Lecture 24: Markov matrices; Fourier series

In this lecture we look at Markov matrices and Fourier series – two applications of eigenvalues and projections.

Eigenvalues of A^T

The eigenvalues of A and the eigenvalues of A^T are the same:

$$(A - \lambda I)^T = A^T - \lambda I,$$

so property 10 of determinants tells us that $\det(A - \lambda I) = \det(A^T - \lambda I)$. If λ is an eigenvalue of A then $\det(A^T - \lambda I) = 0$ and λ is also an eigenvalue of A^T.

Markov matrices

A matrix like:

$$A = \begin{bmatrix} .1 & .01 & .3 \\ .2 & .99 & .3 \\ .7 & 0 & .4 \end{bmatrix}$$

in which all entries are non-negative and each column adds to 1 is called a *Markov matrix*. These requirements come from Markov matrices' use in probability. Squaring or raising a Markov matrix to a power gives us another Markov matrix.

When dealing with systems of differential equations, eigenvectors with the eigenvalue 0 represented steady states. Here we're dealing with powers of matrices and get a steady state when $\lambda = 1$ is an eigenvalue.

The constraint that the columns add to 1 guarantees that 1 is an eigenvalue. All other eigenvalues will be less than 1. Remember that (if A has n independent eigenvectors) the solution to $\mathbf{u}_k = A^k \mathbf{u}_0$ is $\mathbf{u}_k = c_1 \lambda_1^k \mathbf{x}_1 + c_2 \lambda_2^k \mathbf{x}_2 + \cdots + c_n \lambda_n^k \mathbf{x}_n$. If $\lambda_1 = 1$ and all others eigenvalues are less than one the system approaches the steady state $c_1 \mathbf{x}_1$. This is the $\mathbf{x}_1$ component of $\mathbf{u}_0$.

Why does the fact that the columns sum to 1 guarantee that 1 is an eigenvalue? If 1 is an eigenvalue of A, then:

$$A - 1I = \begin{bmatrix} -.9 & .01 & .3 \\ .2 & -.01 & .3 \\ .7 & 0 & -.6 \end{bmatrix}$$

should be singular. Since we've subtracted 1 from each diagonal entry, the sum of the entries in each column of $A - I$ is zero. But then the sum of the rows of $A - I$ must be the zero row, and so $A - I$ is singular. The eigenvector $\mathbf{x}_1$ is in the nullspace of $A - I$ and has eigenvalue 1. It's not very hard to find $\mathbf{x}_1 = \begin{bmatrix} .6 \\ 33 \\ .7 \end{bmatrix}$.

We're studying the equation $\mathbf{u}_{k+1} = A\mathbf{u}_k$ where A is a Markov matrix. For example u_1 might be the population of (number of people in) Massachusetts and u_2 might be the population of California. A might describe what fraction of the population moves from state to state, or the probability of a single person moving. We can't have negative numbers of people, so the entries of A will always be positive. We want to account for all the people in our model, so the columns of A add to $1 = 100\%$.

For example:

$$\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_{t=k+1} = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} \begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_{t=k}$$

assumes that there's a 90% chance that a person in California will stay in California and only a 10% chance that she or he will move, while there's a 20% percent chance that a Massachusetts resident will move to California. If our initial conditions are $\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$, then after one move $\mathbf{u}_1 = A\mathbf{u}_0$ is:

$$\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_1 = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = \begin{bmatrix} 200 \\ 800 \end{bmatrix}.$$

For the next few values of k, the Massachusetts population will decrease and the California population will increase while the total population remains constant at 1000.

To understand the long term behavior of this system we'll need the eigenvectors and eigenvalues of $\begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix}$. We know that one eigenvalue is $\lambda_1 = 1$. Because the trace $.9 + .8 = 1.7$ is the sum of the eigenvalues, we see that $\lambda_2 = .7$.

Next we calculate the eigenvectors:

$$A - \lambda_1 I = \begin{bmatrix} -.1 & .2 \\ .1 & -.2 \end{bmatrix} \mathbf{x}_1 = \mathbf{0},$$

so we choose $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The eigenvalue 1 corresponds to the steady state solution, and $\lambda_2 = .7 < 1$, so the system approaches a limit in which $2/3$ of 1000 people live in California and $1/3$ of 1000 people are in Massachusetts. This will be the limit from any starting vector $\mathbf{u}_0$.

To know how the population is distributed after a finite number of steps we look for an eigenvector corresponding to $\lambda_2 = .7$:

$$A - \lambda_2 I = \begin{bmatrix} .2 & .2 \\ .1 & .1 \end{bmatrix} \mathbf{x}_1 = \mathbf{0},$$

so let $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

From what we learned about difference equations we know that:

$$\mathbf{u}_k = c_1 1^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 (.7)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

When $k = 0$ we have:

$$\mathbf{u}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

so $c_1 = \frac{1000}{3}$ and $c_2 = \frac{2000}{3}$.

In some applications Markov matrices are defined differently – their rows add to 1 rather than their columns. In this case, the calculations are the transpose of everything we've done here.

Fourier series and projections

Expansion with an orthonormal basis

If we have an orthonormal basis $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n$ then we can write any vector $\mathbf{v}$ as $\mathbf{v} = x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + \cdots + x_n\mathbf{q}_n$, where:

$$\mathbf{q}_i^T\mathbf{v} = x_1\mathbf{q}_i^T\mathbf{q}_1 + x_2\mathbf{q}_i^T\mathbf{q}_2 + \cdots + x_n\mathbf{q}_i^T\mathbf{q}_n = x_i.$$

Since $\mathbf{q}_i^T\mathbf{q}_j = 0$ unless $i = j$, this equation gives $x_i = \mathbf{q}_i^T\mathbf{v}$.

In terms of matrices, $\begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{v}$, or $Q\mathbf{x} = \mathbf{v}$. So $\mathbf{x} = Q^{-1}\mathbf{v}$. Because the q_i form an orthonormal basis, $Q^{-1} = Q^T$ and $\mathbf{x} = Q^T\mathbf{v}$. This is another way to see that $x_i = \mathbf{q}_i^T\mathbf{v}$.

Fourier series

The key idea above was that the basis of vectors $\mathbf{q}_i$ was orthonormal. Fourier series are built on this idea. We can describe a function $f(x)$ in terms of trigonometric functions:

$$f(x) = a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + b_2\sin 2x + \cdots.$$

This *Fourier series* is an infinite sum and the previous example was finite, but the two are related by the fact that the cosines and sines in the Fourier series are orthogonal.

We're now working in an infinite dimensional vector space. The vectors in this space are functions and the (orthogonal) basis vectors are 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, ...

What does "orthogonal" mean in this context? How do we compute a dot product or *inner product* in this vector space? For vectors in $\mathbb{R}^n$ the inner product is $\mathbf{v}^T\mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n$. Functions are described by a continuum of values $f(x)$ rather than by a discrete collection of components v_i. The best parallel to the vector dot product is:

$$f^T g = \int_0^{2\pi} f(x)g(x)\,dx.$$

We integrate from 0 to 2π because Fourier series are periodic:

$$f(x) = f(x + 2\pi).$$

The inner product of two basis vectors is zero, as desired. For example,

$$\int_0^{2\pi} \sin x \cos x\,dx = \frac{1}{2}(\sin x)^2\Big|_0^{2\pi} = 0.$$

How do we find a_0, a_1, etc. to find the coordinates or *Fourier coefficients* of a function in this space? The constant term a_0 is the average value of the function. Because we're working with an orthonormal basis, we can use the inner product to find the coefficients a_i.

$$\begin{aligned}\int_0^{2\pi} f(x)\cos x\,dx &= \int_0^{2\pi} (a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + \cdots)\cos x\,dx \\ &= 0 + \int_0^{2\pi} a_1\cos^2 x\,dx + 0 + 0 + \cdots \\ &= a_1\pi.\end{aligned}$$

We conclude that $a_1 = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} f(x)\cos x\,dx$. We can use the same technique to find any of the values a_i.