Jeremy Jakob

Nathan Rogers

Linear Algebra

May 1, 2017

Linear Algebra and its applications in Genetics

* Using principles and techniques learned in this class we can show how certain traits can be passed from generation to generation.
* We can create matrices for different genetic variants, of AA, Aa, and aa varieties and find ways to predict outcomes over long periods of time.
* This is a well-documented study on how genetic variation can occur in cases where environmental change is present and there are selective pressures at play.
* During the industrial revolution the peppered moth which came in two varieties, dark and light colored, experienced selective pressures to retain the darker color because it made it more difficult for predators to see them on the soot stained trees. This made it so that the dominant alleles were darker coloring and the recessive alleles were for the lighter color.
* Every organism has a genotype, which is a genetic makeup of an organism and consists of three pairs:
	+ AA (Dominant)
	+ Aa (Recessive)
	+ aa (Recessive)
* It was discovered that it is possible to model the variation in terms of these alleles using linear algebra.
* With the environmental constraints, it was observed that:
* A moth with genotype AA will have a dark color as its phenotype.
* A moth with genotype Aa will also have a dark color as its phenotype.
* A moth with genotype aa will have a light color as its phenotype.
* There are nine possible permutations:
* Three with AA as the dominant allele
* Three with Aa as the dominant allele
* Three with aa as the dominant allele.
* With the permutations, we can create matrices to calculate the probabilities of a descendent moth
* For the first case where AA is the dominant allele, there are three possible outcomes of cross breeding two moths:
	+ AA with AA
		- The probability the offspring will have genotype AA is 1
		- The probability the offspring will have genotype Aa is 0
		- The probability the offspring will have genotype aa is 0
	+ AA with Aa
		- The probability the offspring will have genotype AA is ½
		- The probability the offspring will have genotype Aa is ½
		- The probability the offspring will have genotype aa is 0
	+ AA with aa
		- The probability the offspring will have genotype AA is 0
		- The probability the offspring will have genotype Aa is 1
		- The probability the offspring will have genotype aa is 0
* Using the information, we can create a matrix to calculate the probability of the genotype of the resulting offspring.
* The columns represent breeding with AA, Aa, and aa, respectively.
* The rows represent the probability of the offspring being AA, Aa, and aa, respectively.
* $A=\left[\begin{matrix}1&1/2&0\\0&1/2&1\\0&0&0\end{matrix}\right]$
* We can find a population distribution of the next generation $x\_{n+1}$ by finding the product of our matrix and the previous generation:
	+ $Ax\_{n}=b$; where $b=x\_{n+1}$
* If we start with an evenly distributed population, $x\_{0}= \left[\begin{matrix}1/3\\1/3\\1/3\end{matrix}\right]$, and want to know the distribution of the next generation $x\_{1}$, we would get:
	+ $\left[\begin{matrix}1&1/2&0\\0&1/2&1\\0&0&0\end{matrix}\right]\left[\begin{matrix}1/3\\1/3\\1/3\end{matrix}\right]=\left[\begin{matrix}1/2\\1/2\\0\end{matrix}\right]$
* If we want to know what the distribution will be in two generations, $x\_{2}$ we can apply the same equation, using $x\_{1}$.
	+ $x\_{2}=Ax\_{1}=\left[\begin{matrix}3/4\\1/4\\0\end{matrix}\right]$
* However, this requires us to calculate the previous generation. What if we wanted to know what the distribution of just the 10th generation? Using this equation, we would have to calculate every previous generation to be able to calculate the 10th.
* However, if we use substitution, we can skip calculating every previous generation.
	+ $x\_{1}=Ax\_{0}$
	+ $x\_{2}=A\left(Ax\_{0}\right)\rightarrow x\_{2}=A^{2}x\_{0}=\left[\begin{matrix}3/4\\1/4\\0\end{matrix}\right]$
* From this we can see that for any generation $x\_{n}=A^{n}x\_{0}$.
* As $n$ approaches ∞, the distribution with $A$ is:
	+ $x\_{\infty }=\left[\begin{matrix}1\\0\\0\end{matrix}\right]$
* Another way to calculate future generations, is to use diagonalization.
* If we show the matrix as the product of an invertible matrix $P$, a diagonal matrix of eigenvalues $D$, and the inverse of $P$, then the our equation becomes:

$$x\_{n}=PD^{n}P^{-1}x\_{0}$$

* From $x\_{n}=PD^{n}P^{-1}x\_{0}$ It follows that:
* P is a matrix composed of the eigenvectors, [v1][v2][v3].
* P = $\left[\begin{matrix}1&1&1\\0&-1&-2\\0&0&1\end{matrix}\right]$
* D is a matrix with the eigenvalues as its diagonals.
* D = $\left[\begin{matrix}1&0&0\\0&1/2&0\\0&0&0\end{matrix}\right]$
* $P^{-1}$ turns out to be P.
* Also, as n approaches infinity D = $\left[\begin{matrix}1&0&0\\0&1/2&0\\0&0&0\end{matrix}\right]$ becomes D = $\left[\begin{matrix}1&0&0\\0&0&0\\0&0&0\end{matrix}\right]$
* And $x\_{n}=PD^{n}P^{-1}x\_{0}$ gives the matrix $\left[\begin{matrix}1\\0\\0\end{matrix}\right]$
* This indicates that as n approaches infinity there is a 100% chance for offspring to have AA, a 0% chance to have Aa and a 0% chance to have aa. This same approach can be used in the other two cases where Aa and aa are the dominant traits.

Works Cited

http://web.csulb.edu/~jchang9/m247/m247\_sp10\_Danial\_John\_Sunny.pdf