*(joint work with Sebastiano Vigna et al.)*

# Axioms for centrality: rank monotonicity for PageRank

**Paolo Boldi**
Università degli Studi di Milano

# TOC

# TOC

- Why centrality

# TOC

- Why centrality
  - Orienteering in the centrality jungle

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices
- Why axioms

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices
- Why axioms
  - Orienteering in the axiom jungle

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices
- Why axioms
  - Orienteering in the axiom jungle
  - Some important axioms

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices
- Why axioms
  - Orienteering in the axiom jungle
  - Some important axioms
- Focus on rank monotonicity for PageRank

# TOC

- Why centrality
  - Orienteering in the centrality jungle
  - Some important centrality indices
- Why axioms
  - Orienteering in the axiom jungle
  - Some important axioms
- Focus on rank monotonicity for PageRank
- Conclusions

# IR System

# IR System

# IR System

*Document Repertoire*

- pages retrieved by Google
- items on sale on Amazon
- members of facebook
- tweets posted by your friends
- photographs on Instagram

*D*

# IR System

$D$

# IR System

$D$

# IR System

$Q$

*Set of Queries* (query language)

- SE query
- product recommendation
- new-friend suggestion
- tweets to be shown

$D$

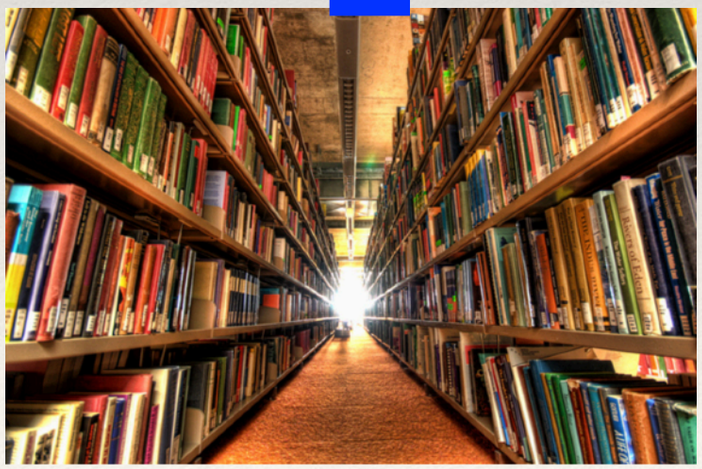
# IR System

$Q$

$D$

# IR System

$Q$

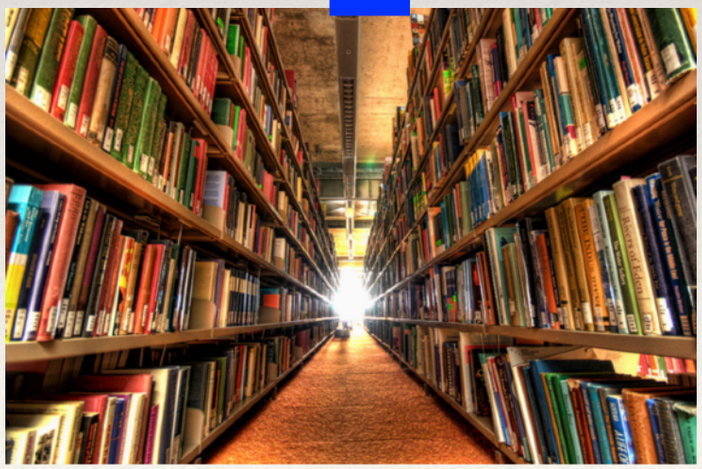
$D$

# IR System

$Q$


*Result*

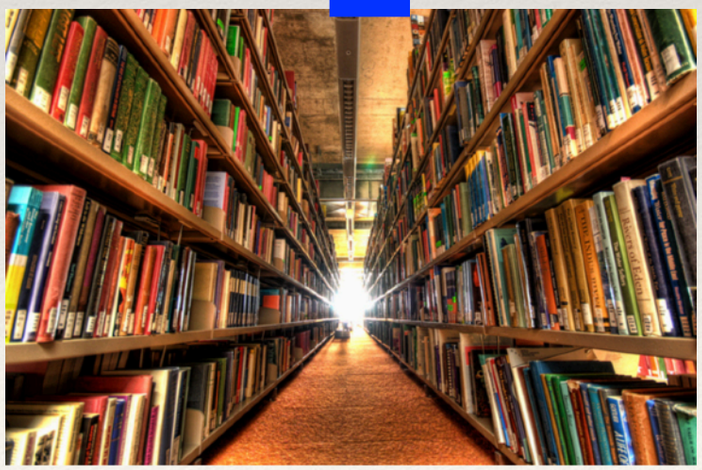
- a selected subset $S \subseteq D$
- with a score (typically: a non-negative real number) assigned to every element of $S$



$D$

# IR System

$Q$

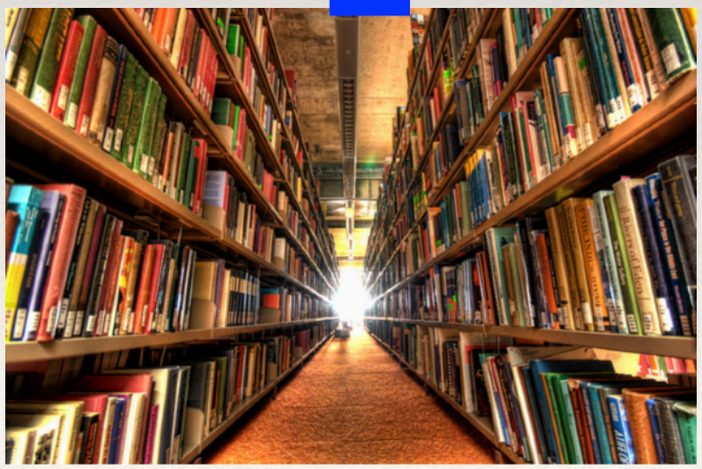
$D$

*Result*

- a selected subset $S \subseteq D$
- with a score (typically: a non-negative real number) assigned to every element of $S$

# IR System: 1st simplification

$Q$



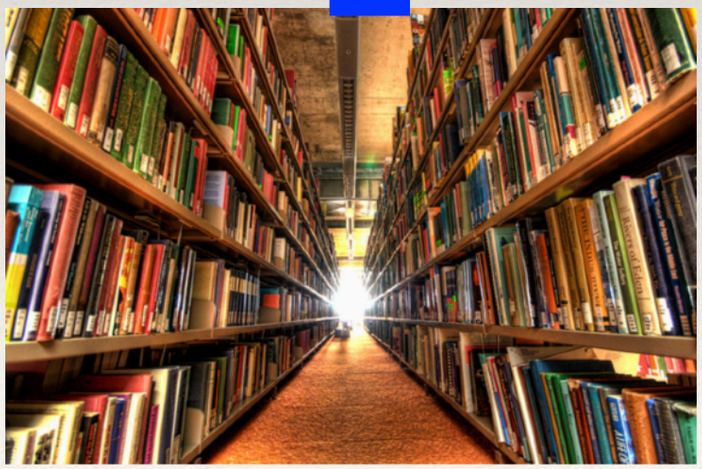
$D$

*Result*

- a selected subset $S \subseteq D$
- with a score (typically: a non-negative real number) assigned to every element of $S$

# IR System: 1st simplification

$Q$

$D$

*Result*

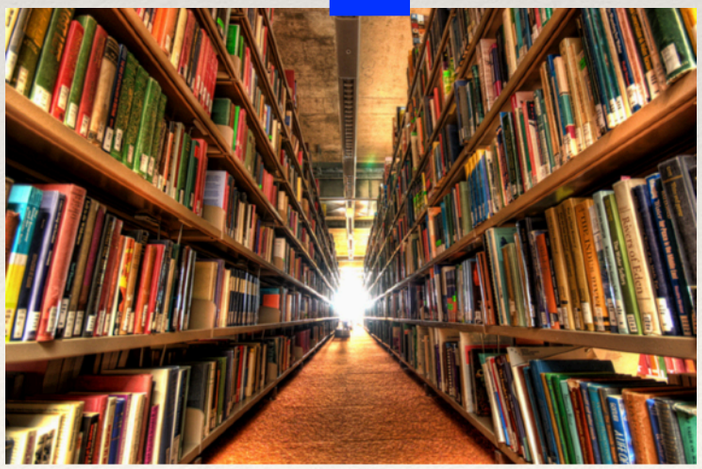
- a selected subset $S \subseteq D$
- with a score (typically: a non-negative real number) assigned to every element of $S$

# IR System: 1st simplification

*Q*


*Result*

- a score assigned to every element of *D*



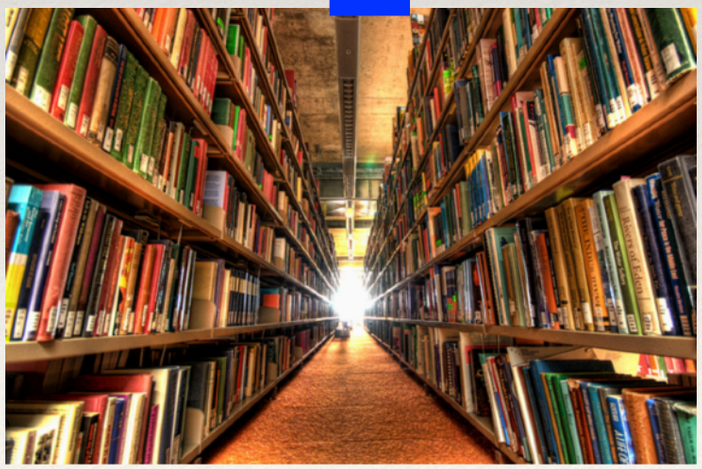
*D*

# IR System: 1st simplification

$Q$


*Result*

- a score assigned to every element of $D$



$D$

The system can be formally represented as a function:

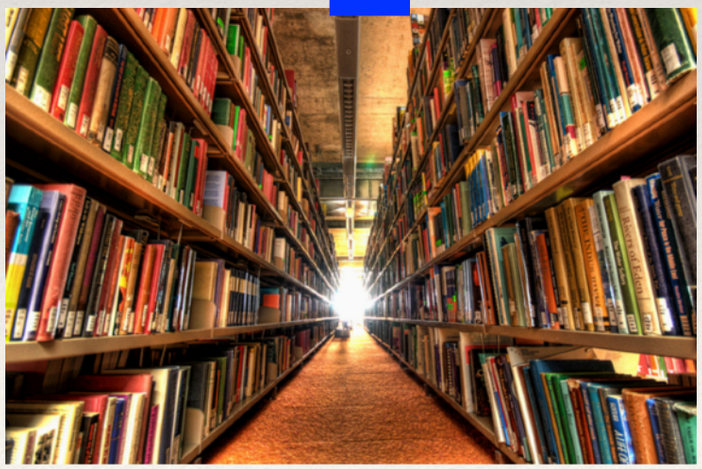
$c: Q \times D \rightarrow \mathbb{R}$

# IR System: 2nd simplification

$Q$

*Result*

- a score assigned to every element of $D$



$D$

The system can be formally represented as a function:

$c: Q \times D \rightarrow \mathbb{R}$

# IR System: 2nd simplification

$Q$

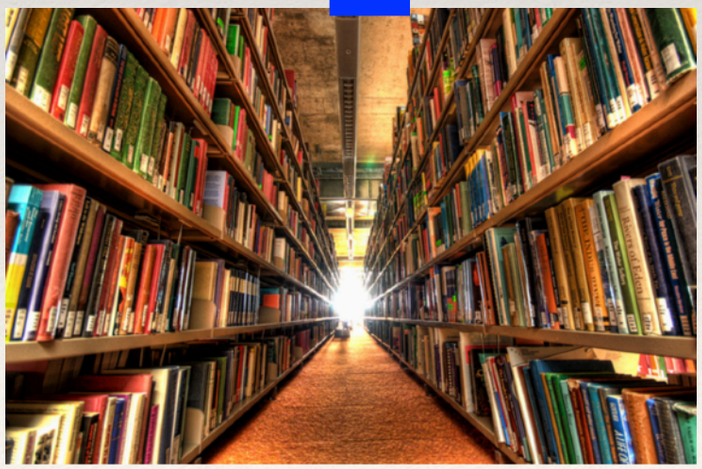
Scores do not depend on the query

*Result*

- a score assigned to every element of $D$

The system can be formally represented as a function:

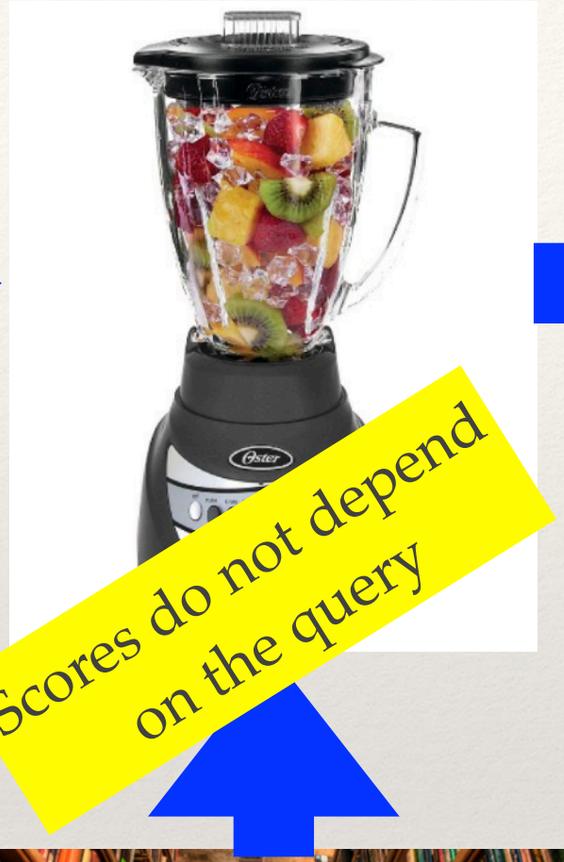
$c: Q \times D \rightarrow \mathbb{R}$



$D$

# IR System: 2nd simplification

$Q$

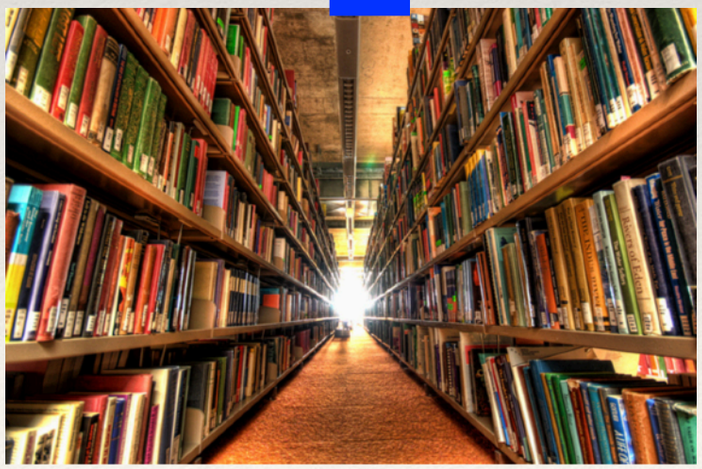



*Result*

- a score assigned to every element of $D$

The system can be formally represented as a function:

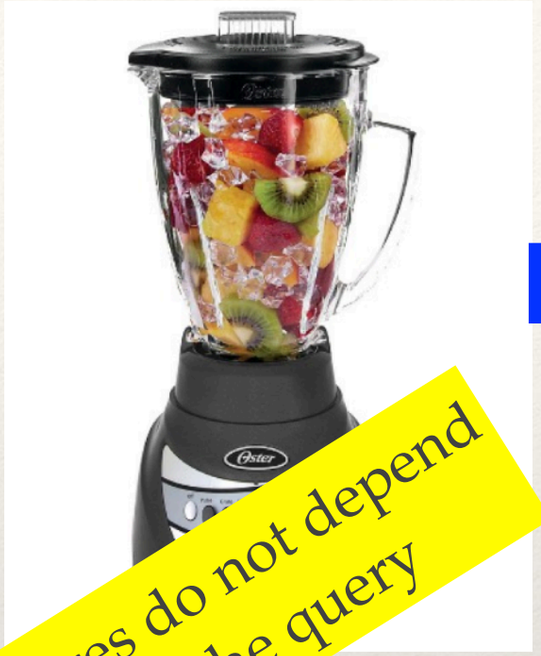
$c: Q \times D \rightarrow \mathbb{R}$



$D$

# IR System: 2nd simplification

$Q$



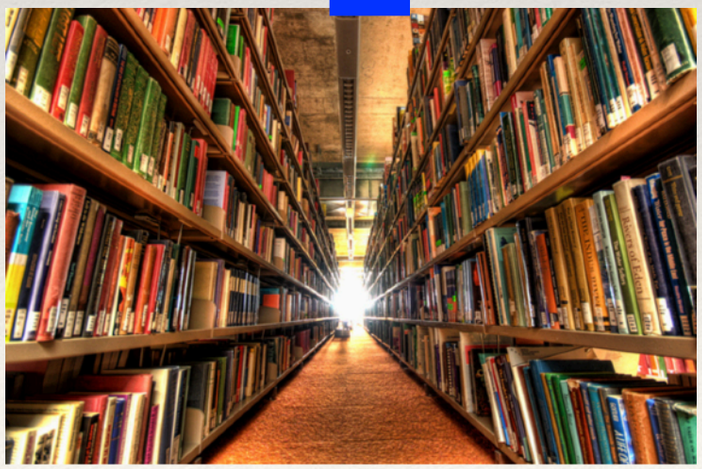
Scores do not depend on the query

*Result*

- a score assigned to every element of $D$

The system can be formally represented as a function:

$c: D \rightarrow \mathbb{R}$



$D$

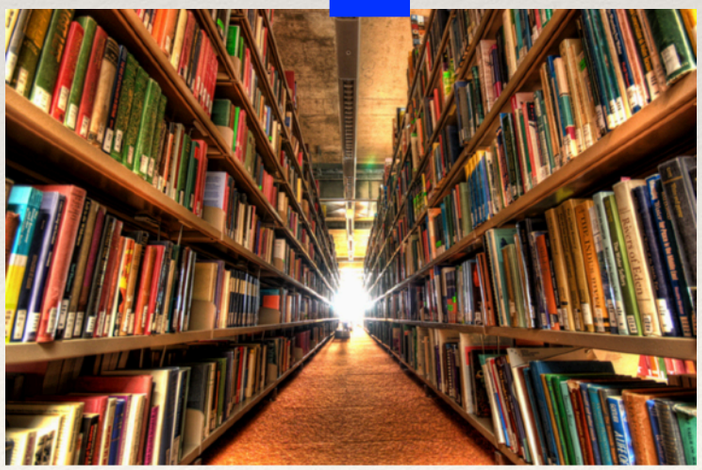
# IR System: 3rd simplification

*Result*

- a score assigned to every element of *D*

The system can be formally represented as a function:

$c: D \rightarrow \mathbb{R}$



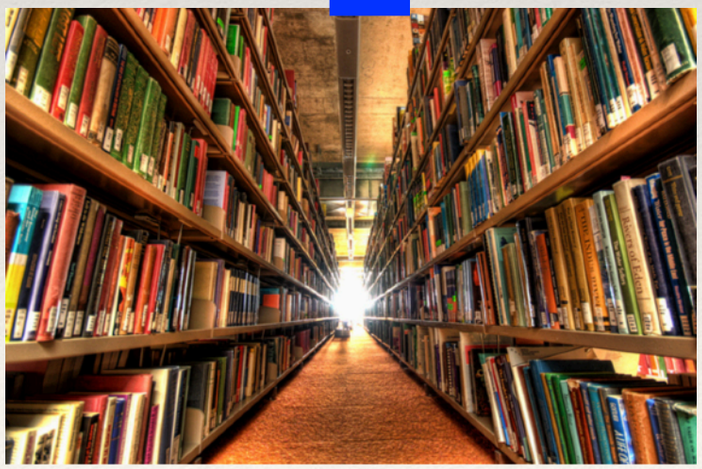
*D*

# IR System: 3rd simplification

Scores depend only on the linkage structure on D

*Result*

- a score assigned to every element of *D*



*D*

The system can be formally represented as a function:

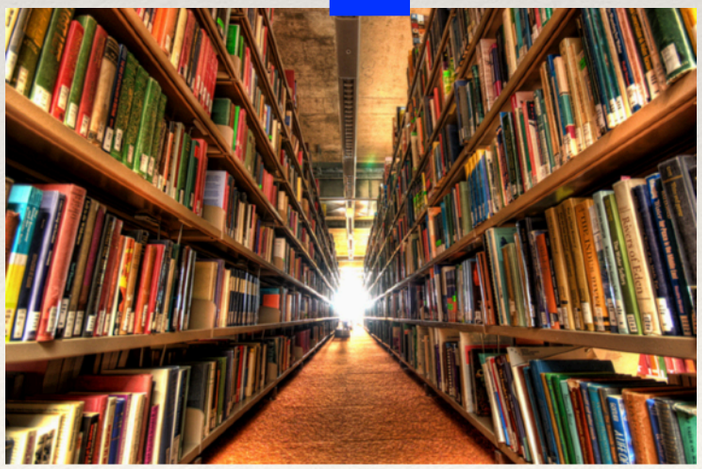
$c: D \rightarrow \mathbb{R}$

# IR System: 3rd simplification

Scores depend only on the linkage structure on D

*Result*

- a score assigned to every element of *D*



*D*

# Centrality

# Centrality

- The system, given a graph $G$ assigns a score to every node of $G$:

$$c_G\colon V_G \to \mathbb{R}$$

# Centrality

- The system, given a graph $G$ assigns a score to every node of $G$:

$$c_G\colon V_G \to \mathbb{R}$$

- The nodes of $G$ are precisely our documents ($V_G = D$)

# Centrality

- The system, given a graph $G$ assigns a score to every node of $G$:

$$c_G\colon V_G \to \mathbb{R}$$

- The nodes of $G$ are precisely our documents ($V_G = D$)
- This is what people refers to as a *centrality index* (or measure, or score, or just "centrality")

# Centrality in social sciences

# Centrality in social sciences

- First works by Bavelas at MIT (1946)

# Centrality in social sciences

- First works by Bavelas at MIT (1946)
- This sparked countless works (Bavelas 1951; Katz 1953; Shaw 1954; Beauchamp 1965; Mackenzie 1966; Burgess 1969; Anthonisse 1971; Czapiel 1974…)

# Centrality in social sciences

- First works by Bavelas at MIT (1946)
- This sparked countless works (Bavelas 1951; Katz 1953; Shaw 1954; Beauchamp 1965; Mackenzie 1966; Burgess 1969; Anthonisse 1971; Czapiel 1974…)
- Brought to CS through IR

# Centrality in social sciences

- First works by Bavelas at MIT (1946)
- This sparked countless works (Bavelas 1951; Katz 1953; Shaw 1954; Beauchamp 1965; Mackenzie 1966; Burgess 1969; Anthonisse 1971; Czapiel 1974…)
- Brought to CS through IR
- Key role in modern IR (=search engines)

# Orienteering in the jungle of centrality indices

# Orienteering in the jungle of centrality indices

- *Path-based indices*, based on the number of *paths* or shortest paths (geodesics) passing through a vertex [betweenness, Katz, …]

# Orienteering in the jungle of centrality indices

- *Path-based indices*, based on the number of *paths* or shortest paths (geodesics) passing through a vertex [betweenness, Katz, …]
- *Spectral indices*, based on some linear-algebra construction [PageRank, Seeley, …]

# Orienteering in the jungle of centrality indices

- *Path-based indices*, based on the number of *paths* or shortest paths (geodesics) passing through a vertex [betweenness, Katz, …]
- *Spectral indices*, based on some linear-algebra construction [PageRank, Seeley, …]
- *Geometric indices*, based on *distances* from a vertex to other vertices [closeness, harmonic, …]

# Orienteering in the jungle of centrality indices

- *Path-based indices*, based on the number of *paths* or shortest paths (geodesics) passing through a vertex [betweenness, Katz, …]
- *Spectral indices*, based on some linear-algebra construction [PageRank, Seeley, …]
- *Geometric indices*, based on *distances* from a vertex to other vertices [closeness, harmonic, …]
- (Actually, the first two families are largely the same, even if that wasn't fully understood for a long time)

# Path-based centralities

# Path-based centralities

- Centrality depends on the paths entering (or passing through) the node

# Path-based centralities

- Centrality depends on the paths entering (or passing through) the node
- Katz's index is a paradigmatic example

# Path-based centralities

- Centrality depends on the paths entering (or passing through) the node
- Katz's index is a paradigmatic example
- Among them: betweenness (Anthonisse 1971), über-popular among social scientists

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

Fraction of shortest paths from $y$ to $z$ passing through $x$

- **Betweenness** centrality:

$$c_{\text{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality:

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality:

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality:

# of paths of length *t* ending in *x*

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

# The path tribe: betweenness and Katz

(Anthonisse 1971; Katz 1953)

- **Betweenness** centrality:

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality:

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

# Spectral centralities (1)

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix
  - First works by Edmund Landau in 1895 on matrices coming from *chess tournaments*

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix
  - First works by Edmund Landau in 1895 on matrices coming from *chess tournaments*
  - Motivation: take a matrix $M$ that in entry $M_{xy}$ has 1 if $x$ won playing with $y$, 0 if he/she lost, 1/2 for a draw

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix
  - First works by Edmund Landau in 1895 on matrices coming from *chess tournaments*
  - Motivation: take a matrix $M$ that in entry $M_{xy}$ has 1 if $x$ won playing with $y$, 0 if he/she lost, 1/2 for a draw
  - $M\mathbf{1}^{\mathrm{T}}$ (the row sums) are a nice score; $M^2\mathbf{1}^{\mathrm{T}}$ even better, but it oscillates: so take score $\boldsymbol{s}$ such that $M\boldsymbol{s}^{\mathrm{T}} = \lambda\boldsymbol{s}^{\mathrm{T}}$

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix
  - First works by Edmund Landau in 1895 on matrices coming from *chess tournaments*
  - Motivation: take a matrix $M$ that in entry $M_{xy}$ has 1 if $x$ won playing with $y$, 0 if he/she lost, 1/2 for a draw
  - $M\mathbf{1}^{\mathrm{T}}$ (the row sums) are a nice score; $M^2\mathbf{1}^{\mathrm{T}}$ even better, but it oscillates: so take score $\boldsymbol{s}$ such that $M\boldsymbol{s}^{\mathrm{T}} = \lambda\boldsymbol{s}^{\mathrm{T}}$
- Berge (1958) extends to general social graphs and develops the theory

# Spectral centralities (1)

- **Eigenvector centrality**: consider the left or right dominant eigenvector of the adjacency matrix
  - First works by Edmund Landau in 1895 on matrices coming from *chess tournaments*
  - Motivation: take a matrix $M$ that in entry $M_{xy}$ has 1 if $x$ won playing with $y$, 0 if he/she lost, 1/2 for a draw
  - $M\mathbf{1}^{\mathrm{T}}$ (the row sums) are a nice score; $M^2\mathbf{1}^{\mathrm{T}}$ even better, but it oscillates: so take score $\boldsymbol{s}$ such that $M\boldsymbol{s}^{\mathrm{T}} = \lambda\boldsymbol{s}^{\mathrm{T}}$
- Berge (1958) extends to general social graphs and develops the theory
- A similar idea was proposed by Seeley to evaluate children popularity

# The spectral tribe: Seeley index

(Seeley 1949)

# The spectral tribe: Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

# The spectral tribe: Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

$$c_{\mathrm{Seeley}}(x) = \sum_{y \to x} \frac{c_{\mathrm{Seeley}}(y)}{d^+(y)}$$

# The spectral tribe: Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

$$c_{\mathrm{Seeley}}(x) = \sum_{y \to x} \frac{c_{\mathrm{Seeley}}(y)}{d^+(y)}$$

- In general it is a left dominant eigenvector of $G_r$

# Spectral centralities (2)

# Spectral centralities (2)

- In 1998, Page, Brin, Motwani and Winograd propose a spectral ranking for the web: **PageRank**

# Spectral centralities (2)

- In 1998, Page, Brin, Motwani and Winograd propose a spectral ranking for the web: **PageRank**
- After some changes in the definition, it stabilizes to a Markov chain $\alpha G_r + (1 - \alpha)\mathbf{1}^T \boldsymbol{v}$

# Spectral centralities (2)

- In 1998, Page, Brin, Motwani and Winograd propose a spectral ranking for the web: **PageRank**
- After some changes in the definition, it stabilizes to a Markov chain $\alpha G_r + (1 - \alpha)\mathbf{1}^T \boldsymbol{v}$
- $G_r$ is Seeley's matrix, $\alpha$ is the *damping factor* and $\boldsymbol{v}$ the *preference vector*

# Spectral centralities (2)

- In 1998, Page, Brin, Motwani and Winograd propose a spectral ranking for the web: **PageRank**
- After some changes in the definition, it stabilizes to a Markov chain $\alpha G_r + (1 - \alpha)\mathbf{1}^T \boldsymbol{v}$
- $G_r$ is Seeley's matrix, $\alpha$ is the *damping factor* and $\boldsymbol{v}$ the *preference vector*
- This is just Katz's index with $\ell_1$-normalization, i.e.,

$$(1 - \alpha)\boldsymbol{v}\sum_{t\geq 0} \alpha^t G_r^t = (1 - \alpha)\boldsymbol{v}(1 - \alpha G_r)^{-1}$$

# The spectral tribe: PageRank

(Brin, Page, Motwani, Winograd 1999)

# The spectral tribe: PageRank
## (Brin, Page, Motwani, Winograd 1999)

- The recursive version of the definition (for uniform preference) is

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(x)}{d^+(x)} + (1-\alpha)v_x$$

# The spectral tribe: PageRank
## (Brin, Page, Motwani, Winograd 1999)

- The recursive version of the definition (for uniform preference) is

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(x)}{d^+(x)} + (1-\alpha)v_x$$

# The spectral tribe: PageRank
## (Brin, Page, Motwani, Winograd 1999)

- The recursive version of the definition (for uniform preference) is

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(x)}{d^+(x)} + (1-\alpha) v_x$$

- … or the dominant eigenvector of the Google matrix

$$\alpha G_r + (1-\alpha)\mathbf{1}^T \mathbf{v}$$

# Geometric centralities and neighbourhood functions

# Geometric centralities and neighbourhood functions

- Define the *distance-count function*

$$D_G(x, t) = \#\{z \mid d_G(z, x) = t\}$$

# Geometric centralities and neighbourhood functions

- Define the *distance-count function*

$$D_G(x, t) = \#\{z \mid d_G(z, x) = t\}$$

- $D_G(x,-)$ is the *distance-count vector* of $x$

# Geometric centralities and neighbourhood functions

- Define the *distance-count function*

$$D_G(x,t) = \#\{z \mid d_G(z,x) = t\}$$

- $D_G(x,\text{-})$ is the *distance-count vector* of $x$
- BTW: in 1-to-1 correspondence with the better known "neighbourhood function"

# Geometric centralities and neighbourhood functions

- Define the *distance-count function*

$$D_G(x,t) = \#\{z \mid d_G(z,x) = t\}$$

- $D_G(x,\text{-})$ is the *distance-count vector* of $x$
- BTW: in 1-to-1 correspondence with the better known "neighbourhood function"
- A geometric centrality is a function of the distance-count vector (i.e., two nodes with the same distance-count vector have the same centrality)

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

Distance from y to x

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- The summation is over all $y$ such that $d(y,x)<\infty$

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- The summation is over all $y$ such that $d(y,x)<\infty$
- **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

# The geometric tribe: closeness and harmonic

(Bavelas 1946; B., Vigna 2013)

- **Closeness** centrality:

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- The summation is over all $y$ such that $d(y,x)<\infty$
- **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- Inspired by (Marchiori, Latora 2000), but may be dated back to (Harris 1954)

# Making sense of centrality

# Making sense of centrality

- Centrality indices can be studied

# Making sense of centrality

- Centrality indices can be studied
  - *individually* (each single centrality index is a world in its own right)

# Making sense of centrality

- Centrality indices can be studied
  - *individually* (each single centrality index is a world in its own right)
  - *comparatively*

# Making sense of centrality

- Centrality indices can be studied
  - *individually* (each single centrality index is a world in its own right)
  - *comparatively*
- Both kinds of studies can be based on

# Making sense of centrality

- Centrality indices can be studied
  - *individually* (each single centrality index is a world in its own right)
  - *comparatively*
- Both kinds of studies can be based on
  - external source of *ground truth*

# Making sense of centrality

- Centrality indices can be studied
  - *individually* (each single centrality index is a world in its own right)
  - *comparatively*
- Both kinds of studies can be based on
  - external source of *ground truth*
  - *axioms* (abstract desirable/undesirable properties)

# Axioms for Centrality

# Axioms for Centrality

- Various attempts, with different flavours: (Sabidussi 1966), (Nieminen 1973), (Kitti 2012), (Brandes *et al.* 2012), (B. & Vigna 2014)

# Axioms for Centrality

- Various attempts, with different flavours: (Sabidussi 1966), (Nieminen 1973), (Kitti 2012), (Brandes *et al*. 2012), (B. & Vigna 2014)
- Sometimes aimed at specific indices (e.g. PageRank): (Chien *et al*. 2004), (Altman and Tennenholtz 2005)

# Orienteering in the jungle of axioms

# Orienteering in the jungle of axioms

- Invariance properties

# Orienteering in the jungle of axioms

- Invariance properties
- Score-dominance properties

# Orienteering in the jungle of axioms

- Invariance properties
- Score-dominance properties
- Rank-dominance properties

# Orienteering in the jungle of axioms

- Invariance properties
- Score-dominance properties
- Rank-dominance properties
- Many other axioms that still need a classification

# Invariance properties

# Invariance properties

# Invariance properties

- Two graphs $G$ and $G'$…

# Invariance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…

# Invariance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

# Invariance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

$$c_G(x) = c_{G'}(x')$$

# Example: Invariance by isomorphism

# Example: Invariance by isomorphism

- If $G$ and $G'$ are isomorphic (via isomorphism $f: G \to G'$)

# Example: Invariance by isomorphism

- If $G$ and $G'$ are isomorphic (via isomorphism $f: G \to G'$)
- and $f(x)=x'$

# Example: Invariance by isomorphism

- If $G$ and $G'$ are isomorphic (via isomorphism $f$: $G \to G'$)
- and $f(x)=x'$

$$c_G(x) = c_{G'}(x')$$

# Example: Invariance by isomorphism

- If $G$ and $G'$ are isomorphic (via isomorphism $f: G \to G'$)
- and $f(x)=x'$

$$c_G(x) = c_{G'}(x')$$

This is **so fundamental** that
it is often given for granted as
part of the notion of *centrality*!

# Example: Invariance by neighbours

# Example: Invariance by neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that

# Example: Invariance by neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$ and $N_G^+(x)=N_G^+(x')$

# Example: Invariance by neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$ and $N_G^+(x)=N_G^+(x')$

$$c_G(x) = c_G(x')$$

# Example: Invariance by neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$ and $N_G^+(x)=N_G^+(x')$

$$c_G(x) = c_G(x')$$

"Two nodes with the same
(in- and out-)neighbours have
the same centrality"

# Invariance by neighbours…

# Invariance by neighbours…

- It is easy to verify that all **geometric centralities** are invariant by neighbours

# Invariance by neighbours…

- It is easy to verify that all **geometric centralities** are invariant by neighbours
- Same for **spectral centralities**

# Example: Invariance by in-neighbours

# Example: Invariance by in-neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that

# Example: Invariance by in-neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$

# Example: Invariance by in-neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$

$$c_G(x) = c_G(x')$$

# Example: Invariance by in-neighbours

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x)=N_G^-(x')$

$$c_G(x) = c_G(x')$$

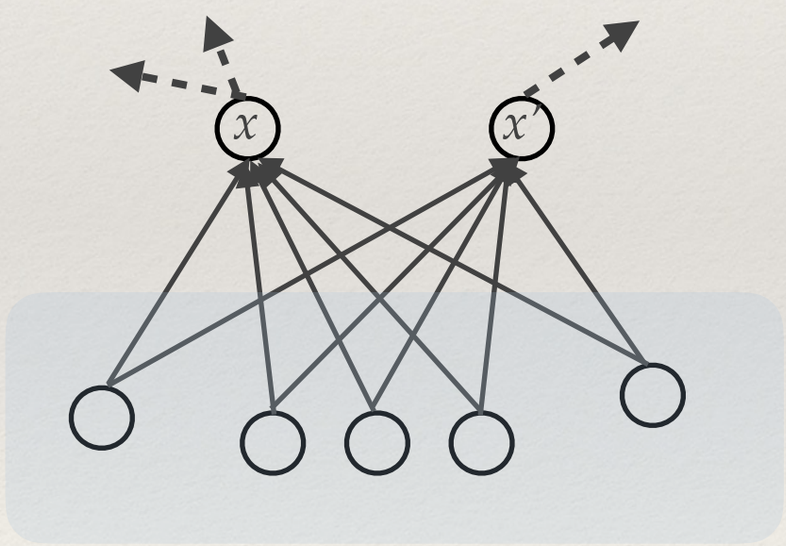
"Two nodes with the same
in-neighbours have
the same centrality"

# Invariance by in-neighbours…

- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours

# Invariance by in-neighbours…

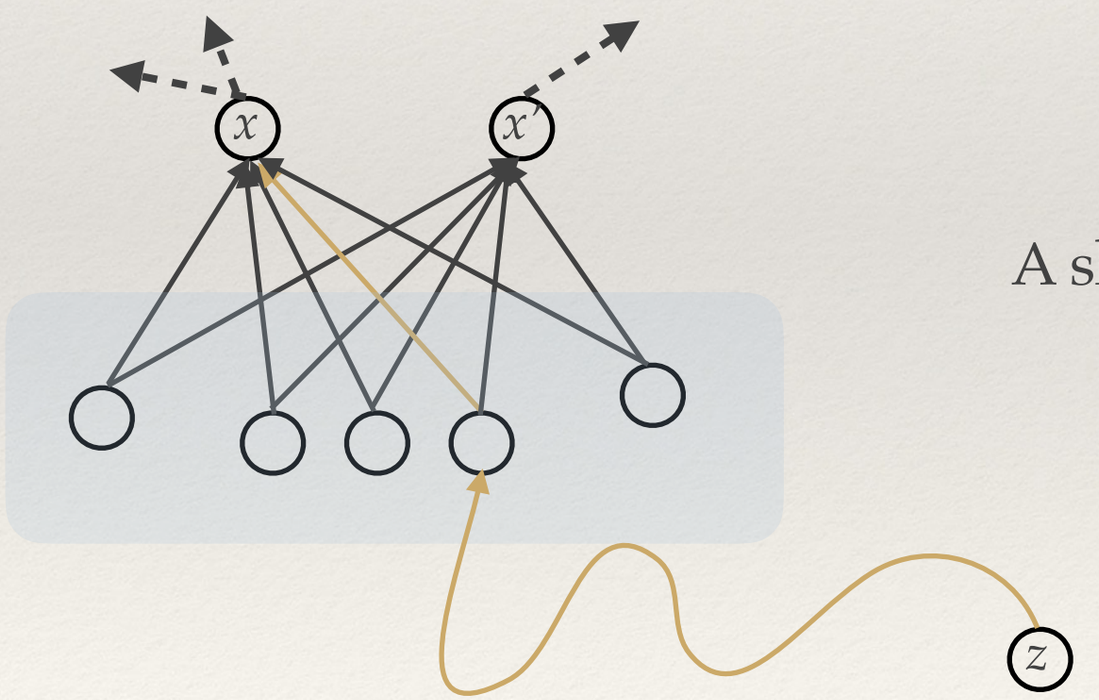
- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours




$$N_G^-(x)=N_G^-(x')$$

# Invariance by in-neighbours…

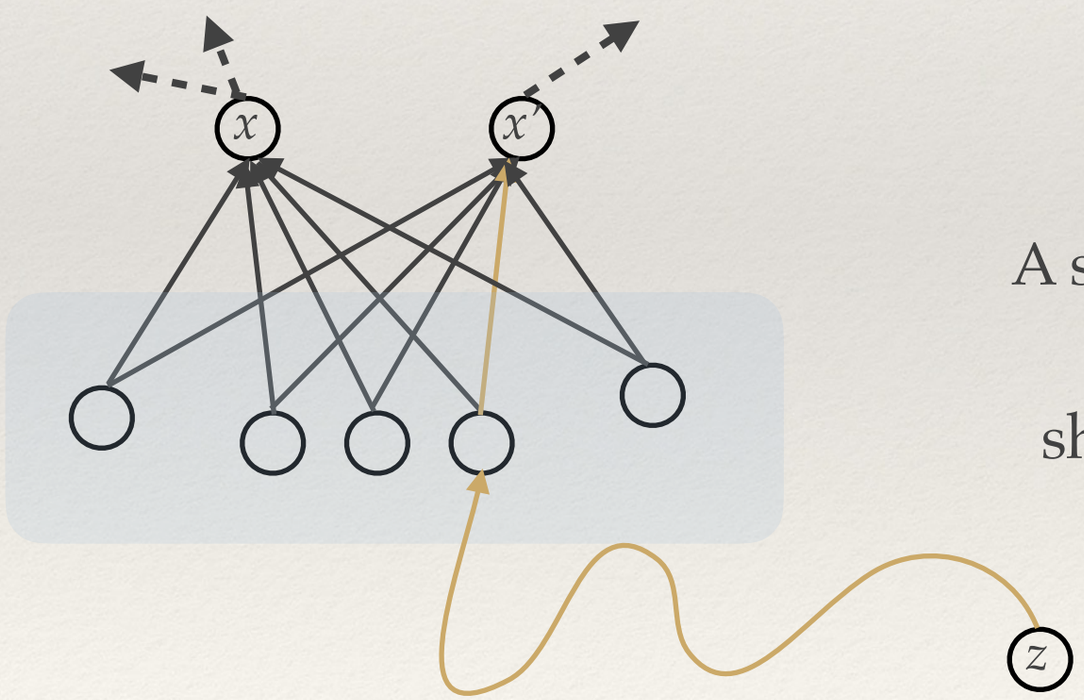
- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours




A shortest path from $z$ to $x$

# Invariance by in-neighbours…

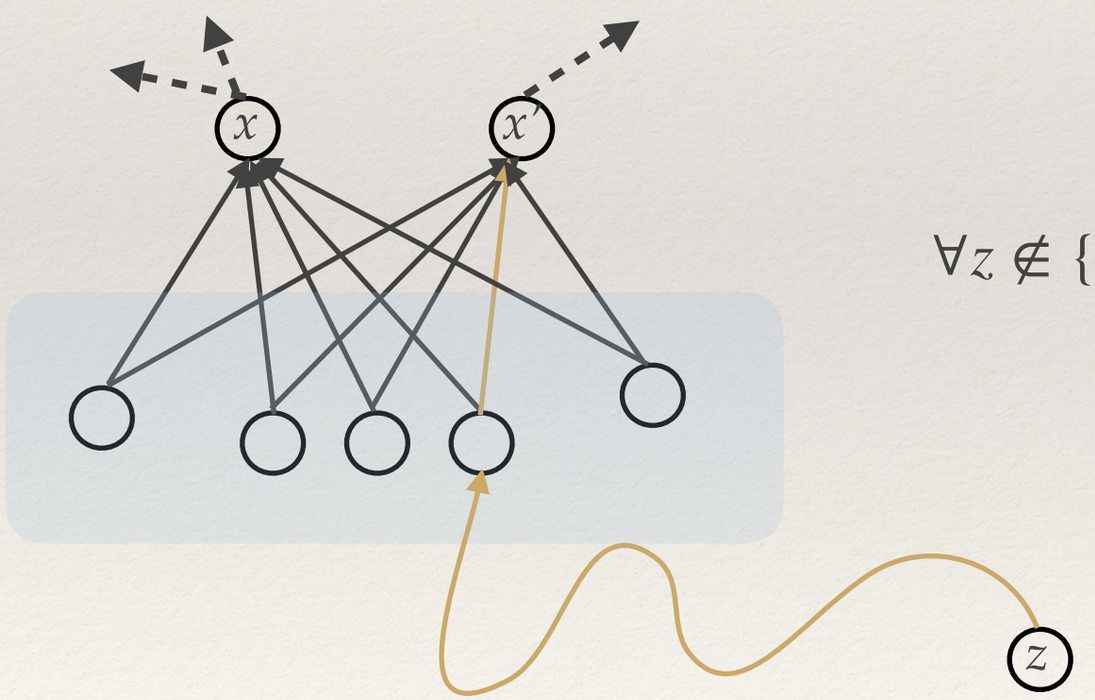
- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours




A shortest path from $z$ to $x$ can be turned into a shortest path from $z$ to $x'$

# Invariance by in-neighbours…

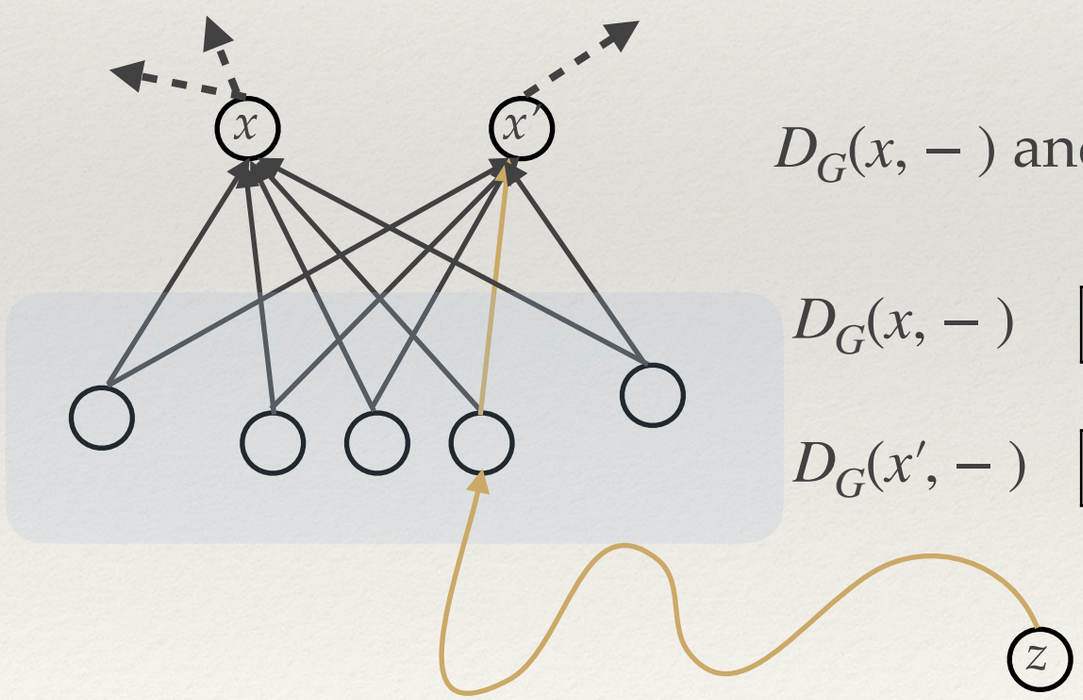
- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours




$$\forall z \notin \{x, x'\} \qquad d_G(z, x) = d_G(z, x')$$

# Invariance by in-neighbours…

- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours



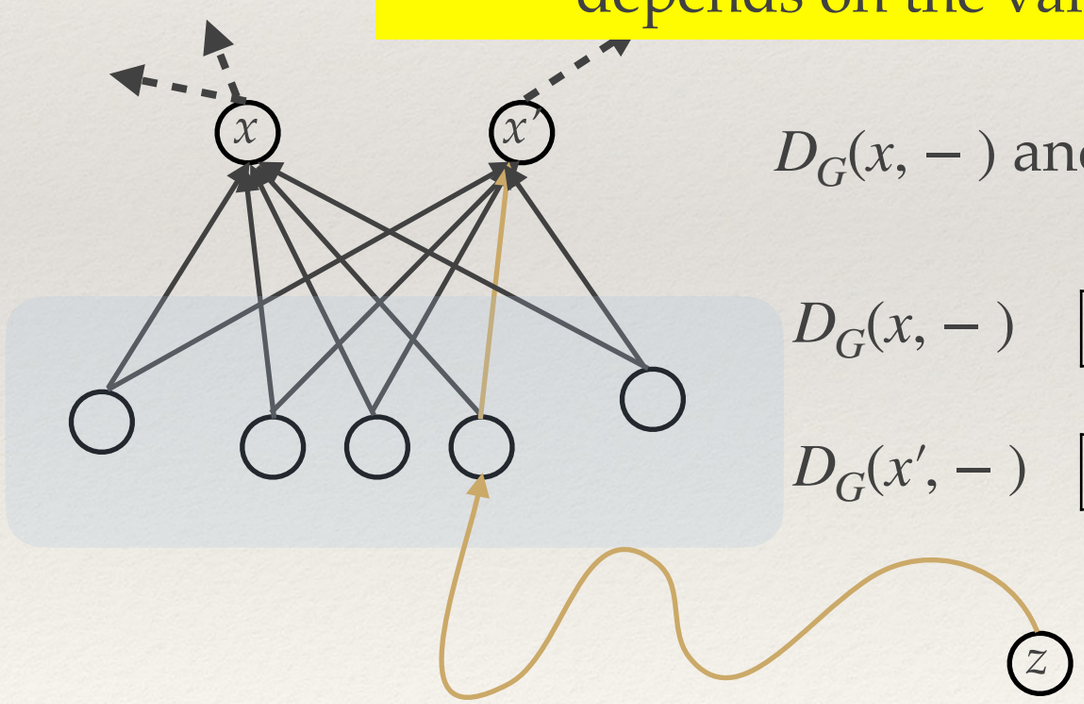

$D_G(x, -)$ and $D_G(x', -)$ are almost the same…

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | … |
|---|---|---|---|---|---|---|---|---|---|---|
| $D_G(x, -)$ | 5 | 0 | 3 | 7 | 2 | 0 | 0 | 4 | 1 | … |
| $D_G(x', -)$ | 5 | 0 | 3 | 6 | 2 | 0 | 0 | 5 | 1 | … |

# Invariance by in-neighbours…

- A superficial observer *may* believe that geometric centralities satisfy invariance by in-neighbours

The difference (+1 in one position, -1 in another position) depends on the values of $d(x,x')$ and $d(x',x)$




$D_G(x, -)$ and $D_G(x', -)$ are almost the same…

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | … |
|---|---|---|---|---|---|---|---|---|---|---|
| $D_G(x, -)$ | 5 | 0 | 3 | 7 | 2 | 0 | 0 | 4 | 1 | … |
| $D_G(x', -)$ | 5 | 0 | 3 | 6 | 2 | 0 | 0 | 5 | 1 | … |

# Invariance by in-neighbours…

# Invariance by in-neighbours…

- So, in general, **geometric centralities** *are not invariant by in-neighbours*

# Invariance by in-neighbours…

- So, in general, **geometric centralities** *are not invariant by in-neighbours*
  - They are on symmetric (i.e. undirected) graphs, though

# Invariance by in-neighbours…

- So, in general, **geometric centralities** *are not invariant by in-neighbours*
  - They are on symmetric (i.e. undirected) graphs, though
- But **spectral centralities** (e.g. PageRank) *are invariant by in-neighbours*

# Score-dominance properties

# Score-dominance properties

# Score-dominance properties

- Two graphs $G$ and $G'$…

# Score-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…

# Score-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

# Score-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

$$c_G(x) \geq c_{G'}(x')$$

# Score-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

$$c_G(x) \geq c_{G'}(x')$$

- Sometimes $>$ is required (*strict dominance*)

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

- Let $G$ be a graph and $x$, $x'$ be two nodes such that

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x) \subseteq N_G^-(x')$

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x) \subseteq N_G^-(x')$

$$c_G(x) \leq c_G(x')$$

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x) \subseteq N_G^-(x')$

$$c_G(x) \leq c_G(x')$$

- **Observe:** if a measure satisfies this property, it is also invariant by in-neighbours

# Example: Dominance by in-neighbours (Schoch and Brandes, 2016)

- Let $G$ be a graph and $x$, $x'$ be two nodes such that
- $N_G^-(x) \subseteq N_G^-(x')$

$$c_G(x) \leq c_G(x')$$

- **Observe:** if a measure satisfies this property, it is also invariant by in-neighbours
- $\Rightarrow$ geometric centralities do not satisfy "dominance by in-neighbours"

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$
- Then

$$c_{G'}(y) > c_G(y)$$

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$
- Then

$$c_{G'}(y) > c_G(y)$$

"Adding one arc towards $y$ (strictly) increases its score"

# Example: Score-dominance by arc addition (a.k.a. *score monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$
- Then

$$c_{G'}(y) > c_G(y)$$

"Adding one arc towards $y$ (strictly) increases its score"

- The weak version (with $\geq$) also makes sense

# Score monotonicity
# ("Axioms for Centrality", B. & Vigna 2014)

| | General | Strongly connected |
|---|---|---|
| Seeley | no | yes |
| PageRank | yes | yes |
| betweenness | no | no |
| Katz | yes | yes |
| closeness | no | yes |
| harmonic | yes | yes |

# Score monotonicity
# ("Axioms for Centrality", B. & Vigna 2014)

| | General | Strongly connected |
|---|---|---|
| Seeley | no | yes |
| PageRank | yes | yes |
| betweenness | no | no |
| Katz | yes | yes |
| closeness | no | yes |
| harmonic | yes | yes |

# PageRank satisfies score monotonicity

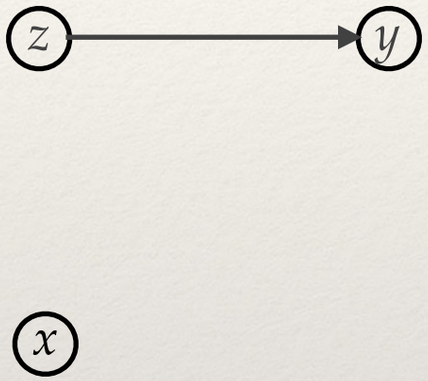
# PageRank satisfies score monotonicity

- Proved by (Chien, Dwork, Kumar, Simon and Sivakumar 2003) for the case when all nodes have nonzero PageRank

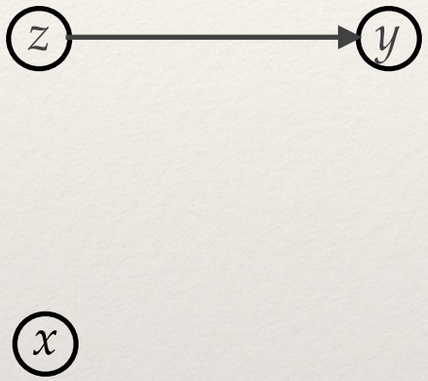
# PageRank satisfies score monotonicity

- Proved by (Chien, Dwork, Kumar, Simon and Sivakumar 2003) for the case when all nodes have nonzero PageRank
- Generalized in (B. and Vigna, 2014) to the case $r_x > 0$

# Closeness does not satisfy score monotonicity

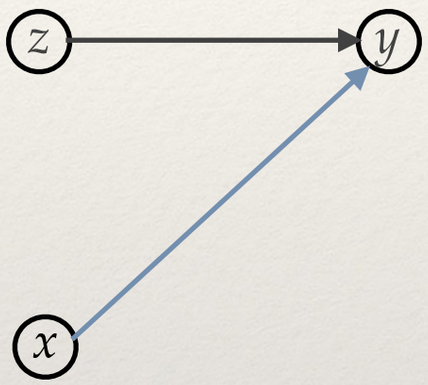


# Closeness does not satisfy score monotonicity



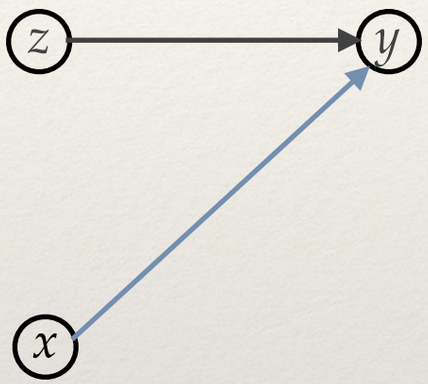

$$c_G(y) = \frac{1}{\sum_t d_G(t, y)} = \frac{1}{1} = 1$$

# Closeness does not satisfy score monotonicity




$$c_G(y) = \frac{1}{\sum_t d_G(t, y)} = \frac{1}{1} = 1$$

# Closeness does not satisfy score monotonicity




$$c_G(y) = \frac{1}{\sum_t d_G(t, y)} = \frac{1}{1} = 1$$

$$c_{G'}(y) = \frac{1}{\sum_t d_{G'}(t, y)} = \frac{1}{1+1} = \frac{1}{2}$$

# Closeness satisfies score monotonicity in the strongly connected case

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\mathrm{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- On strongly connected graphs

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- On strongly connected graphs
  - the summation includes all nodes

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- On strongly connected graphs
  - the summation includes all nodes
  - the distances do not increase after adding the new arc

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- On strongly connected graphs
  - the summation includes all nodes
  - the distances do not increase after adding the new arc
  - at least one distance strictly decreases

# Closeness satisfies score monotonicity in the strongly connected case

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- On strongly connected graphs
  - the summation includes all nodes
  - the distances do not increase after adding the new arc
  - at least one distance strictly decreases
- So closeness centrality *is* score monotone on strongly connected graphs!

# Rank-dominance properties

# Rank-dominance properties

# Rank-dominance properties

- Two graphs $G$ and $G'$…

# Rank-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…

# Rank-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints

# Rank-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints
- The rank of $x'$ in $G'$ is "not less" than the rank of $x$ in $G$

# Rank-dominance properties

- Two graphs $G$ and $G'$…
- …and two nodes $x \in G$ and $x' \in G'$…
- …satisfying some constraints
- The rank of $x'$ in $G'$ is "not less" than the rank of $x$ in $G$

Typically stated on two graphs with the same set of nodes, and for a single node

# Rank-dominance properties revised (weak version)

# Rank-dominance properties revised (weak version)

- Two graphs $G$ and $G'$ *with the same node set* $V$ and node $x \in V$

# Rank-dominance properties revised (weak version)

- Two graphs $G$ and $G'$ *with the same node set* $V$ and node $x \in V$

1. $\forall z \,.\, c_G(x) > c_G(z) \implies c_{G'}(x) > c_{G'}(z)$

# Rank-dominance properties revised (weak version)

- Two graphs $G$ and $G'$ *with the same node set* $V$ and node $x \in V$

1. $\forall z \,.\, c_G(x) > c_G(z) \implies c_{G'}(x) > c_{G'}(z)$
2. $\forall y \,.\, c_G(x) = c_G(y) \implies c_{G'}(x) \geq c_{G'}(y)$

# Rank-dominance properties revised (strict version)

# Rank-dominance properties revised (strict version)

- Two graphs $G$ and $G'$ *with the same node set* $V$ and node $x \in V$

$$\forall z \,.\, c_G(x) \geq c_G(z) \implies c_{G'}(x) > c_{G'}(z)$$

# Example: Rank-dominance by arc addition (a.k.a. *rank monotonicity*)

# Example: Rank-dominance by arc addition (a.k.a. *rank monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$

# Example: Rank-dominance by arc addition (a.k.a. *rank monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$

# Example: Rank-dominance by arc addition (a.k.a. *rank monotonicity*)

- If $G$ is a graph not containing the arc $x \rightarrow y$
- And $G' = G \cup \{x \rightarrow y\}$
- Then, for all $z$

$$\forall z \,.\, c_G(y) > c_G(z) \implies c_{G'}(y) > c_{G'}(z)$$

$$\forall z \,.\, c_G(y) = c_G(z) \implies c_{G'}(y) \geq c_{G'}(z)$$

# Example: Rank-dominance by arc addition (a.k.a. *rank monotonicity*)

- If $G$ is a graph not containing the arc $x \to y$
- And $G' = G \cup \{x \to y\}$
- Then, for all $z$

$$\forall z \,.\, c_G(y) > c_G(z) \implies c_{G'}(y) > c_{G'}(z)$$
$$\forall z \,.\, c_G(y) = c_G(z) \implies c_{G'}(y) \geq c_{G'}(z)$$

- For the strict version, the last $\geq$ should become a $>$

# Rank monotonicity
# ("Rank monotonicity in centrality measures.", B. & Luongo & Vigna 2017)

| | General | Strongly connected |
|---|---|---|
| Seeley | no | yes |
| PageRank† | yes* | yes* |
| betweenness | no | no |
| Katz† | yes* | yes* |
| closeness | no | yes |
| harmonic | yes* | yes* |

† provided that no node has null preference

# Rank monotonicity
# ("Rank monotonicity in centrality measures.", B. & Luongo & Vigna 2017)

| | General | | Strongly connected | |
|---|---|---|---|---|
| Seeley | no | *(no)* | yes | *(yes)* |
| PageRank† | yes* | *(yes)* | yes* | *(yes)* |
| betweenness | no | *(no)* | no | *(no)* |
| Katz† | yes* | *(yes)* | yes* | *(yes)* |
| closeness | no | *(no)* | yes | *(yes)* |
| harmonic | yes* | *(yes)* | yes* | *(yes)* |

† provided that no node has null preference

# Rank monotonicity
# ("Rank monotonicity in centrality measures.", B. & Luongo & Vigna 2017)

| | General | Strongly connected |
|---|---|---|
| Seeley | no | yes |
| PageRank† | yes* | yes* |
| betweenness | no | no |
| Katz† | yes* | yes* |
| closeness | no | yes |
| harmonic | yes* | yes* |

† provided that no node has null preference

# Rank monotonicity
# ("Rank monotonicity in centrality measures.", B. & Luongo & Vigna 2017)

| | General | Strongly connected |
|---|---|---|
| Seeley | no | yes |
| PageRank† | yes* | yes* |
| betweenness | no | no |
| Katz† | yes* | yes* |
| closeness | no | yes |
| harmonic | yes* | yes* |

† provided that no node has null preference

# PageRank and rank monotonicity

# Loose (non-strict) rank monotonicity

# Loose (non-strict) rank monotonicity

- For PageRank, with $G' = G \cup \{x \rightarrow y\}$

$$\forall y \,.\, c_G(y) > c_G(z) \implies c_{G'}(y) > c_{G'}(z)$$
$$\forall y \,.\, c_G(y) = c_G(z) \implies c_{G'}(y) \geq c_{G'}(z)$$

holds (Chien, Dwork, Kumar, Simon & Sivakumar 2004) for *everywhere nonzero score*

# Loose (non-strict) rank monotonicity

- For PageRank, with $G' = G \cup \{x \to y\}$

$$\forall y \,.\, c_G(y) > c_G(z) \implies c_{G'}(y) > c_{G'}(z)$$
$$\forall y \,.\, c_G(y) = c_G(z) \implies c_{G'}(y) \geq c_{G'}(z)$$

  holds (Chien, Dwork, Kumar, Simon & Sivakumar 2004) for *everywhere nonzero score*

- The strict version was proved in (B., Luongo, Vigna 2017)

$$\forall y \,.\, c_G(y) \geq c_G(z) \implies c_{G'}(y) > c_{G'}(z)$$

  for *everywhere nonzero preference*

# Loose vs. strict

# Loose vs. strict

- The proof in (Chien, Dwork, Kumar, Simon & Sivakumar 2004) exploits the fact that the Google matrix is a regular Markov chain

# Loose vs. strict

- The proof in (Chien, Dwork, Kumar, Simon & Sivakumar 2004) exploits the fact that the Google matrix is a regular Markov chain
- (B., Luongo, Vigna 2017) is based on some properties of M-matrices…

# Loose vs. strict

- The proof in (Chien, Dwork, Kumar, Simon & Sivakumar 2004) exploits the fact that the Google matrix is a regular Markov chain
- (B., Luongo, Vigna 2017) is based on some properties of M-matrices…
- …the results have wider applicability (e.g., Katz)

# Damped spectral ranking

# Damped spectral ranking

- Let $M$ be a nonnegative matrix, $0 < \alpha < 1/\rho(M)$, $\mathbf{v}$ a *strictly positive* vector

# Damped spectral ranking

- Let $M$ be a nonnegative matrix, $0 < \alpha < 1/\rho(M)$, $\mathbf{v}$ a *strictly positive* vector
- Then, the centrality vector $\mathbf{r}$ defined by

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

satisfies strict rank monotonicity, suitably generalised to matrices (see below)

# Damped spectral ranking

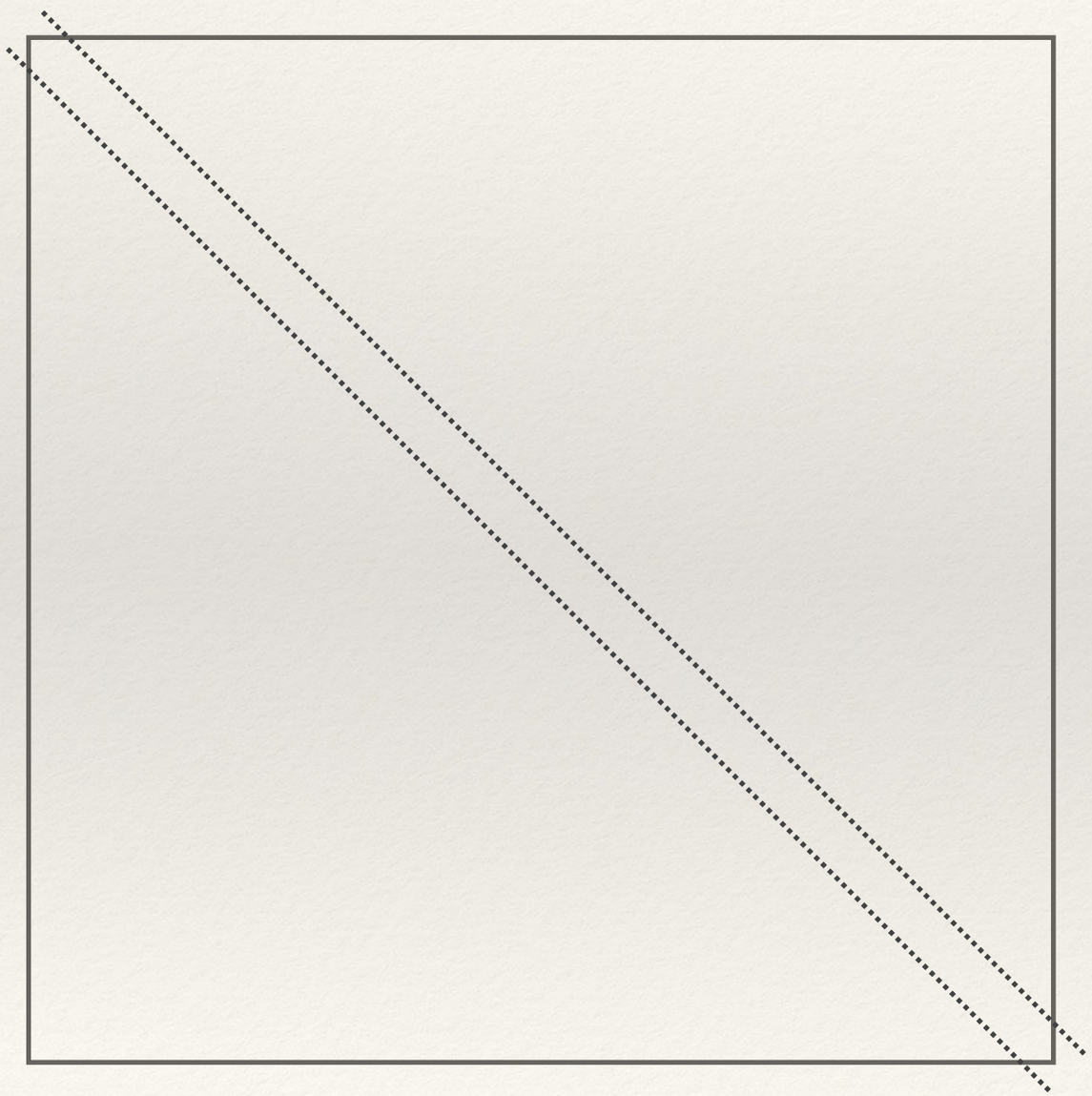
- Let $M$ be a nonnegative matrix, $0 < \alpha < 1/\rho(M)$, $\mathbf{v}$ a *strictly positive* vector
- Then, the centrality vector $\mathbf{r}$ defined by

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

  satisfies strict rank monotonicity, suitably generalised to matrices (see below)
- Applies to PageRank, Katz, …

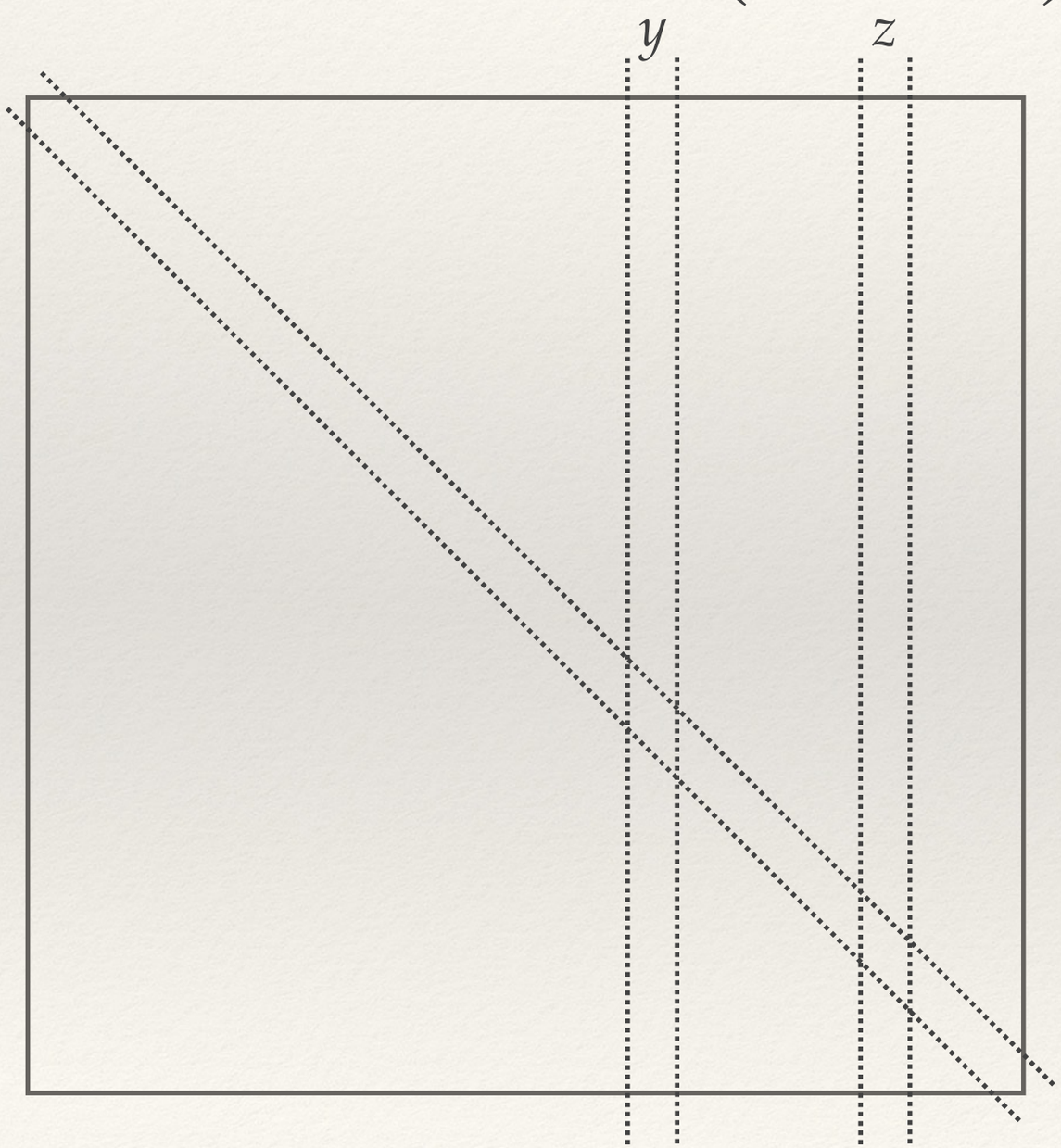
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \quad \mathbf{r} = \mathbf{v}C$$



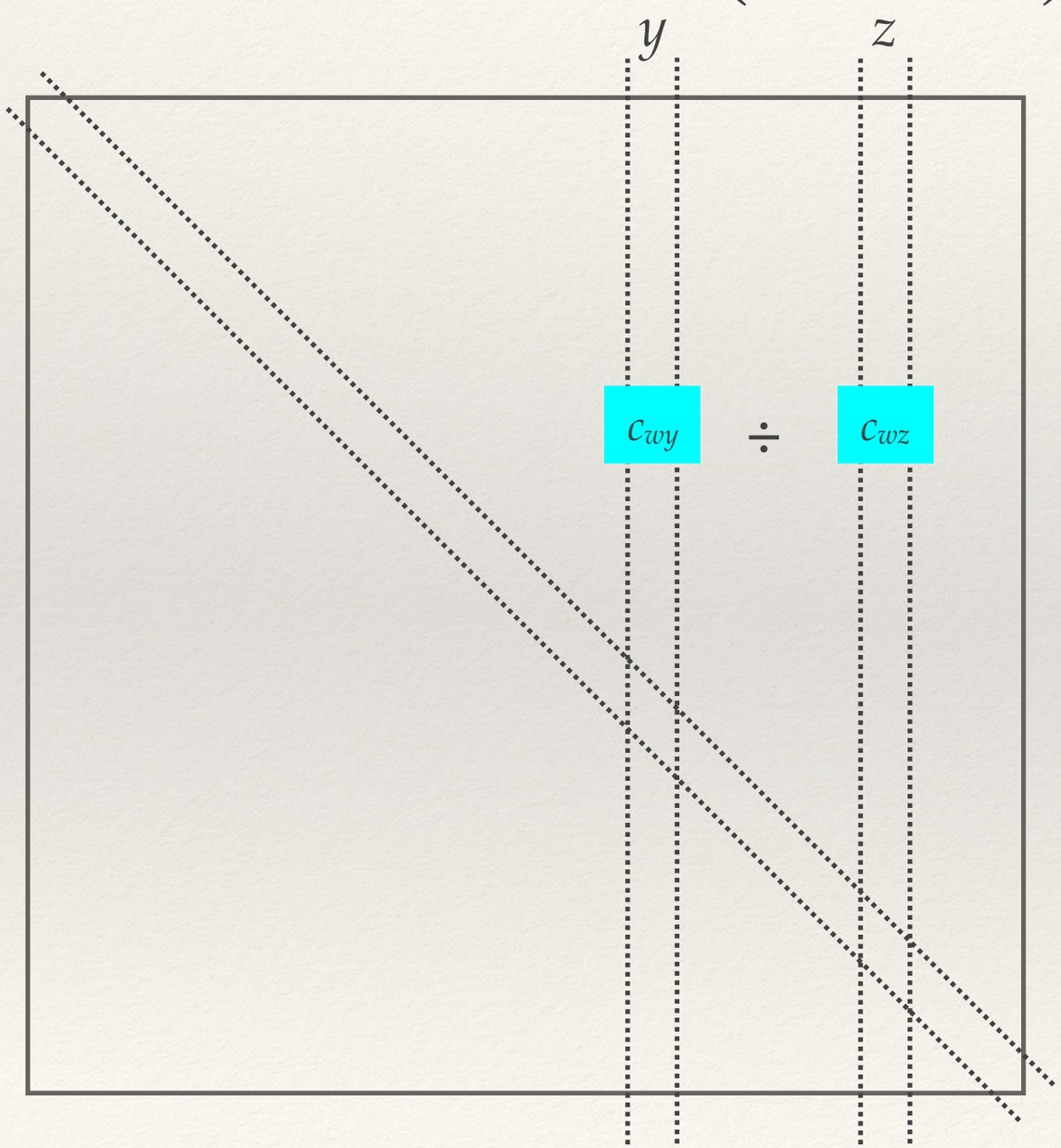
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$



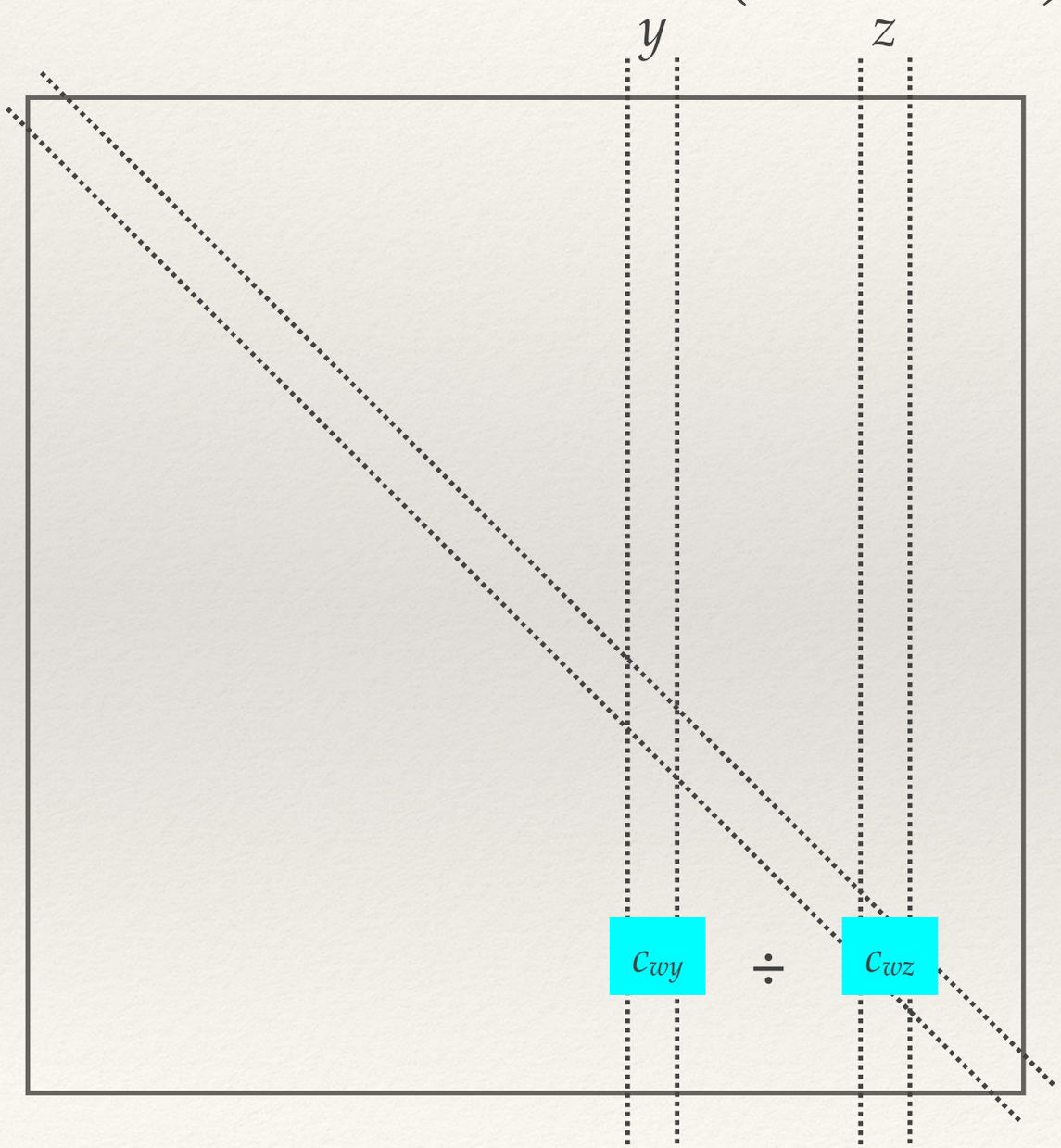
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \quad \mathbf{r} = \mathbf{v}C$$



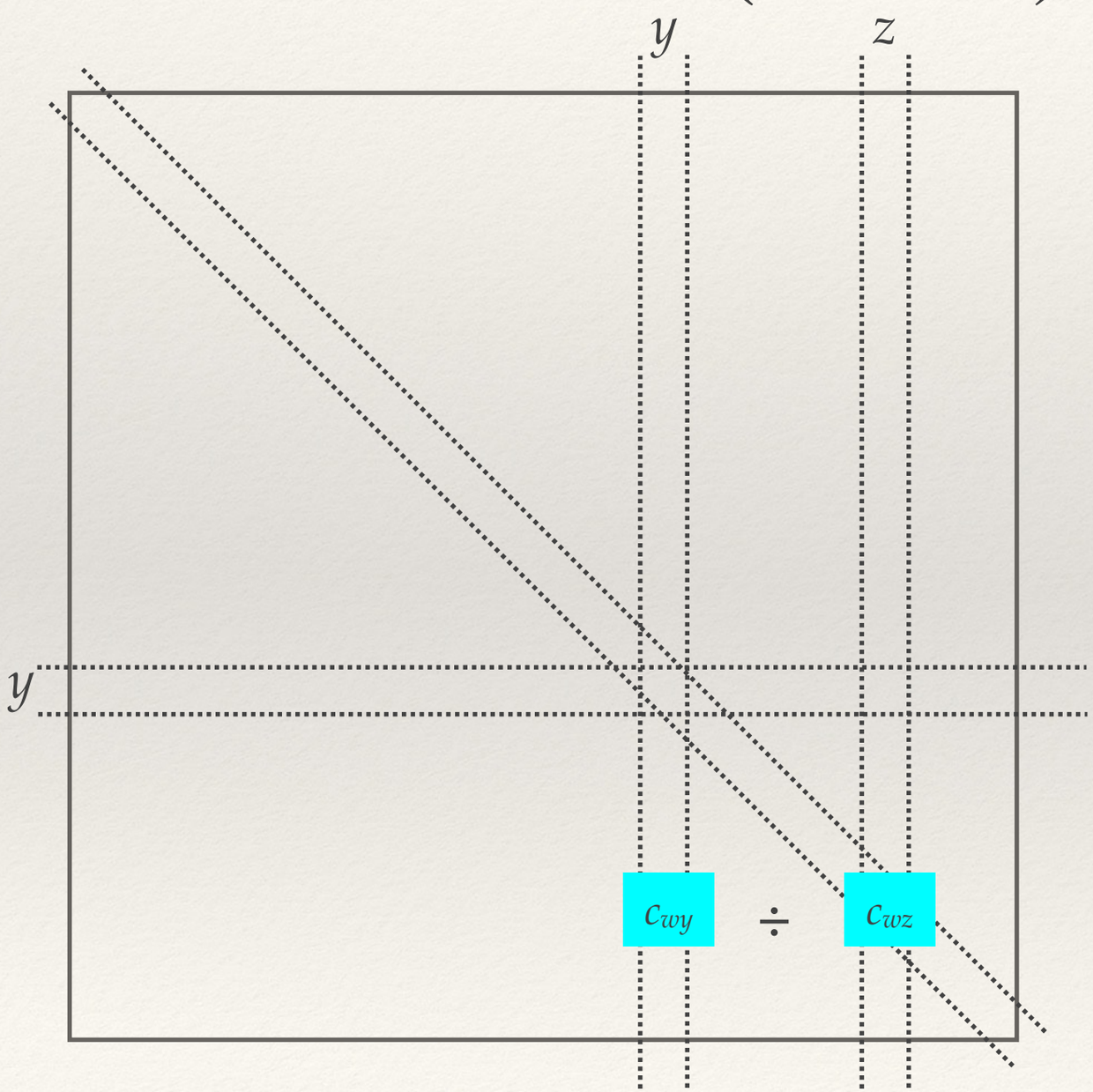
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \quad \mathbf{r} = \mathbf{v}C$$



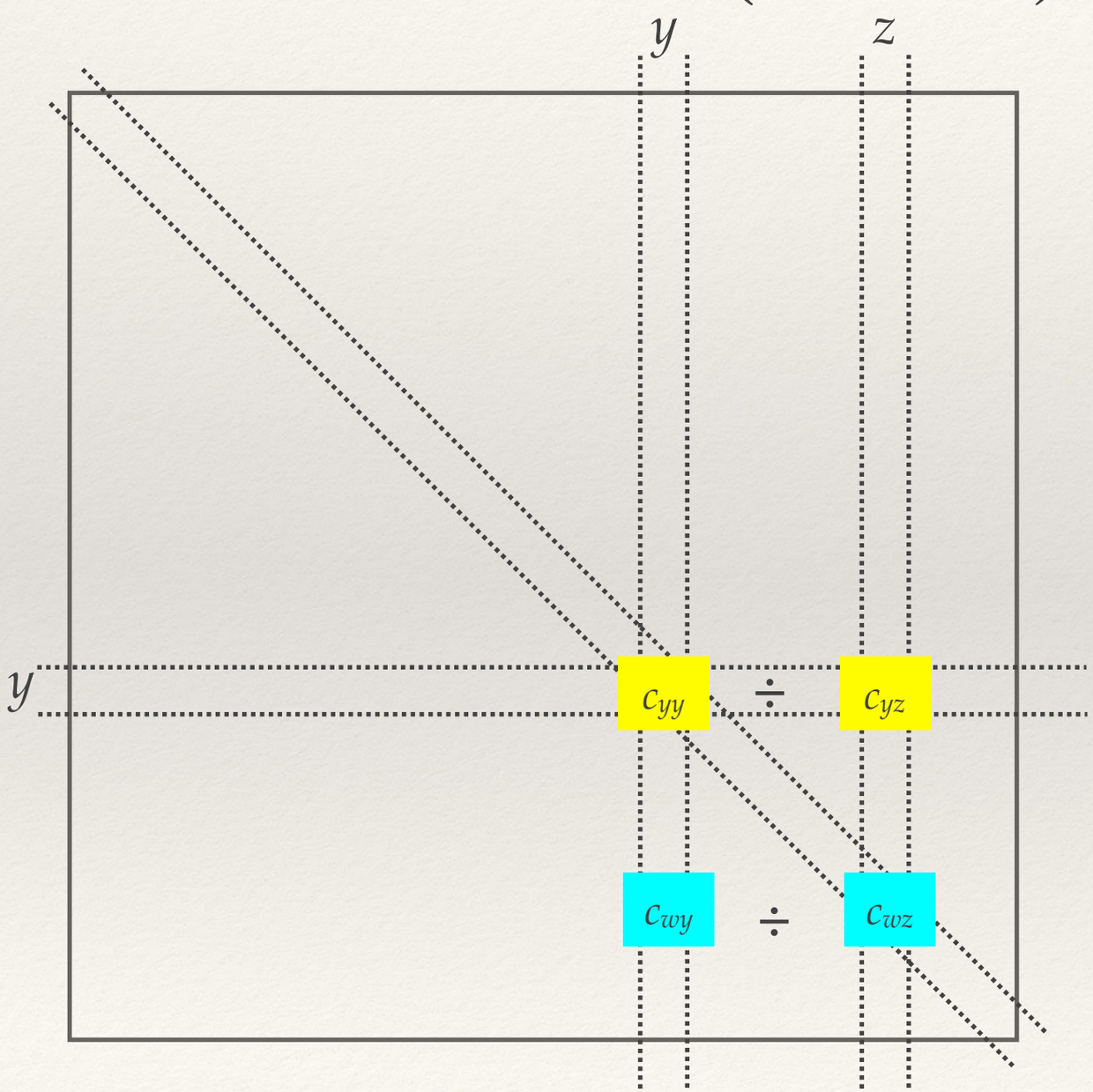
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \quad \mathbf{r} = \mathbf{v}C$$



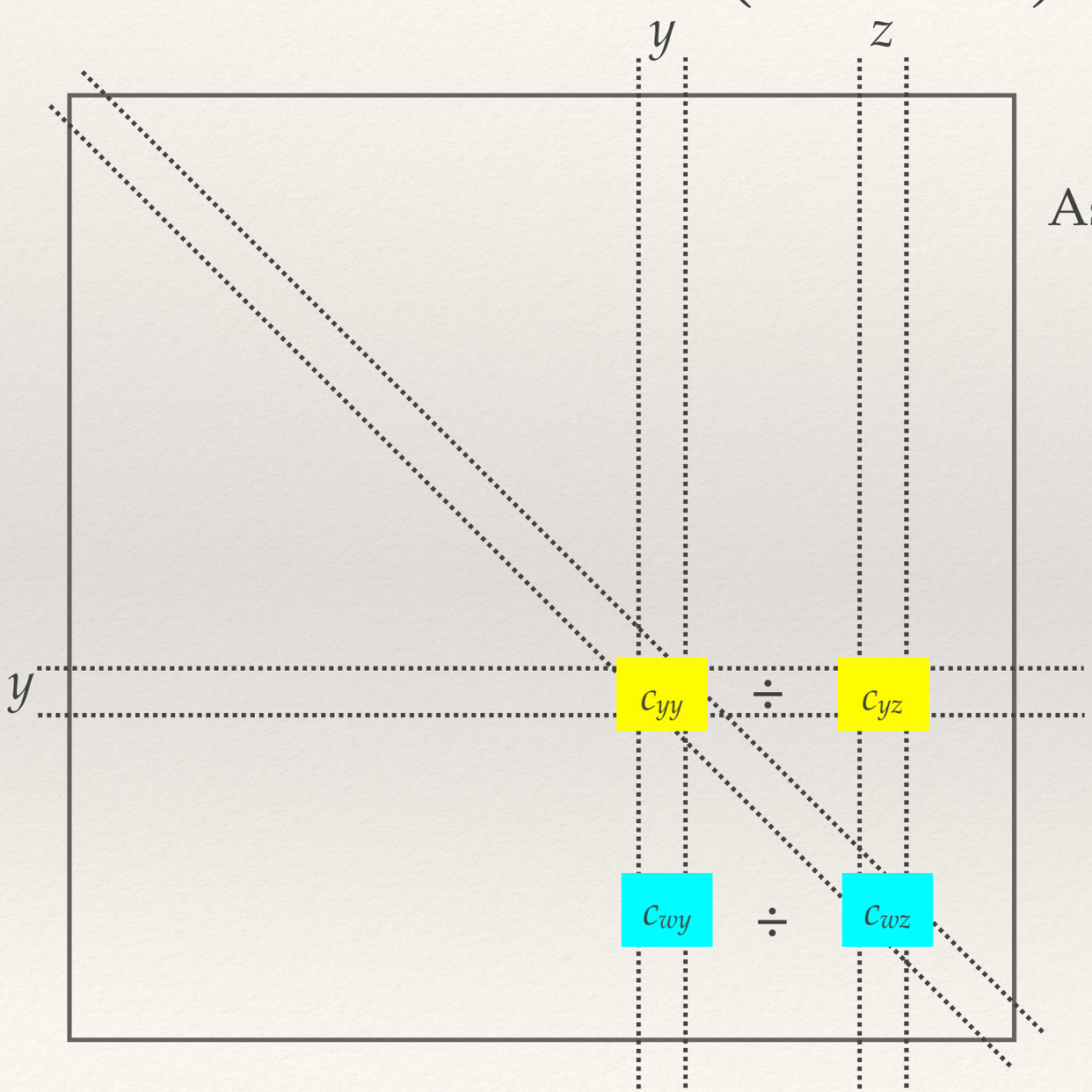
# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$



# Lemma (ext. Willoughby, 1977)

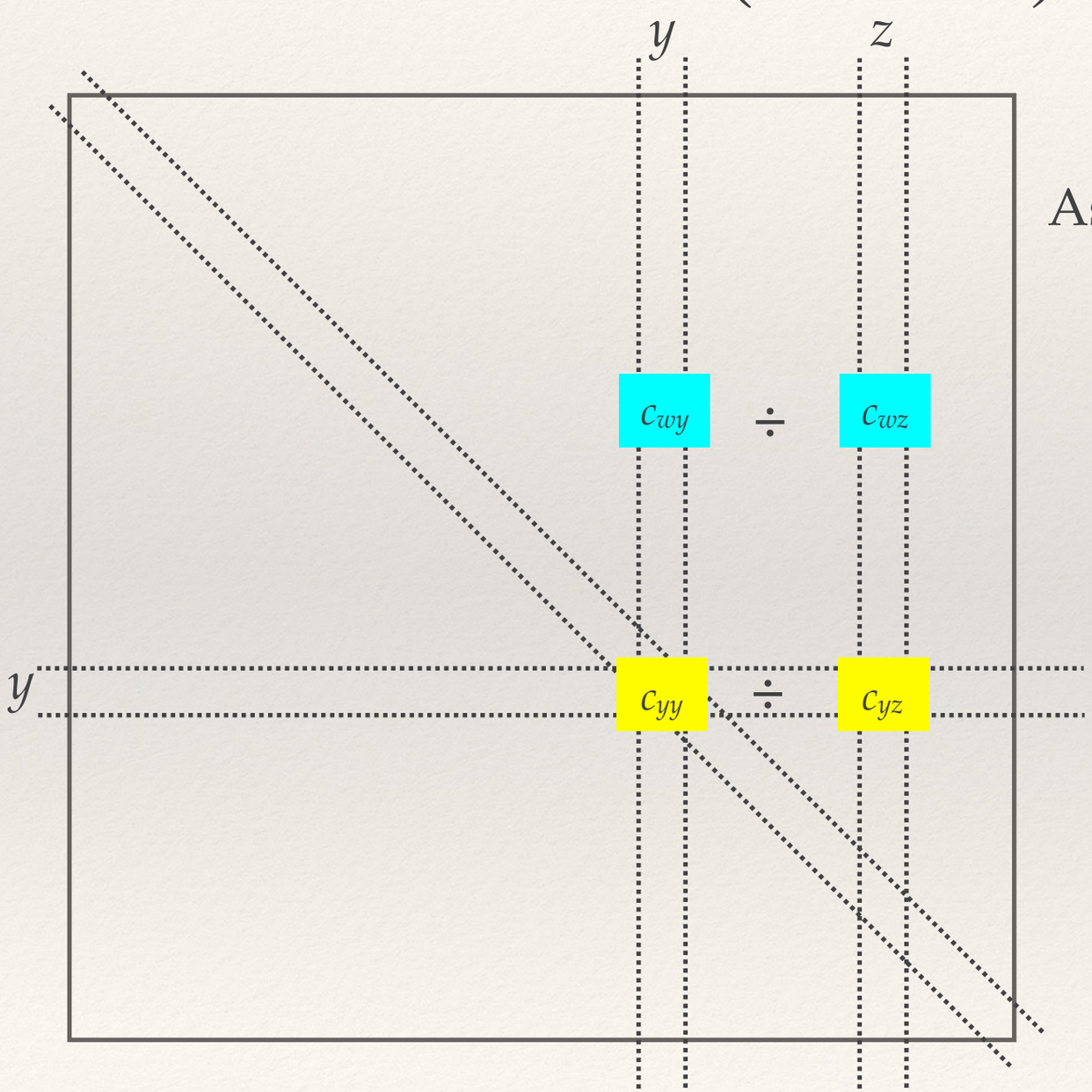
$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

# Lemma (ext. Willoughby, 1977)

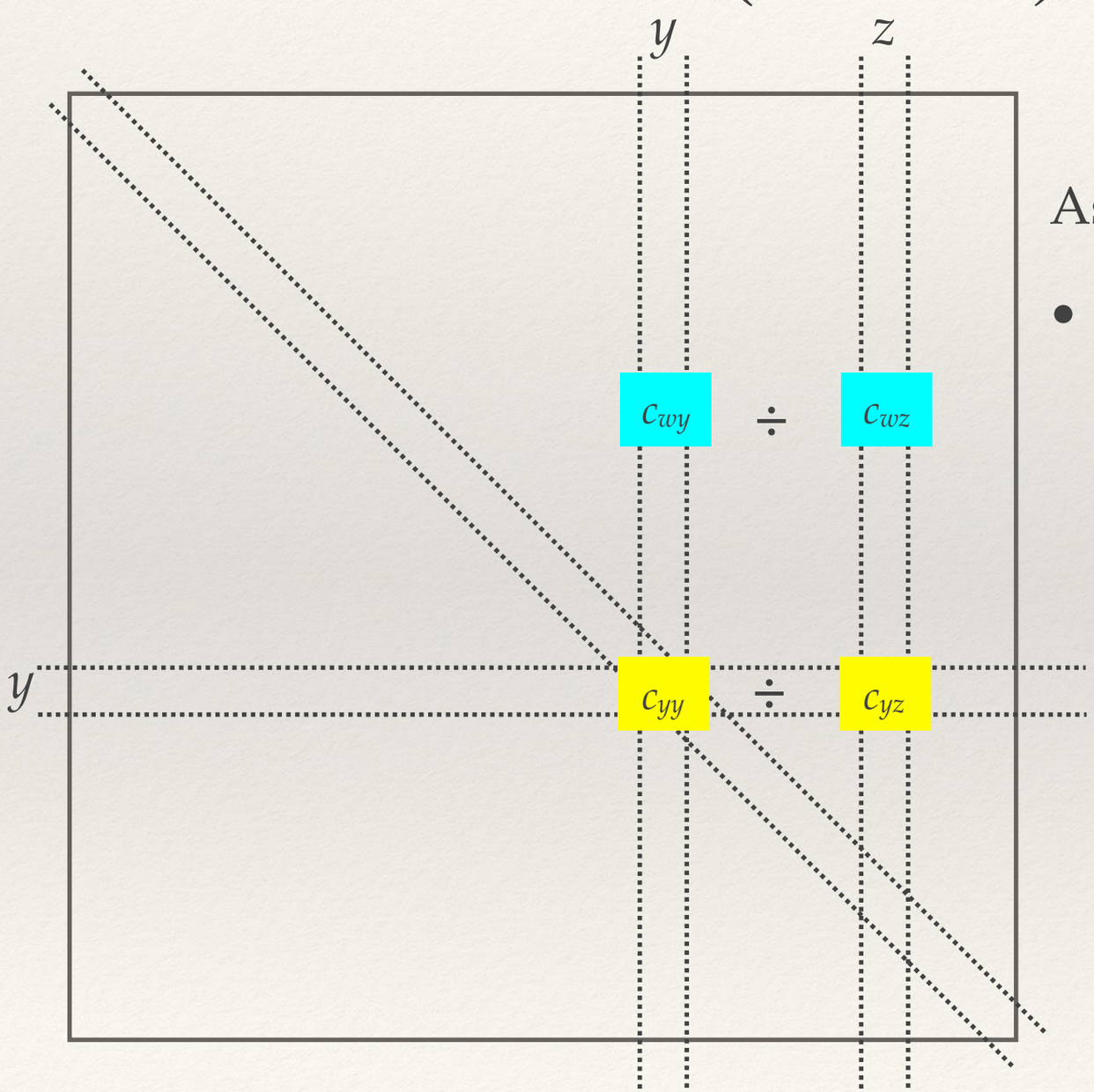
$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$

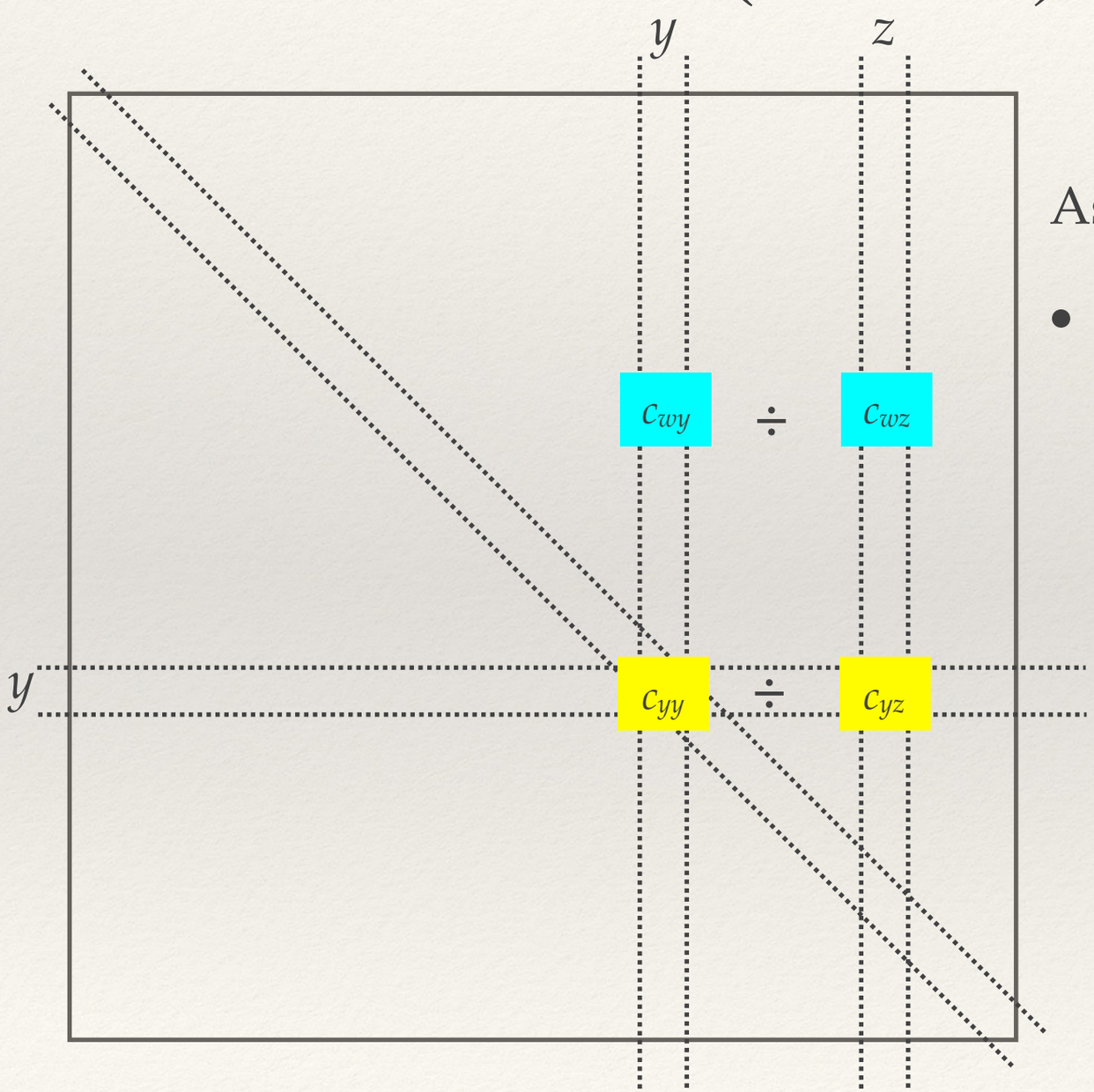



Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

- Then $c_{wy}/c_{wz} \leq q$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$

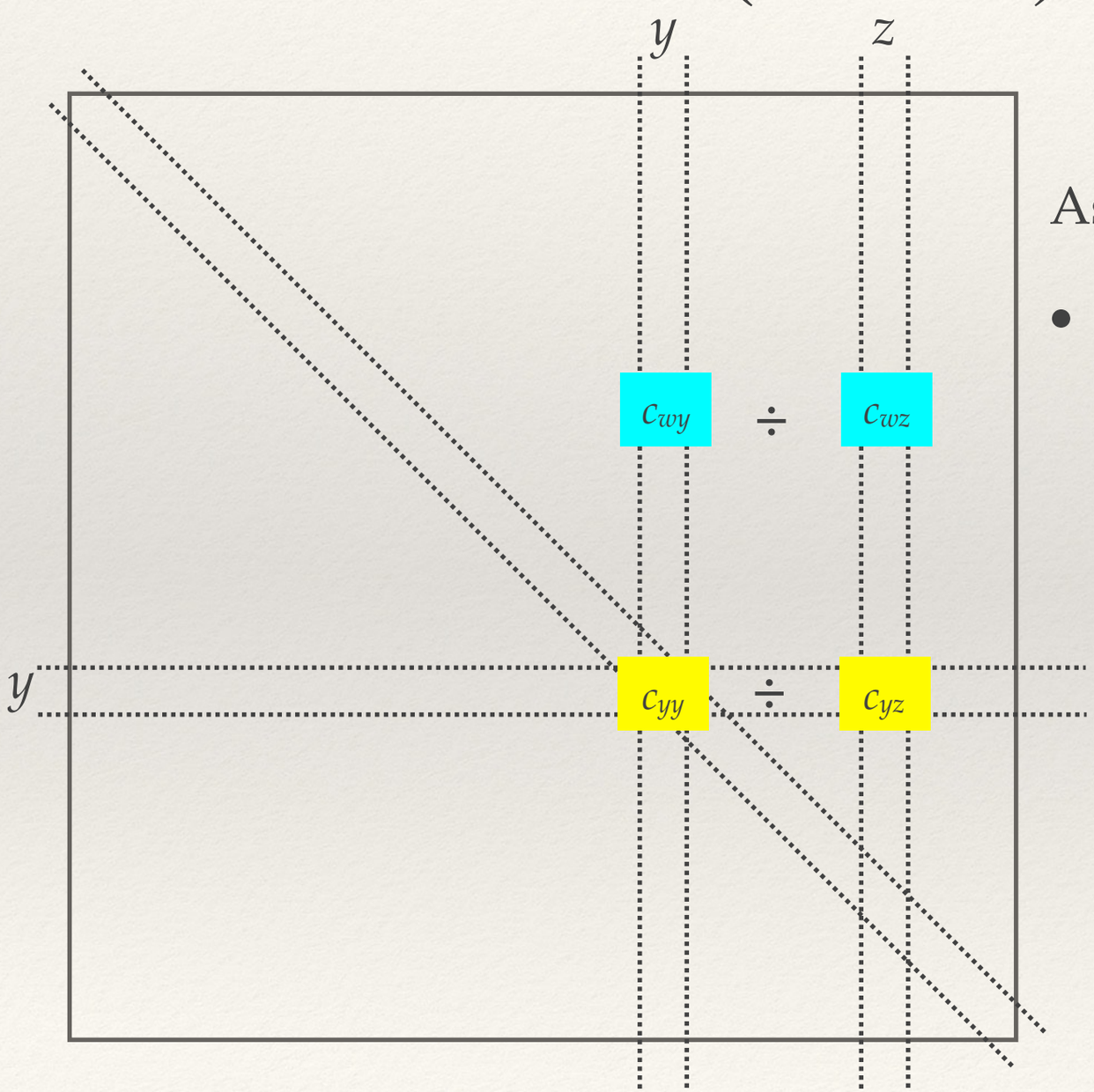



Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

- Then $c_{wy} \leq q\, c_{wz}$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

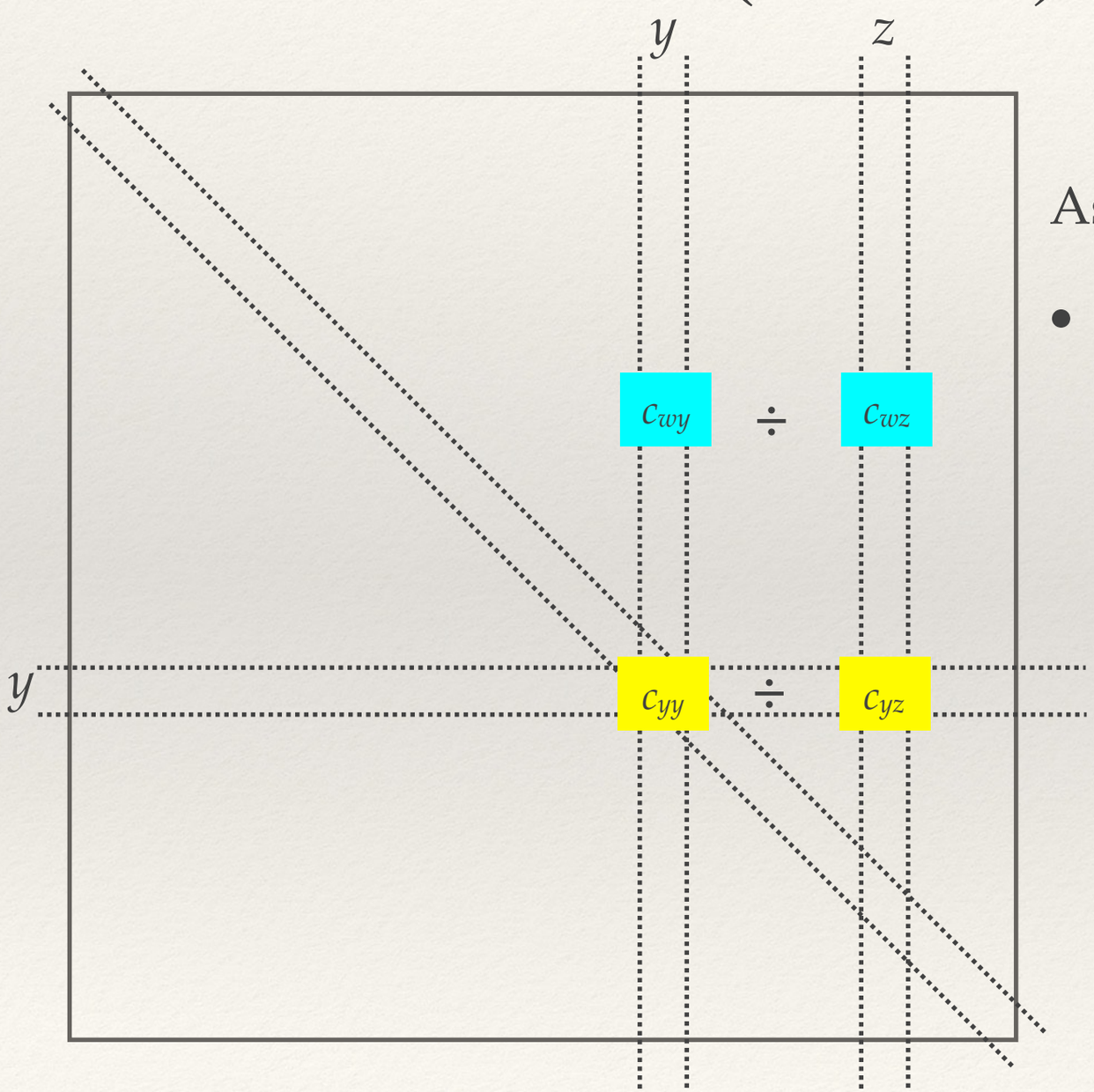
- Then $c_{wy} \leq q\, c_{wz}$

  As a consequence if $q < 1$...

$$r_y = \sum_w v_w c_{wy} \leq$$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

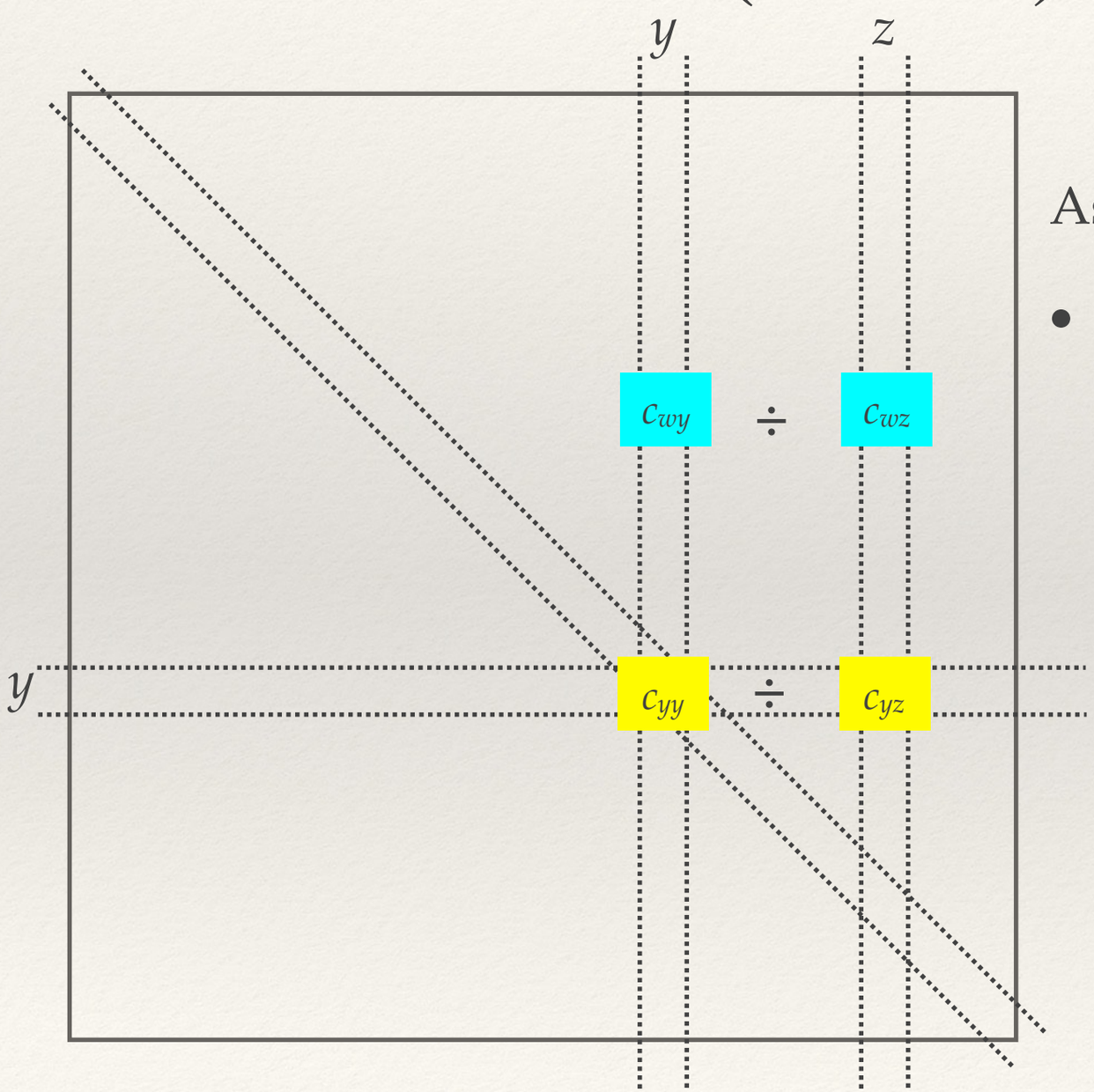
- Then $c_{wy} \leq q\, c_{wz}$

  As a consequence if $q < 1$...

$$r_y = \sum_w v_w c_{wy} \leq \sum_w v_w \frac{c_{wz} c_{yy}}{c_{yz}}$$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

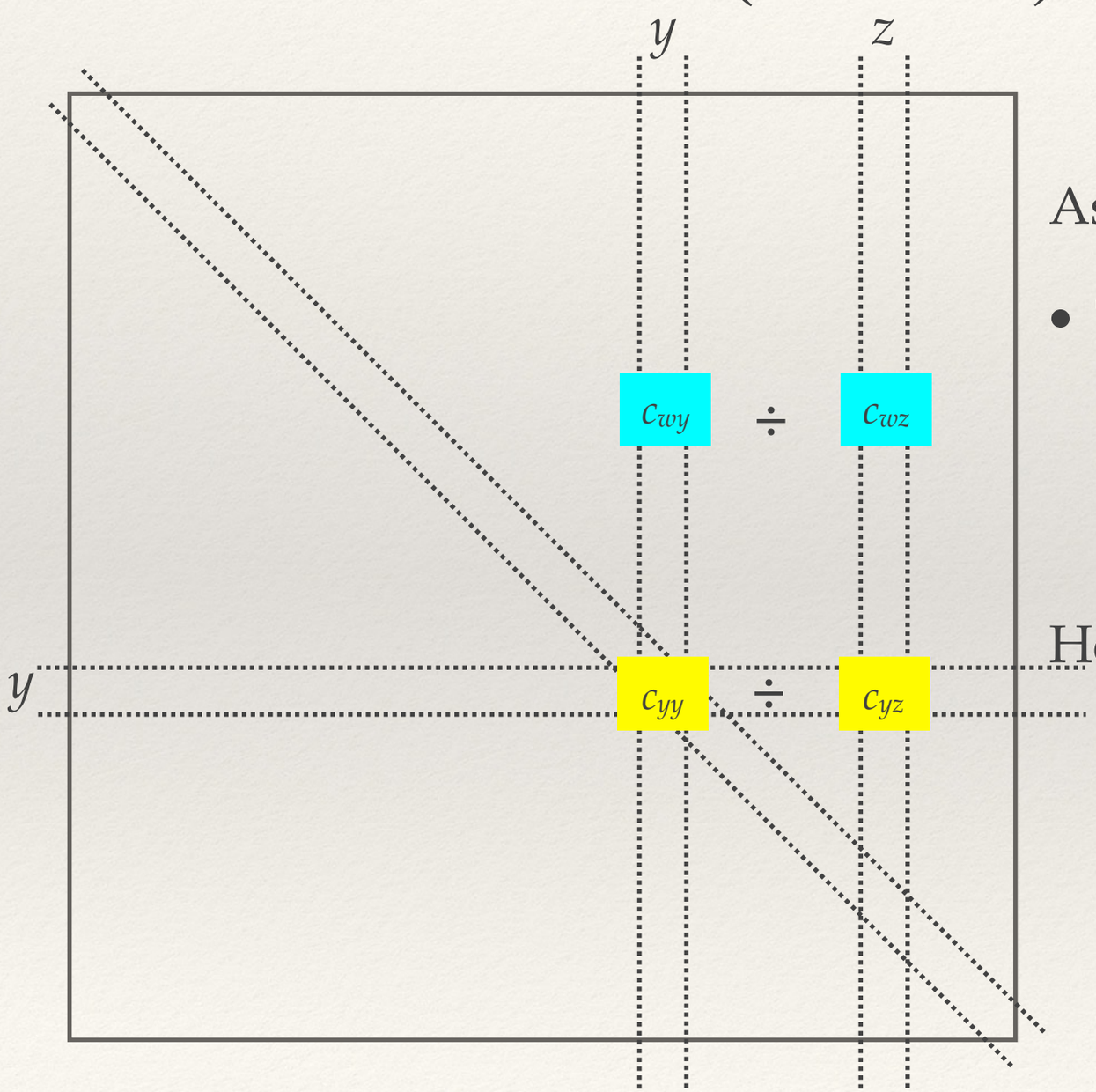
- Then $c_{wy} \leq q\, c_{wz}$

  As a consequence if $q < 1$...

$$r_y = \sum_w v_w c_{wy} \leq \sum_w v_w \frac{c_{wz} c_{yy}}{c_{yz}} \leq \sum_w v_w c_{wz} = r_z$$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

- Then $c_{wy} \leq q\, c_{wz}$

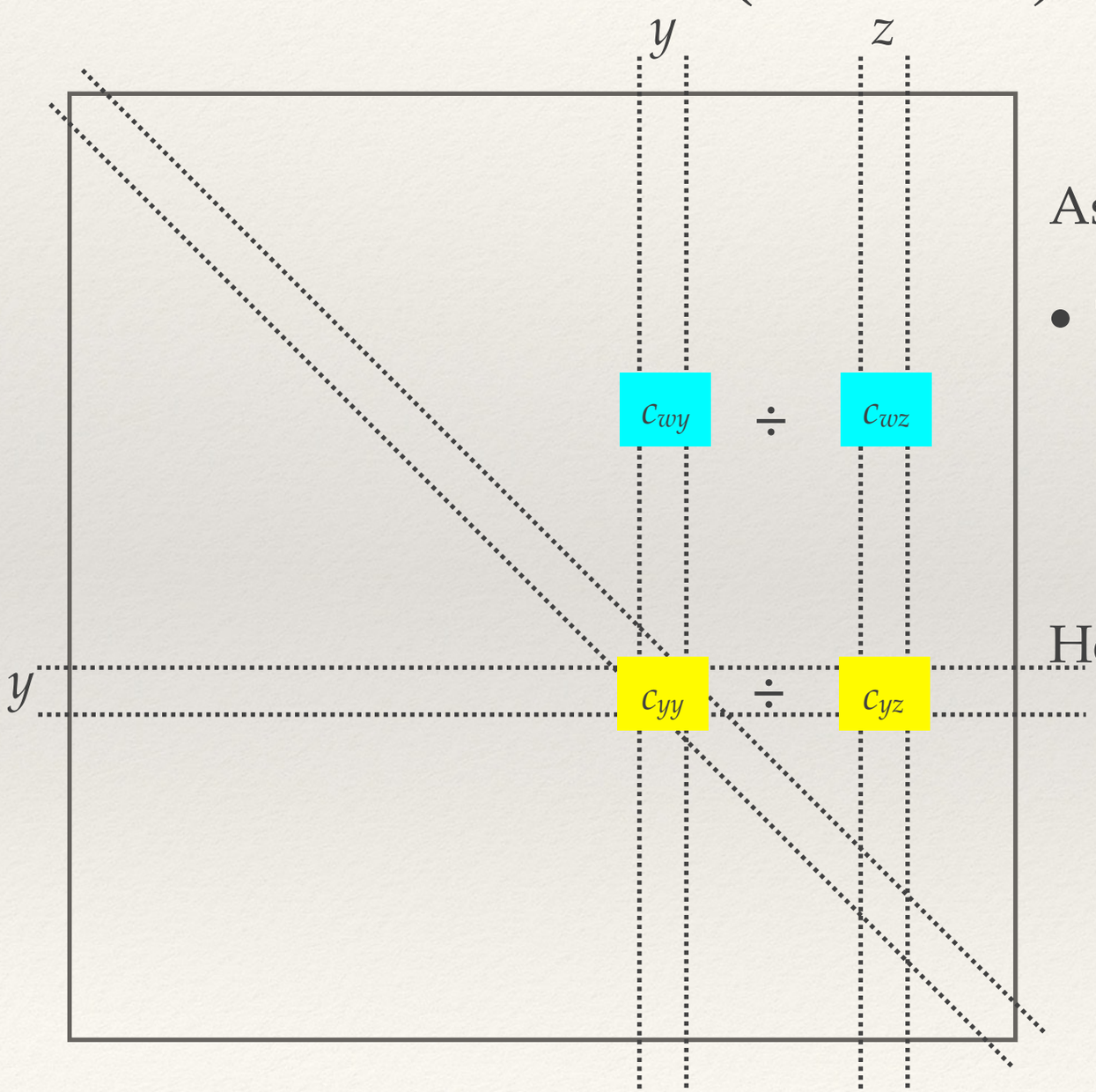
  As a consequence if $q < 1$...

$$r_y = \sum_w v_w c_{wy} \leq \sum_w v_w \frac{c_{wz} c_{yy}}{c_{yz}} \leq \sum_w v_w c_{wz} = r_z$$

Hence:

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

- Then $c_{wy} \leq q\, c_{wz}$

  As a consequence if $q < 1$...

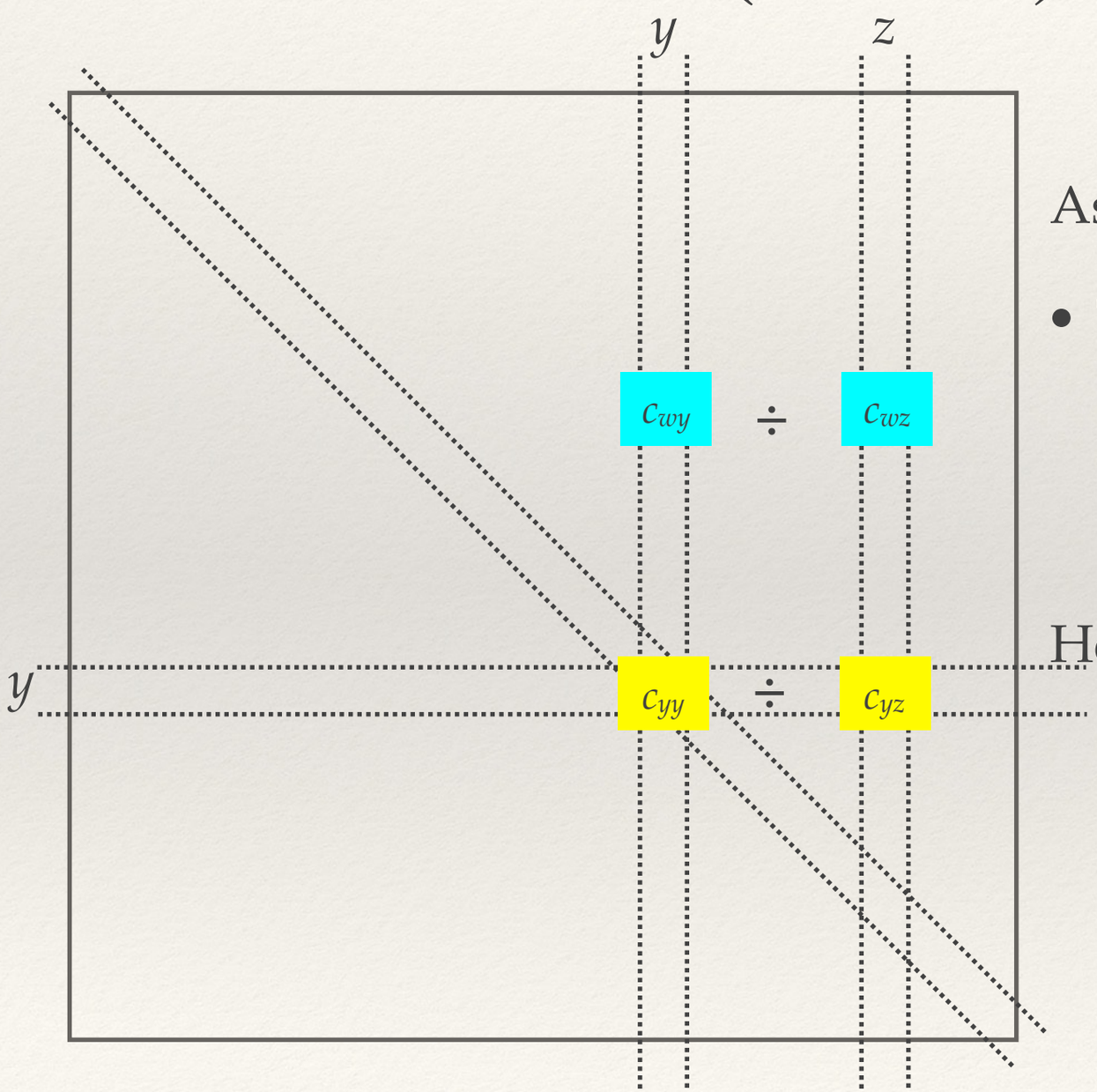
$$r_y = \sum_w v_w c_{wy} \leq \sum_w v_w \frac{c_{wz} c_{yy}}{c_{yz}} \leq \sum_w v_w c_{wz} = r_z$$

Hence:

- If $r_z \leq r_y$ then $q \geq 1$
- If $r_z < r_y$ then $q > 1$

# Lemma (ext. Willoughby, 1977)

$$C = (I - \alpha M)^{-1} \qquad \mathbf{r} = \mathbf{v}C$$




THIS CONDITION IS NECESSARY!

Assume $c_{yz} > 0$ and let $q = c_{yy}/c_{yz}$

- Then $c_{wy} \leq q\, c_{wz}$

  As a consequence if $q < 1$...

$$r_y = \sum_w v_w c_{wy} \leq \sum_w v_w \frac{c_{wz} c_{yy}}{c_{yz}} \leq \sum_w v_w c_{wz} = r_z$$

Hence:

- If $r_z \leq r_y$ then $q \geq 1$
- If $r_z < r_y$ then $q > 1$

# PageRank as a special case of damped spectral ranking

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

only row $x$ is nonzero

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \to y$ we obtain a new matrix $M'$ and

"old" outneighbours of $x$

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

only the $y$-th column is positive

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix}$$

# PageRank as a special case of damped spectral ranking

$$\mathbf{r} = \mathbf{v}(I - \alpha M)^{-1}$$

- In the case of PageRank, $M = G_r$
- When adding the arc $x \rightarrow y$ we obtain a new matrix $M'$ and

$$M' - M = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & -\frac{1}{d(d+1)} & \dots & \frac{1}{d} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{pmatrix} \delta$$

# Rank monotonicity of PageRank (1)

# Rank monotonicity of PageRank (1)

- By the Sherman-Morrison formula:

$$\mathbf{r}' - \mathbf{r} = \kappa\delta\,(I - \alpha M)^{-1}$$

for some suitable positive constant $\kappa$. For simplicity I will assume $\kappa$=1

# Rank monotonicity of PageRank (1)

- By the Sherman-Morrison formula:

$$\mathbf{r}' - \mathbf{r} = \kappa\delta\,(I - \alpha M)^{-1}$$

  for some suitable positive constant $\kappa$. For simplicity I will assume $\kappa$=1

- We need to prove that

$$0 < r_z \leq r_y \text{ implies } r'_z < r'_y$$

# Rank monotonicity of PageRank (1)

- By the Sherman-Morrison formula:

$$\mathbf{r}' - \mathbf{r} = \kappa\delta\,(I - \alpha M)^{-1}$$

  for some suitable positive constant $\kappa$. For simplicity I will assume $\kappa$=1

- We need to prove that

$$0 < r_z \leq r_y \text{ implies } r'_z < r'_y$$

- We will in fact prove that

$$0 < r_z \leq r_y \text{ implies } [\mathbf{r}' - \mathbf{r}]_z < [\mathbf{r}' - \mathbf{r}]_y$$

# Rank monotonicity of PageRank (1)

- By the Sherman-Morrison formula:

$$\mathbf{r}' - \mathbf{r} = \kappa\delta\,(I - \alpha M)^{-1}$$

  for some suitable positive constant $\kappa$. For simplicity I will assume $\kappa$=1

- We need to prove that

$$0 < r_z \leq r_y \text{ implies } r'_z < r'_y$$

- We will in fact prove that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

# Rank monotonicity of PageRank (2)

# Rank monotonicity of PageRank (2)

- We aim at proving that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

# Rank monotonicity of PageRank (2)

- We aim at proving that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

- Let $c_{yz} > 0$ (the other case is easy), and $q = c_{yy}/c_{yz}$

# Rank monotonicity of PageRank (2)

- We aim at proving that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

- Let $c_{yz} > 0$ (the other case is easy), and $q = c_{yy}/c_{yz}$

$$[\delta(1 - \alpha M)^{-1}]_y = \delta_y c_{yy} - \sum_{w \neq y} |\delta_w| c_{wy} \geq \delta_y q c_{yz} - \sum_{w \neq y} q |\delta_w| c_{wz}$$

# Rank monotonicity of PageRank (2)

- We aim at proving that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

- Let $c_{yz} > 0$ (the other case is easy), and $q = c_{yy}/c_{yz}$

$$[\delta(1 - \alpha M)^{-1}]_y = \delta_y c_{yy} - \sum_{w \neq y} |\delta_w| c_{wy} \geq \delta_y q c_{yz} - \sum_{w \neq y} q |\delta_w| c_{wz}$$

- By the Lemma, $r_z \leq r_y \implies q \geq 1$

# Rank monotonicity of PageRank (2)

- We aim at proving that

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z < [\delta(I - \alpha M)^{-1}]_y$$

- Let $c_{yz} > 0$ (the other case is easy), and $q = c_{yy}/c_{yz}$

$$[\delta(1 - \alpha M)^{-1}]_y = \delta_y c_{yy} - \sum_{w \neq y} |\delta_w| c_{wy} \geq \delta_y q c_{yz} - \sum_{w \neq y} q |\delta_w| c_{wz}$$

- By the Lemma, $r_z \leq r_y \implies q \geq 1$

$$\geq \delta_y c_{yz} - \sum_{w \neq y} |\delta_w| c_{wz} = [\delta(1 - \alpha M)^{-1}]_z$$

# Rank monotonicity of PageRank (3)

# Rank monotonicity of PageRank (3)

- We in fact proved only

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z \leq [\delta(I - \alpha M)^{-1}]_y$$

# Rank monotonicity of PageRank (3)

- We in fact proved only

$$0 < r_z \leq r_y \text{ implies } [\delta(I - \alpha M)^{-1}]_z \leq [\delta(I - \alpha M)^{-1}]_y$$

- The strict inequality requires more work…

# Take-home messages

# Take-home messages

- Centrality is **important** and ubiquitous

# Take-home messages

- Centrality is **important** and ubiquitous
- **A jungle of indices**: taxonomies (and generalizations) are of help

# Take-home messages

- Centrality is **important** and ubiquitous
- **A jungle of indices**: taxonomies (and generalizations) are of help
- **Axiomatization** is a good way to make sense of so many indices

# Take-home messages

- Centrality is **important** and ubiquitous
- **A jungle of indices**: taxonomies (and generalizations) are of help
- **Axiomatization** is a good way to make sense of so many indices
- **A jungle of axioms**: taxonomies (and generalizations) are of help

# Take-home messages

- Centrality is **important** and ubiquitous
- **A jungle of indices**: taxonomies (and generalizations) are of help
- **Axiomatization** is a good way to make sense of so many indices
- **A jungle of axioms**: taxonomies (and generalizations) are of help
- Apparently trivial properties fail to hold, or require *a lot of work* to be proved

# Take-home messages

- Centrality is **important** and ubiquitous
- **A jungle of indices**: taxonomies (and generalizations) are of help
- **Axiomatization** is a good way to make sense of so many indices
- **A jungle of axioms**: taxonomies (and generalizations) are of help
- Apparently trivial properties fail to hold, or require *a lot of work* to be proved
- Beware, **it's a wild world out there**

Thanks for your attention!