Lesson 6

Read chapter 3

Perron's theorem for positive matrices

Positive Markov matrices

In a metropolitan area some percentage x of the population lives in the city, some percentage y lives in the suburbs. The total population is constant but every year 90% city population stays in city

10 % city population moves to suburbs.

98 % suburbs population stays in suburbs.

2 % suburbs population moves to city.

How can I represent this info using matrices?

How can I use linear algebra to derive useful info?

Idea : summarize the information using a matrix

$$\begin{array}{c} C \\ S \end{array} \begin{bmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

Two possibilities (states)

$S_1 = C$: in city $\quad S_2 = S$: in suburbs

changing at discrete time intervals : $t_0, t_1, t_2 \cdots$

Probabilities p_{ij} : probability that if in state j some year will be in state i the following year

Something like this is called a Markov process

Ex: Brownian motion

$\begin{bmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{bmatrix}$ is Markov matrix or transition matrix

Def: $A \in \mathbb{R}^{m \times n}$ is positive if all its entries are > 0

$A \in \mathbb{R}^{m \times n}$ is non negative " are ≥ 0

Def: $M \in \mathbb{R}^{n \times n}$ is Markov if it is non negative and all columns add to 1.

Ex $M = \begin{bmatrix} p_{11} & p_{12} & & p_{1n} \\ p_{21} & p_{22} & & p_{2n} \\ p_{31} & p_{32} & & p_{3n} \\ & & & \\ p_{n1} & p_{n2} & & p_{nn} \end{bmatrix}$

p_{ij} : probability that if in state j at time k, then in state i at time $k+1$

FACTS ABOUT MULTIPLICATION:

$$\begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 c_1 + x_2 c_2 + \cdots + x_n c_n$$

$$\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{pmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_n^T \end{pmatrix} = x_1 c_1^T + x_2 c_2^T + \cdots + x_n c_n^T$$

For Markov matrix

$$\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}}_{A} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

<u>Note</u> $A^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

$\boxed{\text{so } \lambda = 1 \text{ is eigenvalue for } A^T \text{ and for } A}$

We know from hw 1 that A and A^T have the same eigenvalues

Therefore A has eigenvalue 1

Notation: $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbb{1}$

Th : if A is a Markov matrix then $\lambda = 1$ is an eigenvalue for A.

Def $V \in \mathbb{R}^n$ is a probability vector if all of its entries are ≥ 0 and add up to $\underline{1}$

Ex $\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$ is a probability vector

Back to our example. Suppose this year (year 0) 70% of the population of the metropolitan area lives in the city and 30% in the suburbs, i.e initial state is $\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$. Consider

$$\begin{bmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{bmatrix} \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.636 \\ 0.364 \end{bmatrix}$$

Still a probability vector, what does it represent?

$$0.9 \times 0.7 + 0.02 \times 0.3 = 0.636$$

$$0.1 \times 0.7 + 0.98 \times 0.3 = 0.364$$

$$\overbrace{0.9 \times 0.7}^{\text{city population that stayed}} + \overbrace{0.02 \times 0.3}^{\text{suburb population that moved to city}} = 0.636 \quad \% \text{ in city in year 1}$$

$$\underbrace{0.1 \times 0.7}_{\text{city population moved to suburbs}} + \underbrace{0.98 \times 0.3}_{\text{suburb pop that stayed}} = 0.364 \quad \% \text{ in suburbs in year 1}$$

so if $w_0 = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$ is initial state

and $M = \begin{bmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{bmatrix}$ is the matrix of the system

then

year zero   year 1

$$M w_0 = w_1$$

What percentage of the population will live in the city in year n ?

year zero

$$M w_0 = w_1$$ year 1

$$M w_1 = M M w_0 = w_2$$ year 2

. . .

What percentage of the population will live in the city in year n ?

$$M^n w_0 = w_n$$ year n

How can I compute this ?

So we want to compute M^n. Is M diagonalizable?

What are the eigenvalues of

$$M = \begin{bmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{bmatrix} \quad ?$$

$\lambda = 1$ with eigenvector $u_1 = \begin{bmatrix} 1/6 \\ 5/6 \end{bmatrix}$ positive positive probability vector

$\lambda_2 = 0.88$ with eigenvector $u_2 = \begin{bmatrix} -1/6 \\ 1/6 \end{bmatrix}$ negative positive

<u>Note</u> : $\underline{|\lambda_2|} < 1$

$$M = P \begin{bmatrix} 1 & 0 \\ 0 & 0.88 \end{bmatrix} P^{-1} \qquad P = \begin{array}{cc} u_1 & u_2 \\ \end{array} \begin{bmatrix} 1/6 & -1/6 \\ 5/6 & 1/6 \end{bmatrix}$$

$$w_n = M^n w_0$$

$$w_n = P \begin{bmatrix} 1^n & 0 \\ 0 & .88^n \end{bmatrix} \underbrace{P^{-1} w_0}_{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}$$

$$w_n = P \begin{bmatrix} 1 & 0 \\ 0 & 0.88^n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} =$$

$$= \underbrace{\begin{bmatrix} u_1 & u_2 \end{bmatrix}}_{P} \begin{bmatrix} c_1 & 0 \\ 0 & c_2\, 0.88^n \end{bmatrix} = c_1 . u_1 + \underbrace{c_2 \cdot 0.88^n u_2}_{\to 0 \text{ when } n \to +\infty}$$

eigenvector for $\lambda = 1$

In the long run the system goes to $c_1 u_1$

What are c_1 and c_2 ?

$$P^{-1}\begin{bmatrix}0.7\\0.3\end{bmatrix}=\begin{bmatrix}c_1\\c_2\end{bmatrix}$$

so

$$P\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}0.7\\0.3\end{bmatrix}$$

Solve this system

$$\left[\begin{array}{cc|c}1/6 & -1/6 & 0.7\\5/6 & 1/6 & 0.3\end{array}\right] \quad \left[\begin{array}{cc|c}1 & -1 & 4.2\\5 & 1 & 1.8\end{array}\right]$$

$$\left[\begin{array}{cc|c}1 & -1 & 4.2\\0 & 6 & -19.2\end{array}\right]$$

$$c_2 = \frac{-19.2}{6}$$

$$c_1 = 4.2 - \frac{19.2}{6} = \frac{25.2-19.2}{6} = 1$$

remember this 1

$$W_n = U_1 + \overset{c_2}{\left(-\frac{19.2}{6}\right)} 0.88^n \, U_2$$

$$= \begin{bmatrix} 1/6 \\ 5/6 \end{bmatrix} - \frac{19.2 \times 0.88^n}{6} \begin{bmatrix} -1/6 \\ 1/6 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{6} \left[1 + \frac{19.2 \times 0.88^n}{6} \right] \\ \frac{1}{6} \left[5 - \frac{19.2 \times 0.88^n}{6} \right] \end{bmatrix}$$

In the long run $W_n \approx U_1 = \begin{bmatrix} 1/6 \\ 5/6 \end{bmatrix}$

$\frac{1}{6}$ of the population will live in the city, $\frac{5}{6}$ in suburbus.

what can we say in general about computing the eigenvalues, eigenvectors and $M^n w_0$ or $\lim_{n \to +\infty} M^n w_0$ for a Markov matrix ?

Th : if $A \in \mathbb{R}^{n \times n}$ is positive Markov then A^k is positive Markov.

Proof :

A^k positive : Assume $A > 0$ then

$A^2 = A \cdot A$ is positive because its entries are sums of products of positive numbers

$A^3 = A^2 \cdot A$ is positive, for the same reason...

A^k Markov : assume $\mathbb{1}^t A = \mathbb{1}^t$ then

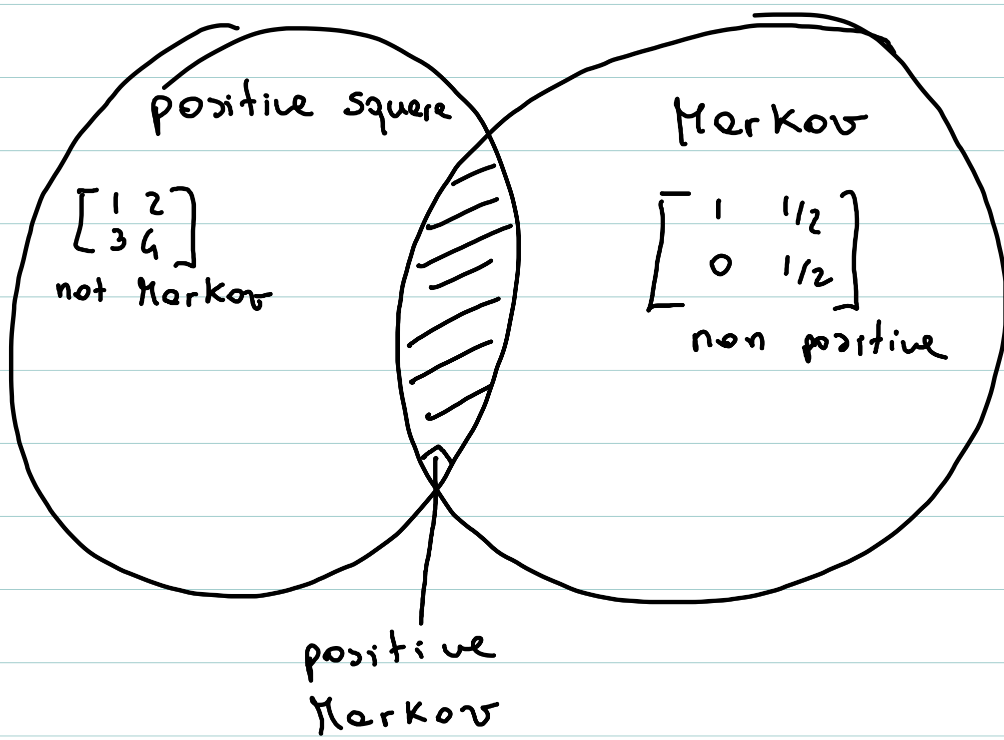
$\mathbb{1}^t A^k = \mathbb{1}^t A \, (A^{k-1}) = \mathbb{1}^t A^{k-1} = \mathbb{1}^t A \, (A^{k-2}) = \mathbb{1}^t A^{k-2} \ldots\ldots$

continuing this way you get $\mathbb{1}^t A^k = \mathbb{1}^t$

Theorem 3.2.4. (**Perron's theorem for positive matrices**) *If* $A > 0$ *then* A *has a* dominant eigenvalue λ_A *with the following properties:*

1. $\lambda_A > 0$ *and it has an eigenvector* $u_A > 0$ *(has all positive entries) called a* dominant eigenvector *of* A.
2. $\mathrm{AM}(\lambda_A) = 1$. $\quad E_{\lambda_A} = \text{span}(u_A)$
3. *If* μ *is another eigenvalue of* A, *then* $|\mu| < \lambda_A$.
4. *If* μ *is a non-dominant eigenvalue of* A, *then it has no eigenvector that is non-negative.*

No proof



Theorem 3.2.5. (**Perron's theorem for positive Markov matrices**) *If A is positive Markov then A has all the properties stated in Theorem 3.2.4, and in addition:*

1. $\lambda_A = 1$ *and thus if μ is another eigenvalue of A, then $|\mu| < 1$.*
2. *If $w_0 \geq 0$ and $w_k = A^k w_0$, then $\lim_{k\to\infty} A^k w_0 = c u_A$ where $c \geq 0$. In other words, the in the long run, w_k tends to a non-negative multiple of u_A.*

[?]

3) if w_0 and u_A are probability vectors, $c=1$

Proof using Perron's th for positive matrices

1) The eigenvalues of A and A^T are the same

A^T is positive and we know $A^T \mathbf{1} = \mathbf{1}$

so $\lambda = 1$ must be the dominant eigenvalue of A^T (why? Because only the dominant eigenvalue of A^T can have a positive eigen vector, by Perron's theorem for positive matrices)

The eigenvalues of A^T and A are the same (see hw 1) so $\lambda = 1$ is the dominant eigenvalue of A.

2) Proof only in the case A is diagonalizable:

$$w_k = A^k w_0 = P D^k P^{-1} w_0 = P D^k \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P \begin{bmatrix} c_1 \lambda_1^k & 0 & \cdots & 0 \\ 0 & c_2 \lambda_2^k & & \\ \vdots & & \ddots & \\ 0 & & & c_n \lambda_n^k \end{bmatrix}$$

$= c_1 \lambda_1^k u_1 + c_2 \lambda_2^k u_2 + \cdots + c_n \lambda_n^k$ $\qquad \lambda_1 = 1$

$= c_1 u_1 +$ stuff that goes to 0 $\qquad |\lambda_i| < 1 \quad i > 1$

u_1 is dominant eigenvector i.e eigenvector for $\lambda = 1$

What can we say about c_1?

$\lim_{k\to+\infty} A^k w_0 = c_1 v_1$ (from previous page)

$* \lim_{k\to+\infty} w_0^T (A^k)^T = c_1 v_1^T$ (Taking transposes)

$\mathbb{1}^T A^k = \mathbb{1}^T$ (because A^k is Markov)

$(A^k)^T \mathbb{1} = \mathbb{1}$ (Taking transposes)

$w_0^T (A^k)^T \mathbb{1} = w_0^T \mathbb{1}$ (multiplying by w_0^T)

$\lim_{k\to+\infty} w_0^T (A^k)^T \mathbb{1} = \lim_{k\to+\infty} w_0^T \mathbb{1}$ (Taking limits)

$c_1 v_1^T \mathbb{1} = w_0^T \mathbb{1}$ (by $*$)

If v_1 and w_0 are probability vectors

$$v_1^T \cdot \mathbb{1} = \underbrace{[x_1 \cdots x_n]}_{v_1^T} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = x_1 + \cdots + x_n = 1$$

$$w_0^T \mathbb{1} = \underbrace{[y_1 \cdots y_n]}_{w_0^T} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = y_1 + y_2 + \cdots + y_n = 1$$

Therefore $c_1 \cdot 1 = 1$