Enforcing Katz and PageRank Centrality Measures in Complex Networks

Computational Aspects of Complex Networks

Università di Roma "Tor Vergata"

Fabio Durastante (fdurastante.github.io)
Università di Pisa,
Dipartimento di Matematica

December 6, 2024

Dipartimento
di Matematica
Università di Pisa

With a little help of my friends

1 Collaborators

Stefano Cipolla

University of Southampton

✉ s.cipolla@soton.ac.uk

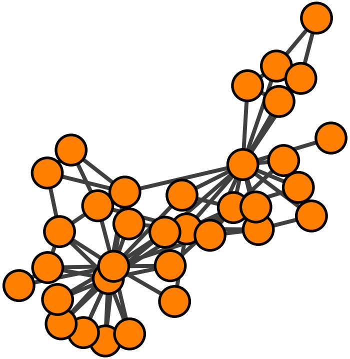
Beatrice Meini

Università di Pisa

✉ beatrice.meini@unipi.it

Table of Contents

Complex Networks and what we want to do with them

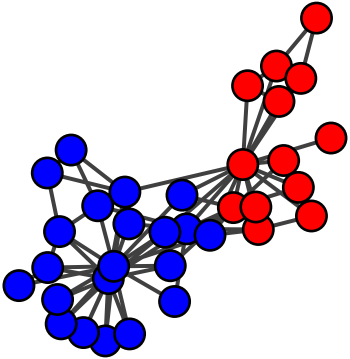


A complex network is a graph with **non-trivial** topological features, neither a structured graph nor a *completely* random graph.

We are interested in tasks in **exploratory data analysis**, that is analyzing the data to summarize their main characteristics:

- Divide the nodes into groups that are in the same community (clustering),
- Find the “most relevant” nodes in the network (centrality),
- Find the “most relevant” edge in the network (edge centrality)
- Individuation of motifs, computation of fluxes, maximum cuts, *etc.*

Complex Networks and what we want to do with them

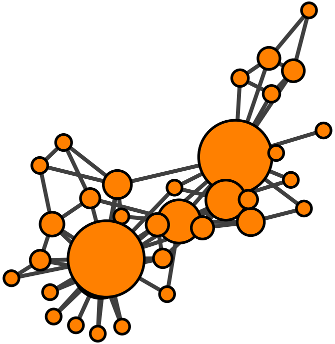


A complex network is a graph with **non-trivial** topological features, neither a structured graph nor a *completely* random graph.

We are interested in tasks in **exploratory data analysis**, that is analyzing the data to summarize their main characteristics:

- Divide the nodes into groups that are in the same community (clustering),
- Find the "most relevant" nodes in the network (centrality),
- Find the "most relevant" edge in the network (edge centrality)
- Individuation of motifs, computation of fluxes, maximum cuts, *etc.*

Complex Networks and what we want to do with them

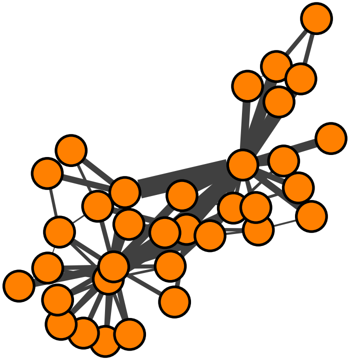


A complex network is a graph with **non-trivial** topological features, neither a structured graph nor a *completely* random graph.

We are interested in tasks in **exploratory data analysis**, that is analyzing the data to summarize their main characteristics:

- Divide the nodes into groups that are in the same community (clustering),
- Find the "most relevant" nodes in the network (centrality),
- Find the "most relevant" edge in the network (edge centrality)
- Individuation of motifs, computation of fluxes, maximum cuts, *etc.*

Complex Networks and what we want to do with them



A complex network is a graph with **non-trivial** topological features, neither a structured graph nor a *completely* random graph.

We are interested in tasks in **exploratory data analysis**, that is analyzing the data to summarize their main characteristics:

- Divide the nodes into groups that are in the same community (clustering),
- Find the “most relevant” nodes in the network (centrality),
- Find the “most relevant” edge in the network (edge centrality)
- Individuation of motifs, computation of fluxes, maximum cuts, *etc.*

Notation

Network

A *network* $\mathcal{G} = (V, E)$ is defined as a pair of sets: a set $V = \{1, 2, \ldots, n\}$ of *nodes* and a set $E \subset V \times V$ of *edges* between them.

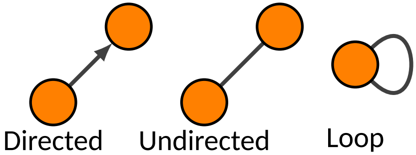
Notation

Network

A *network* $\mathcal{G} = (V, E)$ is defined as a pair of sets: a set $V = \{1, 2, \ldots, n\}$ of *nodes* and a set $E \subset V \times V$ of *edges* between them.

Directed/Undirected

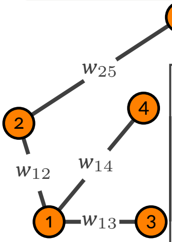
If $\forall\,(i,j) \in E$ then $(j, i) \in E$ the network is said to be *undirected* is *directed* otherwise.



An edge from a node to itself is called a *loop*.

Network

A *network* $\mathcal{G} = (V, E)$ is defined as a pair of sets: a set $V = \{1, 2, \ldots, n\}$ of *nodes* and a set $E \subset V \times V$ of *edges* between them.



$$\begin{bmatrix} 0 & w_{12} & w_{13} & w_{14} & 0 \\ w_{12} & 0 & 0 & 0 & w_{25} \\ w_{13} & 0 & 0 & 0 & 0 \\ w_{14} & 0 & 0 & 0 & 0 \\ 0 & w_{25} & 0 & 0 & 0 \end{bmatrix}$$

Adjacency Matrix

We represent a Network via its *adjacency matrix* $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, entrywise defined as

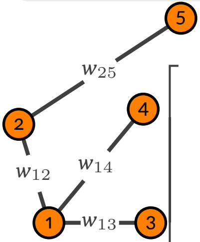
$$a_{ij} = \begin{cases} w_{ij} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

where $w_{ij} > 0$ is the weight of edge (i,j).

Notation

Network

A *network* $\mathcal{G} = (V, E)$ is defined as a pair of sets: a set $V = \{1, 2, \ldots, n\}$ of *nodes* and a set $E \subset V \times V$ of *edges* between them.



$$\begin{bmatrix} 0 & w_{12} & w_{13} & w_{14} & 0 \\ w_{12} & 0 & 0 & 0 & w_{25} \\ w_{13} & 0 & 0 & 0 & 0 \\ w_{14} & 0 & 0 & 0 & 0 \\ 0 & w_{25} & 0 & 0 & 0 \end{bmatrix}$$

Degree Matrix

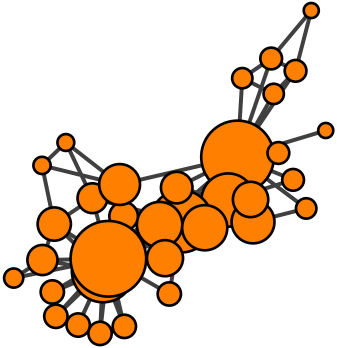
We call degree matrix the diagonal matrix

$$D = \mathrm{diag}(A\mathbf{1}),$$

where $\mathbf{1} = (1, 1, \ldots, 1)^T$.

Given $\alpha > 0$ such that $0 < \alpha\rho(A) < 1$, the Katz centrality of node v_i is the ith entry of the vector $\boldsymbol{\mu} = (I - \alpha A)^{-1}\mathbf{1}$.

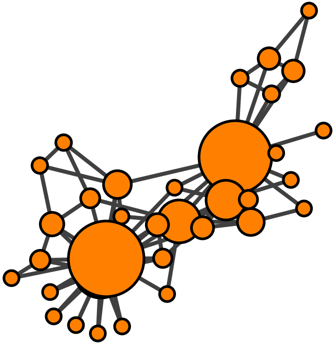
- $A \geq 0 \Rightarrow \boldsymbol{\mu} \geq \mathbf{1}$.



Katz, L. A new status index derived from sociometric analysis. *Psychometrika*. **18**, 39 - 43 (1953)

Given $\alpha \in (0,1)$ a *teletrasportation parameter*, and given $\mathbf{v} > 0$ a *personalization* vector such that $\mathbf{v}^T\mathbf{1} = 1$, the PageRank centrality of node v_i is the ith entry of the stationary vector $\boldsymbol{\pi}$ of the stochastic matrix G^T, where $G = \alpha(D^{-1}A)^T + (1-\alpha)\mathbf{v}\mathbf{1}^T$.

- $G\boldsymbol{\pi} = \boldsymbol{\pi}, \quad \boldsymbol{\pi}^T\mathbf{1} = 1,$
- $(I - \alpha(D^{-1}A)^T)\boldsymbol{\pi} = (1-\alpha)\mathbf{v}.$



Page , L. & Brin, S. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks*. **30**, 107 - 117 (1998)

Gleich, D. PageRank beyond the web. *SIAM Rev.*. **57**, 321-363 (2015)

To the *centralest*

ἐνθ’ ἄλλοις μὲν πᾶσιν ἔην δαίμων, οὐδέ ποθ’ Ἥρη
οὐδὲ Ποσειδάων οὐδὲ γλαυκώπιδι Κούρη,
ἀλλ’ ἔχον ὡς σφιν πρώτων ἀπήχθετο Ἴλιος ἠδὲ
καὶ Πρίαμος καὶ λαὸς Ἀλεξάνδρου ἕνεκ’ ἄτης,
ὃς νέκεσσε θεὰς, ὅτε οἱ μέσον αὐλὸν ἵκοντο,
τὴν δ’ ᾔνησ’ ἥ οἱ πόρε μαχλοσύνην ἀλεγεινήν.

Homer, Iliad 24.25

Can we **change the outcome** of the judgment given by the **centrality measure**?

And this was pleasing to all the others, but never to Hera
nor to Poseidon, nor to the flashing-eyed maiden,
but they remained hostile to sacred Ilios as in the beginning,
and to Priam and to his people, because of Alexander's folly,
he who insulted the goddesses when they came to his inner courtyard
and praised her who provided his grievous lust.

Homer, Iliad 24.25

- Can we **change the outcome** of the judgment given by the **centrality measure**?
- We want to find a **small perturbation** Δ of the matrix A that transforms μ and π into two vectors of our choice.

Table of Contents

Reformulation as an optimization problem

We can formulate the problem we want to solve as:

$$\mathcal{P}^{\text{Katz}}_{\alpha,\beta} : \quad \begin{array}{rl} \min\limits_{\Delta \in \mathbb{S}(A)} & J(\Delta) = \beta \|\Delta\|_F^2 + (1-\beta)\|\Delta\|_1, \\ \text{s.t.} & (I - \alpha(A+\Delta))^{-1}\mathbf{1} = \widehat{\boldsymbol{\mu}}, \\ & A + \Delta \geq 0, \end{array} \qquad 0 < \beta \leq 1.$$

- the objective function weighs between the **magnitude** of the perturbation ($\|\cdot\|_F^2$) and promoting the **sparsity** of the solution ($\|\cdot\|_1$),

Reformulation as an optimization problem

We can formulate the problem we want to solve as:

$$\mathcal{P}^{\text{Katz}}_{\alpha,\beta} : \quad \begin{array}{rl} \min\limits_{\Delta \in \mathbb{S}(A)} & J(\Delta) = \beta\|\Delta\|_F^2 + (1-\beta)\|\Delta\|_1, \\ \text{s.t.} & (I - \alpha(A+\Delta))^{-1}\mathbf{1} = \widehat{\boldsymbol{\mu}}, \\ & A + \Delta \geq 0, \end{array} \qquad 0 < \beta \leq 1.$$

- the objective function weighs between the **magnitude** of the perturbation ($\|\cdot\|_F^2$) and promoting the **sparsity** of the solution ($\|\cdot\|_1$),
- the first constraint requires that the perturbed network has the **desired centrality** $\widehat{\mu}$,

Reformulation as an optimization problem

We can formulate the problem we want to solve as:

$$
\mathcal{P}^{\text{Katz}}_{\alpha,\beta}: \quad \begin{array}{rl} \min\limits_{\Delta \in \mathbb{S}(A)} & J(\Delta) = \beta \|\Delta\|_F^2 + (1-\beta)\|\Delta\|_1, \\ \text{s.t.} & (I - \alpha(A+\Delta))^{-1}\mathbf{1} = \widehat{\boldsymbol{\mu}}, \\ & A + \Delta \geq 0, \end{array} \qquad 0 < \beta \leq 1.
$$

- the objective function weighs between the **magnitude** of the perturbation ($\|\cdot\|_F^2$) and promoting the **sparsity** of the solution ($\|\cdot\|_1$),
- the first constraint requires that the perturbed network has the **desired centrality** $\widehat{\boldsymbol{\mu}}$,
- the second constraint ensures that the network weights still make sense,

Reformulation as an optimization problem

We can formulate the problem we want to solve as:

$$\mathcal{P}^{\text{Katz}}_{\alpha,\beta} : \quad \begin{array}{rl} \displaystyle\min_{\Delta \in \mathbb{S}(A)} & J(\Delta) = \beta\|\Delta\|_F^2 + (1-\beta)\|\Delta\|_1, \\ \text{s.t.} & (I - \alpha(A+\Delta))^{-1}\mathbf{1} = \widehat{\boldsymbol{\mu}}, \\ & A + \Delta \geq 0, \end{array} \qquad 0 < \beta \leq 1.$$

- the objective function weighs between the **magnitude** of the perturbation ($\|\cdot\|_F^2$) and promoting the **sparsity** of the solution ($\|\cdot\|_1$),
- the first constraint requires that the perturbed network has the **desired centrality** $\widehat{\boldsymbol{\mu}}$,
- the second constraint ensures that the network weights still make sense,
- we can have requirements on the sparsity pattern of Δ, e.g., we can require that it has a pattern contained in the sparsity pattern of A: $\mathbb{S}(A)$.

Does this problem have a solution?

Proposition (Cipolla, D., Meini)

Given $\widehat{\boldsymbol{\mu}} \geq \mathbf{1}$, $A \geq 0$ such that $A\mathbf{1} > \mathbf{0}$, and $\alpha > 0$ such that $\rho(A) < 1/\alpha$, then the set of matrices $\Delta \in \mathbb{S}(A)$ such that $(I - \alpha(A + \Delta))\widehat{\boldsymbol{\mu}} - \mathbf{1} = \mathbf{0}$ and $A + \Delta \geq 0$ is non-empty. Moreover, for any such matrix Δ we have $\rho(A + \Delta) < 1/\alpha$.

- In optimization *parlance* this tells us that the **constraints are feasible**.
- If A is *nonnegative* and *irreducible* the condition $A\mathbf{1} > 0$ is automatically satisfied.

Does this problem have a solution?

Proposition (Cipolla, D., Meini)

Given $\widehat{\boldsymbol{\mu}} \geq \mathbf{1}$, $A \geq 0$ such that $A\mathbf{1} > \mathbf{0}$, and $\alpha > 0$ such that $\rho(A) < 1/\alpha$, then the set of matrices $\Delta \in \mathbb{S}(A)$ such that $(I - \alpha(A + \Delta))\widehat{\boldsymbol{\mu}} - \mathbf{1} = \mathbf{0}$ and $A + \Delta \geq 0$ is non-empty. Moreover, for any such matrix Δ we have $\rho(A + \Delta) < 1/\alpha$.

- In optimization *parlance* this tells us that the **constraints are feasible**.
- If A is *nonnegative* and *irreducible* the condition $A\mathbf{1} > 0$ is automatically satisfied.
- The same feasibility result holds if $\Delta \in \mathbb{S}(M)$, for $M \in \mathbb{R}^{n \times n}$ and $M \geq 0$, $M\mathbf{1} > \mathbf{0}$.

Does this problem have a solution?

Proposition (Cipolla, D., Meini)

Given $\widehat{\boldsymbol{\mu}} \geq \mathbf{1}$, $A \geq 0$ such that $A\mathbf{1} > \mathbf{0}$, and $\alpha > 0$ such that $\rho(A) < 1/\alpha$, then the set of matrices $\Delta \in \mathbb{S}(A)$ such that $(I - \alpha(A + \Delta))\widehat{\boldsymbol{\mu}} - \mathbf{1} = \mathbf{0}$ and $A + \Delta \geq 0$ is non-empty. Moreover, for any such matrix Δ we have $\rho(A + \Delta) < 1/\alpha$.

- In optimization *parlance* this tells us that the **constraints are feasible**.
- If A is *nonnegative* and *irreducible* the condition $A\mathbf{1} > 0$ is automatically satisfied.
- The same feasibility result holds if $\Delta \in \mathbb{S}(M)$, for $M \in \mathbb{R}^{n \times n}$ and $M \geq 0$, $M\mathbf{1} > \mathbf{0}$.

What **optimization algorithm** do we actually use to **solve the problem**?

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

$$\begin{aligned} \min \quad & J(\Delta) = \beta \| \operatorname{vec}(\Delta) \|_2^2 + (1-\beta) \| \operatorname{vec}(\Delta) \|_1, \\ \text{s.t.} \quad & (\widehat{\boldsymbol{\mu}}^T \otimes I) \operatorname{vec}(\Delta) = \frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}}, \\ & \operatorname{diag}(\operatorname{vec}(\mathbf{1}\mathbf{1}^T - M \circ \mathbf{1}\mathbf{1}^T)) \operatorname{vec}(\Delta) = \mathbf{0}, \\ & -\operatorname{vec}(A) \leq \operatorname{vec}(\Delta). \end{aligned}$$

- First we **vectorize everything**
  - $\operatorname{vec}(\cdot)$ stacks the column of its matrix argument,
  - $\otimes$ is the Kronecker product and $\operatorname{vec}(AXB) = (B^T \otimes A) \operatorname{vec}(X)$,
  - $\circ$ is the Hadamard product.

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^{n_{P_M}}} \quad & J(\mathbf{x}) = \beta \|P_M^T \mathbf{x}\|_2^2 + (1-\beta)\|P_M^T\mathbf{x}\|_1, \\
\text{s.t.} \quad & (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T \mathbf{x} = \frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}}, \\
& -P_M \operatorname{vec}(A) \leq \mathbf{x}.
\end{aligned}
$$

- First we **vectorize everything**,
- then we **restrict the problem** to the variables associated to the possibly non-zero elements of Δ
  - $P_M \in \mathbb{R}^{n_{P_M} \times n^2}$ is the projector onto the pattern of M,
  - $\mathbf{x} \in \mathbb{R}^{n_{P_M}}$ is defined as $\mathbf{x} = P_M \operatorname{vec}(\Delta)$.

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

$$\begin{aligned} \min_{\bar{\mathbf{x}} \in \mathbb{R}^{n_{P_M}}} \quad & J(\bar{\mathbf{x}}) = \|\bar{\mathbf{x}} - P_M \operatorname{vec}(A))\|_2^2 + \tau \|\bar{\mathbf{x}} - P_M \operatorname{vec}(A)\|_1, \\ \text{s.t.} \quad & (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T \bar{\mathbf{x}} = \frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}} + (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T P_M \operatorname{vec}(A), \\ & \bar{\mathbf{x}} \geq 0. \end{aligned}$$

- First we **vectorize everything**,
- then we **restrict the problem** to the variables associated to the possibly non-zero elements of Δ,
- Define $\bar{\mathbf{x}} = \mathbf{x} + P_M \operatorname{vec}(A)$ and $\tau = (1-\beta)/\beta$.

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

$$\begin{aligned}\min_{\bar{\mathbf{x}}\in\mathbb{R}^{n_{P_M}}} \quad & J(\bar{\mathbf{x}}) = \|\bar{\mathbf{x}} - P_M \operatorname{vec}(A))\|_2^2 + \tau\|\bar{\mathbf{x}} - P_M \operatorname{vec}(A)\|_1, \\ \text{s.t.} \quad & (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T \bar{\mathbf{x}} = \frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}} + (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T P_M \operatorname{vec}(A), \\ & \bar{\mathbf{x}} \geq 0.\end{aligned}$$

- First we **vectorize everything**,
- then we **restrict the problem** to the variables associated to the possibly non-zero elements of Δ,
- Define $\bar{\mathbf{x}} = \mathbf{x} + P_M \operatorname{vec}(A)$ and $\tau = (1-\beta)/\beta$, and get rid of the $\|\cdot\|_1$-norm by adding the nonnegative auxiliary variables:
  - $\ell^+ = \max(\bar{\mathbf{x}} - P_M \operatorname{vec}(A), 0)$ and $\ell^- = \max(-(\bar{\mathbf{x}} - P_M \operatorname{vec}(A)), 0)$,

Formulation as Quadratic Programming (QP) problem

Rewriting the problem requires some formal manipulations

$$\begin{aligned} \min_{\bar{\mathbf{x}} \in \mathbb{R}^{n_{P_M}}} \quad & J(\bar{\mathbf{x}}) = \|\bar{\mathbf{x}} - P_M \operatorname{vec}(A))\|_2^2 + \tau \|\bar{\mathbf{x}} - P_M \operatorname{vec}(A)\|_1, \\ \text{s.t.} \quad & (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T \bar{\mathbf{x}} = \frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}} + (\widehat{\boldsymbol{\mu}}^T \otimes I) P_M^T P_M \operatorname{vec}(A), \\ & \bar{\mathbf{x}} \geq 0. \end{aligned}$$

- First we **vectorize everything**,
- then we **restrict the problem** to the variables associated to the possibly non-zero elements of Δ,
- Define $\bar{\mathbf{x}} = \mathbf{x} + P_M \operatorname{vec}(A)$ and $\tau = (1-\beta)/\beta$, and get rid of the $\|\cdot\|_1$-norm by adding the nonnegative auxiliary variables:
  - $\ell^+ = \max(\bar{\mathbf{x}} - P_M \operatorname{vec}(A), 0)$ and $\ell^- = \max(-(\bar{\mathbf{x}} - P_M \operatorname{vec}(A)), 0)$,
  - $\ell^+ - \ell^- = \bar{\mathbf{x}} - P_M \operatorname{vec}(A)$ and $\|\bar{\mathbf{x}} - P_M \operatorname{vec}(A)\|_1 = \mathbf{1}^T \ell^+ + \mathbf{1}^T \ell^-$

Which finally brings us to the QP problem in standard form:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^{3n_{P_M}}} \quad & \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & L\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq 0, \end{aligned}$$

with:

$$\begin{aligned} Q &= \operatorname{blkdiag}(2I, 0, 0), \quad \mathbf{c} = (-2P_M \operatorname{vec}(A); \tau\mathbf{1}; \tau\mathbf{1}), \\ L &= \begin{bmatrix} (\widehat{\boldsymbol{\mu}}^T \otimes I)P_M^T & 0 & 0 \\ -I & I & -I \end{bmatrix}, \\ \mathbf{b} &= \left(\frac{1}{\alpha}(\widehat{\boldsymbol{\mu}} - \mathbf{1}) - A\widehat{\boldsymbol{\mu}} + (\widehat{\boldsymbol{\mu}}^T \otimes I)P_M^T P_M \operatorname{vec}(A); -P_M \operatorname{vec}(A)\right), \\ \mathbf{x} &= (\bar{\mathbf{x}}; \boldsymbol{\ell}^+; \boldsymbol{\ell}^-) \in \mathbb{R}^{3n_{p_M}}, \quad \tau = (1-\beta)/\beta. \end{aligned}$$

Table of Contents

4 Enforcing PageRank Centrality

For the PageRank problem we have a prescribed $\widehat{\boldsymbol{\pi}}$ and we look for a Δ such that

$$\left(I - \alpha \left((D + \operatorname{diag}(\Delta\mathbf{1}))^{-1}(A + \Delta)\right)^T\right)\widehat{\boldsymbol{\pi}} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1),$$

Enforcing PageRank Centrality

For the PageRank problem we have a prescribed $\widehat{\pi}$ and we look for a Δ such that

$$\left(I - \alpha \left((D + \operatorname{diag}(\Delta \mathbf{1}))^{-1}(A + \Delta)\right)^T\right) \widehat{\boldsymbol{\pi}} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1),$$

⚠ This formulation of the problem is **not linear** in Δ like the one for Katz centrality!

Enforcing PageRank Centrality

For the PageRank problem we have a prescribed $\widehat{\pi}$ and we look for a Δ such that

$$\left(I - \alpha\left((D + \operatorname{diag}(\Delta\mathbf{1}))^{-1}(A + \Delta)\right)^T\right)\widehat{\pi} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1),$$

- This formulation of the problem is **not linear** in Δ like the one for Katz centrality!
- We make the **simplifying assumption** that $\Delta 1 = \mathbf{0}$.

And solve the problem in **two steps**, first we solve

$$\begin{aligned} \min_{\Delta \in \mathbb{S}(A+I)} \quad & \beta\|\Delta\|_F^2 + (1-\beta)\|\operatorname{off-diag}(\Delta)\|_1 \\ \text{s.t.} \quad & \left(I - \alpha(\operatorname{diag}(A\mathbf{1})^{-1}(A+\Delta))^T\right)\widehat{\pi} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1), \\ & \Delta\mathbf{1} = 0 \\ & \operatorname{off-diag}(A+\Delta) \geq 0. \end{aligned} \quad (⚖)$$

Enforcing PageRank Centrality

For the PageRank problem we have a prescribed $\widehat{\pi}$ and we look for a Δ such that

$$\left(I - \alpha\left((D + \operatorname{diag}(\Delta\mathbf{1}))^{-1}(A + \Delta)\right)^T\right)\widehat{\pi} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1),$$

⚠ This formulation of the problem is **not linear** in Δ like the one for Katz centrality!

🔨 We make the **simplifying assumption** that $\Delta 1 = \mathbf{0}$.

And solve the problem in **two steps**, first we solve

$$\begin{aligned} \min_{\Delta \in \mathbb{S}(A+I)} \quad & \beta\|\Delta\|_F^2 + (1-\beta)\|\operatorname{off-diag}(\Delta)\|_1 \\ \text{s.t.} \quad & \left(I - \alpha(\operatorname{diag}(A\mathbf{1})^{-1}(A+\Delta))^T\right)\widehat{\pi} = (1-\alpha)\mathbf{v}, \qquad \alpha \in (0,1), \\ & \Delta\mathbf{1} = 0 \\ & \operatorname{off-diag}(A+\Delta) \geq 0. \end{aligned} \qquad (⚖)$$

Then we use the Δ obtained from (⚖) to compute $\widehat{\alpha} \in (0,1)$ and a stochastic $\widehat{P} \in \mathbb{S}(A+I)$ such that $\widehat{G}^T\widehat{\pi} = \widehat{\pi}$, where $\widehat{G} = \widehat{\alpha}\widehat{P} + (1-\widehat{\alpha})\mathbf{1}\mathbf{v}^T$.

Why this complication?

Proposition (Cipolla, D., Meini) - ⚖

Given $A \geq 0$ irreducible, $\mathbf{v} \geq 0$ such that $\mathbf{v}^T\mathbf{1} = 1$, $\alpha \in (0,1)$, $\widehat{\boldsymbol{\pi}} \geq 0$ such that $\widehat{\boldsymbol{\pi}}^T\mathbf{1} = 1$, then the set of matrices $\Delta \in \mathbb{S}(A+I)$ such that:

- $\Delta\mathbf{1} = 0$,
- the off-diagonal entries of $A + \Delta$ are nonnegative,
- $\tilde{G}^T\widehat{\boldsymbol{\pi}} = \widehat{\boldsymbol{\pi}}$, where $\tilde{G} = \alpha D^{-1}(A+\Delta) + (1-\alpha)\mathbf{1}\mathbf{v}^T$ and $D = \operatorname{diag}(A\mathbf{1})$,

is non-empty.

- ✔ We can prove that the optimization problem (⚖) is feasible.
- 👁 The assumptions on the pattern can be slightly generalized to $\Delta \in \mathbb{S}(M+I)$, and M any non-negative irreducible matrix such that $M\mathbf{1} = A\mathbf{1}$.

Why this complication?

Proposition (Cipolla, D., Meini) - ⚖

Given a matrix Δ satisfying the conditions of Proposition ⚖, define $\theta = \min_i([D^{-1}(A+\Delta)]_{i,i})$. If $\theta \geq 0$, then $D^{-1}(A+\Delta)$ is stochastic. Otherwise, if $\theta < 0$, by setting $\widehat{r} = 1 - \alpha\theta$, then for any $r \geq \widehat{r}$ we have $\widehat{G}^T\widehat{\pi} = \widehat{\pi}$, where

$$\widehat{G} = \widehat{\alpha}\widehat{P} + (1-\widehat{\alpha})\mathbf{1}\mathbf{v}^T,$$

and

$$\widehat{\alpha} = 1 - \frac{1-\alpha}{r}, \quad \widehat{P} = \frac{1}{r-1+\alpha}\left(\alpha D^{-1}(A+\Delta) + (r-1)I\right),$$

with $\widehat{P}$ stochastic.

- ✔ We can then **obtain a solution to the initial problem** under the assumption that we (may) need to modify the teleportation parameter.

An example where things go wrong

To understand where the problem lies, let's look at a small example and look for a feasible Δ of the form $\Delta = D_{\boldsymbol{\sigma}}(A - D)$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{v} = {}^{1}\!/_{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \alpha = 0.75.$$

An example where things go wrong

To understand where the problem lies, let's look at a small example and look for a feasible Δ of the form $\Delta = D_{\boldsymbol{\sigma}}(A - D)$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \ \mathbf{v} = {}^1\!/_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \alpha = 0.75.$$

we have to determine a σ for which the perturbed matrix has the given $\widehat{\boldsymbol{\pi}}$, equivalently,

$$D^{-1}A = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}, \ \mathbf{w} = {}^1\!/_4 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ (I - D^{-1}A + \mathbf{1}\mathbf{w}^T)^{-1} = {}^1\!/_8 \begin{bmatrix} 7 & 2 & -1 \\ 1 & 6 & 1 \\ -1 & 2 & 7 \end{bmatrix}.$$

For $\boldsymbol{\sigma} = \alpha^{-1} D_{\widehat{\boldsymbol{\pi}}}^{-1} (I - D^{-1}A + \mathbf{1}\mathbf{w}^T)^{-T} \left((1-\alpha)\mathbf{v} - (I - \alpha D^{-1}A)^T \widehat{\boldsymbol{\pi}}\right) + \gamma D_{\widehat{\boldsymbol{\pi}}}^{-1}\mathbf{w}$.

An example where things go wrong

To understand where the problem lies, let's look at a small example and look for a feasible Δ of the form $\Delta = D_{\boldsymbol{\sigma}}(A - D)$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, (I - D^{-1}A + \mathbf{1}\mathbf{w}^T)^{-1} = {}^{1}\!/_{8} \begin{bmatrix} 7 & 2 & -1 \\ 1 & 6 & 1 \\ -1 & 2 & 7 \end{bmatrix}.$$

We select $\widehat{\boldsymbol{\pi}}^T = {}^{1}\!/_{3}\,[1, 1, 1]$ and get a feasible matrix Δ for any γ for (⚖), but none of them satisfies $A + \Delta > 0$, e.g.:

$$A + \Delta = {}^{1}\!/_{6} \begin{bmatrix} -5 & 11 & 0 \\ 8 & -4 & 8 \\ 0 & 11 & -5 \end{bmatrix}, \quad \widetilde{G} = {}^{1}\!/_{24} \begin{bmatrix} -13 & 35 & 2 \\ 14 & -4 & 14 \\ 2 & 35 & -13 \end{bmatrix}$$

An example where things go wrong

To understand where the problem lies, let's look at a small example and look for a feasible Δ of the form $\Delta = D_{\boldsymbol{\sigma}}(A - D)$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, (I - D^{-1}A + \mathbf{1}\mathbf{w}^T)^{-1} = {}^1/_8 \begin{bmatrix} 7 & 2 & -1 \\ 1 & 6 & 1 \\ -1 & 2 & 7 \end{bmatrix}.$$

We select $\widehat{\pi}^T = {}^1/_3\,[1, 1, 1]$ and get a feasible matrix Δ for any γ for (⚖), but none of them satisfies $A + \Delta > 0$, e.g.:

$$A + \Delta = {}^1/_6 \begin{bmatrix} -5 & 11 & 0 \\ 8 & -4 & 8 \\ 0 & 11 & -5 \end{bmatrix}, \quad \widetilde{G} = {}^1/_{24} \begin{bmatrix} -13 & 35 & 2 \\ 14 & -4 & 14 \\ 2 & 35 & -13 \end{bmatrix}$$

⚠ $\widetilde{G}$ has eigenvalues $1, -{}^5/_8, -{}^{13}/_8$, i.e., $\rho(\widetilde{G}) > 1$ and it has diagonal negative entries.

Formulation as Quadratic Programming problem

With steps similar to those of the Katz problem, we rewrite the problem in the QP form:

$$\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^{3n_{p_{M+I}}-2n}} \quad & \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\
\text{s.t.} \quad & L\mathbf{x} = \mathbf{b}, \\
& x_i \geq 0, \text{ if } i \in \widehat{\mathcal{C}}, \\
& x_i \text{ free if } i \in \widehat{\mathcal{F}}
\end{aligned}$$

Where we first define

$$\begin{aligned}
K \; &: K\,\text{vec}(\Delta) = \text{vec}(\Delta^T), \\
\widehat{\mathcal{F}} &= \{\bar{k}_i \in \{k_i = i + (i-1)n\}_{i=1}^{n} \;:\; (P_{M+I}\,\text{vec}(\Delta))_{\bar{k}_i} = \Delta_{ii}\}, \\
\widehat{\mathcal{C}} &= \left(\{1,\ldots,n_{p_{M+I}}\} \setminus \mathcal{F}\right) \cup \{n_{p_\Delta}+1,\ldots,3n_{p_\Delta}-2n\} = \mathcal{C} \cup \{n_{p_\Delta}+1,\ldots,3n_{p_\Delta}-2n\}, \\
\mathbf{a} &= P_{M+I}\,\text{diag}(\text{vec}(\mathbf{1}\mathbf{1}^T - I \circ \mathbf{1}\mathbf{1}^T))\,\text{vec}(A).
\end{aligned}$$

Formulation as Quadratic Programming problem

With steps similar to those of the Katz problem, we rewrite the problem in the QP form:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^{3n_{P_{M+I}}-2n}} \quad & \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & L\mathbf{x} = \mathbf{b}, \\ & x_i \geq 0, \text{ if } i \in \widehat{\mathcal{C}}, \\ & x_i \text{ free if } i \in \widehat{\mathcal{F}} \end{aligned}$$

Where

$$Q = \text{blkdiag}(2I, 0, 0),\ \mathbf{c} = (-2\mathbf{a}; \tau\mathbf{1}_{|\mathcal{C}|}; \tau\mathbf{1}_{|\mathcal{C}|}),\ L = \begin{bmatrix} ((\text{diag}(A\mathbf{1})^{-1}\widehat{\boldsymbol{\pi}})^T \otimes I)KP_{M+I}^T & 0 & 0 \\ (\mathbf{1}^T \otimes I)P_{M+I}^T & 0 & 0 \\ -I_{|\mathcal{C}|\times n_{P_{M+I}}} & I_{|\mathcal{C}|} & -I_{|\mathcal{C}|} \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} \frac{1}{\alpha}(\widehat{\boldsymbol{\pi}} - (1-\alpha)\mathbf{v}) - A^T\text{diag}(A\mathbf{1})^{-1}\widehat{\boldsymbol{\pi}} + ((\text{diag}(A\mathbf{1})^{-1}\widehat{\boldsymbol{\pi}})^T \otimes I)KP_{M+I}^T\mathbf{a} \\ (\mathbf{1}^T \otimes I)P_{M+I}^T\mathbf{a} \\ -\mathbf{a}_{\mathcal{C}} \end{bmatrix}.$$

Table of Contents

We solve the optimization problems via the **P**roximal **S**tabilised-**I**nterior **P**oint **M**ethod,[1]:

- It is well-suited for problems characterised by inherent ill-conditioning of the problem's data,
- The **P**roximal-**S**tabilization induces a *Primal-Dual Regularization*, i.e., we solve linear systems with matrix:

$$S_{\rho,\delta} = L(Q + \Theta^{-1} + \rho I)^{-1} L^T + \delta I$$

[1]Cipolla, S., Gondzio, J. & Zanetti, F. A regularized interior point method for sparse optimal transport on graphs. *European J. Oper. Res.*. (2023)

We solve the optimization problems via the **P**roximal **S**tabilised-**I**nterior **P**oint **M**ethod,[1]:

- It is well-suited for problems characterised by inherent ill-conditioning of the problem's data,
- The **P**roximal-**S**tabilization induces a *Primal-Dual Regularization*, i.e., we solve linear systems with matrix:

$$S_{\rho,\delta} = L(Q + \Theta^{-1} + \rho I)^{-1} L^T + \delta I$$

Θ^{-1} is a diagonal IPM iteration dependent matrix responsible for the identification of the *active-variables*,

[1] Cipolla, S., Gondzio, J. & Zanetti, F. A regularized interior point method for sparse optimal transport on graphs. *European J. Oper. Res.*. (2023)

Interior Point Methods – PS-IPM

We solve the optimization problems via the **P**roximal **S**tabilised-**I**nterior **P**oint **M**ethod,[1]:

- It is well-suited for problems characterised by inherent ill-conditioning of the problem's data,
- The **P**roximal-**S**tabilization induces a *Primal-Dual Regularization*, i.e., we solve linear systems with matrix:

$$S_{\rho,\delta} = L(Q + \Theta^{-1} + \rho I)^{-1} L^T + \delta I$$

If we don't use the *Primal-Dual regularization*, i.e., $\rho = \delta = 0$, then the $L(Q + \Theta^{-1})L^T$ matrices have diverging condition numbers as we reach convergence.

[1] Cipolla, S., Gondzio, J. & Zanetti, F. A regularized interior point method for sparse optimal transport on graphs. *European J. Oper. Res.*. (2023)

We solve the optimization problems via the **P**roximal **S**tabilised-**I**nterior **P**oint **M**ethod,[1]:

- It is well-suited for problems characterised by inherent ill-conditioning of the problem's data,
- The **P**roximal-**S**tabilization induces a *Primal-Dual Regularization*, i.e., we solve linear systems with matrix:

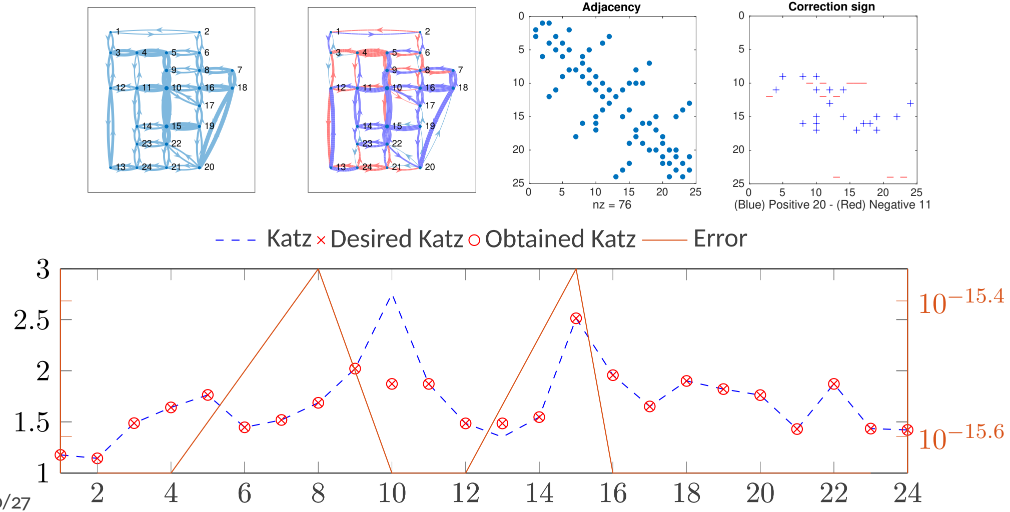
$$S_{\rho,\delta} = L(Q + \Theta^{-1} + \rho I)^{-1} L^T + \delta I$$

💣 We currently **solve** these **linear systems** with a direct method: **Cholesky**.

[1] Cipolla, S., Gondzio, J. & Zanetti, F. A regularized interior point method for sparse optimal transport on graphs. *European J. Oper. Res.*. (2023)

Table of Contents

A small example with Katz: Sioux Falls road network



A larger set of experiments

We look at two scenarios:

S1: $\widehat{\mu}/\widehat{\pi}$ puts the top $10, 20, 30, 40, 50\%$ of the nodes to their averaged value in μ/π;

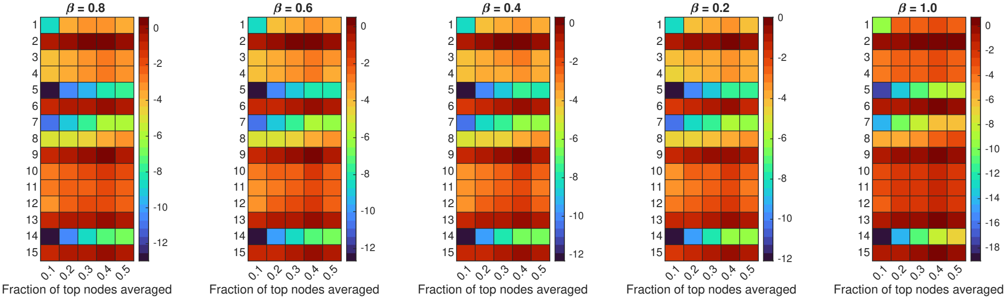
S2: $\widehat{\mu}/\widehat{\pi}$ reverts the rank of the top 10% of nodes in μ/π.

On the following test networks:

	Name	Type	n	nnz		Name	Type	n	nnz
1	EX5	comb. prob.	6545	295680	9	de2010	undir. wtd.	24115	116056
2	PGPgiantcompo	undir. multigr.	10680	48632	10	delaunay_n16	undir.	65536	393150
3	cage10	dir. wtd.	11397	150645	11	fe_4elt2	undir.	11143	65636
4	cage11	dir. wtd.	39082	559722	12	gre_1107	dir. wtd.	1107	5664
5	cs4	undir.	22499	87716	13	nh2010	undir. wtd.	48837	234550
6	ct2010	undir. wtd.	67578	336352	14	uk	undir.	4824	13674
7	cti	undir.	16840	96464	15	vt2010	undir. wtd.	32580	155598
8	data	undir.	2851	30186					

Enforcing Katz in S1

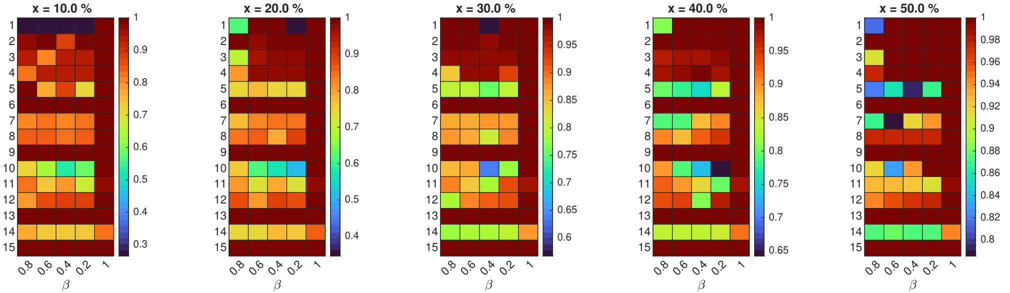
6 Numerical Examples



Value in $\log_{10}$-scale of the relative objective function $J(\Delta)/J(A)$. On the columns, we read the fraction of equalized vertices in increasing order, on the rows, the different test cases.

Enforcing Katz in S1

6 Numerical Examples



Number of nonzero entries scaled by the number of nonzero entries of the original adjacency matrix. On the columns we read the value of the β parameter, on the rows, the different test cases. Each block is obtained for a different percentage of the averaged nodes.

Enforcing Katz in S2 - $\beta = 1$

6 Numerical Examples

	$\text{cond}(LL^T)$	Iter	T (s)	$\|\Delta\|_F/\|A\|_F$	$+$	$-$	nnz Chol.	κ_τ
1]	1.199e+00	5	0.94	1.802e-02	146886	101416	6545	1.00
2]	5.889e+02	15	0.63	1.591e+00	24314	7828	10680	1.00
3]	5.673e+00	12	1.09	1.907e-02	9923	7698	11397	1.00
4]	1.137e+01	12	3.78	1.901e-02	33462	31160	39082	1.00
5]	2.169e+00	5	0.32	6.483e-02	17526	14068	22499	1.00
6]	5.723e+01	24	4.26	3.822e-01	111141	116710	67578	1.00
7]	3.447e+00	5	0.36	2.978e-02	8688	5453	16840	1.00
8]	1.037e+01	5	0.12	8.773e-02	2013	1578	2851	1.00
9]	4.786e+01	21	1.56	5.465e-01	42497	43350	24115	1.00
10]	1.169e+01	6	1.43	1.015e-01	100320	85616	65536	1.00
11]	6.158e+00	5	0.23	1.247e-01	7251	6783	11143	1.00
12]	7.461e+00	10	0.07	9.505e-02	991	2714	1107	1.00
13]	8.208e+01	23	2.95	4.453e-01	89297	92063	48837	1.00
14]	5.676e+00	6	0.08	4.968e-02	1002	669	4824	1.00
15]	8.338e+01	21	1.91	5.337e-01	46689	50026	32580	1.00

Enforcing Katz in S2 - $(1-\beta)/\beta = 100$

	$\mathrm{cond}(LL^T)$	Iter	T (s)	$\|\Delta\|_F/\|A\|_F$	$+$	$-$	`nnz Chol.`	κ_τ
1]	1.512e+02	22	15.23	2.485e+01	26476	18341	597905	1.00
2]	1.838e+03	24	2.84	3.576e+02	2994	4411	107944	1.00
3]	9.129e+01	17	5.45	2.678e+00	1330	2041	312687	1.00
4]	1.176e+02	20	28.98	5.288e+00	3961	7646	1158526	1.00
5]	1.522e+01	18	3.72	2.588e+01	10566	8091	197931	1.00
6]	2.242e+02	32	23.96	1.251e+07	84844	86183	740282	1.00
7]	2.212e+01	17	3.88	1.288e+01	7671	4061	209768	1.00
8]	5.358e+01	14	1.05	2.125e+01	743	637	63223	1.00
9]	1.884e+02	31	8.02	1.244e+07	35519	36245	256227	1.00
10]	8.551e+01	18	17.32	1.073e+02	25598	23428	851836	1.00
11]	4.505e+01	17	2.79	4.766e+01	3149	2935	142415	1.00
12]	4.178e+01	15	0.26	2.683e+00	116	167	12435	1.00
13]	3.151e+02	33	16.85	1.894e+07	60963	61177	517937	1.00
14]	1.896e+01	16	0.65	6.828e+00	799	541	32172	1.00
15]	3.026e+02	29	9.28	2.274e+07	25824	27907	343776	1.00

For the PageRank problem there exists an alternative procedure[2] that finds

$$\begin{aligned} \min_{\Delta \in \mathbb{R}^{n\times n}} \quad & \|\Delta\|_1 \\ \text{s.t.} \quad & \Delta \mathbf{1}_n = \mathbf{0}_n, \\ & \hat{\boldsymbol{\pi}}^\top \Delta = \hat{\boldsymbol{\pi}}^\top (I - G), \\ & \Delta + G \geq 0, \\ & \Delta_{i,j} = 0 \quad \text{for } (i,j) \notin \Omega. \end{aligned}$$

[2] Gillis, N. & Van Dooren, P. Assigning Stationary Distributions to Sparse Stochastic Matrices. *SIAM Journal On Matrix Analysis And Applications*. **45**, 2184-2210 (2024)

For the PageRank problem there exists an alternative procedure[2] which we can apply to

- $G = \alpha(D^{-1}A)^T + (1-\alpha)\mathbf{v}\mathbf{1}^T$,
- to then try and find from the obtained Δ a perturbation on A.

[2] Gillis, N. & Van Dooren, P. Assigning Stationary Distributions to Sparse Stochastic Matrices. *SIAM Journal On Matrix Analysis And Applications*. **45**, 2184-2210 (2024)

For the PageRank problem there exists an alternative procedure[2] which we can apply to

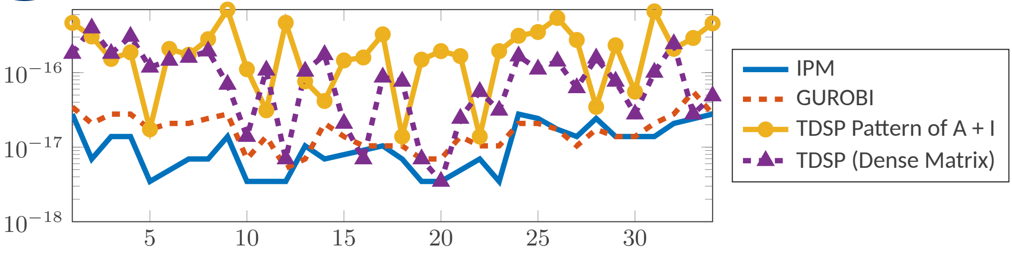
- $G = \alpha(D^{-1}A)^T + (1-\alpha)\mathbf{v}\mathbf{1}^T$,
- to then try and find from the obtained Δ a perturbation on A.

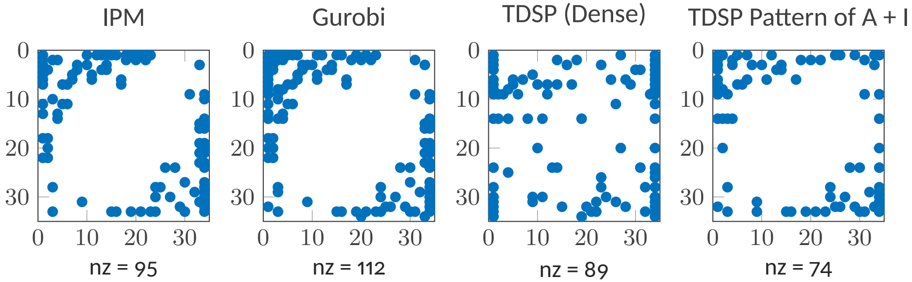
Our approach directly computes the perturbation for the adjacency matrix of the graph A.

[2] Gillis, N. & Van Dooren, P. Assigning Stationary Distributions to Sparse Stochastic Matrices. *SIAM Journal On Matrix Analysis And Applications*. **45**, 2184-2210 (2024)

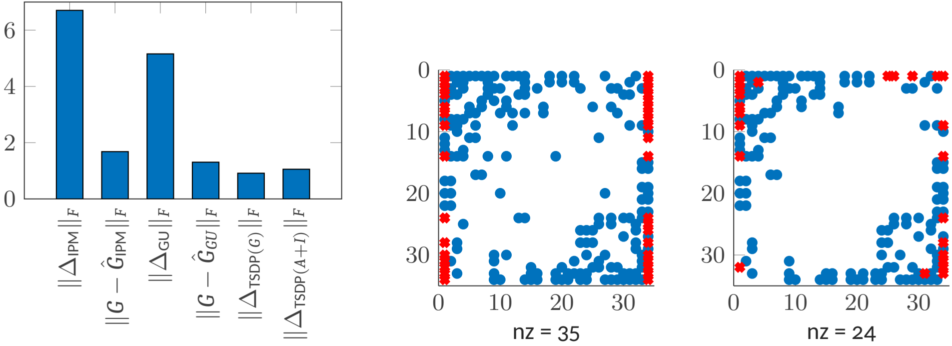
Comparison

6 Numerical Examples





Comparison



- ✗ If we recover the perturbation of A from the TSDP method we fail on the nonnegative requirement!
- The small increase in the perturbation norm is repaid by the preservation of the solution properties.

Conclusions

- ✔ Proved the existence of network modifications to obtain a given centrality index.
- ✔ Proposed a reliable algorithmic framework for computing such modifications.

What do we want to do in the future?

- Can we enforce walk-based centralities[3], e.g., $f(A+\Delta)\mathbf{1} = \widehat{\mathbf{t}}$?
- Can we *control* centralities of dynamical networks?
- Finding iterative strategies for linear systems arising in IPMs.

Where to find details and codes:

- Cipolla, S., D.,F. & Meini, B. Enforcing Katz and PageRank Centrality Measures in Complex Networks. (2024), arXiv:2409.02524.
- Code available at: Cirdans-Home/enforce-katz-and-pagerank

[3] Massei, S. & Tudisco, F. Optimizing network robustness via Krylov subspaces. *ESAIM M2AN*. **58**, 131-155 (2024).

Enforcing Katz and PageRank Centrality Measures in Complex Networks

Thank you for listening!

Any questions?