Universidad de Puerto Rico
Departamento de Matemáticas
Humacao, Puerto Rico 00791

Programa BRIDGES — Modelaje Matemático — Dr. Pablo Negrón

Laboratorio IV: Genética

En este laboratorio estudiaremos la herencia de un rasgo a través de varias generaciones en animales o plantas. El rasgo bajo consideración se asume que se determina por un par de genes los cuales denotamos A y a. Bajo herencia *autosomal* cada individuo en la población, sin importar el sexo, posee dos de estos dos genes. Los posibles pares de genes en un individuo se denotan por AA, Aa, y aa. A este par de genes se le llama el *genotipo* del individuo y el mismo determina como el rasgo en cuestión se manifiesta. Por ejemplo en ciertos tipos de flores, el genotipo determina el color de la flor. En los humanos, el color de los ojos se rige por herencia autosomal. En particular AA y Aa producen ojos marón, mientras que aa produce ojos azules. En este caso decimos que A es *dominante* y que a es *receptivo*.

También existe lo que se llama herencia *X–entrelazada*. En este tipo de herencia el macho de la especie posee uno solo de los dos posibles genes (A ó a), y la hembra posee un par de genes (AA, Aa, ó aa). En los humanos ciertas enfermedades se heredan de acuerdo al modelo de herencia X–entrelazada.

Conceptos Básicos de Algebra Lineal

Pausamos un momento para discutir unos conceptos del algebra de matrices que necesitaremos para la exposición. En el laboratorio anterior, en el contexto de balanceo de reacciones químicas, vimos lo que son *matrices*, i.e., arreglos rectangulares de numeros. A las matrices con una sola columna se les llama usualmente *vectores*. Las matrices se pueden sumar, multiplicar por un numero ó por otra matriz. Ilustramos estas operaciones con un ejemplo. Considere las matrices A, B y el vector x dados por:

$$A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 1 & -4 \\ 6 & -2 \end{bmatrix} \quad , \quad x = \begin{bmatrix} 4 \\ -5 \end{bmatrix}.$$

Entonces

$$A + B = \begin{bmatrix} 2+1 & 3-4 \\ -1+6 & 4-2 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 5 & 2 \end{bmatrix},$$

$$3A = \begin{bmatrix} 3(2) & 3(3) \\ 3(-1) & 3(4) \end{bmatrix} = \begin{bmatrix} 6 & 9 \\ -3 & 12 \end{bmatrix} \quad , \quad -5x = \begin{bmatrix} -5(4) \\ -5(-5) \end{bmatrix} = \begin{bmatrix} -20 \\ 25 \end{bmatrix},$$

$$Ax = \begin{bmatrix} (2)(4) + (3)(-5) \\ (-1)(4) + (4)(-5) \end{bmatrix} = \begin{bmatrix} -7 \\ -24 \end{bmatrix}.$$

Si definimos las matrices A, B y el vector x de arriba en MATLAB, podemos hacer estos cálculos en MATLAB escribiendo $A + B$, $3 * A$, $-5 * x$, y $A * x$. **Nota**: ¡No todas las matrices se pueden multiplicar!

Dada una matriz A, cualquier numero λ y vector $x \neq 0$ tal que $Ax = \lambda x$, se llama un *valor propio* y *vector propio* respectivamente de la matriz A. Por ejemplo:

$$\lambda = 2 \quad , \quad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

es un par de valor propio y vector propio de

$$A = \begin{bmatrix} 2 & 4 \\ 0 & -3 \end{bmatrix},$$

ya que

$$Ax = \begin{bmatrix} (2)(1) + (4)(0) \\ (0)(1) + (-3)(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2x.$$

Ejercicio 0.1. Usando MATLAB calcule las operaciones matriciales indicadas.

a) Halle $A + 3B$ y AB si

$$A = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix}.$$

b) Usando la función `eig` de MATLAB calcule los valores y vectores propios de la matriz

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Herencia Autosomal

En herencia autosomal un individuo hereda un gene de cada uno de sus parientes con igual probabilidad para así formar su propio genotipo. Si uno de los padres tiene el genotipo Aa, entonces el individuo puede heredar de ese padre un gene A ó a con igual probabilidad. Si uno de los padres es del genotipo Aa y el otro aa, entonces siempre hereda un gene del tipo a y el otro puede ser A ó a con igual *probabilidad*. En la Tabla (1) mostramos todos los posibles genotipos con sus probabilidades para un individuo dependiendo de los genotipos de los padres.

Veamos el caso de un agricultor que tiene una población grande de plantas con una cierta distribución de los genotipos AA, Aa, y aa. El agricultor planea establecer un programa de cultivo de plantas donde cada planta de la población siempre se fertiliza con una planta del genotipo AA. Queremos hallar una expresión para la distribución de los tres posibles genotipos en la población de plantas despúes de varias generaciones. Para esto definimos:

$a_n =$ fracción de plantas del genotipo AA en la n–esima generación;

$b_n =$ fracción de plantas del genotipo Aa en la n–esima generación;

$c_n =$ fracción de plantas del genotipo aa en la n–esima generación.

Aquí a_0, b_0, c_0 representan la distribución inicial de de los genotipos. Note que para toda n:

$$a_n + b_n + c_n = 1.$$

Utilizando la Tabla (1) podemos ahora escribir las recurrencias:

$$\begin{array}{rcl} a_n & = & a_{n-1} + \frac{1}{2}b_{n-1}, \\ b_n & = & \frac{1}{2}b_{n-1} + c_{n-1}, \\ c_n & = & 0. \end{array} \tag{1}$$

Tenemos ahora que estas recurrencias se pueden escribir en forma matricial como:

$$x^{(n)} = Mx^{(n-1)} \quad , \quad n = 1, 2, \dots , \tag{2}$$

donde

$$M = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad , \quad x^{(n)} = \begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix} \quad , \quad x^{(n-1)} = \begin{bmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{bmatrix}. \tag{3}$$

Al igual que cuando resolvimos la recurrencia del modelo de crecimiento de células, podemos resolver (2) como:

$$x^{(n)} = M^n x^{(0)} \quad , \quad n = 1, 2, \dots , \tag{4}$$

Genotipo del hijo	Genotipos de los Padres					
	$AA - AA$	$AA - Aa$	$AA - aa$	$Aa - Aa$	$Aa - aa$	$aa - aa$
AA	1	$\frac{1}{2}$	0	$\frac{1}{4}$	0	0
Aa	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0
aa	0	0	0	$\frac{1}{4}$	$\frac{1}{2}$	1

Tabla 1: Probabilidades de los diferentes tipos de genotipos de un individuo dependiendo de los genotipos de los padres bajo herencia autosomal.

donde $x^{(0)}$ representa las distribución inicial de genotipos. Si x_1, x_2, x_3 son vectores propios de M (*linealmente independientes*) correspondientes a los valores propios $\lambda_1, \lambda_2, \lambda_3$, entonces se puede verificar que de (4) obtenemos que:

$$x^{(n)} = \alpha_1 \lambda_1^n x_1 + \alpha_2 \lambda_2^n x_2 \alpha_3 \lambda_3^n x_3, \tag{5}$$

donde $\alpha_1, \alpha_2, \alpha_3$ son constantes tales que

$$x^{(0)} = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3.$$

Podemos ahora usar la función `eig` de MATLAB para buscar los valores y vectores propios de la matriz M en (3). Escribiendo

```
M=[1,0.5,0;0,0.5,1;0,0,0]
[V,D]=eig(M)
```

obtenemos que los vectores propios son aproximadamente:

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad x_2 = \begin{bmatrix} -0.7071 \\ 0.7071 \\ 0 \end{bmatrix} \quad , \quad x_3 = \begin{bmatrix} 0.4082 \\ -0.8165 \\ 0.4082 \end{bmatrix},$$

con los valores propios correspondientes

$$\lambda_1 = 1 \quad , \quad \lambda_2 = 0.5 \quad , \quad \lambda_3 = 0.$$

Usando estos resultados en (5) obtenemos que:

$$x^{(n)} = \begin{bmatrix} \alpha_1 + \alpha_2(-0.7071)(0.5)^n \\ \alpha_2(0.7071)(0.5)^n \\ 0 \end{bmatrix} \quad , \quad n = 1, 2, \ldots ,$$

esto es:

$$a_n = \alpha_1 + \alpha_2(-0.7071)(0.5)^n \quad , \quad b_n = \alpha_2(0.7071)(0.5)^n \quad , \quad c_n = 0 \quad , \quad n = 1, 2, \ldots .$$

De aquí vemos que para n suficientemente grande,

$$a_n \approx \alpha_1 \quad , \quad b_n \approx 0 \quad , \quad c_n = 0,$$

i.e., la población es esencialmente del genotipo AA.

Suponga ahora que en lugar de fertilizar siempre con el genotipo AA, cada planta se fertiliza con una de su mismo genotipo. Definiendo a_n, b_n, c_n como antes y haciendo referencia nuevamente a la Tabla (1), tenemos que ahora

$$x^{(n)} = C^n x^{(0)} \quad , \quad n = 1, 2, \ldots , \tag{6}$$

donde

$$C = \begin{bmatrix} 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & 1 \end{bmatrix}. \tag{7}$$

Usando MATLAB obtenemos que los vectores propios de C son:

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad , \quad x_3 = \begin{bmatrix} -0.4082 \\ 0.8165 \\ -0.4082 \end{bmatrix},$$

con los valores propios correspondientes

$$\lambda_1 = 1 \quad , \quad \lambda_2 = 1 \quad , \quad \lambda_3 = 0.5.$$

Sustituyendo en (5) tenemos que:

$$\begin{aligned} a_n &= \alpha_1 + \alpha_3(-0.4082)(0.5)^n, \\ b_n &= \alpha_3(0.8165)(0.5)^n, \\ c_n &= \alpha_2 + \alpha_3(-0.4082)(0.5)^n \quad , \quad n = 1, 2, \ldots . \end{aligned}$$

De aquí que para n bien grande,

$$a_n \approx \alpha_1 \quad , \quad b_n \approx 0 \quad , \quad c_n \approx \alpha_2,$$

i.e., la población tiende a una con los genotipos AA y aa únicamente.

Ejercicio 0.2. Suponga que las plantas se fertilizan siempre con otra planta del genotipo Aa en lugar del AA como ya trabajamos. Escriba las ecuaciones de recurrencia para este caso y determine los genotipos que sobreviven segń el numero de generaciones se hace bien grande.

Herencia X–Entrelazada

Como mencionamos en la introducción, en herencia X–entrelazada, el macho de la especie posee un solo gene del tipo A ó a y la hembra tiene dos genes con los posibles genotipos AA, Aa, y aa. Un individuo hereda genes de sus parientes de acuerdo a lo siguiente:

- si el hijo es macho, este recibe un solo gene de los dos que posee la madre con igual probabilidad;
- si el hijo es hembra, entonces esta recibe un gene de su padre y uno de los dos genes de la madre con igual probabilidad.

En la Tabla (2) se muestran las probabilidades de los diferentes genotipos del hijo dependiendo el sexo y los genotipos de los parientes.

Consideramos ahora un programa de apareamiento donde se comienza con un macho y una hembra; de sus hijos se selecciona al azar un macho y una hembra y estos se cruzan; seleccionamos un macho y una hembra de los nuevos hijos y los cruzamos, etc.. Este tipo de apareamiento se utiliza con animales para mantener "pura" la especie.

El par original puede tener uno de los siguientes seis posibles pares de de genotipos:

$$(A, AA), (A, Aa), (A, aa), (a, AA), (a, Aa), (a, aa).$$

Los pares de hijos seleccionados sucesivamente generación tras generación pueden igualmente tener uno de estos seis pares con ciertas probabilidades. Definimos ahora las siguientes cantidades:

$a_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (A, AA);

$b_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (A, Aa);

$c_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (A, aa);

$d_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (a, AA);

$e_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (a, Aa);

$f_n =$ probabilidad de que el par de hijos cruzados en la n–esima generación sea del tipo (a, aa);

Genotipo del hijo		Genotipos de los Padres (Padre,Madre)					
		(A, AA)	(A, Aa)	(A, aa)	(a, AA)	(a, Aa)	(a, aa)
Macho	A	1	$\frac{1}{2}$	0	1	$\frac{1}{2}$	0
	a	0	$\frac{1}{2}$	1	0	$\frac{1}{2}$	1
Hembra	AA	1	$\frac{1}{2}$	0	0	0	0
	Aa	0	$\frac{1}{2}$	1	1	$\frac{1}{2}$	0
	aa	0	0	0	0	$\frac{1}{2}$	1

Tabla 2: Probabilidades de los diferentes tipos de genotipos de un individuo dependiendo de los genotipos de los padres bajo herencia X–entrelazada.

Veamos cual es la recurrencia que rige las a_n. La única forma de obtener un par de hijos del tipo (A, AA) es cruzando padres con el tipo (A, AA) ó padres del tipo (A, Aa). Así que la probabilidad de obtener un par de hijos del tipo (A, AA) es:

$$\begin{aligned}
a_n &= (\text{prob. de hijos tipo } (A, AA) \text{ dado que los padres son } (A, AA)) \\
&\quad \times(\text{prob. padres tipo } (A, AA)) \\
&\quad +(\text{prob. de hijos tipo } (A, AA) \text{ dado que los padres son } (A, Aa)) \\
&\quad \times(\text{prob. padres tipo } (A, Aa)) \\
&= (1)(1)a_{n-1} + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) b_{n-1} = a_{n-1} + \frac{1}{4}\, b_{n-1}.
\end{aligned}$$

De igual forma obtenemos las siguientes restantes recursiones:

$$\begin{aligned}
b_n &= \tfrac{1}{4}\, b_{n-1} + d_{n-1} + \tfrac{1}{4}\, e_{n-1}, \\
c_n &= \tfrac{1}{4}\, e_{n-1}, \\
d_n &= \tfrac{1}{4}\, b_{n-1}, \\
e_n &= \tfrac{1}{4}\, b_{n-1} + c_{n-1} + \tfrac{1}{4}\, e_{n-1}, \\
f_n &= \tfrac{1}{4}\, e_{n-1} + f_{n-1}.
\end{aligned}$$

Tenemos ahora con

$$x^{(n)} = \begin{bmatrix} a_b \\ b_n \\ c_n \\ d_n \\ e_n \\ f_n \end{bmatrix} \quad , \quad n = 0, 1, 2, \ldots,$$

que

$$x^{(n)} = M x^{(n-1)} \quad , \quad n = 1, 2, \ldots\ ,$$

donde

$$M = \begin{bmatrix}
1 & \frac{1}{4} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{4} & 0 & 1 & \frac{1}{4} & 0 \\
0 & 0 & 0 & 0 & \frac{1}{4} & 0 \\
0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{4} & 1 & 0 & \frac{1}{4} & 0 \\
0 & 0 & 0 & 0 & \frac{1}{4} & 1
\end{bmatrix}.$$

Usando MATLAB podemos ver que M tiene un valor propio $\lambda = 1$ repetido dos veces con vectores propios:

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

y los otros valores propios son menor de uno en valor absoluto. Tenemos pues que para n bien grande:

$$x^{(n)} \approx \alpha_1 x_1 + \alpha_2 x_2,$$

i.e.,

$$a_n \approx \alpha_2 \quad , \quad b_n \approx 0 \quad , \quad c_n \approx 0 \quad , \quad d_n \approx 0 \quad , \quad e_n \approx 0 \quad , \quad f_n \approx \alpha_2.$$

O sea que despúes de un numero grande de generaciones, el par de hijos seleccionados será del tipo (A, AA) ó (a, aa).

Ejercicio 0.3. Suponga que ninguno de los genotipos femeninos Aa sobreviven para entre–cruzarse. Los únicos posibles pares de genotipos son entonces

$$(A, AA), (A, aa), (a, AA), (a, aa).$$

Halle el sistema de recurrencias que describen la transición de una generación a otra en este caso.

Referencias

[1] Anton, H. and Rorres, C., Elementary Linear Algebra: Applications Version, John Wiley & Sons, Inc., New York, 1994.