• Def: Steady state

Let $\{\vec{x}_0, \dots, \vec{x}_k, \dots\}$ be a Markov chain. We called steady state to the limit

$$\vec{x}_\infty = \lim_{k\to\infty} \vec{x}(k) \qquad \text{(if it exists)}.$$

MATH 312 LECTURE 24 : Markov chains

Remark: Notice that if $\vec{x}_\infty$ exists, then

$$\vec{x}_\infty = \lim_{k\to\infty} \vec{x}(k) = \lim_{k\to\infty} A\vec{x}(k-1) = A \lim_{k\to\infty} \vec{x}(k-1) = A\vec{x}_\infty .$$

That is, $\| A\vec{x}_\infty = \vec{x}_\infty \|$ Eigenvector of A with eigenvalue 1.
(with cols. adding up to 1)

2) (Question) Does a Markov chain always have a steady state?

→ General answer: No.

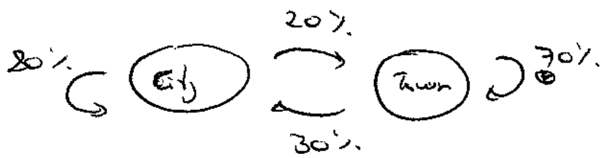
3) If the steady exists, does it depend on the initial configuration?

→ General answer: Yes

However, many systems do have a steady state and it is independent of the initial state.

Example: Find the steady distribution of people in cities and towns, if the initial configuration is

$$\vec{x}(0) = \begin{bmatrix} a \\ 1-a \end{bmatrix} \quad (100a\,\% \text{ in cities}, 100(1-a)\% \text{ in towns}).$$



$$P = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

Sol:

$$\vec{x}_\infty = \lim_{k\to\infty} P^k \vec{x}(0) \quad \to \quad P^k = S D^k S^{-1} \quad (\text{if } P \text{ diagonalizable}).$$

Eigenvalues: $P - \lambda I \rightsquigarrow (0.8-\lambda)(0.7-\lambda) - (0.2)\cdot(0.3) = 0 \Leftrightarrow$

$$\Leftrightarrow (8-10\lambda)(7-10\lambda) - 6 = 0$$

$$\Leftrightarrow 56 + 100\lambda^2 - 150\lambda - 6 = 10\lambda^2 - 15\lambda + 5 = 0 \to \lambda_1 = 1, \ \lambda_2 = 1/2$$

Eigvectors: $P - I = \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix} \to \begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix} \rightsquigarrow \vec{u}_1 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$

not needed: just so that cols. add to 1.

$$P - \tfrac{1}{2} I = \begin{bmatrix} 0.3 & 0.3 \\ 0.2 & 0.2 \end{bmatrix} \to \vec{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

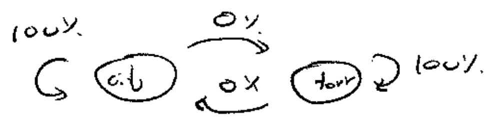
$$P = \begin{bmatrix} 0.6 & 1 \\ 0.4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -0.4 & 0.6 \end{bmatrix} (-1)$$

$$\vec{x}_\infty = P^\infty \vec{x}(0) = \begin{bmatrix}0.6 & 1\\ 0.4 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 0.4 & -0.6\end{bmatrix}\begin{bmatrix}a\\ 1-a\end{bmatrix} =$$

$$= \begin{bmatrix}0.6 & 0\\ 0.4 & 0\end{bmatrix}\begin{bmatrix}a+1-a\\ 0.4a-0.6+0.6a\end{bmatrix} = \begin{bmatrix}0.6\\ 0.4\end{bmatrix} \to \begin{matrix}\text{60\% cities}\\ \text{40\% towns}\end{matrix}$$

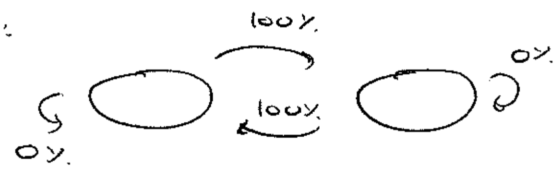
<u>Remark:</u> $\lambda_1 = 1,\ \lambda_2 = \frac{1}{2} \to |\lambda_2| < 1.$

<u>Example 2:</u> $P = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \to$ obviously $\vec{x}_\infty = \vec{x}(0) \to$ so it depends on initial state



$\lambda_1 = 1 = \lambda_2$

<u>Example 3:</u>



$P = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \to \lambda^2 - 1 = 0$

$\hookrightarrow \lambda_1 = 1,\ \lambda_2 = -1.$

There is not steady state!

$\vec{x}(0) = \begin{bmatrix}a\\ 1-a\end{bmatrix} \to \vec{x}(1) = \begin{bmatrix}1-a\\ a\end{bmatrix} \to \vec{x}(2) = \begin{bmatrix}a\\ 1-a\end{bmatrix} \to \ldots$ (unless $a = 0.5$).

Theorem: If A is a Markov matrix with

1) largest eigenvalue $\lambda = 1$ (non repeated)

2) all other eigenvalues $|\lambda| < 1$,

then, for any $\vec{x}(0)$ with ~~positive~~ nonnegative entries and components adding up to 1 (i.e. initial configuration),

we have that

$\rightarrow$ $\vec{x}_\infty = \lim_{k\to\infty} A^k \vec{x}(0)$ is an eigenvector of A for $\lambda = 1$.

$\rightarrow$ $\vec{x}_\infty$ is independent of the choice $\vec{x}(0)$.

Idea of the "proof": Assume A is diagonalizable, so there is a basis of eigenvectors $\{\vec{u}_1, \dots, \vec{u}_n\}$.

Write $\vec{x}(0) = c_1\vec{u}_1 + \dots + c_n\vec{u}_n$, and recall that components of $\vec{x}(0)$ add up to one.

Then,

$$\vec{x}(k) = A^k\vec{x}(0) = c_1 A^k \vec{u}_1 + c_2 A^k \vec{u}_2 + \dots + c_n A^k \vec{u}_n =$$

$$= c_1\vec{u}_1 + c_2\lambda_2^k\vec{u}_2 + \dots + c_n\lambda_n^k\vec{u}_n \Rightarrow$$

$$\Rightarrow \vec{x}_\infty = \lim_{k\to\infty} A^k\vec{x}(0) = c_1\vec{u}_1 + \vec{0} + \dots + \vec{0} = c_1\vec{u}_1.$$

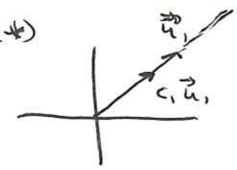
It might seem that $\vec{x}_\infty$ depends on $\vec{x}_0$ through that c_1.

But recall that if $\vec{x}(0)$'s components add up to one, so does $\vec{x}(1), \vec{x}(2), \ldots$ and therefore $\vec{x}_\infty$. Thus,

$\left.\begin{array}{l} \vec{u}_1 \text{ fixed by } A \text{ (is an eigenvector for } \lambda=1) \\ \text{length of } c_1\vec{u}_1 \text{ fixed since its components add up to } 1 \ (*) \end{array}\right\} \rightarrow$

$\rightarrow$ No dependence on $\vec{x}(0)$

(*)



Note that $\vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix}$ gives the slope of the line $m = \frac{u_{12}}{u_{11}}$ (fixed by A).

Then, any vector $\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ on that line is written as $\begin{bmatrix} \alpha_1 \\ m\alpha_1 \end{bmatrix}$

So if components add up to $1 \rightarrow \alpha_1 + m\alpha_1 = 1 \rightarrow \alpha_1 = \frac{1}{1+m}$ (fixed by A.)

• Which matrices satisfy the above theorem? $\rightarrow$ Sufficient condition: A Markov with strictly positive entries.

$\hookrightarrow$ Goal of the rest of chapter. // (pagerank example).

$\downarrow$

(convergence of next theorem)

• Def: Spectral radius of a matrix, $\rho(A)$.

$$\rho(A) = \max\{|\lambda| : \lambda \text{ eigenvalue of } A\}.$$

Example: $\rho\left(\begin{bmatrix} 3+2i & 0 \\ 0 & 1 \end{bmatrix}\right) = \max\{|3+2i|, 1\} = \sqrt{9+4}$

$\rho\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = 0$

Remarks: 1) If A diagonalizable and $\rho(A)=0$, then $A=0$.
($\Leftrightarrow D=0$!)

2) $\rho\left(\frac{1}{\alpha}A\right)=\frac{1}{\alpha}\rho(A)$

Theorem: Perron-Frobenius

Let A be $n\times n$ with all entries strictly positive. Then,

1) $\rho(A)$ is a non-repeated eigenvalue of A.

2) All other eigenvalues satisfy $|\lambda| < \rho(A)$.
(strictly)

3) $\rho(A)$ has an eigenvector with all entries positive.

4) No other eigenvalue of A has a strictly positive eigenvector.
(meaning positive entries).

• $\vec{x}(k) = A\vec{x}(k-1)$ with initial condition $\vec{x}(0)$. [Summary / Review]

If A is diagonalizable, the solution is given by

$\vec{x}(k) = c_1 \lambda_1^k \vec{u}_1 + \dots + c_n \lambda_n^k \vec{u}_n$, where

$$\begin{cases} \vec{u}_1, \dots, \vec{u}_n \text{ are the eigenvectors of } A, \text{ with eigenvalues } \lambda_1, \dots, \lambda_n. \\ c_1, \dots, c_n \text{ are determined by } \vec{x}(0). \end{cases}$$

(indeed, $\vec{x}(0)\Big|_{\mathcal{U}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, with $\mathcal{U}$ the basis of eigenvectors).

• Important result: If A is Markov with $\lambda = 1$ non repeated eigenvalue and all other satisfy $|\lambda| < 1$, then

$\vec{x}(k) = \vec{u}_1 + c_2 \lambda_2^k \vec{u}_2 + \dots + c_n \lambda_n^k \vec{u}_n$, and therefore (Why?)

$\boxed{\vec{x}_\infty = \vec{u}_1}$ does not depend on the initial state. (theorem)

• Remark: In general, anything can happen:

1) No steady state: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

2) Steady state non unique: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ($\vec{x}_\infty$ depends on $\vec{x}(0)$).

- Goal: To find sufficient conditions that guarantee that

  1) $\lambda = 1$ is eigenvalue of A.

  2) $|\lambda| < 1$ for all the rest.

In that case, we can ensure that there exist a unique steady, independent of $\vec{x}(0)$.

- <u>Theorem</u>: If A is Markov with positive entries, then

  1) $\lambda = 1$ eig. of A.

  2) $|\lambda| < 1$ for all other.

We will need to prove first Perron-Frobenius theorem.

From now on, we will say

→ A is positive if all the entries are positive, i.e., $a_{ij} > 0$.

→ $\vec{u}$ " " " " the components " ", i.e., $u_i > 0$

→ $\vec{x} > \vec{y}$ if $x_i > y_i$ for all i

<u>Perron-Frobenius Theorem</u>: Let A $n\times n$ positive matrix (i.e., $a_{ij}>0$).

Then,

1) $\rho(A)$ is a non-repeated eigenvalue of A.

2) All other eigenvalues satisfy $|\lambda|<\rho(A)$.

3) $\rho(A)$ has ~~a~~ a positive eigenvector.

4) No other eigenvalue has a positive eigenvector.

|| Not required ||

We will split the proof in several parts.

<u>Lemma 1</u>: Let $\vec{x},\vec{y}$ vectors such that $\vec{x}\geq\vec{y}$ but $\vec{x}\neq\vec{y}$, and A positive. Then

$$A\vec{x}>A\vec{y}. \text{ In particular, } \exists\, \varepsilon>0 \,/\, A\vec{x}>(1+\varepsilon)A\vec{y}.$$

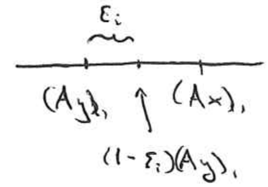
<u>Proof</u>: (not done in class).

① $\vec{x}\geq\vec{y} \Rightarrow x_i-y_i\geq 0 \Rightarrow \cancel{A(\vec{x}-\vec{y})}$

$\hookrightarrow$ $\cancel{A(\vec{x}-\vec{y})}$ .. $(A(\vec{x}-\vec{y}))_i=\sum_{j=1}^{n}a_{ij}(x_j-y_j)\geq$

$\geq(\min_{j=1..n} a_{ij})\sum_{j=1}^{n}(x_j-y_j)>0$ since $\vec{x}\neq\vec{y}$, and $a_{ij}>0$.

② $A\vec{x}>A\vec{y} \Rightarrow A\vec{x}>(1+\varepsilon)A\vec{y}$:



Take $\varepsilon_i=\frac{1}{2}(A\vec{x}-A\vec{y})_i$ and $\varepsilon=\min_{i=1,n}\varepsilon_i$ //

• **Prop 1**: $\rho(A)$ is eigenvalue with a positive eigenvector. (not done in class).

Proof:

Consider the numbers $t \geq 0$ / $A\vec{x} \geq t\vec{x}$ for some nonnegative $\vec{x} \neq \vec{0}$.

Think of the biggest of those numbers, t_{max}: $A\vec{x} \geq t_{max}\vec{x}$.

That is, if we increase t_{max}, there is no $\vec{x}$ verifying the inequality.

Let's prove that, indeed, for this t_{max}, the equality holds:

Assume the contrary: $\cancel{A\vec{x} > t_{max}\vec{x}}$ for some $\vec{x} \neq 0$. $\exists i$ / $(A\vec{x})_i > t_{max} x_i$

Then,

$$A^2\vec{x} > t_{max} A\vec{x} \overset{\text{Lemma 1}}{\Rightarrow} A(A\vec{x}) > (1+\varepsilon)t_{max}(A\vec{x}) \quad \rightarrow\!\leftarrow.$$

So $A\vec{x} = t_{max}\vec{x}$, and t_{max} is eigenvalue.

→ Is t_{max} the larger eigenvalue? I.e, $\rho(A) = t_{max}$?

→ Is $\vec{x}$ corresponding to $t_{max} = \rho(A)$ positive?

↳ Easy: $\left.\begin{array}{l}\vec{x} \geq \vec{0} \\ \vec{x} \neq \vec{0}\end{array}\right| \underset{\text{Lemma 1}}{\Rightarrow} A\vec{x} > \vec{0}$. Since $t_{max}\vec{x} = A\vec{x} > \vec{0} \Rightarrow \vec{x} > 0$!

↳ Suppose λ is another eigenvalue with $\vec{u}$ eigenvector.

$A\vec{u} = \lambda\vec{u}$. Then $|\lambda\vec{u}| = |\lambda||\vec{u}| = |A\vec{u}| = |\sum a_{ij}u_j| \leq \sum |a_{ij}u_j| \overset{A \text{ positive}}{=} A|\vec{u}| \Rightarrow$

$\Rightarrow A|\vec{u}| \geq |\lambda| \underbrace{|\vec{u}|}_{\text{nonnegative}} \Rightarrow |\lambda|$ ~~...~~ cannot be bigger than t_{max}.

- **Prop. 2:** $\rho(A)$ is non repeated eigenvalue of A. (not done in class)

Proof: We won't really prove this. We will only prove that its eigenspace has dimension 1, i.e., that $N(A-\rho(A)I)=1$.

Suppose $\vec{w}>\vec{0}$ is eigenvector of $\rho(A)$, and $\vec{u}$ another l.i. eigenvector of $\rho(A)$.

Then we can find α such that $\vec{w}-\alpha\vec{u}\geq 0$ with one entry $=0$.

Let $\alpha_i=\frac{w_i}{u_i}$ for those $\vec{u}_i\neq 0$. Let $\alpha=\min \alpha_i$.

Then $\vec{w}-\alpha\vec{u}$ has at least one entry zero. Not all zero since $\vec{w}$ and $\vec{u}$ are l.i.

Using Lemma 1, $A(\vec{w}-\alpha\vec{u})>0$, and so

$$A\vec{w}-\alpha A\vec{u}=\rho(A)\vec{w}-\alpha\rho(A)A\vec{u}=\rho(A)(\vec{w}-\alpha\vec{u})>0 \rightsquigarrow \vec{w}-\alpha\vec{u}>0 \text{ ↯}$$

(at least one entry $=0$)

Remark: $\rho(A)\neq 0$ due to Prop 1.

Prop 3: All other eig. $|\lambda| < \rho(A)$. (not done in class)

Proof.

First, $|\lambda| \leq \rho(A)$ by def. of $\rho(A)$. So just need to prove that

If $|\lambda| = \rho(A)$, then $\lambda = \rho(A)$ (not negative, not complex).

· Let $\lambda \neq \rho(A)$, $|\lambda| = \rho(A)$, with eigvtor $\vec{u}$: $A\vec{u} = \lambda\vec{u}$.

Then, $|A\vec{u}| = |\lambda\vec{u}| = \overbrace{|\lambda|}^{=\rho(A)}|\vec{u}|$

$|A\vec{u}| = |\sum a_{ij}u_j| \leq \sum |a_{ij}u_j| = A|\vec{u}|$ $\Bigg\} \Rightarrow A|\vec{u}| \geq \rho(A)|\vec{u}| \underset{\text{prop 1)}}{\Rightarrow}$

$\Rightarrow A|\vec{u}| = \rho(A)|\vec{u}| > 0$

Thus,

$(A|\vec{u}|)_1 = \sum a_{1j}|u_j| = \rho(A)|u_1| > 0 \quad (1)$

and

$(A\vec{u})_1 = (\lambda\vec{u})_1 = \sum a_{1j}u_j \Rightarrow |(\lambda\vec{u})_1| = \rho(A)|u_1| = |\sum a_{1j}u_j| \quad (2)$

That is, $\boxed{\sum a_{1j}|u_j| = |\sum a_{1j}u_j|}$

This implies that

$a_{11}|u_1| + a_{12}|u_2| + \dots + a_{1n}|u_n| = |a_{11}u_1 + \dots + a_{1n}u_n| \overset{(A}{\Rightarrow}$

$\Rightarrow a_{1j}u_j = \alpha_j a_{11}u_1$

(*) We are using that for $z_1, \dots, z_n \neq 0$ complex numbers with

$|z_1 + \dots + z_n| = |z_1| + \dots + |z_n|$, then each z_j is a positive multiple of z_1.

Therefore,

$$u_j = \left(\alpha_j \frac{a_{11}}{a_{1j}}\right) u_1 \Rightarrow \vec{u} = \underbrace{\begin{bmatrix} 1 \\ \alpha_2 a_{11}/a_{12} \\ \vdots \\ \alpha_n a_{11}/a_{1n} \end{bmatrix}}_{\vec{w} > 0 \text{ (notice } \vec{u} = u_1 \vec{w})} u_1$$

(not done in class),

- Going back, $\lambda \vec{u} = A\vec{u} \Rightarrow \lambda \vec{w} = A\vec{w} = |A\vec{w}| = |\lambda \vec{w}| = \rho(A)\vec{w} \Rightarrow$

$\Rightarrow \lambda \vec{w} = \rho(A)\vec{w} \Rightarrow \lambda = \rho(A)$.

(first component of $\vec{w}$ is 1)

[Hasta aquí in class]

- Prop 4: No other eigenvalue has a positive eigenvector.

Proof: Assume $\lambda \neq \rho(A)$ with $\vec{u} > 0$. Then $A\vec{u} = \lambda \vec{u}$ (1).

We know from Prop 1 that $\rho(A)$ is eigenvalue of A. Let's see it is also of A^T:

No hecho in class

(Chapt. 5)

Lemma: $\rho(A) = \rho(A^T)$

↳ Indeed, A and A^T have the same eigenvalues! (but different eigenvectors).

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

Let $\vec{w} > 0$ eigvctor corresponding to $\rho(A^T)$ of A^T:

$$A^T \vec{w} = \rho(A) \vec{w} \quad (2)$$

• Multiply (1) by $\vec{w}^T$: $\vec{w}^T A \vec{u} = \lambda \vec{w}^T \vec{u}$

and transpose: $\underbrace{\vec{u}^T A^T \vec{w}}_{(2) \Rightarrow = \rho(A)\vec{w}} = \lambda \vec{w}^T \vec{u} \Rightarrow \rho(A) \vec{w}^T \vec{u} = \lambda \vec{w}^T \vec{u} \Rightarrow$

$\Rightarrow (\rho(A) - \lambda)\, \vec{w}^T \vec{u} = 0 \Rightarrow \rho(A) = \lambda \ \ \not\rightarrow$.

($\vec{w}^T \vec{u} > 0$ since both are > 0)

No hecho in clase

(not required)

• We can now prove that for A Markov with positive entries, then

$\lambda = 1$ is eig. of A (with positive eigenvector ~ steady state).
$|\lambda| < 1$ for all other

This is required (done in class).

Proof: We just need to prove that $\rho(A) = 1$.

→ $\rho(A) \leq 1$: ~~Let λ eig.: $A\vec{v} = \lambda \vec{v}$ ($\vec{v} \neq \vec{0}$)~~ Suppose $\rho(A) > 1$.

Then, pick the corresponding positive eigvctor $\vec{u} > 0$: $A\vec{u} = \rho(A)\vec{u}$

We can scale $\vec{u}$ so that their components add up to 1, but then $A\vec{u}$ also has components that add up to 1, while $\rho(A)\vec{u}$ would add up to $\rho(A) > 1$ $\not\rightarrow$.

Now, we can check that $\rho(A)=1$ is always an eigenvalue for A Markov.

Recall $\rho(A)=\rho(A^T)$ and notice that

$$A^T\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix} = \begin{bmatrix}\text{sum of row 1 of } A^T\\ \vdots \\ \text{sum of row } n \text{ of } A^T\end{bmatrix} = \begin{bmatrix}\text{sum of col 1 of } A\\ \vdots \\ \text{sum of col } n \text{ of } A\end{bmatrix} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix} \Rightarrow$$

required

$\Rightarrow \lambda = 1$ is eigenvalue of A^T, and thus of A.

Not in sol in Qs.

Questions:

~~1) If A has cols. that add up to 1, then $\lambda=1$ is always an eigenvalue.~~

~~True~~

1) A, A^T have same eigvalues and eigvectors.

~~2) A Markov matrix with some entries equal to zero cannot have a unique steady~~

2) If a Markov chain is given by A matrix with some zero entries, then the steady state depends on $\vec{x}(0)$.

3) For $\vec{x}(k) = A\vec{x}(k-1)$, with $\vec{x}(0)$, the solution is
and A diagonalizable

$u = \{\vec{u}_1, \dots, \vec{u}_n\}$ eigvectors of A.

$$\vec{x}(k) = c_1\lambda_1^k\vec{u}_1 + \dots + c_n\lambda_n^k\vec{u}_n, \text{ where } \begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix} = \vec{x}(0)/_u$$

For A positive Markov and $\vec{u}_1$ eigvector with $\lambda=1$, can we pick $\vec{x}(0)$ such that $c_1 = 0$?

4) A matrix whose cols. add up to 1 always has $\lambda = 1$ as eig. (even if some entries are zero).