональной матрицей A имеем $F^tF = T_n^{-1} A^t A T_n$. Поскольку подобные матрицы имеют одинаковые с точностью до перестановки собственные числа, стоящие на диагонали диагональной матрицы A^tA квадраты диагональных элементов матрицы A суть собственные числа матрицы F^tF. □

[1] По поводу вычисления функций от операторов см. алгебраическое добавление из n° 10.4 на стр. 131 ниже, в частности, ?? на стр. ??.

[2] «SVD» является аббревиатурой от английского *singular values decomposition*.

10.2. Инвариантные углы между подпространствами. Рассмотрим в евклидовом пространстве $\mathbb{R}^k$ пару векторных подпространств U, W размерностей $\dim U = n \leqslant m = \dim W$ и обозначим через $\pi: U \to W$ ортогональную проекцию вдоль $W^{\perp}$. Пусть эта проекция имеет сингулярные числа $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$. Так как $|\pi u| = |u| \cdot \cos \measuredangle(\pi u, u)$ для всех $u \in U$, числа $\alpha_i = \cos \varphi_i$ являются косинусами неубывающих углов

$$0 \leqslant \varphi_1 \leqslant \varphi_2 \leqslant \cdots \leqslant \varphi_n \leqslant \pi/2, \quad \varphi_i = \measuredangle(w_i, u_i), \tag{10-2}$$

между векторами $u_1, u_2, \ldots, u_n$ некоторого ортонормального базиса $\boldsymbol{u}$ в U и первыми n векторами такого ортонормального базиса $\boldsymbol{w} = (w_1, w_2, \ldots, w_m)$ в W, что вектор πu_i пропорционален вектору w_i при $1 \leqslant i \leqslant n$, причём по теор. 10.1 этот набор углов не зависит от выбора ортонормального базиса в U, проектирующегося в ортогональный набор векторов из W. Поэтому углы (10-2) называются *инвариантными углами* между подпространствами U, W.

Мы будем обозначать набор инвариантных углов через $\measuredangle(U, W) \stackrel{\text{def}}{=} (\varphi_1, \varphi_2, \ldots, \varphi_n)$.

Предложение 10.1

Максимальное значение скалярного произведения (u, w) всевозможных пар векторов $u \in U$, $w \in W$ единичной длины $|u| = |w| = 1$ равно максимальному сингулярному числу ортогональной проекции $\pi: U \to W$ вдоль $W^{\perp}$. Минимальный угол $\measuredangle(u, w)$ между ненулевыми векторами $u \in U$, $w \in W$ достигается на сингулярном направлении u_1 проекции π с максимальным коэффициентом растяжения α_1 и его ортогональной проекцией $w_1 = \pi(u_1)/\alpha_1$.

Доказательство. Достаточно доказать второе утверждение, первое является его переформулировкой. Пусть ортонормальные базисы $u_1, \ldots, u_n \in U$ и $w_1, \ldots, w_m \in W$ таковы, что

$$\pi u_i = \alpha_i w_i \quad \text{при } 1 \leqslant i \leqslant n\,.$$

Так как $u_i = w_i + w_i'$ для некоторых векторов $w_1', \ldots, w_n' \in W^{\perp}$, мы имеем при всех $i \neq j$ соотношения ортогональности $(u_i, w_j) = 0$, из которых в силу неравенства Коши – Буняковского – Шварца вытекает, что для любых $u = \sum x_i u_i$ и $w = \sum y_j w_j$ длины $|u| = |w| = 1$

$$\begin{aligned}(u, w) = \sum_{i=1}^{n} x_i y_i (u_i, w_i) = \sum_{i=1}^{n} \alpha_i x_i y_i \leqslant \alpha_1 \sum_{i=1}^{n} x_i y_i \leqslant \\ \leqslant \alpha_1 \sqrt{x_1^2 + \cdots + x_n^2} \sqrt{y_1^2 + \cdots + y_n^2} \leqslant \alpha_1 |u|\,|w| = \alpha_1 = (u_1, w_1)\,.\end{aligned}$$

□

Упражнение 10.7. Выведите существование минимального угла $\measuredangle(u, w)$ между ненулевыми векторами $u \in U$, $w \in W$ из компактности сферы и непрерывности скалярного произведения.

Пример 10.3 (угол между прямой и подпространством)

Если пространство $U = \mathbb{R}\,u$ одномерно, то оно само и является единственным сингулярным направлением ортогональной проекции $\pi: U \to W$. Если евклидова длина $|u| = 1$, то

$$\alpha = \cos \measuredangle(u, W) = \cos \measuredangle(u, u_W) = \sqrt{(u, u_W)}\,, \quad \text{где} \quad u_W = \pi(u)\,,$$

является единственным сингулярным числом. По предл. 10.1 угол $\measuredangle(u, u_W)$ является минимальным среди углов $\measuredangle(u, w)$ по всем ненулевым $w \in W$. Два крайних значения $\measuredangle(u, u_W) = 0$ и

$\measuredangle(u, u_W) = \pi/2$ возникают при $u_W = u$ и $u_W = 0$, т. е. при $u \in W$ и $u \in W^\perp$. Во втором случае вектор u перпендикулярен сразу всем ненулевым векторам из W. Если же угол $\measuredangle(u, u_W)$ ненулевой и острый, то он является абсолютно минимальным, т. е. для всех не пропорциональных u_W векторов $w \in W$ выполняется строгое неравенство $\measuredangle(u, w) > \measuredangle(u, u_W)$.

Упражнение 10.8. Выведите все эти факты из свойств ортогональной проекции[1] без ссылок на предл. 10.1.

Если в пространстве W задан какой-нибудь[2] базис $\boldsymbol{w} = (w_1, \dots, w_m)$, то согласно форм. (1-16) на стр. 11 ортогональная проекция вектора u на W вычисляется по формуле $u_W = \sum_{i=1}^m w_i\,(w_i^\times, u)$, где $\boldsymbol{w}^\times = (w_1^\times, \dots, w_m^\times) = \boldsymbol{w}\,G_{\boldsymbol{w}}^{-1}$ — евклидово двойственный к $\boldsymbol{e}$ базис[3] в W. Поэтому

$$\cos^2 \measuredangle(u, W) = (u_W, u) = \sum_{i=1}^m (w_i^\times, u)(w_i, u) = G_{u\boldsymbol{w}^\times} G_{\boldsymbol{w}u} = G_{u\boldsymbol{w}}\,G_{\boldsymbol{w}}^{-1} G_{\boldsymbol{w}u}\,,$$

где взаимные матрицы Грама $G_{u\boldsymbol{w}}$, $G_{u\boldsymbol{w}^\times}$ и $G_{\boldsymbol{w}u}$ суть две строки и столбец, составленные из скалярных произведений (u, w_j), $(u, w_j^\times)$ и (w_i, u) соответственно.

Альтернативный способ отыскания угла $\measuredangle(u, W)$ заключается в том, что в натянутом на векторы $u, w_1, \dots, w_m$ параллелепипеде длина опущенной из вершины u на грань $w_1, \dots, w_m$ высоты равна, с одной стороны, произведению $|u|\,\sin \measuredangle(u, W)$, а с другой стороны — отношению объёма параллелепипеда к объёму грани, откуда

$$\sin \measuredangle(u, W) = \sqrt{\frac{\det G_{u, w_1, \dots, w_m}}{(u, u)\,\det G_{w_1, \dots, w_m}}}\,,$$

где через $G_{u, w_1, \dots, w_m}$ и $G_{w_1, \dots, w_m}$ обозначены матрицы Грама набора векторов $u, w_1, \dots, w_m$ и набора векторов $w_1, \dots, w_m$ соответственно.

Пример 10.4 (явные формулы в общем случае)
Если в пространствах U и W заданы (не обязательно ортонормальные) базисы

$$\boldsymbol{u} = (u_1, \dots, u_n) = \boldsymbol{e}\,C_{\boldsymbol{eu}} \quad \text{и} \quad \boldsymbol{w} = (w_1, \dots, w_m) = \boldsymbol{e}\,C_{\boldsymbol{ew}}\,,$$

где $\boldsymbol{e}$ — стандартный ортонормальный базис в $\mathbb{R}^n$, то набор ортогональных проекций

$$\boldsymbol{u}_W = (u_{1W}, \dots, u_{nW})$$

базисных векторов пространства U на пространство W выражается через базис $\boldsymbol{w}$ по формуле[4] $\boldsymbol{u}_W = \boldsymbol{w}\left(\boldsymbol{w}^{\times t} \cdot \boldsymbol{u}\right) = \boldsymbol{w}\,G_{\boldsymbol{w}}^{-1} G_{\boldsymbol{wu}}$, где $G_{\boldsymbol{wu}} = C_{\boldsymbol{ew}}^t C_{\boldsymbol{eu}}$ — взаимная матрица Грама[5] наборов $\boldsymbol{w}$ и $\boldsymbol{u}$. Таким образом, проектор $\pi : U \to W$ имеет в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ матрицу

$$\Pi_{\boldsymbol{wu}} = G_{\boldsymbol{w}}^{-1} G_{\boldsymbol{wu}} = G_{\boldsymbol{w}}^{-1} C_{\boldsymbol{ew}}^t C_{\boldsymbol{eu}}\,.$$

[1] См. предл. 1.4 на стр. 11.

[2] Не обязательно ортонормальный.

[3] См. n° 1.3.1 на стр. 8.

[4] Как и выше, точкой обозначается произведение матриц *из векторов*, при вычислении которого векторы перемножаются скалярно. Обратите внимание, что левое произведение в формуле $\boldsymbol{w}\left(\boldsymbol{w}^{\times t} \cdot \boldsymbol{u}\right)$ это произведение матрицы из векторов на *числовую* матрицу $\boldsymbol{w}^{\times t} \cdot \boldsymbol{u}$, и его не следует путать со скалярным произведением матриц из векторов: равенство «$\boldsymbol{w}\left(\boldsymbol{w}^{\times t} \cdot \boldsymbol{u}\right) = \left(\boldsymbol{w} \cdot \boldsymbol{w}^{\times t}\right)\boldsymbol{u}$» категорически неверно!

[5] См. n° 1.2.1 на стр. 5.

Согласно форм. (10-1) на стр. 121 евклидово сопряжённый к нему оператор имеет в тех же базисах матрицу $\Pi^{\times}_{uw} = G_u^{-1}\Pi^t_{wu}G_w = G_u^{-1}G^t_{wu} = G_u^{-1}G_{uw}$. Тем самым, квадраты косинусов инвариантных углов $\measuredangle(U, W)$ суть собственные числа симметричной матрицы

$$\Pi^{\times}_{uw}\Pi_{wu} = G_u^{-1}G_{uw}G_w^{-1}G_{wu} = G_{u^\times w^\times}G_{wu}\,.$$

Пример 10.5 (индуктивное геометрическое описание инвариантных углов)
Ортонормальные базисы в U и W, пригодные для вычисления $\measuredangle(U, W)$, можно получить при помощи следующей индуктивной геометрической процедуры. Сначала выберем произвольный ортонормальный базис $u_1, \dots, u_{i-1}$ в пересечении $U \cap W$ и положим $V_i = (U \cap W)^{\perp}$, $U_i = U \cap V_i$, $W_i = W \cap V_i$. Также положим $w_\nu = u_\nu$ при $\nu \leqslant i - 1$. По предл. 10.1 (или по упр. 10.7) угол $\measuredangle(u, w)$ между переменными векторами $u \in U_i$, $w \in W_i$ единичной длины $|u| = |w| = 1$ достигает своего ненулевого минимума на некоторой паре векторов u_i, w_i. Добавим векторы u_i и w_i в уже имеющиеся базисы $u_1, \dots, u_{i-1}$ и $w_1, \dots, w_{i-1}$, обозначим через $V_{i+1} \subset V_i$ ортогональное дополнение к плоскости, порождённой векторами u_i и w_i, положим $U_{i+1} = U_i \cap V_{i+1}$, $W_{i+1} = W_i \cap V_{i+1}$ и продолжим по индукции.

Предложение 10.2
Пара векторных подпространств U', W' евклидова пространства тогда и только тогда переводится ортогональным линейным преобразованием в пару подпространств U'', W'', когда

$$\dim U' = \dim U''\,, \quad \dim W' = \dim W'' \quad \text{и} \quad \measuredangle(U', W') = \measuredangle(U'', W'') \tag{10-3}$$

Доказательство. Необходимость равенств (10-3) очевидна. Если они выполняются, то

$$\dim(U' + W') = \dim(U'' + W'') \quad \text{и} \quad \dim(U' \cap W') = \dim(U'' \cap W'')\,.$$

Поэтому существует такое ортогональное преобразование g объемлющего евклидова пространства, что $g(U'+W') = U''+W''$, $g(U'\cap W') = U''\cap W''$ и $g(W') = W''$. Тем самым, можно считать, что $W' = W''$, объемлющее евклидово пространство является прямой суммой пространства $W = W' = W''$ и его ортогонального дополнения $W^{\perp}$, а подпространства $U', U'' \subset W \oplus W^{\perp}$ имеют нулевое пересечение с W и размерность $\dim U' = \dim U'' = \dim W^{\perp}$. В этом случае ортогональные проекции подпространств U' и U'' на $W^{\perp}$ вдоль W являются линейными изоморфизмами. Согласно предыдущему, в пространствах U' и U'' имеются ортонормальные базисы из векторов вида $u_i' = w_i' + v_i'$ и $u_i'' = w_i'' + v_i''$, где векторы w_i' и w_i'' составляют части двух ортонормальных базисов пространства W, векторы v_i' и v_i'' образуют два (возможно, не ортонормальных) базиса в $W^{\perp}$, и при всех i и всех $i \neq j$ выполняются соотношения

$$(u_i', w_i') = \alpha_i = (u_i'', w_i'') \quad \text{и} \quad (u_i', w_j') = 0 = (u_i'', w_j'')\,,$$

из которых вытекает, что базисы пространства $W^{\perp}$, состоящие из векторов $v_i' = u_i' - w_i'$ и из векторов $v_i'' = u_i'' - w_i''$ оба ортогональны и имеют одинаковые скалярные квадраты базисных векторов $(v_i', v_i') = (v_i'', v_i'') = 2 - 2\alpha_i$. Поэтому линейное преобразование пространства $W \oplus W^{\perp}$, переводящее w_i' в w_i'', а v_i' — в v_i'' является ортогональным. Оно переводит подпространство U' в подпространство U''. □

10.3. Алгебраическое дополнение I: аннулирующие многочлены. Напомню, что линейный оператор $F : V \to V$, действующий в векторном пространстве V над произвольным полем $\Bbbk$, можно подставить вместо переменной t в любой многочлен $f(t) = a_0 + a_1 t + \dots + a_m t^m \in \Bbbk[t]$. Результатом такой подстановки по определению является линейный оператор

$$f(F) = a_0 \mathrm{Id}_V + a_1 F + a_2 F^2 + \dots + a_m F^m \in \mathrm{End}(V)\,.$$

Подставляя фиксированный оператор $F \in \mathrm{End}\,V$ во всевозможные многочлены $f \in \Bbbk[t]$, мы получаем гомоморфизм $\Bbbk$-алгебр $\mathrm{ev}_F : \Bbbk[t] \to \mathrm{End}(V)$, $f \mapsto f(F)$, который называется *гомоморфизмом вычисления* многочленов на операторе F. Многочлены f, лежащие в ядре этого гомоморфизма, т. е. такие, что $f(F) = 0$, называются *аннулирующими* оператор F. Если $\dim V < \infty$, алгебра $\mathrm{End}\,V$ конечномерна как векторное пространство над $\Bbbk$, а алгебра $\Bbbk[t]$ бесконечномерна. Поэтому $\ker \mathrm{ev}_F \neq 0$, т. е. любой оператор на конечномерном пространстве аннулируется некоторым ненулевым многочленом. Такой многочлен минимальной степени со старшим коэффициентом 1 обозначается μ_F и называется *минимальным многочленом* оператора F. Поскольку все идеалы кольца $\Bbbk[t]$ главные, ядро $\ker \mathrm{ev}_F = (\mu_F)$ состоит из всех многочленов, делящихся на μ_F. Следующий пример показывает, что $\deg \mu_F \leqslant \dim V$.

Пример 10.6 (тождество Гамильнотна – Кэли)
Обозначим через $K = \mathbb{Z}[a_{ij}]$ кольцо многочленов с целыми коэффициентами от n^2 коммутирующих переменных a_{ij}, а через $K[t]$ — кольцо многочленов с коэффициентами из K от ещё одной переменной t, коммутирующей со всеми переменными a_{ij}. Кольцо $n \times n$ матриц $\mathrm{Mat}_n(K[t])$ с элементами из кольца многочленов $K[t]$ совпадает с кольцом многочленов $\mathrm{Mat}_n(K)[t]$ с коэффициентами в кольце матриц $\mathrm{Mat}_n(K)$, поскольку каждую матрицу, в клетках которой стоят многочлены от t, можно записать как многочлен от t с матричными коэффициентами и наоборот. Например,

$$\begin{pmatrix} 3t^2 + 2t & t^3 - 1 \\ 2t + 3 & t^3 + t - 1 \end{pmatrix} = t^3 \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + t^2 \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} + t \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 3 & -1 \end{pmatrix}.$$

Обозначим через $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ матрицу из элементов a_{ij}. В кольце $\mathrm{Mat}_n(K[t])$ выполняется равенство[1]

$$\det(tE - A) \cdot E = (tE - A)(tE - A)^\vee\,, \tag{10-4}$$

в котором E — единичная матрица размера $n \times n$, а $(tE - A)^\vee$ — присоединённая к $(tE - A)$ матрица, транспонированная к матрице, составленной из алгебраических дополнений к элементам матрицы $tE - A$. Многочлен $\chi_A(t) = \det(tE - A) = t^n - \mathrm{tr}(A) t^{n-1} + \dots + (-1)^n \det A \in K[t]$ называется *характеристическим* многочленом матрицы A. Переписывая (10-4) в виде равенства между многочленами от t с коэффициентами в кольце матриц $\mathrm{Mat}_n(K)$:

$$t^n E - \mathrm{tr}(A)\, t^{n-1} E + \dots + (-1)^n \det A\, E = (tE - A)\left(t^m A^\vee_m + \dots + t A^\vee_1 + A^\vee_0\right)\,,$$

где $A^\vee_i \in \mathrm{Mat}_n(K)$ — некоторые матрицы, и подставляя в него $t = A$, получаем в кольце $\mathrm{Mat}_n(K)$ равенство $\chi_A(A) \cdot E = 0$, откуда $\chi_A(A) = 0$. Таким образом, каждый линейный оператор F аннулируется своим характеристическим многочленом $\chi_F(t) = \det(tE - F)$.

[1] Это частный случай соотношения Лапласа из форм. (3-13) на стр. 32, возникающий при $m = 1$, когда $\Lambda^m A = A$, а $\Lambda^{n-m} A = A^\vee = \widehat{A}^t$ — матрица, транспонированная к матрице, состоящей из алгебраических дополнений к элементам матрицы A.

Теорема 10.2 (теорема о разложении)
Пусть линейный оператор $F : V \to V$ на произвольном[1] векторном пространстве V над любым полем $\Bbbk$ аннулируется многочленом $q \in \Bbbk[t]$, который раскладывается над полем $\Bbbk$ в произведение $q = q_1 \cdot q_2 \cdot \dots \cdot q_r$ попарно взаимно простых многочленов $q_i \in \Bbbk[t]$. Положим $Q_j = q/q_j$. Тогда $\ker q_j(F) = \operatorname{im} Q_j(F)$ для каждого j, все эти подпространства F-инвариантны, и пространство V является прямой суммой тех из них, что отличны от нуля.

Доказательство. Так как $q(F) = q_i(F) \circ Q_j(F) = 0$, имеем включение $\operatorname{im} Q_i(F) \subset \ker q_i(F)$. Достаточно показать, что V линейно порождается образами операторов $Q_i(F)$, и что сумма ядер $\ker q_i(F)$ прямая, т. е. $\ker q_i(F) \cap \sum_{j\neq i} \ker q_j(F) = 0$ для всех i. Первое вытекает из того, что н.о.д.$(Q_1, \dots, Q_r) = 1$, а значит, существуют такие $h_1, \dots, h_r \in \Bbbk[t]$, что $1 = \sum Q_j(t)h_j(t)$. Подставляя в это равенство $t = F$ и применяя обе части к произвольному вектору $v \in V$, получаем разложение $v = Ev = \sum Q_j(F)h_j(F)v \in \sum \operatorname{im} Q_j(F)$. Второе вытекает из взаимной простоты q_i и Q_i, в силу которой существуют такие многочлены g и h, что $1 = g(t) \cdot q_i(t) + h(t) \cdot Q_i(t)$. Подставим сюда $t = F$ и применим обе части полученного равенства $E = g(F)\, q_i(F) + h(F) \circ Q_i(F)$ к произвольному вектору $v \in \ker q_i(F) \cap \sum_{j\neq i} \ker q_j$. Так как $\ker q_j(F) \subset \ker Q_i(F)$ при всех $j \neq i$, получим $v = Ev = g(F)\, q_i(F)\, v + h(F)\, Q_i(F)\, v = 0$, что и требовалось. □

Пример 10.7 (проекторы)
Линейный оператор $\pi : V \to V$ называется *идемпотентом* или *проектором*, если он аннулируется многочленом $t^2 - t = t(t-1)$, т. е. удовлетворяет соотношению $\pi^2 = \pi$. По теор. 10.2 образ любого идемпотента $\pi : V \to V$ совпадает с подпространством его неподвижных векторов: $\operatorname{im} \pi = \ker(\pi - \mathrm{Id}_V) = \{v \mid \pi(v) = v\}$, и всё пространство распадается в прямую сумму $V = \ker \pi \oplus \operatorname{im} \pi$. Тем самым, оператор π проектирует V на $\operatorname{im} \pi$ вдоль $\ker \pi$. Отметим, что оператор $\mathrm{Id}_V - \pi$ тоже является идемпотентом и проектирует V на $\ker \pi$ вдоль $\operatorname{im} \pi$. Таким образом, задание прямого разложения $V = U \oplus W$ равносильно заданию пары идемпотентных эндоморфизмов $\pi_1 = \pi_1^2$ и $\pi_2 = \pi_2^2$ пространства V, связанных соотношениями $\pi_1 + \pi_2 = 1$ и $\pi_1\pi_2 = \pi_2\pi_1 = 0$.

Упражнение 10.9. Выведите из этих соотношений, что $\ker \pi_1 = \operatorname{im} \pi_2$ и $\operatorname{im} \pi_1 = \ker \pi_2$.

Пример 10.8 (инволюции)
Линейный оператор $\sigma : V \to V$ называется *инволюцией*, если он аннулируется многочленом $t^2 - 1$, т. е. удовлетворяет соотношению $\sigma^2 = \mathrm{Id}_V$. Тождественная инволюция $\sigma = \mathrm{Id}_V$ называется *тривиальной*. Если $\operatorname{char} \Bbbk \neq 2$, то $t^2 - 1 = (t+1)(t-1) = 0$ является произведением различных линейных множителей. Поэтому над таким полем каждое пространство V с инволюцией σ распадается в прямую сумму $V = V_+ \oplus V_-$ собственных подпространств $V_+ = \ker(\sigma - E)$ и $V_- = \ker(\sigma + E)$ с собственными значениями $+1$ и -1 соответственно. Произвольный вектор $v = v_+ + v_-$ пространства V имеет в этом разложении компоненты

$$v_+ = \frac{v + \sigma v}{2} \in \operatorname{im}(\sigma + \mathrm{Id}_V) = V_+ \quad \text{и} \quad v_- = \frac{v - \sigma v}{2} \in \operatorname{im}(\sigma - \mathrm{Id}_V) = V_- \,.$$

Отметим, в частности, что все инволюции диагонализуемы.

[1] Возможно даже бесконечномерном.

Следствие 10.3 (критерий диагонализуемости)
Линейный оператор $F : V \to V$ на векторном пространстве V над произвольным полем $\Bbbk$ диагонализуем если и только если он аннулируется многочленом, полностью раскладывающимся над полем $\Bbbk$ в произведение попарно различных линейных множителей.

Доказательство. Если F диагонализуем, то $V = \bigoplus_\lambda \ker(F - \lambda E)$, где λ без повторений пробегает все различные собственные числа оператора F. Поэтому произведение $\prod_\lambda (F - \lambda E)$ действует на V нулём, т. е. многочлен $\prod_\lambda (t - \lambda E)$ аннулирует F. Наоборот, если оператор F аннулируется произведением попарно различных двучленов $(t - \lambda)$, то по теор. 10.2 $V = \bigoplus_\lambda \ker(F - \lambda E)$ является прямой суммой собственных подпространств оератора F. □

Следствие 10.4
Если оператор $F : V \to V$ диагонализуем, то его ограничение на любое F-инвариантное подпространство тоже диагонализуемо на этом подпространстве. □

10.3.1. Перестановочные операторы. Если линейные операторы $F, G : V \to V$ на векторном пространстве V над произвольным полем $\Bbbk$ коммутируют друг с другом, то ядро и образ любого многочлена от оператора F переводятся оператором G в себя, поскольку

$$\begin{aligned} f(F)\,v = 0 \quad &\Rightarrow \quad f(F)\,G\,v = G\,f(F)\,v = 0 \\ v = f(F)\,w \quad &\Rightarrow \quad Gv = G\,f(F)\,w = f(F)\,Gw\,. \end{aligned}$$

В частности, все собственные подпространства $V_\lambda = \ker(F - \lambda E)$ и все корневые подпространства $K_\lambda = \ker(\lambda\,\mathrm{Id} - F)^{m_\lambda}$, где m_λ — кратность собственного числа λ оператора F, инвариантны относительно любого перестановочного с F оператора G.

Предложение 10.3
Любое множество коммутирующих операторов на конечномерном векторном пространстве обладает над алгебраически замкнутым полем общим собственным вектором. Если каждый из операторов диагонализуем (над произвольным полем), то их можно одновременно диагонализовать в одном общем для всех операторов базисе.

Доказательство. Индукция по размерности пространства V, на котором действуют операторы. Если все операторы скалярны (что так при $\dim V = 1$), то доказывать нечего — подойдут, соответственно, любой ненулевой вектор и любой базис. Если среди операторов есть хоть один нескалярный оператор F, то над замкнутым полем у него есть собственное подпространство строго меньшей размерности, чем V, а в диагонализуемом случае V является прямой суммой таких собственных подпространств. Каждое собственное подпространство оператора F инвариантно для всех операторов, причём если операторы диагонализуемы на всём пространстве, то их ограничения на собственные подпространства оператора F останутся диагонализуемы по сл. 10.4. Применяя к собственным подпространствам оператора F предположение индукции, получаем требуемое. □

Пример 10.9 (конечные группы операторов)
Если m линейных операторов на конечномерном пространстве V над алгебраически замкнутым полем $\Bbbk$ характеристики $\operatorname{char}\Bbbk > m$ образуют группу G, то каждый из этих операторов аннулируется многочленом $t^m - 1$, который раскладывается в произведение m попарно различных линейных множителей. Поэтому каждый оператор в группе G диагонализуем. Все операторы из группы G одновременно диагонализуются в одном общем базисе если и только если группа G абелева.

10.4. Алгебраическое дополнение II: функции от операторов. В этом разделе мы рассматриваем фиксированный линейный оператор $F : V \to V$, действующий в векторном пространстве V над полем $\mathbb{R}$ или $\mathbb{C}$, которое мы обозначаем через $\mathbb{K}$. Всюду далее мы предполагаем, что оператор F аннулируется некоторым многочленом $\alpha(t) \in \mathbb{K}[t]$, который полностью разлагается над полем $\mathbb{K}$ на линейные множители, т. е.

$$\alpha(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2} \cdots (t-\lambda_s)^{m_s}, \tag{10-5}$$

где $\lambda_i \neq \lambda_j$ при $i \neq j$ и все $m_i \in \mathbb{N}$. Мы полагаем $m = \deg\alpha = m_1 + \dots + m_s$. Алгебра $\mathcal{A}$ над полем $\mathbb{K}$, состоящая из функций $U \to \mathbb{K}$ от переменной t, пробегающей какое-нибудь открытое подмножество $U \subset \mathbb{K}$, содержащее все корни λ аннулирующего оператор F многочлена (10-5), называется *алгебраически вычислимой* на операторе F, если она содержит алгебру многочленов $\mathbb{K}[t]$ и для каждого корня λ кратности k аннулирующего многочлена (10-5) все функции $f \in \mathcal{A}$ определены в точке λ вместе с первыми $k-1$ производными $f^{(\nu)} = \frac{d^\nu}{dt^\nu} f$ и допускают в этой точке разложение Тейлора – Лагранжа вида

$$f(t) = f(\lambda) + \frac{f'(\lambda)}{1!}(t-\lambda) + \dots + \frac{f^{(k-1)}(\lambda)}{(k-1)!}(t-\lambda)^{k-1} + g_\lambda(t)\cdot(t-\lambda)^k, \tag{10-6}$$

где функция $g_\lambda(t)$ тоже лежит в алгебре $\mathcal{A}$.

Например, если оператор F действует на конечномерном векторном пространстве, и все его собственные числа лежат в поле $\mathbb{K}$, то в качестве аннулирующего многочлена α можно взять характеристический многочлен $\alpha(t) = \chi_F(t)$ оператора F, и алгебра $\mathcal{A}$ всех функций $U \to \mathbb{K}$, определённых на каком-нибудь подмножестве $U \subset \mathbb{K}$, содержащем некоторую ε-окрестность каждого собственного числа λ, и представимых в каждой из этих ε-окрестностей суммой абсолютно сходящегося степенного ряда от $(t-\lambda)$, является алгебраически вычислимой на операторе F. Подалгебра этой алгебры, состоящая из всех аналитических функций[1] $\mathbb{K} \to \mathbb{K}$ алгебраически вычислима на любом линейном операторе $F \in \operatorname{End}(V)$, все собственные числа которого лежат в $\mathbb{K}$.

Теорема 10.3

В сделанных выше предположениях каждая алгебраически вычислимая на операторе $F : V \to V$ алгебра функций $\mathcal{A}$ допускает единственный такой гомоморфизм $\mathbb{K}$-алгебр $\operatorname{ev}_F : \mathcal{A} \to \operatorname{End} V$, что $\operatorname{ev}_F(p) = p(F)$ для всех многочленов $p \in \mathbb{K}[t] \subset \mathcal{A}$.

Гомоморфизм $\operatorname{ev}_F : \mathcal{A} \to \operatorname{End} V$ из теор. 10.3 называется *вычислением* функций $f \in \mathcal{A}$ на операторе F. Линейный оператор $\operatorname{ev}_F(f) : V \to V$, в который переходит функция $f \in \mathcal{A}$ при гомоморфизме вычисления, обозначается $f(F)$ и называется *функцией f от оператора F*.

Например, для любого оператора $F \in \operatorname{End}(V)$ с собственными числами из $\mathbb{K}$ определены такие аналитические функции, как e^F или $\sin F$, а для любого оператора $F \in \operatorname{GL}(V)$ с собственными числами из $\mathbb{K}$ — аналитические вне нуля функции, например, $\ln F$ или $\sqrt{F}$, причём алгебраические свойства этих функций в алгебре $\operatorname{End} V$ точно такие же, как у соответствующих числовых функций e^t, $\sin t$, $\ln t$ и $\sqrt{t}$. В частности, все эти функции от оператора F коммутируют друг с другом и с F, а также подчиняются привычным алгебраическим тождествам вроде $\ln F^2 = 2\ln F$ и $\sqrt{F}\sqrt{F} = F$.

[1] Т. е. функций, задаваемых сходящимися всюду в $\mathbb{K}$ степенными рядами.

10.4.1. Интерполяционный многочлен. Если оператор F аннулируется многочленом m-той степени, то все операторы F^k с $k \geqslant m$ являются линейными комбинациями m операторов

$$F^0 = E\,,\ F\,,\ F^2\,,\ \dots\,,\ F^{m-1}\,. \tag{10-7}$$

Поэтому любой оператор, который можно получить из E и F при помощи взятия линейных комбинаций с коэффициентами из $\mathbb{K}$, композиций и переходов к пределу, находится в $\mathbb{K}$-линейной оболочке операторов (10-7), т. е. *является многочленом от F, степени не выше, чем* $m-1$.

Многочлен от переменной t, значение которого на операторе F равно значению $f(F)$ заданной функции $f \in \mathcal{A}$ на операторе F обозначается через $p_{f(F)}(t)$ и называется *интерполяционным многочленом* для вычисления $f(F)$.

Упражнение 10.10. Покажите, что класс многочлена $p_{f(F)}$ в фактор кольце $\mathbb{K}[t]/(\mu_F)$, где μ_F — минимальный многочлен оператора F, однозначно определяется оператором F и функцией $f \in \mathcal{A}$.

Так как оператор F аннулируется многочленом $\alpha(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\dots(t-\lambda_s)^{m_s}$, где $\lambda_i \neq \lambda_j$ при $i \neq j$, пространство V является по теореме о разложении[1] прямой суммой F-инвариантных *корневых подпространств* $K_{\lambda_i} = \ker(F-\lambda_i E)^{m_i}$ этого многочлена. Согласно формуле (10-6) оператор

$$f(F) = f(\lambda)\cdot E + f'(\lambda)\cdot(F-\lambda E) + \dots + \frac{f^{(m_\lambda-1)}(\lambda)}{(m_\lambda-1)!}(F-\lambda E)^{m_\lambda-1} + g_\lambda(F)(F-\lambda E)^{m_\lambda} \tag{10-8}$$

действует на каждом корневом подпространстве K_λ, отвечающем k-кратному корню λ, точно так же, как результат подстановки оператора F в многочлен

$$j_\lambda^{k-1}f(t) = f(\lambda) + f'(\lambda)\cdot(t-\lambda) + \dots + f^{(k-1)}(\lambda)\cdot(t-\lambda)^{k-1}/(k-1)!\ ,$$

класс которого в фактор кольце $\mathbb{K}[t]\big/\big((t-\lambda)^k\big)$ называется $(k-1)$-*струёй* функции $f \in \mathcal{A}$ в точке $\lambda \in \mathbb{K}$. Таким образом, в качестве интерполяционного многочлена для вычисления $f(F)$ подойдёт любой многочлен, имеющий в каждом корне λ кратности k аннулирующего оператор F многочлена $\alpha(t)$ такую же $(k-1)$-струю, как и функция $f(t) \in \mathcal{A}$.

Лемма 10.2

Для любого поля $\Bbbk$, любого набора из s попарно различных чисел $\lambda_1, \dots \lambda_s \in \Bbbk$, любого набора кратностей $m_1, \dots, m_s \in \mathbb{N}$ и произвольного набора из $m = m_1 + \dots + m_s$ значений $\beta_{ij} \in \Bbbk$, где $1 \leqslant i \leqslant s$, $0 \leqslant j \leqslant m_i - 1$, существует единственный такой многочлен p степени не выше, чем $m-1$, что $p^{(j)}(\lambda_i) = \beta_{ij}$ при всех i, j из указанного диапазона, где $p^{(j)} = \frac{d^j p}{dt^j}$ обозначает j-тую производную, и для единообразия обозначений мы полагаем $p^{(0)} \overset{\text{def}}{=} p$.

Доказательство. Отображение $S : \Bbbk[t]_{\leqslant(m-1)} \to \Bbbk^m$ из m-мерного векторного пространства многочленов степени не выше, чем $m-1$, в m-мерное координатное пространство, сопоставляющее многочлену p набор из m значений $p^{(j)}(\lambda_i)$, линейно и инъективно, поскольку для любого ненулевого многочлена, лежащего в его ядре, каждое число λ_i является корнем кратности не менее m_i. Стало быть, отображение S биективно. □

Доказательство теор. 10.3. Если требуемый гомоморфизм $\mathrm{ev}_F : \mathcal{A} \to \mathbb{K}$ существует, то согласно предыдущему для каждой функции $f \in \mathcal{A}$ оператор $f(F)$ действует на пространстве V как единственный многочлен $p_{f(F)}$ степени $\deg p_{f(F)} < \deg \alpha$, имеющий в каждом корне λ кратности k

[1] См. теор. 10.2 на стр. 129.

многочлена α ту же самую $(k-1)$-струю, что и функция f. Тем самым, оператор $f(F) = p_{f(F)}(F)$ определяется по функции f однозначно. Остаётся проверить, что отображение $f \mapsto p_{f(F)}(F)$ является гомоморфизмом $\mathbb{K}$-алгебр. Проверим сначала, что отображение

$$\begin{aligned} J : \mathcal{A} \to \frac{\mathbb{K}[t]}{\left((t-\lambda_1)^{m_1}\right)} \times \dots \times \frac{\mathbb{K}[t]}{\left((t-\lambda_s)^{m_s}\right)} \simeq \frac{\mathbb{K}[t]}{(\alpha)} \\ f \mapsto \left(j^{m_1-1}_{\lambda_1} f, \dots, s^{m_s-1}_{\lambda_s} f\right), \end{aligned} \tag{10-9}$$

сопоставляющее функции $f \in \mathcal{A}$ набор её струй[1] во всех корнях аннулирующего оператор F многочлена α, является гомоморфизмом $\mathbb{K}$-алгебр, т. е. $\mathbb{K}$-линейно и удовлетворяет равенству $J(fg) = J(f)J(g)$. Первое очевидно, второе достаточно проверить для каждой струи j^{m-1}_λ отдельно: по правилу Лейбница $(fg)^{(k)} = \sum_{\nu+\mu=k} \binom{k}{\nu} f^{(\nu)} g^{(\mu)}$ получаем по модулю $(t-\lambda)^m$:

$$\begin{aligned} j^{m-1}_\lambda(fg) = \sum_{k \leqslant m-1} \frac{(t-\lambda)^k}{k!} \sum_{\nu+\mu=k} \frac{k!}{\nu!\,\mu!} f^{(\nu)}(\lambda) g^{(\mu)}(\lambda) = \\ = \sum_{k \leqslant m-1} \sum_{\nu+\mu=k} \frac{f^{(\nu)}(\lambda)}{\nu!}(t-\lambda)^\nu \cdot \frac{g^{(\mu)}(\lambda)}{\mu!}(t-\lambda)^\mu \equiv j^{m-1}_\lambda(f) j^{m-1}_\lambda(g). \end{aligned}$$

Отображение $f \mapsto P_{f(F)}(F)$ является композицией гомоморфизма (10-9) с гомоморфизмом вычисления многочленов $\mathrm{ev}_F : \mathbb{K}[t] \to \operatorname{End} V$, $p \mapsto p(F)$, который корректно пропускается через фактор $\mathbb{K}[t]/(\alpha)$, так как $\alpha(F) = 0$. □

Пример 10.10 (степенная функция и рекуррентные уравнения)
Задача отыскания n-того члена a_n числовой последовательности $z : \mathbb{Z} \to \mathbb{K}$, $n \mapsto z_n$, решающей рекуррентное уравнение $z_n = \alpha_1 z_{n-1} + \alpha_2 z_{n-2} + \dots + \alpha_m z_{n-m}$ с начальным условием $(z_0, \dots, z_{n-1}) = (a_0, \dots, a_{n-1}) \in \mathbb{K}^n$, сводится вычислению n-той степени *матрицы сдвига*

$$S = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_m \\ 1 & 0 & \ddots & \vdots & \alpha_{m-1} \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & \alpha_2 \\ 0 & \cdots & 0 & 1 & \alpha_1 \end{pmatrix}$$

смещающей каждый фрагмент из m последовательных элементов на один шаг вправо:

$$(z_{k+1}, z_{k+2}, \dots, z_{k+m}) \cdot S = (z_{k+2}, z_{k+3}, \dots, z_{k+m+1}),$$

так что член a_n оказывается равным первой координате вектора

$$(a_n, a_{n+1}, \dots, a_{n+m-1}) = (a_0, a_1, \dots, a_{m-1}) \cdot S^n.$$

Согласно сказанному выше, $S^n = p_{S^n}(S)$ является результатом подстановки матрицы S в интерполяционный многочлен $p_{S^n}(t) \in \mathbb{K}[t]$ для вычисления на матрице S *степенной функции*

[1] Мы рассматриваем этот набор как элемент прямого произведения соответствующих колец вычетов, которое по китайской теореме об остатках изоморфно фактор кольцу $\mathbb{K}[t]/(\alpha)$.

$f(t) = t^n$. Степень этого многочлена меньше m, и его коэффициенты находятся решением системы из $m+1$ линейных уравнений на $m+1$ неизвестных. Например, для *последовательности Фиббоначчи*, решающей уравнение $a_n = a_{n-1} + a_{n-2}$ с начальным условием $(a_0, a_1) = (0, 1)$, матрица сдвига

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{имеет} \quad S^n = a\,S + b\,E = \begin{pmatrix} b & a \\ a & a+b \end{pmatrix},$$

поскольку интерполяционный многочлен $p_{S^n}(t) = at + b$ линеен. Тем самым,

$$\left(a_n, a_{n+1}\right) = (0, 1) \cdot S^n = (a, a+b)\,.$$

Характеристический многочлен $\chi_S(t) = t^2 - t\,\mathrm{tr}\,S + \det S = t^2 - t - 1 = (t - \lambda_+)(t - \lambda_-)$ имеет однократные корни $\lambda_\pm = (1 \pm \sqrt{5})\,/\,2$, и функция t^n принимает на них значения $\lambda_\pm^n$. Коэффициенты a, b интерполяционного многочлена находятся из уравнений

$$\begin{cases} a\,\lambda_+ + b = \lambda_+^n \\ a\,\lambda_- + b = \lambda_-^n \end{cases}$$

и первый из них $a = \left(\lambda_+^n - \lambda_-^n\right)/\left(\lambda_+ - \lambda_-\right)$, откуда

$$a_n = a = \frac{\left((1+\sqrt{5})/2\right)^n - \left((1-\sqrt{5})/2\right)^n}{\sqrt{5}}\,.$$

Ответы и указания к некоторым упражнениям

Упр. 10.1. Согласно определениям правых корреляций и двойственного отображения, для всех $u \in U$ и $w \in W$ имеем: $(u, \varepsilon_U^{-1} f^* \varepsilon_W w) = f^* \varepsilon_W w(u) = \varepsilon_W w(fu) = (fu, w)$.

Упр. 10.2. Это вытекает из равенства $(u, f^\times w) = (fu, w)$, выполненного для всех $u \in U$, $w \in W$.

Упр. 10.4. $(u+v, u+v) = (u, u)+2(u, v)+o(|u|)$, $(f(u+v), f(u+v)) = (fu, fu)+2(fu, fv)+o(|u|)$.

Упр. 10.5. Так как оператор $ff^\times$ самосопряжён и биективен, все его собственные числа строго положительны. Поэтому имеется единственный такой самосопряжённый оператор s с положительными собственными значениями, что[1] $s^2 = ff^\times$. Тогда $f = sr$, где $r = s^{-1}f$ ортогонален, поскольку $r^\times r = f^\times s^{-2} f = f^x (ff^\times)^{-1} f = \mathrm{Id}_V$.

Упр. 10.7. Поскольку произведение двух компактов компактно, а функция arccos непрерывна, функция $\varphi = \arccos(u, w)$ достигает минимума на декартовом произведении единичных сфер в U_i и W_i.

Упр. 10.8. Пусть $|u| = 1$. Для любого вектора $w \in W$ длины $|w| = 1$ согласно предл. 1.4 на стр. 11 выполняется равенство $\cos \measuredangle(u, w) = (u, w) = (u_W, w)$. В силу неравенства Коши – Буняковского – Шварца $(u_W, w) \leqslant |u_W| = \cos \measuredangle(u, W)$, где равенство достигается если и только если $u_W = \lambda w$ для некоторого $\lambda \geqslant 0$, т. е. либо когда $\lambda = u_W = 0$, и в этом случае оно имеет место для всех w, либо когда $u_W \neq 0$ и w сонаправлен с u_W.

Упр. 10.10. Пусть многочлены $g, h \in \mathbb{K}[t]$ таковы, что $g(F) = h(F) = f(F)$. Тогда многочлен $g - h$ аннулирует F и, стало быть, делится на минимальный многочлен оператора F.

[1] Так как s и s^2 диагонализуются в одном базисе, оператор s обязан действовать на каждом собственном подпространстве V_λ оператора s^2 умножением на положительный $\sqrt{\lambda}$.