Algebraic graph theory and centrality measures

Giacomo Como, DISMA, Politecnico di Torino

Fabio Fagnani, DISMA, Politecnico di Torino

Weight of walks and powers of the weight matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$\gamma = (i_0, i_1, \dots, i_l)$ walk, $W_\gamma = \prod_{1 \leq h \leq l} W_{i_{h-1} i_h}$

Weight of walks and powers of the weight matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$\gamma = (i_0, i_1, \dots, i_l)$ walk, $W_\gamma = \prod_{1 \le h \le l} W_{i_{h-1} i_h}$

Theorem

$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Weight of walks and powers of the weight matrix

Theorem

$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Weight of walks and powers of the weight matrix

Theorem
$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Proof By induction on $l \geq 1$. $l = 1$ trivial.

Weight of walks and powers of the weight matrix

Theorem

$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Proof By induction on $l \geq 1$. Assume it true for l.

Weight of walks and powers of the weight matrix

Theorem
$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma)=l}} W_\gamma$$

Proof By induction on $l \geq 1$. Assume it true for l.

$$\sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma)=l+1}} W_\gamma$$

Weight of walks and powers of the weight matrix

Theorem
$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Proof By induction on $l \geq 1$. Assume it true for l.

$$\sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l+1}} W_\gamma = \sum_k \sum_{\substack{\tilde{\gamma} \text{ walk from } i \text{ to } k \\ l(\tilde{\gamma}) = l}} W_{\tilde{\gamma}} W_{kj}$$

Weight of walks and powers of the weight matrix

Theorem
$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Proof By induction on $l \geq 1$. Assume it true for l.

$$\sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l+1}} W_\gamma = \sum_k \sum_{\substack{\tilde{\gamma} \text{ walk from } i \text{ to } k \\ l(\tilde{\gamma}) = l}} W_{\tilde{\gamma}} W_{kj}$$

$$= \sum_k (W^l)_{ik} W_{kj} = (W^{l+1})_{ij}$$

Weight of walks and powers of the weight matrix

Theorem

$$(W^l)_{ij} = \sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l}} W_\gamma$$

Proof By induction on $l \geq 1$. Assume it true for l.

$$\sum_{\substack{\gamma \text{ walk from } i \text{ to } j \\ l(\gamma) = l+1}} W_\gamma = \sum_k \sum_{\substack{\tilde{\gamma} \text{ walk from } i \text{ to } k \\ l(\tilde{\gamma}) = l}} W_{\tilde{\gamma}} W_{kj}$$

$$= \sum_k (W^l)_{ik} W_{kj} = (W^{l+1})_{ij}$$

It is proven for $l+1$. By induction result is proven.

Properties of the products of the weight matrix

Theorem

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a graph. Then,

1. $(W^l)_{ij} > 0$ *if and only if there exists a walk of length* l *from* i *to* j*;*
2. $\mathcal{G}$ *is connected iff for every* $i, j \in \mathcal{V}$*, there exists* $l > 0$ *such that* $(W^l)_{ij} > 0$*.*
3. $\mathcal{G}$ *is connected and aperiodic iff there exists* $N > 0$ *such that* $(W^N)_{ij} > 0$ *for every* $i, j \in \mathcal{V}$*.*

Properties of the products of the weight matrix

Theorem

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a graph. Then,

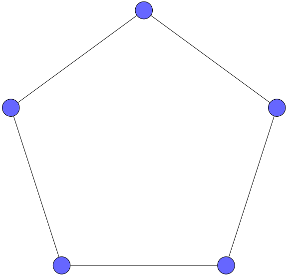
1. $(W^l)_{ij} > 0$ *if and only if there exists a walk of length l from i to j;*
2. $\mathcal{G}$ *is connected iff for every $i, j \in \mathcal{V}$, there exists $l > 0$ such that $(W^l)_{ij} > 0$.*
3. $\mathcal{G}$ *is connected and aperiodic iff there exists $N > 0$ such that $(W^N)_{ij} > 0$ for every $i, j \in \mathcal{V}$.*

Comments on the proof:

- 1. and 2. are consequences of previous theorem.
- 3. is more involved.

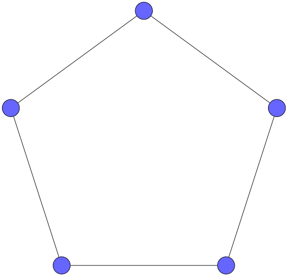
Examples

C_5



Examples

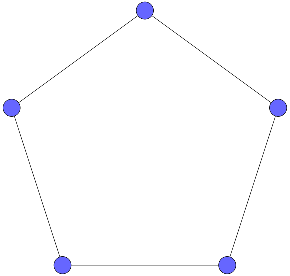
C_5



$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Examples

C_5

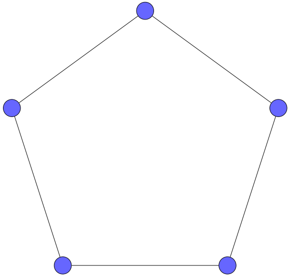


$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$W^2 = \begin{pmatrix} 2 & 0 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 & 1 \\ 1 & 0 & 2 & 0 & 1 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 2 \end{pmatrix}$$

Examples

C_5

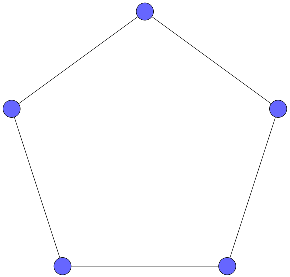


$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$W^3 = \begin{pmatrix} 0 & 3 & 1 & 1 & 3 \\ 3 & 0 & 3 & 1 & 1 \\ 1 & 3 & 0 & 3 & 1 \\ 1 & 1 & 3 & 0 & 3 \\ 3 & 1 & 1 & 3 & 0 \end{pmatrix}$$

Examples

C_5

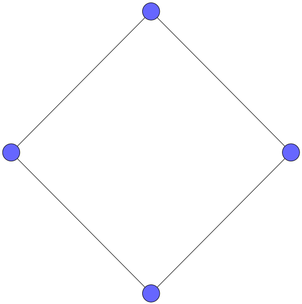


$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$W^4 = \begin{pmatrix} 6 & 1 & 4 & 4 & 1 \\ 1 & 6 & 1 & 4 & 4 \\ 4 & 1 & 6 & 1 & 4 \\ 4 & 4 & 1 & 6 & 1 \\ 1 & 4 & 4 & 1 & 6 \end{pmatrix}$$

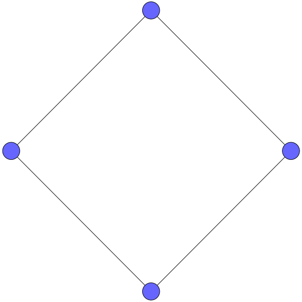
Examples

C_4



Examples

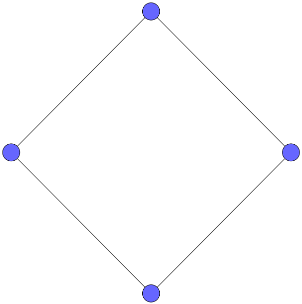
C_4



$$W = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Examples

C_4

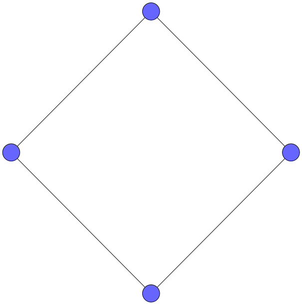


$$W = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

$$W^2 = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$

Examples

C_4

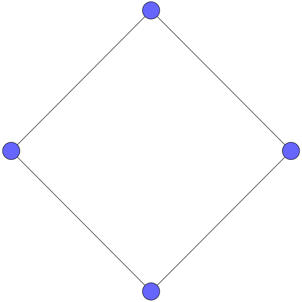


$$W = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

$$W^3 = \begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}$$

Examples

C_4



$$W = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

$$W^4 = \begin{pmatrix} 8 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \\ 8 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \end{pmatrix}$$

Counting objects in simple graphs

Corollary

For a simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, we have that:

(i) $(W^2)_{ii} = w_i$ *for every* $i \in \mathcal{V}$;

(ii) $Tr(W^2) = |\mathcal{E}|$;

(iii) $Tr(W^3) = 6 \cdot$ *number of triangles.*

The normalized weight matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph, $w_i = \sum_j W_{ij}$ out-degrees

$P_{ij} = w_i^{-1} W_{ij}$ *normalized weight matrix* of $\mathcal{G}$.

More compactly, $P = D^{-1}W$ where D is diagonal with $D_i = w_i$

P is a *stochastic* matrix: $P_{ij} \geq 0$, $P\mathbb{1} = \mathbb{1}$

The normalized weight matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph, $w_i = \sum_j W_{ij}$ out-degrees

$P_{ij} = w_i^{-1} W_{ij}$ *normalized weight matrix* of $\mathcal{G}$.

More compactly, $P = D^{-1} W$ where D is diagonal with $D_i = w_i$

P is a *stochastic* matrix: $P_{ij} \geq 0$, $P\mathbb{1} = \mathbb{1}$

- Topology of $\mathcal{G} \leftrightarrow$ Spectral properties of P;
- Through P we can describe interesting dynamical systems over $\mathcal{G}$;
- P can be interpreted as the transition matrix of a Markov chain, a *random walk* over $\mathcal{G}$.

The Laplacian matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph, D diagonal matrix with $D_i = w_i$ out-degrees

$$L = D - W$$

$$L_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ w_i - W_{ii} & \text{if } i = j \end{cases}$$

The Laplacian matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph, D diagonal matrix with $D_i = w_i$ out-degrees

$$L = D - W$$

$$L_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ w_i - W_{ii} & \text{if } i = j \end{cases}$$

$-L$ is called a Metzler matrix (non-negative off diagonal), $L\mathbb{1} = 0$

The Laplacian matrix

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph, D diagonal matrix with $D_i = w_i$ out-degrees

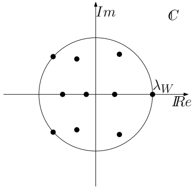
$$L = D - W$$

$$L_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ w_i - W_{ii} & \text{if } i = j \end{cases}$$

$-L$ is called a Metzler matrix (non-negative off diagonal), $L\mathbb{1} = 0$

- Topology of $\mathcal{G} \leftrightarrow$ Spectral properties of L;
- Through L we can describe interesting dynamical systems over $\mathcal{G}$ in continuous time;

Perron-Frobenius (PF) theory



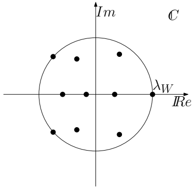
$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$$

Theorem (Perron-Frobenius)

There exists $\lambda_W \geq 0$ and non-negative vectors $x \neq 0$, $y \neq 0$ s.t.

- $Wx = \lambda_W x, \;\; W'y = \lambda_W y$;
- *every eigenvalue μ of W is such that $|\mu| \leq \lambda_W$;*

Perron-Frobenius (PF) theory



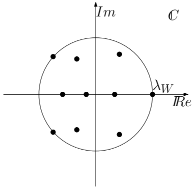
$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$$

Theorem (**Perron-Frobenius**)

There exists $\lambda_W \geq 0$ and non-negative vectors $x \neq 0$, $y \neq 0$ s.t.

- $Wx = \lambda_W x, \quad W'y = \lambda_W y;$
- *every eigenvalue μ of W is such that $|\mu| \leq \lambda_W$;*
- *If $\mathcal{G}$ is connected, then λ_W is simple;*

Perron-Frobenius (PF) theory



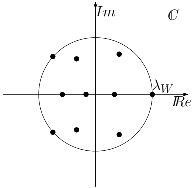
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

Theorem (**Perron-Frobenius**)

There exists $\lambda_W \geq 0$ and non-negative vectors $x \neq 0$, $y \neq 0$ s.t.

- $Wx = \lambda_W x, \;\; W'y = \lambda_W y;$
- *every eigenvalue μ of W is such that $|\mu| \leq \lambda_W$;*
- *If $\mathcal{G}$ is connected, then λ_W is simple;*
- *If $\mathcal{G}$ is connected and aperiodic, then every eigenvalue $\mu \neq \lambda_W$ is s.t. $|\mu| < \lambda_W$.*

Perron-Frobenius (PF) theory



$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

Theorem (**Perron-Frobenius**)

There exists $\lambda_W \geq 0$ and non-negative vectors $x \neq 0$, $y \neq 0$ s.t.

- $Wx = \lambda_W x, \;\; W'y = \lambda_W y;$
- *every eigenvalue μ of W is such that $|\mu| \leq \lambda_W$;*
- *If $\mathcal{G}$ is connected, then λ_W is simple;*
- *If $\mathcal{G}$ is connected and aperiodic, then every eigenvalue $\mu \neq \lambda_W$ is s.t. $|\mu| < \lambda_W$.*

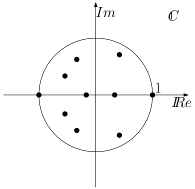
λ_W *dominant eigenvalue* of W. Spectral radius $\rho(W) = \lambda_W$.

PF theory for stochastic matrices

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$P = D^{-1}W$ normalized weight matrix

PF theory for stochastic matrices



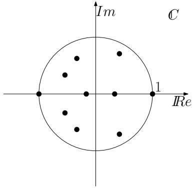
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$P = D^{-1}W$ normalized weight matrix

Theorem (Spectral properties of stochastic matrices)

- $\lambda_P = 1$;
- $P\mathbb{1} = \mathbb{1}$, $\exists \pi \geq 0$ *s.t.* $\mathbb{1}'\pi = 1$ *and* $P'\pi = \pi$;
- *every eigenvalue* μ *of* P *is such that* $|\mu| \leq 1$;

PF theory for stochastic matrices



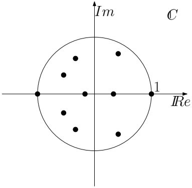
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$P = D^{-1}W$ normalized weight matrix

Theorem (Spectral properties of stochastic matrices)

- $\lambda_P = 1$;
- $P\mathbb{1} = \mathbb{1}$, $\exists \pi \geq 0$ *s.t.* $\mathbb{1}'\pi = 1$ *and* $P'\pi = \pi$;
- *every eigenvalue* μ *of* P *is such that* $|\mu| \leq 1$;
- *If* $\mathcal{G}$ *is connected, then* 1 *is simple and* $\pi_i > 0$ *for all* i;
- *If* $\mathcal{G}$ *is connected and aperiodic, then every eigenvalue* $\mu \neq 1$ *is s.t.* $|\mu| < 1$.

PF theory for stochastic matrices



$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$P = D^{-1}W$ normalized weight matrix

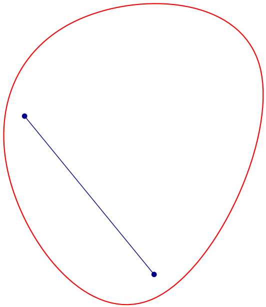
Theorem (Spectral properties of stochastic matrices)

- $\lambda_P = 1$;
- $P\mathbb{1} = \mathbb{1}$, $\exists \pi \geq 0$ *s.t.* $\mathbb{1}'\pi = 1$ *and* $P'\pi = \pi$;
- *every eigenvalue* μ *of* P *is such that* $|\mu| \leq 1$;
- *If* $\mathcal{G}$ *is connected, then* 1 *is simple and* $\pi_i > 0$ *for all* i;
- *If* $\mathcal{G}$ *is connected and aperiodic, then every eigenvalue* $\mu \neq 1$ *is s.t.* $|\mu| < 1$.

π invariant distribution

A digression: convex sets...

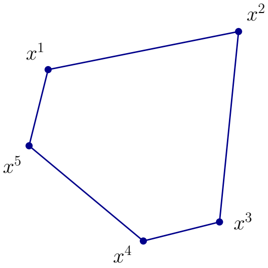
A subset $\mathcal{K} \subseteq \mathbb{R}^n$ is called *convex* if given any two points in $\mathcal{K}$, the segment joining them is all inside $\mathcal{K}$



polytopes...

A *polytope* is a convex subset that can be obtained by taking convex combinations of $k+1$ vectors $x^1, \ldots, x^{k+1} \in \mathbb{R}^n$,

$$\mathcal{K} = \{x = \sum_i \lambda_i x^i \mid \lambda_i \geq 0, \sum_i \lambda_i = 1\}$$



x^i extremal points of the polytope.

Simplexes

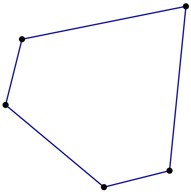
A *k-simplex* is a convex subset that can be obtained by taking convex combinations of $k+1$ vectors $x^1, \dots, x^{k+1} \in \mathbb{R}^n$, s.t. $x^i - x^1$ are independent vectors

$$\mathcal{K} = \{x = \sum_i \lambda_i x^i \mid \lambda_i \geq 0, \ \sum_i \lambda_i = 1\}$$

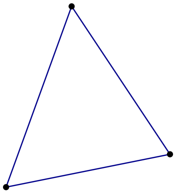
Simplexes

A *k-simplex* is a convex subset that can be obtained by taking convex combinations of $k+1$ vectors $x^1, \dots, x^{k+1} \in \mathbb{R}^n$, s.t. $x^i - x^1$ are independent vectors

$$\mathcal{K} = \{x = \sum_i \lambda_i x^i \mid \lambda_i \geq 0,\ \sum_i \lambda_i = 1\}$$



polytope, not a symplex

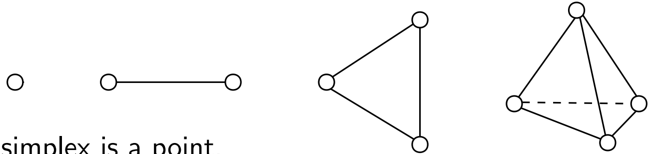


symplex

Simplexes

A *k-simplex* is a convex subset that can be obtained by taking convex combinations of $k+1$ vectors $x^1, \dots, x^{k+1} \in \mathbb{R}^n$, s.t. $x^i - x^1$ are independent vectors

$$\mathcal{K} = \{x = \sum_i \lambda_i x^i \mid \lambda_i \geq 0, \ \sum_i \lambda_i = 1\}$$



- 0-simplex is a point,
- 1-simplex is a segment,
- 2-simplex is a triangle,
- ...

Simplexes

Example: $x^i = e^i \in \mathbb{R}^n$ canonical basis

$$\mathcal{K} = \{x = (\lambda_1, \ldots, \lambda_n); \mid \lambda_i \geq 0, \sum_i \lambda_i = 1\}$$

$(n-1)$-simplex of probability vectors in $\mathbb{R}^n$

Simplexes

Example: $x^i = e^i \in \mathbb{R}^n$ canonical basis

$$\mathcal{K} = \{x = (\lambda_1, \dots, \lambda_n); \mid \lambda_i \geq 0, \sum_i \lambda_i = 1\}$$

$(n-1)$-simplex of probability vectors in $\mathbb{R}^n$

P stochastic matrix: $P_{ij} \geq 0$, $\sum_j P_{ij} = 1$

$$p \in \mathcal{K} \;\Rightarrow\; P' p \in \mathcal{K}$$

Simplexes

Example: $x^i = e^i \in \mathbb{R}^n$ canonical basis

$$\mathcal{K} = \{x = (\lambda_1, \dots, \lambda_n); \mid \lambda_i \geq 0,\ \sum_i \lambda_i = 1\}$$

$(n-1)$-simplex of probability vectors in $\mathbb{R}^n$

P stochastic matrix: $P_{ij} \geq 0$, $\sum_j P_{ij} = 1$

$$p \in \mathcal{K} \ \Rightarrow \ P'p \in \mathcal{K}$$

Perron-Frobenius theory $\Rightarrow \exists$ a fixed point $\pi \in \mathcal{K}$: $\quad P'\pi = \pi$

PF theory for stochastic matrices

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A more refined result.

PF theory for stochastic matrices

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A more refined result.

$\mathcal{H}_{\mathcal{G}}$ condensation graph, $s_{\mathcal{G}}$ number of sinks in $\mathcal{H}_{\mathcal{G}}$.

Theorem

- *Invariant distributions* π

$$\pi \geq 0,\ \mathbb{1}'\pi = 1,\ P'\pi = \pi$$

  form a simplex in $\mathbb{R}^{\mathcal{V}}$ *with* $s_{\mathcal{G}}$ *vertices.*
- *For every sink component with nodes* $\mathcal{W}$*, there exists an invariant distribution* π *such that* $\pi_i > 0$ *if and only if* $i \in \mathcal{W}$*.*
- *The invariant distribution is unique if and only if* $s_{\mathcal{G}} = 1$*.*

The special case of balanced graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A remarkable fact.

Suppose $\mathcal{G}$ is balanced ($w_i = w_i^-$ for every $i \in \mathcal{V}$). Then,

The special case of balanced graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A remarkable fact.

Suppose $\mathcal{G}$ is balanced ($w_i = w_i^-$ for every $i \in \mathcal{V}$). Then,

$$\sum_i w_i P_{ij} = \sum_i w_i w_i^{-1} W_{ij} = \sum_i W_{ij} = w_j^- = w_j$$

The special case of balanced graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A remarkable fact.

Suppose $\mathcal{G}$ is balanced ($w_i = w_i^-$ for every $i \in \mathcal{V}$). Then,

$$\sum_i w_i P_{ij} = \sum_i w_i w_i^{-1} W_{ij} = \sum_i W_{ij} = w_j^- = w_j$$

$$P'w = w$$

The special case of balanced graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $P = D^{-1}W$.

A remarkable fact.

Suppose $\mathcal{G}$ is balanced ($w_i = w_i^-$ for every $i \in \mathcal{V}$). Then,

$$\sum_i w_i P_{ij} = \sum_i w_i w_i^{-1} W_{ij} = \sum_i W_{ij} = w_j^- = w_j$$

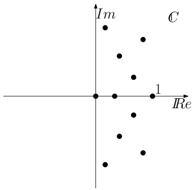
$$P'w = w$$

- the out-degree vector w is an eigenvector of eigenvalue 1
- $\pi = \frac{1}{|\mathcal{E}|} w$ is an invariant distribution of $\mathcal{G}$!
- If the graph is connected, this is the unique invariant distribiution.

Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$L = D - W$ Laplacian matrix



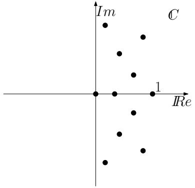
Theorem (Spectral properties of the Laplacian)

- *$L\mathbb{1} = 0$, 0 is an eigenvalue of L*
- *$\exists \bar{\pi} \geq 0$ s.t. $\mathbb{1}'\bar{\pi} = 1$ and $L'\bar{\pi} = 0$;*
- *All other eigenvalues λ have $\mathcal{R}e(\lambda) > 0$;*
- *If $\mathcal{G}$ is connected, then 0 is simple and $\bar{\pi}_i > 0$ for all i;*
- *$L'\bar{y} = 0 \Leftrightarrow P'(D\bar{y}) = (D\bar{y})$;*

Laplacian

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

$L = D - W$ Laplacian matrix

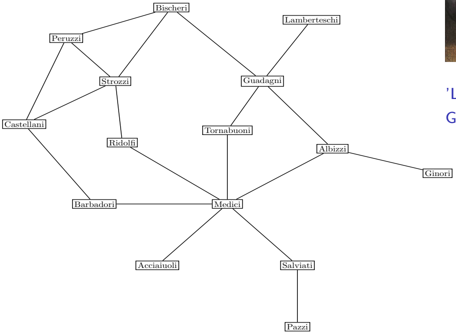


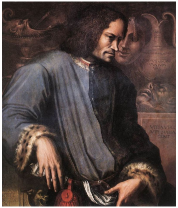
Theorem (Spectral properties of the Laplacian)

- *$L\mathbb{1} = 0$, 0 is an eigenvalue of L*
- *$\exists \bar{\pi} \geq 0$ s.t. $\mathbb{1}'\bar{\pi} = 1$ and $L'\bar{\pi} = 0$;*
- *All other eigenvalues λ have $\mathcal{R}e(\lambda) > 0$;*
- *If $\mathcal{G}$ is connected, then 0 is simple and $\bar{\pi}_i > 0$ for all i;*
- *$L'\bar{y} = 0 \Leftrightarrow P'(D\bar{y}) = (D\bar{y})$;*

$\bar{\pi}$ Laplace-invariant distribution

Who is the most central node?

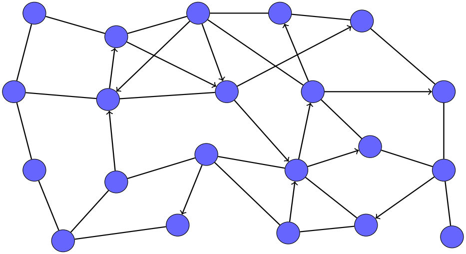




'Lorenzo de' Medici'
G. Vasari, 1534

Network centralities

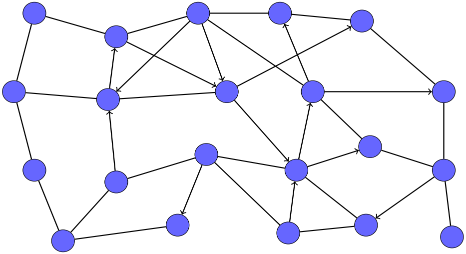
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



$z_i \geq 0$ centrality of node i

Degree centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

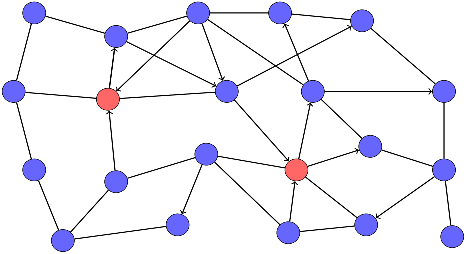


$z_i = w_i^-$ in-degree of node i *degree centrality*

Example: number of citations of an article, of followers on Twitter

Degree centrality

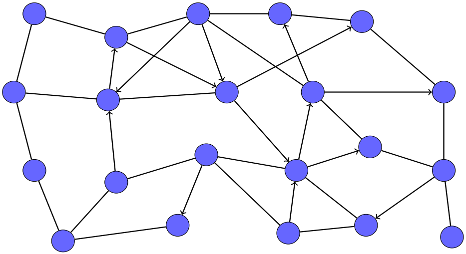
$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$$



The two nodes with the highest degree centrality.

Eigenvector centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

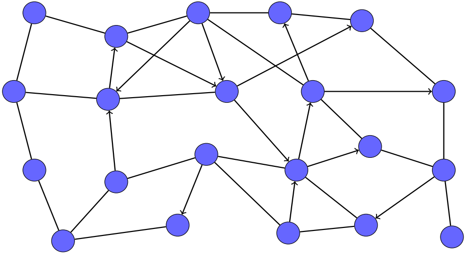


Drawback of degree-centrality: all incoming links are considered the same.

A different approach: $z_i \propto \sum_{j \in N_i^-} z_j$

Eigenvector centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

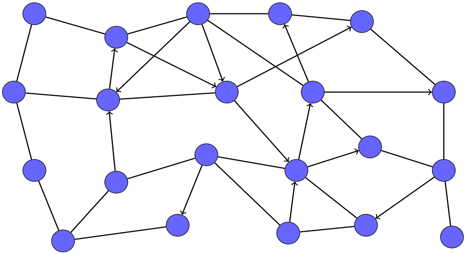


Drawback of degree-centrality: all incoming links are considered the same.

A different approach: $z_i \propto \sum_{j \in N_i^-} z_j = \sum_j W_{ji} z_j$

Eigenvector centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



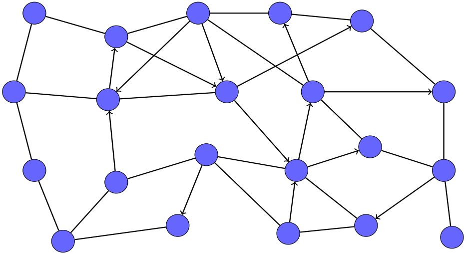
Drawback of degree-centrality: all incoming links are considered the same.

A different approach: $z_i \propto \sum_{j \in N_i^-} z_j = \sum_j W_{ji} z_j$

$\lambda z = W' z$ with $\lambda = \lambda_W$ *eigenvector centrality*

Eigenvector centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



$\lambda z = W'z$ with $\lambda = \lambda_W$

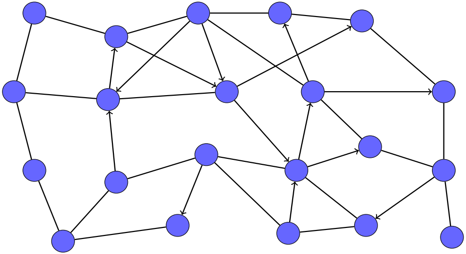
Remark: If all nodes have the same in-degree: $w_i^- = \delta$ for all i

$W'\mathbb{1} = \delta\mathbb{1} \Rightarrow z = \mathbb{1}$

Drawback: node j contributes proportional to its out-degree w_j

Invariant distribution centrality

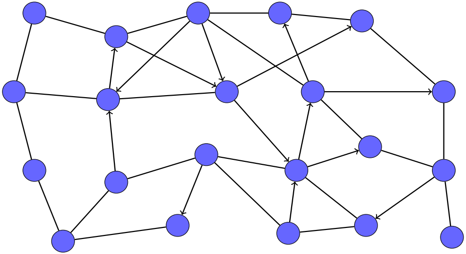
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



Another centrality measure: $z_i \propto \sum_{j \in N_i^-} \frac{1}{w_j} z_j$

Invariant distribution centrality

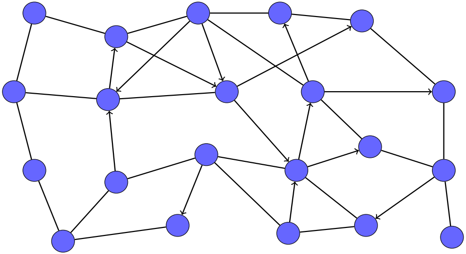
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



Another centrality measure: $z_i \propto \sum_{j \in N_i^-} \frac{1}{w_j} z_j = \sum_j P_{ji} z_j$

Invariant distribution centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

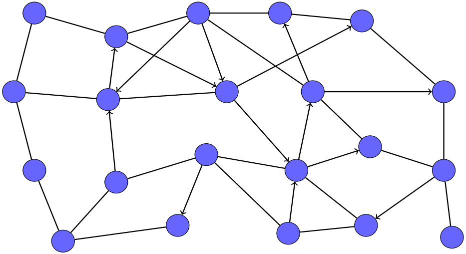


Another centrality measure: $z_i \propto \sum_{j \in N_i^-} \frac{1}{w_j} z_j = \sum_j P_{ji} z_j$

$z = P'z \Rightarrow$ *$z = \pi$ invariant distribution centrality*

Invariant distribution centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

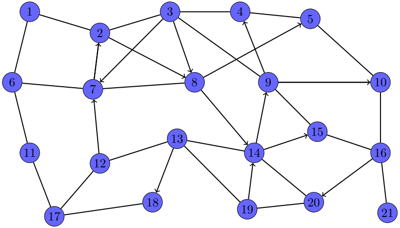


Another centrality measure: $z_i \propto \sum_{j \in N_i^-} \frac{1}{w_j} z_j = \sum_j P_{ji} z_j$

$z = P'z \Rightarrow$ $z = \pi$ *invariant distribution centrality*

If $\mathcal{G}$ is balanced, $\pi = w = w^- \Rightarrow$ inv. dist. centr. $=$ deg centr.

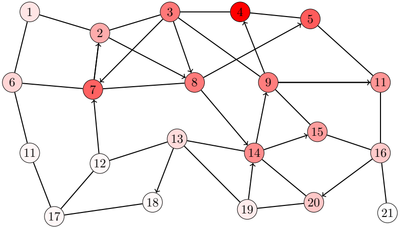
A comparison of the various centralities



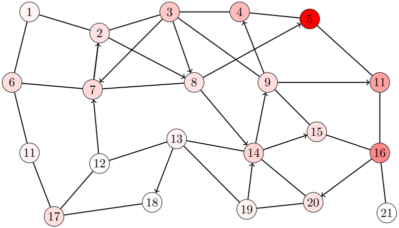
	Deg	Eig	Inv. dist.
1	0.0345	0.0348	0.0313
2	0.0517	0.0581	0.0451
3	0.0517	0.0664	0.0613
4	0.0517	0.0689	0.0680
5	0.0517	0.0680	0.0869
6	0.0517	0.0430	0.0490
7	0.0690	0.0678	0.0514
8	0.0517	0.0661	0.0444
9	0.0517	0.0659	0.0491
10	0.0517	0.0627	0.0761
11	0.0345	0.0226	0.0324
12	0.0345	0.0215	0.0240
13	0.0517	0.0399	0.0317
14	0.0690	0.0640	0.0548
15	0.0517	0.0613	0.0464
16	0.0517	0.0484	0.0817
17	0.0517	0.0225	0.0481
18	0.0345	0.0215	0.0240
19	0.0345	0.0307	0.0300
20	0.0517	0.0492	0.0441
21	0.0172	0.0166	0.0204

A comparison of the various centralities

Eigenvalue centrality



Invariant distribution centrality



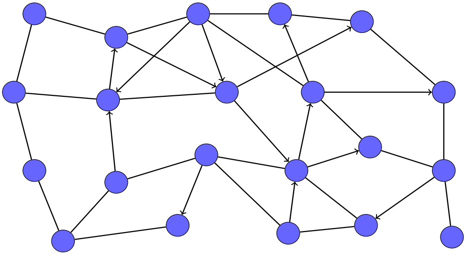
Drawbacks of eigenvalue and invariant distribution centr.

In general, they are not uniquely defined if the graph has more than one sink component

if the graph has just one sink component, only nodes in that component have non zero centrality

Katz centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



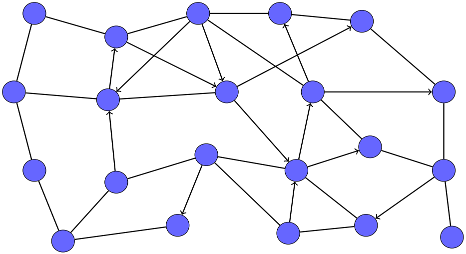
A convex comb. of *network centrality* and *intrinsic centrality*

$$z = \frac{1-\beta}{\lambda_W} W' z + \beta \mu$$

$z = (I - (1-\beta)\lambda_W^{-1} W')^{-1} \beta \mu$ *Katz centrality*

Bonacich centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



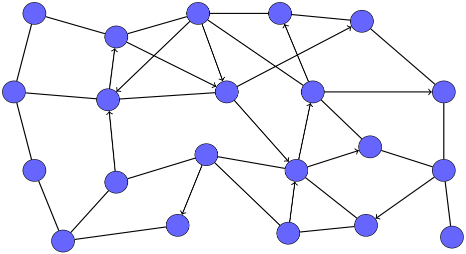
A convex comb. of *network centrality* and *intrinsic centrality*

$$z = (1-\beta)P'z + \beta\mu$$

$z = (I - (1-\beta)P')^{-1}\beta\mu$ *Bonacich centrality*

Bonacich centrality

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$



A convex comb. of *network centrality* and *intrinsic centrality*

$$z = (1-\beta)P'z + \beta\mu$$

$z = (I - (1-\beta)P')^{-1}\beta\mu$ *Bonacich centrality*

Page-rank (Google): $\mu = n^{-1}\mathbb{1}$, $\beta \sim 0.15$

The structure of Bonacich centrality

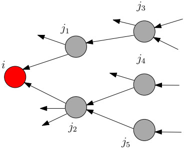
$$z = (I - (1-\beta)P')^{-1}\beta\mu$$

$$z_i = \beta \sum_{k=0}^{\infty} (1-\beta)^k \sum_j \mu_j (P^k)_{ji}$$

The structure of Bonacich centrality

$$z = (I - (1-\beta)P')^{-1}\beta\mu$$

$$z_i = \beta \sum_{k=0}^{\infty}(1-\beta)^k \sum_j \mu_j (P^k)_{ji}$$

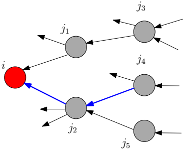


$$= \beta\left[\mu_i + (1-\beta)(\mu_{j_1}P_{j_1 i} + \mu_{j_2}P_{j_2 i}) + (1-\beta)^2(\mu_{j_3}P_{j_3 j_1}P_{j_1 i} + \mu_{j_4}P_{j_4 j_2}P_{j_2 i} + \mu_{j_5}P_{j_5 j_2}P_{j_2 i})\right] + \cdots$$

The structure of Bonacich centrality

$$z = (I - (1-\beta)P')^{-1}\beta\mu$$

$$z_i = \beta \sum_{k=0}^{\infty} (1-\beta)^k \sum_j \mu_j (P^k)_{ji}$$

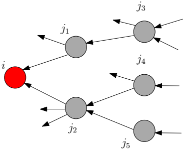


$$= \beta \left[\mu_i + (1-\beta)(\mu_{j_1} P_{j_1 i} + \mu_{j_2} P_{j_2 i}) + (1-\beta)^2(\mu_{j_3} P_{j_3 j_1} P_{j_1 i} + \mu_{j_4} P_{j_4 j_2} P_{j_2 i} + \mu_{j_5} P_{j_5 j_2} P_{j_2 i})\right] + \cdots$$

The structure of Bonacich centrality

$$z = (I - (1-\beta)P')^{-1}\beta\mu$$

$$z_i = \beta \sum_{k=0}^{\infty} (1-\beta)^k \sum_j \mu_j (P^k)_{ji}$$



$$= \beta \left[\mu_i + (1-\beta)(\mu_{j_1} P_{j_1 i} + \mu_{j_2} P_{j_2 i}) + (1-\beta)^2 (\mu_{j_3} P_{j_3 j_1} P_{j_1 i} + \mu_{j_4} P_{j_4 j_2} P_{j_2 i} + \mu_{j_5} P_{j_5 j_2} P_{j_2 i})\right] + \cdots$$

- The Bonacich centrality is uniquely defined for every graph $\mathcal{G}$
- $\beta = 0 \;\Rightarrow\;$ Intrinsic centrality (no network)
- $\beta \to 1 \;\Rightarrow\;$ Invariant distribution centrality (only network)

The various centralities

Eigenvector centrality	Invariant distribution
$z = \frac{1}{\lambda_W} W' z$	$z = P' z$
Katz centrality	Bonacich centrality
$z = \frac{1-\beta}{\lambda_W} W' z + \beta \mu$	$z = (1-\beta) P' z + \beta \mu$

An application in economy

The Review of Economic Statistics

VOLUME XVIII AUGUST, 1936 NUMBER 3

QUANTITATIVE INPUT AND OUTPUT RELATIONS IN THE ECONOMIC SYSTEM OF THE UNITED STATES

INTRODUCTION

The statistical study presented in the following pages may be best defined as an attempt to construct, on the basis of available statistical materials, a *Tableau Economique* of the United States for the year 1919.[1]

One hundred and fifty years ago, when Quesnay first published his famous schema, his contemporaries and disciples acclaimed it as the greatest "invention" since Newton's laws. The idea of general interdependence existing among the various parts of the economic system has become by now the very foundation of economic analysis. And yet, when it comes to the practical application of this theoretical tool, modern economists must rely exactly as Quesnay did upon fictitious numerical examples. What would be the present state of the theory and policy of international trade if, instead of actual balances of foreign trade, the economist had to base his analysis upon assumed numerical set-ups, supplemented by scattered items of actual statistical information? This is the situation in which the student of economics finds himself at present when he faces a problem of national production, consumption, and distribution. Despite the remarkable increase in the volume of primary statistical data, the proverbial boxes of theoretical assumptions are in this respect as empty as ever. Considerable progress has been achieved in the field of national income statistics. The economic balance of some of the most important branches of the national economy, particularly that of agriculture, has been studied with much success. Thus the ground has been prepared, at least in part, for a more complete analysis of the interrelations of the whole economic system. Nevertheless, the difficulty of the task still remaining can hardly be exaggerated.

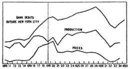
[1] The research project, a summary of the results of which are presented in this paper, has been financed by the Harvard University Committee on Research in the Social Sciences. Mr. Maynard C. Heins worked on it as a full-time assistant for four years. Without his able collaboration, the statistical task could hardly have been accomplished.

The publication of this preliminary survey is prompted by the conviction that the inevitable path of any empirical research is that of trial and error.

Governmental publications constitute the main source of primary statistical information used in this study. Additional data were gathered from trade publications, and in some instances the results of special investigations have been utilized. In many cases, use was made of the work of the National Bureau of Economic Research on national income.

At the time that this study was initiated (1932), the publication of the detailed results of the 1929 Census was still far from complete. As a result, the Census of 1919 had to be used. It is because of this fact that the entire investigation is based on 1919 data.

CHART 1.—SERIES INDICATIVE OF BUSINESS CONDITIONS



The general business conditions prevailing during that year are described in W. L. Thorp's *Business Annals* in the following terms:

> Revival; prosperity.
> Uncertainty gives way to extraordinary activity, late spring; building revival; enormous output of new securities; speculation; steel, coal, and railroad shopmen's strikes, autumn; commodity prices rise; active foreign trade.
> Money eases slightly but tightens late in year; stock exchange booms, railroads reaching peak, May, and industrials, November; falling bond prices; embargo on gold export removed, June.
> Large wheat, fair cotton and corn crops; prices very high.[2]

[2] *Business Annals* (National Bureau of Economic Research, New York, 1926), p. 143.

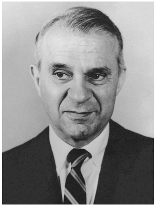
[105]

Wassily Leontief

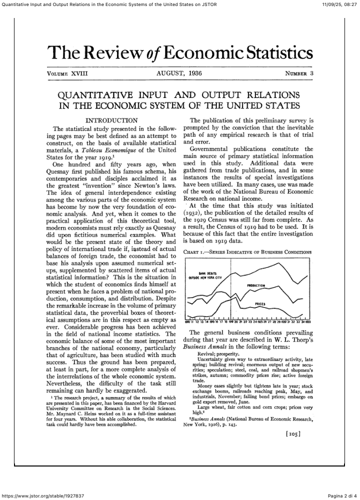
russian born

american naturalized

economist



An application in economy



The Review of Economic Statistics

VOLUME XVIII AUGUST, 1936 NUMBER 3

QUANTITATIVE INPUT AND OUTPUT RELATIONS IN THE ECONOMIC SYSTEM OF THE UNITED STATES

INTRODUCTION

The statistical study presented in the following pages may be best defined as an attempt to construct, on the basis of available statistical materials, a *Tableau Economique* of the United States for the year 1919.[1]

One hundred and fifty years ago, when Quesnay first published his famous schema, his contemporaries and disciples acclaimed it as the greatest "invention" since Newton's laws. The idea of general interdependence existing among the various parts of the economic system has become by now the very foundation of economic analysis. And yet, when it comes to the practical application of this theoretical tool, modern economists must rely exactly as Quesnay did upon fictitious numerical examples. What would be the present state of the theory and policy of international trade if, instead of actual balances of foreign trade, the economist had to base his analysis upon assumed numerical set-ups, supplemented by scattered items of actual statistical information? This is the situation in which the student of economics finds himself at present when he faces a problem of national production, consumption, and distribution. Despite the remarkable increase in the volume of primary statistical data, the proverbial boxes of theoretical assumptions are in this respect as empty as ever. Considerable progress has been achieved in the field of national income statistics. The economic balance of some of the most important branches of the national economy, particularly that of agriculture, has been studied with much success. Thus the ground has been prepared, at least in part, for a more complete analysis of the interrelations of the whole economic system. Nevertheless, the difficulty of the task still remaining can hardly be exaggerated.

[1] The research project, a summary of the results of which are presented in this paper, has been financed by the Harvard University Committee on Research in the Social Sciences. Mr. Maynard C. Heins worked on it as a full-time assistant for four years. Without his able collaboration, the statistical task could hardly have been accomplished.

The publication of this preliminary survey is prompted by the conviction that the inevitable path of any empirical research is that of trial and error.

Governmental publications constitute the main source of primary statistical information used in this study. Additional data were gathered from trade publications, and in some instances the results of special investigations have been utilized. In many cases, use was made of the work of the National Bureau of Economic Research on national income.

At the time that this study was initiated (1932), the publication of the detailed results of the 1929 Census was still far from complete. As a result, the Census of 1919 had to be used. It is because of this fact that the entire investigation is based on 1919 data.

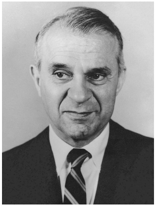
CHART 1.—SERIES INDICATIVE OF BUSINESS CONDITIONS

The general business conditions prevailing during that year are described in W. L. Thorp's *Business Annals* in the following terms:

> Revival; prosperity.
>
> Uncertainty gives way to extraordinary activity, late spring; building revival; enormous output of new securities; speculation; steel, coal, and railroad shopmen's strikes, autumn; commodity prices rise; active foreign trade.
>
> Money eases slightly but tightens late in year; stock exchange booms, railroads reaching peak, May, and industrials, November; falling bond prices; embargo on gold export removed, June.
>
> Large wheat, fair cotton and corn crops; prices very high.[2]

[2] *Business Annals* (National Bureau of Economic Research, New York, 1926), p. 143.

Wassily Leontief
russian born
american naturalized
economist



- The paper was initially a bit unappreciated, at the time of the Great Depression, eventually it got more and more attention, till the Nobel prize in 1973.

An application in economy

The Review *of* Economic Statistics

VOLUME XVIII | AUGUST, 1936 | NUMBER 3

QUANTITATIVE INPUT AND OUTPUT RELATIONS IN THE ECONOMIC SYSTEM OF THE UNITED STATES

INTRODUCTION

The statistical study presented in the following pages may be best defined as an attempt to construct, on the basis of available statistical materials, a *Tableau Economique* of the United States for the year 1919.[1]

One hundred and fifty years ago, when Quesnay first published his famous schema, his contemporaries and disciples acclaimed it as the greatest "invention" since Newton's laws. The idea of general interdependence existing among the various parts of the economic system has become by now the very foundation of economic analysis. And yet, when it comes to the practical application of this theoretical tool, modern economists must rely exactly as Quesnay did upon fictitious numerical examples. What would be the present state of the theory and policy of international trade if, instead of actual balances of foreign trade, the economist had to base his analysis upon assumed numerical set-ups, supplemented by scattered items of actual statistical information? This is the situation in which the student of economics finds himself at present when he faces a problem of national production, consumption, and distribution. Despite the remarkable increase in the volume of primary statistical data, the proverbial boxes of theoretical assumptions are in this respect as empty as ever. Considerable progress has been achieved in the field of national income statistics. The economic balance of some of the most important branches of the national economy, particularly that of agriculture, has been studied with much success. Thus the ground has been prepared, at least in part, for a more complete analysis of the interrelations of the whole economic system. Nevertheless, the difficulty of the task still remaining can hardly be exaggerated.

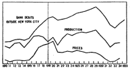
[1] The research project, a summary of the results of which are presented in this paper, has been financed by the Harvard University Committee on Research in the Social Sciences. Mr. Maynard C. Heins worked on it as a full-time assistant for four years. Without his able collaboration, the statistical task could hardly have been accomplished.

The publication of this preliminary survey is prompted by the conviction that the inevitable path of any empirical research is that of trial and error.

Governmental publications constitute the main source of primary statistical information used in this study. Additional data were gathered from trade publications, and in some instances the results of special investigations have been utilized. In many cases, use was made of the work of the National Bureau of Economic Research on national income.

At the time that this study was initiated (1932), the publication of the detailed results of the 1929 Census was still far from complete. As a result, the Census of 1919 had to be used. It is because of this fact that the entire investigation is based on 1919 data.

CHART 1.—SERIES INDICATIVE OF BUSINESS CONDITIONS

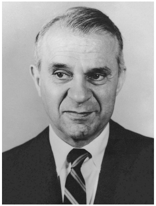


The general business conditions prevailing during that year are described in W. L. Thorp's *Business Annals* in the following terms:

> Revival; prosperity.
>
> Uncertainty gives way to extraordinary activity, late spring; building revival; enormous output of new securities; speculation; steel, coal, and railroad shopmen's strikes, autumn; commodity prices rise; active foreign trade.
>
> Money eases slightly but tightens late in year; stock exchange booms, railroads reaching peak, May, and industrials, November; falling bond prices; embargo on gold export removed, June.
>
> Large wheat, fair cotton and corn crops; prices very high.[2]

[2] *Business Annals* (National Bureau of Economic Research, New York, 1926), p. 143.

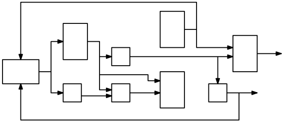
[105]



Wassily Leontief
russian born
american naturalized
economist

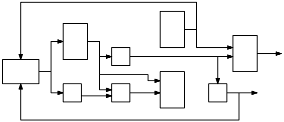
- The paper was initially a bit unappreciated, at the time of the Great Depression, eventually it got more and more attention, till the Nobel prize in 1973.
- It depicts a general economy as a set of productive sectors interconnected by input/output relations.

An application in economy: the open Leontief model



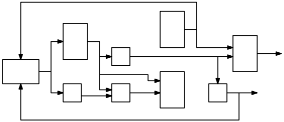
- $\mathcal{V}$ set of sectors (aggregates of firms) each producing just one good.

An application in economy: the open Leontief model



- $\mathcal{V}$ set of sectors (aggregates of firms) each producing just one good.
- Fixed proportions production: W_{ij} quantity of good i needed to produce a unit of product j.
- For i to produce a quantity x_i, a quantity $x_{ki} = W_{ki}x_i$ is needed from each other sector k

An application in economy: the open Leontief model



- $\mathcal{V}$ set of sectors (aggregates of firms) each producing just one good.
- Fixed proportions production: W_{ij} quantity of good i needed to produce a unit of product j.
- For i to produce a quantity x_i, a quantity $x_{ki} = W_{ki}x_i$ is needed from each other sector k
- Open Leontief model with i/o matrix $W \to \mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$
- External consumers create a demand of a quantity d_i of good i $\forall i \in \mathcal{V}$

Given the demand vector d, is the production system capable of exactly cover such demand?

An application in economy: input/output Leontief model

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, $d \in \mathbb{R}_+^{\mathcal{V}}$

Market clearing conditions $x_i = \sum_j x_{ij} + d_i = \sum_j W_{ij} x_j + d_i$

Equivalently, $x = Wx + d$ or $(I - W)x = d$.

The model is called *feasible* if for every $d \geq 0$, there exists a solution $x \geq 0$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Leftarrow$: $\rho(W) < 1$, $x = (I - W)^{-1} d$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Leftarrow$: $\rho(W) < 1$, $x = (I - W)^{-1}d = \sum_{k=0}^{+\infty} W^k d \geq 0$ model is feasible.

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Rightarrow$:

- Feasibility $\Rightarrow$ For $d = \mathbb{1}$ $\exists x > 0$ (all components positive) such that $x - Wx = \mathbb{1}$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Rightarrow$:

- Feasibility $\Rightarrow$ For $d = \mathbb{1}$ $\exists x > 0$ (all components positive) such that $x - Wx = \mathbb{1}$
- PF theorem $\Rightarrow$ $\exists y$ nonnegative non null such that $y'W = \rho(W)y$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Rightarrow$:

- Feasibility $\Rightarrow$ For $d = \mathbb{1}$ $\exists x > 0$ (all components positive) such that $x - Wx = \mathbb{1}$
- PF theorem $\Rightarrow$ $\exists y$ nonnegative non null such that $y'W = \rho(W)y$
- $y'\mathbb{1} = y'x - y'Wx = y'x - \rho(W)y'x = (1 - \rho(W))y'x$

An application in economy: input/output Leontief model

Theorem

An open Leontief model with i/o matrix W is feasible if and only if $\rho(W) < 1$

$\Rightarrow$:

- Feasibility $\Rightarrow$ For $d = \mathbb{1}$ $\exists x > 0$ (all components positive) such that $x - Wx = \mathbb{1}$
- PF theorem $\Rightarrow$ $\exists y$ nonnegative non null such that $y'W = \rho(W)y$
- $y'\mathbb{1} = y'x - y'Wx = y'x - \rho(W)y'x = (1 - \rho(W))y'x$
- $y'\mathbb{1} > 0$, $y'x > 0 \Rightarrow 1 - \rho(W) = y'\mathbb{1}/y'x > 0$

Some references

- A. Bermann, R. J. Plemmons; *Nonnegative matrices in the mathematical sciences*, SIAM, 1993.
- L. Katz; A new status index derived from sociometric analysis, *Psychometrica*, 18, pp. 3-43, 1953.
- P. Bonacich; Power and Centrality: A Family of Measures, *American Journal of Sociology*, 1987.
- W. W. Leontief; Quantitative Input and Output Relations in the Economic Systems of the United States, *The Review of Economic Statistics*, Vol XVIII (n. 3), pp. 105-125, 1936.