Markov

September 7, 2017

1 Markov matrices

A matrix A is a **Markov matrix** if

- Its entries are all ≥ 0
- Each **column**'s entries **sum to 1**

Typicaly, a Markov matrix's entries represent **transition probabilities** from one state to another.
For example, consider the 2×2 Markov matrix:

```
In [1]: A = [0.9 0.2
             0.1 0.8]

Out[1]: 2×2 Array{Float64,2}:
         0.9  0.2
         0.1  0.8
```

Let us suppose that this represents the fraction of people switching majors each year between math and English literature.
Let

$$x = \begin{pmatrix} m \\ e \end{pmatrix}$$

represent the number of math majors m and English majors e. Suppose that each year, 10% of math majors and 20% of English majors switch majors. After one year, the new number of math and English majors is:

$$m' = 0.9m + 0.2ee' = 0.1m + 0.8e$$

But this is equivalent to a matrix multiplication! i.e. the numbers x' of majors after one year is

$$x' = Ax$$

Note that the two Markov properties are critical: we never have negative numbers of majors (or negative probabilities), and the probabilities must sum to 1 (the net number of majors is not changing: we're not including new students or people that graduate in this silly model).

1.1 Eigenvalues of Markov matrices

There are two key questions about Markov matrices that can be answered by analysis of their eigenvalues:

- Is there a **steady state**?
- i.e. is there an $x_0 \neq 0$ such that $Ax_0 = x_0$?
- i.e. is there $\lambda_0 = 1$ eigenvector x_0?

- Does the system **tend toward a steady state?**
- i.e. does $A^n x \to$ multiple of x_0 as $n \to \infty$?
- i.e. is $\lambda = 1$ the **largest** $|\lambda|$?

The answers are **YES** for **any Markov** matrix A, and **YES** for any *positive* Markov matrix (Markov matrices with entries > 0, not just ≥ 0). For *any* Markov matrix, all of the λ satisfy $|\lambda| \leq 1$, but if there are zero entries in the matrix we *may* have multiple $|\lambda| = 1$ eigenvalues (though this doesn't happen often in practical Markov problems).

```
In [2]: eigvals(A)
```

```
Out[2]: 2-element Array{Float64,1}:
         1.0
         0.7
```

To see why, the key idea is to write the columns-sum-to-one property of Markov matrices in linear-algebra terms. It is equivalent to the statement:

$$\underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \end{pmatrix}}_{o^T} A = o^T$$

since this is just the operation that sums all of the rows of A. Equivalently, if we transpose both sides:

$$A^T o = o$$

i.e. o is an eigenvector of A^T (called a **left eigenvector of A**) with eigenvalue $\lambda = 1$.

But since A and A^T have the **same eigenvalues** (they have the same characteristic polynomial $\det(A - \lambda I) = \det(A^T - \lambda I)$ because transposed don't change determinants), this means that A **also has an eigenvalue 1** but with a **different eigenvector**.

```
In [3]: o = [1,1]
        o' * A
```

```
Out[3]: 1×2 Array{Float64,2}:
         1.0  1.0
```

```
In [4]: A' * o
```

```
Out[4]: 2-element Array{Float64,1}:
         1.0
         1.0
```

The eigenvector of A with eigenvalue 1 must be a basis for $N(A - I)$:

```
In [5]: A - 1*I
```

```
Out[5]: 2×2 Array{Float64,2}:
         -0.1   0.2
          0.1  -0.2
```

By inspection, $A - I$ is singular here: the second column is -2 times the first. So, $x_0 = (2, 1)$ is a basis for its nullspace, and is the steady state:

```
In [6]: (A - I) * [2,1]
```

```
Out[6]: 2-element Array{Float64,1}:
         5.55112e-17
         5.55112e-17
```

Let's check if some arbitrary starting vector $(3,0)$ tends towards the steady state:

```
In [7]: using Interact
        @manipulate for n in slider(0:100,value=0)
            A^n * [3,0]
        end
```

```
Interact.Slider{Int64}(Signal{Int64}(0, nactions=1),"",0,0:100,"horizontal",true,"d",true)
```

```
Out[7]: 2-element Array{Float64,1}:
         3.0
         0.0
```

Yes! In fact, it tends to exactly $(2,1)$, because the other eigenvalue is < 1 (and hence that eigenvector component decays exponentially fast).

An interesting property is that the **sum of the vector components is conserved** when we multiply by a Markov matrix. Given a vector x, $o^T x$ is the sum of its components. But $o^T A = o^T$, so:

$$o^T Ax = o^T x = o^T A^n x$$

for any n! This is why $(3,0)$ must tend to $(2,1)$, and not to any other multiple of $(2,1)$, because both of them sum to 3. (The "number of majors" is conserved in this problem.)

1.2 Why no eigenvalues > 1?

Why are all $|\lambda| \leq 1$ for a Markov matrix?

The key fact is that the **product AB of two Markov matrices A and B is also Markov**. Reasons:

- If A and B have nonnegative entries, AB does as well: matrix multiplication uses only $\times$ and $+$, and can't introduce a minus sign.
- If $o^T A = o^T$ and $o^T B = o^T$ (both have columns summing to 1), then $o^T AB = o^T B = o^T$: the columns of AB sum to 1.

For example, A^n is a Markov matrix for any n if A is Markov.

Now, if there were an eigenvalue $|\lambda| > 1$, the matrix A^n would have to *blow up exponentially* as $n \to \infty$ (since the matrix times that eigenvector, or any vector with a nonzero component of that eigenvector, would blow up). But since A^n is Markov, all of its entries must be between 0 and 1. It can't blow up! So we must have all $|\lambda| \leq 1$.

```
In [8]: A^100
```

```
Out[8]: 2×2 Array{Float64,2}:
         0.666667  0.666667
         0.333333  0.333333
```

(In fact, A^n is pretty boring for large n: it just takes in any vector and redistributes it to the steady state.)

1.3 Can there be more than one steady state?

We have just showed that we have *at least one* eigenvalue $\lambda = 1$, and that *all* eigenvalues satisfy $|\lambda| \le 1$. But can there be *more than one* independent eigenvector with $\lambda = 1$?

Yes! For example, the **identity matrix** I is a Markov matrix, and *all* of its eigenvectors have eigenvalue 1. Since $Ix = x$ for *any* x, *every vector is a steady state* for I!

But this does not usually happen for *interesting* Markov matrices coming from real problems. In fact, there is a theorem:

- If all the entries of a Markov matrix are > 0 (not just ≥ 0), then *exactly one* of its eigenvalues $\lambda = 1$ (that eigenvalue has "multiplicity 1": $N(A - I)$ is one-dimensional), and **all other eigenvalues have** $|\lambda| < 1$. There is a *unique steady state* (up to an overall scale factor).

I'm not going to prove this in 18.06, however.

1.4 Can the solutions oscillate?

If you have a Markov matrix with zero entries, then there might be more than one eigenvalue with $|\lambda| = 1$, but these additional solutions might be *oscillating* solutions rather than steady states.

For example, consider the permutation matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

that simply swaps the first and second entries of any 2-component vector.

If $x = (1, 0)$, then $P^n x$ will oscillate forever, never reaching a steady state! It simply oscillates between $(1, 0)$ (for even n) and $(0, 1)$ (for odd n):

```
In [9]: P = [0 1
             1 0]

Out[9]: 2×2 Array{Int64,2}:
         0  1
         1  0
```

```
In [10]: [P^n * [1,0] for n = 0:5]

Out[10]: 6-element Array{Array{Int64,1},1}:
          [1,0]
          [0,1]
          [1,0]
          [0,1]
          [1,0]
          [0,1]
```

But this is a Markov matrix, so all $|\lambda|$ are ≤ 1:

```
In [11]: eigvals(P)

Out[11]: 2-element Array{Float64,1}:
          -1.0
           1.0
```

The problem is that the $\lambda = -1$ eigenvalue corresponds to an oscillating solution:

$$P^n \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (-1)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

for the eigenvector $(1,-1)$.

The steady state still exists, corresponding to the eigenvector $(1,1)$:

$$P^n \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

```
In [12]: eig(P)[2] # the eigenvectors
```

```
Out[12]: 2×2 Array{Float64,2}:
          -0.707107  0.707107
           0.707107  0.707107
```

Since $(1,0) = [(1,1)+(1,-1)]/2$, we have:

$$P^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right]$$

which alternates between $(1,0)$ and $(0,1)$.

1.5 Another example

Let's generate a random 5x5 Markov matrix:

```
In [13]: M = rand(5,5) # random entries in [0,1]
```

```
Out[13]: 5×5 Array{Float64,2}:
          0.629234   0.686579  0.181409   0.140297  0.172192
          0.423335   0.345782  0.983745   0.609905  0.16914
          0.966665   0.996079  0.0627286  0.372804  0.443826
          0.0259794  0.260556  0.123107   0.383512  0.763029
          0.968473   0.673163  0.618425   0.434404  0.954041
```

```
In [14]: sum(M,1) # not Markov yet
```

```
Out[14]: 1×5 Array{Float64,2}:
          3.01369  2.96216  1.96941  1.94092  2.50223
```

```
In [15]: M = M ./ sum(M,1)
```

```
Out[15]: 5×5 Array{Float64,2}:
          0.208792    0.231783   0.0921133  0.0722838  0.0688154
          0.140471    0.116733   0.499511   0.314235   0.0675956
          0.320758    0.336268   0.0318514  0.192076   0.177372
          0.00862046  0.0879614  0.0625094  0.197593   0.30494
          0.321358    0.227254   0.314015   0.223813   0.381277
```

```
In [16]: sum(M,1)
```

```
Out[16]: 1×5 Array{Float64,2}:
          1.0  1.0  1.0  1.0  1.0
```

```
In [17]: eigvals(M)
```

```
Out[17]: 5-element Array{Complex{Float64},1}:
                  1.0+0.0im
            -0.310196+0.0im
             0.151282+0.139568im
             0.151282-0.139568im
          -0.0561211+0.0im
```

In [18]:
```
abs.(eigvals(M))
```

Out[18]:
```
5-element Array{Float64,1}:
 1.0
 0.310196
 0.205829
 0.205829
 0.0561211
```

In [19]:
```
x = rand(5)
x = x / sum(x)
```

Out[19]:
```
5-element Array{Float64,1}:
 0.176392
 0.119515
 0.188738
 0.231065
 0.28429
```

In [20]:
```
M^100 * x
```

Out[20]:
```
5-element Array{Float64,1}:
 0.126478
 0.212976
 0.20221
 0.155533
 0.302803
```

In [21]:
```
λ, X = eig(M)
X[:,1] / sum(X[:,1])
```

Out[21]:
```
5-element Array{Complex{Float64},1}:
 0.126478+0.0im
 0.212976+0.0im
  0.20221+0.0im
 0.155533+0.0im
 0.302803+0.0im
```

In [ ]: