Math 225 lecture 5 sept 15^{th} 2023

<u>Goal</u>: Students should be able to list characteristics of a Markov matrix, probability vector, and steady state vector. In addition to being able to find a steady state vector.

Section 4.6

<u>Class Q</u>: What makes a vector "steady state" for a Markov process?

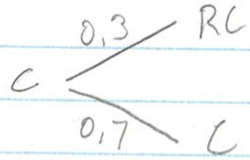
application of diagonalization, Markov process

Ex: this time of year its cold (C) or really colder (RC)

lets assume C today means tommorow will be C 70% of the time & RC the other 30%.

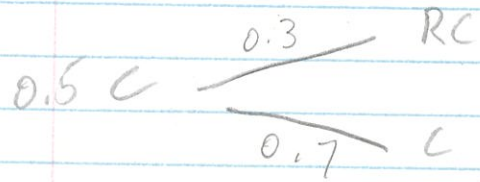
similarly if RC today tommorow will be C 20% of the time and RC the other 80%.

visually if its C today what is the weather tommorow?



suppose therse a 50% chance of C today (& 50% RC) then overall theres what chance of C tommorow?

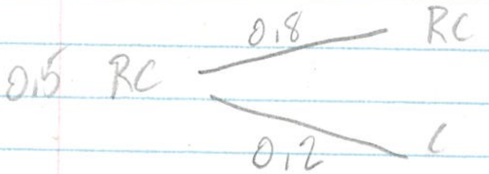
Today tomorrow



$(0.5)(0.7)$

$+$ $= 0.45$ or 45% chance of C tommorow

$(0.5)(0.2)$



There must be a better way! use linear algebra

let $A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$ this is a "Markov matrix" i.e. a square matrix with entries ≥ 0 and whose columns sum to one

aside

Notation: a vector $v = \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}$ is called a probability vector

if i) $p_i \geq 0$ ii) $p_1 + \dots + p_n = 1$ meaning a Markov matrix is a matrix with columns which are probability vectors.

let $v_0 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be the initial probabilities. (i.e. prob of weather today x_1% C x_2% RC)

$A v_0 = \begin{bmatrix} 0.3x_1 + 0.2x_2 \\ 0.7x_1 + 0.8x_2 \end{bmatrix}$ which is also a probability vector which is the prob. of the weather tommorow!

this leads us to $A^k v_0$ = probability of weather in k days given v_0 today

hard to compute? use eigen basis/diagonalization

suppose we write $v_0 = c_1 w_1 + c_2 w_2$ for w_1, w_2 eigen vetors of A with eigen values λ_1, λ_2

☆ then $A^k v_0 = A^k(c_1 w_1 + c_2 w_2) = c_1 A^k w_1 + c_2 A^k w_2 = c_1 \lambda_1^k w_1 + c_2 \lambda_2^k w_2$

return to example

what are the eigen values of A? $A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$

$$P(\lambda) = \begin{vmatrix} 0.7-\lambda & 0.2 \\ 0.3 & 0.8-\lambda \end{vmatrix} = 0.56 - 1.5\lambda + \lambda^2 - 0.06$$

$$= \lambda^2 - \tfrac{3}{2}\lambda + \tfrac{1}{2} = (\lambda - 1)(\lambda - \tfrac{1}{2})$$

$\therefore \lambda = 1, \lambda = \frac{1}{2}$

by last class all distinct eigenvalues $\Rightarrow$ all alg mults $=1$ $\Rightarrow$ all geo mults $=1$

$\Rightarrow$ A is diagonalizable

by diagonalization theorem

C

$$E_1 = null(A - 1I) = \begin{bmatrix} -0.3 & 0.2 \\ 0.3 & -0.2 \end{bmatrix} \xrightarrow{REF} \begin{bmatrix} -0.3 & 0.2 \\ 0 & 0 \end{bmatrix}$$

making $\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} = W_1$ an eigen vector & a basis for E_1

$E_{1/2} = null(A - \frac{1}{2}I) = null\left(\begin{bmatrix} 0.2 & 0.2 \\ 0.3 & 0.3 \end{bmatrix}\right)$ by inspection $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector and a basis of $E_{1/2}$

$\therefore P = \begin{bmatrix} 0.2 & 1 \\ 0.3 & -1 \end{bmatrix}$ diagonalizes A & $\left\{\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$ is a basis of eigenvectors.

lets compute specific examples $K=1$, $K=2$, $K=$ big beginning with $v_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (it is C today)

$$A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

2-ways $A^2\begin{bmatrix} 1 \\ 0 \end{bmatrix} = AA\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix}$

or $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $c_1 = 2$ $c_2 = 0.6$

same as expected

$$A^2\begin{bmatrix} 1 \\ 0 \end{bmatrix} = A^2\left(2\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} + 0.6\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = 2A^2\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} + 0.6A^2\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= 2(1^2)\begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} + 0.6\left(\tfrac{1}{2}\right)^2\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} + \begin{bmatrix} 0.15 \\ -0.15 \end{bmatrix} = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix}$$

Kilroy

$$A^k\begin{bmatrix}1\\0\end{bmatrix} = 2A^k\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.6A^k\begin{bmatrix}1\\-1\end{bmatrix} = 2(1)^k\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.6\left(\tfrac{1}{2}\right)^k\begin{bmatrix}1\\-1\end{bmatrix}$$

if K is "Big" $\Rightarrow$ limit as $k \to \infty$ which gives $2\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.6\left(\frac{1}{2^k}\right)\begin{bmatrix}1\\-1\end{bmatrix}$

$$= \begin{bmatrix}0.4\\0.6\end{bmatrix} + \begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0.4\\0.6\end{bmatrix}$$

this 40% chance C
& 60% chance RC
is called a "steady state"

we see this as $A\begin{bmatrix}0.4\\0.6\end{bmatrix} = \begin{bmatrix}0.4\\0.6\end{bmatrix}$

was it a fluke? try $v_0 = \begin{bmatrix}0.5\\0.5\end{bmatrix} = 2\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.1\begin{bmatrix}1\\-1\end{bmatrix}$

$$\lim_{k\to\infty} A^k\begin{bmatrix}0.5\\0.5\end{bmatrix} = \lim_{k\to\infty} 2A^k\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.1A^k\begin{bmatrix}1\\-1\end{bmatrix} = \lim_{k\to\infty} 2(1)^k\begin{bmatrix}0.2\\0.3\end{bmatrix} + 0.1\left(\frac{1}{2^k}\right)\begin{bmatrix}1\\-1\end{bmatrix}$$

try a few more if you'd like!

$$= \begin{bmatrix}0.4\\0.6\end{bmatrix} + \begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0.4\\0.6\end{bmatrix}$$

we again reach the same steady state!

basic properties of Markov matrices

1. it has $1 = \lambda_1$ as an eigen value and $|\lambda_i| < \lambda_1 = 1$ for all other e.vals λ_i

2. if v_0 is a probability vector then $\lim_{k\to\infty} A^k v_0$ converges to a vector w s.t.

   i) w is an e.vec of A with e.val 1
   ii) w is a probability vector
   iii) w does <u>not</u> depend on the starting choice of v_0

Such a vector w is called the steady state vector

In summary

- A an $n \times n$ matrix is called a Markov matrix if
  i) all of its entries are > 0 (sometimes we can relax this to ≥ 0)
  ii) the sum of the entries of each column is 1. ie. each column is a probability vector

- Markov matrices have a steady state vector W which is
  i) probability
  ii) an eigenvector of A with eigenvalue 1
  iii) equal to $\lim_{k \to \infty} A^k V_0$ for any probability vector V_0

It is possible to prove these facts. the proofs make use of diagonalization, and essentially ammount to repeating our previous example in more generality - with a few additional small arguments.

If your interested we can go over them in office hours or you can see ____________