Journal Of Harmonized Research (JOHR)

Journal Of Harmonized Research in Applied Sciences
6(4), 2018, 236-240

ISSN 2321 – 7456

Original Research Article

APPLICATIONS OF LINEAR ALGEBRA IN GENETICS

M. Naga Maruthi Kumari[1], Kavya.H.U[2] and Krupashree .Y.N[3]

Research and Development Centre, Reva University, Bengaluru

Abstract: The present paper is focusing on the Mathematical concepts from Linear Algebra including Eigen values and eigen vectors and disgonalization to predict the genotype distribution. This paper is based on finding the genotype distribution for n^{th} generation in Autosomal inheritance.

Key Words: Eigen values, Eigen vectors, Diagonalization, Autosomal inheritance, Genotype, genotype distribution.

1. Introduction: Linear algebra is also quite useful in variety of real world problems, including genetics. Genetics is the branch of biology that deals with the heredity and variation of organisms.

Gregor Mendel is known as father of genetics. [Who Studied trait inheritance by controlled breeding of ~30,000 pea plants between 1856 – 1863].

To find the genotype distribution of a particular trait in a population. After any number of generation from only genotype distribution of the initial population.

Jared Kirkham (2001) worked on genotypic distribution of flower population in n^{th} generation considering auto somal inheritance. John gallego, Sunit kambli and Daniel lee worked on the probability of the peppered moth to survive and predict genotype distribution in future generation. [Due to the industrial revolution the environment will get pollute in which the moths are affected by this].

If we know the distribution of the present generation, we can use the transition matrix A. of the population to find the genotype distribution of the future generation.

Here we find genotype distribution of n^{th} generation by using linear algebraic concepts.

Let x_o be the initial genotype distribution. A be the transition matrix for the genotype of the given population.

One year later, the genotype of the population x_1 is equal to product of transition matrix A and the initial matrix x_o.

$$X_1 = AX_o$$

Two year later, the genotype of population can be

For Correspondence:
suma@reva.edu.in.
Received on: August 2018
Accepted after revision: September 2018
DOI: 10.30876/JOHR.6.4.2018.236-240

$$X_2 = AX_1 = A(AX_o)$$
$$= A^2X_o$$

Three years later, the genotype of the population can be

$$X_3 = AX_2 = A(A^2Xo)$$
$$= A^3Xo$$

And n years later,

$$X_n = A^nX_o$$

A common problem in this field finding the probability of certain genotype after a given number of years.

Suppose we want to study the fractions of 3 genotypes in n^{th} generation of cows in terms of initial genotype fractions.

2. Pegegree Analysis: In this paper we will focus on autosomal inheritance

Definition 2.1: [Mendelian and complex disease models]

Autosomal inheritance is a pattern of inheritance in which the transmission of traits depends on the presence of the alleles on the autosomes

Definition 2.2: [Wilhelm Johannsen's 1909]

Genotype is the pair of alleles at locus [example: AA, Aa, aa].

Phenotype is the observable characteristic or trait

Definition 2.3 : [David C.Lay 5^{th} edition]

Diagonalization: for the given transition square matrix A is said to be diagnalizable

If $A = PDP^{-1,}$ where

P = modle matrix obtained by the eigen vectors

P^{-1} = inverse of an modle matrix

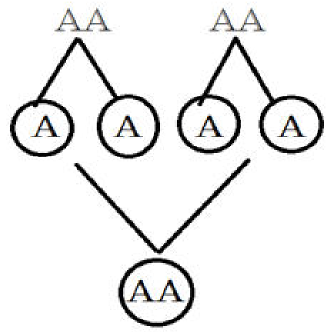
D = diagonal matrix

Definition 2.4: Pedigree analysis: the chart of the family histories that show the phenotypes and family relationships of the individuals.

3. **Genotypes in the n^{th} generation:** To find the fractions of three genotypes in n^{th} generation, we have done a prediction by the following method:

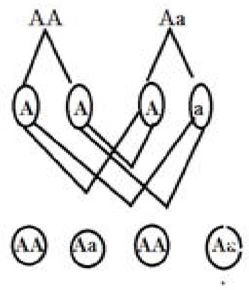
Result 3.1: If cows with genotype AA can produce better quality milk than other genotypes .if we interested to discovering the fraction of offspring cows with genotype AA. If we chose to cross only genotype AA with other genotypes, then the probabilities of offspring bring which kind of Geno types? AA, Aa or aa?

Solution:



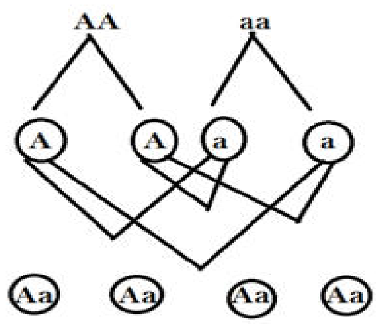
Genotype	AA	Aa	aa
probability	1	0	0

Therefore, the probabilities of an offspring to be AA, Aa, aa are 1, 0, 0.



Geno type	AA	Aa	aa
probability	½	½	0

The offspring will have half chance to be of genotype AA and half chance the genotype Aa Therefore, probabilities of AA, Aa and aa respectively are 1/2, ½, 0



Genotype	AA	Aa	aa
probability	0	1	0

This will always results in genotype Aa.
Therefore, the probabilities of genotype of genotypes AA, Aa, and aa respectively are 0, 1, 0 respectively.
The following matrix A is the result of the previous observation.

$$A=\begin{bmatrix}1 & 1/2 & 0\\ 0 & 1/2 & 1\\ 0 & 0 & 0\end{bmatrix}$$

Assume that the initial population of cows made up of an equal number of each genotype.
Therefore, initial distribution vector X_0 is given

$$X_0 = \begin{bmatrix}1/3\\ 1/3\\ 1/3\end{bmatrix}$$

After one year the distribution is

$$X_1 = AX_0 = \begin{bmatrix}1/3(1/2+1)\\ 1/3(1+1/2)\\ 0\end{bmatrix}$$

After another year passed by the distribution vector can obtained as follows

$$X_2 = AX_1 = A\,(AX_0) = \begin{bmatrix}3/4\\ 1/4\\ 0\end{bmatrix}$$

Suppose if we interested to see the number of cows with genotype AA , Aa, and aa after 30 generations , one way to compute A^2X_0 in this way we may get some error during multiplication.
Therefore , by using the diagonalization of the matrix method we can reduce the highest power of the matrix A. And we know that matrix A can be written as

$$A = PDP^{-1}$$

Where P = modle matrix obtained by the eigen vectors.
P^{-1} = inverse of an modle matrix
D = diagonal matrix
Therefore, A^n = can be written as

$$A^n = PDP^{-1} \quad \text{for } n=1,2,3......$$

$$D^n = \begin{bmatrix}
\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \lambda_3 & . & . & . & . & . & . & 0\\
0 & 0 & . & \lambda_4 & . & . & . & . & . & 0\\
0 & 0 & . & . & \lambda_5 & . & . & . & . & 0\\
0 & 0 & . & . & . & . & . & . & . & 0\\
0 & 0 & . & . & . & . & . & . & . & 0\\
0 & 0 & . & . & . & . & . & . & . & 0\\
. & . & . & . & . & . & . & . & . & 0\\
. & . & . & . & . & . & . & 0 & 0 & \lambda_n
\end{bmatrix}.$$

For the matrix A the eigen values are

$$\lambda 1 = 1,\quad \lambda 2 = \frac{1}{2},\quad \lambda 3 = 0$$

The corresponding eigen vectors are

$$V1 = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix},\quad V2 = \begin{bmatrix}1\\ -1\\ 0\end{bmatrix},\quad v3 = \begin{bmatrix}1\\ -2\\ 1\end{bmatrix}$$

Therefore, the diagonal matrix

$$D=\begin{bmatrix}1 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & 0\end{bmatrix}$$

$$P=\begin{bmatrix}1 & 1 & 1\\ 0 & -1 & -2\\ 0 & 0 & 1\end{bmatrix}$$

$$P^{-1} = \begin{bmatrix}1 & 0 & 0\\ 1 & -1 & 0\\ 1 & -2 & 0\end{bmatrix}$$

$PDP^{-1}X_0$

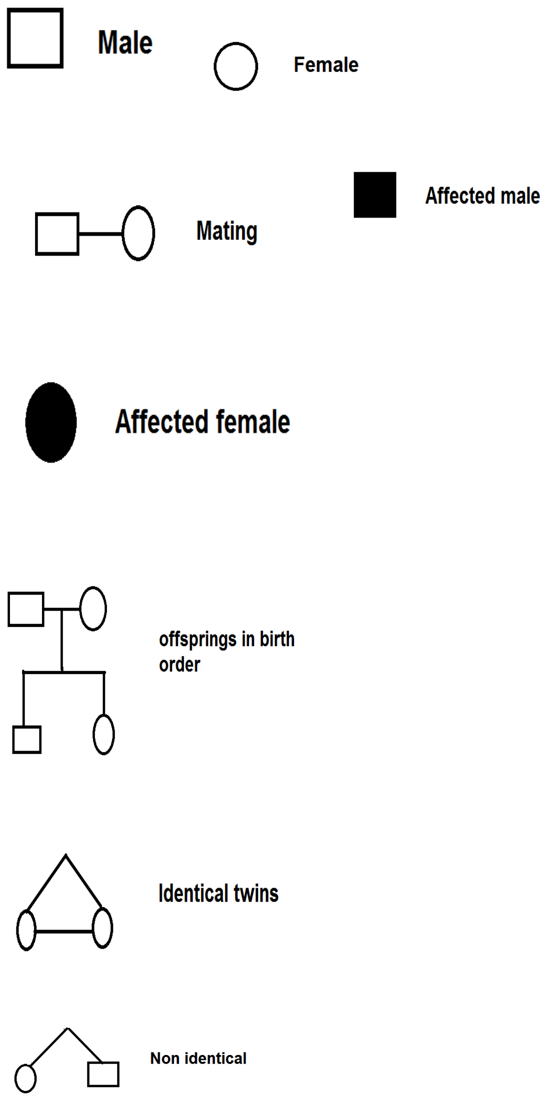
$$= \begin{bmatrix}1 & 1 & 1\\ 0 & -1 & -2\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 1 & -1 & 0\\ 1 & -2 & 0\end{bmatrix}\begin{bmatrix}1/3\\ 1/3\\ 1/3\end{bmatrix}$$

$$= \begin{bmatrix}1/6\\ 0\\ 0\end{bmatrix}$$

Therefore, by calculating eigen values and eigen vectors the genotypic distribution of cows with nth generation can be caluculated.

Result 3.2: Pedigree analysis

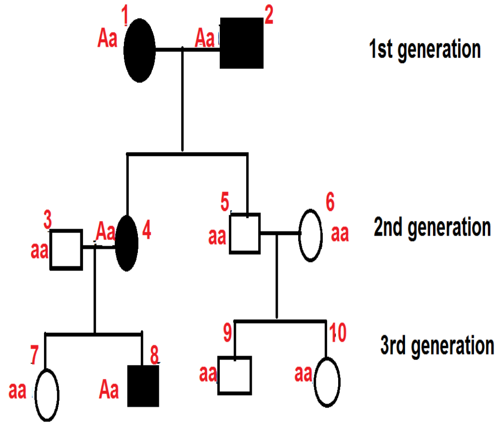
The following symbols used to represent the different aspects of the pedigree:



Non identical twins are slanting lines but without horizontal lines.

Autosomal dominant gene disease: "The individual with any of the dominant allele present will be affected"

The following pedigree chart represents the affected individuals, by taking this as an example we will find the probability distribution of an affected individuals at nth generation.



In the first generation both the parents are affected but in the second generation one offspring is unaffected this is typical nature of the dominant pedigree.

For the autosomal dominant trait diseased gene is A.

The three chances of genotypes are:

AA - diseased; Aa - diseased; aa - healthy

Therefore, the healthy individuals in the above pedigree chart should have genotype aa rest of the affected individuals should have genotype either Aa or AA.

Case 1: if 7^{th} individual in the above pedigree chart marries 9^{th} individual with genotype aa then the probability of the individual to be affected are zero (0). Because neither of them are affected (they does not contain any dominant allele).

7/9	A	a
a	Aa	aa
a	Aa	aa

The probability distribution of affected individuals is

Genotype	AA	Aa	aa
probability	0	0	0

Case 2: If 8th individual in the above chart marries 10th

8/10	A	a
A	Aa	Aa
a	Aa	aa

Aa - diseased; aa - not diseased

Therefore, the probability of affected individuals are 2/4 = 1/2 = 50%

The probability distribution is

Genotype	AA	Aa	aa
probability	0	1/2	0

Case 3: if 10th individual marries some affected individual with genotype AA then,

10/ with some AA	A	A
a	Aa	Aa
a	Aa	Aa

The probabilities are 4/4 = 1

The probability distribution is

Genotype	AA	Aa	aa
Probability	0	1	0

The initial matrix is $X_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$

The eigen values for the above transition matrix M are

$\lambda 1 = 0$; $\lambda 2 = 0$; $\lambda 3 = ½$

The corresponding eigen vectors are,

$V1 = V2 = V3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Therefore, $PDnP^{-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Therefore, the individuals are not affected.

4. Conclusion:

4.1: one of the main uses of diagonalization is we can easily find the genotypic distribution of the nth generation by computing the highest power of the matrix.

4.2: one of the major applications to find the genotype distribution in an nth generation in problem number 3.1 is that, the off springs in future generation can produce better quality of milk.

4.3: Pedigree analysis is used for various purposes like criminal, legal, scientific etc.....

4.4: Pedigree is used to determine the mode of transmission of a genetic diseases whether it is dominant or recessive or whether it is x-linked or y-linked or autosomal or it is maternal.

4.5: Doctors and scientist have used pedigrees to study human genetics long before the technologies existed to sequence DNA.

5 . Reference:

1) Principles of Mendelian inheritance
2) Linear algebra in genetics, Jared Kirkham math 308 project [2001]
3) Linear algebra and genetics John Gallego, Sunit Kambli, Daniel LeeMath 247 – Dr. Jen Mei Chang
4) Matrix approach to genetics -
5) Linear algebra and its applications David.C.Lay (5th edition)
6) Elementary linear algebra [Howard anton, chris rorres]