

MATRIX POWERS AND MARKOV MATRICES (11/02/2020)

Review:

Remember that an eigenvector/eigenvalue is a vector \vec{x} and scalar λ such that for matrix A , we find $A\vec{x} = \lambda x$. When we diagonalize a matrix, we take a matrix with columns of eigenvectors X and a matrix with the eigenvalues on the diagonal Λ to show $A = X\Lambda X^{-1}$. We will use these definitions to talk about matrix powers, and then use matrix powers to define and explore Markov matrices.

Matrix Powers:

Think about the $n \times n$ matrix A and the diagonalization $X\Lambda X^{-1}$. Consider A^2 . It follows:

$$A^2 = (X\Lambda X^{-1})^2 = (X\Lambda X^{-1})(X\Lambda X^{-1}) = X\Lambda\Lambda X^{-1} = X\Lambda^2 X^{-1}$$

Furthermore, a quick induction proof will allow us to generalize this pattern! We find that for any $m \in \mathbb{N}^+$, $A^m = X\Lambda^m X^{-1}$. This is easier to solve than A^m because Λ is a diagonalized matrix! We see that Λ^m is just the same thing as Λ but with each individual element taken to the m -th power. This can also be easily found via a quick induction proof.

A Markov Example:

Using this ability to calculate matrix powers, we can define something really cool – Markov matrices! To do so, consider this little example. Let's say there are only two countries: France and Germany. Every year, 95% of people in France stay in France, and 5% of people in France leave to Germany, while 90% of people in Germany stay in Germany, and 10% of people in Germany leave to France. Considering an arbitrary number of initial Germans and initial French, we see we can develop a recurrence relation! Let's say $u_F(0)$ French people and $u_G(0)$ German people are our initial values. It follows then that after one year, $u_F(i) = 0.95u_F(0) + 0.1u_G(0)$ and that $u_G(i) = 0.05u_F(i) + 0.9u_G(i)$. Looking at these numbers, we can generate a simple recurrence relation! Using matrix multiplication, we see that:

$$\begin{pmatrix} u_F(i+1) \\ u_G(i+1) \end{pmatrix} = \begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix} \begin{pmatrix} u_F(i) \\ u_G(i) \end{pmatrix}$$

Now we have this nice recursive formula. But recursive formulas are really ugly and hard to work with. Also, it's such a simple equation that we should be able to find something better:

$$\begin{pmatrix} u_F(i) \\ u_G(i) \end{pmatrix} = \begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix}^i \begin{pmatrix} u_F(0) \\ u_G(0) \end{pmatrix}$$

Now, we can easily find our populations over time with a nice equation! Through repeated multiplication, we can keep going and keep finding new values each year. However, before continuing we must note an important property about this matrix and formula. For all i , $u_F(i) + u_G(i) = u_F(0) + u_G(0)$. This is because nothing is lost or gained in our model. The total amount of people in France and Germany always stays the same.

Taking the computation again and again, we see that we will always reach an equilibrium:

$$\begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix}^\infty \begin{pmatrix} u_F(0) \\ u_G(0) \end{pmatrix} = k \begin{pmatrix} 2 \\ 1 \end{pmatrix}, k \in \mathbb{R}$$

This vector, $(2, 1)$, is what we will refer to as a steady state. No matter what initial values we give it, we are going to find that we'll always end up with this same end behavior! We can try to find the steady state using something we are all too familiar with – eigenvectors! Our steady

state when multiplied by the Markov matrix should equal itself. Otherwise it is not "steady!" Thus, we can construct the formula for steady state (s_F, s_G) :

$$\begin{pmatrix} s_F \\ s_G \end{pmatrix} = \begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix} \begin{pmatrix} s_F \\ s_G \end{pmatrix}$$

This formula looks an awful lot like an eigenvector/eigenvalue equation. Actually, this is one, with the eigenvalue being 1. So, if we solve for the eigenvector \vec{x} given the eigenvalue $\lambda = 1$, we have our steady state!

We know that our Markov matrix is 2×2 . Thus, why is the eigenvalue 1 going to be used, while the other eigenvalues are ignored? It all comes down to the fact that 1 is the largest eigenvalue. Think about our change of basis and diagonalization of Markov matrices – we were able to write $M\vec{v}$ as the sum of a linear transformation of the eigenvalues times the eigenvectors. As we take M to a higher and higher power, it simply results in taking the eigenvalues to a higher and higher power! Thus, the largest eigenvalue will exponentially dominate the equation, and we only need to bother with the largest eigenvalue for large enough powers.

Defining Markov Matrices

So, you may notice that while we started using the idea of Markov matrices, we never actually defined them! Well, a Markov matrix M has two central properties. First, every item of M is greater than or equal to zero. Second, the sum of the elements in each column must be one. Using this properties, it follows that for a vector of only ones $\vec{1}$, that $M^T \vec{1} = \vec{1}$. Furthermore, Markov matrices describe probability maps. Each column represents the population in a single "space," and each row represents the probability each population will move to that specific "space."

Positive Markov Matrices

What's nice about Markov matrices is that they will almost always "work out" just like our example. As long our Markov matrix M has eigenvector \vec{x}_1 and corresponding eigenvalue $\lambda = 1$, and all other eigenvalues $|\lambda| < 1$, then $k\vec{x}_1$ is the attracting steady state! Here, k is the total initial population, and the sum of the elements of \vec{x}_1 add to one (for scaling). Attracting steady state means that no matter what population we start with, we get this end steady state. The properties described in the last paragraph are always satisfied by positive Markov matrices. These matrices are Markov matrices with no zero entries. From just these properties, we can conclude that the largest eigenvalue must be 1, and thus we have an attractive steady state.

Conclusion/Looking Forward:

We have defined a lot of things today and showed a lot of really cool patterns! However, this is still not anywhere close to a "full" understanding of Markov matrices. We will begin to prove all of these properties of Markov matrices next class. Also, we never applied the findings of the "Matrix Powers" section to Markov matrices. We will show why they are important in next class as well!