Perron's Theorem

November 17, 2018

Perron’s theorem is about square matrices in which all entries are positive real numbers.

Such matrices will be called positive matrices.

Proved by Oskar Perron - 1907.

Further extended by Frobenius -1912.

We use the following notation:

If A is an $n \times n$ matrix, then the ijth entry will be denoted by

$$a_{ij}.$$

We will write $A = [a_{ij}]$.

A will be called a positive if $a_{ij} > 0$ for all i, j.

We shall write

$$A > 0 \iff \text{A is positive.}$$

A will be called non-negative if $a_{ij} \geq 0$ for all i, j.

We shall write

$$A \geq 0 \iff \text{A is non-negative.}$$

Positive matrices are non-negative matrices.

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is non-negative but not positive.

Let A be an $n \times n$ real matrix.

$f(x) := \det(xI - A)$ is called the characteristic polynomial of A.

f is monic. (Meaning: Coefficient of $x^n = 1$)

$\deg(f) = n$.

There exist $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that $f(\lambda_i) = 0$.

There exist non-zero vectors $x^1, \ldots, x^n$ in $\mathbb{C}^n$ such that

$$Ax^i = \lambda_i x^i.$$

λ_i are eigenvalues of A.

x^i are eigenvectors of A.

Define
$$\rho(A) := \max\{|\lambda_1|, \ldots, |\lambda_n|\}.$$

$\rho(A)$ is called the spectral radius of A.

Note: $\rho(A) \geq 0$.

If A is non-zero, can we say that $\rho(A) > 0$?

Find 2×2 matrices such that:

Spectral radius is an eigenvalue.

Spectral radius is not an eigenvalue.

We will say that a vector$x$ in $\mathbb{R}^n$ is positive if all the components are positive.

Similarly, we have non-negative vectors.

For example, $(1, 2, 8) \in \mathbb{R}^3$ is a positive vector whereas $(1, 2, 0)$ is not positive.

$(1, 2, -1)$ is not a non-negative vector.

Vectors in $\mathbb{R}^n$ will be regarded as row/column vectors depending on the context.

What is the meaning for Some vector is not non-negative?

Theorem (Perron)

Let $A > 0$. Then,

(i) $\rho(A)$ *is an eigenvalue of* A.

(ii) *There exists* $y > 0$ *such that*

$$Ay = \rho(A)y \quad y > 0.$$

(iii) *Geometric multipilcity of* $\rho(A)$ *is* 1.

(iv) *Suppose* $Ax = \lambda x$ *for some* $x > 0$. *Then,* x *and* y *are linearly dependent.*

(v) *If* $Av = \lambda v$ *for some* $v > 0$, *then* $\lambda = \rho(A)$.

Example: Consider the following $n \times n$ matrix.

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

- Verify Perron's theorem.

To prove Perron's theorem, we need some tools.

First, we need to measure the distance between any two elements in $\mathbb{R}^n$.

• Take $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$.

• $\|x - y\|$ will be the notation used to denote the distance between x and y.

• $\|x\|$ is the distance between x and the origin.

•Definition: $\|x\| := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

If we know how to calculate the distance, then we can define the convergence of a sequence in $\mathbb{R}^n$.

Consider a sequence of vectors $\{x^1, x^2, x^3 \dots\}$ in $\mathbb{R}^n$.

Write the sequence: $\{x^k\}_{k=1}^{\infty}$.

We say that

$$\lim_{k \to \infty} x^k = p$$

if for each $\epsilon > 0$, there exists a positive integer M such that

$$k \geq M \quad \Rightarrow \quad \|x^k - p\| < \epsilon$$

•Show that the sequence $\{(1,0), (-1,0), (1,0), \dots\}$ is not convergent.

Closed sets in $\mathbb{R}^n$:

A set $F \subseteq \mathbb{R}^n$ is called closed if any convergent sequence $\{x^m\}$ in F satisfies the following:

$$\lim_{m \to \infty} x^m = x \quad \Rightarrow \quad x \in F.$$

Examples:

$\mathbb{R}^n_+$ is a closed set in $\mathbb{R}^n$.

$\mathbb{R}^n_{++}$ is not closed. Why?

• BOUNDED SETS.

A set E in $\mathbb{R}^n$ is bounded if there exists $K > 0$ such that

$$x \in E \Rightarrow \quad \|x\| \leq K.$$

• COMPACT SETS

A set K in $\mathbb{R}^n$ is said to be compact iff K is closed and bounded.

Unbounded sets: Sets that are not bounded.

• Show that

$$\Delta := \{x \in \mathbb{R}^n : x \geq 0, \quad \sum_{i=1}^{n} x_i \leq 1\}$$

is compact.

$E := (0,1) \times [0,1]$ is not closed in $\mathbb{R}^2$.

$\{(1/m, 1/2)\}_{m=1}^{\infty}$ is a sequence in E.

Limit of this sequence is the vector $(0, 1/2)$.

But this is not a point in E.

$E := \{(y_1, y_2) \in \mathbb{R}^2 : y_1 \geq 0, \;\; y_2 \geq 0\}$ is closed but not compact.

$\{(n, n) : n \in N\}$ is an unbounded sequence in E.

Convex sets

Let S be a closed set in $\mathbb{R}^n$. It is called convex iff

$$x \in S, \quad y \in S \quad \Rightarrow \quad \frac{x+y}{2} \in S.$$

- Let
$$\Delta := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}.$$
Show that Δ is convex.

-
$$\nabla := \{(x_1, x_2) : x_!^2 + x_2^2 = 1\}.$$
is not convex.

First assertion (Proof)

First step in the proof:

Some positive number is always an eigenvalue of A, by using the assumption that all entries of A are positive.

The proof of this will be completed by using Brouwer's fixed point theorem.

Theorem (Brouwer)

Let $\Delta \subseteq \mathbb{R}^n$ be a compact and convex set. Suppose $f : \Delta \to \Delta$ is a continuous function. Then there exists $x \in \Delta$ such that

$$f(x) = x.$$

If $n = 1$, then the proof is easy.

Otherwise, proof is not straightforward. (One way is by using degree theoretic techniques.)

If $A > 0$, then for any non-negative vector $0 \neq y$, $Ay > 0$. (Why?)

This is not true if we just assume that $A \geq 0$.

Example?

To apply Brouwer's theorem, we need a set Δ which is compact and convex and we need a continuous function $f : \Delta \to \Delta$.

f must be associated to A.

Define

$$\Delta := \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0\}.$$

Show that Δ is compact and convex.

• Construction of a continuous self-map.

We are given an $n \times n$ matrix A such that $A > 0$.

For $x \in \Delta$, let $g(x)$ be defined by

$$g(x) := (Ax)_1 + (Ax)_2 + \ldots + (Ax)_n.$$

$(Ax)_i > 0$ for all i.

$g(x) \neq 0$ for all $x \in \Delta$.

If $A \geq 0$, and if g is as above, then $g(p)$ may be 0 for some p. For example, consider

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Put $e_1 = (1, 0)^T$.

Then, $Ae_1 = 0$.

$e_1 \in \Delta$.

$Ae_1 = 0 \quad \Rightarrow \quad (Ae_1)_1 + (Ae_1)_2 = 0$.

So, $g(e_1) = 0$.

Define $f : \Delta \to \mathbb{R}^n$ by

$$f(x) = \frac{1}{g(x)} Ax.$$

- f is continuous.
- $$f(x)_1 + f(x)_2 + \cdots + f(x)_n = 1.$$

Each $f(x)_i \geq 0$.

So, $x \in \Delta \Rightarrow \quad f(x) \in \Delta$.

Hence, $f : \Delta \to \Delta$.

We can apply Brouwer's theorem to f and conclude the following:

There exists $y \in \Delta$ such that

$$f(y) = y.$$

But $f(y) = \frac{1}{g(y)} Ay$.

So, $f(y) = y$ implies

$$Ay = g(y)y.$$

- $y \in \Delta$.
- y is an eigenvector of A.
- $g(y) > 0$ is an eigenvalue of A.
- $y > 0$ (Why?)

To this end, we have shown the following:

Lemma

If $A > 0$, then there exists a positive real number δ and $y > 0$ such that

$$Ay = \delta y.$$

To complete the first assertion of Perron's theorem, we shall prove the following claim:

Claim: $\delta = \rho(A)$.

By definition, $\delta \leq \rho(A)$.

$$Ay = \delta y, \;\; y > 0, \;\; \delta > 0.$$

Apply the above lemma to A^T.

- There exists $u \in \mathbb{R}^n$ and $\alpha > 0$ such that

$$A^T u = \alpha_2 u.$$

We first show that $\alpha_2 = \delta$.

$$u^T A = \alpha_2 u^T.$$

Post-multiply by y.

$$u^T A y = \alpha_2 u^T y. \tag{1}$$

Also,

$$Ay = \delta y, \quad y > 0.$$

Since $u, y > 0, \quad u^T y > 0$.

By (1),

$$u^T \delta y = \alpha_2 u^T y$$

So,

$$(\alpha_2 - \delta) u^T y = 0.$$

This gives $\alpha_2 = \delta$.

So far, we have shown the following:

Lemma

Let A be an $n \times n$ positive matrix. Then, there exists $\delta > 0$ and $x, u \in \mathbb{R}^n$ such that

$$Ax = \delta x, \;\; A^T u = \delta u, \;\; x > 0, \; u > 0.$$

Next step:

Claim: $\delta = \rho(A)$.

We shall prove this claim.

Let $x := (x_1, \ldots, x_n)^T$ and $y := (y_1, \ldots, y_n)^T)$ be any two vectors in $\mathbb{R}^n$.

We write

$$x \geq y \iff (x - y)_i \geq 0 \;\; \forall i = 1 : n.$$

Equivalently,

$$x_i \geq y_i \;\; \forall i = 1 : n.$$

We now define the following:

Given a vector $x = (x_1, \dots, x_n)$ in $\mathbb{C}^n$, we define

$$m : \mathbb{C}^n \to \mathbb{R}^n$$

by

$$m(x) := (|x_1|, \dots, |x_n|).$$

Note: $m(x)$ is a non-negative vector in $\mathbb{R}^n$.

i.e. $m(x) \geq 0$ for all $x \in \mathbb{C}^n$.

For any $\eta \in \mathbb{C}$, $m(\eta x) = |\eta| m(x)$.

Remark

If $u \in \mathbb{C}^n$, then

$$Am(u) \geq m(Au).$$

Proof:

$$m(u) = (|u_1|, \ldots, |u_n|).$$

$$Am(u) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} |u_1| \\ |u_2| \\ \vdots \\ |u_n| \end{bmatrix}.$$

Let $v := Am(u)$.

$$v_1 = a_{11}|u_1| + a_{12}|u_2| + \cdots + a_{1n}|u_{1n}|.$$

$$v_2 = a_{21}|u_1| + a_{22}|u_2| + \cdots + a_{2n}|u_{1n}|.$$

$$\vdots$$

$$v_n = a_{n1}|u_1| + a_{n2}|u_2| + \cdots + a_{nn}|u_{1n}|.$$

Let $w = Au$.

$$w_1 = a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_{1n}$$
$$w_2 = a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_{1n}$$
$$\vdots$$
$$w_n = a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_{1n}.$$

By Triangle inequality

$$
\begin{aligned}
(Am(u))_1 = v_1 &\geq |a_{11}u_1 + a_{12}|u_{12} + \cdots + a_{1n}u_{1n}| \\
&= |w_1| \\
&= |(Au)_1|.
\end{aligned} \tag{2}
$$

Continuing this,

$$(Am(u))_i \geq |(Au)_i| \forall \;\; i.$$

Thus,

$$Am(u) - m(Au) \geq 0 \;\; \forall u \in \mathbb{C}^n.$$

Consider an eigenvalue of A.

Let $Ap = \mu p, \quad 0 \neq p \in \mathbb{C}^n$.

Now,

$$\begin{aligned} Am(p) &\geq m(Ap) \\ &= m(\mu p) \\ &= |\mu| p. \end{aligned}$$

Now use the following:

$$A^T u = \delta u, \quad u > 0.$$

Pre-multiply by u^T in

$$Am(p) \geq |\mu| p.$$

This gives

$$u^T Am(p) \geq u^T |\mu| m(p).$$

Use $u^T A = \delta u^T$.

Now,

$$\delta u^T m(p) \geq |\mu| u^T m(p).$$

Note: $u^T m(p) > 0$.

So, $\delta \geq |\mu|$.

For any eigenvalue of μ of A, we have

$$\delta \geq |\mu|.$$

So, $\delta \geq \rho$.

Thus, $\delta = \rho(A)$. The proof is complete.

So far, we have proved the following:

If $A > 0$, then there exists $x \in \mathbb{R}^n$ such that

$$Ax = \rho(A)x, \quad x > 0.$$

Third assertion:

Geometric multiplicity of $\rho(A)$ is one.

Meaning:

$$\text{nullspace}(A - \rho(A)I) = \{x \in \mathbb{C}^n : Ax = \rho(A)x\}.$$

To show: $\dim(\text{nullspace}(A) - \rho(A)I) = 1$.

Let $x, y \in \text{nullspace}(A - \rho(A)I)$.

Claim: x and y are linearly dependent.

Case 1: Assume $y \in \mathbb{R}^n$. Suppose x and y are LI.

There exists $\alpha \in \mathbb{R}$ such that

$$x - \alpha y \geq 0, \quad (x - \alpha y)_i = 0 \text{ for some } i.$$

Because x, y are LI, $x - \alpha y \neq 0$.

$x - \alpha y$ is an eigenvector of A.

As $A > 0$, $A(x - \alpha y) > 0$.

This is a contradiction, as for some i, $(x - \alpha y)_i = 0$.

Case 2: Suppose

$$y \in \mathbb{C}^n, \quad Ay = \rho(A)y.$$

Let $p := realpart(y)$ and $q := Imaginarypart(y)$.

$$A(p + iq) = \rho(A)p + i\rho(A)q.$$

$$Ap = \rho(A)p$$
$$Aq = \rho(A)q.$$

Apply case 1 to get:
$p = \delta_1 x$ and $q = \delta_2 x$.

y is a multiple of x.

Proof is complete.

We have thus shown the following:

If $A > 0$, then there exists $x > 0$ such that $Ax = \rho(A)x$. Further, geometric multiplicity of $\rho(A)$ is one.

Final assertion:
Suppose $Aw = \mu w$ and $w > 0$.

Claim: $\mu = \rho(A)$.

Let $A^T u = \rho(A)u$, where $u > 0$.

$$w^T A^T = \mu w^T.$$

This gives

$$w^T A^T u = \rho(A) w^T u = \mu w^T u.$$

So, $\rho(A) = \mu$.

Matrices for which spectral radius is an eigenvalue

Suppose A is symmetric.

Assume all the entries in A are positive.

Then, proving that spectral radius is an eigenvalue is easy.

(i) All eigenvalues are real.

(ii)
$$\lambda_{max}(A) := \max\{x^T Ax : \|x\| = 1\}.$$

(iii) $\lambda_{max}(A) = \rho(A)$.

Positive definite matrices

A matrix $A \in \mathbb{R}^{n \times n}$ is positive definite if it is symmetric and $x^T A x > 0$ for all $0 \neq x \in \mathbb{R}^n$.

A is positive definite if and only if all the eigenvalues of A are positive.

$A = B^2$ for some $B \in \mathbb{R}^{n \times n}$.

These matrices have spectral radius as an eigenvalue. (Trivial).

Copositive matrices

Let A be an $n \times n$ matrix. Then, A is said to be copositive if:

$$x \geq 0 \quad \Rightarrow \quad x^T A x \geq 0.$$

• Non-negative matrices are copositive matrices.

• Positive semidefinite matrices are copositive matrices.

• Sum of a non-negative matrix and a positive semidefinite matrix is a copositive matrix.

• All symmetric copositive matrices upto order 4 can be written as a sum: $N + P$

• For higher orders: Not true.

Theorem (Haynsworth and Hoffman)

Let A be an $n \times n$ symmetric matrix. Suppose A is copositive. Then, $\rho(A)$ is an eigenvalue.

Given a vector $x \in \mathbb{R}^n$, it can be written as

$$x = x^+ - x^-.$$

$$x^+ = \max(x, 0) \quad x^- = \max(-x, 0).$$

$$\max(x, 0) = (\max(x_1, 0), \max(x_2, 0), \ldots, \max(x_n, 0))$$
$$\max(-x, 0) = (\max(-x_1, 0), \max(-x_2, 0), \ldots, \max(-x_n, 0))$$

Consider $x = (2, -3)$ in $\mathbb{R}^2$.

$x^+ = (2, 0)$ and $x^- = (0, 3)$.

x^+ and x^- are non-negative vectors.

x^+ and x^- are orthogonal.

Recall:

$$m(x) = (|x_1|, |x_2|, \ldots, |x_n|).$$

$$m(x) = x^+ + x^-.$$

If $x \in \mathbb{R}^n$, then $\|x\| = \|m(x)\|$.

$$\begin{aligned} \|x\|^2 &= \|x^+ - x^-\|^2 \\ &= \langle x^+, x^+ \rangle + \langle x^-, x^- \rangle - 2\langle x^+, x^- \rangle \\ &= \|x^+\|^2 + \|x^-\|^2 \\ &= \|x^+ + x^-\|^2 \\ &= \|m(x)\|^2. \end{aligned} \tag{3}$$

Lemma

If A is symmetric and copositive, then for any vector $x \in \mathbb{R}^n$,

$$\langle Am(x), m(x) \rangle + \langle Ax, x \rangle \geq 0.$$

Our aim is to show that $\rho(A)$ is an eigenvalue of A.

In otherwords, $\rho(A) = \lambda_{\max}(A)$.

If

$$\lambda_{\max}(A) + \lambda_{\min}(A) \geq 0,$$

then $\rho(A) = \lambda_{\max}(A)$.

Proof.

$$\lambda_{\max}(A) = \max\{x^T A x : \|x\| = 1\}.$$
$$\lambda_{\min}(A) = \min\{x^T A x : \|x\| = 1\}.$$

Let

$$\lambda_{\max}(A) = q^T A q.$$
$$\lambda_{\min}(A) = p^T A p.$$

Note:

$$\|p\| = 1, \quad \|q\| = 1.$$

□

$$\langle Ap, p\rangle + \langle Am(p), m(p)\rangle \geq 0.$$

$$\begin{aligned} -\langle Ap, p\rangle &\leq \langle Am(p), m(p)\rangle \\ &\leq \langle Aq, q\rangle \end{aligned} \tag{4}$$

So,

$$\langle Aq, q\rangle + \langle Ap, p\rangle \geq 0.$$

This means that $\lambda_{\max}(A) + \lambda_{\min}(A) \geq 0$.

Proof is complete.