On the ubiquity of Perron theorem

José Antonio de la Peña
UNAM and CONACyT

EMALCA 2010, Quito, Ecuador.
Survey of the talk.

- The aim of the talk is to describe the ubiquitous Perron-Frobenius theorem (PF in the sequel), and discuss some connections with diverse areas, such as:
  1. topology (Brouwer fixed-point theorem)
  2. Graph theory (spectral graph theory)
  3. probability theory (finite-state Markov chains)
  4. mathematical economy (Leontieff model)
  5. other applications (Google...)
  6. Some recent results on social networks.
We first state a special case of the theorem, due to Perron.

In the sequel, we shall use the notation $B > 0$ (resp. $B \geq 0$) for any (possibly even rectangular) matrix with positive (resp. non-negative) entries.

The German mathematician Oskar Perron proved in 1907:

**Theorem:** Let $A = ((a_{ij})) > 0$ be a square matrix, and let $r(A) = \sup \{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$ be the spectral radius of $A$.

Then

- $r(A)$ is an eigenvalue of $A$ of (algebraic, hence also geometric) multiplicity one, and (a suitably scaled version $v$ of)
- the corresponding eigenvector has strictly positive entries;
- $|\lambda| < r(A)$ for all eigenvalues of $A$.
On the proof.

Proof: Let $\Delta$ be the standard simplex in $\mathbb{R}^n$. Observe that for any $v \in \Delta$ if we set $r_v := \min (Av)_i / v_i$ we get $r_v v \leq Av$. Take now

$$r' := \sup \{ r > 0 : \text{there exists } w \geq 0 \text{ such that } rw \leq Aw \} ;$$

Similarly, for $r^v := \max (Av)_i / v_i$ we get $r^v v \geq Av$. Take

$$r' := \inf \{ r > 0 : \text{there exists } v \geq 0 \text{ such that } rv \geq Av \}.$$

Show: $r' = r(A) = r''$
The extension by Frobenius.

The Frobenius extension of the theorem relaxes the strict positivity assumption on $A$. Call $A$ irred
cucibile if it satisfies either of the following equivalent conditions:

(i) there does not exist a permutation matrix $P$ such that $PAP^T$ has the form $A' \oplus A''$ for some matrices $A'$ and $A''$ of strictly smaller size;

(ii) $\forall \ i, j \ \exists \ m > 0$ such that $(A^m)_{ij} > 0$.

Important example: given a graph $G$ define $A(G)=(a_{ij})$ the adjacency matrix, by $a_{ij} = 1$ (or 0) if $i \neq j$ (or not).

Then $A(G)$ is irreducible if and only if $G$ is connected.
The Perron-Frobenius theorem.

Theorem: Suppose $A \geq 0$ is an irreducible square matrix. Then

- (i) $r(A)$ is an eigenvalue of $A$ of (algebraic, hence also geometric) multiplicity one, and (a suitably scaled version $v$ of) the corresponding eigenvector has strictly positive entries;
- (ii) The only non-negative eigenvectors of $A$ are multiples of $v$;
- (iii) If there are $k$ eigenvalues $\lambda$ of $A$ such that $|\lambda| = r(A)$ then the set of eigenvalues of $A$ is invariant under a rotation about the origin by $2\pi k$. 
Relation with Brouwer fixed point theorem.

Brouwer Fixed Point Theorem: any continuous function
\[ h : B^n \rightarrow B^n \] where \( B^n \) is the unit n-dimensional ball
has a fixed point.

- Proof of PF using BFP theorem:
  If \( A \geq 0 \) is irreducible, and if \( v \in \Delta \), then \( Av \neq 0 \).
  Hence \( |Av| \neq 0 \) and we may define
  \[ h : \Delta \rightarrow \Delta, \text{ by } h(v) = \frac{Av}{|Av|} \]
  which is continuous. Let \( v \) denote the fixed point guaranteed by Brouwer.
- Conversely, BFP theorem can be proved using PF theorem.
Matrices with invariant cones

The standard simplex $\Delta$ in $\mathbb{R}^n$ is an example of a cone:

(a) Closed set and $0 \in \Delta$;
(b) additively invariant $\Delta + \Delta \subseteq \Delta$;
(c) Invariant by multiplication with positive scalars;
(d) Contains a basis of $\mathbb{R}^n$ but $\Delta \cap (-\Delta) = \{0\}$

walls of cones are cones in $\mathbb{R}^{n-1}$
Let $\Delta$ be a cone in $\mathbb{R}^n$ and $A$ a $n \times n$ matrix. We say that $A$ leaves $\Delta$ (properly) invariant if $A(\Delta) \subseteq \Delta$ (but no wall of $\Delta$ is invariant).

**Theorem:** Suppose $A$ leaves a cone properly invariant. Then

- (i) $r(A)$ is an eigenvalue of $A$ of (algebraic, hence also geometric) multiplicity one, and (a suitably scaled version $v$ of) the corresponding eigenvector has strictly positive entries;
- (ii) The only non-negative eigenvectors of $A$ are multiples of $v$;
- (iii) If there are $k$ eigenvalues $\lambda$ of $A$ such that $|\lambda| = r(A)$ then the set of eigenvalues of $A$ is invariant under a rotation about the origin by $2\pi k$. 
Let $A=(a_{ij})$ be a real $n \times n$-matrix,
The **spectrum** of $A$ is the set of eigenvalues of $A$, ie
the roots of the **characteristic polynomial**:
$$\varphi_A(t) = \det (t \text{id}_n - A) = t^n - a_1 t^{n-1} - \ldots - a_{n-1} t + a_n.$$ 
Hence:
- $a_1 = -\text{tr } A$ and $a_n = (-1)^n \det (A)$.
- If $A$ is **symmetric** then the eigenvalues are real numbers:
  $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n = r(A)$;
The weight of the vertices

adjacency matrix

\[
\begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
a & a & a \\
0 & 0 & 1 & 0
\end{pmatrix}
\]

= \begin{pmatrix} 1 \\ 1 \\ a \\ 1 \end{pmatrix}

\[
a \cdot \begin{pmatrix}
1 \\
1 \\
a \\
1
\end{pmatrix}
\]

\[
a^2 = 3 \quad \text{and} \quad a > 1; \quad \text{weights:} \quad \sqrt{3} \quad 1 \\
1 \quad 1
\]

eigenvector corresponding to the Perron-eigenvalue \( a \).
Weights of the nodes of a network

Each node in a network has a different ‘weight’ according to the role of the node at the connectivity of the network. This weight is determined by the Perron Theorem. Which node (the red or the blue has higher weight?

A famous application of the Perron algorithm occurs at the Google browser.
Consider the path $P_n$:

\[
1 \quad \ldots \quad 2 \quad \ldots \quad n
\]

Lemma:

For $j = 1, \ldots, n$ let $x^{(j)} = (x_k^{(j)}) \in \mathbb{R}^n$ with

\[
x_k^{(j)} = \sin \left( \frac{jk\pi}{n+1} \right), \quad k = 1, \ldots, n.
\]

Then the set \( \{x^{(1)}, \ldots, x^{(n)}\} \) forms a basis of $\mathbb{R}^n$ consisting of eigenvectors respective eigenvalues

\[
\lambda_j = 2 \cos \left( \frac{j\pi}{n+1} \right).
\]

Moreover, we have $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. 
Another example.

The characteristic polynomial of $A(G)$ is

$$x^{10} - 17 x^8 - 14 x^7 + 90 x^6 + 144 x^5 - 89 x^4 - 338 x^3 - 253 x^2 - 56 x + 4 = 0$$

with roots $r=3.24,\ 2.72,\ 2,\ 0.56,\ -1$ (with multiplicity 4), $-2 \ y \ -2.52$. The Perron eigenvector (corresponding to $r$) is depicted explicitly. The red vertex is the most central in the network.

Observe that $D=4$ is the Dunbar number and $c(S) = 3$, we should have 

$$c(S) \leq r \leq D.$$
Let $A=(a_{ij})$ be a non-negative $n \times n$-matrix associated to a weighted graph $(G,a)$ with $G$ connected.

The $s$-th power $A^s=(a_{ij}(s))$ satisfies that:

- $a_{ij}(s)$= walks of length $s$ from $i$ to $j$;
- $r= \lim_{s \to \infty} s\sqrt{a_{ij}(s)}$, for any choice of $i$ and $j$.

**Corollary:** the sequence of matrices $A^s$, $s \in \mathbb{N}$, converges if and only if $r(A) < 1$ or ($r(A)=1$ is the unique eigenvalue with norm 1).

Moreover:

- $\min \{ \sum_k a_{ik} : i \} \leq r \leq \max \{ \sum_k a_{ik} : i \}$
- if $H$ is a subgraph of $G$, then $r(H) \leq r(G)$
Coloring a map.

Two countries are joined by an edge if they share a border. Each edge should have extremes of different colors! Four colors are enough in this case (always?).

The minimal number of colors to color a graph $G$ is the chromatic number $C(G)$.

Theorem: $C(G) \leq r + 1$.

Proof: there is a subgraph $H$ with $\text{val}_{\text{min}}(H) \geq C(G)-1$. Then $C(G)-1 \leq r(H) \leq r(G)$.
Stochastic process.

0 ≤ A=(aij) is **stochastic** if Σ_j a_ij =1 for all i. That is exactly when 1 is an eigenvalue with eigenvector (1,...,1)^T. The number a_ij is interpreted as the probability of the system to transit from state i to the state j.

**Proposition:** The limit A^\infty = \lim A^s exists and all rows of A^\infty are (p_1,...,p_n), where p_i is the probability of the system to be at state i.

**Example:** a rat in a chamber moves to any other chamber with probability b and stays with probability a.

\[
\begin{pmatrix}
a & b & b & b \\
b & a & b & b \\
b & b & a & b \\
b & b & b & a \\
\end{pmatrix}
\]

and all p_i= ¼
The factory producing $P_j$ requires to buy from other factories $P_i$ a total of $a_{ij}$ units (per unit of production). If the production is $x=(x_j)$, the production remaining for internal profit of the factories is: $y = x - Ax$.

What should be the production in order that the profit $y > 0$?

**Theorem (Leontief, 1958):** Given $y > 0$ there is always $x > 0$ such that $y = x - Ax > 0$ if and only if $r(A) < 1$.

Indeed, the condition $r(A)< 1$ implies that the inverse of $(id-A)$ is $A + A^2 + A^3 + \ldots$ a well-defined non-negative matrix.
An algebraic application: Galois coverings.

Let $G$ be a graph (finite or infinite with bounded degree). Let $\mathcal{G}$ be a group acting on $G$ and $p: G \rightarrow G/\mathcal{G}$ a quotient (=Galois covering defined by $\mathcal{G}$).

If $G$ is finite the eigenvalues of $G$ and $G/\mathcal{G}$ are the same (with different multiplicities).

In case $G$ is infinite we observe that: $r(G) = \lim r(F)$, where $F$ runs over all finite full subgraphs of $G$. Then:

$$r(G/\mathcal{G}) \leq r(G) \leq r(G/\mathcal{G})^2.$$ 

Theorem (JAP-Takane, 1992): $r(G/\mathcal{G}) = r(G)$ if and only if the group $\mathcal{G}$ is amenable.
Two examples:

- $g = \text{free non-commutative}$
- $g = \text{integers}$
Growth of groups::

Let $G$ be a group with a finite set $F = F^{-1}$ of generators. We consider the elements $B(n)$ ($=n$-th ball) generated by at most $n$ elements in $F$.

$B(n) \subseteq B(n+1)$ determines the growth.

$d(B(n)) = \text{border of } B(n)$

G is *amenable* if $\lim |d(B(n))| / |B(n)| = 0$

$B(1) = 5$, $d(B(1)) = 4$, $B(2) = B(1) + 3$, $d(B(2)) = 12$

$\lim |d(B(n))| / |B(n)| = \frac{1}{2}$

the free group in $x, y$ is not amenable
Networks
Why networks?

Creating qualitative new levels of structure.

The neural system: from cells to brains.
Scientific journals.

The history of scientific journals dates from 1665, when the French *Journal des sçavans* and the English *Philosophical Transactions of the Royal Society* first began systematically publishing research results. Over a thousand, not long lasting journals were founded in the 18th century, and the number has increased rapidly after that.
The network of scientific journals.

Scientific journals are dots at the net and citations are the edges. Color stand for disciplinary classification.
Social networks: the Dunbar number.

- Group size is a function of relative neocortical volume in non-human primates.
- Extrapolation from this regression equation yields a predicted group size for modern humans very similar to that of certain hunter-gatherer and traditional horticulturalist societies. Groups of similar size are also found in other large-scale forms of contemporary and historical societies.
- Among primates, the cohesion of groups is maintained by social grooming; the time devoted to social grooming is linearly related to group size among the monkeys and apes.

<table>
<thead>
<tr>
<th>Primate</th>
<th>max. group size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibbons</td>
<td>15</td>
</tr>
<tr>
<td>Gorillas</td>
<td>33</td>
</tr>
<tr>
<td>Orang utan</td>
<td>50</td>
</tr>
<tr>
<td>Chimpanze</td>
<td>65</td>
</tr>
<tr>
<td>Humans</td>
<td>150</td>
</tr>
</tbody>
</table>
The Dunbar number in human societies.

Let $S$ be the graph of a social network, with $c(x)$ denoting the number of neighbors of $x$ in $S$, $c(x) \leq D$ (Dunbar 150).

A clan is a society where all individuals have close bonds and share norms, values and property. As a graph a clan is a complete graph $K_s$ with $s \leq D$.

Examples:
- extended families;
- military unities (around 150 soldiers since roman times);
- scientific subdisciplines (around 150-200 experts);
- tribal societies (Yanomamo in Brazil, bushpeople in Southafrica).
Social networks: Tribal societies.

In a tribal society $S$ each individual belongs to exactly a clan in $S$. Then:

- $d(S) =$ average of $d(x)$ for $x$ in $S$ and $D(S)$ is the Dunbar number of $S$, we get $d(S) \leq r(S) \leq D(S)$;
- if $C$ is a clan with $m$ members in $S$, then $m-1 \leq r(S)$;
- if $S$ is a tribal society, the number of clans satisfies $\text{cl}(S) \leq r(S)^2$ and the size of the society is $n(S) \leq r(S)^3$;
- this would predict at most 22,000 individuals at yanomamo society. Changnon reported 15,000 inhabitants;
- max. distance between vertices dimeter $d(S)$, satisfies $n(S) \leq D(S)^{d(S)/2}$ (small world property).
Small worlds

- Small-world
  - [Watts, Strogatz]
    - 6 degrees of separation
    - Small diameter

- Effective diameter:
  - Distance at which 90% of pairs of nodes are reachable
Turán's theorem.

Let $G$ be any subgraph of $K_n$ such that $G$ is $K_{r+1}$-free. Then the number of edges in $G$ is at most

$$\frac{1}{2} \left(1 - \frac{1}{r}\right)n(n-1)$$

An equivalent formulation is the following:
Among the $n$-vertex simple graphs with no $(r + 1)$-cliques, $T(n,r)$ has the maximum number of edges.

$T(7,3)$:

As a special case, for $r = 2$, one obtains Mantel's theorem: The maximum number of edges in an $n$-vertex triangle-free graph is $\leq \frac{n^2}{4}$. 
Cliques and the trace function.

**Proposition.** We have \( \text{tr} (A^3) \leq \sum_{i=1}^{n} d(i)^2 \).

1.2. **Complete graphs.** The complete graph \( K_r \) has characteristic polynomial \( P_{K_r}(t) = (t + 1)^{r-1}(t - r + 1) \), that is, the eigenvalues are \(-1\), with multiplicity \( r - 1 \), and \( r - 1 \). Therefore \( 2e(K_r) = \text{tr} (A(K_r)^2) = (r - 1)^2 + (r - 1) = (r - 1)r \) and \( 6t(G) = \text{tr} (A(K_r)^3) = (r - 1)^3 - (r - 1) = (r - 2)(r - 1)r \).

**Proposition.** Let \( G \) be a graph as above accepting an \( r \)-clique \( K_r \), then the following holds:

1. \( \text{tr} (A^3) \geq (r - 2)(r - 1)r \);
2. the spectral radius \( r - 1 \leq \lambda_n \), equality holds if and only if \( G = K_r \);
3. for \( 1 \leq i \leq r \) and \( n - r \leq j \leq n \) we have \( \lambda_i \leq -1 \leq \lambda_j \).
A local version of Turán’s theorem.

**Theorem 1.** Let $A$ be the adjacency matrix of $G$ and choose $r \geq 3$ with the property that

$$\text{tr} \ (A^3) \geq \frac{1}{2D} (1 - \frac{1}{r-2}) \sum_{i=1}^{n} d(i)^2$$

then the following holds:

(a) $G$ contains a complete subgraph $K_r$;
(b) there are at least $\frac{n \cdot 1}{12 \cdot D}$ vertices of $G$ belonging to $r$-cliques.
THANK YOU

GRACIAS.