

ÁLGEBRA Y GENÉTICA

Jesús Hernando Pérez

Profesor Universidad Sergio Arboleda

Bogotá D.C, Colombia

jhppelusa@gmail.com

Stefany Moreno

Estudiante Instituto Alberto Merani

Bogotá D.C, Colombia

smogam@gmail.com

Resumen

El objetivo de este documento es presentar una introducción a las relaciones que existen entre el Álgebra y la Genética, con algunos ejemplos de aplicaciones en ambos campos. Aunque se dan algunos conceptos biológicos es necesario que el lector consulte otros materiales.

Un pequeño marco biológico ...

Todos los organismos tienen material genético propio que está contenido en una molécula denominada ADN. A parte de almacenar la información, esta molécula tiene dos funciones: La primera es expresar la información para la supervivencia del organismo y la segunda es transmitir la información a las siguientes generaciones.

En este documento se introducirá una estructura algebraica para cada uno de los dos procesos, de tal manera se puedan estudiar las características matemáticas de un álgebra que define un proceso biológico.

El primer proceso a tratar se denomina síntesis de proteínas y el segundo se conoce como herencia genética.

1. Antecedentes:¹

Acerca de aplicaciones del Álgebra al primer proceso como antecedentes tenemos un solo texto cuyo nombre es “*Un álgebra de Genes para una Estructura Algebraica del código genético (Gene Algebra from a Genetic Code Algebraic Structure)*”. Este consiste básicamente en tomar el conjunto de los 64 codones del código genético (Tripletes de las bases: A, U, C, G, como estos AUG, CCA) y “copiar” dentro de este la operación de una estructura matemática con el mismo número de elementos (64) de tal manera que cada codón este asociado a un número y viceversa. De esta forma se pueden operar codones ya que en el fondo se están operando los números asociados a estos.

Respecto a la segunda aplicación el conjunto de antecedentes se puede ubicar en una línea del tiempo desde 1856 hasta nuestros días, en la cual se hacen explícitas las contribuciones hechas por cada investigación al tema. A su vez esta línea está dividida en dos partes que indican un hecho que marcó el estudio de las aplicaciones matemáticas al estudio de la herencia genética: La caracterización de la herencia como un álgebra. Se indicará el orden de este proceso histórico a continuación: Las aplicaciones de la matemática

¹Algunos de los conceptos que se utilizan en los antecedentes estarán definidos más adelante, en los apartados correspondientes a cada tema.

al estudio genético empezaron en 1856 cuando en un monasterio Mendel hizo un estudio sobre la herencia en un tipo de guisantes. Él encontró que las características pasaban a la siguiente generación con una probabilidad determinada matemáticamente y que se seguía un proceso regular tras cada generación. Después de Mendel otros autores como Jennings, Serebrovskij y Glivenko aportaron a su trabajo al darle un carácter un poco más matemático y básicamente al buscar nuevos casos diferentes a los trabajados por el monje austriaco. El corte mencionado anteriormente se efectúa en 1939 cuando Ethington plantea una estructura más compleja que en los anteriores casos conocida como la estructura de álgebra. Es decir, lo que hizo este investigador fue pasar de la búsqueda y análisis de algunos elementos matemáticos en la herencia a caracterizarla en su totalidad como un álgebra. Después del trabajo hecho por este autor considerado el “padre” de las aplicaciones del álgebra a la genética, los demás empezaron tanto a estudiar las características de la herencia desde las propiedades matemáticas de un álgebra cómo a estudiar distintos tipos de fenómenos genéticos que propiciaron el surgimiento de distintos tipos de álgebras.

2. Un álgebra de codones

2.1. Conceptos básicos en biología

Como ya se había dicho el proceso de expresión del material genético se conoce como síntesis de proteínas, pero antes de entrar mas a fondo se presentarán algunas características del ADN.

El ADN es una molécula de doble hélice que esta compuesta entre otros elementos por cuatro bases nitrogenadas: Adenina (A), Guanina (G), Citosina (C) y Timina (T). Estas bases se unen por pares (A y T, C y G) y forman la molécula en su totalidad. La mayoría de organismos vivos tienen ADN y este guarda la información genética de cada uno en particular.

La síntesis de proteínas tiene dos pasos fundamentales (A parte de la replicación de ADN): La Transcripción y la Traducción. En el primero de ellos y por cuestiones de seguridad (Debido a la importancia del ADN es mejor trabajar con una copia) la información codificada en ADN se pasa a ARN (Una molécula similar, que entre otras diferencias en vez de tener Timina tiene Uracilo). En el segundo paso se toma dicho ARN y se codifica una proteína: Para empezar hay que aclarar que cada triplete de bases codifica para un aminoácido (Por ejemplo UCG codifica para Serina). Entonces en esta parte del proceso se toman de a tres bases, se determina el aminoácido para el cual codifican y así se va formando una cadena de aminoácidos. Dicha cadena es la que forma una proteína. Para que comience el proceso de traducción es necesario tener el codon AUG y para que termine se puede tener o UAA o UAG.

2.2. Construcción de un álgebra de codones

Tenemos que todos los posibles codones del ARN son 64 ($4 \times 4 \times 4$). Con esta información se toma una estructura matemática de 64 elementos, en este caso Z_{64} y se hace una asociación en la cual cada codón se corresponde con un número dentro de dicha estructura.

Para hacer esta asociación se tienen en cuenta algunos axiomas que nos permiten definir exactamente el número al que esta asociado cada base: De acá obtenemos dos posibles asociaciones: $(A, 0)$, $(C, 1)$, $(G, 2)$, $(U, 3)$ y $(U, 0)$, $(G, 1)$, $(C, 2)$, $(A, 3)$. A continuación presentamos la tabla correspondiente a la primera asociación:

		A			C			G			U		
		No	⁽¹⁾	⁽²⁾									
A	0	AAA	K	16	ACA	T	32	AGA	R	48	AUA	I	A
	1	AAC	N	17	ACC	T	33	AGC	S	49	AUC	I	C
	2	AAG	K	18	ACG	T	34	AGG	R	50	AUG	M	G
	3	AAU	N	19	ACU	T	35	AGU	S	51	AUU	I	U
C	4	CAA	Q	20	CCA	P	36	CGA	R	52	CUA	L	A
	5	CAC	H	21	CCC	P	37	CGC	R	53	CUC	L	C
	6	CAG	Q	22	CCG	P	38	CGG	R	54	CUG	L	G
	7	CAU	H	23	CCU	P	39	CGU	R	55	CUU	L	U
G	8	GAA	E	24	GCA	A	40	GGA	G	56	GUA	V	A
	9	GAC	D	25	GCC	A	41	GGC	G	57	GUC	V	C
	10	GAG	E	26	GCG	A	42	GGG	G	58	GUG	V	G
	11	GAU	D	27	GCU	A	43	GGU	G	59	GUU	V	U
U	12	UAA	-	28	UCA	S	44	UGA	-	60	UUA	L	A
	13	UAC	Y	29	UCC	S	45	UGC	C	61	UUC	F	C
	14	UAG	-	30	UCG	S	46	UGG	W	62	UUG	L	G
	15	UAU	Y	31	UCU	S	47	UGU	C	63	UUU	F	U

Al hacer dicha correspondencia se pueden encontrar las siguientes regularidades:

- El Codón de inicio (AUG, metionina) es inverso a los dos codones de parada (UAG en la tabla de la primera asociación y UAA en la tabla de la segunda asociación).
- Los Codones Hidrofilicos (Afines al agua) son inversos a los codones Hidrofóbicos (No afines al agua). Lo que muestra una simetría en la tabla.
- Cuando dos codones codifican para el mismo aminoácido tienen la misma paridad.

2.3. Aplicaciones: resistencia a Inhibidores del VIH

En este momento el interrogante que surge es hacia los usos de plantear dichas asociaciones: Hasta el momento se ha buscado crear un inhibidor para el virus del VIH que impida que este se reproduzca por todo el cuerpo. El problema es que al reproducirse el virus esta mutando (Cambia su material genético. Para nosotros será como un cambio de codón. Por ejemplo de ACU a ACC) constantemente y a veces el inhibidor pierde su efecto. Como estamos analizando cambios de codones podemos referirnos más exactamente a una mutación como un cambio de números, al hacer esto obtenemos las siguientes conclusiones frente a mutaciones que generan que el inhibidor pierda su acción:

- Cuando el codon resultante de la mutación aumenta el orden del codon inicial la mutación genera resistencia cruzada.
- Cuando el codon resultante de la mutación cambia de paridad respecto al codon inicial la mutación genera resistencia cruzada.

Lo que define un método para determinar si una mutación en el VIH genera resistencia a un inhibidor determinado.

3. Un álgebra para la herencia genética

3.1. Conceptos básicos en biología

En este apartado se trabajara la herencia genética. Para entender este proceso el primer concepto importante a tener en cuenta es el de gen: Este es un fragmento de ADN y se constituye en la unidad básica de la herencia. Cada gen tiene un conjunto de alelos que representan las distintas formas fenotípicas (Expresión física). Por ejemplo para el gen que codifica para color de ojos tenemos que los distintos alelos pueden ser para ojos cafés, negros, azules, verdes, . . . La herencia genética en cierto tipo de organismos esta ligada a la reproducción sexual. En esta se tiene que dos padres se cruzan y transmiten su información genética al individuo resultante. Para un gen determinado tenemos dos posibles casos: Que los alelos del padre y de la madre sean los mismos o que sean diferentes. En el caso en que son los mismos se expresa el alelo compartido y en el caso en que son diferentes introduciremos los conceptos de dominancia y recesividad: Un alelo es dominante frente a otro cuando a pesar de que ambos se encuentran presentes físicamente solo se expresa uno (El concepto de recesividad se puede construir partiendo de este).

Por último es necesario tener en cuenta que hay células que tienen un sola copia del material genético y hay células que tienen dos copias. Las primeras se conocen como gametos y las segundas como cigotos. De esta forma podemos definir dos tipos de álgebras: Las Álgebras Gaméticas y las Álgebras Zigóticas. Estas como se verá más adelante darán lugar a las Álgebras con Realización Genética.

3.2. Construcción de las álgebras con realización genética

3.2.1. Álgebras gaméticas

Tomemos dos alelos $\{a_0, a_1\}$. Partiendo de que estos tienen la misma probabilidad de expresarse tenemos los siguientes resultados frente a los fenotipos:

\times	a_0	a_1
a_0	a_0	$\frac{1}{2}(a_0 + a_1)$
a_1	$\frac{1}{2}(a_0 + a_1)$	a_1

3.2.2. Álgebras zigóticas

Tomemos $\{a_0a_0, a_0a_1, a_1a_1\}$. Partiendo de que estos tienen la misma probabilidad de expresarse tenemos los siguientes resultados frente a los fenotipos:

\times	a_0a_0	a_0a_1	a_1a_1
a_0a_0	a_0a_0	$\frac{1}{2}(a_0a_0 + a_0a_1)$	a_0a_1
a_0a_1	$\frac{1}{2}(a_0a_0 + a_0a_1)$	$\frac{1}{4}a_0a_0 + \frac{1}{2}a_0a_1 + \frac{1}{4}a_1a_1$	$\frac{1}{2}(a_1a_1 + a_0a_1)$
a_1a_1	a_0a_1	$\frac{1}{2}(a_1a_1 + a_0a_1)$	a_1a_1

3.2.3. Álgebras con realización genética

Supongamos que A es un álgebra sobre R de dimensión n . Tomemos $\{a_0, a_1, \dots, a_n\}$. Tenemos que:

$$a_i a_j = \sum_{k=1}^n c_{ijk} a_k \quad 0 \leq c_{ijk} \leq 1.$$

Con

$$\sum_{k=1}^n c_{ijk} = 1 \quad \forall i, j, k = 1, \dots, n$$

Al hacer estas construcciones una de las primeras conclusiones encontradas es frente al carácter no asociativo de la herencia genética. Este se demostrara mediante el siguiente ejemplo:

$$(A \times a) \times a = \frac{1}{2}(A + a) \times a = \frac{1}{2}\left(\frac{1}{2}(A + a) + a\right) = \frac{1}{4}A + \frac{1}{4}a + \frac{1}{2}a = \frac{1}{4}A + \frac{3}{4}a$$

$$(A \times a) \times a = A \times a = \frac{1}{2}(A + a).$$

La relación entre la matemática y la Biología aun es un campo que no se ha desarrollado mucho pero que tiene un gran potencial por delante. Si se quiere conocer mas a fondo el trabajo se puede consultar la pagina www.usa.edu.co en el link Semicírculo.

Bibliografía

- [1] AUBRY, G., *Algebraic Approach to Population Genetics*. École Polytechnique Fédérale de Lausanne, July 2001.
- [2] ETHERINGTON, I., *Genetic Algebras*. Proc, Roy. Soc. Edinburg 59 (1939) 242 - 258.
- [3] GONSHOR, H., *Special Train Algebras arising in Genetics*. Proc, Edinburg Math. Soc. (2) 14 (1965) 333 - 338.
- [4] REED, M., *Algebraic structures in genetic inheritance American Mathematical Society*. Volume 34, Number 2, April 1997, Pages 107-130.

- [5] SÁNCHEZ, R., Y OTROS, *Gene Algebra from a Genetic Code Algebraic Structure*. Research Institute of Tropical Roots, Tuber Crops and Banana (INIVIT). Biotechnology Group. Santo Domingo, Villa Clara, Cuba.
- [6] WÔRZ-BUSEKROS, A., *Algebras in Genetics*. Springer Verlag, Berlin Heidelberg, New York, Lecture Notes in Biomathematics, 36. 1980.