Mathematical Programming 24 (1982) 39–54
North-Holland Publishing Company

RANDOM POLYTOPES: THEIR DEFINITION, GENERATION AND AGGREGATE PROPERTIES

Jerrold H. MAY

Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

Robert L. SMITH

Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI 48109, U.S.A.

Received 23 January 1981
Revised manuscript received 26 October 1981

The definition of random polytope adopted in this paper restricts consideration to those probability measures satisfying two properties. First, the measure must induce an absolutely continuous distribution over the positions of the bounding hyperplanes of the random polytope; and second, it must result in every point in the space being equally as likely as any other point of lying within the random polytope. An efficient Monte Carlo method for their computer generation is presented together with analytical formulas characterizing their aggregate properties. In particular, it is shown that the expected number of extreme points for such random polytopes increases monotonically in the number of constraints to the limiting case of a polytope topologically equivalent to a hypercube. The implied upper bound of 2^n where n is the dimensionality of the space is significantly less than McMullen's attainable bound on the maximal number of vertices even for a moderate number of constraints.

Key words: Random Polytopes, Linear Programming, Problem Generation, Aggregate Polytope Properties.

1. Introduction

We consider, in this paper, the problem of generating and characterizing random feasible regions of linearly constrained mathematical programs. In particular, a large class of such random feasible regions (or convex polytopes) is operationally defined, efficient Monte Carlo procedures for their generation are discussed, and analytical formulas characterizing their aggregate properties are presented.

The motivation for this study is two-fold. First, with the increasing proliferation of new and competing mathematical programming algorithms, sound procedures for their comparative testing have taken on new importance. Many of these evaluation schemes rely on the results of applying the algorithms to a series of randomly generated test problems. There is concern, however, about biases that may be inadvertently created by an uncritical introduction of randomness [7]. The emphasis here is on the term inadvertant, since special structure consistent with the class of real world problems being simulated is

welcome, and, indeed, necessary. We introduce here extremely efficient procedures that generate geometrically unbiased feasible regions. Controlled special structure may then be induced through standard rejection techniques. Second, turning to the observed performance with real world problems of mathematical programming algorithms in general and the simplex method in particular, it remains to this day an open question as to why they perform so well [10]. As Liebling [18] has remarked, one can view this performance as the result of a vast simulation provided by real world testing. The success of the simplex algorithm in particular is remarkable for the following reason. Klee and Minty [14] demonstrated that the simplex method may at times search every extreme point of the polytope of feasible solutions before arriving at the optimum. Moreover, McMullen [20] has shown that the maximal number of extreme points that a convex polytope may attain is equal to that of the class of cyclic polytopes, and this number grows exponentially in the size of the problem. These results, and others, have led to speculation that real world convex polytopes must be much better behaved, on the average. The mean number of extreme points for the class of random polytopes considered in this paper is supportive of this conjecture. In particular, we show that the expected number of vertices is bounded from above by 2^n for any number of constraints, where n is the dimensionality of the problem.

The general class of random polytopes considered in this paper is defined in Section 2. We make it as general as possible, with any special characteristics being motivated by geometric notions of randomness. Section 3 discusses an efficient Monte Carlo procedure for the generation of random polytopes. Their aggregate properties are demonstrated in Section 4. In particular, explicit analytical formulas are derived for the expected number of vertices, edges, facets, etc. These results are essentially distribution free. In Section 5, we propose a special class of random polytopes that are defined to be those whose random properties are preserved under Euclidean transformations. Their generation leads to a choice of distribution that is geometrically random.

2. A definition for random polytopes

An operational definition of random polytopes is employed in this section to indirectly define the class or population of random polytopes to be considered. More precisely, we restrict the allowable distributions in the Monte Carlo generation of random polytopes.

As mentioned in the Introduction, a general convex polytope [25], hereafter referred to as a *polytope*,[1] is the intersection of a finite number of closed

[1] What we have referred to as a polytope is more conventionally termed a polyhedral set [12]. The term polytope is usually reserved for bounded polyhedral sets.

half-spaces. Algebraically, $P \subseteq R^n$ is a polytope, if and only if it can be represented by

$$P = \{x \in R^n : a^i \cdot x \leq b_i, a^i \in R^n, b_i \in R, \quad i = 1, 2, \ldots, m\}.$$

Note that P may be unbounded.

In order to define the notion of a random polytope, we first need to parameterize the class of all polytopes. This could be done via (a^i, b_i), $i = 1, 2, \ldots, m$. However, this parameterization is not one-to-one since (ka^i, kb_i), $i = 1, 2, \ldots, m$, for all $k > 0$, refers to the same polytope. We need our parameterization to be one-to-one in Section 4, and so we adopt the one-to-one parameterization

$$s = ((p^1, i_1), (p^2, i_2), \ldots, (p^m, i_m)) \in S = (R^n \times \{0, 1\})^m$$

corresponding to

$$P = \{x \in R^n : p^j \cdot x \gtreqless \|p^j\|^2, p^j \in R^n, j = 1, 2, \ldots, m\}$$

where the sense of the jth inequality is $\leq$ for $i_j = 0$ and $\geq$ for $i_j = 1$. Geometrically, p^j is the foot of the perpendicular from the origin to the jth constraining hyperplane (Fig. 1).

The random polytopes being considered may now be defined by imposing a multivariate probability distribution F over the parameter space S. It is clear that some $s \in S$ correspond to empty polytopes (e.g., $s = ((p^1, 0), (p^2, 0), (p^3, 1))$ in Fig. 1). To rule out these pathological cases, we define F as the conditional probability distribution over S given that the polytope generated by the probability distribution F_0 over S is non-empty. The properties we assume for F_0 and, by implication, for F are stated below.[2]

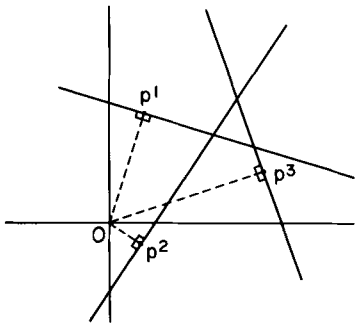

Fig. 1

[2] Letting $\bar{S}$ be the set of all $s \in S$ corresponding to feasible polytopes, F assigns zero measure to any set in $S - \bar{S}$. Restricting F to $\bar{S}$ and considering any event $E \subseteq \bar{S}$, we get $F(E) = kF_0(E)$ where $k = 1/F_0(\bar{S})$.

Assumptions on the generating probability distribution F_0

(a) The multivariate distribution over $p = (p^1, p^2, \ldots, p^m) \in R^{mn}$ is absolutely continuous with respect to Lebesgue measure.

(b) The multivariate distribution over $(i_1, i_2, \ldots, i_m) \in \{0, 1\}^m$ is given by $\mathbf{P}(i_j = 0) = \mathbf{P}(i_j = 1) = \frac{1}{2}$ independently[3] for all $j = 1, 2, \ldots, m$ and independently of p.

Assumption (a) is motivated by the need to rule out degeneracy and other exceptional alignments of constraining hyperplanes. The polytopes generated are, in fact, with probability one so-called simple polytopes, i.e., polytopes whose vertices are formed by exactly n facets [12]. This is a necessary assumption for the analysis in Section 4, where the aggregate properties of these polytopes are investigated. The effect of this assumption is to also, with probability one, rule out sparsity and integer constraint coefficients. Assumption (b) in the presence of assumption (a) is equivalent to assumption (b′):

(b′) Every point is equally as likely as any other point of lying in the random polytope generated.

To deny (b) in the presence of (a) necessarily leads to a procedure for selecting regions that is biased towards (or against) those regions covering exceptional points (e.g., the origin). Procedures for generating random test problems often force feasibility of the origin to assure non-emptiness of the region generated. For the class of simple product mix problems this is justified and necessary on conceptual grounds. On the other hand, for the class of diet problems, for example, the origin should, with probability one, be excluded. We view the class of random problems defined by assumptions (a), (b) as an aggregate class, subsuming these and other special subclasses. There is, accordingly, no reason to favor or disfavor some points over others.

The equivalence of (b) and (b′) is seen by conditioning on a realized configuration of the hyperplanes corresponding to $p^1, p^2, \ldots, p^m$. Let x_0 be any point. Then we want the probability that x_0 lies in the region selected by the choice of $i_1, i_2, \ldots, i_m$ given that the region is non-empty. But this is simply $k(\frac{1}{2})^m$ by assumption (b), where k is one over the probability that the region generated by F_0 is non-empty. The reverse implication is argued similarly.

An important corollary of the above analysis is that *for a particular realized configuration of hyperplanes, each polytope so created is equally likely to be selected.* This follows from choosing a representative interior point of each polytope and noting that each point is equally likely to lie in the polytope generated. Hence assumption (b) amounts to a simple random selection of a polytope from the population of realized polytopes.

[3] $i_1, i_2, \ldots, i_m$ are dependent under the induced distribution F since, under F, the polytope generated is necessarily non-empty.

3. Efficient generation of random polytopes

An operational definition of the class of random polytopes considered in this paper was given in the last section. We turn now to designing an efficient Monte Carlo procedure for their generation for a fixed F_0 and hence F.

A direct implementation of the definition of Section 2 would be the following iterative procedure. First, generate $p^1, p^2, \ldots, p^m$ according to the multivariate distribution induced by F. Second, generate m uniform [0, 1] random variables $u_1, u_2, \ldots, u_m$ and if $u_j \in [0, \frac{1}{2})$ select $\leq$ for constraint j and otherwise select $\geq$. Third, determine whether the resulting polytope is non-empty. If so, we are done and if not, return to the second step.

Unfortunately, the procedure is not practical. First, Step 3 requires an effort equivalent to solving a linear program of m constraints in R^n. Second, the expected number of iterations through Step 3 required to attain a non-empty region grows exponentially in the number of constraints m. For example, for 100 lines in the plane, the expected number of iterations[4] is in the order of 10^{26}.

We could avoid the computational difficulties of generating a feasible polytope altogether by forcing feasibility through a selection of $\leq$ for all constraints. Such a procedure does not, however, represent a simple random selection from the regions generated and therefore violates our definition of a randomly generated polytope. This is not a minor point, because a choice of the region covering the origin biases the choice toward larger regions in an effect similar to the inspection paradox of renewal theory [9]. A more sophisticated approach would be to attempt to prune the tree of inequality choices. This can also introduce subtle biases in region selection. For example, in a depth first search where a node is fathomed if it corresponds to an infeasible partial selection of constraint directions, the ultimate selection is biased toward those regions bounded by constraints forming early branches of the tree.

The algorithm we propose for region selection is based on an assignment of vertices of the partition to regions so that corresponding to each region is a unique vertex. A simple random selection of a vertex, which is easily accomplished, then corresponds to a simple random selection of a region.

We first enclose all vertices of the partition in a hypercube (see Fig. 2). The intersection of the interior of the hypercube and the partition then forms a collection of *bounded* regions corresponding in a one-to-one fashion with the original regions of the partition. The assignment of vertices to these bounded regions proceeds by randomly generating a hyperplane H_0 which is successively passed through vertices of the larger partition that includes the $2n$ bounding hyperplanes. The region associated with a given vertex is then that unique region lying wholly within a fixed half-space of H_0 (see [19] for a proof that this region

[4]The expected number of iterations is, in general, given by $((\sum_{i=0}^{n}\binom{m}{i})/2^m)^{-1}$. The result follows from the fact that $\sum_{i=0}^{n}\binom{m}{i}$ represents the number of regions created (see Section 4).

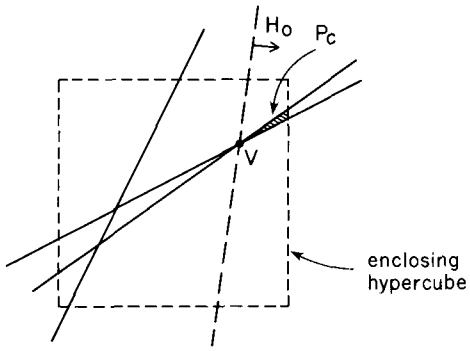

Fig. 2

is unique with probability one). It is clear that each of the bounded regions is with probability one assigned one and only one such vertex since every bounded linear program whose constraints are in general intersection has a unique optimal extreme point solution with H_0 representing the objective function.

The procedure then is to randomly generate a set of m hyperplanes H_1, $H_2, \ldots, H_m$ and enclose all vertices so formed with $2n$ hypercube constraints. Next pass a randomly generated hyperplane H_0 through a vertex V simply randomly selected from the set of all vertices formed by the $m+2n$ hyperplanes and identify the unique bounded polytope lying in a fixed half-space of H_0. If that polytope lies within the enclosing hypercube, accept it and the corresponding original polytope. If it lies outside, reject it and repeat the procedure with another simply randomly chosen vertex V. In contrast to the other procedures discussed, the probability of being successful on any given iteration converges to one as the number of hyperplanes m grows large.

The following formal statement of the procedure includes a constructive algorithm for identifying the unique region that lies in the fixed half-space of H_0. This is a non-trivial problem, since the determination of each inequality direction is dependent on the joint configuration of all of the hyperplanes passing through V.

Monte Carlo procedure for the generation of random polytopes

(1) Simply randomly select without replacement any n of the $m+2n$ constraints indexed $i_1, i_2, \ldots, i_n$ and let $B = [p^{i_1}, \ldots, p^{i_n}]^T$ and set $V = B^{-1}[\|p^{i_1}\|^2 \ldots \|p^{i_n}\|^2]^T$.
Note: B^{-1} exists with probability one.

(2) Randomly select $c \in R^n$ and let the corresponding hyperplane H_0 pass through V, i.e., $c^T x = c^T V$.

(3) The unique bounded polytope P_c lying in the half-space $c^T x \geq c^T V$ is then given by the following assignment of inequalities: Let $y^* = c^T B^{-1}$. Then the

i_jth inequality is $\geq$ if $y_j^* > 0$ and $\leq$ if $y_j^* < 0$. Set the directions of all inequalities of hyperplanes not passing through V to make V feasible.

Note: $y_j^* = 0$ will with probability one not occur.

(4) If P_c lies wholly within the enclosing hypercube, accept it and remove the hypercube constraints. Otherwise, reject it and return to (1).

Step (3) is justified by the following argument. By construction, V is optimal for the linear program

$$\begin{aligned} &\text{minimize} && c^{\mathrm{T}}x, \\ &\text{subject to} && (p^{i_j})^{\mathrm{T}}x \pm s_j = \|p^{i_j}\|^2, \quad j = 1, 2, \ldots, n, \\ & && s_j \geq 0, \end{aligned}$$

where the $\pm$ sign in front of s_j is understood to be chosen so that the resulting feasible region is the polytope P_c. By the Strong Duality theorem, there exists an optimal solution y^* to the dual program:

$$\begin{aligned} &\text{maximize} && \sum_{j=1}^{n} \|p^{i_j}\|^2 y_j, \\ &\text{subject to} && B^{\mathrm{T}}y = c, \\ & && \pm y_j \leq 0, \quad j = 1, 2, \ldots, n. \end{aligned}$$

In particular, $y^* = B^{-\mathrm{T}}c$ must be dual feasible. This condition determines the signs of the slacks s_j in the primal problem and hence the inequality directions as in Step (3).

4. Aggregate properties of random polytopes

Aggregate properties, such as the expected number of vertices, for the class of random polytopes we consider are of interest for at least two reasons. First, prior to a possible use of rejection filters to introduce special structure and thereby complex constraint dependencies, most random mathematical programming test problem generators fall within our definition for some implied distribution F. Aggregate properties of the feasible regions so generated can help explain and even predict algorithm performance on test problems. Second, to the extent that our class of random polytopes mimics the statistical characteristics of real world problems, we may be able to gain insight into the observed behavior of algorithms on real world problems. We may also begin to understand why the simplex method performs so phenomenally well.

Renyi and Sulanke [24] and Efron [8] derived formulas for the expected number of vertices of convex hulls of random points in R^2 and R^3. Cover and Efron [6] and Carnal [3] generalized these results to n-dimensions in the asymptotic case where the number of points increases without bound. The

results are not distribution-free. Moreover, they are not directly applicable to our case, since the process of forming a convex hull induces dependencies among the constraints; in fact, the faces of the polytopes so generated are, with probability one, simplices. Liebling [18] indirectly utilized these results in a dual framework, where points correspond to hyperplane constraints, to establish formulas for the expected number of vertices for random polytopes so induced under various special distributions for generating the parametric point set (uncorrelated multivariate normals and so on). Unfortunately, the dual framework used forced a selection of the polytope covering the origin, and the mean number of vertices is therefore not only distribution dependent but peculiar to a biased polytope selection. Schmidt and Mattheiss [27, 28] performed extensive simulation studies on aggregate properties of random polytopes under various distributional assumptions, but, again, the polytopes generated were chosen to cover the origin. Finally, Goudsmidt [11] and Miles [21] considered aggregate properties of random polygons and polyhedra in R^2 and R^3, respectively, as the number of lines and planes dividing space increased without bound. They found that the average polygon, for example, was a quadralateral. Our results generalize these asymptotic properties to n dimensions and, moreover, give aggregate properties for an arbitrary number of dividing hyperplanes.

We now turn to deriving a formula for the expected number of k-dimensional faces of a random polytope in R^n whose generating distribution F_0 satisfies the assumptions of Section 2. The expected number of k-dimensional faces $E_k(m)$ is shown to be dependent only on the dimension of the face k considered and the number of constraints m. The result is, in this sense, distribution-free. Note that for $k = 0$ we get the expected number of vertices, for $k = 1$ the expected number of edges, for $k = n - 1$ the expected number of facets, and for $k = n$ the expected number of non-empty polytopes.[5]

Theorem 1. *The expected number of k-dimensional faces $E_k(m)$ of a random polytope in R^n generated by m constraints in accordance with assumptions* (a), (b) *is given by*

$$E_k(m) = 2^{n-k} \sum_{i=n-k}^{n} \binom{i}{n-k}\binom{m}{i} \Big/ \sum_{i=0}^{n} \binom{m}{i},$$

where $k = 0, 1, 2, \ldots, n$ and $m > n$. Moreover, $E_k(m) \leq \binom{n}{k}2^{n-k}$ for all $m > n$ and $\lim_{m\to\infty} E_k(m) = \binom{n}{k}2^{n-k}$.

Proof. We observed in Section 3 that assumption (b) leads to a simple random selection of a polytope from the population of polytopes formed by the m

[5] It has come to the authors' attention (July 1981) that S.E. Berenguer has independently stated and proven the results in Theorems 1 and 2 below in his unpublished Ph.D. Thesis, "Randomly generated linear programs", University of California, Berkeley, CA (June 1978). These results and others are summarized in I. Ader and S.E. Berenguer, "Random linear programs", ORC Report 81-4, University of California, Berekeley, CA (March 1981).

hyperplanes partitioning R^n. Hence the *expected* number of k-dimensional faces is the *average* number of k-dimensional faces in the population. Now each k-dimensional face of the partition is determined by $n-k$ hyperplanes all passing through a common point and in general intersection with probability one by assumption (a). Hence each k-dimensional face is shared by exactly 2^{n-k} polytopes of the partition with probability one. Let $X_i(k)$ be the number of k-dimensional faces in the ith polytope of the partition where $i = 1, 2, \ldots, N$, and let f_k be the number of k-dimensional faces in the partition. Then

$$E_k(m) = \sum_{i=1}^{N} X_i(k)/N = 2^{n-k} f_k / N.$$

But it is elementary that for partitions formed by m hyperplanes in R^n in general intersection, $f_k = \sum_{i=n-k}^{n} \binom{i}{n-k}\binom{m}{i}$ regardless of the particular configuration (see Buck [2]). Hence

$$E_k(m) = 2^{n-k} \sum_{i=n-k}^{n} \binom{i}{n-k}\binom{m}{i} \Big/ \sum_{j=0}^{n} \binom{m}{j}$$

since $f_n = N$, the number of polytopes in the partition. As for the limiting result, the numerator is of the order $2^{n-k}\binom{n}{n-k}\binom{m}{n}$ for large m while the denominator is of the order $\binom{m}{n}$. Their ratio then converges to $2^{n-k}\binom{n}{n-k}$. The convergence is from below since term by term

$$\left(\sum_{j=0}^{n} \binom{m+1}{j}\right)\binom{l}{n-k}\binom{m}{l} \leq \left(\sum_{j=0}^{n} \binom{m}{j}\right)\binom{l}{n-k}\binom{m+1}{l}$$

for each $l = n-k, n-k+1, \ldots, n$, with strict inequality holding for at least the term with $l = n-k$. Dividing through by $\left(\sum_{j=0}^{n}\binom{m+1}{j}\right)\left(\sum_{j=0}^{n}\binom{m}{j}\right)$ and summing on i from $n-k$ through n, we get

$$\sum_{i=n-k}^{n} \binom{i}{n-k}\binom{m}{i} \Big/ \sum_{j=0}^{n}\binom{m}{j} \leq \sum_{k=n-k}^{n} \binom{i}{n-k}\binom{m+1}{i} \Big/ \sum_{j=0}^{n}\binom{m+1}{j}$$

or $E_k(m) \leq E_k(m+1)$ and the result follows inductively.

The bound on the expected number of k-dimensional faces in Theorem 1 suggests that the worst average case for a random polytope is monotonically approached as the number of constraints grows large. This case is topologically equivalent to an n-dimensional hypercube. In particular, the *average* number of vertices for randomly generated feasible regions is bounded from above by 2^n. This is dramatically less than the theoretically attainable upper bound for the number of vertices of a convex polytope of m constraints $V_{\max}$ where:

$$V_{\max} = \binom{m - [\frac{1}{2}(n+1)]}{m-n} + \binom{m - [\frac{1}{2}(n+2)]}{m-n}$$

with $[x]$ the greatest integer less than or equal to x (see McMullen [20]). To give a numerical illustration, consider the case of $n = 7$ and $m = 19$ constraints (Table

Table 1. Means and bounds for random polytopes

n	m	$E_0(m)$	2^n	V_{max}	$\binom{m}{n}$
7	19	68.48	128	910	50,388

1). The average number of vertices is an order of magnitude less than the maximal number. This suggests, for example, that complete enumeration of extreme points for fairly large linearly constrained problems is not computational unreasonable. Incidentally, the fact that the average number of vertices converges to 2^n as the number of constraining hyperplanes goes to infinity can be seen directly. We established in Section 2 a one-to-one correspondence between the vertices of a partition and a subset of the regions within it. In the limit, the correspondence is over all regions. Since each vertex is shared by exactly 2^n regions, the result follows immediately.

One of the more promising ways of reducing the computational complexity of large scale mathematical programs is to delete redundant constraints prior to the search for an optimum. The resulting reduction in size of the system of linear inequalities can significantly reduce subsequent computational effort. The following theorem indicates that for the class of randomly generated feasible regions we consider, the improvement in efficiency gained by this approach can be expected to increase significantly as the size of the problem increases. That is, the proportion of redundant constraints increases quite rapidly in the number of constraints m.

Theorem 2. *Let P in R^n be a random linearly constrained feasible region generated by m constraints in accordance with assumptions* (a), (b).

(a) *The probability π_m that a constraint is redundant is given by*

$$\pi_m = 1 - \left(2\sum_{i=0}^{n-1}\binom{m-1}{i}\Big/\sum_{i=0}^{n}\binom{m}{i}\right)$$

and

$$\lim_{m\to\infty} \pi_m = 1.$$

(b) *The expected number of non-redundant constraints $C_n(m)$ is given by*

$$C_n(m) = 2m\sum_{i=0}^{n-1}\binom{m-1}{i}\Big/\sum_{i=0}^{n}\binom{m}{i}.$$

Moreover,

$$\lim_{m\to\infty} C_n(m) = 2n.$$

Proof. (a) We want the probability that a randomly chosen constraint is redundant in a randomly selected polytope. Considering the complementary event, and reversing the natural order of selection, we may view the problem in the following way. First, tag a randomly selected constraint, then determine the probability that a randomly selected region is bounded by that constraint. The number of regions bordered by a given constraint is twice the number of new regions created in adding that constraint. But the number of regions created by adding the mth constraint is the number of regions the mth hyperplane is divided into by its intersection with the $m-1$ hyperplanes already in place. But this is nothing more than the number of regions an $(n-1)$-dimensional space is partitioned into by $m-1$ $(n-2)$-dimensional hyperplanes; that is, $f_{n-1} = \sum_{i=0}^{n-1}\binom{m-1}{i}$. Twice this number divided by the number of regions formed in R^n by the m constraints is then the probability that the randomly selected constraint is non-redundant and the result follows. As for the limiting result, $1-\pi_m$ for large m is given by $2\binom{m-1}{n-1}/\binom{m}{n}$ which converges to 0 as $m \to \infty$.

(b) Let Z be the number of non-redundant constraints for a randomly generated feasible region P and let $Z_i = 0$ or 1 as constraint i is redundant or non-redundant for P. Then

$$C_n(m) = EZ = E\sum_{i=1}^{m} Z_i = mEZ_i = m(1-\pi_m)$$

and the first result follows. The limiting result follows from the fact that $C_n(m) \sim 2m\binom{m-1}{n-1}/\binom{m}{n}$ which converges to $2n$ as $m \to \infty$.

5. The class of Euclidean random polytopes

We have until now purposely left the class of random polytopes considered as general as possible. This maximizes the range of real world problems and randomly generated test problems that fall within our classification. Turning now from the descriptive point of view to the normative, although the procedure for deciding constraint inequality directions is specific, the distribution F governing the positions of the hyperplanes has until now been assumed continuous but otherwise arbitrary.

From the standpoint of generating test problems, F should clearly be chosen to agree closely with the distribution exhibited by the class of real world problems for which the algorithm tested will ultimately be applied. The difficulty is that very little is known about the distribution of real world problems. The approach often then used may be described as the generation of hyperplane positions in as unbiased a way as possible followed by a filter that rejects all sampled feasible regions that fail to satisfy various structural properties that the real world problem class is known to possess. Our focus in this section is on how to choose F so that the candidate population of feasible regions generated is 'unbiased'.

In the absence of any structural information, we interpret bias as (Euclidean) geometric bias and our search is then for a distribution F that is geometrically random. For example, the likelihood of obtaining a plane nearly orthogonal to the x_1-axis should be the same as that of a plane nearly orthogonal to the x_2-axis (rotational symmetry). Also, the likelihood of the plane being near the origin should be the same as the likelihood of it being one unit away (translational symmetry). In short, all hyperplane positions and orientations should be equally likely. This property of space homogeneity may be more generally expressed in the following way: *The probability of all geometrically equivalent events*[6] *should be the same.* We can rephrase this definition of being geometrically random for the parameter space as follows.

Definition. A random hyperplane parameterized by p will be said to be *Euclidean random* if the probability measure M generating p is invariant under transformations of the parameter space induced by Euclidean transformations of the solution space.

Various authors have already noted the importance of requiring the distribution over p to be rotation symmetric (see, e.g., Liebling [18], Schmidt and Mattheiss [27], Van Dam and Telgen [7]). In addition, Van Dam and Telgen [7] noted the significant geometric regularity induced by an uncritical introduction of randomness such as choosing the constraint coefficients independently and uniformly over the interval [0, 1] as many mathematical programming test generators do (see, e.g., [1, 4, 16, 17, 29]). However, to our knowledge, no one has noted in this context the equally important requirement of translational symmetry. For example, selecting p uniformly within a hypersphere has rotational symmetry but fails to have translational symmetry.

The following theorem explicitly characterizes a probability density function that is invariant under translations and rotations. The theorem is a generalization to n dimensions of a result proven for $n = 2$ and 3 by Kendall and Moran [13]. Incidentally, Poincaré [22] and Polya [23] showed that the Euclidean invariant measure is unique for the case $n = 2$ and 3.

Theorem 3. *The density function* $f(p) = 1/\|p\|^{n-1}$ *defined over* $B = \{p : \|p\| \leq 1\} \subseteq R^n$ *induces a probability measure* $M(E) = \int_E f(p)\,dp$ *for measurable* $E \subseteq B$ *that is invariant under transformations induced by Euclidean transformations of the solution space.*

Proof. Since any Euclidean transformation can be expressed as the composition of a rotation and a translation, it sufficies to consider them separately. Let S be a

[6] An event in our context is a (measurable) collection of hyperplanes in solution space R^n. Two events are *geometrically equivalent* if there is a rotation and translation of the solution space that makes one set of hyperplanes coincide with the other.

rotation or translation of hyperplanes in R^n and let $E' = T(E)$ be the image of $E \subseteq B$ under the invertible transformation T in parameter space induced by S. We want $f(p)$ such that $M(E') = M(E)$. But $M(E') = \int_E f(p')|J_T(p)|\,dp$ by the change of variable theorem where

$$J_T(p) = \det \frac{\partial(p'_1, \ldots, p'_n)}{\partial(p_1, \ldots, p_n)}$$

is the Jacobian of T and $p' = T(p)$. Hence we require $f(p)$ to be such that $\int_E f(p)\,dp = \int_E f(p')\,|J_T(p)|\,dp$ for all measureable sets E. Then we must have $f(p) = f(p')\,|J_T(p)|$ or, equivalently, we get the following necessary and sufficient condition on $f(p)$:

$$\frac{f(p)}{f(p')} = J_T(p).$$

(Rotations) Since a rotation in solution space induces a rotation in parameter space, $p' = Tp$, where T is orthogonal. Then

$$\frac{f(p)}{f(p')} = \frac{\|p'\|^{n-1}}{\|p\|^{n-1}} = \frac{\|Tp\|^{n-1}}{\|p\|^{n-1}} = \|T\| = 1 = J_T(p).$$

(Translations) Letting the translated $x' = x + c$, we find that the new $p' = (1 + c^{\mathrm{T}}p/\|p\|^2))p$, so that

$$J_T(p) = \frac{\partial(p'_1, p'_2, \ldots, p'_n)}{\partial(p_1, p_2, \ldots, p_n)} =$$

$$= \begin{bmatrix} 1 + \frac{\|p\|^2[c^{\mathrm{T}}p + c_1p_1] - 2(c^{\mathrm{T}}p)p_1^2}{\|p\|^4} & \frac{\|p\|^2c_2p_1 - 2(c^{\mathrm{T}}p)p_1p_2}{\|p\|^4} & \cdots & \frac{\|p\|^2c_np_1 - 2(c^{\mathrm{T}}p)p_1p_n}{\|p\|^4} \\ \frac{\|p\|^2c_1p_2 - 2(c^{\mathrm{T}}p)p_1p_2}{\|p\|^4} & 1 + \frac{\|p\|^2[c^{\mathrm{T}}p + c_2p_2] - 2(c^{\mathrm{T}}p)p_2^2}{\|p\|^4} & \cdots & \frac{\|p\|^2c_np_2 - 2(c^{\mathrm{T}}p)p_2p_n}{\|p\|^4} \\ \vdots & \vdots & & \vdots \\ \frac{\|p\|^2c_1p_n - 2(c^{\mathrm{T}}p)p_1p_n}{\|p\|^4} & \frac{\|p\|^2c_2p_n - 2(c^{\mathrm{T}}p)p_2p_n}{\|p\|^4} & \cdots & 1 + \frac{\|p\|^2[c^{\mathrm{T}}p + c_np_n] - 2(c^{\mathrm{T}}p)p_n^2}{\|p\|^4} \end{bmatrix}$$

Subtracting (p_i/p_{i+1}) times row $i+1$ from row i for $i = 1, 2, \ldots, n-1$, and factoring $k = 1 + (c^{\mathrm{T}}p/\|p\|^2)$ out of the first $n-1$ rows yields

$$k^{n-1} \begin{bmatrix} 1 & -\frac{p_1}{p_2} & 0 & \cdots & 0 \\ 0 & 1 & -\frac{p_2}{p_3} & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{p_{n-1}}{p_n} \\ \frac{\|p\|^2c_1p_n - 2(c^{\mathrm{T}}p)p_1p_n}{\|p\|^4} & \frac{\|p\|^2c_2p_n - 2(c^{\mathrm{T}}p)p_2p_n}{\|p\|^4} & \frac{\|p\|^2c_3p_n - 2(c^{\mathrm{T}}p)p_3p_n}{\|p\|^4} & \cdots & 1 + \frac{\|p\|^2[c^{\mathrm{T}}p + c_np_n] - 2(c^{\mathrm{T}}p)p_n^2}{\|p\|^4} \end{bmatrix}$$

Expanding by the last row gives

$$|J_T(p)| = \left| k^{n-1} \left[\frac{\|p\|^2 c_1 p_1 - 2(c^T p)p_1^2}{\|p\|^4} + \frac{\|p\|^2 c_2 p_2 - 2(c^T p)p_2^2}{\|p\|^4} + \cdots \right.\right.$$
$$\left.\left. + 1 + \frac{\|p\|^2[c^T p + c_n p_n] - 2(c^T p)p_n^2}{\|p\|^4} \right] \right|$$
$$= \left| k^{n-1} \left[\frac{\|p\|^2 c^T p - 2c^T p \|p\|^2 + \|p\|^2 c^T p}{\|p\|^4} + 1 \right] \right| = |k^{n-1}|$$
$$= \left| 1 + \frac{c^T p}{\|p\|^2} \right|^{n-1} = \frac{\|p'\|^{n-1}}{\|p\|^{n-1}} = \frac{f(p)}{f(p')}$$

so that the measure is preserved.

Theorem 3 characterizes the probability distribution that describes Euclidean random hyperplanes. The next theorem spells out a Monte Carlo procedure for generating hyperplanes that satisfy such a distribution.

Theorem 4. *A Euclidean random hyperplane parameterized by p meeting the unit hypersphere B can be obtained by selecting the direction d of p as a point uniformly distributed on B and independently and uniformly selecting its length r along the line segment in B in the direction d.*

Proof. Since any measurable set E can be represented as a countable union of differences of cones, it suffices to show that the theorem is true for a cone E, with vertex at the origin and which is spanned by a finite set of vectors.

Let $e_1, e_2, \ldots, e_n$ be unit vectors along the edges of the cone E. Let

$$D = \left\{ d \mid d = \sum_{i=1}^{n} \lambda_i e_i, \quad \lambda \geq 0, \quad \|\lambda\| = 1 \right\},$$

the set of feasible directions of unit length within E. For $d \in D$, let $r(d) = \max\{r \mid r > 0, rd \in E\}$. Note that the length chosen along d is just $\|p\|$, so that, by a result in [5], we can change to a spherical coordinate system determined by d and r to get

$$M(E) = \int_E \frac{1}{\|p\|^{n-1}} \, dp = \int_D \int_0^{r(d)} \frac{1}{r^{n-1}} r^{n-1} \, d(d) \, dr = \int_D \int_0^{r(d)} d(d) \, dr.$$

Hence the joint density $f(r, d) = 1$ for all r and d. Now the conditional density $f(r \mid d) = f(r, d)/f(d) = 1$ since $f(d) = \int_0^1 f(r, d) \, dr = 1$. The conditional and marginal densities being constant, the result follows.

The Monte Carlo procedure outlined in Theorem 4 requires a random deviate uniformly distributed on the unit sphere (i.e., a random direction). A number of efficient techniques for its generation exist, the simplest of which is to generate n independent normal random variables and divide by their norm (see Knuth [15]).

6. Conclusions

In this paper, we dealt with three issues. First, given a multivariate distribution satisfying certain general assumptions, how does one randomly select a polytope from the regions formed by the generated matrix? We showed that such a selection could be accomplished in an efficient manner by using a procedure exploiting certain primal-dual relationships. Second, we examined summary measures for random polytopes, and showed that the average polytope exhibits the characteristics of an n-dimensional hypercube. This means that, on the average, the number of vertices of a polytope is substantially less than is given by McMullen's attainable upper bound, and may help explain the often observed real-life efficiency of the Simplex algorithm. Third, we showed how to choose a multivariate distribution so as to generate polytopes that satisfy geometric notions of randomness. These notions rule out, for example, using the uniform distribution, and require a density exhibiting both rotational and translational invariance.

References

[1] D. Avis and V. Chvátal, "Notes on Bland's pivoting rule", *Mathematical Programming Study* 8 (1978) 24–34.
[2] R.C. Buck, "Partition of space", *The American Mathematical Monthly* 50 (1943) 541–544.
[3] H. Carnal, "Die konvexe Hülle von N rotationssymmetrische verteilten Punkten", *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 15 (1970) 168–176.
[4] A. Charnes, W. Raike, J.D. Stutz and A.S. Walters, "On generation of test problems for linear programming codes", *Communications of the ACM* 17 (1974) 583–586.
[5] J.R. Charnetski and R.M. Soland, "Statistical measures for linear functions on polytopes", *Operations Research* 24 (1976) 201–204.
[6] T.M. Cover and B. Efron, "Geometrical probability and random points on a hypersphere", *Annals of Mathematical Statistics* 38 (1967) 213–220.
[7] W.B. van Dam and J. Telgen, "Randomly generated polytopes for testing mathematical programming algorithms", Report No. 7929/0, Erasmus University, Rotterdam (1979).
[8] B. Efron, "The convex hull of a random set of points", *Biometrika* 52 (1965) 331–343.
[9] W. Feller, *An introduction to probability theory and its applications* 2, 2nd Ed. (Wiley, New York, 1971) 11–15.
[10] D. Gale, "How to solve linear inequalities", *The American Mathematical Monthly* 76 (1969) 589–599.
[11] S. Goudsmidt, "Random distributions of lines in a plane", *Reviews of Modern Physics* 17 (1945) 494–498.
[12] B. Grünbaum, *Convex polytopes* (Wiley, New York, 1967).
[13] M.G. Kendall and P.A.P. Moran, *Geometrical probability* (Griffin, London, 1963).
[14] V. Klee and G.J. Minty, "How good is the simplex algorithm?", in: O. Shisha, ed., *Proceedings of the Third Symposium on inequalities* (Academic Press, New York, 1972) pp. 159–175.
[15] D.E. Knuth, *The art of computer programming 2: Seminumerical algorithms* (Addison-Wesley, Reading, MA, 1969) 103–113.
[16] H. W. Kuhn and R.E. Quandt, "An experimental study of the simplex method", in: N.C. Metropolis, A.H. Taub, J. Todd and C.B. Tompkins, eds., *Experimental arithmetic, high speed computing and mathematics*, Proceedings of the 15th Symposium in Applied Mathematics of the AMS (American Mathematical Society, Providence, RI, 1963) pp. 107–124.

[17] C.H. Layman and R.P. O'Neill, "A study of the effect of LP parameters on algorithm performance", in: W.W. White, ed., *Computers and mathematical programming* (U.S. Government Printing Office, Washington, DC, 1978) pp. 251–260.

[18] T.M. Liebling, "On the number of iterations of the simplex method", in: R. Henn, H.P. Kunzi and H. Schubert, eds., *Operations Research Verfahren* XVII (1972) 248–264.

[19] J.H. May and R.L. Smith, "Random polytopes: Their definition, generation, and aggregate properties", Technical Report 80-6, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI (1980); also: Working Paper No. 416 Graduate School of Business, University of Pittsburgh, Pittsburgh, PA (1980).

[20] P. McMullen, "The maximum number of faces of a convex polytope", *Mathematika* 17 (1970) 179–184.

[21] R.E. Miles, "Random polygons determined by random lines in a plane, I, II", *Proceedings of the National Academy of Sciences of the USA* 52 (1964), 901–907, 1157–1160.

[22] H. Poincaré, *Calcul des probabilités*, 2nd ed. (Gauthier-Villars, Paris, 1912).

[23] G. Polya, "Über geometrische Wahrscheinlichkeiten", *Sitzungsberichte der Akademie der Wissenschaften Wien* 126 (1917) 319–328.

[24] A. Renyi and R. Sulanke, "Über die konvexe Hülle von n zufällig gewählten Punkten, I, II", *Zeitschrift für Wahrscheinlichkeitshteorie und Verwandte Gebiete 2*, 3 (1963, 1964) 75–84, 138–148.

[25] R.T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, NJ, 1970).

[26] L. Schläfli, "Theorie der vielfachen Kontinuität", *Denkschriften der Schweizerische naturforschende Gesellschaft* 38 (1901) 1–237.

[27] B.K. Schmidt and T.H. Mattheiss, "The probability that a random polytope is bounded", *Mathematics of Operations Research* 2 (1977) 292–296.

[28] B.K. Schmidt and T.H. Mattheiss, "Computational results on an algorithm for finding all vertices of a poltyope", *Mathematical Programming* 18 (1980) 308–329.

[29] S. Zionts, "Some empirical tests of the criss-cross method", *Management Science* 19 (1972) 406–410.