Matrix Cuts

Based on the paper "Cones of matrices and setfunctions, and 0-1 optimization" by L.Lovász and A. Schrijver [1]

Srikanth Jagabathula

EECS, MIT
jskanth@mit.edu
October 21, 2008

The following writeup is simply my attempt at understanding the above mentioned paper. It is not my research; however, it might contain some simplifications or comments. This writeup is written as a part of the Graduate Seminar course offered by Prof. Devavrat Shah.

1 Overview.

This writeup discusses the idea of matrix cuts developed by Lovász and Schrijver to generate a sequence of relaxations of the constraint set polytope in an Integer Program. The idea will then applied to the stable set problem. This document is organized as follows:

1. Motivation, setup and basic notation.
2. Definition of the matrix cut.
3. Geometric intuition.
4. Main theorems related to the properties of the cuts.
5. Stable set problem.

2 Setup of the problem.

Consider an Integer Program (IP): optimization of a linear objective over integer constraints. Specifically, we assume that the variables are restricted to take $0, 1$ values. We know that solving this IP is equivalent to optimizing the linear objective over the convex hull $\mathscr{C}$ of the feasible solutions. In the "nicest" of cases, $\mathscr{C}$ has a polynomial number of facets and the IP can be solved in polynomial time. In more general cases, $\mathscr{C}$ has an exponentially many facets. A class of such problems can still be solved in polynomial time if the polytope $\mathscr{C}$ is "nice." There are different notions of "niceness"; the one that is most popular is the existence of a polynomial time separation algorithm.

There are many combinatorial problems that have such nice polytopes. So far, several *ad hoc* methods have been used, exploiting the combinatorial structure of the problem. However, there are two general approaches: *Gomory-Chvátal cuts* and *Projection representation*. In this writeup we will concentrate on the projection representation.

The method of projection representation is based on the observation that the projection of a polytope may have more facets than the polytope itself. More specifically, even if a polytope P has exponentially many facets, we may be able to find a lifting P' in a higher dimension such that P' has only a polynomial number of facets and P is its projection. Once such a P' has been constructed, it is clear that the linear objective can be optimized in polynomial time because optimizing over P is the same as optimizing over P' with the cost corresponding to the new variables set to zero.

Until now, this approach has been applied to several combinatorial optimization problems. But the methods used are not general and try to exploit the special structure of the problem. This paper provides a general method to determine such a lifting.

The approach developed in this paper combines the ideas of cuts and projections, resulting in a sequence of relaxations of the constraint set. The outline of the general procedure for the combinatorial optimization problems is as follows: Let P denote a polytope with integer vertices. We start with a polytope P_1 that contains P, contains no other integer points apart from the vertices of P, and is nice. Optimizing over P_1 provides an approximation. This polytope P_1 is then *cut*, resulting in a polytope P_2 such that P_2 is contained in P_1 and contains P. Continuing this procedure results in a sequence of cuts. This procedure should typically halt in a finite number of steps. In this writeup we discuss a method of cuts called matrix cuts.

Before we describe the details of matrix cuts, we will first make some definitions.

2.1 Notation and Definitions.

We define formally what we mean by a cut. For that let P be a polytope with $0,1$ vertices. A *relaxation* of P is denoted by P_R and is defined as a polytope that contains P and is such that the only $0,1$ points it contains are the vertices of P. A *cut* is an operator/function from the space of relaxations of P to itself. It is denoted by Φ and $\Phi(P_R)$ is a relaxation that is contained inside P_R. In this paper, we will define a cut called the matrix cut.

The dimension of the polytopes being considered will be denoted by n. It will be convenient to homogenize and view the n dimensional convex sets as cones in the $n+1$ dimensional space. With any n dimensional convex set $\mathscr{C}$ we can associate an $n+1$ dimensional cone K, defined as $K = \{(\lambda, \lambda x) : \lambda \in \mathbb{R}_+, x \in \mathscr{C}\}$, where $\mathbb{R}_+$ is the set of non-negative real numbers. This correspondence is one-one because the convex set $\mathscr{C}$ is given by the intersection of K and the hyperplane $x_0 = 1$; here we have indexed the components of $x \in \mathbb{R}^{n+1}$ from 0 to n. For brevity, we will denote this correspondence by H, i.e., we write $H(\mathscr{C}) = K$ and $H^{-1}(K) = \mathscr{C}$.

For any $n+1$ dimensional cone K we denote by K° the cone spanned by the $0,1$ vectors in K. We denote by Q the cone spanned by all the $0,1$ vectors with $x_0 = 1$. In other words, Q is the $n+1$ dimensional cone corresponding to the n dimensional unit cube. Given any cone K we denote its dual by K^* and define it as:

$$K^* = \left\{u \in \mathbb{R}^{n+1} : \langle u, x\rangle \geq 0,\ \forall\, x \in K\right\}$$

The dual of a cone can be thought of in two ways: (a) intersection of half spaces (b) union of normals. Thinking of $\langle x, y\rangle \geq 0$ with x fixed and y variable, K^* can be thought of as the intersection of halfspaces with normals lying in K. This is the first interpretation. The other way is thinking of $\langle x, y\rangle \geq 0$ with y fixed and x variable. Now, K^* will be the union of normals such that the non-negative halfspace contains K. Using any of these interpretations, it is easy to note that Q^* is the conic hull of e_i, $i = 0, 1, 2, \dots, n$ and $f_i = e_i - e_0$, $i = 1, 2, \dots, n$; here e_i denotes the i^{th} unit vector.

We will now formally define a matrix cut.

3 Definition of a matrix cut.

We will define the cut for a cone $K \subseteq Q$. The cut will be obtained by first lifting the $n+1$ dimensional cone to the space of $(n+1)\times(n+1)$ dimensional matrices, and then projecting it back to the $n+1$ dimensional space. Since we are lifting the cone to the space of matrices, this procedure is termed matrix cut. We define the procedure for the intersection of two cones $K_1 \cap K_2$, where $K_1 \subseteq Q$ and $K_2 \subseteq Q$. Since we will only be interested in the cut of a cone K, we can think of K_1 as K and K_2 as Q. The reasons for defining the cut for the intersection of two cones and not just for a cone are technical, and will become apparent later.

We first define the lifting. Given cones K_1 and K_2, we define the cone $M(K_1, K_2) \subseteq \mathbb{R}^{(n+1)\times(n+1)}$ as the set of all $(n+1)\times(n+1)$ matrices $Y = (y_{ij})$ satisfying:

(i) Y is symmetric.

(ii) $y_{ii} = y_{0i}$ for all $1 \leq i \leq n$.

(iii) $\langle u, Yv\rangle \geq 0$ for every $u \in K_1^*$ and $v \in K_2^*$.

Note that (iii) can also be written as

(iii')$YK_2^* \subseteq K_1$.

We will slowly parse each of the conditions to get an intuitive understanding, but before that we will define the projection. We define $N(K_1, K_2) = \{Ye_0 : Y \in M(K_1, K_2)\}$. Clearly, $N(K_1, K_2)$ is an $(n+1)$ dimensional cone. We will define $N(K_1, K_2)$ as the cut of $K_1 \cap K_2$.

Now, to intuition. The cut should contain all the $0, 1$ vectors of $K_1 \cap K_2$. Therefore, let's concentrate on the $0, 1$ vectors. Roughly speaking, the lift and projection should be such that the $0, 1$ vectors are mapped to themselves. For that, let x be a $0, 1$ vector in $K_1 \cap K_2$ and let $Y = xx^T$. Clearly, Y satisfies the properties (i), (ii) and (iii) stated above. It is also easy to observe that the three constraints are also satisfied by the conic hull of matrices of the form xx^T, with x a $0, 1$ vector in $K_1 \cap K_2$. We would be happy if we can restrict $M(K_1, K_2)$ to just the conic hull. But that would have exponential facets in general. Therefore, we limit ourselves to the three constraints, and this will turn out to be a good trade-off between the complexity and optimal cut.

This finishes the definition of the matrix cut. Of course, we still have to prove that $N(K_1, K_2)$ obtained is indeed a cut. Moreover, the geometric meaning of the matrix cut is still not very clear. For that, we will first provide a geometric interpretation of the matrix cut and then prove some of its useful properties.

4 Geometry of the cuts.

As mentioned earlier, we are interested in the cuts of a cone K. Therefore, from this section on we will take K_1 as K and K_2 as Q. For brevity, we will write $M(K,Q)$ simply as $M(K)$ and $N(K,Q)$ as $N(K)$.

Let H_i denote the hyperplane $\left\{x \in \mathbb{R}^{n+1} : x_i = 0\right\}$ and G_i the hyperplane $\left\{x \in \mathbb{R}^{n+1} : x_i = x_0\right\}$, for $1 \leq i \leq n$. The hyperplanes H_i and G_i are the facets of Q. We have the following lemma:

Lemma 1. *For every convex cone $K \subseteq Q$ and every $1 \leq i \leq n$,*

$$N(K) \subseteq (K \cap H_i) + (K \cap G_i).$$

The proof of this lemma is quite straightforward and we will come back to it soon. Before that, let's understand the implications of this lemma. Note that if K does not intersect G_i then $N(K) \subseteq H_i$. Similarly, if K meets two opposite facets of Q only at zero, then $N(K) = \{0\}$. It follows from lemma 1 that:

$$N(K) \subseteq \bigcap_i ((K \cap G_i) + (K \cap H_i))$$

We now take a quick example to understand this better.

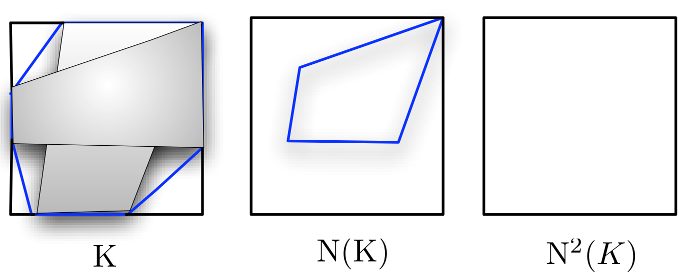

Fig. 1. Illustration of the geometric interpretation of matrix cuts

Figure 1 illustrates the matrix cuts on a convex cone K. Here $n = 2$. The figure shows that intersection of K and the hyperplane $\{x_0 = 1\}$. This depicts a geometric construction of the matrix cut.

We now prove Lemma 1.

Proof of Lemma 1. Consider any $x \in N(k)$. We can write it as Ye_0 for some $Y \in M(K)$. Let use denote the i^{th} column of Y by y_i, for $0 \leq i \leq n$. Thus, $x = y_0$. For any $1 \leq i \leq n$, we can write $y_0 = (y_0 - y_i) + y_i$. Let u denote $y_0 - y_i$ and v denote y_i. Since $y_{0i} = y_{ii}$, we have $u_i = 0$ and $v_0 = v_i$. Therefore, $u \in H_i$ and $v \in G_i$. This completes the proof of the lemma. □.

5 Properties of the matrix cuts.

We now state the three most important properties of the cuts. The first property states that the above definition of matrix cuts is indeed a cut. The second property states that starting from K, after at most n cuts we get back K°. The last property states that if there K has a polynomial time weak separation oracle ‡ then $N(K)$ has a weak separation oracle too. We will only sketch the proofs of the properties.

Lemma 2. $K^\circ \subseteq N(K) \subseteq K$.

Proof. Let x be any non-zero $0, 1$ vector in K. Since $K \subseteq Q$, $x_0 = 1$. As earlier, it is easy to check that $Y = xx^T \in M(K)$. Hence, $x = Ye_0 \in N(K)$. Thus, $K^\circ \subseteq N(k)$.

$N(K) \subseteq K$ follows immediately from $YQ^* \subseteq K$ and $e_0 \in Q^*$. □.

Lemma 3. $N^n(K) = K^\circ$.

Proof. We will only sketch the proof of this lemma. For any $1 \leq t \leq n$, let F be a $n - t$ dimensional facet of the n dimensional unit cube. Let F'

‡ A weak separation oracle is a version of the separation oracle which allows for numerical errors: its input is a vector $x \in \mathbb{R}^{n+1}$ and a rational number $\epsilon > 0$, and it either return an assertion that the Euclidean distance of x from K is at most ϵ, or returns a vector w such that $\|w\|_{\ell_2} \geq 1$, $\langle w, x \rangle \leq \epsilon$, and the distance of w from K^* is at most ϵ. If the cone K has $0, 1$ vertices, then we can strengthen a weak separation oracle to a strong separation oracle in polynomial time.

denote the $n+1$ dimensional cone corresponding to the union of all the $n-t$ dimensional facets that are parallel to F. We will prove by induction that $N^t \subseteq \text{cone}(K \cap F')$.

Note that the case $t = 1$ is equivalent to lemma 1. For $t = n$, this is just the statement of the theorem. Induction can be carried out in a manner similar to the proof of lemma 1 and we shall skip it here. □.

Lemma 4. *Suppose that we have a weak separation oracle for K. Then the weak separation problem for $N(K)$ can be solved in polynomial time.*

Proof. Suppose that we have a weak separation oracle for K. We will prove that we have a weak separation oracle for $M(K)$. In fact, if Y is a matrix in $M(K)$ then we can trivially check conditions (i) and (ii). The third condition states that $YQ^* \subseteq K$. Since Q^* has a polynomial number of extreme rays and K has a polynomial time weak separation oracle, it follows that the third condition can be checked in polynomial time.

Since $M(K)$ has a polynomial time weak separation oracle, even its projection $N(K)$ has a weak separation oracle; this follows from the general results proved in [2]. □.

6 Applications to the stable set problem.

We will now apply matrix cuts to determine the classes of graphs for which the stable set problem can be solved in polynomial time.

Let $G = (V, E)$ be a graph with no isolated nodes. Let $\alpha(G)$ denote the maximum size of any stable set of nodes of G. For each $A \subseteq V$, let $\chi^A \in \mathbb{R}^V$ denote its incidence vector. The stable set polytope is defined as:

$$\text{STAB}(G) = \text{conv}\left\{\chi^A : A \text{ is stable}\right\}.$$

The polytope STAB(G) satisfies the following inequalities:

$$x_i \geq 0, \quad \text{for each } i \in V \tag{1}$$

and

$$x_i + x_j \leq 1 \quad \text{for each } ij \in E. \tag{2}$$

(1) and (2) are called respectively the non-negativity and edge constraints. The solution set of the (1) and (2) is denoted by FRAC(G) and is in general much larger than STAB(G). FRAC(G) and STAB(G) are equal if and only if G is bipartite.

There are other classes of inequalities that are satisfied by STAB(G). Some important classes are *clique* constraints and *odd hole* constraints. The clique constraints are defined as: for every clique B, we have:

$$\sum_{i \in B} x_i \leq 1. \tag{3}$$

Graphs for which (1) and (2) are sufficient to describe STAB(G) are called *perfect*. The odd hole constraints are defined as: if C induces a chordless odd cycle in G then

$$\sum_{i \in C} \leq \frac{1}{2}(|C| - 1). \tag{4}$$

Graphs for which (1), (2) and (4) are sufficient to describe STAB(G) are called *t-perfect*.

It has been proved that the stable set problem can be solved in polynomial time for bipartite, perfect and t-perfect graphs.

Taking P as FRAC(G), with K denoting the corresponding cone, we can apply the theory developed in the last section to determine $N(K)$. We denote $N(K)$ by $N(G)$. Since FRAC(G) is polynomial time separable, it follows that $N^r(G)$ is polynomial time separable for fixed r.

We will now state an important theorem:

Theorem 1. *The polytope $N(G)$ is exactly the solution set of the non-negativity, edge and odd-hole constraints.*

We will skip the details of the proof. This theorem gives us an alternative and more general method to prove the stable set problem for t-prefect graphs can be solved in polynomial time.

There are several other results that can be proved for the stable set problem using this methodology.

7 Conclusion.

In this paper, Lovász and Schrijver have developed the method of matrix cuts and applied it to the problems of stable sets in graphs and set functions. Their method is based on the idea that, in general, projection of a polytope has more facets than the polytope itself. More specifically, a polytope with a polynomial number of facets can have exponentially many facets in its projection. Using this idea they introduced matrix cuts. Carrying out the matrix cuts repeatedly results in a sequence of relaxations. This procedure is guaranteed to end in n steps.

The methodology they develop is more general and encompasses most of the *ad hoc* techniques.

References

1. L.Lovász and A. Schrijver, "Cones of matrices and setfunctions and 0-1 optimization," *Siam Journal on Optimization*, vol. 1, pp. 166–190, 1991.
2. L. L. M. Grötschel and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization.* Springer, 1988.