Fundamental Domains for Integer Programs with Symmetries

Eric J. Friedman

Cornell University, Ithaca, NY 14850,
ejf27@cornell.edu,
WWW home page: http://www.people.cornell.edu/pages/ejf27/

Abstract. We define a fundamental domain for a linear programming relaxation of a combinatorial integer program which is symmetric under a group action. We then describe a straightforward way to construct fundamental domains defined by the maximization of a linear function. The computation of this fundamental domain is at worst polynomial in the size of the group; however, for the symmetric group, which has exponential size, we show how to compute separation in polynomial time (in the size of the integer program).
Fundamental domains are a simple and flexible approach to reducing the computation difficulties that often arise in integer programs with symmetries. Their construction is closely related to the constructions of orbitopes (by Kaibel and Pfetsch), but more general and easier to analyze, although the computations required may be somewhat more complex.

1 Introduction

Combinatorial integer programs with symmetries arise in many standard problem formulations. Unfortunately, these symmetries often make the problems difficult to solve with integer programming algorithms. They cause difficulties with branching, since each branch will have many symmetric copies, and with linear programming relaxations, which become much less informative due to the symmetry of the feasible region.

For example, in a simple bin packing problem with multiple bins of the same size one often uses the variable x_{ij} to represent whether item i is in bin j. However, if bins j and k are the same size then any solution x is equivalent to the solution x' when x' is derived from x by exchanging the two columns, j and k. One way of resolving these problems is to restrict the search space to eliminate the additional equivalent copies of a solution. This can be done either by adding additional constraints to the integer program [2, 1, 3, 8] or by modifying the branch and bound or branch and cut algorithms [6, 7].

In this paper, we consider the problem of removing the multiple symmetrically equivalent solutions. We present a new approach which is simple, flexible and extremely general. In addition the analysis of our methods appears to much simpler than most existing approaches.

Our approach consists of constructing a polytope for a "minimal fundamental domain", which is a subset of the feasible region and contains only a single "representative" from each equivalence class of symmetrically equivalent extreme points.

Our work was motivated by Kaibel and Pfetch's [5] recent study of orbitopes. In that paper they considered a face of a fundamental domain generated by a lexicographic ordering. They provided a complete description for orbitopes for the cyclic and symmetric groups under packing and partitioning constraints.

In this paper, we consider a different approach to this problem: finding fundamental domains defined by maximizing a linear function. This leads to simple constructions and straightforward proofs. It also allows these techniques to be extended to more complex settings and provides flexibility in implementation.

For example, consider a bin packing problem with multiple bins. Previous methods have considered the case when all bins are identical, in which the problem is invariant under the full symmetric group; however, our methods apply to arbitrary sets of identical bins, e.g., three bins of size 10, six bins of size 14 and one bin of size 22. In addition, our methods extend directly to covering problems without the combinatorial complexities that arise in the related orbitopes.

Our methods also apply to other groups, such as cyclic groups, which arise in transportation scheduling problems [10] or even problems for which several different group actions are combined. For example, consider a periodic bus scheduling problem with multiple bus sizes. This problem is invariant under the exchange of equal capacity buses (symmetric groups) and under time transformations (cyclic group).

In fact, our methods apply straightforwardly to any linear group action acting on mixed-integer programs, e.g., symmetries can involve continuous variables.

In this paper, we present the general theory of fundamental domains. For notational simplicity we restrict to binary linear programs, but note that all our methods extend easily to general mixed-integer programs.

In the following section we provide the basic construction and then in Section 3 discuss the separation problem for the cyclic and symmetric groups. Section 4 compares fundamental domains to orbitopes, Section 5 discusses combinations of groups and Section 6 considers the linear optimization criterion used for generating the fundamental domains. We conclude in Section 7.

2 Group Actions and Fundamental Domains

Let G be a finite group and given a set $X \subset \Re^n$ consider a group action $\phi_g : X \to X$. A group action must satisfy $\phi_{g \circ g'} = \phi_g \circ \phi_{g'}$ for all $g, g' \in G$. Given $x \in X$ define the orbit of x, $orb(x)$ to be the set $orb(x) = \{\phi_g(x) \mid \forall g \in G\}$. A (generalized) fundamental domain of X is a subset $F \subset X$ such that its orbit $orb(F) = X$, where $orb(F) = \bigcup_{y \in F} orb(y)$. A fundamental domain is minimal if in addition, there is no closed subset of F that is also a fundamental domain.

To specialize to polytopes, for simplicity assume that $X \subseteq [0, 1]^n$. Let $Ext(X)$ be the extreme points of X, which are assumed to be integral. In addition, we will require that for all $g \in G$, the group action is an affine transformation of X. Thus, for each $g \in G$ we can assume that $\phi_g = A_g + b_g$ where A_g is an nxn matrix and b_g is an n-vector. (Note that our approach easily generalizes to arbitrary polytopes corresponding to general mixed-integer programs.)

We first note a basic fact about affine group actions of finite groups.

Lemma 1. *Let G be a finite group and $\phi : G \times X \to X$ be an affine group action of G. Then $\forall g \in G$ the determinant of A_g has the absolute value of 1.*

Proof: Since $\phi_{g^{-1}} = (\phi_g)^{-1}$ the action ϕ_g is invertible so A_g must have nonzero determinant. In addition, since G is finite and for all $g \in G$, the composition of ϕ_g with itself k times, $(\phi_g)^k = \phi_{g^k} = \phi_{g'}$, for some $g' \in G$. Now the determinant of $\phi_{g'}$ must satisfy $det(\phi_{g'}) = det(\phi_g)^k$, so unless $|det(\phi_g)| = 1$, $(\phi_g)^k$ will be different for all k, contradicting the assumption that G is finite. □

Given an "ordering vector" $c \in \Re_n$, we define the fundamental domain of X, F_c, with respect to G by

$$F_c = \{x \in X \mid c^t x \geq c^t \phi_g x \quad \forall g \in G\}$$

Lemma 2. *For any ordering vector c, the fundamental domain, F_c is a polytope.*

Proof: The fundamental domain is defined by a finite set of affine inequalities.
□

For example, consider the case with $X = [0, 1]^2$ where G is the additive group Z_2 with elements $\{0, 1\}$, and $0 + 0 = 0$, $0 + 1 = 1$, and $1 + 1 = 0$. Define the action of this group by setting ϕ_0 to be the identity and ϕ_1 being the exchange operator, $\phi_1(x_1, x_2) = (x_2, x_1)$. Let $c = (2, 1)$. Then

$$F_c = \{x \in X \mid 2x_1 + x_2 \geq x_1 + 2x_2\},$$

which implies that

$$F_c = \{x \in X \mid x_1 \geq x_2\}.$$

Thus, F_c is the convex hull of the extreme points $(0, 0)$, $(1, 1)$ and $(1, 0)$. Different choices of c can lead to different fundamental domains. For example if $c = (1, 2)$ then the fundamental domain now includes $(0, 1)$ instead of $(1, 0)$.

First we note that a fundamental domain always contains a "representative" for each extreme point of X.

Theorem 1. *Let $x \in Ext(X)$. For any ordering vector c, there exists a $g \in G$ such that $\phi_g x \in Ext(F_c)$.*

Proof: This follows immediately from the definition of F_c since for each $x \in Ext(X)$ must have at least one largest element in it's orbit, $c^t \phi_g x \;\; \forall g \in G$, since $|G|$ is finite. □

Note that, unlike orbitopes [5], there can exist extreme points of F_c which are not integral. For example, consider the case with $X = [0,1]^2$ and $G = Z_2$, where ϕ_1 inverts the first element of x, $\phi_1 = (x_1, x_2) = (1 - x_1, x_2)$. Let $c = (2, 1)$. Then

$$F_c = \{x \in X \mid 2x_1 + x_2 \geq 2(1 - x_1) + x_2\},$$

which implies that

$$F_c = \{x \in X \mid x_1 \geq 1/2\}$$

which has $(1/2, 0)$ as an extreme point.

Note that a fundamental domain generated in this way need not be minimal. For example, when $c = 0$ we get $F_c = X$. However, even if c is nontrivial, the fundamental domain need not be minimal.

First we show that there is a universal ordering vector $\hat{c}$ which generates minimal fundamental domains.

Theorem 2. *Let $\hat{c} = (2^{n-1}, 2^{n-2}, \ldots, 2, 1)$ be the "universal ordering vector". Then $F_{\hat{c}}$ will be minimal.*

Proof: First note that $F_{\hat{c}}$ contains a unique element of any orbit of an extreme point. This follows because $\hat{c}$ induces a lexicographic order on extreme points.

Next, we note that $F_{\hat{c}}$ must be full dimensional. i.e., the same dimensionality as X. This is because $Orb(F_{\hat{c}}) = X$ and each $Orb(x)$ contains a finite number of points.

Suppose that for some point $x \in F_{\hat{c}}$ there exists some $g \in G$ such that $\hat{c}^t \phi_g x = \hat{c}^t x$ and $\phi_g x \neq x$. However, this implies that the constraint from ϕ_g is tight, so unless the constraint is trivial $(0 \geq 0)$ x will not be an interior point.

Since $x \in X$ and X is convex, we can write $x = \sum_{j=1}^{n+1} \alpha_j w^j$ where $\alpha \geq 0$, $\sum_{j=1}^{n+1} \alpha_j = 1$ and w^j are all extreme points of x. Since $\phi_g x \neq x$ and $\phi_g x = \sum_{j=1}^{n+1} \alpha_j \phi_g w^j$, there exists at least one j such that $w^j \neq \phi_g w^j$, and call this extreme point v.

Since $\hat{c}^t y \neq \hat{c}^t y'$ for any pair of extreme point $y \neq y'$ this implies that $\hat{c}^t v \neq \hat{c}^t \phi_g v$ which implies that the constraint is not trivial, since ϕ_g is affine. □

Note that the universal ordering vector $\hat{c}$ requires $\Omega(n)$ bits per element, on average. As we show in in Section 6 the need for a large number of bits seems common and provides a potential drawback to this approach.

3 Separation for the Cyclic and Symmetric Groups

Two of the most common groups arising is practice are the cyclic and symmetric groups. The cyclic groups of order k can be represented simply the additive group of integers modulo k and are denoted by Z_k. These are generated by a single element $1 \in Z_k$. The most natural action can be most easily described by viewing $x \in X$ as a matrix with r rows and k columns, and note that $n = rk$. Then the action of ϕ_1 is given by cyclicly moving the first k columns of this matrix, i.e, the first column becomes the second, the second becomes the third, column $k-1$ becomes column k and column k becomes column 1. Let $A = A_1$ be the matrix representation of ϕ_1 and note that $b_1 = 0$. Since $|G| = n$ the fundamental domain can be concisely represented by

$$F_c = \{x \in X \mid c^t x \geq c^t M^j x \quad j = 1..k-1\}$$

and clearly given a point $x \in X$ but $x \notin F_c$ one can find a separating hyperplane by checking all $k-1$ inequalities.

Theorem 3. *For the cyclic group (as described above), given a point $x \in X$ but $x \notin F_c$ one can find a separating hyperplane between x and F_c in $O(ns)$ time where s is the average number of bits in an element of c.*

The symmetric group is more complicated. As above, consider the vector x as a matrix, but in this case the group S_k is the set of all permutations of k elements and note that $|G| = k!$ which is exponential in k. Now, the group action consists of permutations of the first k columns of the matrix representation of x and the fundamental domain requires $k!$ additional inequalities. However, one can find a separating hyperplane efficiently as follows.

For simplicity, assume that $c > 0$. Construct a bipartite graph where the one set of vertices represents each of the first k columns of x (in the current ordering) and the second set represents the first k columns of c. Let the value of an edge from i to j be the inner product of the i'th column of x and the j'th column of c. Then the maximum matching gives the optimal permutation. The separating hyperplane is simply given by the constraint related to this permutation. Since a maximum matching can be computed k^3 operations, we have proven the following theorem.

Theorem 4. *For the symmetric group action (as described above), given a point $x \in X$ but $x \notin F_c$ one can find a separating hyperplane between x and F_c in $O((n + k^3)s)$ time where s is the average number of bits in an element of c.*

We note separation for orbitopes [5] can be done in linear time and thus our approach (for $r >> k$) is slower when the ordering vector uses a large number of bits.

4 Partitioning, Packing, Covering and Relations to Orbitopes

Now we discuss some important applications in which symmetry arises. Consider an optimization problem where there are r objects which must be put into k groups. Let x_{ij} be the variable that indicates that item i is in group j. Thus, the j'th column identifies the elements that are in group j. Problems can then be classified into three classes: partitioning (in which each item is in exactly one group), packing (in which each item is in at most one group), and covering (where each item is in at least one group).

When groups are identical, as in many combinatorial graph theory problems (such as coloring or partitioning), the IP is invariant under the full symmetry group of column permutations. Thus, our results from the previous section provide polynomial representations that remove much of the redundancy in the natural formulations.

However, in periodic scheduling problems, the same matrices arise but are only invariant under the cyclic group.

These problems are the subject of study by Kaibel and Pfetch [5] and the motivation behind orbitopes. Orbitopes are constructed by taking the convex hull of the set of $x \in Ext(X)$ which are maximal under the lexicographic ordering. While orbitopes are more refined than minimal fundamental domains, their analysis is significantly more complicated. In particular, the orbitopes for the covering problems appear to be quite complex and their explicit construction is not known. However, as can be seen from the analysis in the previous sections, the minimal fundamental domains can be easily characterized in all of these cases.

In fact, given the simplicity of their construction, our analysis easily extends to a wide variety of group actions.

5 Products of Groups

In this section we show that our analysis can be easily extended to special cases involving products of groups.

Given a group G and with a group action ϕ_g^G define the null space of the action to be the set of indices for which

$$(\phi_g^G x)_i = x_i \quad \forall g \in F, x \in X.$$

Define the active space of the action to be the complement of the null space.

Now consider a second group action, H, ϕ^H such that the active space of H does not intersect the active space of G. The if we define the product action GH, ϕ^{GH} where GH is the direct product of the two groups, so an element of GH is the pair (g, h) with $g \in G$ and $h \in H$. The action is then given by $\phi_{(g,h)}^{GH} = \phi_g^G \phi_h^H$ and note that this is equal to $\phi_h^H \phi_g^G$ since the actions ϕ^G and ϕ^H must commute.

Then, the fundamental domain of the product action is simply the intersection of the fundamental domains and thus the required number of constraints in only $(|G| - 1) + (|H| - 1)$ instead of $(|G| - 1)(|H| - 1)$.

Theorem 5. *If active spaces of a set of group actions do not intersect then the fundamental domain of the product action is the intersection of the fundamental domains of the individual actions.*

For example, in the case where pairs of groups $\{(1, 2), (3, 4) \cdots, (n-1, n)\}$ are interchangeable, the product action has $2^{n/2}$ constraints while the representation of the fundamental domain only requires $n/2$ constraints using the above result.

It appears that non-intersection of the group actions, although quite stringent, is necessary for these simplifications. One natural conjecture, that commutativity of the group actions is sufficient can be seen to be false from the following example.

Consider $X = [0,1]^2$ and the action of G is interchanging the two components, $\phi_1^G(x_1, x_2) = (x_2, x_1)$ while H flips both bits, $\phi_1^H(x_1, x_2) = (1 - x_1, 1 - x_2)$. It is easy to see that the two group actions commute; However, the constraints for the two fundamental domains when taken separately with $c = (2, 1)$ are:

$$G: \ 2x_1 + x_2 \geq x_1 + 2x_2 \quad \rightarrow x_1 \geq x_2$$

$$H: \ 2x_1 + x_2 \geq 2(1 - x_1) + (1 - x_2) \quad \rightarrow 4x_1 + 2x_2 \geq 3$$

However, the constraint for the joint action, $\phi_{(1,1)}^{GH}$ is

$$2x_1 + x_2 \geq 2(1 - x_2) + (1 - x_1) \quad \rightarrow x_1 + x_2 \geq 1$$

which removes additional points from the intersection of the two separate group actions.

6 Ordering Vectors

The universal ordering vector $\hat{c}$ requires $\Omega(n)$ bits which can be problematic with standard IP solvers. As we shall see, it is often true that large number of bits are necessary to generate minimal fundamental domains.

Consider the space with variables $x_{ij} \in X$ considered earlier with r rows and $k = t$ columns. Thus $X = [0,1]^{kr}$. We consider the action of the symmetric group that exchanges columns.

Note that when $r = 1$ cost vector with $c_{1j} = k - j$ generates minimal fundamental domains under the symmetric group. However, for larger values of r significantly more bits are required

Theorem 6. *Let $c \geq 0$ be an ordering vector which generates a minimal fundamental domain for the action of the symmetric (or cyclic) group on $X = [0,1]^{kr}$. Then there exists some i, j such that c_{ij} uses more than r/k^2 bits.*

Proof: For $j \neq j'$ define $d_i^{jj'} = c_{ij} - c_{ij'}$ and let $\emptyset \subset S_j \subset \{1, 2, \ldots r\}$. Then define $x(S)$ by $x_{ij} = 1$ if $i \in S_j$ and 0 otherwise. Define $C^{jj'}(S_{\hat{j}}) = \sum_{i \in S_{\hat{j}}} d_i^{jj'}$.

We claim that if there exists some $S = (S_1, S_2, \ldots, S_k)$ such that $C^{jj'}(S_j) = C^{jj'}(S_{j'})$ and $S_j \neq S_{j'}$ for all $j \neq j'$ then the fundamental domain will not be minimal, because $c^t x(S)$ was constructed to be invariant under the symmetric group. To see this note that the elements of the group that exchange two columns generate all elements of the group and the construction guarantees that $c^t x(S)$ is invariant under any number of column exchanges.

Now assume that all components in c use less than b bits. This implies that $0 \leq C^{jj'}(S_{\hat{j}}) \leq r2^b$. Now define $C(S_j) \in \{0, 2^b - 1\}^{k(k-1)/2}$ by $C(S_j)_{jj'} =$

$C^{jj'}(S_j)$ for $0 < j < j' \leq k$. Note that there are $(2^b)^{k(k-1)/2}$ points in $\{0, 2^b - 1\}^{k(k-1)/2}$ and there are $2^r - 1$ subsets of $\{1, 2, \ldots r\}$. In order to guarantee that some point in $\{0, 2^b - 1\}^{k(k-1)/2}$ is covered by at least K subsets it is necessary that $2^r - 1 > (k-1)2^{bk(k-1)/2}$ which we relax to $2^r > k2^{bk^2}$. So if $b < r/k^2 - \log_2(k)$ there must be two extreme points of X which which are maximal under c and in the same orbit, proving the theorem. □

7 Conclusions

We have provided a direct method for finding a minimal fundamental group for a group action on a binary polytope. First we note that our methods can be easily extended to arbitrary polytopes and group actions. The only impediment to complete generality is the need to compute separation which might not be efficiently computable in some cases.

While this problem is of theoretical interest, it remains to be seen whether it is of practical value in solving real integer programs. However, recent results on the use symmetries in solving packing and partitioning problems [4, 9], suggest fundamental domains might prove useful.

8 Acknowledgements

I would like to thank Volker Kaibel, Adrian Lewis, Marc Pfetsch and David Shmoys for helpful comments. This work was partially supported by NSF grant ITR-0325453.

References

1. C. Barnhart, E. L. Johnson, G. L. Nemhauser, M.W. P. Savelsbergh, and P. H. Vance. Branch and price: Column generation for solving huge integer programs. *Operations Research*, 46:316–329, 1998.
2. M. S. Bazaraa and O. Kirca. A branch-and-bound heuristic for solving the quadratic assignment problem. *Naval Research Logistics Quarterly*, 30:287304, 1983.
3. S. Holm and M. Srensen. The optimal graph partitioning problem: Solution method based on reducing symmetric nature and combinatorial cuts. *OR Spectrum*, 15:18, 1993.
4. V. Kaibel, M. Peinhardt, and M. Pfetsch. Orbitopal fixing. In *Proc. 12th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, Lecture Notes in Computer Science (forthcoming), 2007.
5. V. Kaibel and M. Pfetsch. Packing and partitioning orbitopes. *Math. Program., Ser. A (in press, DOI: 10.1007/s10107-006-0081-5)*, 2006.
6. F. Margot. Pruning by isomorphism in branch-and-cut. *Mathematical Programming*, 94(1):71–90, 2002.
7. F. Margot. Small covering designs by branch-and-cut. *Mathematical Programming*, 94(2):207–220, 2003.

8. Isabel Méndez-Díaz and Paula Zabala. A branch-and-cut algorithm for graph coloring. *Discrete Applied Mathematics*, 154(5):826847, 2006.
9. J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio. Orbital branching. In *Proc. 12th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, Lecture Notes in Computer Science (forthcoming), 2007.
10. P. Serafini and W. Ukovich. A mathematical model for periodic scheduling problems. *SIAM J. Discrete Math*, 2(4):550–581, 1989.