Nr. 11a

Karl Heinz Borgwardt, Universität Augsburg,
Institut für Mathematik
Memminger Str. 6
D - 8900 Augsburg

Average Behaviour of the Simplex-Method: Some Improvements in the Analysis of the Rotation - Symmetry - Model

Acknowledgement

This paper reports on a talk which was given at the 12th International Symposium on Mathematical Programming in Boston, Massachusetts, USA, on August 6, 1985.
The paper explains some improvements as announced in the title. It does not contain proofs. Detailed and full-length proofs for these improvements and for the whole derivation of the analysis based on the rotation-symmetry-model as well as an historical overview can be found in the book "A Probabilistic Analysis of the Simplex-Method" which will be published by Springer-Verlag in 1986. It will appear in the series "Algorithms and Combinatorics" editet by Korte, Lovasz and Graham.

Abstract

During the last four years the polynomiality of the average number of pivot steps required by the Simplex-Method was proven under two different stochastic models: The Sign-Invariance-Model and the Rotation-Symmetry-Model. The Sign-Invariance-Model allows high probabilities for emptyness of the feasible region, for redundancy and for unboundedness of the objective and leads to very optimistic results on the average number of pivot steps. The Rotation-Symmetry-Model reflects more pessimistic assumptions on the Real-World-Distribution of Linear Programming Problems, because every generated problem has a feasible point. The evaluation of the average behaviour leads to a higher size of steps. It is an open question whether this size can still be diminished. The talk will report on some recent improvements in the analysis of that model (e.g. Phase I-results, sign-constraints) and some generalizations. In addition, some open problems shall be discussed.

Introduction and Motivation

We are interested in the number of pivot steps which are required by the Simplex-Method for the solution of problems

$$\begin{array}{lll} \text{Maximize} & v^T x & \text{where } x,\ v \in \mathbb{R}^n \\ \text{subject to} & Ax \leq b & A \in \mathbb{R}^{(m,n)},\ m \geq n. \end{array}$$

In particular, we concentrate on the following question: How many steps are required on the average if (m,n) is fixed and if v, A, b are somehow distributed random vectors?

In the past some successful approaches to such an analysis of the average behaviour could be done for parametric variants resp. for variants based on the parametric Simplex-Algorithm. Examples are

Shadow-Vertex-Algorithm (used by Borgwardt, Haimovich)
Lemke's Algorithm (used by Smale)
Lexicographic Lemke-Algorithm (used by Todd, Adler/Megiddo)
Constraint-By-Constraint-Method (Adler/Karp/Shamir).

Whereas we observe a great similarity in the choice of the variant, the stochastic models used for the analysis differ tremendously.

The most important models are

I a rotation-invariance model (used by Borgwardt 1977, 1978,81,82)

II a permutation-invariance model (used by Smale 1982, 1983 and Blair 1983)

III a sign-invariance model (used by Haimovich,Todd,Adler/Megiddo,Adler/Karp/Shamir all 83-84)

Polynomial upper bounds for the average number of steps could be derived under models I and III.

The results for III showed a much lower size (quadratic in the lower dimension, i.e. n^2 in our case).

The sign-invariance model can be explained as follows.

Consider that the input data v, A, b are fixed.
Now we generate 2^m problems out of the generic problem by left-multiplication of A resp. b with so-called sign matrices

$$S := \begin{pmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \end{pmatrix} \qquad S \in \mathbb{R}^{(m,m)}$$

a sign matrix is a diagonal matrix whose diagonal elements have the absolute value 1.

Since there are 2^m such sign-matrices, we obtain an instance class of 2^m problems

Maximize $v^T x$
subject to $SA\,x \leq Sb$.

Now the main assumption of that model is that all of these 2^m generated problems should have equal probability.

A geometrical interpretation gives additional insight. We have (in the generic problem) m linear inequalities

$$\begin{array}{l} a_1^T x \leq b_1 \\ \vdots \qquad \vdots \\ a_m^T x \leq b_m \end{array}$$

Now we assume that each of these inequalities can be "flipped" or inverted into $\geq$ independently and with probability $\frac{1}{2}$.

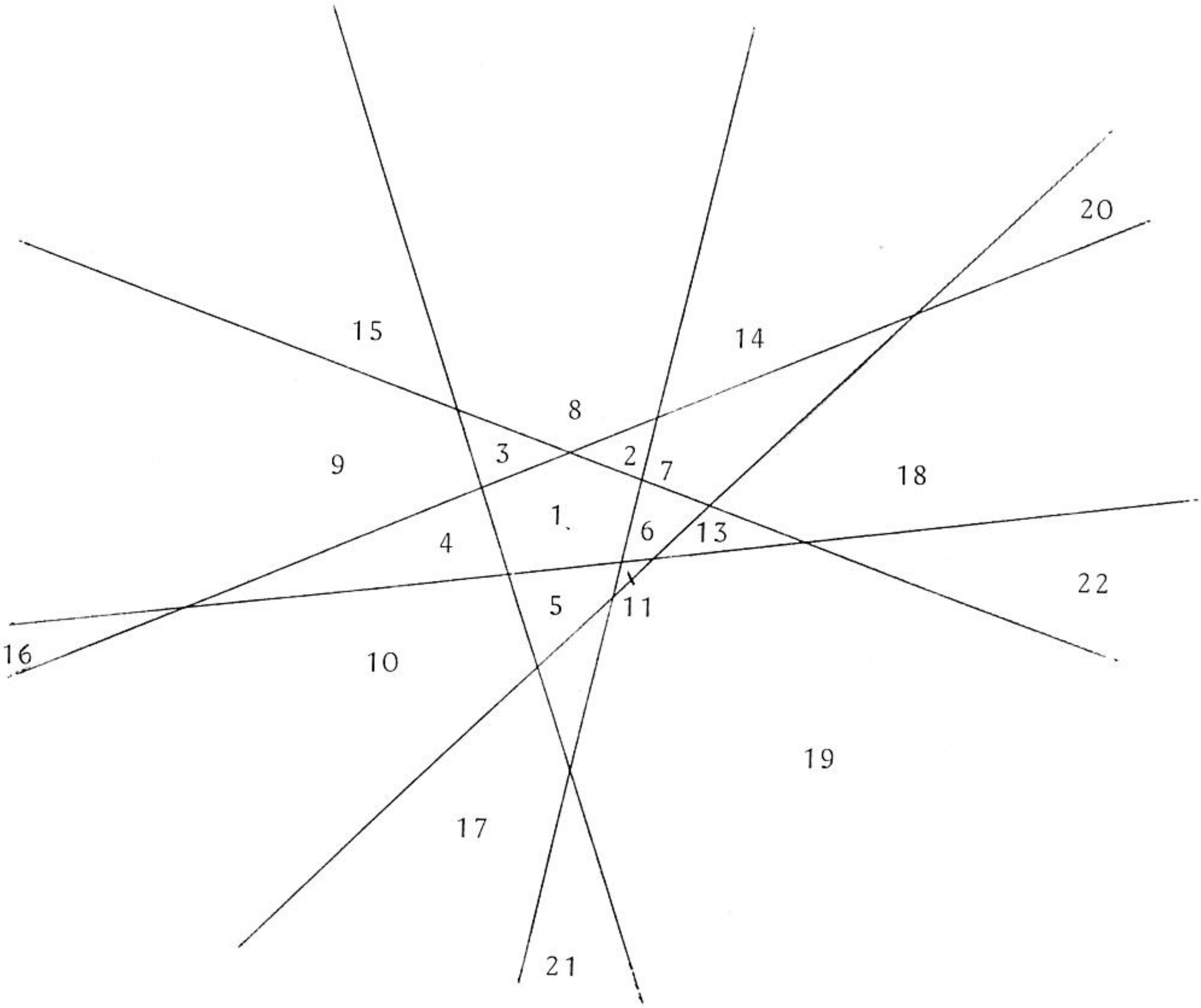
Again, we obtain a class of 2^m problems, each having a probability of 2^{-m} for realization.

It is clear that a feasible inner point of one of these problems cannot be feasible in any other problem, since at least one inequality must have been flipped.

And we observe that (under weak nondegeneracy assumptions) $\binom{m}{o} + \binom{m}{1} + \ldots + \binom{m}{n}$ problems have feasible points. But this number is (for $m >> n$) far less than 2^m. In the latter case the share of feasible problems becomes very small (tending to 0). Since all problems are counted in the analysis of the expected number of steps, we can expect a small size of steps. In the opposite case, namely $m \simeq n$, i.e., $m - n << n$, we have similar difficulties. Here most of the problems are unbounded and do not have an optimal solution.

As infeasibility the nonexistence of an optimal solution can be detected easily and quickly. So also this situation yields a very optimistic estimation.

The following figure shows the typical situation for $m = 6$ and $n = 2$. $2^6 = 64$ problems are generated, only 22 of them have feasible regions.



The main results under that model are due to Todd, Adler/Megiddo, Adler/Karp/Shamir and say that

<u>Theorem</u>

$E_{m,n}(s) \leq C \text{ Min } (m^2, n^2)$, where $C \in \mathbb{R}$, $C > 0$

and where $E_{m,n}$ denotes the expected value for m restrictions and n variables and s denotes the number of pivot steps.

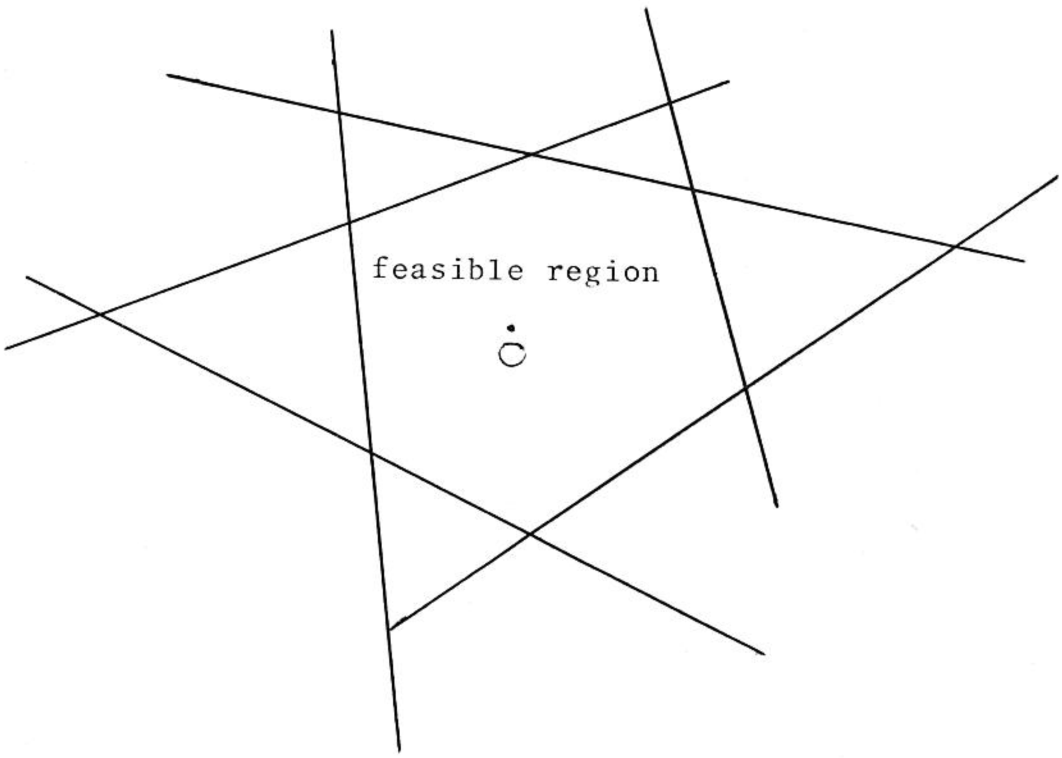
The Rotation - Invariance - Model

The above mentioned dangers of infeasibility and high probability for unboundedness of the objective are not valid under the stochastic model of the author, the rotation-invariance-model. Here we deal with problems of the following kind

$$\begin{array}{lll} \text{Maximize} & v^T x & a_i \in \mathbb{R}^n \\ \text{subject to} & a_1^T x \leq 1, \ldots, a_m^T x \leq 1 & m \geq n \end{array}$$

where $a_1, \ldots, a_m$, v are distributed independently, identically and symmetrically under rotations over $\mathbb{R}^n \setminus \{0\}$.

Here 0 is feasible in any case. A typical situation for $m = 6$ and $n = 2$ would be



The distribution under consideration is characterized by its radial part. The rotation-symmetry leaves that freedom to choose the radial-distribution as desired. Particularly the redundancy rate can be chosen very high (close to 1 as shall be discussed later) as well as extremely small (0 for uniform distribution on the unit sphere or boundary of the unit ball).

Under this stochastic model the author has derived a polynomial upper bound for the average number of steps.

Result of 1981 for Phase II (Borgwardt)

$$E_{m,n}(s) \leq n^3 m^{\frac{1}{n-1}} e\pi \left(\frac{\pi}{2} + \frac{1}{e}\right)$$

The purpose of this talk is to report on some generalizations, improvements and open questions. All these results are not yet published, there is no dramatic breakthrough, but our knowledge about the analysis of that model may get completer.

Some Generalizations

Note that our published results hold under somehow more general conditions than postulated above. Consider problems of the following type

Maximize $v^T x$

subject to $a_1^T(x-x_0) \leq b_1, \ldots, a_m^T(x-x_0) \leq b_m$

where x_0 is a given, known point of $\mathbb{R}^n$,

where $a_1, \ldots, a_m, v$ are distributed as above and where the values on the right side $b_1, \ldots, b_m$ are independently and identically distributed positive random variables over $\mathbb{R}$.

(Of course $a_1, \ldots, a_n$, v, $b_1 \ldots, b_m$ shall be independent).

Since all of the b_i's are positive, x_o is a known feasible point. A problem-formulation of this kind can easily be reduced to our original formulation.

1) We apply a linear transformation $\bar{x} = x - x_o$ and obtain the problem

   Maximize $v^T\bar{x} + v^T x_o$

   subject to $a_1{}^T\bar{x} \leq b_1, \ldots, a_m{}^T\bar{x} \leq b_m$

2) We divide by the right hand sides $\bar{a}_i = \frac{1}{b_i} a_i$

   and get a "normalized" problem

   Maximize $v^T\bar{x}$

   subject to $\bar{a}_1{}^T \bar{x} \leq 1, \ldots, \bar{a}_m{}^T \bar{x} \leq 1$,

   since $v^T x_o$ is constant.

The new vectors $\bar{a}_1, \ldots \bar{a}_m$ satisfy all the required properties of the original model, they are distributed independently, identically and symmetrically under rotations. So we have the original situation.

Note that here the knowledge of x_o is necessary. It is still an open problem whether we can generalize our results to the case where the problem is feasible, but no feasible point is known. We have got some ideas about that, but we are not yet through.

A dual interpretation and redundancy

In the analysis of the author the parametric variant was translated into the shadow-vertex-algorithm in order to provide a dual interpretation of the algorithm. This interpretation has the advantage that we can deal directly with the input data when we want to evaluate the probability that a typical basic solution is feasible and situated on the Simplex-Path.

The only candidates for being Simplex-Vertices are the basic solutions. At a basic solution x_Δ are n (dimension) of m (number of restrictions) restrictions active.

Let $\Delta = \{\Delta^1, \ldots, \Delta^n\} \subset \{1, \ldots, m\}$ be such an n-element-subset
and let

x_Δ be the solution of the system of equations

$$a_{\Delta^1}^T x = 1, \ldots\ldots, a_{\Delta^n}^T x = 1.$$

The following lemma is the main tool for the analysis

Lemma

1) A basic solution x_Δ is a vertex of the feasible region X if and only if

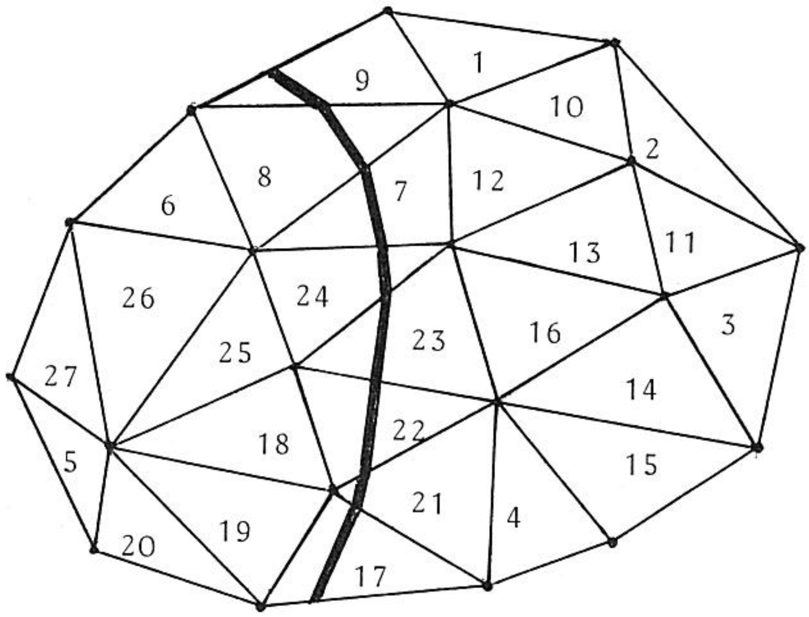
   CH $(a_{\Delta^1}, \ldots, a_{\Delta^n})$ is a facet of CH $(o, a_1 \ldots, a_m)$.
   (Here CH stands for convex hull).

2) A vertex x_Δ lies on the shadow-vertex-Simplex-Path only if

$$CH\ (a_{\Delta 1}, \ldots\ a_{\Delta n}) \cap \ \text{span}\ (u,v) \neq \emptyset$$

with a fixed $u \in \mathbb{R}^n \setminus \{o\}$

The following figure gives an illustration of the efficiency of these two conditions.



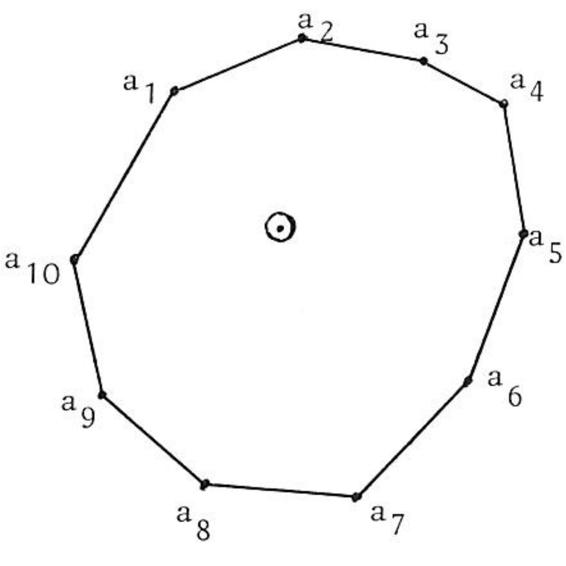
Here we see the upper half of a 3-dimensional polyhedron CH $(0, a_1, \ldots, a_m)$. Its surface is covered by 27 visible facets (triangles since $n = 3$), although more then 20 points (restrictions) are involved. Only eight of these facets are intersected by a two-dimensional plane span (u,v).

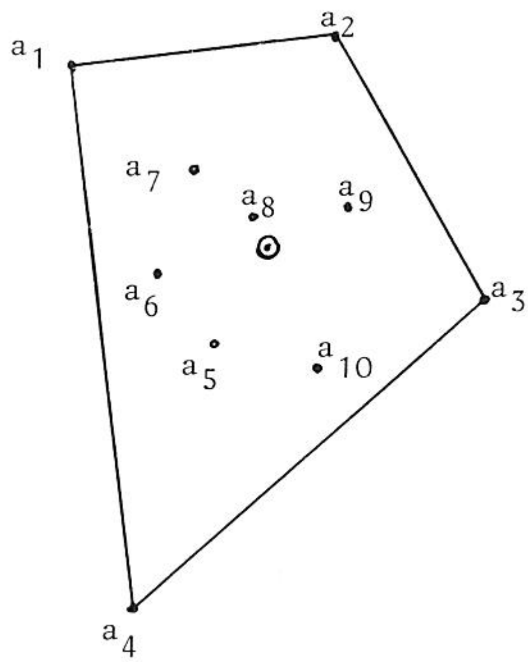
These facets correspond uniquely to the vertices on the Simplex-Path under consideration.

The dual interpretation also allows a simple explanation of redundancy.

Lemma

If $a_i \in \mathrm{Int}\,(\mathrm{CH}\,(o,a_1,\ldots,a_m))$, then $a_i{}^T x \leq 1$ is redundant. The following figure shows a configuration with low redundancy rate (as for uniform distribution on the unit sphere) and a configuration with high redundancy rate (as for Gaussian distribution or the so-called W. Schmidt-examples).





The number of facets or boundary simplices depends in a dramatic way on the redundancy rate.

For high redundancy rate the size of the average number of steps is surprisingly small. Here we refer to a famous result of W. Schmidt under our model.

Theorem (W. Schmidt 1968)

If $P(\| a_i \| \geq r) = \frac{1}{r}$ for $r \to \infty$, then

$E_{m,n}(V) \leq C(n)$ for $m \to \infty$, n fixed.

Here V denotes the number of facets.

In 1979 we could generalize that result to the case where

$P(\| a_i \| \leq r) = \frac{1}{\Pi(r)}$ for $r \to \infty$ and where Π is a polynomial.
A new result in this direction could be derived recently.

Theorem

If for $r \to \infty$ $P(\| a_i \| \geq r)$ behaves like $\frac{k}{r^l}$

(where $l = n^2, k > 0$) then a $\varphi \in \mathbb{R}$ exists such that $\limsup\limits_{m \to \infty} E_{m,n}(s) \leq n^{5/2} \cdot \varphi$ for fixed n.

Note that here we have got such a low-size result even in the case that we have feasible problems in each case.

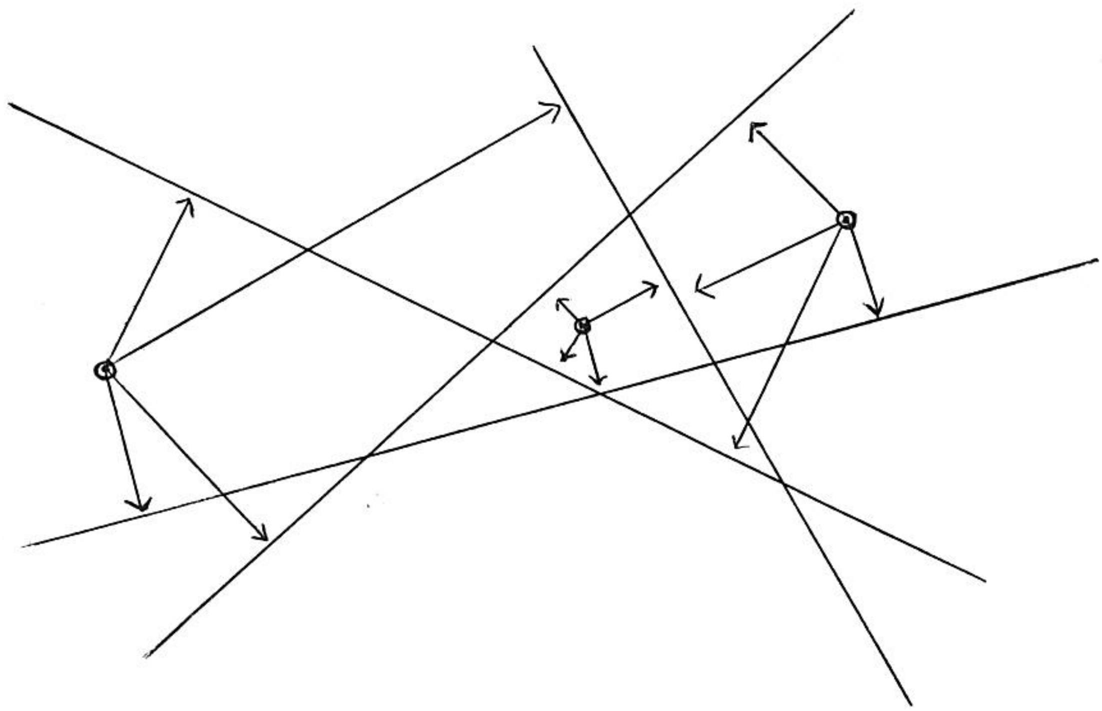
This result should be compared with a by-result in the analysis of the sign-invariance-model.
(Adler/Megiddo and Adler/Karp/Shamir) remarked that the order of steps for their complete method is O $(n^{5/2})$ on the average.

However, Adler/Karp/Shamir have shown that even the expected number of vertices of x under the sign-invariance model is bounded from above by a value $\mathcal{C}(n)$ which is independent of m.

So it seems to be very interesting to study the possible implicit relations and similarities between the stochastic assumptions of the sign-invariance model and the distributions of the W. Schmidt-type.

The most important question is:

> Let us "observe" the restrictions from a "typical" feasible point in the sign-invariance-model.
> Do the distances between this point and the restricting hyperplanes and the directions simulate a distribution as in the rotation-symmetry-model?



We can neither prove such a conjectured relationship nor define precisely what we mean by "typical" point.

At this point we should remark that these W.Schmidt-examples are a very special case of the rotation-symmetry-model. For almost all distributions according to our model we have lower bounds which are increasing to infinity with $m \to \infty$.

Result for the Phase I-Method

In my published paper (1982 b)I gave an algorithm for Phase I which should exploit the Phase II - result and which was constructed in such a manner that the stochastic assumptions for the shadow-vertex-algorithm are met. The (rather crucial) method works as follows.

Take Π_k for the orthogonal projection from $\mathbb{R}^n$ into $\mathbb{R}^k$ and denote by I_k the problem

$$I_k \quad \begin{array}{ll} \text{Maximize} & \Pi_k(v)^T \, \Pi_k(x) \\ \text{s.t.} & \Pi_k(a_i)^T \, \Pi_k(x) \leq 1 \quad \text{for } i = 1, \ldots, m. \end{array}$$

and let X_k be the feasible set for I_k.

Our complete method proceeds as follows:

Initialization: Find a vertex of X_2 and solve I_2.
If this is impossible, STOP

Typical Step ($2 < k \leq n$):

Take the solution $\bar{x} \in \mathbb{R}^{k-1}$ of I_{k-1} and augment it with a o in its k-th component. Then $\binom{\bar{x}}{o}$ lies on an edge of X_k. Start the shadow-vertex-algorithm from a vertex adjacent to that edge. Find a solution for I_k in X_k. If such a solution does not exist, STOP.

The resulting vertex of the last step is then the desired optimal solution. The total number of steps was estimated in the following simple way in (1982 b)

$$E_{m,n}\ (s^t) \leq \sum_{k=2}^{n} \text{expected number of steps for } I_k$$

$$\leq \sum_{k=2}^{n} k\ e\pi \left(\frac{1}{e} + \frac{\pi}{2}\right) m^{\frac{1}{k-1}}$$

$$\leq n^2\ (n+1)^2\ \frac{e\pi}{4} \left(\frac{\pi}{2} + \frac{1}{e}\right) m.$$

Since k is small at the begin of the sumation, we lost the small growth in m as derived for Phase II.

But this order of size can be diminished significantly. The reason is that our Phase II - estimation is based on the worst distributions (concerning redundancy and average behaviour as the uniform distribution of the unit sphere).

Note that the distributions of $\Pi_k\ (a_i)$ (for $k \subset n$) are marginal distributions. A consequence is that only a subclass of the rotational - symmetric distributions over $\mathbb{R}^k$ can occur as the result of a projection from $\mathbb{R}^n$ into $\mathbb{R}^k$. The "bad distributions", where the probability is concentrated at the boundary of the support, are avoided here. The density of the marginal distributions is much more concentrated in the interior of the support. The latter situation leads to a higher redundancy rate and to a much better behaviour of the algorithm. This gives the motivation for a hope that the term m can be improved.

And indeed, if we insert the representation as a marginal distribution into our integrals, we are able to reduce the cases $k = 2, \ldots, n - 1$ to the case $k = n$.

So we obtain a better bound of the type

New result

$$E_{m,n}(s^t) \leqq m^{\frac{1}{n-1}} (n + 1)^4 \; \frac{2}{5} \pi \left(1 + \frac{e\pi}{2}\right).$$

Problems with nonnegativity constraints

Another drawback in our theory -as published- was our unability to analyze problems with sign-restrictions, because here we have n restrictions which are fixed and do destroy the property of rotational symmetry.

Now we have found a way to deal with such problems. We have to include the sign-restrictions into the set of the "normal" restrictions. Therefore, we describe

$x_i \geq o$ by the condition $- \rho\, e_i{}^T x \leqq 1$ for all $\rho \leqq o$.

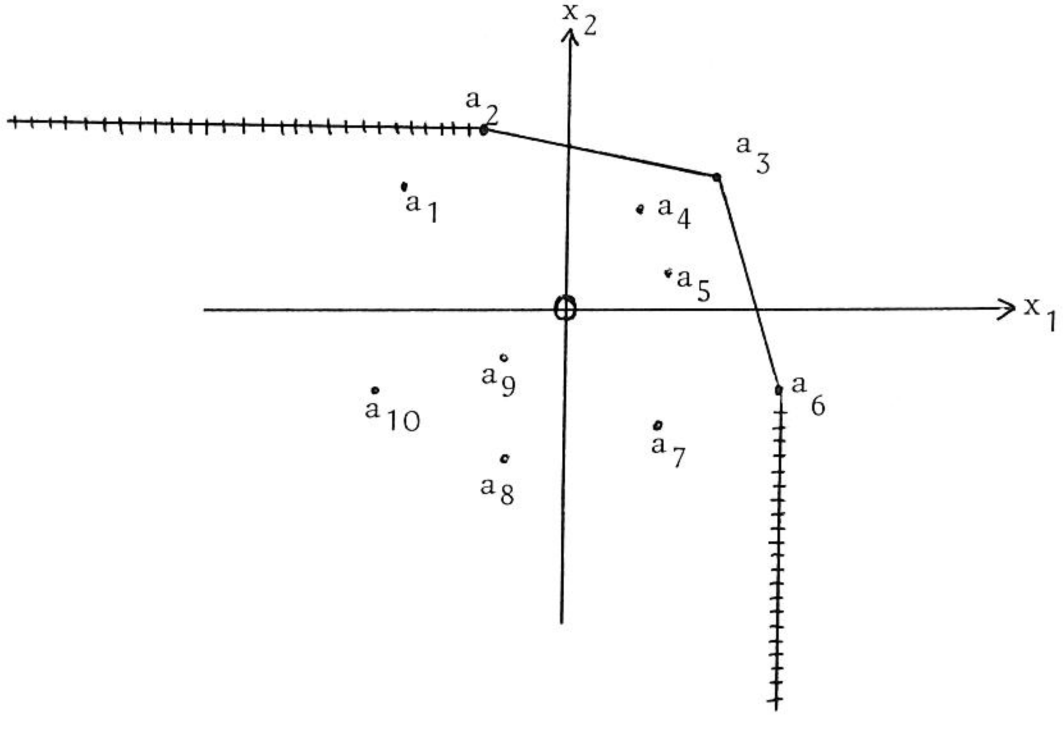
The dual polyhedron CH $(o, a_1, \ldots, a_m)$ is replaced by the set CH $(o, a_1, \ldots, a_m)$ + CC $(-e_1, \ldots, -e_n)$
(CC means convex cone),
and the boundary simplices CH $(a_{\Delta^1}, \ldots, a_{\Delta^n})$ by the boundary simplex-cones
CH $(a_{\Delta^1}, \ldots, a_{\Delta^k})$ + CC $(-e_{k+1}, \ldots, -e_n)$.

For the solution algorithm we define J_k $(k = 2, \ldots, n)$ as the original problem with additional constraints
$x_{k+1} = \ldots\ldots = x_n = o$.

We <u>initialize</u> by solving J_2 as above.

In the <u>typical step</u> we start from a solution of J_{k-1}
and solve J_k with our shadow-vertex-algorithm.
If this is impossible, we STOP.

Again, only boundary simplex-cones correspond to vertices of X and only such vertices are visited on the path whose boundary simplex cones are intersected by a two-dimensional plane.



So we obtain our result even for such problems

$$E_{m,n}(s^t) \leqq (n+1)^4 \ \frac{2}{5}\pi \left(1+\frac{e\pi}{2}\right).$$

Let me finish with a personal opinion.
Until now we have got an impression why the Simplex-Method works so efficiently, but we are far from understanding it completely.

[1] Adler, I., Karp, R. & Shamir, R. [1983a]: *A Family of Simplex Variants Solving an $m \times d$ Linear Program in Expected Number of Pivot Steps Depending on d Only,* University of California, Computer Science Division, Berkeley, December 1983.

[2] Adler, I., Karp, R. & Shamir, R. [1983b]: *A Simplex Variant Solving an $m \times d$ Linear Program in* $0(\min(m^2, d^2))$ *Expected Number of Pivot Steps,* University of California, Computer Science Division, Berkeley, December 1983.

[3] Adler, I. & Meggido, N. [1983]: *A Simplex Algorithm where the Average Number of Steps is Bounded Between two Qudratic Functions of the Smaller Dimension,* Department of Industrial Engineering and Operations Research, University of California, Berkeley, California, December 1983.

[4] Avis, D. & Chvatal, V. [1978]: *Notes on Bland's Pivoting Rule,* Mathematical Programming Study 8 (1978), 24–34.

[5] Blair C. [1983]: *Random Linear Programs with Many Variables and Few Constraints,* College of Commerce and Business Administration, University of Illinois at Urbana, Champaign, April 1983.

[6] Borgwardt, K. H. [1977a]: *Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung,* Dissertation Universität Kaiserslautern.

[7] Borgwardt, K. H. [1977b]: *Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung,* Operations Research Verfahren 28 (1977), 332–345.

[8] Borgwardt, K. H. [1978]: *Zum Rechenaufwand von Simplexverfahren,* Operations Research Verfahren 31 (1978), 83–97.

[9] Borgwardt, K. H. [1979]: *Die asymptotische Ordnung der mittleren Schrittzahl von Simplexverfahren,* Methods of Operations Research 37 (1979), 81–95.

[10] Borgwardt, K. H. [1981]: *The Expected Number of Pivot Steps Required by a Certain Variant of the Simplex Method is Polynomial,* Methods of Operations Research 43 (1981), 35–41.

[11] Borgwardt, K. H. [1982a]: *Some Distribution-Independent Results About the Asymptotic Order of the Average Number of Pivot Steps of the Simplex Method,* Mathematics of Operations Research 7 (1982), 441–462.

[12] Borgwardt, K. H. [1982b]: *The Average Number of Pivot Steps Required by the Simplex-Method is Polynomial,* Zeitschrift für Operations Research 26 (1982), 157–177.

[13] Carnal, H. [1970]: *Die konvexe Hülle von n rotationssymmetrisch verteilten Punkten,* Zeitschrift für Wahrscheinlichkeitsrechnung und verwandte Gebiete 15 (1970), 168–176.

[14] Dantzig, G. [1966]: *Lineare Programmierung und Erweiterungen*, Springer Verlag, Berlin, 1966.

[15] Efron, B. [1965]: *The Convex Hull of a Random Set of Points*, Biometrica 52 (3) and (4) (1965), 331–345.

[16] Goldfarb, D. [1983]: *Worst Case Complexity of the Shadow Vertex Simplex Algorithm*, Columbia University, School of Engineering and Applied Science, May 1983.

[17] Goldfarb, G. & Sit, W. J. [1979]: *Worst Case Behavior of the Steepest Edge Simplex Method*, Discrete Applied Mathematics 1 (1979), 277–285.

[18] Haimovich, M. [1983]: *The Simplex Algorithm is Very Good! — On The Expected Number of Pivot Steps and Related Properties of Random Linear Programs*, 415 Uris Hall, Columbia University, New York, April 1983.

[19] Jeroslow, R. G. [1973]: *The Simplex Algorithm with the Pivot-Rule of Maximizing Criterion Improvement*, Discrete Mathematics (1973), 367–377.

[20] Kelly, D. G. & Tolle, J. W. [1979]: *Expected Number of Vertices of a Random Convex Polytope*, University of North Carolina at Chapel Hill, 1979.

[21] Klee, V. & Minty, G. [1971]: *How Good is the Simplex-Algorithm?*, Inequalities III, O. Shisha (ed.), Academic Press, New York, 1971.

[22] Liebling, T. [1972]: *On the Number of Iterations of the Simplex Method*, ETH Zürich, Institut für Operations Research, 1972.

[23] Lindberg, P. O. [1981]: *A Note on Random LP-Problems*, Dept. of Mathematics, Royal Institute of Technology, Stockholm Schweden, 1981.

[24] May, J. H. & Smith, R. L. [1982]: *Random Polytopes: Their Definition, Generation and Aggregate Properties*, Mathematical Programming 24 (1982), 39–54.

[25] Megiddo, N. [1983]: *Improved Asymptotic Analysis of the Average Number of Steps Performed by the Self-Dual Simplex Algorithm*, Dept. of Computer Science, Stanford University, September 1983.

[26] Raynaud, H. [1970]: *Sur l'Enveloppe Convexe des Nuages de Points Aléatoires dand R^n*, Journal of Applied Probability 7 (1970), 35–48.

[27] Renyi, A. & Sulanke, R. [1963]: *Über die konvexe Hülle von n zufällig gewählten Punkten I*, Zeitschrift für Warsch. Verw. Gebiete 2 (1963), 75–84.

[28] Renyi, A. & Sulanke, R. [1964]: *Über die konvexe Hülle von n zufällig gewählten Punkten II*, Zeitschrift für Wahrscheinlichkeitstheorie 3 (1964), 138–147.

[29] Renyi, A. & Sulanke, R. [1968]: *Zufällige konvexe Polygone in einem Ringgebiet*, Zeitschrift für Wahrscheinlichkeitstheorie 9 (1968), 146–157.

[30] Schmidt, W. M. [1968]: *Some Results in Probabilistic Geometry*, Zeitschrift für Warsch. Verw. Gebiete 9 (1968), 146–157.

[31] Smale, S. [1982]: *The Problem of the Average Speed of the Simplex Method*, Proceedings of the 11th International Symposium on Mathematical Programming, Universität Bonn, August 1982, 530–539.

[32] Shamir, R. [1984]: *The Efficiency of the Simplex Method: A Survey*, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, May 1984.

[33] Smale, S. [1983]: *On the Average Speed of the Simplex Method*, Mathematical Programming 27 (1983), 241–262.

[34] Todd, M. J. [1983]: *Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complementarity and Linear Programming Problems*, Technical Report No. 595, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, November 1983.