A doubly non-negative relaxation for modularity density maximization

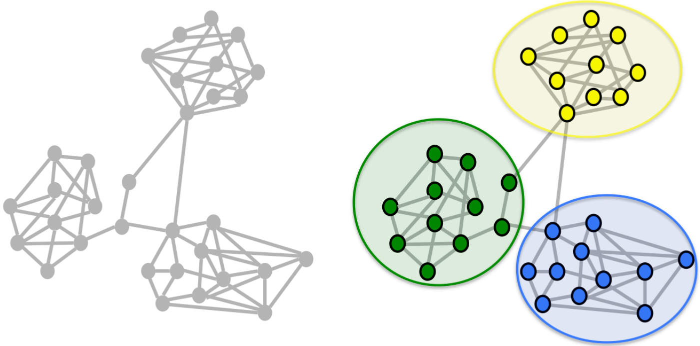
Yoichi Izunaga[†] Tomomi Matsui[‡] Yoshitsugu Yamamoto[†]

[†]University of Tsukuba
[‡]Tokyo Institute of Technology

November 22, 2015

Introduction (Community detection)

Community detection is grouping nodes of a graph into several parts:



- each part (community) consists of tightly connected nodes
- communities are loosely connected each other

Introduction (Quality measure)

Ratio cut (Cheng and Wei '91)

Normalized cut (Shi and Malik '00)

Min-max cut (Ding et al. '01)

Modularity (Newman and Girvan '04)

- degeneracy
- resolution limit (Fortunato and Barthelemy '07)
- NP-hard (Brandes et al. '08)

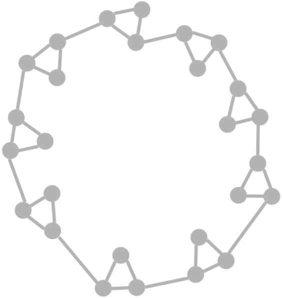
Modularity density (Li et al. '08)

- avoids the resolution limit
- NP-hard ?

Introduction (Quality measure)

Resolution limit

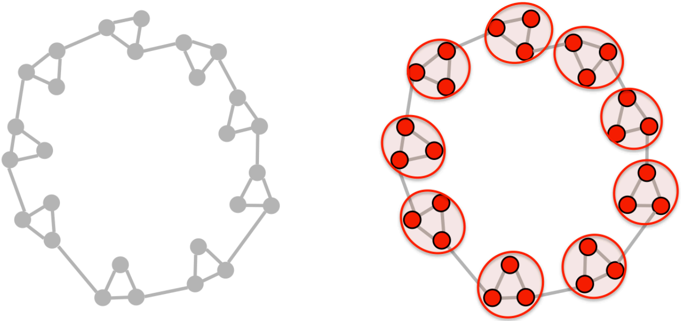
leaves small communities not identified and hidden inside larger ones



Introduction (Quality measure)

Resolution limit

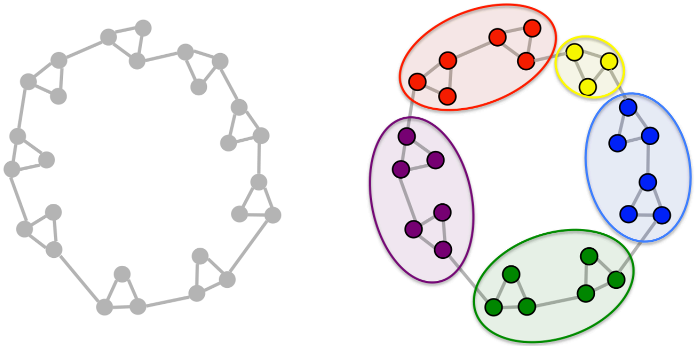
leaves small communities not identified and hidden inside larger ones



Introduction (Quality measure)

Resolution limit

leaves small communities not identified and hidden inside larger ones



Introduction (Quality measure)

Ratio cut (Cheng and Wei '91)

Normalized cut (Shi and Malik '00)

Min-max cut (Ding et al. '01)

Modularity (Newman and Girvan '04)

- degeneracy
- resolution limit (Fortunato and Barthelemy '07)
- NP-hard (Brandes et al. '08)

Modularity density (Li et al. '08)

- avoids the resolution limit
- NP-hard ?

Introduction (Modularity & Modularity density)

- undirected graph $G = (V, E)$ ($n = |V|$, $m = |E|$)
- $E(C, C') = \{ \{i, j\} \in E \mid i \in C,\ j \in C' \}$ for $C,\ C' \subseteq V$
  (when $C = C'$, we abbreviate it to $E(C)$)
- Π : a partition of the node set V

Modularity

$$\mathsf{M}(\Pi) = \sum_{C \in \Pi} \left(\frac{|E(C)|}{m} - \left(\frac{\sum_{C' \in \Pi} |E(C, C')|}{2m} \right)^2 \right)$$

Modularity density

$$\mathsf{MD}(\Pi) = \sum_{C \in \Pi} \left(\frac{2\,|E(C)| - \sum_{C' \in \Pi} |E(C, C')|}{|C|} \right)$$

Modularity density maximization

$$
\text{(P)} \quad \left| \begin{array}{ll}
\text{max.} & \displaystyle\sum_{p \in T} \left(\frac{2 \sum_{i \in V} \sum_{j \in V} a_{ij} x_{ip} x_{jp} - \sum_{i \in V} d_i x_{ip}}{\sum_{i \in V} x_{ip}} \right) \\
\text{s.t.} & \displaystyle\sum_{p \in T} x_{ip} = 1 \qquad (i \in V) \\
& \displaystyle\sum_{i \in V} x_{ip} \geq 1 \qquad (p \in T) \\
& x_{ip} \in \{0, 1\} \qquad (i \in V, p \in T).
\end{array} \right.
$$

- $T = \{1, \ldots, t\}$: index set of communities (t is unknown a priori)
- $A = (a_{ij})_{i,j \in V}$: adjacency matrix of G
- d_i : degree of node i (i.e., $d_i = \sum_{j \in V} a_{ij}$)
- x_{ip} : decision variable

$$
x_{ip} = \begin{cases} 1 & (i \in C_p) \\ 0 & (i \notin C_p) \end{cases}
$$

Introduction (Overview)

Costa '15

- formulated the problem as Mixed-Integer-Linear-Programming (MILP)
- made use of the *McCormick inequalities*
  $\Rightarrow$ need to solve an auxiliary problem
- solved the instances up to $n = 40$ by branch-and-bound alg.

Izunaga, Matsui, and Yamamoto

- show that the problem can be modeled as 0-1SDP
  $\Rightarrow$ does not require the number of communities t
- solve a relaxation problem to obtain an upper bound
- develop a heuristics to obtain a lower bound

Mixed-Integer-Linear-Programming (MILP)

MILP formulation

$$
\begin{array}{lll}
\text{max.} & \displaystyle\sum_{p \in T} \alpha_p & \\
\text{s.t.} & \displaystyle\sum_{p \in T} x_{ip} = 1 & (i \in V) \\
& \displaystyle\sum_{i \in V} x_{ip} \geq 1 & (p \in T) \\
& y_{ijp} \leq x_{ip},\ y_{ijp} \leq x_{jp} & (\{i,j\} \in E, p \in T) \\
& \displaystyle\sum_{i \in V} \gamma_{ip} \leq 4 \sum_{\{i,j\} \in E} y_{ijp} - \sum_{i \in V} d_i x_{ip} & (p \in T) \\
& L_\alpha x_{ip} \leq \gamma_{ip} \leq U_\alpha x_{ip} & (i \in V, p \in T) \\
& \alpha_p - U_\alpha(1 - x_{ip}) \leq \gamma_{ip} \leq \alpha_p - L_\alpha(1 - x_{ip}) & (i \in V, p \in T) \\
& x_{ip} \in \{0,1\} & (i \in V, p \in T) \\
& y_{ijp} \in \mathbb{R} & (\{i,j\} \in E, p \in T) \\
& L_\alpha \leq \alpha_p \leq U_\alpha & (p \in T) \\
& \gamma_{ip} \in \mathbb{R} & (i \in V, p \in T).
\end{array}
$$

0-1 semidefinite programming (0-1SDP)

$$\mathcal{S}_n^+ = \{\, Y \in \mathbb{R}^{n\times n} \mid Y^\top = Y,\, \forall \boldsymbol{d} \in \mathbb{R}^n, \boldsymbol{d}^\top Y \boldsymbol{d} \geq 0 \,\}$$
$$\mathcal{N}_n = \{\, Y \in \mathbb{R}^{n\times n} \mid Y^\top = Y,\, \forall i, j,\, y_{ij} \geq 0 \,\}$$
$$A = (a_{ij})_{i,j\in V}$$
$$D = \mathrm{Diag}(d_1, \dots, d_n) \in \mathbb{R}^{n\times n}$$
$$\boldsymbol{e}_k = (1, \dots, 1)^\top \in \mathbb{R}^k$$

Introducing a matrix $X \in \{\, 0, 1 \,\}^{n\times t}$, we have the following problem:

Matrix representation

$$(\mathrm{P}) \quad \begin{array}{|lll} \text{max.} & \mathrm{Tr}((2A - D)Z) \\ \text{s.t.} & X\boldsymbol{e}_t = \boldsymbol{e}_n \\ & X^\top \boldsymbol{e}_n \geq \boldsymbol{e}_t \\ & Z = X(X^\top X)^{-1} X^\top \\ & X \in \{\, 0, 1 \,\}^{n\times t}. \end{array}$$

$$\sum_{p\in T} x_{ip} = 1\,(i \in V) \;\Leftrightarrow\; X\boldsymbol{e}_t = \boldsymbol{e}_n$$

$$\sum_{i\in V} x_{ip} \geq 1\,(p \in T) \;\Leftrightarrow\; X^\top \boldsymbol{e}_n \geq \boldsymbol{e}_t$$

- (X, Z) is feasible for the problem (P)
  $\Rightarrow\ Z\boldsymbol{e}_n = \boldsymbol{e}_n,\ Z^2 = Z,\ Z \in \mathcal{N}_n$

Matrix representation

$$
\text{(P)}\quad \left|\ \begin{array}{ll} \text{max.} & \text{Tr}((2A - D)Z) \\ \text{s.t.} & X\boldsymbol{e}_t = \boldsymbol{e}_n \\ & X^\top \boldsymbol{e}_n \geq \boldsymbol{e}_t \\ & Z = X(X^\top X)^{-1}X^\top \\ & X \in \{0,1\}^{n\times t}. \end{array} \right.
$$

0-1SDP formulation

$$
(\bar{\text{P}})\quad \left|\ \begin{array}{ll} \text{max.} & \text{Tr}((2A - D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z^2 = Z \\ & Z \in \mathcal{N}_n. \end{array} \right.
$$

Lemma 1.

For any feasible solution Z of $(\bar{\text{P}})$, we can construct a feasible solution X which satisfies $Z = X(X^\top X)^{-1}X^\top$

$\Rightarrow$ the problem $(\bar{\text{P}})$ is equivalent to (P)

sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V,\ z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1,\ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1} \boldsymbol{e} = \boldsymbol{e},\ Z_{\bar{\mathcal{I}}_1}^2 = Z,\ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$

then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

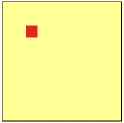
sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V, \ z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1, \ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1} \boldsymbol{e} = \boldsymbol{e}, \ Z^2_{\bar{\mathcal{I}}_1} = Z, \ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \dots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

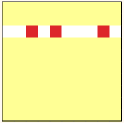
sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V,\ z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1,\ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1}\boldsymbol{e} = \boldsymbol{e},\ Z_{\bar{\mathcal{I}}_1}^2 = Z,\ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1}X^\top$.

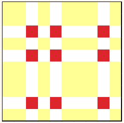
sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V,\ \ z_{i_1 i_1} = \max\{z_{ij} \mid i,j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i,j \in \mathcal{I}_1,\ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1}\boldsymbol{e} = \boldsymbol{e},\ Z_{\bar{\mathcal{I}}_1}^2 = Z,\ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

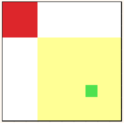
sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V,\ z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1,\ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1} \boldsymbol{e} = \boldsymbol{e},\ Z^2_{\bar{\mathcal{I}}_1} = Z,\ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

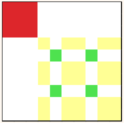
sketch of proof

$Z^2 = Z \;\Rightarrow\; Z \succeq 0 \;\Rightarrow\; \exists i_1 \in V,\;\; z_{i_1 i_1} = \max\{z_{ij} \mid i,j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i,j \in \mathcal{I}_1,\; z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1}\boldsymbol{e} = \boldsymbol{e},\; Z_{\bar{\mathcal{I}}_1}^2 = Z,\; Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1}X^\top$.

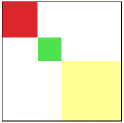
sketch of proof

$Z^2 = Z \Rightarrow Z \succeq 0 \Rightarrow \exists i_1 \in V,\ z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1,\ z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1} \boldsymbol{e} = \boldsymbol{e},\ Z_{\bar{\mathcal{I}}_1}^2 = Z,\ Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

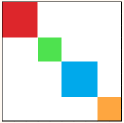
sketch of proof

$Z^2 = Z \;\Rightarrow\; Z \succeq 0 \;\Rightarrow\; \exists i_1 \in V,\;\; z_{i_1 i_1} = \max\{z_{ij} \mid i, j \in V\}$.
Let $\mathcal{I}_1 = \{j \in V \mid z_{i_1 j} > 0\}$, then $\forall i, j \in \mathcal{I}_1,\; z_{ij} = 1/|\mathcal{I}_1|$.
By using an appropriate permutation matrix P, we obtain

$$P^\top Z P = \begin{pmatrix} Z_{\mathcal{I}_1} & O \\ O & Z_{\bar{\mathcal{I}}_1} \end{pmatrix}, \quad \text{where } \bar{\mathcal{I}}_1 = V \setminus \mathcal{I}_1.$$

The sub-matrix $Z_{\bar{\mathcal{I}}_1}$ satisfies that $Z_{\bar{\mathcal{I}}_1}\boldsymbol{e} = \boldsymbol{e},\; Z_{\bar{\mathcal{I}}_1}^2 = Z,\; Z_{\bar{\mathcal{I}}_1} \in \mathcal{N}$.
Repeating the process described above, we can convert Z to a block diagonal matrix $P^\top Z P = \mathrm{Diag}(Z_{\bar{\mathcal{I}}_1}, \ldots, Z_{\bar{\mathcal{I}}_t})$.
We construct a matrix $X = (x_{ip})$ such that

$$x_{ip} = \begin{cases} 1 & (i \in \mathcal{I}_p) \\ 0 & (i \notin \mathcal{I}_p) \end{cases},$$



then X is feasible for (P) and $Z = X(X^\top X)^{-1} X^\top$.

0-1 semidefinite programming (0-1SDP)

0-1SDP formulation

$$
(\bar{\mathsf{P}}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A - D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z^2 = Z \\ & Z \in \mathcal{N}_n. \end{array} \right.
$$

- Laplacian: $L = D - A \in \mathcal{S}_n^+$
- $Z^2 = Z$
  $\Rightarrow \forall i,\ \lambda_i \in \{0, 1\}$.
  (λ_i : eigenvalue of Z)

- the objective function is linear with respect to Z
- the idempotence constraint makes the problem difficult
  $\Rightarrow$ relax the constraint $Z^2 = Z$ to a more tractable constraint

$$
D - 2A = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix} \notin \mathcal{S}_3^+
$$

0-1 semidefinite programming (0-1SDP)

0-1SDP formulation

$$(\bar{\mathsf{P}}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A-D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z^2 = Z \\ & Z \in \mathcal{N}_n. \end{array} \right.$$

- Laplacian: $L = D - A \in \mathcal{S}_n^+$
- $Z^2 = Z$ $\Rightarrow \forall i,\ \lambda_i \in \{0, 1\}$. (λ_i : eigenvalue of Z)

- the objective function is linear with respect to Z
- the idempotence constraint makes the problem difficult $\Rightarrow$ relax the constraint $Z^2 = Z$ to a more tractable constraint

$$D-2A = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix} \notin \mathcal{S}_3^+$$

Doubly Non-Negative relaxation

- DNN relaxation
  $\Rightarrow$ provides a tight bound for combinatorial optimization problems

0-1SDP formulation

$$(\bar{\mathrm{P}}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A-D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z^2 = Z \\ & Z \in \mathcal{N}_n. \end{array} \right.$$

DNN relaxation

$$(\mathrm{DNN}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A-D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z \in \mathcal{S}_n^+ \cap \mathcal{N}_n. \end{array} \right.$$

- the interior-point method solves the problem over a symmetric cone efficiently
- we cannot directly apply the interior-point method to solve (DNN) since doubly non-negative cone is not symmetric

$$Z \in \mathcal{S}_n^+ \cap \mathcal{N}_n \Leftrightarrow \begin{pmatrix} Z & O \\ O & \mathrm{Diag}(\mathrm{vec}(Z)) \end{pmatrix} \in \mathcal{S}_{n+n^2}^+$$

Valid inequality

Lemma 2.

The following inequalities are valid for $(\bar{\mathsf{P}})$

$$z_{ii} \geq z_{ij} \quad (i, j \in V).$$

DNN relaxation

$$(\mathrm{DNN}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A - D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & Z \in \mathcal{S}_n^+ \cap \mathcal{N}_n. \end{array} \right.$$

DNN with valid inequalities

$$(\overline{\mathrm{DNN}}) \quad \left| \begin{array}{ll} \text{max.} & \mathrm{Tr}((2A - D)Z) \\ \text{s.t.} & Z\boldsymbol{e}_n = \boldsymbol{e}_n \\ & z_{ii} \geq z_{ij} \ (i, j \in V) \\ & Z \in \mathcal{S}_n^+ \cap \mathcal{N}_n. \end{array} \right.$$

Permutation based on spectrum

Z^* : solution of the relaxation problem

- $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$: eigenvalues of Z^*
- $\boldsymbol{u}_i \in \mathbb{R}^n$: eigenvector corresponding to λ_i

Permuting the rows and columns of Z^* consistent with the decreasing order of values of elements of $\boldsymbol{u}_2$, we have

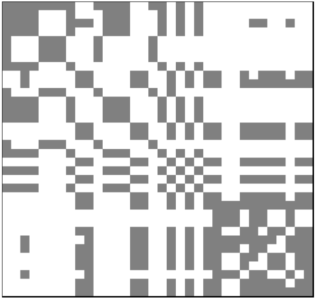


Figure: Original matrix

Permutation based on spectrum

Z^* : solution of the relaxation problem

- $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$: eigenvalues of Z^*
- $\boldsymbol{u}_i \in \mathbb{R}^n$: eigenvector corresponding to λ_i

Permuting the rows and columns of Z^* consistent with the decreasing order of values of elements of $\boldsymbol{u}_2$, we have

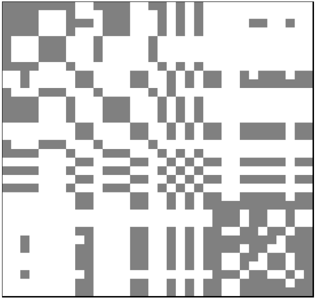


Figure: Original matrix

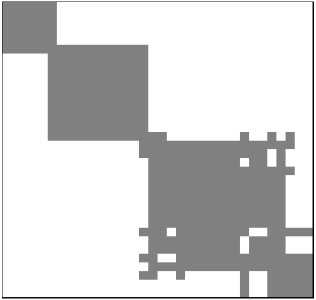


Figure: Permuted matrix

$\overline{V}$: sequence consistent with the decreasing order of $\boldsymbol{u}_2$
(we write $\overline{V} = [1 \ldots n]$ for the sake of simplicity)

- $q(k, \ell) = \dfrac{2 \sum_{i=k}^{\ell} \sum_{j=k}^{\ell} a_{ij} - \sum_{i=k}^{\ell} d_i}{\ell - (k-1)}$ for k, ℓ of $\overline{V}$ with $k \leq \ell$
- $\mu(s)$: the maximum value that is achieved by partitioning of [1. . . s] into several consecutive subsequences
(assume $\mu(0) = 0$ for notational convenience)

Recursive equation

$$\mu(s) = \max\{\, \mu(k) + q(k+1, s) \mid k \in \{\, 0, 1, \ldots, s-1 \,\}\,\}.$$

$\Rightarrow \quad \mu(1) = q(1,1)$
$\mu(2) = \max\{\, q(1,2),\ \mu(1) + q(2,2) \,\}$
$\mu(3) = \max\{\, q(1,3),\ \mu(1) + q(2,3),\ \mu(2) + q(3,3) \,\} \cdots$

$\overline{V}$: sequence consistent with the decreasing order of $\boldsymbol{u}_2$
(we write $\overline{V} = [1 \ldots n]$ for the sake of simplicity)

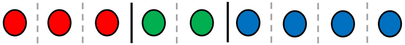
- $q(k,\ell) = \dfrac{2\sum_{i=k}^{\ell}\sum_{j=k}^{\ell} a_{ij} - \sum_{i=k}^{\ell} d_i}{\ell - (k-1)}$ for k, ℓ of $\overline{V}$ with $k \leq \ell$
- $\mu(s)$: the maximum value that is achieved by partitioning of [1. . . s] into several consecutive subsequences
(assume $\mu(0) = 0$ for notational convenience)

Recursive equation

$$\mu(s) = \max\{\, \mu(k) + q(k+1, s) \mid k \in \{\, 0, 1, \ldots, s-1 \,\}\,\}.$$

$$\Rightarrow \quad \begin{aligned} \mu(1) &= q(1,1) \\ \mu(2) &= \max\{\, q(1,2),\ \mu(1) + q(2,2) \,\} \\ \mu(3) &= \max\{\, q(1,3),\ \mu(1) + q(2,3),\ \mu(2) + q(3,3) \,\} \cdots \end{aligned}$$

$\overline{V}$: sequence consistent with the decreasing order of $\boldsymbol{u}_2$
(we write $\overline{V} = [1 \ldots n]$ for the sake of simplicity)



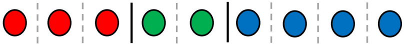
- $q(k, \ell) = \dfrac{2 \sum_{i=k}^{\ell} \sum_{j=k}^{\ell} a_{ij} - \sum_{i=k}^{\ell} d_i}{\ell - (k-1)}$ for k, ℓ of $\overline{V}$ with $k \leq \ell$
- $\mu(s)$: the maximum value that is achieved by partitioning of [1. . . s] into several consecutive subsequences
(assume $\mu(0) = 0$ for notational convenience)

Recursive equation

$$\mu(s) = \max\{\, \mu(k) + q(k+1, s) \mid k \in \{\, 0, 1, \ldots, s-1 \,\} \,\}.$$

$$\Rightarrow \quad \mu(1) = q(1,1)$$
$$\mu(2) = \max\{\, q(1,2),\ \mu(1) + q(2,2) \,\}$$
$$\mu(3) = \max\{\, q(1,3),\ \mu(1) + q(2,3),\ \mu(2) + q(3,3) \,\} \cdots$$

$\overline{V}$: sequence consistent with the decreasing order of $\boldsymbol{u}_2$
(we write $\overline{V} = [1 \dots n]$ for the sake of simplicity)



- $q(k,\ell) = \dfrac{2\sum_{i=k}^{\ell}\sum_{j=k}^{\ell} a_{ij} - \sum_{i=k}^{\ell} d_i}{\ell - (k-1)}$ for k, ℓ of $\overline{V}$ with $k \leq \ell$
- $\mu(s)$: the maximum value that is achieved by partitioning of [1. . . s] into several consecutive subsequences
  (assume $\mu(0) = 0$ for notational convenience)

Recursive equation

$$\mu(s) = \max\{\, \mu(k) + q(k+1, s) \mid k \in \{\, 0, 1, \dots, s-1 \,\}\,\}.$$

$$\Rightarrow \quad \mu(1) = q(1,1)$$
$$\mu(2) = \max\{\, q(1,2),\ \mu(1) + q(2,2) \,\}$$
$$\mu(3) = \max\{\, q(1,3),\ \mu(1) + q(2,3),\ \mu(2) + q(3,3) \,\} \cdots$$

Computational experiment

- Computational environment
  CPU : Intel Core i7 3.70 GHz
  Memory : 32.0 GB
  SDP Solver : SeDuMi 1.2
  MILP Solver : Gurobi 6.0.0
- The instances we tested :

ID	name	n	m	t	OPT
1	Strike	24	38	4	8.8611
2	Karate	34	78	3	7.8451
3	Mexico	35	117	3	8.7180
4	Sawmill	36	62	4	8.6233
5	Dolphins	62	159	5	12.1252[1]
6	Books	105	441	7	21.9652[1]

[1] the best lower bound reported in Costa et al. '15

Table: Comparison of obtained lower and upper bounds

	(DNN)		($\overline{\text{DNN}}$)		(MILP)	
ID	UB	LB	UB	LB	UB	LB
1	9.5808	8.8611	9.3049	8.8611	8.8611	8.8611
2	8.9548	7.8424	8.4141	7.8451	7.8451	7.8451
3	10.3151	8.5580	9.9570	8.5227	8.7180	8.7180
4	10.5048	7.0486	10.0311	7.3587	8.6223	8.6233
5	15.0218	9.8286	14.3552	11.4610	17.1252	12.1252
6	26.5387	20.2470	24.7749	20.3150	56.8739	21.0815

Table: Comparison of computational time in seconds

ID	(DNN)	($\overline{\text{DNN}}$)	(MILP)
1	1.05	3.54	0.50
2	5.83	36.04	0.74
3	7.64	43.48	7.84
4	7.75	54.21	6.10
5	316.61	1681.81	OT[2]
6	4626.11	60437.45	OT[2]

[2] more than 10,000 seconds

Conclusion

Conclusion

- We proved the equivalence between the modularity density maximization and 0-1SDP
- obtained a tight upper bound by DNN relaxation
- developed a heuristics to obtain a lower bound

However, there is no theoretical validity of using the second largest eigenvector. Here still remains room for further research.

Thank you for your attention.

Conic programming

- $\mathcal{K}$: a nonempty closed convex cone
- $\langle \cdot, \cdot \rangle$: an inner product
- $\mathcal{K}^*$: the dual cone of $\mathcal{K}$, i.e., $\mathcal{K}^* = \{ \boldsymbol{x} \mid \forall \boldsymbol{y} \in \mathcal{K}, \langle \boldsymbol{x}, \boldsymbol{y} \rangle \geq 0 \}$
- $A : \mathbb{R}^n \to \mathbb{R}^m$: a linear operator
- A^* : the adjoint operator of A, i.e., $\langle A\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, A^*\boldsymbol{y} \rangle$

Primal

$$\left| \begin{array}{ll} \text{min.} & \langle \boldsymbol{c}, \boldsymbol{x} \rangle \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathcal{K}. \end{array} \right.$$

Dual

$$\left| \begin{array}{ll} \text{max.} & \langle \boldsymbol{b}, \boldsymbol{y} \rangle \\ \text{s.t.} & \boldsymbol{c} - A^*\boldsymbol{y} \in \mathcal{K}^*. \end{array} \right.$$

Roughly speaking, $\mathcal{K}$ is called a symmetric cone if $\mathcal{K}^* = \mathcal{K}$.

- symmetric cones :
  non-negative orthant $\mathbb{R}^n_+$, semidefinite cone $\mathcal{S}^+_n$, second-order cone, etc.

Conic programming

- Copositive cone $\mathcal{C}_n = \{\, Y \in \mathbb{R}^{n \times n} \mid Y^\top = Y, \forall \boldsymbol{d} \in \mathbb{R}^n_+, \boldsymbol{d}^\top Y \boldsymbol{d} \geq 0 \,\}$
- Completely positive cone $\mathcal{C}_n^* = \text{conv}(\{\, \boldsymbol{y}\boldsymbol{y}^\top \mid \boldsymbol{y} \in \mathbb{R}^n_+ \,\})$
- Doubly non-negative cone $\mathcal{S}_n^+ \cap \mathcal{N}_n$

Properties

- $(\mathcal{S}_n^+ \cap \mathcal{N}_n)^* = \mathcal{S}_n^+ + \mathcal{N}_n$
- $\mathcal{C}_n^* \subseteq \mathcal{S}_n^+ \cap \mathcal{N}_n \subseteq \mathcal{S}_n^+ \subseteq \mathcal{S}_n^+ + \mathcal{N}_n \subseteq \mathcal{C}_n$
- $\mathcal{C}_n^* = \mathcal{S}_n^+ \cap \mathcal{N}_n \subseteq \mathcal{S}_n^+ \subseteq \mathcal{S}_n^+ + \mathcal{N}_n = \mathcal{C}_n$ for $n \leq 4$

Strong results on $\mathcal{C}_n$, $\mathcal{C}_n^*$

- the maximum clique number:

$$\min\{\, \alpha \in \mathbb{N} \mid \alpha(E - A) - E \in \mathcal{C}_n \,\} \text{ where } E = \boldsymbol{e}\boldsymbol{e}^\top.$$

- non-convex quadratic programming:

$$\min\{\, \mathsf{Tr}(QX) \mid \mathsf{Tr}(EX) = 1,\ X \in \mathcal{C}_n^* \,\} \text{ where } E = \boldsymbol{e}\boldsymbol{e}^\top.$$