

UNIVERSIDADE FEDERAL DO PARANÁ
CNPQ – PROCESSO 442132/2018-2
ANA FLAVIA LOPES

ESTRUTURA ALGÉBRICA DA HERANÇA GENÉTICA

Relatório Final de Iniciação Científica associado ao projeto Meninas na Matemática: procuram-se Arletes e desenvolvido no período de Março/2019 a Fevereiro/2020, sob a orientação da Profa. Dra. Maria Eugenia Martin.

CURITIBA

2020

RESUMO

O trabalho de iniciação científica "*Estrutura Algébrica da Herança Genética*" faz parte do projeto "*Meninas na Matemática: Procuram-se Arletes*" o qual conta com o apoio financeiro do CNPq através da chamada CNPq/MCTIC N. 31/2018, "*Meninas nas Ciências Exatas, Engenharias e Computação*". Sob a coordenação da Professora Doutora Elizabeth Wegner Karas, o projeto nasce de experiências pessoais de mulheres que foram influenciadas e motivadas por professoras mulheres para seguir carreira nas ciências exatas e que se sentem com responsabilidade social em influenciar e motivar outras meninas para a carreira da Matemática. Participam do projeto cinco escolas públicas de ensino fundamental e/ou médio de Curitiba e região metropolitana, com cinco professores(as) responsáveis e quinze bolsistas de iniciação científica júnior, além de uma equipe de docentes do Departamento de Matemática da UFPR e três alunas bolsistas de iniciação científica da UFPR.

O objetivo deste trabalho é ser uma introdução ao estudo das estruturas algébricas não-associativas que surgem na genética, uma das tantas aplicações da matemática que tem significado biológico. Uma álgebra genética é uma álgebra usada para modelar a herança genética. Estas álgebras frequentemente têm uma base que corresponde aos gametas geneticamente diferentes e as constantes estruturais da álgebra codificam as probabilidades de produzir descendência de vários tipos. As leis da herança são codificadas como propriedades da álgebra. Neste trabalho apresentaremos a interação entre a estrutura puramente matemática e as propriedades genéticas correspondentes.

Palavras-chaves: álgebras não-associativas, álgebras genéticas, álgebras de Jordan, herança genética, genética de populações.

INTRODUÇÃO

A compreensão moderna da herança genética foi iniciada com as teorias de Charles Darwin, mas foi o monge agostiniano Gregor Mendel quem, em 1856, começou a descobrir a natureza matemática do assunto. De fato, o simbolismo que Mendel usou para descrever seus primeiros resultados é algebricamente sugestivo.

Mais de oito décadas depois, o matemático inglês Ivor Etherington introduziu a linguagem formal da álgebra abstrata para o estudo da genética em seu artigo "Genetic Algebras", veja [1].

O objetivo de Etherington era encontrar uma definição de *álgebra genética* que fosse geral o suficiente para abranger a maioria das álgebras

que surgem na genética, mas específica o suficiente para estudar matematicamente os intrincados detalhes da estrutura. Ele definiu duas classes de álgebras não-associativas, as básicas e as álgebras de trem, as quais surgem naturalmente no simbolismo da genética. Também introduziu o conceito de álgebras de trem especiais as quais são definidas por sua estrutura e não por qualquer tipo de equação de recorrência.

Estas últimas foram a primeira tentativa de Etherington de especializar o assunto de álgebras genéticas em um tamanho administrável. No entanto, em [3], ele apontou a falha de sua própria definição. Embora as álgebras gaméticas para os tipos fundamentais de herança simétrica sejam álgebras de trem especiais, suas duplicações comutativas (as álgebras zigóticas) podem falhar em serem álgebras de trem especiais. Portanto, ainda era necessária uma definição mais geral.

Menos de uma década após o surgimento dos primeiros trabalhos de Etherington, R. D. Schafer em seu artigo "*Structure of Genetic Algebras*" [10], forneceu um outro candidato para a definição de uma álgebra genética e estudou essas álgebras em termos de suas álgebras de transformação. O conceito foi mais satisfatório que o conceito de álgebra de trem especial em dois pontos: a estrutura da álgebra não é postulada e álgebras genéticas são fechadas em relação a duplicação comutativa. Segue desse último fato que a definição de álgebra genética de Schafer inclui, não apenas a álgebra gamética simétrica fundamental mas também as álgebras zigóticas e copular obtidas dela pelo processo de duplicação comutativa. Por outro lado, esse conceito foi restritivo o suficiente para se deduzir uma teoria de estrutura transparente para álgebras genéticas.

Embora a definição de Schafer tenha conseguido preencher a lacuna entre álgebras de trem e álgebras de trem especiais, a impraticabilidade de verificar a definição ainda era um problema. Por esse motivo, outra definição foi procurada. Em 1971, vinte e dois anos após o artigo original de Schafer aparecer, H. Gonshor em [4] deu uma nova definição de álgebra genética e mostrou que é equivalente à de Schafer, mas tem o benefício de ser muito mais simples de utilizar e verificar.

Matematicamente, as álgebras que surgem na genética são estruturas interessantes. Em geral são álgebras comutativas, porém não são associativas e estão relacionadas com uma classe de álgebras não-associativas bem conhecida que é a classe das álgebras de Jordan. Estas últimas foram introduzidas em 1932 pelo físico alemão Pascual Jordan em uma tentativa de reformular o cenário da mecânica quântica, em termos do produto de Jordan ao invés do produto associativo. A relação entre as álgebras genéticas e as álgebras de Jordan foi observada pela primeira vez por Schafer em [10]. Ele provou que as álgebras gaméticas

e zigóticas para a herança mendeliana simples para dois alelos são álgebras de Jordan especiais. Esses resultados foram generalizados por P. Holgate em [6]. Ele provou os resultados de Schafer relativos a álgebras de Jordan por métodos que não fazem uso de álgebras de transformações, e que portanto acomodam o caso multialélico mais facilmente. O principal objetivo do trabalho de Holgate é maximizar a inter-relação entre o formalismo algébrico e a situação genética correspondente.

Este trabalho será dividido em quatro seções. Na primeira estudaremos os conceitos e resultados básicos sobre álgebras não-associativas que se farão necessários para o entendimento das álgebras genéticas desde o ponto de vista matemático, assim como também apresentaremos os conceitos preliminares relativos à genética. Para tanto, as referências principais são o livro “*An Introduction to Nonassociative Algebras*”, [11], de R. D. Schafer e o livro “*Introdução à genética*”, [5], de R. C. Griffiths et al.

Nas Seções 2 e 3, daremos uma visão geral das álgebras que surgem na genética e estudaremos as propriedades e relações básicas dessas estruturas que têm significado genético. Apresentaremos a interação entre a estrutura puramente matemática e as propriedades genéticas correspondentes. As referências mais abrangentes para a pesquisa matemática feita nesta área são [1], [10], [4], [12] e o artigo “*Algebraic structure of genetic inheritance*”, [9], de M. L. Reed, o qual constitui um completo resumo da evolução deste tópico de pesquisa.

Finalmente, na Seção 4 apresentaremos a definição de álgebras de Jordan e álgebras de Jordan especiais, além de resultados clássicos sobre a estrutura desta classe de álgebras não-associativas. Faremos uma compilação dos resultados obtidos por Schafer e Holgate relativos à estrutura das álgebras genéticas que também são álgebras de Jordan.

1. PRELIMINARES

1.1. Introdução às álgebras não-associativas.

Definição 1. Uma **álgebra** sobre o corpo \mathbf{k} é um \mathbf{k} -espaço vetorial \mathcal{A} munido de um produto \cdot que satisfaz:

- (1) $(a + b) \cdot c = a \cdot b + a \cdot c$ e $a \cdot (b + c) = a \cdot b + a \cdot c$
- (2) $\alpha(a \cdot b) = (\alpha a) \cdot b = a \cdot (\alpha b)$

quaisquer que sejam $a, b, c \in \mathcal{A}$ e $\alpha \in \mathbf{k}$.

Chamaremos “álgebra sobre o corpo \mathbf{k} ” de **\mathbf{k} -álgebra** ou, simplesmente, de **álgebra** quando não houver ambiguidade.

Uma álgebra diz-se **comutativa** se $a \cdot b = b \cdot a$ quaisquer que sejam $a, b \in \mathcal{A}$. Diremos que uma álgebra \mathcal{A} é **associativa** se a lei associativa

$$(1.1) \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

é satisfeita para todos $a, b, c \in \mathcal{A}$.

Na literatura geral, uma álgebra (em nosso sentido) é comumente chamada de **álgebra não-associativa** para enfatizar que a identidade (1.1) não está sendo assumida.

Um subespaço vetorial $\mathcal{S} \subseteq \mathcal{A}$ que satisfaz $\mathcal{S} \cdot \mathcal{S} \subseteq \mathcal{S}$ é chamado de **subálgebra** de \mathcal{A} . Se um subespaço $I \subseteq \mathcal{A}$ satisfaz a condição mais forte $\mathcal{A} \cdot I \subseteq I$ então I é chamado de **ideal** da álgebra \mathcal{A} .

Um **homomorfismo** f de uma álgebra \mathcal{A} em uma álgebra \mathcal{A}' sobre um mesmo corpo \mathbf{k} é uma transformação linear $f : \mathcal{A} \rightarrow \mathcal{A}'$ tal que $f(a \cdot b) = f(a) \circ f(b)$ para todo $a, b \in \mathcal{A}$ onde \circ denota a multiplicação em \mathcal{A}' . O **kernel** $\ker(f)$ é um ideal de \mathcal{A} .

Diremos que duas álgebras \mathcal{A} e \mathcal{A}' são **isomorfas** (e denotamos por $\mathcal{A} \simeq \mathcal{A}'$) se existir um homomorfismo bijetor entre elas. Neste caso o homomorfismo é chamado de **isomorfismo**.

Dizemos que \mathcal{A} é uma **álgebra com unidade** se existe um elemento $1 \in \mathcal{A}$ tal que $1 \cdot a = a \cdot 1 = a$ para todo $a \in \mathcal{A}$. Se um tal elemento existe ele é único e é chamado de **unidade** de \mathcal{A} .

No que segue \mathcal{A} representará uma álgebra de dimensão n sobre \mathbf{k} e $B = \{e_1, e_2, \dots, e_n\}$ uma base de \mathcal{A} , assim todo elemento $a \in \mathcal{A}$ é uma combinação linear

$$a = \sum_{i=1}^n \alpha_i e_i \quad \text{com } \alpha_i \in \mathbf{k}.$$

O produto de dois elementos quaisquer de \mathcal{A} é definido por

$$\left(\sum_{i=1}^n \alpha_i e_i \right) \cdot \left(\sum_{j=1}^n \beta_j e_j \right) = \sum_{i,j=1}^n (\alpha_i \beta_j) (e_i \cdot e_j)$$

juntamente com uma tabela de multiplicação dos elementos da base:

$$(1.2) \quad e_i \cdot e_j = \sum_{k=1}^n \gamma_{ijk} e_k \quad \text{onde } \gamma_{ijk} \in \mathbf{k}.$$

Organizaremos as n^2 equações (1.2) na forma de uma tabela:

	e_1	\dots	e_j	\dots	e_n
e_1	\vdots				
\vdots	\vdots				
e_i	\dots	\dots	$\sum \gamma_{ijk} e_k$	\dots	\dots
\vdots	\vdots				
e_n	\vdots				

A multiplicação em \mathcal{A} fica completamente determinada pelos n^3 escalares $\gamma_{ijk} \in \mathbf{k}$, chamados de **constantes estruturais** da álgebra \mathcal{A} em relação à base B .

Dada uma álgebra comutativa \mathcal{A} de dimensão n sobre \mathbf{k} podemos obter a partir dela uma outra álgebra comutativa \mathcal{A}' de dimensão $\frac{1}{2}n(n+1)$ sobre \mathbf{k} através do que chamamos de **processo de duplicação comutativa** da seguinte forma: seja $\{e_1, e_2, \dots, e_n\}$ uma base de \mathcal{A} com tabela de multiplicação

$$e_i \cdot e_j = \sum_{k=1}^n \gamma_{ijk} e_k \quad (i, j = 1, \dots, n).$$

A **duplicação** \mathcal{A}' de \mathcal{A} é definida como a álgebra comutativa de dimensão $\frac{1}{2}n(n+1)$ sobre \mathbf{k} com elementos básicos

$$\{v_{i,j} \mid i \leq j \text{ e } i, j = 1, 2, \dots, n\}$$

satisfazendo

$$v_{ij} \cdot v_{rs} = \sum_{k,t} \gamma_{ijk} \gamma_{rst} v_{kt} \quad (i \leq j; r \leq s; i, j, r, s, k, t = 1, \dots, n)$$

onde identificamos $v_{tk} = v_{kt}$ para $t > k$.

Seja \mathcal{A} uma álgebra não-associativa comutativa, existem várias maneiras de definir as potências de um elemento $a \in \mathcal{A}$. A saber,

- **potências principais** $a^1 = a$ e $a^i = a^{i-1} \cdot a$
- **potências plenárias** $a^{[1]} = a$ e $a^{[i]} = a^{[i-1]} \cdot a^{[i-1]}$.

Dizemos que um elemento $a \in \mathcal{A}$ é

- um **elemento idempotente** se $a^2 = a$,
- um **elemento nilpotente** se $a^k = 0$ para algum $k \in \mathbb{N}$.

Dizemos que uma álgebra \mathcal{A} é **nilpotente** no caso em que exista um inteiro k tal que qualquer produto $a_1 \cdot a_2 \cdot \dots \cdot a_k$ de k elementos de \mathcal{A} , não importa como se associem, é zero.

Observe que se uma álgebra \mathcal{A} é nilpotente então todo $a \in \mathcal{A}$ é um elemento nilpotente, mas a recíproca não é verdade.

Uma álgebra \mathcal{A} é dita **de potências associativas** no caso em que toda subálgebra gerada por um único elemento é associativa. Nestas álgebras tem-se a seguinte relação nas potências principais $a^i \cdot a^j = a^{i+j}$ para todo elemento $a \in \mathcal{A}$.

1.2. Alguns conceitos da genética. Para um melhor entendimento das seções subseqüentes vamos a introduzir os conceitos da biologia que se farão necessários neste trabalho.

A vida na Terra é representada por todos os organismos que atualmente vivem no planeta. Uma das propriedades mais fascinantes da vida é o fato de que ela, por si própria, se regenera a cada geração, a partir de uma única célula. Tal regeneração ocorre desde a origem da vida, e cada organismo existente hoje na Terra, dos menores, como as bactérias, aos maiores, como as baleias, resulta de milhões de ciclos de regeneração. Essa observação simples levou os biólogos a pesquisarem que tipo de *informação* há dentro dessas células singulares que lhes dá a capacidade de reconstruir um organismo adulto complexo. A partir daí, a questão era: “o que constitui a **informação biológica**?”

Primeiramente, precisamos definir genética. Em termos abrangentes, **genética** é o estudo de todos os aspectos dos genes. Por sua vez, **genes** são definidos como as unidades fundamentais da informação biológica.

Desde o início do século XX, os cientistas argumentavam que, tanto em animais como em vegetais, a informação está obrigatoriamente nos cromossomos. Os **cromossomos** foram considerados prováveis portadores da informação, por serem transmitidos intactos de uma geração à seguinte, mediante divisões nucleares precisamente orquestradas, denominadas **meiose** e **mitose**.

Na década de 1940, várias correntes de pesquisa mostraram que o elemento que contém a informação biológica nos cromossomos é a molécula de DNA. A estrutura molecular detalhada do DNA foi elucidada por J. Watson e F. Crick na década de 1950, que deduziram a partir dessa estrutura que o DNA contém a informação escrita em um código genético. O **DNA** é uma série linear de quatro estruturas moleculares denominados **nucleotídios**. A sequência específica de nucleotídios constitui a linguagem do código. O DNA, como parte do cromossomo, é transmitido intacto de uma geração para a seguinte, de modo que todas as células em cada geração têm o mesmo conjunto de DNA com a mesma informação contida na sequência de nucleotídios.

Cada gene em um cromossomo pode tomar diferentes formas. Essas formas são chamadas de **alelos**. Por exemplo, o gene que determina o tipo sanguíneo em humanos tem três alelos diferentes: A , B ou O .

A disposição dos genes nos cromossomos é fixa e essas posições são chamadas de **locus**.

Na fertilização, os **gametas** (células sexuais que carregam um único conjunto de cromossomos) se fundem de maneira aleatória, independentemente de qual dos alelos ele leva. Um ovócito fertilizado, a primeira célula que se desenvolve no indivíduo da prole, é chamado de **zigoto**.

Pode-se definir um **fenótipo** como uma forma de uma característica. Geralmente, em genética, as combinações alélicas subjacentes aos fenótipos são chamadas de **genótipos**.

O conjunto completo de informação genética de um organismo, codificado no seu DNA, constitui o seu **genoma**. Organismos **diploides**, são aqueles cujos núcleos contêm duas cópias completas do genoma e, assim, dois conjuntos idênticos de cromossomos.

Uma **população** é um grupo de indivíduos da mesma espécie. A genética de populações analisa a quantidade e a distribuição da variação genética nas populações e as forças que controlam essa variação.

O **pool gênico** é um conceito fundamental para o estudo da variação genética nas populações: é o somatório de todos os alelos nos membros reprodutivos de uma população em determinado momento. Podemos descrever a variação em uma população em termos de frequências genotípicas e alélicas.

A genética mendeliana constitui a base principal da genética clássica. Trata-se de um conjunto de princípios relacionados à transmissão hereditária das características de um organismo a seus descendentes.

- (1) A **lei da segregação** afirma que alelos determinam uma característica na formação dos gametas. Os progenitores segregam os alelos aleatoriamente de modo que cada gameta possui apenas um deles. Ou seja, cada descendente possui um alelo vindo de um progenitor, e um segundo vindo do outro progenitor.
- (2) A **lei da dominância** afirma que todos os descendentes de progenitores que diferem em apenas uma característica terão o fenótipo de apenas um deles, o dominante. A outra característica é chamada recessiva, e só se manifesta caso ambos os fatores sejam recessivos.
- (3) A **lei da segregação independente** afirma que uma característica é herdada independentemente de outra característica. Ou seja, durante a formação dos gametas, os alelos que determinam cada uma das características são segregados independentemente.

As leis de Mendel, embora expliquem como os genes são transmitidos dos genitores para a prole nos casos de cruzamentos controlados e heredogramas conhecidos, são insuficientes para entendermos a transmissão de genes de uma geração para a seguinte nas populações naturais, nas quais nem todos os indivíduos geram prole, nem toda a prole sobrevive.

Consideremos a lei da segregação simples de Mendel para um único gene com dois alelos: A e a . A modo de exemplo, o gene A que determina a presença de chifres em bovinos possui dois alelos: A e a . O alelo a determina a ausência de chifres (nos machos) e o alelo A determina a presença de chifres.

Nesse caso, dois gametas que se fundem (ou reproduzem) para formar um zigoto dão a tabela de multiplicação mostrada na Tabela 1, que na biologia é chamada de **quadrado de Punnett**.

	A	a
A	AA	Aa
a	aA	aa

TABELA 1. Alelos passando de gametas a zigotos.

Os zigotos AA e aa são chamados **homozigotos**, já que carregam duas cópias do mesmo alelo. Neste caso, não há dúvidas sobre qual informação genética será herdada na próxima geração; isto é, AA passará o alelo A e aa passará a . No entanto, os zigotos Aa e aA (que são equivalentes) possuem dois alelos diferentes. Esses zigotos são chamados **heterozigotos**. As regras da herança mendeliana simples indicam que a próxima geração herdará A ou a com a mesma frequência.

2. ÁLGEBRAS GENÉTICAS

Uma álgebra genética é uma álgebra (possivelmente não-associativa) usada para modelar a herança genética. A definição mais geral de uma álgebra que poderia ter significado genético é a de uma **álgebra com realização genética**, i.e., uma álgebra \mathcal{A} sobre \mathbb{R} que tem uma base $\{e_1, \dots, e_n\}$ e uma tabela de multiplicação

$$e_i \cdot e_j = \sum_{k=1}^n \gamma_{ijk} e_k,$$

de modo que as constantes estruturais satisfaçam

$$0 \leq \gamma_{ijk} \leq 1 \quad \text{e} \quad \sum_{k=1}^n \gamma_{ijk} = 1,$$

para todo $i, j, k = 1, \dots, n$.

Uma tal base é chamada de **base natural** para \mathcal{A} . Interpretamos os elementos da base natural $\{e_1, e_2, \dots, e_n\}$ como sendo n gametas distintos de uma população. O acasalamento aleatório dos gametas e_i e e_j produzirá gametas e_k com uma certa probabilidade γ_{ijk} .

Um elemento x numa álgebra com realização genética \mathcal{A} representa uma população se sua expressão como combinação linear dos elementos da base

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$$

satisfaz $0 \leq \alpha_i \leq 1$, para todo $i = 1, \dots, n$ e $\sum_{i=1}^n \alpha_i = 1$. Neste caso, os escalares $\alpha_i \in \mathbb{R}$ representam a percentagem da população x que carrega o alelo e_i .

A multiplicação de dois tais elementos representa o acasalamento aleatório entre duas populações.

Exemplo 2. Considere novamente a herança mendeliana simples para um único gene com dois alelos: A e a . Então, quando dois gametas se reproduzem, uma multiplicação é induzida, o que indica como a informação hereditária será passada para a próxima geração, de acordo com o quadrado de Punnett da Tabela 1. Isso dá origem à álgebra de dimensão 2 sobre \mathbb{R} , com base $\{A, a\}$ e com tabela de multiplicação:

\times	A	a
A	A	$\frac{1}{2}A + \frac{1}{2}a$
a	$\frac{1}{2}a + \frac{1}{2}A$	a

TABELA 2. Tabela de multiplicação da álgebra gamética \mathcal{G}_2 .

As constantes estruturais representam a probabilidade do alelo ser passado para a próxima geração. Chamamos esta álgebra de **álgebra gamética da herança mendeliana simples para um gene com dois alelos** e a denotamos por \mathcal{G}_2 . Note que \mathcal{G}_2 é uma álgebra com realização genética comutativa.

Observe que no exemplo anterior os produtos $AA = A$ e $aa = a$ são expressões do fato que, se ambos os gametas carregarem o mesmo alelo, os descendentes herdarão o mesmo alelo, isso é traduzido na linguagem da álgebra como “os elementos A e a são idempotentes”.

Exemplo 3. Ainda considerando a herança mendeliana simples para um único gene com dois alelos: A e a . Se os dois gametas se fundem, os zigotos terão três possíveis genótipos: AA , aa e Aa . Quando dois zigotos se reproduzem outra multiplicação é induzida dando origem à

álgebra de dimensão 3 sobre \mathbb{R} , com base $\{AA, aa, Aa\}$ e com tabela de multiplicação:

\times	AA	Aa	aa
AA	AA	$\frac{1}{2}AA + \frac{1}{2}Aa$	Aa
Aa	$\frac{1}{2}AA + \frac{1}{2}Aa$	$\frac{1}{4}AA + \frac{1}{2}Aa + \frac{1}{4}aa$	$\frac{1}{2}Aa + \frac{1}{2}aa$
aa	Aa	$\frac{1}{2}Aa + \frac{1}{2}aa$	aa

TABELA 3. Tabela de multiplicação da álgebra zigótica \mathcal{Z}_2 .

Esta álgebra é chamada de **álgebra zigótica para a herança mendeliana simples com dois alelos** e será denotada por \mathcal{Z}_2 . Observe que \mathcal{Z}_2 também é uma álgebra com realização genética.

As álgebras dos exemplos anteriores podem ser generalizadas para o caso multialélico.

Definição 4. Uma **álgebra gamética** \mathcal{G}_k correspondente a k alelos a_0, \dots, a_{k-1} em um gene, tem tabela de multiplicação dada por

$$a_i \cdot a_j = \frac{1}{2}(a_i + a_j)$$

para todo $i, j = 0, \dots, k-1$.

Podemos considerar uma mudança de base em \mathcal{G}_k fazendo $c_0 = a_0$ e $c_i = a_0 - a_i$ com $i \neq 0$ (chamamos esta base de **base canônica**) para a qual a tabela de multiplicação fica

$$c_0^2 = c_0, \quad c_0 \cdot c_i = \frac{1}{2}c_i, \quad c_i \cdot c_j = 0 \quad (i, j \neq 0).$$

Dada uma álgebra gamética \mathcal{G}_k podemos construir uma nova álgebra a partir dela através do processo de duplicação comutativa detalhado na Seção 1. Quando dois gametas se fundem o resultado será um zigoto, então se fundirmos os elementos da base canônica de uma álgebra gamética obteremos os elementos da base $d_{ij} = (c_i, c_j)$ de uma nova álgebra \mathcal{Z}_k , que chamaremos de **álgebra zigótica**. A tabela de multiplicação da álgebra zigótica relativa à base $\{d_{i,j} \mid i \leq j \text{ e } i, j = 1, 2, \dots, n\}$ é a seguinte:

$$d_{00}^2 = d_{00}, \quad d_{00} \cdot d_{0i} = \frac{1}{2}d_{0i}, \quad d_{0i} \cdot d_{0j} = \frac{1}{4}d_{ij} \quad (i, j \neq 0)$$

e os outros produtos são zero.

Podemos novamente aplicar o processo de duplicação comutativa começando com uma álgebra zigótica \mathcal{Z}_k . O resultado será outra álgebra

com relevância genética que chamamos de **álgebra copular** e denotamos por \mathcal{C}_k .

Em particular, para o caso dialélico, a álgebra copular \mathcal{C}_2 tem base $\{v, p_1, p_2, p_3, p_4, p_5\}$ e tabela de multiplicação

\times	v	p_1	p_2	p_3	p_4	p_5
v	v	$\frac{1}{2}p_1$	$\frac{1}{4}p_3$	0	0	0
p_1	$\frac{1}{2}p_1$	$\frac{1}{4}p_2$	$\frac{1}{8}p_4$	0	0	0
p_2	$\frac{1}{4}p_3$	$\frac{1}{8}p_4$	$\frac{1}{16}p_5$	0	0	0
p_3	0	0	0	0	0	0
p_4	0	0	0	0	0	0
p_5	0	0	0	0	0	0

TABELA 4. Tabela de multiplicação da álgebra copular \mathcal{C}_2 .

O significado genético dessa álgebra é que seus elementos, que são pares não ordenados de zigotos, representam os tipos de acasalamento de uma população.

Matematicamente, as álgebras que surgem na genética são estruturalmente interessantes: em geral são álgebras comutativas, porém não são associativas.

De fato, se a população P acasalar com a população Q e, em seguida, a população resultante acasalar com R , a população resultante não é a mesma que a população resultante do acasalamento de P com a população obtida do acasalamento de Q e R , i.e.:

$$(2.1) \quad (P \times Q) \times R \neq P \times (Q \times R).$$

No entanto, elas não são necessariamente álgebras de Lie, de Jordan ou alternativas mas pertencem a uma classe especial de álgebras não-associativas, as chamadas álgebras báricas.

Definição 5. Uma álgebra \mathcal{A} é chamada de **álgebra bárica** se possui uma representação de dimensão 1, ou seja, se admite um homomorfismo de álgebras não trivial

$$\omega : \mathcal{A} \rightarrow \mathbf{k}$$

chamado de **função peso**.

A função peso de uma álgebra bárica não necessariamente é única. A seguinte proposição fornece uma condição para a unicidade de ω .

Proposição 6. [9] *Seja (\mathcal{A}, ω) uma álgebra bárica. Se todos os elementos de $\ker \omega$ são nilpotentes, então ω é univocamente determinado.*

Matematicamente, a existência de um elemento idempotente numa álgebra fornece uma decomposição em soma direta da álgebra. No caso da álgebra ser bária temos o seguinte resultado:

Proposição 7. [9] *Seja (\mathcal{A}, ω) uma álgebra bária. Suponha que \mathcal{A} contém um idempotente e tal que $\omega(e) = 1$, então $\mathcal{A} = \mathbf{k}e \oplus \ker \omega$.*

Além de sua importância matemática, elementos idempotentes também tem um significado genético. Se uma população x satisfaz $x^2 = x$, isto significa que o equilíbrio genético foi alcançado após uma geração de cruzamentos aleatórios dentro da população x . Ou seja, a população x^2 tem a mesma carga genética que a população inicial x .

Um conceito que se relaciona com o de álgebra bária é o de álgebra dibária. Mas para poder introduzir este conceito precisamos apresentar previamente a seguinte álgebra:

Definição 8. Seja $\mathcal{V} = \langle f, m \rangle$ uma álgebra comutativa de dimensão 2 sobre \mathbb{R} com tabela de multiplicação: $f^2 = m^2 = 0$ e $f \cdot m = \frac{1}{2}(f + m)$. Então \mathcal{V} é chamada de **álgebra de diferenciação sexual**.

Definição 9. Uma álgebra \mathcal{A} será chamada de **dibária** se admite um homomorfismo sobrejetor $\Psi : \mathcal{A} \rightarrow \mathcal{V}$ na álgebra de diferenciação sexual \mathcal{V} .

A justificativa do termo “dibária” é dada pela seguinte proposição.

Proposição 10. *Se uma álgebra \mathcal{A} é dibária, então \mathcal{A}^2 é uma álgebra bária.*

De fato, como \mathcal{A} é dibária existe um homomorfismo sobrejetor $\Psi : \mathcal{A} \rightarrow \mathcal{V}$. Logo $\Psi(\mathcal{A}) = \mathcal{V}$ o que, pelas propriedades de homomorfismo, implica que $\Psi(\mathcal{A}^2) = \mathcal{V}^2$. Mas note que $\mathcal{V}^2 \simeq \mathbb{R}$, logo \mathcal{A}^2 é uma álgebra bária pois $\Psi : \mathcal{A}^2 \rightarrow \mathbb{R}$ é um homomorfismo sobrejetor.

No caso de uma álgebra \mathcal{A} com realização genética, fixada sua base natural $\{e_1, \dots, e_n\}$, definimos a função peso $\omega : \mathcal{A} \rightarrow \mathbb{R}$ por $\omega(e_i) = 1$, assim se $x = \sum_{i=1}^n \alpha_i e_i$ e $y = \sum_{i=1}^n \beta_i e_i$ então $\omega(x) = \sum_{i=1}^n \alpha_i$ e $\omega(y) = \sum_{i=1}^n \beta_i$ então

$$\begin{aligned}
x \cdot y &= \left(\sum_{i=1}^n \alpha_i e_i \right) \cdot \left(\sum_{j=1}^n \beta_j e_j \right) \\
&= \sum_{i=1}^n \sum_{j=1}^n (\alpha_i \beta_j) e_i \cdot e_j \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (\alpha_i \beta_j \gamma_{ijk}) e_k.
\end{aligned}$$

Aplicando ω temos

$$\begin{aligned}
\omega(x \cdot y) &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (\alpha_i \beta_j \gamma_{ijk}) \\
&= \left(\sum_{i=1}^n \alpha_i \right) \left(\sum_{j=1}^n \beta_j \right) \underbrace{\left(\sum_{k=1}^n \gamma_{ijk} \right)}_1 \\
&= \left(\sum_{i=1}^n \alpha_i \right) \left(\sum_{j=1}^n \beta_j \right) \\
&= \omega(x) \omega(y).
\end{aligned}$$

Logo ω é um homomorfismo, o que mostra que:

Teorema 11. *Toda álgebra com realização genética é uma álgebra bária.*

Outra classe de álgebra com significado genético são as chamadas álgebras de trem as quais foram introduzidas por Etherington em [1].

Definição 12. Uma álgebra bária comutativa (\mathcal{A}, ω) de dimensão n sobre um corpo \mathbf{k} é chamada de **álgebra de trem** se existem elementos $\beta_1, \dots, \beta_{r-1} \in \mathbf{k}$ tal que

$$(2.2) \quad x^r + \beta_1 \omega(x) x^{r-1} + \beta_2 \omega^2(x) x^{r-2} + \dots + \beta_{r-1} \omega^{r-1}(x) x = 0$$

para todo $x \in \mathcal{A}$, onde x^k são potências principais de x .

O polinômio $p(x) \in \mathbf{k}(\alpha_1, \dots, \alpha_n)(x)$ de menor grau que satisfaz (2.2) é chamado de *polinômio posto* de \mathcal{A} e ele também tem relevância genética.

Como o polinômio posto anula todos os elementos da álgebra, igualando o polinômio a 0 obtemos uma equação que fornece uma relação de recorrência entre as potências principais de todos os elementos da

álgebra \mathcal{A} . Se o polinômio posto tem grau r , então essa relação de recorrência indica como o pool gênico da r -ésima geração (na sequência das potências principais) está relacionado aos pools gênicos das gerações anteriores

$$(2.3) \quad x^r = -\beta_1\omega(x)x^{r-1} - \beta_2\omega^2(x)x^{r-2} - \dots - \beta_{r-1}\omega^{r-1}(x)x$$

Exemplo 13. Em [2], Etherington mostrou que uma álgebra de trem de dimensão k sobre um corpo de característica diferente de 2 e tal que o grau do polinômio posto seja igual a 2, é isomorfa a \mathcal{G}_k .

Suponha agora que $x \in \ker \omega$ então $\omega(x) = 0$ e logo (2.3) torna-se $x^r = 0$, ou seja, x é um elemento nilpotente. Isto implica, pela Proposição 6, que numa álgebra de trem \mathcal{A} a função peso ω é única.

Observe que o fato de todo elemento de $\ker \omega$ ser nilpotente não implica que a subálgebra $\ker \omega \subseteq \mathcal{A}$ seja nilpotente, esta condição motiva a seguinte definição:

Definição 14. Uma álgebra bária comutativa (\mathcal{A}, ω) de dimensão n sobre um corpo \mathbf{k} é chamada de **álgebra de trem especial** se $N = \ker \omega$ é nilpotente e todas as subálgebras $N^i \subseteq \mathcal{A}$ definidas indutivamente por $N^1 = N$, $N^i = N^{i-1}N$ para $i = 2, 3, \dots$ são ideais de \mathcal{A} .

O conceito de álgebra de trem especial foi a primeira tentativa de Etherington de fornecer um candidato à definição do que seria chamado de *álgebra genética*. O conceito deveria ser geral o suficiente para abranger a maioria das álgebras que surgem na genética. No entanto, a definição de Etherington tinha uma falha pois, embora as álgebras gaméticas sejam álgebras de trem especiais, suas duplicações comutativas (as álgebras zigóticas, copulares, etc.) não necessariamente o são. Portanto, ainda era necessária uma definição mais geral.

A definição adequada foi finalmente introduzida por Schafer em seu artigo [10] de 1949. Mas, antes de apresentá-la precisamos de alguns conceitos prévios.

Consideremos \mathcal{A} uma álgebra comutativa não-associativa sobre um corpo \mathbf{k} . Para $x \in \mathcal{A}$, seja $R_x : \mathcal{A} \rightarrow \mathcal{A}$ a multiplicação à direita por x . Como \mathcal{A} é comutativa então R_x é equivalente a L_x , a multiplicação à esquerda por x . Logo, a álgebra de transformação $T(\mathcal{A})$ de \mathcal{A} consiste de todos os polinômios de operadores de multiplicação à direita em \mathcal{A} e o operador identidade. Portanto, qualquer transformação $T \in T(\mathcal{A})$ pode ser representada como

$$(2.4) \quad T = \alpha I + f(R_{x_1}, \dots, R_{x_s})$$

onde $\alpha \in \mathbf{k}$, I é a identidade de \mathcal{A} e f um polinômio.

Se (\mathcal{A}, ω) for uma álgebra bária, então $T(\mathcal{A})$ é também uma álgebra bária. Uma função peso para $T(\mathcal{A})$ é definida por

$$(2.5) \quad \theta(T) = \omega(T(u)),$$

onde u é um elemento de peso 1 de \mathcal{A} .

Com isso já podemos apresentar a definição de álgebra genética segundo Schafer.

Definição 15. (SCHAFFER-1949, [10]) Seja (\mathcal{A}, ω) uma álgebra bária comutativa de dimensão n sobre um corpo \mathbf{k} , então \mathcal{A} é uma **álgebra genética** se a função característica $|\lambda I - T|$ de qualquer transformação $T = \alpha I + f(R_{x_1}, \dots, R_{x_s}) \in T(\mathcal{A})$ é uma função de $\omega(x_1), \dots, \omega(x_s)$.

Schafer provou que, aplicando a definição à transformação $T = R_x$, todas as álgebras genéticas (de acordo com sua definição) são também álgebras de trem.

Teorema 16. [10] *Uma álgebra genética \mathcal{A} sobre \mathbf{k} é uma álgebra de trem.*

E mais ainda,

Teorema 17. [10] *Uma álgebra de trem especial \mathcal{A} sobre \mathbf{k} é uma álgebra genética.*

Em relação à estrutura dessas álgebras, Schafer provou o seguinte resultado:

Teorema 18. [10] *O kernel $\ker \omega$ da função peso ω de uma álgebra genética \mathcal{A} sobre \mathbf{k} é nilpotente.*

A vantagem do conceito de álgebra genética de Schafer sobre o conceito de álgebra de trem especial de Etherington está em que álgebras genéticas são fechadas sob duplicação comutativa.

Teorema 19. [10] *A duplicação \mathcal{A}' de uma álgebra genética \mathcal{A} sobre \mathbf{k} é também uma álgebra genética.*

Embora a definição de Schafer tenha conseguido preencher a lacuna entre álgebras de trem e álgebras de trem especiais, a impraticabilidade de verificar a definição ainda era um problema. Por esse motivo, outra definição foi buscada. Em 1971, H. Gonshor, em seu artigo [4], deu uma definição que é equivalente à de Schafer, mas tem o benefício de ser muito mais fácil de ser utilizada e verificada.

Definição 20. (GONSHOR-1971, [4]) Uma álgebra comutativa \mathcal{A} de dimensão n sobre um corpo \mathbf{k} é uma **álgebra genética** se existe uma

base $\{e_0, e_1, \dots, e_{n-1}\}$ com tabela de multiplicação

$$e_i \cdot e_j = \sum_{k=0}^{n-1} \lambda_{ijk} e_k$$

tal que as constantes estruturais satisfaçam

$$\begin{aligned} \lambda_{000} &= 1 \\ \lambda_{0jk} &= 0 \quad \text{para } k < j \\ \lambda_{ijk} &= 0 \quad \text{para } i, j > 0 \text{ e } k \leq \max\{i, j\}. \end{aligned}$$

Tal base é chamada de **base canônica** de \mathcal{A} .

Como toda álgebra genética é uma álgebra de trem sabemos que a função peso ω existe e é única. De fato, Gonshor em [4] mostra que necessariamente $\omega(e_0) = 1$ e $\omega(e_i) = 0$ para $i > 0$.

Além disso ele justifica o uso do mesmo termo utilizado por Schafer provando o teorema a seguir.

Teorema 21. [4, Teorema 2.1] *Uma álgebra é genética no sentido de Gonshor se e somente se é genética no sentido de Schafer.*

Com a definição de Gonshor é fácil verificar que a álgebra gamética \mathcal{G}_k é uma álgebra genética e logo, pelo Teorema 19, todas suas duplicações \mathcal{Z}_k e \mathcal{C}_k são também álgebras genéticas.

Em particular, Schafer provou que \mathcal{G}_2 é uma álgebra de trem especial. Entretanto, a álgebra copular \mathcal{C}_2 , obtida duplicando a álgebra gamética duas vezes, não é uma álgebra de trem especial. De fato, $(\ker \omega)^2$ não é um ideal da álgebra. Com isso fica provado que:

- álgebras de trem especiais não são fechadas sob duplicação comutativa, e
- a recíproca do Teorema 17 não é verdadeira, i.e., existem álgebras genéticas que não são de trem especiais.

Além disso, Schafer construiu um contra-exemplo para a recíproca do Teorema 16, ou seja mostrou a existência de álgebras de trem que não são álgebras genéticas:

Considere \mathbf{k} um corpo de característica 2, então o quadrado de qualquer elemento da álgebra comutativa \mathcal{N} com base $\{v_1, v_2, v_3\}$ e tabela de multiplicação

\times	v_1	v_2	v_3
v_1	0	v_3	v_2
v_2	v_3	0	v_1
v_3	v_2	v_1	0

é zero. Seja \mathcal{A} a álgebra obtida por adjunção de um elemento unidade 1 a \mathcal{N} . Isto é, $\mathcal{A} = \mathbf{k}1 \oplus \mathcal{N}$ é a álgebra com base $\{1, v_1, v_2, v_3\}$ e tabela de multiplicação

\times	1	v_1	v_2	v_3
1	1	v_1	v_2	v_3
v_1	v_1	0	v_3	v_2
v_2	v_2	v_3	0	v_1
v_3	v_3	v_2	v_1	0

Se $x \in \mathcal{A}$ então $x = \alpha 1 + z$ com $\alpha \in \mathbf{k}$ e $z \in \mathcal{N}$. Logo $\omega : \mathcal{A} \rightarrow \mathbf{k}$ dada por $\omega(x) = \alpha$ é uma função peso para \mathcal{A} e $\mathcal{N} = \ker \omega$. Temos ainda que $z^2 = (x - \alpha 1)^2 = x^2 + \alpha^2 1 = 0$ o que implica que $x^3 + \alpha^2 x = 0$ e portanto $p(x) = x^3 + \alpha^2 x$ é um polinômio que anula-se em todo elemento de \mathcal{A} o que faz de \mathcal{A} uma álgebra de trem.

Porém, como $\mathcal{N} = \ker \omega$ não é nilpotente (pois $\mathcal{N} = \mathcal{N}^2$) segue que \mathcal{A} não é uma álgebra genética (veja o Teorema 18).

Finalmente, todas as álgebras com significado genético que apresentamos neste capítulo estão relacionadas na Figura 2.1.

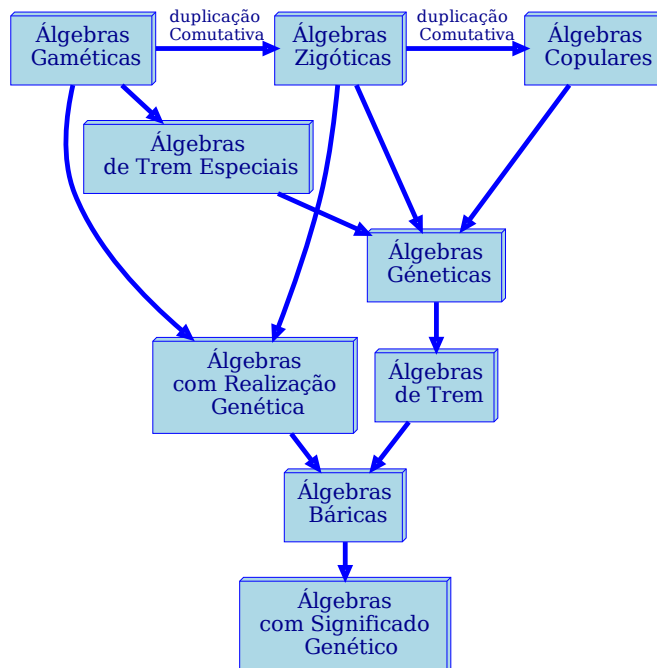


FIGURA 2.1. Álgebras com significado genético

3. APLICAÇÕES

3.1. Autofertilização. Um dos ramos da genética que estudamos neste trabalho é chamado de genética de populações, ou seja, o estudo de como as populações evoluem através das gerações.

Para um dado sistema de acasalamento, o objetivo final da genética de populações é determinar a distribuição de tipos genéticos na n -ésima geração filial, com base nos tipos genéticos presentes na população original, bem como determinar a distribuição de equilíbrio, se houver.

Consideremos o sistema de herança genética chamado de autofertilização. A **autofertilização** ocorre principalmente no mundo das plantas em que um mesmo indivíduo é capaz de gerar ambos gametas femininos e masculinos, como é o caso do amendoim, tomate, feijão. Na reprodução por autofertilização não ocorre troca de gametas entre indivíduos. É o cruzamento do indivíduo com ele próprio.

Seja $P = \alpha AA + \beta Aa + \gamma aa$ é um elemento da álgebra zigótica \mathcal{Z}_2 para a herança mendeliana simples para dois alelos (veja Exemplo 3). O nosso interesse não é estudar o que acontece quando P se acasala aleatoriamente com outras populações, focaremos nas gerações filiais resultantes após repetidas autofertilizações. Utilizando as regras da multiplicação da álgebra zigótica, a primeira geração filial $F_1 = P \times P$ depois da autofertilização será:

$$\begin{aligned} F_1 &= \alpha(AA \times AA) + \beta(Aa \times Aa) + \gamma(aa \times aa) \\ &= \alpha AA + \beta\left(\frac{1}{4}AA + \frac{1}{2}Aa + \frac{1}{4}aa\right) + \gamma aa \\ &= \left(\alpha + \frac{1}{4}\beta\right) AA + \frac{1}{2}\beta Aa + \left(\frac{1}{4}\beta + \gamma\right) aa. \end{aligned}$$

Seguindo com esse procedimento, obtemos a segunda geração filial

$$F_2 = F_1 \times F_1 = \left(\alpha + \frac{3}{8}\beta\right) AA + \frac{1}{4}\beta Aa + \left(\frac{3}{8}\beta + \gamma\right) aa.$$

Assim, a distribuição de tipos genéticos na n -ésima geração filial será

$$F_n = \left(\alpha + \frac{1}{2}\beta - \frac{1}{2^{n+1}}\beta\right) AA + \frac{1}{2^n}\beta Aa + \left(\frac{1}{2}\beta + \gamma - \frac{1}{2^{n+1}}\beta\right) aa.$$

Conforme n aumenta F_n se aproxima da distribuição de equilíbrio

$$\left(\alpha + \frac{1}{2}\beta\right) AA + \left(\gamma + \frac{1}{2}\beta\right) aa.$$

Observe que na distribuição de equilíbrio o heterozigoto Aa não aparece, portanto, a autofertilização repetida matará os heterozigotos.

Desta forma, a compreensão deste mecanismo (a autofertilização) proporciona sua aplicação em diversos programas de melhoramento genético, para obter linhagens de indivíduos com melhor e maior uniformidade.

A Figura 3.1 mostra o aumento dos homozigotos em gerações sucessivas de vários sistemas de endogamia, i.e., de união entre indivíduos aparentados, geneticamente semelhantes. A autofertilização (curva de número 1) leva a um aumento muito rápido dos homozigotos, atingindo por exemplo o 90% da população logo na terceira geração filial.

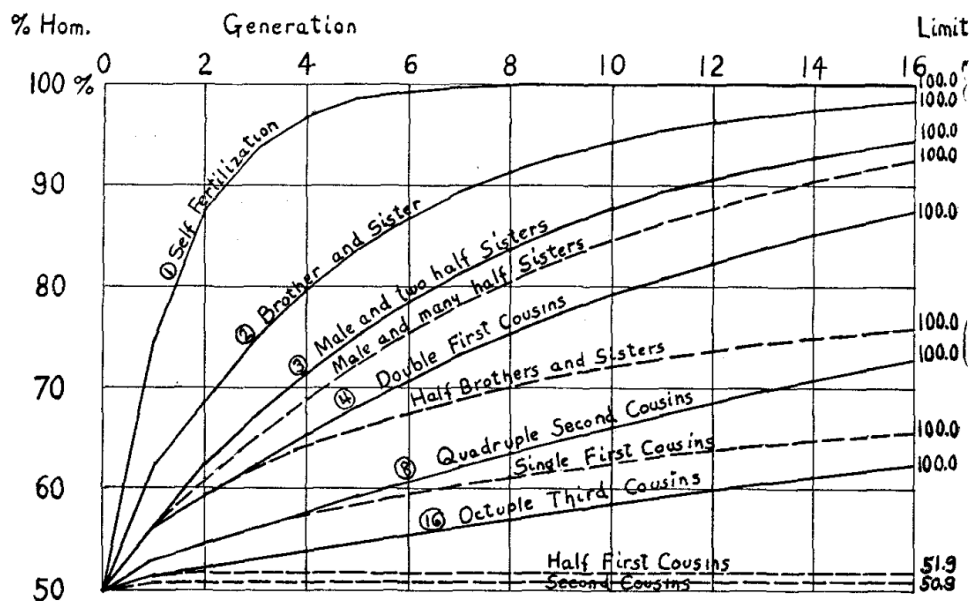


FIGURA 3.1. Aumento dos homozigotos em vários sistemas de endogamia.

3.2. Herança ligada ao sexo. Nos exemplos que vimos até agora a herança foi sempre "simétrica" no sentido que o sexo de um dos pais não afetava a distribuição dos tipos gaméticos produzidos. A simetria da herança pode ser perturbada pelo cruzamento desigual entre homens e mulheres. Da seguinte forma:

Considere uma população diploide de organismos femininos e masculinos. Os masculinos serão heterogaméticos, carregando cromossomos X e Y enquanto que os femininos são homogaméticos, ou seja, carregam dois cromossomos X . Assuma que os alelos A e a estão em um locus ligado ao sexo no cromossomo X . Os genótipos femininos

são $AA := X^A X^A$, $Aa := X^A X^a$ e $aa := X^a X^a$ e os masculinos são $AY := X^A Y$ e $aY := X^a Y$.

Para entender melhor vejamos um exemplo:

Exemplo 22. Nos humanos o daltonismo é um exemplo de característica ligada ao sexo. Trata-se de uma doença causada por um gene recessivo localizado no cromossomo X , na qual há uma alteração da visão de cores, que incapacita o indivíduo a distinguir certas cores. Em geral, há uma confusão entre as cores verde e vermelho.

Seja X^A alelo dominante que determina a visão normal e X^a alelo recessivo que carrega a doença, então temos

SEXO	GENÓTIPO	FENÓTIPO
Feminino	$X^A X^A$	Normal
Feminino	$X^A X^a$	Normal*
Feminino	$X^a X^a$	Daltônica
Masculino	$X^A Y$	Normal
Masculino	$X^a Y$	Daltônico

No caso assinalado por *, a mulher possui o gene para daltonismo, mas a doença não se manifesta, ela é portadora. Então para a descendência feminina ser daltônica ela precisa receber o gene do daltonismo tanto da mãe quanto do pai. Já a descendência masculina daltônica recebe o gene da mãe (pais não passam daltonismo para filhos).

Do ponto de vista da álgebra, assumimos que a segregação mendeliana simples ocorre em ambos os sexos, i.e.

	A	a	Y
A	A	$\frac{1}{2}A + \frac{1}{2}a$	$\frac{1}{2}A + \frac{1}{2}Y$
a	$\frac{1}{2}A + \frac{1}{2}a$	a	$\frac{1}{2}a + \frac{1}{2}Y$

TABELA 5. Multiplicação gamética para a herança ligada ao sexo.

Considere a \mathbb{R} -álgebra \mathcal{S}_2 com base $\{AA, Aa, aa, AY, aY\}$ e tabela de multiplicação como na Tabela 6.

Aqui o produto de dois elementos básicos, correspondentes aos zigotos, representa a descendência desses zigotos. Como apenas zigotos de diferentes sexos acasalam, todos os produtos dos elementos da base correspondentes a zigotos do mesmo sexo são iguais a zero. Os demais produtos formam as taxas de segregação. A álgebra \mathcal{S}_2 é chamada de

\times	AA	Aa	aa	AY	aY
AA	0	0	0	$\frac{1}{2}AA + \frac{1}{2}AY$	$\frac{1}{2}Aa + \frac{1}{2}AY$
Aa		0	0	$\frac{1}{4}AA + \frac{1}{4}Aa + \frac{1}{4}AY + \frac{1}{4}aY$	$\frac{1}{4}Aa + \frac{1}{4}aa + \frac{1}{4}AY + \frac{1}{4}aY$
aa			0	$\frac{1}{2}Aa + \frac{1}{2}aY$	$\frac{1}{2}aa + \frac{1}{2}aY$
AY				0	0
aY					0

TABELA 6. Tabela de multiplicação da álgebra zigótica para herança ligada ao sexo, \mathcal{S}_2 .

álgebra zigótica para herança ligada ao sexo para dois alelos com taxa de segregação Mendeliana simples.

Vamos introduzir agora a definição geral de uma álgebra zigótica para a herança ligada ao sexo com mais de dois tipos gaméticos. Sejam a_0, a_1, \dots, a_n os diferentes tipos gaméticos no cromossomo X . Existem $(n+1)^2$ genótipos femininos $a_i a_j$ que denotaremos por a_{ij} para todo $i, j = 0, \dots, n$ e os genótipos masculinos são dados por $a_i Y$ para $i = 0, \dots, n$. Assumimos que o genótipo feminino a_{ij} produz gametas a_k com probabilidade γ_{ijk} enquanto que o genótipo masculino $a_i Y$ produz a_k com probabilidade γ_{ik} . Na ausência de seleção e se os genótipos masculinos produzem gametas X e Y com a mesma frequência, então

$$\sum_{k=0}^n \gamma_{ijk} = 1, \quad \text{e} \quad \sum_{k=0}^n \gamma_{ik} = \frac{1}{2}.$$

Definição 23. Seja \mathcal{S}_{n+1} o espaço vetorial sobre \mathbb{R} com base $\{a_{ij}, a_i Y \mid i, j = 0, \dots, n\}$ de dimensão $(n+1)^2 + n + 1$. Definimos em \mathcal{S}_{n+1} a multiplicação

$$a_{ij} \cdot a_l Y = a_l Y \cdot a_{ij} = \sum_{k,r=0}^n \gamma_{ijk} \gamma_{lkr} a_{lr} + \frac{1}{2} \sum_{k=0}^n \gamma_{ijk} a_k Y,$$

$$a_{ij} \cdot a_l p = a_l Y \cdot a_p Y = 0,$$

para $i, j, l, p = 0, \dots, n$. Então \mathcal{S}_{n+1} é a **álgebra zigótica para a herança ligada ao sexo assimétrica com tipos gaméticos** a_0, \dots, a_n .

Enquanto a modelagem do fenômeno da herança ligada ao sexo dessa maneira parece bastante direta, existem diferenças significativas na estrutura dessas álgebras quando comparadas à herança genética, que é simétrica em relação ao sexo. Algumas dessas diferenças são enumeradas nos seguintes resultados:

Proposição 24. *A álgebra zigótica \mathcal{S}_{n+1} é uma álgebra comutativa, não-associativa e não é uma álgebra bária.*

De fato, o quadrado de alguns elementos da base é zero o que não permite definir uma função peso não nula.

Proposição 25. *A álgebra zigótica \mathcal{S}_{n+1} não é duplicação comutativa de uma álgebra gamética.*

Veja, por exemplo, no caso de \mathcal{S}_2 que $AA \times AA \neq AA$.

Proposição 26. *A álgebra \mathcal{S}_{n+1} é uma álgebra dibárica.*

O homomorfismo sobrejetor que faz de \mathcal{S}_{n+1} uma álgebra dibárica é $\Psi : \mathcal{S}_{n+1} \rightarrow \mathcal{V}$ é definido por $\Psi(a_{ij}) = f$ e $\Psi(a_i Y) = m$ para $i, j = 0, \dots, n$.

Ou seja, Ψ leva os zigotos femininos da base no elemento f (feminino) da álgebra de diferenciação sexual e os masculinos em m .

O fato de \mathcal{S}_{n+1} ser dibárica implica, pela Proposição 10, que seu quadrado \mathcal{S}_{n+1}^2 é uma álgebra bárica.

4. RELAÇÃO COM AS ÁLGEBRAS DE JORDAN

Uma classe de álgebras não-associativas bem conhecida é a classe das álgebras de Jordan. Estas últimas foram introduzidas em 1932 pelo físico alemão Pascual Jordan em uma tentativa de reformular o cenário da mecânica quântica, em termos do produto de Jordan ao invés do produto associativo. Este produto é comutativo e satisfaz uma versão debilitada da associatividade denominada “*identidade de Jordan*”.

Definição 27. Uma **álgebra de Jordan** \mathcal{J} é uma álgebra sobre um corpo \mathbf{k} na qual a multiplicação é comutativa e satisfaz

$$(4.1) \quad (a^2 \cdot b) \cdot a = a^2 \cdot (b \cdot a)$$

para todo $a, b \in \mathcal{J}$.

Seja \mathcal{A} uma \mathbf{k} -álgebra na qual a multiplicação \cdot é associativa e \mathbf{k} um corpo de característica diferente de 2. Denote por $\mathcal{A}^{(+)} = (\mathcal{A}, \odot)$ a álgebra obtida de \mathcal{A} substituindo a multiplicação por \odot definida por

$$x \odot y = \frac{1}{2}(x \cdot y + y \cdot x).$$

$\mathcal{A}^{(+)}$ é claramente comutativa e satisfaz (4.1), logo é uma álgebra de Jordan.

Definição 28. Uma álgebra isomorfa a $\mathcal{A}^{(+)}$ ou a uma subálgebra dela é chamada de **álgebra de Jordan especial**.

Não todas as álgebras de Jordan são especiais, nesse caso dizemos que a álgebra de Jordan é **excepcional**.

Um resultado bem conhecido na teoria de estrutura das álgebras de Jordan é o seguinte:

Proposição 29. [11] *Toda álgebra de Jordan é de potências associativas.*

Além disso, se adicionarmos a hipótese da álgebra ter dimensão finita pode-se mostrar também que:

Proposição 30. [11] *Se \mathcal{J} é uma álgebra de Jordan de dimensão finita tal que todo elemento é nilpotente então \mathcal{J} é uma álgebra nilpotente.*

Seja $e \in \mathcal{J}$ um elemento idempotente numa álgebra de Jordan então \mathcal{J} admite decomposição em soma direta de espaços vetoriais

$$\mathcal{J} = \mathcal{J}_1(e) \oplus \mathcal{J}_{\frac{1}{2}}(e) \oplus \mathcal{J}_0(e),$$

onde $\mathcal{J}_i(e) = \{x \in \mathcal{J} \mid e \cdot x = x \cdot e = ix\}$ com $i = 0, \frac{1}{2}, 1$.

Esta decomposição em soma de subespaços é chamada **decomposição de Peirce da álgebra de Jordan \mathcal{J} relativa ao idempotente e** . Os subespaços $\mathcal{J}_i(e)$ são chamados de **componentes** (ou subespaços) **de Peirce** da álgebra \mathcal{J} e suas propriedades multiplicativas são as seguintes:

$$\begin{aligned} \mathcal{J}_1(e)^2 &\subseteq \mathcal{J}_1(e), & \mathcal{J}_1(e) \cdot \mathcal{J}_0(e) &= (0), & \mathcal{J}_0(e)^2 &\subseteq \mathcal{J}_0(e), \\ \mathcal{J}_0(e) \cdot \mathcal{J}_{\frac{1}{2}}(e) &\subseteq \mathcal{J}_{\frac{1}{2}}(e), & \mathcal{J}_1(e) \cdot \mathcal{J}_{\frac{1}{2}}(e) &\subseteq \mathcal{J}_{\frac{1}{2}}(e), \\ \mathcal{J}_{\frac{1}{2}}(e)^2 &\subseteq \mathcal{J}_0(e) + \mathcal{J}_1(e). \end{aligned}$$

A relação entre as álgebras genéticas e as álgebras de Jordan foi observada pela primeira vez por R. D. Schafer, em seu artigo "*Structure of Genetic Algebras*" (veja [10]) de 1949. Nesse artigo Schafer provou que as álgebras gaméticas e zigóticas para a herança mendeliana simples para dois alelos, \mathcal{G}_2 e \mathcal{Z}_2 , são álgebras de Jordan especiais. De fato, elas correspondem às álgebras de Jordan de dimensões 2 e 3 sobre \mathbb{R} , denotadas por \mathcal{B}_2 e \mathcal{J}_{19} , respectivamente, no artigo [7] de M. E. Martin e I. Kashuba em que é apresentada a classificação completa a menos de isomorfismos de todas as álgebras de Jordan reais de dimensão menor ou igual a 3.

Essa propriedade estende-se também aos casos multialélicos, como foi mostrado por Holgate.

Proposição 31. [6] *A álgebra gamética \mathcal{G}_k e a álgebra zigótica \mathcal{Z}_k são álgebras de Jordan.*

Nos esquemas de acasalamento mostrados em (2.1), isso significa que as populações finais têm as mesmas proporções genéticas se P é o filho do acasalamento de R consigo mesmo.

Observamos que em 2013, M. E. Martin mostrou em [8, Teo 4.65] que as álgebras gaméticas \mathcal{G}_k para cada $k \geq 2$ são álgebras rígidas dentro da variedade Jor_k das álgebras de Jordan de dimensão k sobre um corpo de característica diferente de 2. Isto significa que elas determinam uma componente irredutível da variedade e, segue também de [8, Teo 4.65], que tal componente tem de dimensão k .

Holgate ainda mostrou que \mathcal{G}_k é sempre uma álgebra de Jordan especial, pois é isomorfa à álgebra $\mathcal{A}^{(+)}$ onde \mathcal{A} é a álgebra associativa com base $\{a_0, \dots, a_{k-1}\}$ e multiplicação $a_i \cdot a_j = a_i$. Porém, somente o caso dialético da álgebra zigótica foi provado ser uma álgebra de Jordan especial.

Em virtude da Proposição 29 sabemos que, numa álgebra de Jordan, potências de um único elemento são associativas. Logo, a álgebra copular de herança mendeliana simples \mathcal{C}_2 não é uma álgebra de Jordan pois a potência $p_1^4 = 0$, enquanto que $p_1^2 p_1^2 = 2^{-8} p_5 \neq 0$.

Estas considerações nos levam a uma análise dessas álgebras genéticas \mathcal{A} que são ao mesmo tempo álgebras de Jordan. Sobre a estrutura dessas álgebras tem-se o seguinte resultado:

Teorema 32. [10] *Sejam \mathbf{k} um corpo de característica diferente de 2, \mathcal{A} uma álgebra genética que é uma álgebra de Jordan sobre \mathbf{k} . Então \mathcal{A} contém um idempotente e , e as diferentes raízes da função característica $|\lambda I - T|$ de $T \in T(\mathcal{A})$ em (2.4) são no máximo três:*

$$\alpha, \quad \theta(T), \quad \alpha + f\left(\frac{1}{2}\omega(x_1), \frac{1}{2}\omega(x_2), \dots\right),$$

onde θ é a função peso (2.5) da álgebra bárica $T(\mathcal{A})$. A multiplicidade dessas raízes são as dimensões sobre \mathbf{k} das componentes de Peirce $\mathcal{A}_0(e)$, $\mathcal{A}_1(e)$ e $\mathcal{A}_{\frac{1}{2}}(e)$, respetivamente.

CONCLUSÃO

Este trabalho é uma introdução ao estudo das estruturas algébricas não-associativas que surgem na genética, uma das tantas aplicações da matemática que tem significado biológico. O objetivo principal é capacitar ao leitor na pesquisa e no análise de informações sobre o tema, permitindo assim uma ampliação do espectro do seu conhecimento.

Diversas perspectivas de pesquisa na área de álgebra e em especial na área de álgebras não-associativas são apresentadas. A multidisciplinaridade e a conseqüente multiplicidade de aspectos envolvidos, permitem ao leitor uma enriquecedora conexão com ideias e técnicas de diferentes áreas da matemática e da biologia.

REFERÊNCIAS

- [1] I. M. H. ETHERINGTON, *Genetic algebras*, Proc. Roy. Soc. Edinburgh, 59 (1939), pp. 242–258.
- [2] ———, *Commutative train algebras of ranks 2 and 3*, Journal of the London Mathematical Society, 15 (1940), pp. 136–149.
- [3] ———, *Duplication of linear algebras*, Proc. Edinburgh Math. Soc., 6 (1941), pp. 222–230.
- [4] H. GONSHOR, *Contributions to genetic algebras*, Proc. Edinburgh Math. Soc., 17 (1971), pp. 289–298.
- [5] A. J. F. GRIFFITHS, J. H. MILLER, D. T. SUZUKI, R. C. LEWONTIN, AND W. M. GELBART, *Introdução à genética*, Guanabara Koogan, Rio de Janeiro, 2002.
- [6] P. HOLGATE, *Jordan algebras arising in population genetics*, Proceedings of the Edinburgh Mathematical Society, 15 (1967), pp. 291–294.
- [7] I. KASHUBA AND M. E. MARTIN, *The variety of three-dimensional real Jordan algebras*, Journal of Algebra and its Applications, 15 (2016), p. 1650158 (17 pages).
- [8] M. E. MARTIN, *Deformações e isotopias de álgebras de Jordan*, tese de doutorado, Instituto de Matemática e Estatística, Universidade de São Paulo, IME-USP, São Paulo, SP, 2013.
- [9] M. L. REED, *Algebraic structure of genetic inheritance*, Bulletin (New Series) of the American Mathematical Society, 34 (1997), pp. 107–130.
- [10] R. D. SCHAFER, *Structure of genetic algebras*, American Journal of Mathematics, 71 (1949), pp. 121–135.
- [11] ———, *An Introduction to Nonassociative Algebras*, Department of Mathematics at Oklahoma State University, 1961.
- [12] A. WÖRZ-BUSEKROS, *Algebras in genetics*, Lecture Notes in Biomathematics 36, Springer-Verlag, Berlin, 1980.