František Katrnoška

Genetické algebry. Věnováno prof. RNDr. Ladislavu Procházkovi, DrSc., k jeho 75. narozeninám

Pokroky matematiky, fyziky a astronomie, Vol. 50 (2005), No. 1, 62–74

Persistent URL: `http://dml.cz/dmlcz/141254`

Genetické algebry

Věnováno prof. RNDr. Ladislavu Procházkovi, DrSc., k jeho 75. narozeninám.

František Katrnoška, Praha

1. Úvod

Základní koncepce klasické genetiky vznikla roku 1865, když J. G. Mendel zdůvodnil získané výsledky svých pokusů pomocí matematické statistiky. Zpočátku se v genetice používaly pouze elementární matematické metody. Příkladem toho je matematické zdůvodnění Hardyova-Weinbergova zákona formulovaného nezávisle slavným anglickým matematikem G. H. Hardym a německým lékařem W. Weinbergem roku 1908 (viz [18]–[19]). Tento zákon se stal základem populační genetiky. V současnosti již nejsou matematické metody používané v genetice tak elementární. Často se používají metody teorie pravděpodobnosti, statistiky [1], [18], [22], náhodných procesů [22], [31], nelineární diferenciální a diferenční rovnice [11], diferenciální operátory [26], neasociativní algebry [4]–[10], [20]–[26], [29], [32] a též se používají topologické disciplíny k řešení některých problémů vznikajících v genetice [24]. Tato práce se týká užití teorie algeber v genetice. Algebry, které budou v této práci uvedeny, jsou většinou komutativní, nikoliv však asociativní, a není možné je úplně zařadit do některé ze známých tříd neasociativních algeber (tj. algeber Lieových, alternativních nebo Jordanových [14]). Splňují-li však algebry další podmínky, je možné je použít k řešení některých problémů vznikajících v genetice (viz [9], [12]). Je však třeba říci, že pojem genetické algebry není přesně vyhraněn, neboť se u různých autorů liší. Pionýrskými pracemi v uvedeném směru jsou články S. N. Bernsteina [2], V. I. Glivenka [8], I. M. H. Etheringtona [4]–[7]. Pozdější práce P. Holgatea [12], [13], Ju. I. Ljubiče [20]–[24], R. Raffina [25], O. Reiersöla [26], R. D. Schafera [29], na ně navazující, mají již složitější charakter, matematika v nich používaná je obtížnější. Pojem genetické algebry se poprvé objevil v práci [4] I. M. H. Etheringtona v roce 1939, i když již předtím S. N. Bernstein [2] v roce 1923 a V. I. Glivenko [8] roku 1936 pojem algebry ve stejném smyslu použili. I. M. H. Etherington zavedl v [4] pojem algebry barické a train algebry a roku 1941 v [5] pojem speciální train algebry, kterými se zabývají též H. Gonshor [9], Ju. I. Ljubič [22], [24] a další autoři. Teorie genetických algeber je v současnosti velmi rozsáhlá. Tento článek je jen stručnou informací o některých výsledcích aplikací genetických algeber v genetice.

Doc. RNDr. František Katrnoška, CSc. (1929), Ústav matematiky VŠCHT, Technická 5, 166 28 Praha 6.

2. Pojem algebry

Nejprve uvedeme některé základní pojmy. Nechť F je komutativní těleso s jednotkovým prvkem e.

Definice 1. *F-algebrou A* nazveme lineární prostor A nad tělesem F, v němž je definována binární operace násobení $\circ$, která má následující vlastnosti:

$$(x+y)\circ z = x\circ z + y\circ z, \quad x,y,z\in A, \tag{2.1}$$
$$z\circ(x+y) = z\circ x + z\circ y, \quad x,y,z\in A,$$
$$\alpha(x\circ y) = (\alpha x)\circ y = x\circ(\alpha y), \quad \alpha\in F, \quad x,y\in A, \tag{2.2}$$

F-algebra je *asociativní*, jestliže platí

$$(x\circ y)\circ z = x\circ(y\circ z) \quad \text{pro} \;\; x,y,z\in A, \tag{2.3}$$

F-algebra A je *neasociativní*, jestliže (2.3) neplatí. Řekneme, že F-algebra A je *komutativní*, platí-li

$$x\circ y = y\circ x, \quad x,y\in A, \tag{2.4}$$

neplatí-li (2.4), je F-algebra *nekomutativní*.

Pojmy podalgebry, ideálu algebry, idempotentu algebry, homomorfismu, jádra homomorfismu, které čtenář zná ze studia algebraických struktur, jsou nezávislé na jejich asociativnosti. Definice těchto pojmů jsou pro algebry v podstatě stejné jako pro okruhy. Jejich znalost se u čtenáře předpokládá. V článku nebudou definovány.

Jestliže A je algebra nad tělesem F, pak *dimenzí algebry A* nazveme dimenzi jejího lineárního prostoru. Vzhledem k tomu, že genetické algebry mají konečnou dimenzi, budeme se dále zabývat algebrami konečněrozměrnými.

Pro ilustraci uveďme dva jednoduché příklady.

Příklad 1. Vektorový prostor F^n $(n\in N)$ nad tělesem F, v němž operace násobení je definována po souřadnicích, je komutativní asociativní F-algebra dimenze n.

Příklad 2. Množina $M_n(F)$ všech (n,n)-matic nad tělesem F s obvykle definovanými maticovými operacemi (sčítání, násobení) a skalárním násobkem je nekomutativní asociativní F-algebrou dimenze n^2.

Ukážeme si nyní způsob konstrukce algeber konečné dimenze. Nechť A je algebra nad komutativním tělesem F. Předpokládejme, že $n=\dim A$. Označme prvky báze lineárního prostoru algebry A symboly $u_1,u_2,\dots,u_n$. Strukturu F-algebry A můžeme zadat, definujeme-li součiny prvků báze F-algebry A následujícím způsobem

$$u_i\circ u_j = \sum_{k=1}^{n} a_{ij}^k u_k, \quad a_{ij}^k\in F, \quad 1\leqq i,j\leqq n. \tag{2.5}$$

Výše uvedené vztahy můžeme nyní prodloužit na všechny prvky F-algebry A tímto způsobem:

Nechť $b, c \in A$,

$$b = \sum_{i=1}^{n} b_i u_i, \quad c = \sum_{j=1}^{n} c_j u_j.$$

Potom užitím vztahu (2.5) dostáváme

$$b \circ c = \Big(\sum_{i=1}^{n} b_i u_i\Big) \circ \Big(\sum_{j=1}^{n} c_j u_j\Big) = \sum_{k=1}^{n}\Big(\sum_{i,j=1}^{n} b_i c_j a_{ij}^{k}\Big) u_k. \tag{2.6}$$

Prvky a_{ij}^k vyskytující se v předešlých vztazích se nazývají *strukturními konstantami* algebry A. Počet strukturních konstant této F-algebry A je n^3. Takto zkonstruované algebry ovšem nemusí být asociativní.

Ukážeme si nyní ještě dobře známý příklad.

Příklad 3. Nechť F je komutativní těleso s jednotkovým prvkem e. Symbolem F^3 označme třírozměrný lineární prostor nad F. Za bázi tohoto prostoru vezměme vektory $u_1 = (e, 0, 0)$, $u_2 = (0, e, 0)$, $u_3 = (0, 0, e)$. Definujme nyní v prostoru F^3 součin prvků

$$u_i \circ u_k = \sum_{j=1}^{3} a_{ik}^{j} u_j, \quad i, k = 1, 2, 3.$$

Strukturní konstanty a_{ik}^j určíme takto:

$$\begin{aligned} a_{ik}^{j} &= e, && \text{jestliže } (i,k,j) \text{ je sudou permutací indexů } (1,2,3),\\ a_{ik}^{j} &= -e, && \text{jestliže } (i,k,j) \text{ je lichou permutací indexů } (1,2,3),\\ a_{ik}^{j} &= 0, && \text{jestliže } \emptyset \neq \{i,k,j\} \subsetneqq \{1,2,3\}. \end{aligned} \tag{2.7}$$

Nechť b, c jsou libovolné prvky F^3, kde

$$b = b_1u_1 + b_2u_2 + b_3u_3, \quad c = c_1u_1 + c_2u_2 + c_3u_3, \quad b_i, c_i \in F, \quad i = 1, 2, 3.$$

Užitím vztahů (2.7) lze dokázat, že platí

$$b \circ c = \begin{vmatrix} u_1, & u_2, & u_3 \\ b_1, & b_2, & b_3 \\ c_1, & c_2, & c_3 \end{vmatrix}. \tag{2.8}$$

Součin prvků b, c je tedy známý vektorový součin vektorů b, c, který není komutativní ani asociativní.

3. Algebry vyskytující se v genetice

V tomto odstavci se budeme zabývat algebrami, které vznikly v období let 1920 až 1980. Velice dobrý přehled o tom může čtenář získat v knize [32]. Byly to zejména průkopnické práce [2], [4]–[8], v nichž byly definovány algebry, které úzce souvisely s populační genetikou.

V průběhu dalších let byly zavedeny algebry stochastické, algebry Bernsteinovy [21] a další algebry, jejichž vznik byl iniciován problémy genetiky. Všechny tyto algebry byly zařazeny mezi genetické algebry.

Zaveďme nyní některé důležité pojmy.

Populací nazveme skupinu jednotlivců určitého druhu, nacházející se na určitém místě. Populace je *panmiktická*, je-li v ní zcela zaručen náhodný výběr při křížení partnerů. Dále budeme předpokládat, že v populaci nedochází k žádným migracím ani mutacím a že průměrná plodnost všech jedinců je stejná. *Genotyp* je souhrn všech genů určitého jednotlivce.

Nechť panmiktická populace má n genotypů $g_1, g_2, \dots, g_n$ a nechť genotypy mužských a ženských jedinců vyskytujících se v populaci jsou identické. Označme symbolem $x_i \in R$, $i = 1, 2, \dots, n$, pravděpodobnost výskytu jedinců genotypu g_i v dané populaci. Pak stav populace je určen vektorem $x = (x_1, x_2, \dots, x_n)$, kde

$$x_i \geqq 0, \quad i = 1, 2, \dots, n, \quad \text{a} \quad \sum_{i=1}^{n} x_i = 1.$$

Takový vektor se nazývá *stochastický vektor*. Pro konstrukci algebry obsahující všechny stochastické vektory je tedy vhodné vzít za základ eukleidovský lineární prostor R^n.

Nad tímto prostorem nyní zkonstruujeme genetickou algebru. Proto nejprve určíme strukturní konstanty algebry. Tyto konstanty musí mít tzv. genetické zdůvodnění. K tomuto účelu si připomeneme některé další důležité pojmy. Při tzv. dialelním křížení se sleduje shoda nebo rozdíl mezi znaky rodičů a potomků. V našem případě toto můžeme vyjádřit určením pravděpodobností výskytu jednotlivých genotypů potomků. Z těchto důvodů označme symbolem p_{ik}^{j} pravděpodobnost narození potomka genotypu g_j rodičům s genotypy g_i, resp. g_k, $i, j, k \in \{1, 2, \dots, n\}$. Zřejmě platí

$$p_{ik}^{j} = p_{ki}^{j} \geqq 0, \quad i, j, k \in \{1, 2, \dots, n\}, \tag{3.1}$$

$$\sum_{j=1}^{n} p_{ik}^{j} = 1 \quad \text{pro každé} \ \ i, k \in \{1, 2, \dots, n\}. \tag{3.2}$$

Čísla p_{ik}^{j} splňující podmínky (3.1), (3.2) definujeme jako strukturní konstanty algebry, kterou označíme symbolem $G^n(R)$. Zmiňme se nyní podrobněji o konstrukci této algebry.

Nechť A_n je n-rozměrný vektorový prostor nad R. Předpokládejme, že jeho báze je tvořena genotypy $g_1, g_2, \dots, g_n$ nějaké (neselektivní) populace[1]), jejíž strukturní konstanty dědičnosti p_{ik}^{j} splňují podmínky (3.1) a (3.2). Operaci násobení prvků báze g_i a g_k pro $i, k = 1, \dots, n$ definujeme vztahem podobným (2.5). Rozšíříme-li operaci násobení na všechny prvky A_n, zkonstruujeme tímto způsobem komutativní algebru R-algebru $G^n(R)$.

---

[1]) Selekce a její význam v populaci je popsána v [11], [31].

Ukažme nyní, jaká je genetická interpretace algebry $G^n(R)$. Jednotlivec populace g může být reprezentován pomocí stochastického vektoru algebry $G^n(R)$, tj.

$$g = \sum_{i=1}^{n} \alpha_i g_i, \quad 0 \leqq \alpha_i \leqq 1 \ \text{ pro } \ i = 1, \dots, n, \quad \sum_{i=1}^{n} \alpha_i = 1.$$

Křížením jedinců g a g' této populace vznikne potomek, kterého lze reprezentovat vektorem tvaru

$$g \circ g' = \Big(\sum_{i=1}^{n} \alpha_i g_i\Big) \circ \Big(\sum_{i=1}^{n} \alpha'_i g_i\Big) = \sum_{i,k=1}^{n} \alpha_i \alpha'_k (g_i \circ g_k) = \sum_{j=1}^{n} \Big(\sum_{i,k=1}^{n} p_{ik}^{j} \alpha_i \alpha'_k\Big) g_j.$$

Zavedeme-li nyní označení $\beta_j = \sum_{i,k=1}^{n} p_{ik}^{j} \alpha_i \alpha'_k$, pak platí $0 \leqq \beta_j \leqq 1$ pro $j = 1, \dots, n$ a $\sum_{j=1}^{n} \beta_j = 1$. Odtud plyne, že potomek jedinců g a g' reprezentovaný vektorem $g \circ g'$ má distribuci pravděpodobnosti genotypů dané populace rovnou koeficientům β_j.

Jak již bylo řečeno v úvodu této práce, není v případě algebry $G^n(R)$ ustálená terminologie. V knize [32] je algebra $G^n(R)$ nazývána *algebrou s genetickou realizací* nebo *gametickou algebrou*, zatímco Ju. I. Ljubič ji v práci [22] nazval *stochastickou algebrou.* Ke vztahu (3.1) ještě poznamenejme, že jeho platnost je zdůvodněna výsledky zjištěnými experimentálně, tj. že zygoty $g_i \times g_k$, $g_k \times g_i$, $i \neq k$, produkují gamety g_j, $j = 1, 2, \dots, n$ (tj. pohlavní buňky), se stejnými pravděpodobnostmi. Důsledkem vztahů (3.1) je, že algebry vyskytující se v genetice jsou komutativní. Nejsou však asociativní, jak ukazuje následující příklad.

Většina genů existuje alespoň ve dvou rozdílných formách, které označujeme jako *alely* (tj. alelomorfní geny).

Příklad 4. Uvažujme populaci mající alely $\boldsymbol{A}$ a $\boldsymbol{a}$, které ztotožníme s bázi lineárního prostoru dimenze 2. Nad ním zkonstruujeme algebru, definujeme-li násobení prvků báze $\boldsymbol{A}$, $\boldsymbol{a}$ následujícím způsobem

$$\boldsymbol{A} \circ \boldsymbol{A} = \boldsymbol{A}, \quad \boldsymbol{a} \circ \boldsymbol{a} = \boldsymbol{a}, \quad \boldsymbol{A} \circ \boldsymbol{a} = \boldsymbol{a} \circ \boldsymbol{A} = \frac{1}{2}\boldsymbol{A} + \frac{1}{2}\boldsymbol{a}. \tag{3.3}$$

Operaci $\circ$ rozšíříme nyní na všechny prvky lineárního prostoru. Tímto způsobem jsme zkonstruovali algebru dimenze 2, která se nazývá *mendelovskou algebrou gamet* (viz Ljubič [20], [24]). Ze vztahů (3.3) plyne, že tato algebra je komutativní. Není však asociativní, neboť platí

$$(\boldsymbol{A} \circ \boldsymbol{A}) \circ \boldsymbol{a} = \frac{1}{2}\boldsymbol{A} + \frac{1}{2}\boldsymbol{a}, \quad \boldsymbol{A} \circ (\boldsymbol{A} \circ \boldsymbol{a}) = \frac{3}{4}\boldsymbol{A} + \frac{1}{4}\boldsymbol{a},$$

tj. $(\boldsymbol{A} \circ \boldsymbol{A}) \circ \boldsymbol{a} \neq \boldsymbol{A} \circ (\boldsymbol{A} \circ \boldsymbol{a})$.

4. Obecná evoluce populace

Vývoj každé populace je vždy provázen určitými změnami jejích stavů. Tyto změny stavů popisují a určují změny pravděpodobností výskytů různých genotypů populace. Nedochází-li u populace ke změnám těchto pravděpodobností v dalších generacích, pak populace je v *evoluční rovnováze.* Nyní vznikají dvě otázky.

1. Má vždy populace stav, kdy je v evoluční rovnováze?
2. Jestliže je populace v evoluční rovnováze, jak lze tuto rovnováhu zjistit?

Odpověď na tyto otázky je předložena v tomto odstavci. Nejprve však uvedeme některé důležité pojmy, které dále použijeme.

Množinu

$$\Delta^{n-1} = \left\{(x_1, x_2, \dots, x_n) \in R^n\colon \quad 0 \leqq x_i \leqq 1, \quad i = 1, 2, \dots, n, \quad \sum_{i=1}^{n} x_i = 1\right\}$$

nazveme $(n-1)$-*simplexem.* Odtud plyne, že prvky n-simplexu jsou stochastické vektory, které budeme v dalším textu nazývat též *stavy populace.* Pro vyšetřování změn stavů populace byl zaveden operátor, který byl nazván *evolučním operátorem.* Jeho definice je následující:

Definice 2 (viz [24])**.** Evoluční operátor je zobrazení $V\colon \Delta^{n-1} \to \Delta^{n-1}$ definované předpisem

$$\big(V(x)\big)_j = \sum_{i,k=1}^{n} p_{ik}^j x_i x_k, \quad j = 1, \dots, n, \tag{4.1}$$

kde jednotlivé složky vektoru $x = (x_1, x_2, \dots, x_n)$ odpovídají různým alelám a p_{ik}^j jsou *strukturní konstanty dědičnosti* splňující podmínky (3.1) a (3.2). Číslo n se nazývá *dimenze populace.*

Z genetického hlediska je důležité vědět, jaké podmínky musí populace splňovat, aby u ní nastal rovnovážný stav. Matematicky to lze vyjádřit následujícím způsobem:

Rovnovážným stavem populace nazýváme takový její stav $x \in \Delta^{n-1}$, pro který platí $V(x) = x$. Rovnovážné stavy populace jsou tedy pevné body operátoru V. Tím se otázka existence pevných bodů operátoru V stává topologickou záležitostí.

Uvedeme nyní větu, která byla dokázána již roku 1910 L. E. L. Brouwerem v práci *Über eineindeutige stetige Transformationen von Flächen in sich* v časopise Math. Annalen.

Věta 1 (Brouwerova)**.** *Jestliže T je spojité zobrazení uzavřené jednotkové koule $S = \{x \in R^n\colon \|x\| \leqq 1\}$ do S, pak existuje prvek $x_0 \in S$ takový, že $T(x_0) = x_0$.*

Prvek x_0 splňující podmínku předešlé věty se nazývá *pevným bodem* zobrazení T.

Brouwerova věta byla v pozdějších letech 1930–1935 zobecněna J. Schauderem a A. N. Tichonovem. Tato věta má následující tvar:

Věta 2 (Schauderova-Tichonovova)**.** *Nechť K je konvexní, kompaktní podmnožina lokálně konvexního lineárního topologického prostoru X. Jestliže $f\colon K \to K$ je spojité zobrazení, pak existuje prvek $x_0 \in K$ takový, že $f(x_0) = x_0$.*

Jestliže V je evoluční operátor populace, pak lze snadno dokázat, že pro $x_0 \in \Delta^{n-1}$ platí $V(x_0) \in \Delta^{n-1}$ a že V je spojité zobrazení. Jak již bylo uvedeno, n-simplex Δ^{n-1} je konvexní a kompaktní podmnožina lokálně konvexního lineárního topologického prostoru R^n. Protože zobrazení V splňuje podmínky věty 2, existuje pevný bod x_0 operátoru V (tj. že platí $V(x_0) = x_0$). Tím je dokázáno, že populace má vždy rovnovážný stav. Podle [20] či [24] je množina $E(V)$ všech pevných bodů operátoru (rovnovážných stavů populace) kompaktní. Platí-li pro populaci $E(V) = \Delta^{n-1}$, nazývá se tato populace *identicky rovnovážnou.*

Zabývejme se nyní způsobem zjištění rovnovážných stavů populace (tj. pevných bodů operátoru V). K tomuto účelu zkonstruujeme algebru, kterou lze zařadit mezi algebry genetické. Lineárním prostorem této algebry prohlásíme prostor R^n s bází $e_i = (0, \dots, 1, \dots, 0)$, $i = 1, 2, \dots, n$. V prostoru R^n definujeme nyní násobení prvků báze $e_1, e_2, \dots, e_n$ následujícím předpisem

$$e_i \circ e_k = \sum_{j=1}^{n} p_{ik}^j e_j, \quad 1 \leqq i, k \leqq n, \tag{4.2}$$

kde p_{ik}^j jsou konstanty dědičnosti populace, které prohlásíme za strukturní konstanty konstruované algebry. Nechť $x \in R^n$, $x = \sum\limits_{j=1}^{n} x_j e_j$. Užitím (4.2) dostáváme

$$\begin{aligned} x^2 = x \circ x &= \left(\sum_{i=1}^{n} x_i e_i \right) \circ \left(\sum_{k=1}^{n} x_k e_k \right) = \sum_{i,k=1}^{n} x_i x_k (e_i \circ e_k) = \\ &= \sum_{i,k=1}^{n} x_i x_k \left(\sum_{j=1}^{n} p_{ik}^j e_j \right) = \sum_{j=1}^{n} \left(\sum_{i,k=1}^{n} p_{ik}^j x_i x_k \right) e_j. \end{aligned}$$

Vzhledem k (4.1) dostáváme, že

$$x^2 = x \circ x = \sum_{j=1}^{n} (V(x))_j e_j = V(x).$$

Odtud plyne, že pro $x \in R^n$ v konstruované algebře platí $V(x) = x^2$. Za předpokladu, že prvek x je rovnovážný stav populace, platí zároveň i $V(x) = x$. Proto v tomto případě dostáváme, že $x^2 = x$, tj. rovnovážné stavy populace jsou idempotentní prvky námi zkonstruované algebry, která se obvykle nazývá *evoluční algebrou* (viz [24], [32]). Tuto algebru označíme symbolem A_V.

Evoluční operátor V splňuje *Bernsteinův princip*, jestliže $V^2 = V$ (tj. když V je idempotentní operátor). Odpovídající biologická interpretace je, že populace, jejíž evoluční operátor splňuje Bernsteinův princip, zůstává v evoluční rovnováze v příští generaci. Bernsteinův princip je tím vlastně určité zobecnění Hardyova-Weinbergova zákona.

5. Barické a Jordanovy algebry

Barické algebry byly zavedeny I. M. H. Etheringtonem v práci [4] v souvislosti s genetickými algebrami. Uveďme nejprve některé důležité pojmy a definice.

Definice 3 (viz [22]). Nechť A je komutativní algebra nad tělesem reálných čísel R. *Charakterem algebry* A nazveme nenulový lineární funkcionál $f\colon A \to R$, pro který platí $f(x \circ y) = f(x)f(y)$, $x, y \in A$.

Příklad 5. Nechť X je kompaktní Hausdorffův prostor. Jestliže $C(X)$ je algebra všech spojitých reálných funkcí na množině X, jestliže $x_0 \in X$ je daný bod a definujeme-li funkci $F_{x_0}\colon C(X) \to R$ předpisem

$$F_{x_0}(f) = f(x_0), \quad f \in C(X),$$

potom funkce F_{x_0} je charakter algebry $C(X)$.

Ukážeme příklad komutativní algebry, která nemá charakter.

Příklad 6. Vezměme reálný trojrozměrný lineární prostor R^3 s bází tvořenou vektory $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Operaci násobení vektorů tvořících bázi prostoru R^3 zavedeme následujícím způsobem: jestliže $i \in \{1, 2, 3\}$, pak položíme $e_i^2 = 0$ a dále pak definujeme

$$\begin{aligned} e_1 \circ e_2 &= e_2 \circ e_1 = e_3, \\ e_1 \circ e_3 &= e_3 \circ e_1 = e_2, \\ e_2 \circ e_3 &= e_3 \circ e_2 = e_1. \end{aligned}$$

Lineární prostor R^3 s takto definovaným násobením je komutativní algebra dimenze 3, která nemá jednotkový prvek, není asociativní, neboť například $(e_1 \circ e_1) \circ e_2 = 0$, $e_1 \circ (e_1 \circ e_2) = e_2$. Tuto algebru označíme symbolem A_3. Nyní ukážeme, že algebra A_3 nemá charakter. Kdyby lineární funkcionál f splňoval podmínku danou definicí 3, pak by pro vektory báze e_i, $i = 1, 2, 3$, platilo $f(e_i^2) = f(e_i)f(e_i) = 0 \Longrightarrow f(e_i) = 0$, $i = 1, 2, 3$. Kdyby $x \in A_3$, $x = x_1e_1 + x_2e_2 + x_3e_3$, pak by odtud vyplývalo, že $f(x) = 0$, tj. funkcionál f by byl nulový a nemohl by pak být charakterem algebry.

Poznámka. Jestliže $b, c \in A_3$, $b = b_1e_1 + b_2e_2 + b_3e_3$, $c = c_1e_1 + c_2e_2 + c_3e_3$, pak lze ukázat, že platí

$$b \circ c = \operatorname{per} \begin{vmatrix} e_1, & e_2, & e_3 \\ b_1, & b_2, & b_3 \\ c_1, & c_2, & c_3 \end{vmatrix}, \tag{5.1}$$

kde symbol $\operatorname{per} M$ označuje tzv. *permanent* matice M. Definice permanentu se liší od definice determinantu pouze tím, že znaménka všech součinů tvaru $e_ib_jc_k$ jsou kladná (viz [28]).

Charakter algebry je tedy homomorfismus algebry na algebru R všech reálných čísel, a je zároveň jedinou možnou reprezentací neasociativních algeber.

Poznamenejme ještě, že charakter algebry bývá též nazýván lineárně-multiplikativním funkcionálem (viz [33], [34]).

Uveďme nyní nutnou a postačující podmínku pro existenci charakteru na komutativní reálné algebře.

Věta 3 (viz [22]). *Pro existenci charakteru na komutativní algebře A je nutné a stačí, aby v této algebře A existoval ideál I takový, že $I \not\supset A^2$ a* $\operatorname{codim} I = 1$.

Je obecně známo, že systém všech charakterů algebry je lineárně nezávislý. V této souvislosti byla dokázána následující věta:

Věta 4. *Jestliže A je algebra nad tělesem reálných čísel, která má dimenzi n, pak počet všech různých charakterů této algebry může být roven nejvýše n.*

Má-li algebra A charakter f, pak tento charakter se nazývá též *váhou algebry*.

Definice 4 (viz [4]). Algebru mající charakter nazýváme *barickou algebrou*.

Jestliže A je R-algebra s genetickou realizací, jejíž bázi tvoří gametické typy $g_1, g_2 \ldots, g_n$, pak A je barická algebra (viz [32]). Uveďme nyní příklad barické algebry, která má určitý význam i v kvantové mechanice.

Příklad 7 (viz [17]). Označme symbolem $L_n(R)$ množinu všech čtvercových matic řádu n s reálnými prvky, u nichž součty všech prvků jednotlivých řádků i jednotlivých sloupců jsou rovny témuž reálnému číslu. Matice tohoto tvaru jsou zobecněním známých *latinských čtverců* (viz [28]).

Lze ukázat, že pro $n > 2$ je $L_n(R)$ vzhledem k obvyklým maticovým operacím nekomutativní, asociativní R-algebrou s jednotkovým prvkem. Označíme-li symbolem f zobrazení $f\colon L_n(R) \to R$ definované předpisem $f(M) = \sum_{k=1}^{n} a_{ik}$, kde $M \in L_n(R)$, $i = 1, 2, \ldots, n$, $M = (a_{ik})$, pak f je charakter algebry $L_n(R)$, a tato algebra je tedy barická. Některé vlastnosti barických algeber jsou uvedeny v [13], [24] a [29].

Poznámka. V práci [17] F. Katrnoška zavedl pojem R-latinského čtverce, který je zobecněním latinského čtverce nad N. Prvky R-latinských čtverců jsou prvky komutativního okruhu R s jednotkovým prvkem.

Čtenář zná pojem magického čtverce, který je speciálním případem latinského čtverce. V této souvislosti je třeba se zmínit o nedávném významném výsledku dosaženém v tomto směru N. Revoyem (viz [27]), který zkonstruoval pomocí počítače jako první magickou krychli řádu 5. V této krychli jsou rozmístěna všechna přirozená čísla od 1 do 125 do jednotlivých řádků, sloupců, vrstev analogickým způsobem jako u magických čtverců. Přitom Revoy musel otestovat celkem 80 000 krychlí.

V šedesátých letech minulého století byly studovány genetické neasociativní algebry, které byly původně zavedeny roku 1932 fyzikem P. Jordanem (viz [15], [16]) pro objasnění formalismu kvantové mechaniky. V pozdějších letech se ukázalo, že algebry tohoto typu umožňují určit řešení komplikovaných situací vznikajících v genetice. První výsledky jsou obsaženy v pracích [12] a [29]. Uvedeme nyní definici této algebry, která byla nazvána podle svého autora algebrou Jordanovou.

Definice 5 (viz [29]). *Jordanovou algebrou* se nazývá komutativní algebra J nad tělesem všech reálných čísel R, která splňuje následující podmínku

$$(x \circ y) \circ x^2 = x \circ (y \circ x^2), \quad x, y \in J. \tag{5.2}$$

Vztah (5.2) se nazývá Jordanova identita.

Příklad 8. Nechť $G^n(R)$ je R-algebra, jejíž bázi tvoří n alel $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$. Multiplikace vektorů tvořících bázi algebry $G^n(R)$ je určena následujícím způsobem

$$\boldsymbol{a}_i \circ \boldsymbol{a}_j = \frac{1}{2}\,(\boldsymbol{a}_i + \boldsymbol{a}_j), \quad i, j = 1, \ldots, n. \tag{5.3}$$

Algebra $G^n(R)$ je barická. Charakter této algebry je zobrazení $f\colon G^n(R) \to R$ definované předpisem $f(\boldsymbol{a}) = \sum_{i=1}^{n} \alpha_i$, jestliže $\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i$. Užitím (5.3) lze dokázat, že platí $\boldsymbol{a} \circ \boldsymbol{a} = f(\boldsymbol{a})\boldsymbol{a}$. Odtud vyplývá, že prvek algebry $G^n(R)$, který má váhu rovnou 1, je idempotentní. Genetická aplikace tohoto posledního algebraického výsledku znamená, že pokud v populaci neexistuje selekce, pak pravděpodobnosti výskytu alel $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ jsou ve všech následujících generacích stále stejné.

Algebra $G^n(R)$, jak snadno plyne z (5.3), není asociativní. Ukážeme však, že je algebrou Jordana. Vzhledem k tomu, že pro $\boldsymbol{a}, \boldsymbol{b} \in G^n(R)$ platí

$$\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{b} \circ \boldsymbol{a}, \quad \boldsymbol{a} \circ \boldsymbol{a} = f(\boldsymbol{a})\boldsymbol{a},$$

dostáváme

$$\begin{aligned} \big[(\boldsymbol{a} \circ \boldsymbol{a}) \circ \boldsymbol{b}\big] \circ \boldsymbol{a} &= \big[\big(f(\boldsymbol{a})\boldsymbol{a}\big) \circ \boldsymbol{b}\big] \circ \boldsymbol{a} = f(\boldsymbol{a})[\boldsymbol{a} \circ \boldsymbol{b}] \circ \boldsymbol{a} = \\ &= \big(f(\boldsymbol{a})\boldsymbol{a}\big) \circ (\boldsymbol{a} \circ \boldsymbol{b}) = \big(f(\boldsymbol{a})\boldsymbol{a}\big) \circ (\boldsymbol{b} \circ \boldsymbol{a}) = \\ &= (\boldsymbol{a} \circ \boldsymbol{a}) \circ (\boldsymbol{b} \circ \boldsymbol{a}). \end{aligned}$$

Zavedeme-li v algebře $G^n(R)$ normu vztahem

$$\|\boldsymbol{a}\| = \sum_{i=1}^{n} |\alpha_i|$$

pro $\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i$, pak lze dokázat, že algebra $G^n(R)$ s touto normou je Banachova (neasociativní) algebra.[2])

Příklad 9. V algebře $L_n(R)$ zaveďme novou multiplikativní operaci

$$\circ\colon\; L_n(R) \times L_n(R) \to L_n(R)$$

vztahem

$$M_1 \circ M_2 = \frac{1}{2}\,(M_1 M_2 + M_2 M_1), \quad M_1, M_2 \in L_n(R). \tag{5.4}$$

Množina $L_n(R)$ opatřená operací $\circ$ je komutativní, neasociativní algebrou Jordana, jejíž charakter f je shodný s charakterem původní algebry $L_n(R)$ uvedené v příkladu 7. Tato algebra je tedy barická. Nezměněny zůstávají i všechny idempotentní prvky této nové algebry.

Další informace o aplikacích Jordanových algeber v genetice může čtenář nalézt v [12], [29], [32]. Jordanovy algebry mají použití i ve fyzice elementárních částic.

---

[2]) Toto tvrzení platí obecněji i pro každou stochastickou nebo evoluční barickou algebru (viz [32]).

Poznámka. Má-li matice $M \in L_n(R)$ všechny své prvky nezáporné, potom pro $f(M) > 0$ je matice $B = [1/f(M)]\,M$ dvojitě stochastická (angl. double stochastic). Množina všech dvojitě stochastických matic tvoří n-simplex Δ^n algebry $L_n(R)$. Z Birkhoffovy věty [28] plyne, že Δ^n je konvexní podmnožina algebry $L_n(R)$ taková, že pro $A \in \Delta^n$ platí $A = \sum_{i=1}^{k} \lambda_i P_i$, kde P_i, $i = 1, 2, \ldots, k$, jsou matice permutací, $\lambda_i \geqq 0$, $\sum_{i=1}^{k} \lambda_i = 1$, $k \leqq n^2 - 2n + 2$. Přitom matice P_i tvoří vrcholy n-simplexu Δ^n (viz [28]).

Uveďme ještě jedno tvrzení (viz [32]), které zdůvodňuje fakt známý genetikům, že pokud v populaci neexistuje selekce v dalších generacích, pak ve všech generacích této populace jsou poměry všech genotypů stejné.

Věta 5. *Nechť A je komutativní barická algebra s bází $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, jejíž charakter f splňuje podmínku $f(a) = \sum_{i=1}^{n} \alpha_i$, kde $\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i$. Pak platí $\boldsymbol{a}^2 = \boldsymbol{a} \circ \boldsymbol{a} = f(\boldsymbol{a})\boldsymbol{a}$ pro $\boldsymbol{a} \in A$.*

Důsledek. *Jestliže prvek $\boldsymbol{a} \in A$ má váhu $f(\boldsymbol{a}) = 1$, pak platí $\boldsymbol{a}^2 = \boldsymbol{a}$, tj. $\boldsymbol{a}$ je idempotentní prvek algebry A.*

6. Závěr

Podrobný výklad aplikací genetických algeber v genetice lze najít např. v [32]. Zde se omezíme jen na několik poznámek. Většina algeber používaných v genetice jsou algebry barické a některé z nich mají bázi, jejíž multiplikační tabulka pro operaci násobení má trojúhelníkový tvar. Tyto algebry se nazývají G-algebrami. Někteří autoři je nazývají jednoduše *genetickými algebrami*. Jejich definice je následující:

Definice 6. *G-algebrou* nazveme lineární prostor R^n s bází $v_0, v_1, \ldots, v_{n-1}$ a s charakterem f splňujícím podmínku

$$f(v_0) = I, \qquad f(v_i) = 0, \quad i = 1, 2, \ldots, n-1.$$

Multiplikační koeficienty prvků báze α_{00}^k, α_{ji}^k mají v této algebře následující tvar

$$v_0^2 = v_0 + \sum_{k=1}^{n} \alpha_{00}^k v_k, \quad v_0 v_i = \sum_{k=i}^{n} \alpha_{0i}^k v_k, \quad v_j v_i = \sum_{k \geqq i,j} \alpha_{ji}^k v_k,$$
$$1 \leqq i, j, k \leqq n-1.$$

Prvky báze $v_1, v_2, \ldots, v_{n-1}$ generují ideál, který se nazývá *barideálem.*

Vlastnosti G-algeber mají v evoluční a populační genetice hluboký význam.

Závěrem lze konstatovat, že vnitřní strukturu a charakterizaci genetických algeber ve svých pracích popsali I. Etherington [4]–[7], P. Holgate [12], Ju. I. Ljubič [20], [21], R. Raffin [25], O. Reiersöl [26], R. D. Schafer [30] a další.

Speciálními barickými algebrami jsou *algebry Bernsteinovy.* Jsou to takové barické algebry A s vahou f, které splňují následující vlastnost

$$(x^2)^2 = f^2(x) \cdot x^2, \quad x \in A.$$

Pomocí dalších vlastností Bernsteinových algeber lze zdůvodnit platnost Hardyova-Weinbergova zákona pro polyalelické lokusy. Viz Ljubič [21], [24].

V knize [32] A. Wörz-Busekros byla dokázána následující věta.

Věta 6. *Každá Bernsteinova algebra má vždy alespoň jeden idempotentní prvek.*

Celou řadu výsledků týkajících se Bernsteinových algeber uvedl ve své práci [20] Ju. I. Ljubič.

V současné době jsou genetické algebry vyšetřovány dvěma základními způsoby. Buď na základě úzké souvislosti genetických algeber s jejich odpovídajícími algebrami matic (viz Schafer [29]), nebo užitím výsledků teorie neasociativních algeber (viz Holgate [12], [13]). Druhá možnost dává jednodušší metodu k prokázání úzkého vzájemného vztahu mezi algebraickým formalismem a jemu odpovídající genetickou realitou.

Závěrem této práce uvedeme některé dosud ještě otevřené problémy týkající se genetických algeber.

1. Bylo by velmi žádoucí vytvořit teorii, která by umožnila kvalitní matematický popis vývoje, resp. evoluce populace s více lokusy.
2. Zobecnit pojem genetické algebry tak, aby její definice nevylučovala existenci selekce.
3. Pokusit se o konstrukci modelů biologických systémů, které by byly určeny a definovány pomocí více různých tříd genetických algeber.

Poděkování. Autor děkuje Mgr. Heleně Holovské za technické zpracování a jazykovou úpravu tohoto příspěvku. Práce byla podpořena grantem MŠMT: Výzkumný záměr MSM 6046137305.

L i t e r a t u r a

[1] Beier, W.: *Biophysik.* VEB Georg Thieme, Leipzig 1960.

[2] Bernstein, S.: *Demonstration mathematique de la loi d'heredité de Mendel.* C. R. Acad. Sci. Paris *177* (1923), 528–531.

[3] Drozd Kiričenko, Ju.: *Koněčnomernyje algebry.* Kiev 1980.

[4] Etherington, I. M. H.: *Genetic algebras.* Proc. Roy. Soc. Edinburgh *59* (1939), 242 až 258.

[5] Etherington, I. M. H.: *Special train algebras.* Quart. J. Math. *12* (1941), 1–8.

[6] Etherington, I. M. H.: *Nonassociative algebra and the symbolism of genetics.* Proc. Roy Soc. Edinburgh *61* (1941), 24–42.

[7] Etherington, I. M. H.: *Duplication of linear algebras.* Proc. Edinburgh Math. Soc. Ser. 2 *6* (1941), 222–230.

[8] Glivenko, V. I.: *Mendělevskaja algebra.* Dokl. Akad. Nauk *13* (1936), 371–372.

[9] Gonshor, H.: *Special train algebras arising in genetics.* Proc. Edinburgh Math. Soc. Ser. 2 *12* (1960), 41–63.

[10] Heuch, I.: *Genetic algebras considered as elements in a vector space.* Siam J. Appl. *35* (1978), 695–703.

[11] Hogben, L.: *An introduction to mathematical genetics.* Norton, New York 1946.

[12] Holgate, P.: *Jordan algebras arising in population genetics.* Proc. Edinburgh Math. Soc. *15* (1967), 291–294.

[13] Holgate, P.: *Characterization of genetic algebras.* J. London Math. Soc. *6* (1972), 169–174.

[14] Hungerford, T. V.: *Algebra.* Springer-Verlag, New York-Heidelberg-Berlin 1980.

[15] Jordan, P.: *Über Verallgeneinerungsmöglichkeiten des Formalismus der Quantenmechanik.* Nachr. Ges. Wiss. Göttingen (1933), 209–214.

[16] Jordan, P., Neumann, J., Wigner, E.: *On an algebraic generalization of the quantum mechanical formalism.* Ann. Math. *35* (1934), 29–64.

[17] Katrnoška, F: *Logics that are generated by idempotents.* Lobachevskii J. Math. *15* (2004), 11–19.

[18] Katrnoška, F., Křížek, M.: *Genetický kód a teorie monoidů aneb 50 let od objevu struktury DNA.* PMFA *48* (2003), 207–222.

[19] Katrnoška, F., Křížek, M.: *Významné milníky matematické genetiky.* Matematika–fyzika–informatika *14* (2004/05), 1–12.

[20] Ljubič, Ju. I.: *Osnovnyje poňatia i teoremy evoljucionnoj genetiki svobodnych populacii.* Uspekhi Mat. Nauk *161* (1971), 51–116.

[21] Ljubič, Ju. I.: *Bernštejnovskije algebry.* Uspekhi Mat. Nauk *32* (1977), 261–262.

[22] Ljubič, Ju. I.: *Stochastičeskije algebry i někotoryje primeněnija v matematičeskoj genetike.* In: Matematičeskije Metody v Biologii, Naukova dumka, Kiev 1977, 119–131.

[23] Ljubič, Ju. I.: *Algebraic methods in evolutionary genetics.* Biomath. J. *20* (1978), 511 až 529.

[24] Ljubič, Ju. I.: *Matematičeskije struktury v populacionnoj genetike.* Naukova dumka, Kiev 1983.

[25] Raffin, R.: *Axiomatization des algebres genetiques.* Bull. V. Sci., Acad. Roy. Belg. *17* (1951), 359–366.

[26] Reiersöl, O.: *Genetic algebras studied recursively and by means of differential operators.* Math. Scand. (1962), 25–44.

[27] Revoy, N.: *Cube magique, La perfection a été atteinte.* Science et Vie, Mars (2004), 66–68.

[28] Rybnikov, K. A.: *Vvěděnije v kombinatornyj analiz.* Izd. Moskovskogo Universiteta, Moskva 1985.

[29] Schafer, R. D.: *Structure of genetic algebras.* Amer. J. Math. *71* (1949), 121–135.

[30] Schafer, R. D.: *An introduction to nonassociative algebras.* Academic Press, New York-London 1966.

[31] Thompson, J. S. Thompson, M. W.: *Genetics in medicine.* W. B. Saunders Company, Philadelphia 1986, 2001.

[32] Wörz-Busekros, A.: *Algebras in genetics.* Springer-Verlag, Berlin 1980.

[33] Želazko, W.: *Banach algebras.* Elsevier, Amsterdam 1973.

[34] Želazko, W.: *A characterization of multiplicative linear functionals in complex Banach algebras.* Studia Math. *30* (1968), 83–85.