RANDOMLY GENERATED POLYTOPES FOR TESTING MATHEMATICAL PROGRAMMING ALGORITHMS

W.B. van DAM

Department of Industrial Engineering, University of Technology, Eindhoven, The Netherlands

J.B.G. FRENK

Econometric Institute, Erasmus University, Rotterdam, The Netherlands

J. TELGEN

Department of Applied Mathematics, Rabobank Nederland,
Zeist, The Netherlands

Randomly generated polytopes are used frequently to test and compare algorithms for a variety of mathematical programming problems. These polytopes are constructed by generating linear inequality constraints with coefficients drawn independently from a distribution such as the uniform or the normal.

It is noted that this class of 'random' polytopes has a special property: the angles between the hyperplanes, though dependent on the specific distribution used, tend to be equal when the dimension of the space increases.

Obviously this structure of 'random' polytopes may bias test results.

Key words: Random Polytopes, Linear Inequalities, Testing and Comparing Algorithms.

1. Introduction

Testing and comparing algorithms still is an aspect of mathematical programming that has received little attention. In spite of efforts by the Committee On Algorithms (COAL) of the Mathematical Programming Society, a generally accepted treatment of this topic has not yet been formulated (see e.g. [8, 11, 12]). One of the difficulties involved in devising a uniform and sound methodology for testing and comparing algorithms stems from the test problems to be used. The pros and cons of 'randomly generated' versus 'real world' test problems have been the subject of many lengthy and vivid discussions. Here we do not take a stand in that argument, but point out a special property of some 'randomly generated' problems, that may bias test results.

We consider convex polytopes P, constructed as the intersection of linear halfspaces

$$\alpha_i x \leq b_i, \quad i = 1, 2, \ldots, m,$$

i.e.

$$P = \{x \in \mathbb{R}^n \mid \alpha_i x \leq b_i, \quad i = 1, \ldots, m\}$$

with $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{R}^n$.

Such polytopes are used in testing and comparing linear, quadratic, integer and some nonlinear programming algorithms. Usually the coefficients a_{ij} are generated by drawing from either a uniform or from a normal distribution. [4, 7, 9, 10, 13, 16, 18, 20, 21, 23].

In Section 2 we list some intuitive evidence that the angles between the hyperplanes corresponding to the constraints are not uniformly distributed. We indicate reasons why these angles tend to be equal. Moreover we show that this behaviour becomes stronger as the dimension of the space increases. These arguments are formalized for a broad class of probability distributions in Section 3. There we derive analytical expressions for the limit of the mean and variance of the cosine of the angle between two hyperplanes. In Section 4 we consider the rate at which the generated angles approach the asymptotic results. This analytically derived rate of convergence is $O(1/n)$. The final section contains a discussion on the implications of our results.

Since the difficulties arise for all examined probability distributions we advise care in drawing conclusions based on empirical results on 'random' polytopes.

2. Some preliminary observations

No matter which definition of a 'random' polytope is adopted (see e.g. [6, 15, 16, 17, 21]) the randomness of this polytope should also be reflected in the distribution of the angles between the hyperplanes determining the polytope.

It seems reasonable to demand that these angles behave according to a uniform distribution on the relevant interval. As it turns out, such properties are not exhibited if the coefficients are drawn from most simple distributions. To make this intuitively clear we will illustrate this point for the uniform distribution and the normal distribution.

First consider the uniform distribution on the interval (p, q) with $0 < p < q$. Independently drawing coefficients $a_{i1}, \ldots, a_{in}$ from a uniform distribution is equivalent to selecting points α_i from a hypercube. Each point determines the direction of the normal vector to a constraint. Therefore the angle between two constraints (hyperplanes) is determined by two points. As illustrated in Fig. 1 only small angles are possible; these are also not uniformly distributed.

Now consider the normal distribution $N(0, 1)$; In this case we can get every angle, but still the size of the angle is not uniformly distributed for $n \geq 3$. To make this intuitively clear consider Fig. 2 $(n = 3)$. Without loss of generality (use the independence assumption) we may assume the first normal vector we have generated is $\alpha_0 \in \mathbb{R}^3$. The set of (normalized) vectors that make an angle φ with α_0 is a circle of which the radius depends on φ. Since the probability density
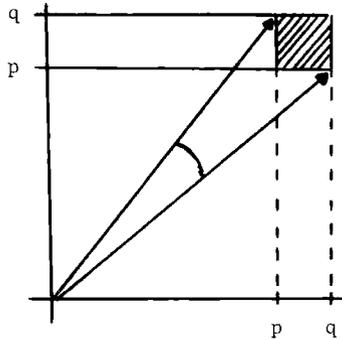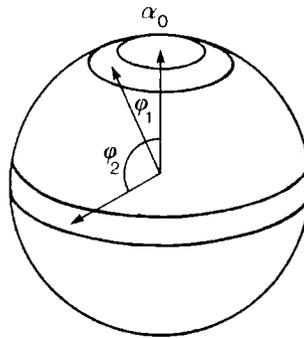
Fig. 1.

Fig. 2.

function on the sphere is uniform the probability of an angle $\varphi \in (\varphi_1 - \epsilon, \varphi_1 + \epsilon)$ is not equal to the probability of an angle $\varphi \in (\varphi_2 - \epsilon, \varphi_2 + \epsilon)$.

Anderson [2, p. 64] states the density f_n of the cosine of the angle between two constraints (hyperplanes with dimension $n - 1$)

$$f_n(x) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}(n-1))} (1 - x^2)^{(n-3)/2} \quad \text{for } n \geq 2.$$

This implies (see [2, p. 65]) that

$$E(\cos(\text{angle})) = 0$$

and

$$\text{var}(\cos(\text{angle})) = \frac{1}{n} \quad \forall n \geq 2.$$

3. Mean and variance of the angles between hyperplanes

In this section we derive the mean and variance of the angles between hyperplanes over *all* possible polytopes generated in the way described in the

previous section. It is important to note that we do not attempt to determine these moments for the angles between hyperplanes in *one* polytope. This allows us to work with stochastically independent realizations of cos c_{kl}, where we denote by c_{kl} the angle between the hyperplanes corresponding to constraints k and l. (In one polytope the angles c_{kl}, c_{kh} and c_{hl} are obviously dependent.) We consider the angle between hyperplanes formed by generating i.i.d. coefficients a_{kj} and a_{lj} from the same probability space Ω (with probability measure P), *where the abbreviation i.i.d. stands for independent identically distributed.*

To indicate realizations of the random variables a_{ij} $(i = k, l)$ we will use $a_{ij}(\omega)$ and to indicate the dependence of c_{kl} on n we will use $c_{kl;n}$. Thus,

$$\cos c_{kl;n}(\omega) = \frac{\sum_{j=1}^{n} a_{kj}(\omega) a_{lj}(\omega)}{\sqrt{\sum_{j=1}^{n} a_{kj}^2(\omega) \sum_{j=1}^{n} a_{lj}^2(\omega)}}. \tag{3.1}$$

Define

$$\beta_{ij} :\overset{\triangle}{=} E(a_{ij}^2) = \int_{\Omega} a_{ij}^2(\omega) \, dP(\omega), \quad i = k, l$$

and

$$\beta_{kl;j} :\overset{\triangle}{=} E(a_{kj} a_{lj}) = \int_{\Omega} a_{kj}(\omega) a_{lj}(\omega) \, dP(\omega). \tag{3.2}$$

Since $\{a_{ij}\}_{j \in \mathbb{N}, \, i = k, l}$ are i.i.d. random variables we can drop the subscript j in (3.2).

Lemma 1. *For i.i.d. random variables a_{ij} we have:*

$$P \left(\lim_{n \to \infty} \cos c_{kl;n}(\omega) = \frac{\beta_{kl}}{\sqrt{\beta_k \cdot \beta_l}} \right) = 1. \tag{3.3}$$

Proof. Since the a_{ij}'s are i.i.d. random variables we may apply the strong law of large numbers [5, p. 250].

Hence for

$$S_{ni}(\omega) = \sum_{j=1}^{n} a_{ij}^2(\omega) \quad (i = k, l) \quad \text{and} \quad S_{nkl} = \sum_{j=1}^{n} a_{kj}(\omega) a_{lj}(\omega)$$

we have

$$P \left(\lim_{n \to \infty} \frac{1}{n} S_{ni}(\omega) = \beta_i \right) = 1, \quad i = k, l,$$

$$P \left(\lim_{n \to \infty} \frac{1}{n} S_{nkl}(\omega) = \beta_{kl} \right) = 1.$$

This implies (3.3). □

As a corollary we have:

Lemma 2. *For i.i.d. random variables a_{ij} we have:*

$$\lim_{n \to \infty} E(\cos c_{kl;n}) = \frac{\beta_{kl}}{\sqrt{\beta_k \cdot \beta_l}},$$

$$\lim_{n \to \infty} \text{Var}(\cos c_{kl;n}) = 0.$$

Proof. Using Lebesgue's bounded convergence theorem [5, p. 180] and the fact that $|\cos c_{kl;n}| \leq 1 \ \forall n \in \mathbb{N}$ yields:

$$\lim_{n \to \infty} E(\cos c_{kl;n}) = E\left(\frac{\beta_{kl}}{\sqrt{\beta_k \cdot \beta_l}}\right) = \frac{\beta_{kl}}{\sqrt{\beta_k \cdot \beta_l}}.$$

Since $\text{Var}(\cos c_{kl;n}) = E(\cos^2 c_{kl;n}) - (E(\cos c_{kl;n}))^2$ we can use a similar reasoning to get the second result. □

Note that these lemmas do not depend on the specific probability distribution F. The only assumption we have to make is $\beta_i < \infty$ ($i = k, l$) to avoid pathological cases.

Finally we mention a result that follows trivially from Lemma 1.

$$P\left(\lim_{n \to \infty} c_{kl;n} = \arccos \frac{\beta_{kl}}{\sqrt{\beta_k \cdot \beta_l}}\right) = 1.$$

4. Rate of convergence

In the last section we derived analytical expressions for the limit of the mean and variance of the cosine of the angle $c_{kl;n}$ in case $n \to \infty$. In this section we are interested in the rate of convergence. This rate of convergence yields a notion of the discrepancy between mean and variance and their respective limits.

The next lemma (which is a special case of Theorem 1) gives us this rate in case the items are uniformly distributed.

Lemma 3. *Suppose $\{a_{ij}\}_{j \in \mathbb{N}, i=k, l}$ are drawn independently from a uniform distribution on the interval (p, q). This implies:*

$$E(\cos c_{kl;n}) = \frac{3(q^2 - p^2)(q + p)}{4(q^3 - p^3)} + O\left(\frac{1}{n}\right)$$

and　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(4.1)

$$\text{Var}(\cos c_{kl;n}) = O\left(\frac{1}{n}\right).$$

Proof. The probability density function of (a_{ij}) for every $j \in \mathbb{N}$ is given

by:

$$f(a_{ij}) = (q - p)^{-1}, \quad p < a_{ij} < q, \quad i = k, j.$$

Hence

$$\beta_k = \beta_l = \frac{q^3 - p^3}{3(q - p)}, \quad \beta_{kl} = \tfrac{1}{4}(q + p)^2$$

and applying Theorem 1 with

$$B(x_1, x_2, x_3) = \frac{x_3}{\sqrt{x_1 \cdot x_2}}$$

yields the stated result. $\square$

We now give the proof of the main theorem of this section.

Theorem 1. *Let $B : \mathbb{R}^3 \to \mathbb{R}$ be a bounded function with continuous first and second order partial derivatives in a neighbourhood of $\beta = (\beta_k, \beta_l, \beta_{kl})$ and let a_{ij} ($j = 1, \ldots, n$; $i = k, l$) be i.i.d. random variables drawn from a distribution with finite fourth moment, and let*

$$G_n(\omega) :\triangleq B \left(\frac{1}{n} S_{nl}(\omega), \frac{1}{n} S_{nk}(\omega), \frac{1}{n} S_{nkl}(\omega) \right).$$

Then

$$E(G_n) = B(\beta) + O\left(\frac{1}{n}\right)$$

and

$$\mathrm{Var}(G_n) = O\left(\frac{1}{n}\right).$$

Proof. Define $I = \{k, l, kl\}$ and

$$A_{n;i} :\triangleq \left\{ \omega \in \Omega \;\middle|\; \left| \frac{1}{n} S_{ni}(\omega) - \beta_i \right| > \epsilon \right\}, \quad i \in I.$$

Using Chebychev's inequality we have for any $\epsilon > 0$ an M such that

$$P(A_{n;i}) \le \frac{M}{\epsilon^2 \cdot n}, \quad i \in I.$$

Hence for $Z_n :\triangleq A_{n;k} \cup A_{n;l} \cup A_{n;kl}$

$$P(Z_n) \le \frac{3M}{\epsilon^2 \cdot n}.$$

Since G_n is uniformly bounded by a constant C we obtain

$$\left| \int_{Z_n} G_n(\omega) \, \mathrm{d}P(\omega) \right| \le \frac{3M \cdot C}{\epsilon^2 \cdot n}.$$

This implies $\forall n \in \mathbb{N}$

$$E(G_n) = \int_{Z_n^c} G_n(\omega)\, \mathrm{d}P(\omega) + \mathrm{O}\!\left(\frac{1}{n}\right) \quad \text{with } Z_n^c :\triangleq \text{the complement of } Z_n.$$

(4.2)

Choosing ϵ sufficiently small we can apply Taylor's formula for functions of several variables [3, p. 123]. Hence we obtain for all $\omega \in Z_n^c$ with $D_i :\triangleq$ partial derivative of a function with respect to the ith component and $\theta(\omega) :\triangleq$ some intermediate point

$$G_n(\omega) = B(\beta) + \sum_{i \in I} D_i B(\beta)\left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right)$$

$$+ \sum_{i \in I} \sum_{j \in I} D_i D_j B(\theta(\omega))\left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right)\left(\frac{1}{n} S_{nj}(\omega) - \beta_j\right). \tag{4.3}$$

Using the Cauchy–Schwarz inequality yields

$$\left|\int_{Z_n} \left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right) \mathrm{d}P(\omega)\right| \le$$

$$\le \left[P(Z_n) \int_{\Omega} \left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right)^2 \mathrm{d}P(\omega)\right]^{1/2} = \mathrm{O}\!\left(\frac{1}{n}\right), \quad i \in I.$$

Thus

$$\left|\int_{Z_n^c} \left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right) \mathrm{d}P(\omega)\right| = \left|\int_{Z_n} \left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right) \mathrm{d}P(\omega)\right| = \mathrm{O}\!\left(\frac{1}{n}\right)$$

and this implies

$$\left|\int_{Z_n^c} \sum_{i \in I} D_i B(\beta)\left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right) \mathrm{d}P(\omega)\right| = \mathrm{O}\!\left(\frac{1}{n}\right). \tag{4.4}$$

Also

$$\left|\int_{Z_n^c} D_{ij} B(\theta(\omega))\left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right)\left(\frac{1}{n} S_{nj}(\omega) - \beta_j\right) \mathrm{d}P(\omega)\right| \le$$

$$\le \left(\int_{Z_n^c} (D_{ij} B(\theta(\omega)))^2 \left(\frac{1}{n} S_{ni}(\omega) - \beta_i\right)^2 \mathrm{d}P(\omega)\right)^{1/2}$$

$$\times \left(\int_{Z_n^c} \left(\frac{1}{n} S_{nj}(\omega) - \beta_j\right)^2 \mathrm{d}P(\omega)\right)^{1/2}$$

$$\leq M_1 \left(\int\limits_{Z_n^c} \left(\frac{1}{n} S_{ni}(\omega) - \beta_i \right)^2 dP(\omega) \right)^{1/2} \left(\int\limits_{Z_n^c} \left(\frac{1}{n} S_{nj} - \beta_j \right)^2 dP(\omega) \right)^{1/2}$$

$$= O\left(\frac{1}{n} \right), \quad i, j \in I. \tag{4.5}$$

Combining (4.2), ... , (4.5) yields

$$E(G_n) = B(\beta) + O\left(\frac{1}{n} \right).$$

We can apply the same reasoning to B^2 and hence

$$E(G_n^2) = B^2(\beta) + O\left(\frac{1}{n} \right).$$

Then we obtain

$$\mathrm{Var}(G_n) = E(G_n^2) - (E(G_n))^2$$

$$= B^2(\beta) + O\left(\frac{1}{n} \right) - \left(B(\beta) + O\left(\frac{1}{n} \right) \right)^2$$

$$= O\left(\frac{1}{n} \right). \quad \square$$

To illustrate the implications of this theorem we performed some Monte-Carlo experiments. Coefficients a_{kj} and a_{lj} were drawn independently from a uniform distribution on $(-1, 1)$ and $\cos c_{kl}$ was computed according to (3.1) for $n = 10, 50$ and 100. Sample sizes were 1000 in each case. The results are shown in Fig. 3.
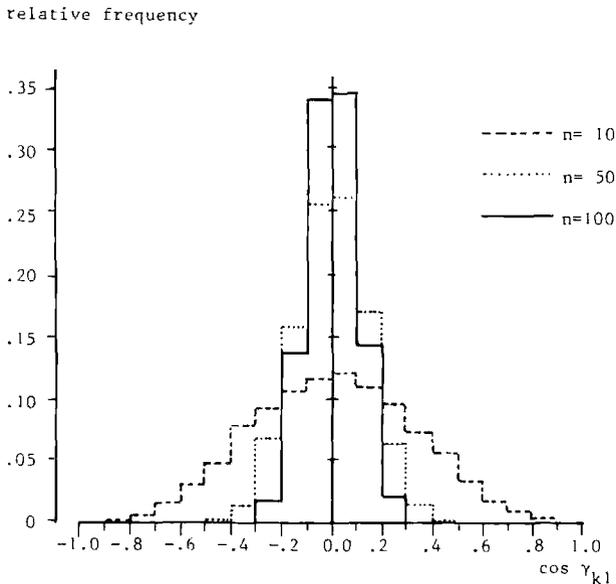

Fig. 3.

5. Concluding remarks

We have shown that in 'random' polytopes the angles between the hyper-planes exhibit some surprising asymptotic behaviour. If we were to use such polytopes in testing and comparing mathematical programming algorithms we might get biased results. A simple observation supports this claim: it is well known that Khachian's polynomial algorithm [13] and any other relaxation method [1, 19] require at most m iterations to find a solution to a system of m linear inequalities whenever the hyperplanes corresponding to these inequalities are mutually orthogonal. Now, although the performance of these algorithms is generally considered to be very poor, tests on certain polytopes, obtained by drawing from a uniform or normal distribution with mean zero, could show a very good performance.

It is important to note that it is possible to generate hyperplanes and con-sequently angles with a prespecified asymptotic cosine v. As an example consider the uniform distribution on $(p, p + 1)$, then from (4.1) we see we should take

$$p = \tfrac{1}{2} + \tfrac{1}{2} \sqrt{\frac{v}{3 - 3v}},$$

Since the angles tend to be equal for large n, the effects are strengthened in higher dimensions.

One might argue that we have considered hyperplanes instead of polytopes. In a strict sense that is true, but for most mathematical programming algorithms that distinction is not relevant. Angles between hyperplanes that do not intersect in a facet of the feasible region are just as important as the other angles in the nonfeasible stages of the algorithm (e.g. a phase I in LP). Therefore we can omit any reference to right hand side values and the determination of redundant constraints ([13] and [22]).

It is noted that the results of Section 4 also hold for discontinuous probability distributions with finite fourth moment. Therefore similar results apply to α_i's which are not full dense.

Unfortunately there does not seem to be a way out of difficulties sketched in this paper. Therefore we have to conclude with the remark that 'random' polytopes constructed by generating all coefficients of a linear inequality system by drawing from a simple distribution, such as the uniform or the normal, have special geometric structure. This obliges the user of such problems to take these properties into consideration.

References

[1] S. Agmon, "The relaxation method for linear inequalities", *Canadian Journal of Mathematics* 6 (1954) 382–392.

[2] T.W. Anderson, *An introduction to multivariate statistical analysis* (Wiley, New York, 1958).

[3] T.M. Apostol, *Mathematical Analysis* (Addison–Wesley, London, 1960).

[4] D. Avis and V. Chvatal, "Notes on Bland's pivoting rule", *Mathematical Programming Study* 8 (1978) 24–34.

[5] P. Billingsley, *Probability and measure* (Wiley, New York, 1979).

[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, "Some computational experiments with a primal-dual surrogate simplex method", Report 7828 Econometric Institute, Erasmus University (Rotterdam, 1978).

[8] R.S. Dembo and J.M. Mulvey, "On the analysis and comparison of mathematical programming techniques, *Proceedings of the bicentennial conference on mathematical programming* (Gaithersburg, MD, 1976).

[9] R.S. Garfinkel and P.L. Yu, "A primal–dual surrogate simplex algorithm", Working Paper 22, College of Business Administration, University of Tennessee (Knoxville, TN, 1975).

[10] S. Holm and D. Klein, "Size reduction of linear programs with special structure", Working Paper, Odense University (Odense, Denmark, 1975).

[11] J.P. Ignizio, "On the establishment of standards for comparing algorithm performance", *Interfaces* 2 (1971) 8–11.

[12] R.H. Jackson and J.M. Mulvey, "A critical review of methods for comparing mathematical programming algorithms and software 1951–1977", presented at TIMS XXIII. (Athens, 1977).

[13] M.H. Karwan, J. Telgen and S. Zionts, *Redundancy in mathematical programming* (Springer-Verlag, New York, Heidelberg, 1982).

[14] L.G. Khachian, "A polynomial algorithm in linear programming", *Soviet Mathematics Doklady* 20 (1979) 191–194.

[15] T.M. Liebling, "On the number of iterations of the simplex method", *Operations Research Verfahren* 23 (1972) 264–284.

[16] J.H. May and R.L. Smith, "Definition and generation of random polyhedra", Working Paper 390, Graduate School of Business, University of Pittsburgh (Pittsburgh, PA, 1980).

[17] 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).

[18] W.M. Michaels and R.P. O'Neill, "User's guide for LPGENR", Department of Computer Science, Louisiana State University (Baton Rouge, LA, 1975).

[19] T.S. Motzkin and I.J. Schoenberg, "The relaxation method for linear inequalities", *Canadian Journal of Mathematics* 6 (1954) 393–404.

[20] R.E. Quandt and H.W. Kuhn, "On upper bounds for the number of iterations in solving linear programs", *Operations Research* 12 (1964) 161–165.

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

[22] J. Telgen, *Redundancy and linear programs* (Mathematical Centre, Amsterdam, 1979).

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