MAU34804

Lecture 16

2026-03-09

# Perron Matrices

Consider the following three conditions on an $n \times n$ matrix $A$:
- $A \geq 0$, i.e. $a_{j,k} \geq 0$ for all $1 \leq j, k \leq n$.
- $\sum_{j=1}^{n} a_{j,k} = 1$ or, equivalently, $(1, 1, \ldots, 1)$ is an eigenvector of $A^T$.
- For each $1 \leq j, k \leq n$ there's an $i$ such that the $j, k$'th entry of $A^i$ is positive.

The first and second conditions together define Markov matrices. We proved some results about these, but needed to add the third condition to get stronger results.
What happens if we assume only the first and third conditions, without the second?
No row or column of $A$ can be identically zero, since otherwise the corresponding row or column of $A^m$ would be identically zero for all positive $m$.
We'll be interested in the action of $A$ on the simplex

$$\Delta_{\mathbf{w}} = \left\{ \mathbf{p} \in \mathbf{R}_+^n : \mathbf{w} \cdot \mathbf{p} = 1 \right\} = \left\{ \mathbf{p} \in \mathbf{R}^n : p_0 \geq 0, \ldots, p_n \geq 0, \sum_{j=0}^{n} w_j p_j = 1 \right\}.$$

where $\mathbf{w} \gg \mathbf{0}$, i.e. $w_j > 0$ for each $j$.

# Action on the simplex $\Delta_{\mathbf{w}}$

$\frac{a_{i,k}}{w_k} \geq \min_{1\leq j\leq n} \frac{a_{i,j}}{w_j}$ so if $\mathbf{p} \in \Delta_{\mathbf{w}}$ then

$$\sum_{k=1}^{n} a_{i,k} p_k \geq \sum_{k=1}^{n} \min_{1\leq j\leq n} \frac{a_{i,j}}{w_j} w_k p_k = \min_{1\leq j\leq n} \frac{a_{i,j}}{w_j} \sum_{k=1}^{p} w_k p_k = \min_{1\leq j\leq n} \frac{a_{i,j}}{w_j} \geq 0$$

and

$$\begin{aligned}\sum_{j=1}^{n}\sum_{k=1}^{n} w_j a_{j,k} p_k &= \sum_{k=1}^{n}\sum_{j=1}^{n} \frac{w_j a_{j,k}}{w_k} w_k p_k \geq \min_{1\leq l\leq n} \frac{\sum_{j=1}^{n} w_j a_{j,l}}{w_l} \sum_{k=1}^{n} w_k p_k \\ &= \min_{1\leq l\leq n} \frac{\sum_{j=1}^{n} w_j a_{j,l}}{w_l} \sum_{k=1}^{n} w_k p_k \geq \min_{1\leq l\leq n} \max_{1\leq j\leq n} \frac{w_j a_{j,l}}{w_l} > 0.\end{aligned}$$

We can therefore define a function $f_{\mathbf{w}} \colon \Delta_{\mathbf{w}} \to \Delta_{\mathbf{w}}$ by $f(\mathbf{p}) = \mathbf{q}$, where

$$q_i = \frac{\sum_{k=1}^{n} a_{i,k} p_k}{\sum_{j=1}^{n}\sum_{k=1}^{n} w_j a_{j,k} p_k}.$$

# Fixed points and eigenvectors

Consider the function $f_{\mathbf{w}}(\mathbf{p}) = \mathbf{q}$, where

$$q_i = \frac{\sum_{k=1}^{n} a_{i,k} p_k}{\sum_{j=1}^{n} \sum_{k=1}^{n} w_j a_{j,k} p_k}.$$

If $\mathbf{p} \in \Delta_{\mathbf{w}}$ is a fixed point of this function then

$$p_i = \frac{\sum_{k=1}^{n} a_{i,k} p_k}{\sum_{j=1}^{n} \sum_{k=1}^{n} w_j a_{j,k} p_k}$$

or

$$A\mathbf{p} = \mu \mathbf{p}, \quad \mu = \sum_{j=1}^{n} \sum_{k=1}^{n} w_j a_{j,k} p_k,$$

so $\mathbf{p}$ is an eigenvector of $A$ with eigenvalue $\mu > 0$ .
Any positive multiple $\mathbf{v}$ of $\mathbf{p}$ is also an eigenvector with eigenvalue $\mu$, with $\mathbf{v} > \mathbf{0}$.

# Fixed points and eigenvectors, continued

Conversely, suppose $\mathbf{v} > \mathbf{0}$ is an eigenvector of $A$ with eigenvalue $\mu$. Let

$$\mathbf{p} = \frac{1}{\mathbf{w} \cdot \mathbf{v}} \mathbf{v}.$$

Then $\mathbf{w} \cdot \mathbf{p} = 1$ so $\mathbf{p} \in \Delta_{\mathbf{w}}$. Also $\mathbf{p}$ is an eigenvector of $A$ with eigenvalue $\mu$.
Then

$$q_i = \frac{\sum_{k=1}^{n} a_{i,k} p_k}{\sum_{j=1}^{n} \sum_{k=1}^{n} w_j a_{j,k} p_k} = \frac{\mu p_i}{\sum_{j=1}^{n} w_j \mu p_j} = p_i$$

so $f_{\mathbf{w}}(\mathbf{p}) = \mathbf{p}$, i.e. $\mathbf{p}$ is a fixed point of $f_{\mathbf{w}}$.
So the eigenvectors of $A$ in $\mathbf{R}_{+}^{n}$ are precisely the positive multiples of the fixed points of $f_{\mathbf{w}}$.

# Brouwer

The function $f_{\mathbf{w}}$ defined by $f(\mathbf{p}) = \mathbf{q}$, where $q_i = \frac{\sum_{k=1}^{n} a_{i,k} p_k}{\sum_{j=1}^{n} \sum_{k=1}^{n} w_j a_{j,k} p_k}$ is continuous, so by Brouwer there is a $\mathbf{b} \in \Delta_{\mathbf{w}}$ such that $f_{\mathbf{w}}(\mathbf{b}) = \mathbf{b}$.

This fixed point in fact lies in the interior of $\Delta_{\mathbf{w}}$. We can see this as follows.

Split $\{1, \ldots, n\}$ into $J = \{i \colon b_i = 0\}$ and $K = \{i \colon b_i > 0\}$.

Then $\sum_{k=1}^{n} a_{j,k} b_k = \mu b_j = 0$ for $j \in J$, which is possible only if $a_{j,k} = 0$ for all $k \in K$.

Suppose $P$ and $Q$ are matrices such that $p_{j,k} = 0$ and $q_{j,k} = 0$ for all $j \in J$ and $k \in K$ and let $R = PQ$. If $j \in J$ and $k \in K$

$$r_{j,k} = \sum_{i=1}^{n} p_{j,i} q_{i,k} = \sum_{i \in J} p_{j,i} q_{i,k} + \sum_{i \in K} p_{j,i} q_{i,k}$$

and the first sum is zero because $q_{i,k} = 0$ for each summand and the second sum is zero because $p_{j,i} = 0$ for each summand, so $r_{j,k} = 0$ for all $j \in J$ and $k \in K$.

By induction we can then prove that all powers of $A$ have this property.

Is this consistent with our assumptions about $A$? Yes, but only if $J = \varnothing$! So $\mathbf{b}$ lies in the interior of $\Delta_{\mathbf{w}}$.

# Duality

We could apply the same analysis to $A^T$ in place of $A$ to get a $\mathbf{c} \gg \mathbf{0}$ and a $\nu > 0$ such that

$$\nu c_k = \sum_{j=1}^{n} c_j a_{j,k}.$$

What is the relation between $\mu$ and $\nu$?

$$\mu \sum_{j=1}^{n} c_j b_j = \sum_{j=1}^{n} c_j \sum_{k=1}^{n} a_{j,k} b_k = \sum_{j=1}^{n} \sum_{k=1}^{n} c_j a_{j,k} b_k = \sum_{k=1}^{n} \sum_{j=1}^{n} c_j a_{j,k} b_k$$

$$\nu \sum_{j=1}^{n} c_j b_j = \nu \sum_{k=1}^{n} c_k b_k = \sum_{k=1}^{n} \left( \sum_{j=1}^{n} c_j a_{j,k} \right) b_k = \sum_{k=1}^{n} \sum_{j=1}^{n} c_j a_{j,k} b_k$$

$\sum_{j=1}^{n} c_j b_j > 0$ so we can conclude that $\mu = \nu$.
What is the relation between $\mathbf{b}$ and $\mathbf{c}$? None, in general.

# A related matrix

We got **b** and **c** from Brouwer so there's no guarantee they're unique. We will see that they are, but for now we'll just choose some eigenvector $\mathbf{c} \gg \mathbf{0}$ with eigenvalue $\mu > 0$ and use it to define a matrix $B$ via

$$b_{j,k} = \frac{c_j a_{j,k}}{\mu c_k},$$

or, equivalently, $B = \mu^{-1} C A C^{-1}$, where $C$ is a diagonal matrix whose $j$'th row, $j$'th column is $c_j$.
It's clear that $B \geq O$ and an easy calculation gives

$$\sum_{j=1}^{n} b_{j,k} = \sum_{j=1}^{n} \frac{c_j a_{j,k}}{\mu c_k} = \frac{\sum_{j=1}^{n} c_j a_{j,k}}{\mu c_k} = \frac{\mu c_k}{\mu c_k} = 1$$

so $(1, \ldots, 1)$ is an eigenvector of $B^T$ with eigenvalue 1 and $B$ is a Markov matrix.
Now $B^i = \mu^i C A^i C^{-1}$ so the condition on $A$ that for each $1 \leq j, k \leq n$ there's an $i$ such that the $j, k$'th entry of $A^i$ is positive implies the same condition on $B$.

# Uniqueness of $\mathbf{b}$ and $\mathbf{c}$

We've now seen that $B = \mu^{-1}CAC^{-1}$ is a Markov matrix satisfying the additional condition which allowed us to prove that left multiplication by $B$ has a unique fixed point in the standard simplex in $\mathbf{R}^n$, which is equivalent to $f_{\mathbf{c}}$ having a unique fixed point in the simplex $\Delta_{\mathbf{c}}$.

This in turn is equivalent to $A$ having a unique positive eigenvector, up to multiplication by positive constants, which must be $\mathbf{b}$.

So $\mathbf{b}$ is uniquely determined, but $A^T$ satisfies the same conditions as $A$ so $A^T$ also has a unique positive eigenvector, up to multiplication by positive constants, which must be $\mathbf{c}$.

# Banach

We saw earlier that all Markov matrices are Lipschitz with Lipschitz constant 1,

$$d_1(B\mathbf{y}, B\mathbf{z}) \leq d_1(\mathbf{y}, \mathbf{z}).$$

Two proofs were given in lecture. The second, more complicated, proof gave a Lipschitz constant strictly less than 1 under some hypotheses which our $B$ satisfies. It only applied to the action of $A$ on the standard simplex though. The first proof works on all of $\mathbf{R}^n$.

Suppose $|\lambda| > \mu$, so that $\left|\frac{\mu}{\lambda}\right| < 1$, and consider the function

$$g_{\mathbf{x}}(\mathbf{y}) = \mathbf{x} + \frac{\mu}{\lambda} B\mathbf{y},$$

which depends on a parameter $\mathbf{x}$.

Now, using the translation invariance and scaling properties of $d_1$ we find

$$d_1(g_{\mathbf{x}}(\mathbf{y}), g_{\mathbf{x}}(\mathbf{z})) = \left|\frac{\mu}{\lambda}\right| d_1(B\mathbf{y}, B\mathbf{z}) \leq \left|\frac{\mu}{\lambda}\right| d_1(\mathbf{y}, \mathbf{z}).$$

The Banach fixed point theorem therefore applies and gives us a $\mathbf{y}$ such that $g_{\mathbf{x}}(\mathbf{y}) = \mathbf{y}$.

# Banach, continued

For every $\mathbf{x}$ we have a $\mathbf{y}$ such that $g_{\mathbf{x}}(\mathbf{y}) = \mathbf{y}$, i.e. such that $\mathbf{x} + \frac{\mu}{\lambda} B\mathbf{y} = \mathbf{y}$.
In other words, the matrix $I - \frac{\mu}{\lambda} B$ is invertible.
This was on the assumption that $|\lambda| > \mu$.
It follows that $B$ has no eigenvalues with absolute value greater than $\mu$.
The Banach fixed point theorem gives us more than this though. It tells us that the fixed point $\mathbf{y}$ of $g_{\mathbf{x}}$ is given by

$$\lim_{n\to\infty} \mathbf{y}_n, \quad \mathbf{y}_{n+1} = \mathbf{x} + \frac{\mu}{\lambda} B\mathbf{y}_n.$$

We can choose any $\mathbf{y}_0$ and will choose $\mathbf{y}_0 = \mathbf{x}$. Then an easy induction gives

$$\mathbf{y}_n = \sum_{i=0}^{n} \frac{\mu^i}{\lambda^i} B^i \mathbf{x}.$$

So

$$\left(I - \frac{\mu}{\lambda} B\right)^{-1} \mathbf{x} = \sum_{i=0}^{\infty} \frac{\mu^i}{\lambda^i} B^i \mathbf{x}.$$

# Geometric series

From

$$\left(I - \frac{\mu}{\lambda}B\right)^{-1} = \sum_{i=0}^{\infty} \frac{\mu^i}{\lambda^i} B^i$$

we get

$$(\lambda I - A)^{-1} = \lambda^{-1} C^{-1} \left(I - \frac{\mu}{\lambda}B\right)^{-1} C = \sum_{i=0}^{\infty} \frac{\mu^i}{\lambda^{i+1}} C^{-1} B^i C = \sum_{i=0}^{\infty} \frac{1}{\lambda^{i+1}} A^i.$$

This is what we would expect, formally, but what's useful here is the convergence.