Robust convex relaxation for the clique and densest k-subgraph problems

Brendan Ames

Institute for Mathematics and its Applications
University of Minnesota

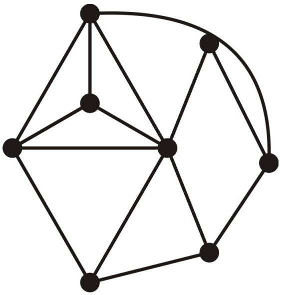
IMA Postdoc Seminar
Tuesday, May 22, 2012

Cliques of a graph

- Given graph $G = (V, E)$, a clique of G is a pairwise adjacent subset of V.
- $C \subseteq V$ is a clique of G if $uv \in E$ for all $u, v \in C$.
- The subgraph $G(C)$ induced by C is complete.

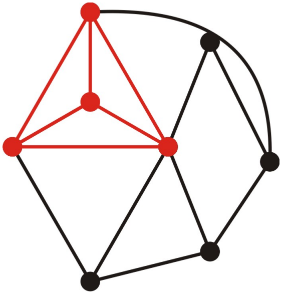
Cliques of a graph

- Given graph $G = (V, E)$, a clique of G is a pairwise adjacent subset of V.
- $C \subseteq V$ is a clique of G if $uv \in E$ for all $u, v \in C$.
- The subgraph $G(C)$ induced by C is complete.



Cliques of a graph

- Given graph $G = (V, E)$, a clique of G is a pairwise adjacent subset of V.
- $C \subseteq V$ is a clique of G if $uv \in E$ for all $u, v \in C$.
- The subgraph $G(C)$ induced by C is complete.



The Maximum Clique problem

- Optimization version: Find the clique of G of maximum size. Size of the largest clique is the clique number $\omega(G)$.
- Decision version: Given graph G, integer k: does G contain a clique of cardinality at least k.
- Complexity: NP-complete, cannot approximate within a ratio of $N^{1-\epsilon}$ for any $\epsilon > 0$.
- Many applications: communication, biological, and social networks.

The planted case

- Hardness results are worst case.
- There should be instances we should be able to solve efficiently.
- In particular, if G has a clique of size k, we should be able to find it if k is large.

The planted case

- Hardness results are worst case.
- There should be instances we should be able to solve efficiently.
- In particular, if G has a clique of size k, we should be able to find it if k is large.
- Alon et al. 1998, Feige and Krauthgamer 2000: if $k \geq \Omega(\sqrt{N})$ and all other edges are added independently at random then we can find the maximum clique in polynomial time.

A more general model?

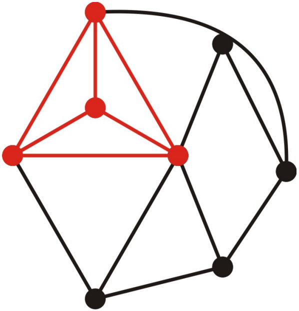
- These recovery guarantees rely heavily on the fact that G is an undirected graph.
  - symmetry of A_G, the fact that a stable set of $\overline{G}$ is a clique of G.
- Ideally, would like an approach that translates to finding other “clique-like” objects with minimal effort.
- for example: the maximum biclique of a bipartite graph

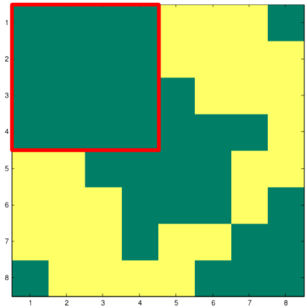
Cliques and low-rank matrices

- Every clique C (with characteristic vector $\mathbf{v}$) of the graph $G = (V, E)$ defines a rank-one matrix by $X = \mathbf{v}\mathbf{v}^T$.
- Moreover, nonzero entries of X form a $|C| \times |C|$ rank-one block in $A_G + I$.

Cliques and low-rank matrices

- Every clique C (with characteristic vector $\mathbf{v}$) of the graph $G = (V, E)$ defines a rank-one matrix by $X = \mathbf{v}\mathbf{v}^T$.
- Moreover, nonzero entries of X form a $|C| \times |C|$ rank-one block in $A_G + I$.





Clique as rank minimization

- G has a clique of cardinality at least k if and only if there exists rank-one symmetric binary matrix X such that

$$\sum \sum X_{ij} \geq k^2$$

$$X_{ij} = 0 \quad \forall\, ij \notin E,\ i \neq j.$$

- Otherwise $\omega(G) < k$.

Clique as rank minimization

- G has a clique of cardinality at least k if and only if there exists rank-one symmetric binary matrix X such that

$$\sum \sum X_{ij} \geq k^2$$

$$X_{ij} = 0 \quad \forall\, ij \notin E,\ i \neq j.$$

- Otherwise $\omega(G) < k$.

- Therefore clique is equivalent to the rank minimization problem:

$$\min_{\substack{X \in \{0,1\}^{V \times V} \\ X \in \Sigma^V}} \left\{ rank(X) : \mathbf{e}^T X \mathbf{e} \geq k^2, X_{ij} = 0 \text{ if } (i,j) \in \tilde{E} \right\}$$

where $\tilde{E} = V \times V - \left\{ E \cup \{(u,u) : u \in V\} \right\}$.

Rank minimization

- Affine rank minimization problem: find matrix with minimum rank satisfying linear constraints:

$$\min\{rank(X) : \mathcal{A}(X) = \mathbf{b}\}.$$

  Well-known to be NP-hard.

- Relax $rank(X)$ with nuclear norm $\|X\|_*$:

$$rank(X) = \|\sigma(X)\|_0, \quad \|X\|_* = \|\sigma(X)\|_1.$$

- If $\mathcal{A}$ satisfies certain “niceness” conditions then the minimum nuclear norm solution is the minimum rank solution.

Nuclear norm relaxation of Clique

- We solve the convex relaxation

$$\min\left\{\|X\|_* : \mathbf{e}^T X \mathbf{e} \geq k^2, X_{ij} = 0 \text{ if } (i,j) \in \tilde{E}\right\} \quad (\mathbf{NNR})$$

- Does the linear operator defining the constraints satisfy RIP/incoherence/null space conditions?

Failure of RIP

- Consider the constraints

$$X_{ij} = 0 \text{ if } (i, j) \in \tilde{E}.$$

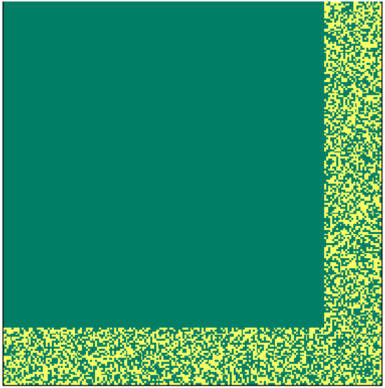
- Only sample information from entries corresponding to nonadjacent nodes: not evenly distributed among $V \times V$!

Failure of RIP

- Consider the constraints

$$X_{ij} = 0 \text{ if } (i,j) \in \tilde{E}.$$

- Only sample information from entries corresponding to nonadjacent nodes: not evenly distributed among $V \times V$!

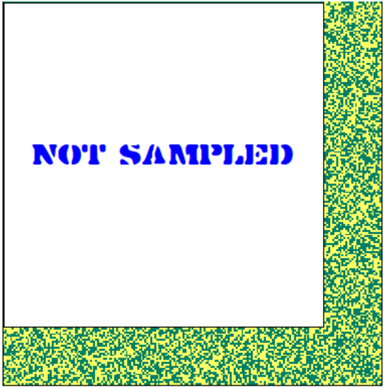


Failure of RIP

- Consider the constraints

$$X_{ij} = 0 \text{ if } (i, j) \in \tilde{E}.$$

- Only sample information from entries corresponding to nonadjacent nodes: not evenly distributed among $V \times V$!



Why the relaxation works

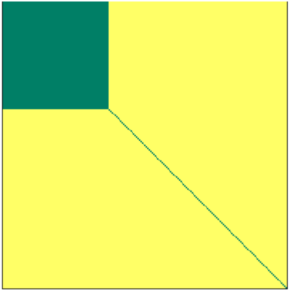
- Low-rank matrix completion fails when the matrix to be recovered lies in the null space of the sampling operator.
- In this case: can't distinguish from the all zero matrix 0.
- For our problem: solutions are bounded away from 0 by the constraint $\mathbf{e}^T X \mathbf{e} \geq k^2$.
- We want the sum constraint to be satisfied using only the clique entries, and all other entries can be set to 0.

The planted case

- Construction:
  - Add all potential edges between nodes in vertex set V^* of size k.
  - Then some of the remaining potential edges are added as noise: either at random or deterministically by an adversary.
  - By construction, V^* is a clique of G (called a planted or hidden clique).

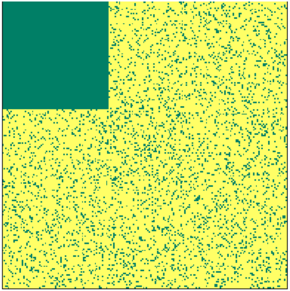
The planted case

- Construction:
  - Add all potential edges between nodes in vertex set V^* of size k.
  - Then some of the remaining potential edges are added as noise: either at random or deterministically by an adversary.
  - By construction, V^* is a clique of G (called a planted or hidden clique).



The planted case

- Construction:
  - Add all potential edges between nodes in vertex set V^* of size k.
  - Then some of the remaining potential edges are added as noise: either at random or deterministically by an adversary.
  - By construction, V^* is a clique of G (called a planted or hidden clique).



Recovery guarantee (Random case)

Theorem

- *Suppose that noise edges are added independently with fixed probability p.*
- *There exists scalar $c > 0$ such that if*

$$k \geq c\sqrt{N}$$

*then V^* is the unique maximum clique of G and $X^* = \mathbf{v}\mathbf{v}^T$ is the unique optimal solution of (**NNR**) with probability tending exponentially to 1 as $N \to \infty$.*

Recovery guarantee (Adversarial case)

Theorem

Suppose that the adversary is allowed to add

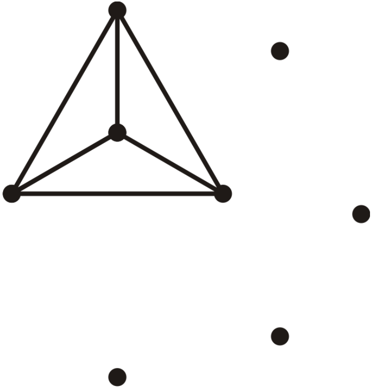
(1) *at most* δk *edges from any node in* $V - V^*$ *to* V^* *for some* $\delta \in [0, 1)$*, and*

(2) *at most* ck^2 *edges total for some scalar* $c > 0$ *depending only on* δ*.*

Then V^* *is the unique maximum clique of* G *and* X^* *is the unique optimal solution of* (**NNR**).

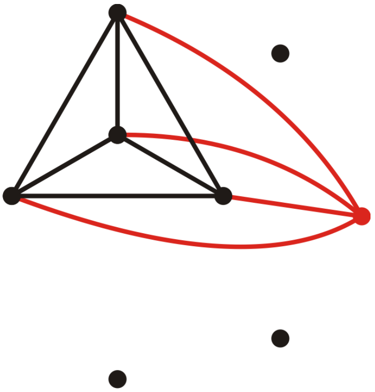
Optimality of the adversarial bounds

- Suppose that adversary can add k edges from $v \in V - V^*$ to V^*.



Optimality of the adversarial bounds

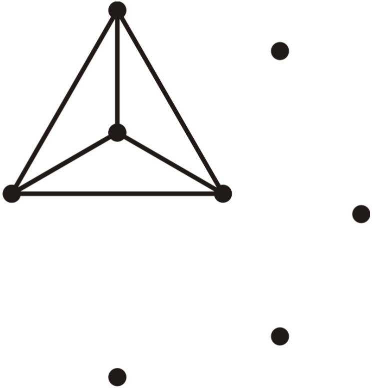
- Suppose that adversary can add k edges from $v \in V - V^*$ to V^*.
- Then can expand the planted clique V^* to $V^* \cup \{v\}$.



Optimality of the adversarial bounds

- Suppose that adversary can add k edges from $v \in V - V^*$ to V^*.
- Then can expand the planted clique V^* to $V^* \cup \{v\}$.

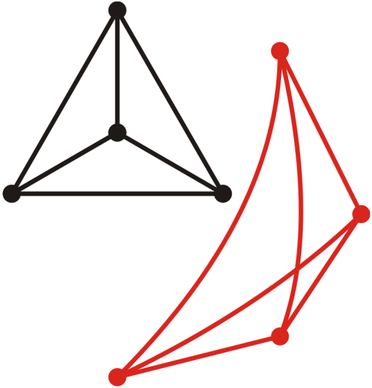
- Suppose the adversary can add $k(k-1)/2$ edges.



Optimality of the adversarial bounds

- Suppose that adversary can add k edges from $v \in V - V^*$ to V^*.
- Then can expand the planted clique V^* to $V^* \cup \{v\}$.

- Suppose the adversary can add $k(k-1)/2$ edges.
- Can make a new clique of size k.



Proof Idea

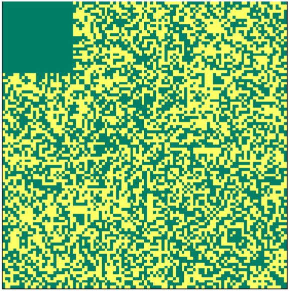
- We will apply KKT conditions and SDP duality to derive conditions ensuring optimality and uniqueness of X^*.
- Propose a choice of multipliers corresponding to X^*.
- Use matrix norm bounds to establish that these multipliers satisfy the optimality and uniqueness conditions (with high probability).

What happens if edges are deleted?

- These guaranteees do not tolerate edge deletion noise.
- Suppose the graph is corrupted so that edge uv is deleted for some $u, v \in V^*$.
- Then V^* is not a clique and X^* is not feasible for (**NNR**).

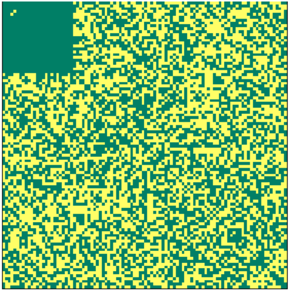
What happens if edges are deleted?

- These guaranteees do not tolerate edge deletion noise.
- Suppose the graph is corrupted so that edge uv is deleted for some $u, v \in V^*$.
- Then V^* is not a clique and X^* is not feasible for (**NNR**).



What happens if edges are deleted?

- These guaranteees do not tolerate edge deletion noise.
- Suppose the graph is corrupted so that edge uv is deleted for some $u, v \in V^*$.
- Then V^* is not a clique and X^* is not feasible for (**NNR**).



The densest k-subgraph problem

- Want to find a dense subgraph of size k, not necessarily a clique.
- Densest k-subgraph problem (DKS): Given a graph G, find subgraph $H \subseteq G$ on k nodes with maximum density:

$$d(H) = \frac{|E(H)|}{|V(H)|} = \frac{|E(H)|}{k}.$$

- NP-hard: proof is by reduction to Clique; hard to approximate
- Maximizing $d(H)$ is equivalent to maximizing $|E(H)|$ over all k-node subgraphs.

Duality of density and number of missing edges

- Let $V^* \subseteq V$ be a k-subset with characteristic vector $\mathbf{v}$.
- Introduce a new variable Y: acts as a correction for entries of $X = \mathbf{v}\mathbf{v}^T$ that should be 0:

$$Y_{ij} = \begin{cases} -X_{ij}, & \text{if } ij \in \tilde{E} \\ 0, & \text{otherwise.} \end{cases}$$

Duality of density and number of missing edges

- Let $V^* \subseteq V$ be a k-subset with characteristic vector $\mathbf{v}$.
- Introduce a new variable Y: acts as a correction for entries of $X = \mathbf{v}\mathbf{v}^T$ that should be 0:

$$Y_{ij} = \begin{cases} -X_{ij}, & \text{if } ij \in \tilde{E} \\ 0, & \text{otherwise.} \end{cases}$$

- If V^* is almost a clique then $G(V^*)$ should be very dense and Y should be very sparse.
- Cardinality of Y acts as a dual of density of $G(V^*)$:

$$|E(G(V^*))| = \binom{k}{2} - \frac{\|Y^*\|_0}{2}$$

Formulation as sparse plus low-rank decomposition

- Can formulate (**DKS**) as

$$\begin{aligned} \mathit{min} \quad & \mathit{rank}(X) + \gamma \|Y\|_0 \\ \mathit{st} \quad & \mathbf{e}^T X \mathbf{e} = k^2 \\ & X_{ij} + Y_{ij} = 0 \text{ if } ij \in \tilde{E} \\ & X \in \{0,1\}^{V \times V} \\ & X \in \Sigma^V \end{aligned}$$

where γ is a regularization parameter.

Formulation as sparse plus low-rank decomposition

- Can formulate (**DKS**) as

$$
\begin{aligned}
\mathit{min} \quad & \|X\|_* + \gamma \|Y\|_1 \\
\mathit{st} \quad & \mathbf{e}^T X \mathbf{e} = k^2 \\
& X_{ij} + Y_{ij} = 0 \text{ if } ij \in \tilde{E} \\
& X \in [0,1]^{V \times V}
\end{aligned}
$$

  where γ is a regularization parameter.
- Relax $\|Y\|_0$ using the ℓ_1-norm $\|Y\|_1$, $\mathit{rank}(X)$ with the nuclear norm $\|X\|_*$

Robust PCA

- Chandrasekaran et al 2009, Candès et al 2009, Doan and Vavasis 2010, Chen et al. 2011, Oymak and Hassibi 2011: under certain assumptions on the support of S, column/row space of L, γ and the set of observed entries Ω, we have

$$\begin{aligned}(L^*, S^*) &= argmin\{rank(L) + \gamma\|S\|_0 : P_\Omega(L+S) = P_\Omega(M)\}\\ &= argmin\{\|L\|_* + \gamma\|S\|_1 : P_\Omega(L+S) = P_\Omega(M)\}.\end{aligned}$$

Robust PCA

- Chandrasekaran et al 2009, Candès et al 2009, Doan and Vavasis 2010, Chen et al. 2011, Oymak and Hassibi 2011: under certain assumptions on the support of S, column/row space of L, γ and the set of observed entries Ω, we have

$$\begin{aligned}(L^*, S^*) &= argmin\{rank(L) + \gamma\|S\|_0 : P_\Omega(L+S) = P_\Omega(M)\} \\ &= argmin\{\|L\|_* + \gamma\|S\|_1 : P_\Omega(L+S) = P_\Omega(M)\}.\end{aligned}$$

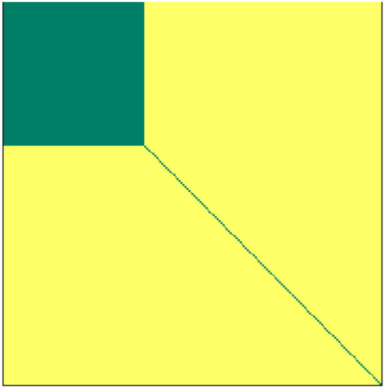
- Existing guarantees do not apply for our problem:
  - Sampled entries are far more common outside of the planted clique.
  - Have nonsampling linear constraints.

Planted case

- Start with N nodes V.
- Add all edges between nodes in $V^* \subseteq V$.
- Add noise:
  - Add some of the remaining potential edges.
  - Delete some edges in $V^* \times V^*$.

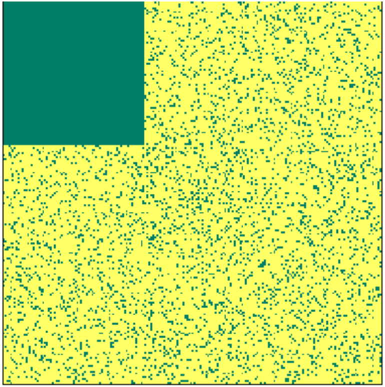
Planted case

- Start with N nodes V.
- Add all edges between nodes in $V^* \subseteq V$.
- Add noise:
  - Add some of the remaining potential edges.
  - Delete some edges in $V^* \times V^*$.



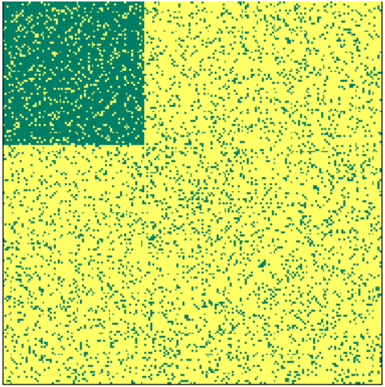
Planted case

- Start with N nodes V.
- Add all edges between nodes in $V^* \subseteq V$.
- Add noise:
  - Add some of the remaining potential edges.
  - Delete some edges in $V^* \times V^*$.



Planted case

- Start with N nodes V.
- Add all edges between nodes in $V^* \subseteq V$.
- Add noise:
  - Add some of the remaining potential edges.
  - Delete some edges in $V^* \times V^*$.



Adversarial case guarantee

Theorem

- *Suppose the adversary is allowed to add and delete edges so that at most*
  - $\delta_1 k$ *edges incident at* v *are deleted for all* $v \in V^*$,
  - $\delta_2 k$ *edges added from any* $v \in V - V^*$ *to* V^*,
  - $O(k^2)$ *edges deleted in* $V^* \times V^*$ *in total, and*
  - $O(k^2)$ *edges are added in total,*

  for some $\delta_1, \delta_2 \in (0, 1/2)$.

- *Then* $G(V^*)$ *is the unique maximum density* k*-subgraph of* G *and* $(X^*, Y^*) = (\mathbf{v}\mathbf{v}^T, -P_{\tilde{E}}(\mathbf{v}\mathbf{v}^T))$ *is the unique optimal solution of* $(\mathbf{DKSR})$ *with regularization parameter*

$$\gamma = \frac{1}{(1 - 2\delta_1)k}.$$

Random case guarantee

Theorem

- *Suppose that each edge in $V^* \times V^*$ is deleted independently with fixed probability q and each potential edge is added independently with probability p such that*

$$p + q < 1.$$

- *Then there exist scalars $c_1, c_2 > 0$ depending on p, q such that if*

$$k \geq c_1 \sqrt{N}$$

*then $G(V^*)$ is the unique maximum density k-subgraph of G and (X^*, Y^*) is the unique optimal solution of (**DKSR**) for regularization parameter $\gamma = c_2/k$ with probability tending exponentially to 1 as $N \to \infty$.*

Conclusion

- Proposed new heuristics for the Clique and Densest k-subgraph problems.
- Established theoretical guarantees for exact recovery.
- Open problems:
  - How to efficiently solve the relaxations?
  - Are the random bounds tight? Can we relax $\Omega(N^{1/2})$ to $\Omega(N^{1/2-\epsilon})$?
- References:

  - B. Ames and S. Vavasis. Nuclear norm minimization for the planted clique and biclique problems. *Mathematical Programming*, 129(1):121, 2011.
  - B. Ames. Convex relaxation for the densest k-subgraph problem. *In preparation.* Prepreprint available from: ima.umn.edu/∼bpames/hidden/DKS_notes.pdf