Nuclear Norm Minimization for the Planted Clique and Biclique Problems

Brendan P.W. Ames Stephen A. Vavasis

Department of Combinatorics and Optimization
University of Waterloo

Graduate Student Research Conference
Thursday, April 30, 2009

Outline

Outline

Outline

Affine Rank Minimization

- **Affine rank minimization** refers to optimization problems of the form

$$\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s.t.} & \mathcal{A}(X) = b \end{array}$$

  for linear $\mathcal{A} : \mathbf{R}^{m \times n} \to \mathbf{R}^p$, $b \in \mathbf{R}^p$.
- Known to be **NP-hard**.

Nuclear Norm Relaxation

- Heuristic suggested by **Recht-Fazel-Parrillo (2007)**.
- Replace rank (X) with $\|X\|_* = \sigma_1(X) + \sigma_2(X) + \cdots + \sigma_{\min m,n}(X)$.
- Obtain the convex optimization problem

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & \mathcal{A}(X) = b. \end{array}$$

- For every integer $1 \leq r \leq m$, the **r-restricted isometry constant** is the smallest number $\delta_r(\mathcal{A})$ such that

$$(1 - \delta_r(\mathcal{A}))\|X\|_F \leq \|\mathcal{A}(X)\| \leq (1 + \delta_r(A))\|X\|_F.$$

- **RFP:**
  - Suppose rank $(X_0) = r$, $\delta_{5r}(\mathcal{A}) < 1/10$ and $b = \mathcal{A}(X_0)$.
  - Then the unique optimal solution to the nuclear norm relaxation is X_0.

- For every integer $1 \leq r \leq m$, the **r-restricted isometry constant** is the smallest number $\delta_r(\mathcal{A})$ such that

$$(1 - \delta_r(\mathcal{A}))\|X\|_F \leq \|\mathcal{A}(X)\| \leq (1 + \delta_r(A))\|X\|_F.$$

- **RFP:**
  - Suppose $\operatorname{rank}(X_0) = r$, $\delta_{5r}(\mathcal{A}) < 1/10$ and $b = \mathcal{A}(X_0)$.
  - Then the unique optimal solution to the nuclear norm relaxation is X_0.
- **Candès and Recht (2008):**
  - Suppose X_0 is sampled from the random orthogonal model such that $\operatorname{rank}(X_0) = r$, $\mathcal{A}$ samples m entries of X_0 uniformly at random and $b = \mathcal{A}(X_0)$.
  - Then X_0 is the unique optimal solution of the nuclear norm relaxation if $m = O(n^{5/4}\log(n))$.

Compressive Sensing

- Rank minimization can be thought of as a generalization of **vector cardinality minimization**:

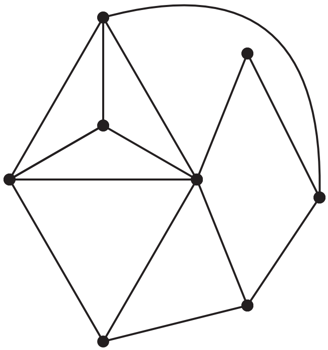
$$\begin{array}{ll} \min & \#\text{ nonzero entries of } x \\ \text{s.t.} & Ax = b \end{array}$$

  for $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$.
- Problem is NP-hard.
- Can be relaxed using the l_1-norm: $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$.
- For certain classes of random A and $b = Ax_0$, then the l_1-relaxation recovers exact solution x_0 to the original problem.

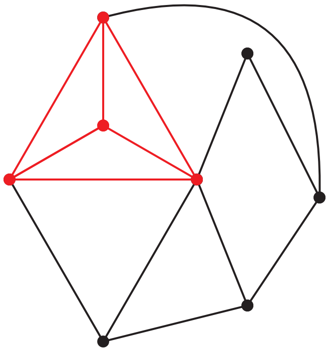
Maximum clique problem

- Given graph $G = (V, E)$, want to find the largest complete subgraph (clique) of G.
- One of the original NP-hard problems.
- Can be used as a simple model for data mining problems.

Example



Example



Formulation as affine rank minimization

- The adjacency matrix of the graph $H' = H \cup \{vv : v \in V(H)\}$ has rank equal to one if and only if H is a complete graph.
- Thus, a clique K of G containing n vertices can be found by solving

$$\begin{aligned} \min \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & \sum_{i \in V} \sum_{j \in V} X_{ij} \geq n^2, \\ & X_{ij} = 0 \text{ if } (i,j) \notin E \text{ and } i \neq j, \\ & X \in [0,1]^{V \times V}. \end{aligned}$$

- Still NP-hard.

Nuclear norm relaxation of max clique

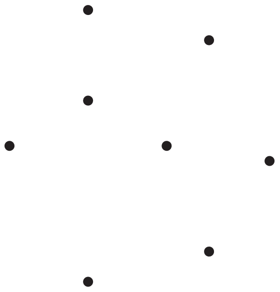
- We consider the convex relaxation

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & \displaystyle\sum_{i \in V} \sum_{j \in V} X_{ij} \geq n^2, \\ & X_{ij} = 0 \text{ if } (i,j) \notin E \text{ and } i \neq j. \end{array}$$

- **Our result:** If G consists of a clique on n vertices plus some diversionary edges, then this convex relaxation recovers the maximum clique of G for sufficiently large n.

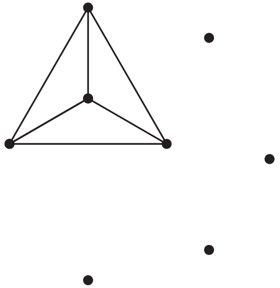
The Adversarial Case

- Suppose $G = (V, E)$ is generated as follows:



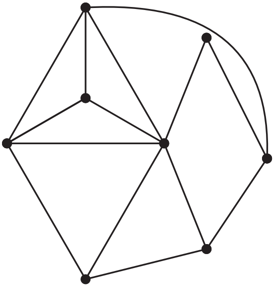
The Adversarial Case

- Suppose $G = (V, E)$ is generated as follows:
  - For $V^* \subset V$, for all $i, j \in V^*$ add (i, j) to E.



The Adversarial Case

- Suppose $G = (V, E)$ is generated as follows:
  - For $V^* \subset V$, for all $i, j \in V^*$ add (i, j) to E.
  - An adversary chooses a number of the remaining potential edges to be added to the graph.

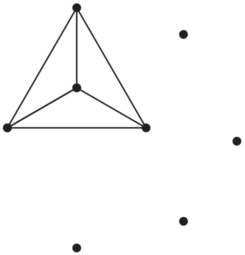


Theorem

- *Suppose that every vertex in* $V - V^*$ *is adjacent to at most* δn *vertices in* V^* *for* $0 < \delta < 1$.
- *Then* G *can contain at most* $O(n^2)$ *edges other than those in* $V^* \times V^*$ *and* K_{V^*} *will still be the maximum clique of* G.

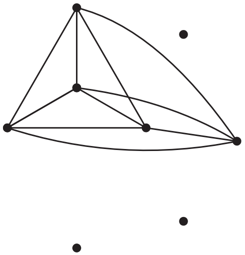
Example

- Suppose an adversary can add n edges from $v \in V - V^*$ to K_{V^*}.



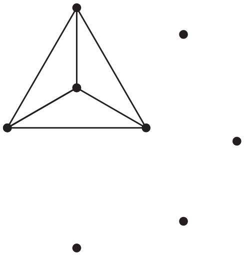
Example

- Suppose an adversary can add n edges from $v \in V - V^*$ to K_{V^*}.
- Then $K_{V^*} \cup \{v\}$ is a bigger clique.



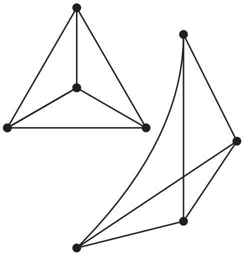
Example

- Suppose adversary can add n^2 edges not in the clique.



Example

- Suppose adversary can add n^2 edges not in the clique.
- Can form an additional clique of size n^2.



The Random Case

- Form $G = (V, E)$ as follows:
  - For $V^* \subset V$, for all $i, j \in V^*$ add (i, j) to E.
  - Add each remaining potential edge to the graph independently with probability p.

The Random Case

- Form $G = (V, E)$ as follows:
  - For $V^* \subset V$, for all $i, j \in V^*$ add (i, j) to E.
  - Add each remaining potential edge to the graph independently with probability p.

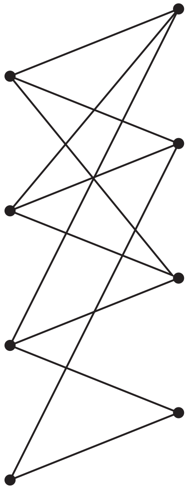
Theorem

- *Suppose $n > \alpha\sqrt{|V|}$ for sufficiently large $\alpha > 0$.*
- *Then K_{V^*} is the unique maximum clique of G and X^* is the optimal solution of the nuclear norm relaxation with probability exponentially close to 1.*

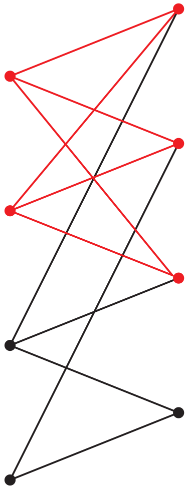
The Maximum Edge Biclique Problem

- Given a bipartite graph $G = (U, V, E)$, the **maximum edge biclique problem** is to find a complete bipartite subgraph of G with maximum number of edges.
- **Peeters (2003):** MEBP is NP-hard.

Example



Example



Nuclear norm relaxation of MEBP

- A biclique G containing mn edges can be found by solving

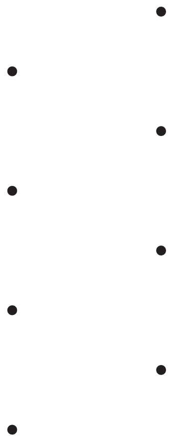
$$\begin{aligned} \min \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & \sum_{i \in U} \sum_{j \in V} X_{ij} \geq mn, \\ & X_{ij} = 0 \text{ if } (i,j) \notin E \\ & X \in [0,1]^{U \times V}. \end{aligned}$$

- We consider the convex relaxation

$$\begin{aligned} \min \quad & \|X\|_* \\ \text{s.t.} \quad & \sum_{i \in U} \sum_{j \in V} X_{ij} \geq mn, \\ & X_{ij} = 0 \text{ if } (i,j) \notin E. \end{aligned}$$

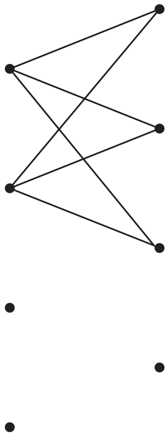
The Adversarial Case

- Suppose bipartite $G = (U, V, E)$ is generated as follows:



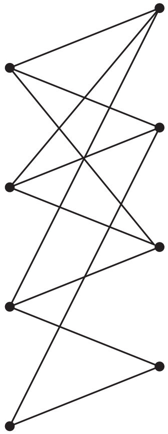
The Adversarial Case

- Suppose bipartite $G = (U, V, E)$ is generated as follows:
  - First add a biclique $U^* \times V^*$ with $U^* \subseteq U$, $V^* \subseteq V$. Let $|U^*| = m$, $|V^*| = n$.



The Adversarial Case

- Suppose bipartite $G = (U, V, E)$ is generated as follows:
  - First add a biclique $U^* \times V^*$ with $U^* \subseteq U$, $V^* \subseteq V$. Let $|U^*| = m$, $|V^*| = n$.
  - An adversary chooses a number of the remaining potential edges to be added to the graph.



Theorem

Suppose that

- *G contains at most r edges not in $U^* \times V^*$,*
- *each vertex of $V - V^*$ is adjacent to at most αm vertices of U^* for $\alpha \in (0, 1)$, and*
- *each vertex of $U - U^*$ is adjacent to at most βn vertices of V^* for $\beta \in (0, 1)$.*

Then there exists constant $c \in (0, 1)$ depending on α, β such that $U^* \times V^*$ is the unique maximum edge biclique of G if $r < cmn$.

- As with the maximum clique problem, this bound is the best possible.

The Random Case

- Suppose bipartite $G = (U, V, E)$ such that $|V| = N$, $|U| = \lceil yN \rceil$ for some $y > 0$ is generated as follows:
    - First add biclique $U^* \times V^*$ with $|V^*| = n$, $|U^*| = \lceil zn \rceil$ for some $z > 0$.
    - Add each remaining potential edge to the graph with probability p (independently).

The Random Case

- Suppose bipartite $G = (U, V, E)$ such that $|V| = N$, $|U| = \lceil yN \rceil$ for some $y > 0$ is generated as follows:
  - First add biclique $U^* \times V^*$ with $|V^*| = n$, $|U^*| = \lceil zn \rceil$ for some $z > 0$.
  - Add each remaining potential edge to the graph with probability p (independently).

Theorem

If $n = \Omega(\sqrt{N})$ then $U^* \times V^*$ is the unique maximum edge biclique of G with probability tending exponentially to 1 as $N \to \infty$ and $X^* = \bar{u}\bar{v}^T$ is the unique solution to the nuclear norm relaxation.

Experimental Results

- For each $n = 1, 2, \ldots, N = 25$, and $G^* \in \mathbf{R}^{N \times N}$ defined by

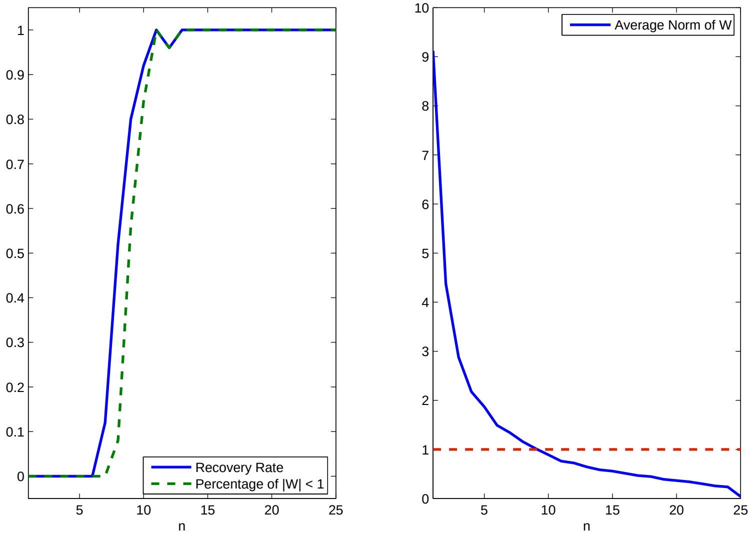
$$G^*_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\} \\ 0 & \text{otherwise} \end{cases}$$

  we repeated the following procedure 25 times:
  - An $N \times N$ matrix $G = G^* + D$ was generated where we take $D_{ij} = 0$ if $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\}$ and

$$D_{ij} = \begin{cases} 1 & \text{with probability } p = 1/2 \\ 0 & \text{with probability } p = 1/2 \end{cases}$$

    otherwise.
  - For each G, $X^* = \arg\min \left\{ \|X\|_* : \sum X_{ij} \geq n^2, \quad X_{ij} = 0 \text{ if } G_{ij} = 0 \right\}$ was obtained using the SDP solver SeDuMi.
  - We declared G^* to be recovered if $\|X^* - G^*\|_F / \|G^*\|_F < 10^{-3}$.

Experimental Results



Conclusions and Future Work

- Have shown that for two hard combinatorial problems, nuclear norm minimization returns the exact solution in poly-time in certain cases.
- Future work:
  - Apply the same technique to other hard problems (NMF, clustering).
  - Specialized algorithms.