Matematický ústav Slezské univerzity v Opavě
Učební texty k přednášce ALGEBRA II, letní semestr 2000/2001
Michal Marvan

18. První rozklad lineární transformace

Úmluva. V této přednášce V je vektorový prostor (obvykle konečněrozměrný) nad polem P a $f : V \longrightarrow V$ je lineární transformace.

První (primární) rozklad lineární transformace f je indukován rozkladem anulujícího (např. charakteristického) polynomu na nesoudělné součinitele. Geometricky jde o rozložení prostoru V na přímý součet tzv. invariantních podprostorů, s čímž je spojeno uvedení matice transformace f do blokově diagonálního tvaru. V případě $f : P^n \to P^n$, $f(u) = Au$, kde A je čtvercová matice, dostáváme jako výsledek blokově diagonální matici podobnou matici A.

1. Invariantní podprostory

Je-li $U \subseteq V$ podprostor, pak symbolem fU označujeme jeho obraz při zobrazení f, to jest, podprostor $\{ f(u) \mid u \in U \}$.

Definice. Podprostor $U \subseteq V$ se nazývá *invariantní* (vzhledem k lineární transformaci f), když $fU \subseteq U$, tj. když pro každé $u \in U$ je $f(u) \in U$.

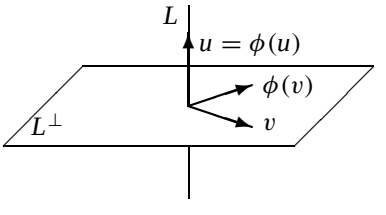
Je-li U invariantní podprostor, pak zobrazení $U \longrightarrow U$, zadané předpisem $u \mapsto f(u)$, nazýváme *restrikce* (česky *ohraničení*) lineárního zobrazení f na invariantní podprostor U. Značí se $f|_U : U \to U$ a je zřejmě opět lineární (ověřte).

Příklad. (1) Celý prostor V a nulový podprostor jsou invariantní podprostory.

(2) Je-li $f : v \mapsto cv$, pak je každý podprostor invariantní.

(3) Je-li u vlastní vektor s vlastní hodnotou c, pak $[\![u]\!]$ je invariantní podprostor a $f|_{[\![u]\!]}$ je zobrazení $v \mapsto cv$.

(4) Uvažujme o rotaci ϕ v prostoru E^3 kolem osy L procházející počátkem 0 o úhel $\alpha \in (0, 2\pi)$. Invariantní podporostory jsou nulový podprostor $\{0\}$, osa rotace L, její ortogonální doplněk $L^\perp$ a celý prostor E^3. Libovolný vektor $u \in L$ se zobrazí sám na sebe, proto $\phi|_L$ je identické zobrazení id_L. Libovolný vektor $v \in L^\perp$ zůstane v rovině $L^\perp$ a $\phi|_{L^\perp}$ je otáčení roviny $L^\perp$ o úhel α.



(5) Uvažujme o zrcadlení ζ v prostoru E^3 vzhledem k rovině U procházející počátkem 0. Invariantní podporostory jsou nulový podprostor $\{0\}$, rovina U a každý její podprostor $V \subseteq U$, ortogonální doplněk $U^\perp$ a celý prostor E^3. Zobrazení $\zeta|_V$ je identické zobrazení id_V. Zobrazení $\zeta|_{U^\perp}$ je zrcadlení přímky $U^\perp$ vzhledem k počátku 0.

Cvičení. (1) Jednorozměrný podprostor $[\![u]\!]$, $u \neq 0$, je invariantní právě tehdy, když u je vlastní vektor. Dokažte.

(2) Průnik a součet invariantních podprostorů jsou invariantní podprostory. Dokažte.

(3) Ker f je invariantní podprostor. Dokažte. Co je $f|_{\mathrm{Ker}\, f}$?
(4) Im f je invariantní podprostor. Dokažte.
(5) Buď $v \in V$ libovolný vektor. Dokažte, že $[\![v,\ f(v),\ f(f(v)),\ f(f(f(v))), \ldots]\!]$ je invariantní podprostor.
(6) Nechť lineární transformace $f, g : V \to V$ komutují, to jest, $f \circ g = g \circ f$. Buď $U \subseteq V$ invariantní podprostor vzhledem k transformaci f. Pak je gU též invariantní podprostor vzhledem k transformaci f.

Rozklad prostoru V na přímý součet invariantních podprostorů vede ke zjednodušení matice transformace f.

Tvrzení. *Nechť existuje přímý rozklad $V = U_1 \dotplus U_2$, kde U_1, U_2 jsou invariantní podprostory. Zvolme nějakou bázi $e_1, \ldots, e_m$ v podprostoru U_1 a nějakou bázi $e_{m+1}, \ldots, e_n$ v podprostoru U_2. Pak $e_1, \ldots, e_n$ je báze v prostoru V a transformace f v ní má matici tvaru*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,m+1} & \cdots & a_{nn} \end{pmatrix}.$$

Označíme-li

$$A_1 = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \vdots & & \vdots \\ a_{n,m+1} & \cdots & a_{nn} \end{pmatrix}.$$

pak A_i je matice lineární transformace $f|_{U_i}$.

Důkaz. Jako cvičení ověřte, že $e_1, \ldots, e_n$ je báze v prostoru V.

Ohledně matice A víme, že prvních m jejích sloupců je tvořeno souřadnicemi vektorů $f(e_1), \ldots, f(e_m)$ v bázi $e_1, \ldots, e_n$. Vektory $f(e_1), \ldots, f(e_m)$ ovšem leží v podprostoru U_1 s bazí $e_1, \ldots, e_m$, takže zbývající bázové vektory $e_{m+1}, \ldots, e_n$ budou mít nulové koeficienty. Submatice A_1 je pak maticí zobrazení $f|_U$ v bázi $e_1, \ldots, e_m$ (ověřte).

Zbytek analogicky.

O shora uvedené matici A říkáme, že je v blokově diagonálním tvaru s bloky A_1, A_2 na diagonále. Stručně zapisujeme

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

Říkáme též, že A je *přímý součet submatic* A_1 a A_2.

Podobně se v případě přímého součtu $V = U_1 \dot{+} \cdots \dot{+} U_n$ invariantních podprostorů U_1, $\ldots$, U_n, matice zobrazení f rozpadá na přímý součet submatic $A_1, \ldots A_n$ odpovídajících lineárním zobrazením $f|_{U_1}, \ldots, f|_{U_n}$:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix},$$

O matici A pak rovněž pravíme, že je blokově diagonální.

V ideálním případě lze prostor V rozložit na přímý součet jednorozměrných invariantních podprostorů, generovaných vlastními vektory, což je nám již známý případ diagonalizovatelné matice.

2. Algebraická struktura na množině lineárních zobrazení

Na množině

$$\mathrm{Hom}_P(V, V) = \{\, f : V \to V \mid f \text{ je lineární zobrazení nad polem } P \,\}$$

existuje bohatá algebraická struktura.

Především, $\mathrm{Hom}_P(V, V)$ je monoid vzhledem k binární operaci „$\circ$“ skládání transformací a s neutrálním prvkem id_V (ověřte). Avšak $\mathrm{Hom}_P(V, V)$, a obecněji $\mathrm{Hom}_P(V, V')$, má též strukturu vektorového prostoru.

Definice. Buďte $f, g : V \to V'$ lineární zobrazení vektorových prostorů nad polem P.

(1) Zobrazení $f + g : V \to V'$, zadané předpisem $u \mapsto f(u) + g(u)$ pro libovolný vektor $u \in V$, se nazývá *součet* lineárních zobrazení f a g.

(2) Je-li $c \in P$ skalár, pak zobrazení $cf : V \to V'$, zadané předpisem $u \mapsto c \cdot f(u)$ pro libovolný vektor $u \in V$, se nazývá *c-násobek* zobrazení f.

Platí tedy

$$(f + g)(u) = f(u) + g(u),$$
$$(cf)(u) = c \cdot f(u)$$

pro každé $u \in V$.

Tvrzení. *Buďte $f, g : V \to V'$ lineární zobrazení vektorových prostorů V, V' nad polem P, buď $c \in P$ skalár. Pak jsou zobrazení $f + g$, cf lineární.*

Důkaz. Cvičení.

Algebraická struktura na množině $\mathrm{Hom}_P(V, V)$ tedy zahrnuje binární operace sčítání $+$ a skládání $\circ$, operace násobení skalárem a dva neutrální prvky: 0 pro sčítání a id pro skládání.

Tvrzení. *Buď V vektorový prostor nad polem P. Pak* $\mathrm{Hom}_P(V,V)$ *je monoid a současně vektorový prostor nad polem P a platí*

$$f\circ(g+h) = f\circ g + f\circ h,$$
$$f\circ(cg) = c(f\circ g),$$
$$(f+g)\circ h = f\circ g + g\circ h,$$
$$(cf)\circ g = c(f\circ g).$$

pro libovolná $f,g,h\in\mathrm{Hom}_P(V,V)$ a $c\in P$.

Důkaz. Cvičení.

Algebraická struktura, která je současně monoid i vektorový prostor nad polem P a platí pro ni identity uvedené v předchozím tvrzení, se nazývá *asociativní P-algebra*. Tudíž, $\mathrm{Hom}_P(V,V)$ je asociativní P-algebra.

Jiný příklad asociativní P-algebry dává množina $\mathrm{gl}(n,P)$ všech čtvercových matic typu n/n nad polem P vzhledem k binárním operacím násobení a sčítání matic a k operaci násobení skalárem (ověřte).

Tvrzení. *Buď V konečněrozměrný vektorový prostor nad polem P, buď $e_1,\dots,e_n$ jeho báze. Buďte f, g lineární transformace $V\to V$, Buďte A, B jejich matice vzhledem k bázi $e_1,\dots,e_n$, buď c skalár. Pak platí*

(a) *$A+B$ je matice součtu transformací $f+g$.*
(b) *cA je matice lineární transformace cf,*
(c) *$A\cdot B$ je matice lineární transformace $f\circ g$.*

Důkaz. Cvičení.

Podle uvedeného tvrzení je v konečněrozměrném případě P-algebra $\mathrm{Hom}_P(V,V)$ izomorfní P-algebře $\mathrm{gl}(n,P)$.

Pro libovolné přirozené číslo k ještě zaveďme lineární transformaci $f^k : V\to V$ předpisem $f^k(v) = \underbrace{f(f(\cdots f}_{k}(v)\cdots))$ pro libovolný vektor $v\in V$, tj. $f^k = \underbrace{f\circ\cdots\circ f}_{k}$.

Definice. Buď $f : V\to V$ lineární transformace vektorového prostoru V nad polem P. Buď $p = a_m x^m + \cdots + a_1 x + a_0 \in P[x]$ polynom. Položme $p(f) = a_m f^m + \cdots + a_1 f + a_0\,\mathrm{id}$. Říkáme, že lineární zobrazení $p(f)$ vzniklo dosazením lineárního zobrazení f do polynomu p. Hodnota takového zobrazení na vektoru $v\in V$ se zapisuje $p(f)(v)$.

Příklad. Nechť $p = x^2 - 2x + 2$. Pak pro libovolnou lineární transformaci $f : V\to V$ máme $p(f) = f^2 - 2f + 2\,\mathrm{id}$ a pro libovolný vektor $v\in V$ máme $p(f)(v) = f(f(v)) - 2f(v) + 2v$.

Tvrzení. *Buď A matice lineární transformace f vzhledem k nějaké bázi $e_1,\dots,e_n$ prostoru V. Položme $p(A) = a_m A^m + \cdots + a_1 A + a_0 E$, kde E je jednotková matice stejného rozměru jako matice A. Pak je $p(A)$ maticí lineárního zobrazení $p(f)$ vzhledem k bázi $e_1,\dots,e_n$.*

Důkaz. Cvičení.

Tvrzení. *Jsou-li $p, q \in P[x]$ dva polynomy, pak platí*

$$(p+q)(f) = p(f) + q(f), \qquad (pq)(f) = p(f) \circ q(f),$$
$$(p+q)(A) = p(A) + q(A), \qquad (pq)(A) = p(A)\, q(A)$$

pro libovolnou lineární transformaci f resp. libovolnou čtvercovou matici A.

Důkaz. Cvičení.

Důsledek. *Pro libovolnou lineární transformaci f a polynomy p, q máme*

$$p(f) \circ q(f) = q(f) \circ p(f)$$

(říkáme, že $p(f)$ a $q(f)$ komutují). Podobně pro libovolnou čtvercovou matici A máme

$$p(A)\, q(A) = q(A)\, p(A),$$

tj. matice získané dosazením A do různých polynomů též komutují.

3. Anulující polynom a první rozklad

Definice. Nechť $p \in P[x]$, $p \neq 0$. Řekneme, že p je *anulující polynom* čtvercové matice A, jestliže $p(A) = 0$. Podobně, p je anulující polynom lineární transformace f, jestliže $p(f) = 0$.

Později uvidíme, že všechny čtvercové matice i všechny lineární transformace konečněrozměrného vektorového prostoru V mají anulující polynom. Nyní se budeme zabývat příslušným rozkladem prostoru V. Odpovídá rozkladu anulujícího polynomu na ireducibilní činitele.

Tvrzení. *Buď q anulující polynom lineární transformace $f : V \to V$. Nechť existuje rozklad $q = q_1 q_2 \cdots q_n$, kde polynomy $q_1, \ldots, q_n$ jsou po dvou nesoudělné ($D(q_i, q_j) = 1$ pro $i \neq j$). Uvažujme o lineárním zobrazení $q_i(f) : V \to V$, označme $U_i = \operatorname{Ker} q_i(f)$, $i = 1, \ldots, n$. Pak platí:*
(i) *Každý podprostor U_i je invariantní;*
(ii) $V = U_1 \dotplus \cdots \dotplus U_n$;
(iii) *polynom q_i je anulujícím polynomem transformace $f|_{U_i}$, pro každé $i = 1, \ldots, n$.*

Důkaz. (i) Nechť $u \in U_i$, tj. $q_i(f)(u) = 0$. Počítejme:

$$q_i(f)\big(f(u)\big) = \big(q_i(f) \circ f\big)(u) = \big(f \circ q_i(f)\big)(u) = f\big(q_i(f)(u)\big) = f(0) = 0$$

(použili jsme tvrzení, že f a $q_i(f)$ spolu komutují). Tudíž, $f(u) \in U_i$.

(ii) Nejdříve případ $n = 2$. Nechť tedy $f = q_1 q_2$ a $D(q_1, q_2) = 1$. Pak existují polynomy p_1, p_2 takové, že $1 = q_1 p_1 + q_2 p_2$. Dosazením zobrazení f získáme rovnost

$$\mathrm{id} = q_1(f) \circ p_1(f) + q_2(f) \circ p_2(f),$$

takže pro libovolný vektor $v \in V$ platí

$$v = q_1(f)\big(p_1(f)(v)\big) + q_2(f)\big(p_2(f)(v)\big). \qquad (*)$$

Ukažme, že první sčítanec $q_1(f)\big(p_1(f)(v)\big)$ z $(*)$ leží v U_2. Označíme-li $w = p_1(f)(v)$, stačí ověřit, že $w \in \operatorname{Ker} q_2(f)$:

$$q_2\big(q_1(f)(w)\big) = (q_2 \circ q_1)(f)(w) = q(f)(w) = 0,$$

protože q je anulující polynom pro f. Podobně se ukáže, že druhý ze sčítanců leží v U_1. Tudíž, $v \in U_1 + U_2$. Protože v byl libovolný vektor z V, máme $V = U_1 + U_2$.

Ukažme ještě, že $U_1 \cap U_2 = 0$. Nechť tedy $v \in U_1 \cap U_2$, tj. $q_1(f)(v) = 0$ a $q_2(f)(v) = 0$. Rovnost $(*)$ platí i po záměně $q \leftrightarrow p$ (protože $p_i(f)$ a $q_i(f)$ spolu komutují), načež

$$v = p_1(f)\big(q_1(f)(v)\big) + p_2(f)\big(q_2(f)(v)\big) = p_1(f)(0) + p_2(f)(0) = 0,$$

protože p_1, p_2 jsou lineární zobrazení. Dokázali jsme tedy, že $V = U_1 \dotplus U_2$.

Obecný případ $n > 2$ se dokáže indukcí (cvičení).

(iii) Polynom q_i je anulujícím polynomem transformace $f|_{U_i}$, protože $\operatorname{Ker} q_i(f) = U_i$, načež $\operatorname{Ker} q_i(f|_{U_i}) = \operatorname{Ker}\big(q_i(f)\big|_{U_i}\big) = U_i \cap \operatorname{Ker} q_i(f) = U_i$, a tedy $q_i(f|_{U_i}) = 0$.

K nalezení právě uvedeného rozkladu musíme znát alespoň jeden anulující polynom.

Věta Hamilton–Cayleyova. *Charakteristický polynom čtvercové matice nad P je jejím anulujícím polynomem.*

Důkaz. Pro libovolnou čtvercovou matici B jsme kdysi odvodili vztah $B \cdot \operatorname{adj} B = \det B \cdot E$. Dosaďme za B matici $A - xE$:

$$(A - xE) \cdot \operatorname{adj}(A - xE) = \chi_A(x) \cdot E.$$

Je-li matice A typu n/n, pak je její charakteristický polynom χ_A polynomem stupně n, řekněme $\chi_A = c_n x^n + \cdots + c_1 x + c_0$. Dále je (z definice adjungované matice) jasné, že prvky matice $\operatorname{adj}(A - xE)$ jsou polynomy stupně $n - 1$ v x. Sdružíme-li sčítance s týmiž mocninami x, získáme vyjádření $\operatorname{adj}(A - xE) = C_{n-1}x^{n-1} + \cdots + C_1 x + C_0$, kde C_i jsou čtvercové matice typu n/n.

Po dosazení máme

$$(A - xE) \cdot (C_{n-1}x^{n-1} + \cdots + C_1 x + C_0) = (c_n x^n + \cdots + c_1 x + c_0) \cdot E,$$

tj.

$$\begin{aligned} &-C_{n-1}x^n + (AC_{n-1} - C_{n-2})x^{n-1} + \cdots + (AC_1 - C_0)x + AC_0 \\ &\qquad = c_n E x^n + \cdots + c_1 E x + c_0 E. \end{aligned}$$

Porovnáním koeficientů u stejných mocnin x obdržíme

$$\begin{aligned}
-C_{n-1} &= c_n E,\\
-C_{n-2} + AC_{n-1} &= c_{n-1},\\
&\vdots\\
-C_0 + AC_1 &= c_1\\
AC_0 &= c_0.
\end{aligned}$$

Vynásobíme-li i-tou rovnost i-tou mocninou A^i matice A a vzniklé rovnosti sečteme, získáme

$$0 = c_n A^n + c_{n-1}A^{n-1} + c_1 A + c_0,$$

což se mělo dokázat.

Důsledek. *Charakteristický polynom lineární transformace $f : V \to V$ je jejím anulujícím polynomem.*

Příklad. Nechť

$$A = \begin{pmatrix} 1 & -1 & 1\\ 0 & 2 & -1\\ -1 & 0 & 2 \end{pmatrix}.$$

Charakteristický polynom je $x^3 - 5x^2 + 9x - 5 = (x-1)(x^2 - 4x + 5)$ (ověřte). Vidíme, že polynom χ_A je součinem nesoudělným polynomů $q_1 = x - 1$ a $q_2 = x^2 - 4x + 5$.

Matice A představuje lineární zobrazení $\alpha : \mathbf{R}^3 \to \mathbf{R}^3$, $u \mapsto Au$. Počítejme $U_1 = \operatorname{Ker} q_1(\alpha) = \operatorname{Ker}(\alpha - \mathrm{id})$. Jádro $\operatorname{Ker}(\alpha - \mathrm{id})$ vypočteme řešením rovnice $(\alpha - \mathrm{id})(u) = 0$, což je homogenní soustava s maticí

$$q_1(A) = A - E = \begin{pmatrix} 1 & -1 & 1\\ 0 & 2 & -1\\ -1 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1\\ 0 & 1 & -1\\ -1 & 0 & 1 \end{pmatrix}$$

a fundamentálním řešením $e_1 = (1, 1, 1)$ (ověřte). Dostáváme jednorozměrný invariantní podprostor

$$U_1 = \operatorname{Ker} q_1(\alpha) = [\![(1, 1, 1)]\!].$$

Podobně $U_2 = \operatorname{Ker} q_2(\alpha) = \operatorname{Ker}(\alpha^2 - 4\alpha + 5\,\mathrm{id})$. Toto jádro vypočteme řešením rovnice

$$(\alpha^2 - 4\alpha + 5\,\mathrm{id})(u) = 0,$$

což je homogenní soustava s maticí

$$q_2(A) = A^2 - 4A + 5E = \begin{pmatrix} 1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 1 & 0 \end{pmatrix}$$

a fundamentálním řešením $e_2 = (0, 0, 1)$, $e_3 = (1, -1, 0)$ (ověřte). Dostáváme dvourozměrný invariantní podprostor

$$U_2 = \operatorname{Ker} q_2(\alpha) = [\![(0, 0, 1), (1, -1, 0)]\!],$$

Získali jsme (a) přímý rozklad $\mathbf{R}^3 = U_1 \dot{+} U_2$ na invariantní podprostory U_1 a U_2;
(b) bázi e_1, e_2, e_3 s maticí přechodu

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

v níž má zobrazení α blokově diagonální matici

$$Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

4. Minimální polynom

Příklad. Matice

$$A = \begin{pmatrix} -1 & 0 & 2 \\ -2 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

má anulující polynom $q = x^2 - 1 = (x+1)(x-1)$ (ověřte). Charakteristický polynom matice A je $\chi_A = x^3 - x^2 - x + 1 = (x+1)(x-1)^2$.

Poslední příklad ukazuje, že charakteristický polynom může mít netriviálního dělitele, který je rovněž anulujícím polynomem. Ukažme, že mezi anulujícími polynomy existuje jeden, který dělí všechny ostatní.

Definice. Anulující polynom se nazývá *minimální polynom*, je-li normovaný a nejmenšího stupně ze všech anulujících polynomů.

Tvrzení. *Každý anulující polynom je dělitelný minimálním polynomem.*

Důkaz. Buď f anulující polynom matice A, buď g minimální polynom matice A. Dělme se zbytkem: $f = qg + r$, kde $r = 0$ nebo $\deg r < \deg g$. V případě $r = 0$ jsme hotovi. Připusťme opak, tj. $r \neq 0$. Potom

$$0 = f(A) = q(A)g(A) + r(A) = r(A),$$

protože $g(A) = 0$. Tudíž, r je anulující polynom nižšího stupně než polynom g, a to je spor.

Důsledek. *Ke každé lineární transformaci konečněrozměrného vektorového prostoru V resp. ke každé čtvercové matici A existuje minimální polynom a je jediný.*

Cvičení. Dokažte jednoznačnost minimálního polynomu.