By Ruize Luo

Proposition 1. *Prove that* $\exists v \in \mathbb{R}^n$ *such that* $\forall w \in \mathbb{R}^n$ *with entries summing to 1 and* $\forall \epsilon > 0$, $\exists N \geq 1$ *such that* $\forall k \geq N$, $\|A^k w - v\| < \epsilon$, *where* $\|(x1, \cdots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$ *and A is a left stochastic matrix.*

Proof 1.

Firstly we would like to show that every eigenvalue of A has modulus less than or equal to 1, i.e. $|\lambda| \leq 1$.

PROOF Suppose $\lambda \in \mathbb{C}$ *is an eigenvalue of* A^T *and X is a corresponding eigenvector. (Because a matrix A and its transpose* A^T *has the same eigenvalues) Then we have:*

$$A^T X = \lambda X$$

Choose k such that $|x_j| \leq |x_k|\ \forall j \in [1, n]$. *Then take the* k^{th} *component of each side of the equation:*

$$\sum_{i=1}^{n} a_{ik} x_i = \lambda x_k$$

Hence

$$|\lambda x_k| = |\lambda| \cdot |x_k| = |\sum_{i=1}^{n} a_{ik} x_i| \leq \sum_{i=1}^{n} a_{ik} |x_i| \leq \sum_{i=1}^{n} a_{ik} |x_k|$$

Because the sum of entries in each column of A equals 1, $\sum_{i=1}^{n} a_{ij} = 1\ \forall j \in [1, n]$.

Hence

$$|\lambda x_k| = |\lambda| \cdot |x_k| \leq \sum_{i=1}^{n} a_{ik} |x_k| = |x_k|$$

Therefore $|\lambda| \leq 1$ ■

Then we want to show that the multiplicity of eigenvalue 1 *is exactly* 1.

PROOF Suppose $|\lambda| = 1$, $A^T X = \lambda X$, $X \neq 0$
Then by using the inequalities from the proof above:

$$|\lambda x_k| = |\sum_{i=1}^{n} a_{ik} x_i| \leq \sum_{i=1}^{n} a_{ik}|x_i| \leq \sum_{i=1}^{n} a_{ik}|x_k|$$
$$\Rightarrow |x_k| = |\sum_{i=1}^{n} a_{ik} x_i| \leq \sum_{i=1}^{n} a_{ik}|x_i| \leq \sum_{i=1}^{n} a_{ik}|x_k| = |x_k|$$

The above inequalities gives that $|x_j| = |x_k|$ *for* $j \in [1, n]$. *In addition, as equality holds in the triangle inequality section, this means that all complex numbers* $a_{ik}x_i$ *lie on the same direction. Consequently, it implies that either* $x_j = x_k$ *or* $x_j = -x_k$ *for all* $j \in [1, n]$. *Hence eigenvector* X *is unique and* $X = x_k \cdot (1, 1, \cdots, 1)$. *Therefore eigenvalue* 1 *is unique.* ∎

Now matrix A can form a Jordan form by applying change of basis matrix to it:

$$PAP^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & J(\lambda_1) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J(\lambda_m) \end{bmatrix}$$

While all the $J(\lambda_i)$*s are Jordan Blocks with* $\lambda = \lambda_i$ *and* $|\lambda_i| < 1$.
If we are to take the power of PAP^{-1}, *we need to show that wile we are taking the power of all Jordan Blocks, the entries in those Jordan Blocks tends to be zero. I would like to prove this by proving this lemma:*

$$J_n(\lambda)^m = \begin{bmatrix} \lambda^m & m\lambda^{m-1} & \cdots & \cdots & \binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^m & m\lambda^{m-1} & \cdots & \binom{m}{n-2}\lambda^{m-n+2} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ & & & \lambda^m & m\lambda^{m-1} \\ 0 & \cdots & \cdots & 0 & \lambda^m \end{bmatrix}$$

(Note that if there are'n enough terms to fill each row, all entry after 1 would be zero for every row). and then show that when $m \to \infty$, *all the entries in the Jordan Block approaches to* 0.

PROOF Firstly, when $m = 2$,

$$J_n(\lambda)^2 = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} \lambda^2 & \lambda+\lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda^2 & \lambda+\lambda & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ \vdots & & & & \lambda^2 & \lambda+\lambda \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda^2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda^2 & 2\lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda^2 & 2\lambda & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ \vdots & & & & \lambda^2 & 2\lambda \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda^2 \end{bmatrix}$$

So it holds for $J_n(\lambda)^2$. Now suppose that it is true for $J_n(\lambda)^m$, then

$$J_n(\lambda)^{m+1} = \begin{bmatrix} \lambda^m & m\lambda^{m-1} & \cdots & \cdots & \binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^m & m\lambda^{m-1} & \cdots & \binom{m}{n-2}\lambda^{m-n+2} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ & & & \lambda^m & m\lambda^{m-1} \\ 0 & \cdots & \cdots & 0 & \lambda^m \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} \lambda^{m+1} & \lambda^m + \lambda m\lambda^{m-1} & \cdots & \cdots & \binom{m}{n-2}\lambda^{m-n+2} + \lambda\binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^{m+1} & \lambda^m + \lambda m\lambda^{m-1} & & \binom{m}{n-3}\lambda^{m-n+3} + \lambda\binom{m}{n-2}\lambda^{m-n+2} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \lambda^{m+1} & \lambda^m + m\lambda^m \\ 0 & \cdots & \cdots & 0 & \lambda^{m+1} \end{bmatrix}$$

Since combinatory has this property: $\binom{m}{n-2} + \binom{m}{n-1} = \binom{m+1}{n-1}$ *for all* $m > n$*, the matrix above is equal to:*

$$\begin{bmatrix} \lambda^{m+1} & (m+1)\lambda^m & \cdots & \cdots & \binom{m+1}{n-1}\lambda^{m-n+2} \\ 0 & \lambda^m & (m+1)\lambda^m & \cdots & \binom{m+1}{n-2}\lambda^{m-n+3} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ & & & \lambda^{m+1} & (m+1)\lambda^m \\ 0 & \cdots & \cdots & 0 & \lambda^{m+1} \end{bmatrix} = J_n(\lambda)^{m+1}$$

Therefore

$$J_n(\lambda)^m = \begin{bmatrix} \lambda^m & m\lambda^{m-1} & \cdots & \cdots & \binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^m & m\lambda^{m-1} & \cdots & \binom{m}{n-2}\lambda^{m-n+2} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ & & & \lambda^m & m\lambda^{m-1} \\ 0 & \cdots & \cdots & 0 & \lambda^m \end{bmatrix}$$

is true for all $m > 2$. ■

With the lemma above, it is easy to show that all entries in $J_n(\lambda)^m$ *approaches* 0 *as* m *approaches infinity.*

PROOF It is clear that $\lim_{m\to\infty} \lambda^m = 0$.
Now, for any other non-zero entries,

$$\begin{aligned} &\lim_{m\to\infty} \binom{m}{k-1}\lambda^{m-k+1} \\ =& \lim_{m\to\infty} \frac{m!}{(k-1)!(m-k+1)!}\lambda^{m-k+1} \\ =& \lim_{m\to\infty} \frac{m(m-1)\cdots(m-k+2)}{(k-1)!}\lambda^{m-k+1} \end{aligned}$$

Because $|\lambda| < 1$*, pick* p, q *with* $p < q$ *such that* $\lambda = \frac{p}{q}$*. Then the limit becomes:*

$$\begin{aligned} &\lim_{m\to\infty} \frac{m(m-1)\cdots(m-k+2)}{(k-1)!}(\frac{p}{q})^{m-k+1} \\ =& \lim_{m\to\infty} \frac{m(m-1)\cdots(m-k+2)}{(k-1)!(\frac{q}{p})^{m-k+1}} \end{aligned}$$

Because when $m \to \infty$, $(k-1)!$ is fixed, now we consider this case where the numerator approaches infinity, $(\frac{q}{p})^{m-k+1}$ in the denominator also goes to infinity. Hence in this case we can apply L'Hôpital's rule to it. Because $m(m-1)\cdots(m-k+2)$ on the numerator has m^{k-1} as the highest order term, we apply L'Hôpital's rule $k-1$ times to this equation, and then the only term left on the numerator would be $(k-1)!$, which is fixed.
However, after we applied L'Hôpital's rule $k-1$ times, the denominator still has the $(\frac{q}{p})^{m-k+1}$ term because it is exponential. And when $m \to \infty$, $(\frac{q}{p})^{m-k+1} \to \infty$ on the denominator, and hence,

$$\lim_{m\to\infty} \binom{m}{k-1} \lambda^{m-k+1} = 0$$

Consequently, because the largest term tends to zero, the Jordan Block $J_n(\lambda)^m$ tends to zero. ■

Then we want to show that as n goes to infinity, $(PAP^{-1})^n$ converges to a fixed matrix.

PROOF

$$\lim_{n\to\infty} (PAP^{-1})^n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

However, $(PAP^{-1})^n = PA^nP^{-1}$, hence

$$\lim_{n\to\infty} PA^nP^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

$$\Rightarrow \lim_{n\to\infty} A^n = P^{-1} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} P$$

Here P, P^{-1} are the change of basis matrix which changes $[A]_e^e$ to $[A]_j^j$ (let e be the standard basis and j be the Jordan Basis), hence $P = [I]_e^j$ and $P^{-1} = [I]_j^e$. ■

Finally we want to show that $P^{-1}MP$ sends every vector whose entries sum 1 to the right eigenvector of A, $v = (v_1, \cdots, v_n)$. Where $M = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$

PROOF Let J be the Jordan From of A, namely PAP^{-1}. Note that:

$$\begin{aligned} PAP^{-1} &= J \\ \Rightarrow PA &= JP \end{aligned}$$

Hence the rows of P are the left eigenvectors of A. Because $J = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & J(\lambda_1) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J(\lambda_m) \end{bmatrix}$*, hence the first row of P must be the left eigenvector of A corresponding to eigenvalue 1. As we have shown in the first lemma in this problem, the right eigenvector of A^T for $\lambda = 1$ is $(1, \cdots, 1)$; and we know that the right eigenvector of A^T is the left eigenvector of A, hence the left eigenvector of A for $\lambda = 1$ is $(1, \cdots, 1)$. Consequently, the first row of P is $(1 \cdots 1)$.Hence*

$$P = \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ & & & \\ & * & & \\ & & & \end{bmatrix}$$

Note also that:

$$\begin{aligned} AP^{-1} &= P^{-1}J \\ \Rightarrow AP^{-1} &= P^{-1}J \end{aligned}$$

Hence the columns of P^{-1} consist of components of right eigenvectors of A. Similarly, the first column would be the one corresponding to eigenvalue 1, namely $v = (v_1, \cdots, v_n)$. Hence

$$P^{-1} = \begin{bmatrix} v_1 & & \\ v_2 & & \\ \vdots & * & \\ v_n & & \end{bmatrix}$$

Now let's calculate $P^{-1}MP$.

$$
\begin{aligned}
P^{-1}MP &= (P^{-1}M)P \\
&= \begin{bmatrix} v_1 & & \\ v_2 & & \\ \vdots & * & \\ v_n & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} P = \begin{bmatrix} v_1 & & \\ v_2 & & \\ \vdots & 0 & \\ v_n & & \end{bmatrix} P \\
&= \begin{bmatrix} v_1 & & \\ v_2 & & \\ \vdots & 0 & \\ v_n & & \end{bmatrix} \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ & & & \\ & * & & \\ & & & \end{bmatrix} = \begin{bmatrix} v_1 & v_1 & \cdots & v_1 \\ v_2 & v_2 & \cdots & v_2 \\ \vdots & \vdots & & \vdots \\ v_n & v_n & \cdots & v_n \end{bmatrix}
\end{aligned}
$$

Now take $w \in \mathbb{R}^n$ whose entries summing to 1 and apply $P^{-1}MP$ to it. Let $w = (w_1, \cdots, w_n)$

$$
P^{-1}MPw = \begin{bmatrix} v_1 & v_1 & \cdots & v_1 \\ v_2 & v_2 & \cdots & v_2 \\ \vdots & \vdots & & \vdots \\ v_n & v_n & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 \sum_{j=1}^n w_j \\ v_2 \sum_{j=1}^n w_j \\ \vdots \\ v_n \sum_{j=1}^n w_j \end{bmatrix}
$$

Because $\sum\limits_{j=1}^{n} w_j = 1$, hence

$$
P^{-1}MPw = \begin{bmatrix} v_1 \sum_{j=1}^n w_j \\ v_2 \sum_{j=1}^n w_j \\ \vdots \\ v_n \sum_{j=1}^n w_j \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v^T
$$

Therefore $P^{-1}MP$ sends any $w \in \mathbb{R}^n$ whose entries summing to 1 to the right eigenvector v. ■

Because $\lim_{n\to\infty} A^n = P^{-1}MP$, hence $\lim_{n\to\infty} A^n w = P^{-1}MPw = v$.

Therefore there exist v, namely the right eigenvector of stochastic matrix A, such that for any $w \in \mathbb{R}^n$ with entries summing to 1, for all $\epsilon > 0$, there exist $N > 1$ such that for all $k > N$, $\|A^k w - v\| < \epsilon$

Quod Erat Demonstrandum!