KHACHIAN'S ALGORITHM FOR LINEAR PROGRAMMING

by

Peter Gács and Laszlo Lovász

STAN-CS-79-750
July 1979

DEPARTMENT OF COMPUTER SCIENCE
School of Humanities and Sciences
STANFORD UNIVERSITY

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited
<table>
<thead>
<tr>
<th>REPORT DOCUMENTATION PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REPORT NUMBER</td>
</tr>
<tr>
<td>2. GOVT ACCESSION NO.</td>
</tr>
<tr>
<td>4. TITLE (and Subtitle)</td>
</tr>
<tr>
<td>5. TYPE OF REPORT &amp; PERIOD COVERED</td>
</tr>
<tr>
<td>6. PERFORMING ORG. REPORT NUMBER</td>
</tr>
<tr>
<td>7. AUTHOR(S)</td>
</tr>
<tr>
<td>8. CONTRACT OR GRANT NUMBER(S)</td>
</tr>
<tr>
<td>9. PERFORMING ORGANIZATION NAME AND ADDRESS</td>
</tr>
<tr>
<td>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</td>
</tr>
<tr>
<td>11. CONTROLLING OFFICE NAME AND ADDRESS</td>
</tr>
<tr>
<td>12. REPORT DATE</td>
</tr>
<tr>
<td>13. NUMBER OF PAGES</td>
</tr>
<tr>
<td>14. MONITORING AGENCY NAME &amp; ADDRESS</td>
</tr>
<tr>
<td>15. SECURITY CLASS. (of this report)</td>
</tr>
<tr>
<td>16. DISTRIBUTION STATEMENT (of this Report)</td>
</tr>
<tr>
<td>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</td>
</tr>
<tr>
<td>18. SUPPLEMENTARY NOTES</td>
</tr>
<tr>
<td>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</td>
</tr>
<tr>
<td>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</td>
</tr>
</tbody>
</table>
L. G. Khachian's algorithm to check the solvability of a system of linear inequalities with integral coefficients is described. The running time of the algorithm is polynomial in the number of digits of the coefficients. It can be applied to solve linear programs in polynomial time.
Khachian's Algorithm for Linear Programming

Peter Dács and Laszlo Lovász

Computer Science Department
Stanford University
Stanford, California 94305

Abstract.

L. G. Khachian's algorithm to check the solvability of a system of linear inequalities with integral coefficients is described. The running time of the algorithm is polynomial in the number of digits of the coefficients. It can be applied to solve linear programs in polynomial time.

This research was supported in part by National Science Foundation grant MCS-77-23738 and by Office of Naval Research contract N00014-76-C-0330. Reproduction in whole or in part is permitted for any purpose of the United States government.
L. G. Khachian [Doklady Akademii Nauk SSSR, 1979, Vol. 244, No. 5, 1093-1096] published a polynomial-bounded algorithm to solve linear programming. These are some notes on this paper. We have ignored his considerations which concern the precision of real computations, in order to make the underlying idea clearer, on the other hand, proofs which are missing from his paper are given in an appendix.

Let

\[ a_i x < b_i \quad (i = 1, \ldots, m, \ a_i \in \mathbb{Z}^n, \ b_i \in \mathbb{Z}) \]

be a system of strict linear inequalities with integral coefficients. We present an algorithm which decides whether or not (1) is solvable, and yields a solution if it is.

Define

\[ L = \sum_{i,j} \log(|a_{ij}| + 1) + \sum_{i} \log(|b_i| + 1) + \log nm + 1 \]

L is the space needed to state the problem.

The Algorithm.

We define a sequence \( x_0, x_1, \ldots \in \mathbb{R}^n \) and a sequence of symmetric positive definite matrices \( A_0, A_1, \ldots \) recursively as follows.

\( x_0 = 0, \ A_0 = 2^n I \). Assume that \( (x_k, A_k) \) is defined. Check if \( x_k \)
is a solution of (1). If it is, stop. If not, pick any inequality in (1) which is violated:

\[ a_i x_k \not\geq b_i \]

and set
\[ x_{k+1} = x_k - \frac{1}{n+1} \frac{A_k a_1}{\sqrt{A_k^T a_1 A_k a_1}} \]
\[ A_{k+1} = \frac{n^2}{n^2-1} \left( A_k - \frac{n}{n+1} \frac{(A_k a_1)(A_k a_1)^T}{a_1^T A_k a_1} \right) \]

(Note that the multiplication of vector \( A_k a_1 \) with itself in the second term results in an \( n \times n \) matrix.)

In practice, we will compute only certain approximations of \( x_k \) and \( A_k \) by decimals of a certain precision. It can be shown that approximations within \( \exp(-10nL) \) preserve the validity of the following theorem.

**Theorem.** If the algorithm stops, \( x_k \) is a solution of (1). If the algorithm does not stop in \( 6n^2L \) steps, then (1) is not solvable.

The first assertion is, of course, just a repetition of the stopping rule for the algorithm. To prove the crucial second statement, we shall need a series of lemmas, along with a geometric description of what's happening.

Let \( x_0 \in \mathbb{R}^n \) and \( A \) a positive definite matrix. Then
\[ (x-x_0)^T A^{-1} (x-x_0) \leq 1 \]
defines an ellipsoid \( E = (x, A) \) with center \( x \). Let \( a \in \mathbb{R}^n \), \( a \neq 0 \).
Then we shall denote by \( E^a \) the ellipsoid \( (x_0', A') \), where
\[ x_0' = x_0 - \frac{1}{n+1} A \frac{a}{\sqrt{a^T A a}} \]
\[ A' = \frac{n^2}{n^2 - 1} \left( A - \frac{2}{n+1} \frac{(Aa)(Aa)^T}{a^T A a} \right) \]

We shall denote the semi-ellipsoid
\[ E \cap \{ x : (x-x_0)a \leq 0 \} \]
by \( \frac{1}{2} E_a \).

Let us remark (although this is not needed in the proof) that geometrically this construction means the following. Take a hyperplane \( ax = d \), \( d < ax_0 \), which is tangent to \( E \) at point \( y \). Then
\[ x_0 - y = A \frac{a}{\sqrt{a^T A a}} \]

Now \( E^n \) will be the (unique) ellipsoid which touches the hyperplane \( ax = d \) at \( y \) and intersects the hyperplane \( ax = ax_0 \) in the same ellipsoid as \( E \).

So here come the lemmas. The first three are facts of number-theoretic nature which probably are familiar to many people who have investigated the complexity of algorithmic problems in linear algebra.

We use the notation \( |x|_\infty = \max_1 x_i \), \( |x|_2 = \sqrt{\sum_1 x_i^2} \).

**Lemma 1.** Every vertex \( v \) of the polyhedron
\[
\begin{align*}
\begin{array}{cc}
a_i & b_i \\
(1 = 1, \ldots, m)
\end{array}
\end{align*}
\]
\[ x \geq 0 \]
satisfies \( |v|_\infty < 2^L/n \), and its entries are rational numbers with denominator at most \( 2^L \).

**Lemma 2.** If (1) has a solution, then the volume of its solutions inside the cube \( |x_1| \leq 2^L \) is at least \( 2^{-nL} \).
Lemma 3. Suppose that the system
\[ a_i x < b_i + 2^{-L} \quad (i = 1, \ldots, m) \]
has a solution. Then
\[ a_i x < b_i \quad (i = 1, \ldots, m) \]
has a solution.

Lemma 4. \( \frac{1}{2} E_a \subset E^a \).

Lemma 5. \( \lambda(E^a) = c(n) \lambda(E) \),
where
\[ c(n) = \left( \frac{n^2}{n^2 - 1} \right)^{n-1/2} \quad \frac{n}{n+1} < e^{-\frac{1}{2(n+1)}} \]
and \( \lambda(X) \) is the volume of the set \( X \).

The proof of the theorem is quite easy now. Suppose that the procedure does not stop after \( k = 6n^2 L \) steps, and yet (1) is solvable. Then by Lemma 2, the set \( P \) of its solutions \( x \) inside \( E_0 \) has \( \lambda(P) \geq 2^{-nL} \). By Lemma 4, \( P \subset E_k \). But by Lemma 5,
\[ \lambda(E_k) < e^{-(k/2(n+1))} \lambda(E_0) < e^{-(k/2(n+1))} 2^{2Ln} < 2^{-nL} \]
a contradiction.

If one would like to decide the solvability of a system of the form
\[ (2) \quad a_i x \leq b_i \quad (i = 1, \ldots, n) \]
then we may consider instead the system
\[ (3) \quad [2^L] a_i x < [2^L] b_i + 1 \quad (i = 1, \ldots, n) \].
By Lemma 3, this is solvable iff (2) is solvable.

If we want to solve a linear programming problem

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to} & \quad Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

then consider the system of inequalities

\[
\begin{align*}
c^T x &= b^T y \\
Ax &\leq b \\
x &\geq 0 \\
A^T y &\geq c \\
y &\geq 0.
\end{align*}
\]

This is solvable iff the original program has a feasible solution and a finite optimum, and for any solution \((x, y)\) of this system, \(x\) is an optimal solution of the program.
Appendix

Proof of Lemma 1. Let \( v = (v_1, \ldots, v_n) \). By Cramer's rule, each \( v_i \) can be expressed as

\[ v_i = \frac{D_i}{D}, \]

where \( D_i \) and \( D \) are determinants whose entries are 0, 1, \( a_{ij} \) or \( b_1 \). Hence \( D \) and \( D_i \) are integers, and

\[ |D| \geq 1, \]

\[ |D| \leq \prod (\text{norms of row vectors}) \]

\[ < 2^L/n < 2^L/n, \]

and the same holds for the \( D_i \)'s. This implies the assertion.

Proof of Lemma 2. We may assume that (1) has a solution \( x_0 > 0 \). So the polyhedron

\[
\begin{align*}
a_i x & \leq b_i \quad (i = 1, \ldots, m) \\
x & \geq 0
\end{align*}
\]

has an interior point. Since it contains no line, it also has a vertex \( v = (v_1, \ldots, v_n) \). By Lemma 1, we know that \( v_i < 2^L/n < \lfloor 2^L \rfloor \). It follows that the polyhedron (4) has an interior point \( x = (x_1, \ldots, x_n) \) with \( x_j < \lfloor 2^L \rfloor \), and so the polytope

\[
\begin{align*}
a_i x & \leq b_i \quad (i = 1, \ldots, m) \\
x & \geq 0 \\
x_j & \leq \lfloor 2^L \rfloor \quad (j = 1, \ldots, n)
\end{align*}
\]

has an interior point. Hence, it has \( n+1 \) vertices \( v_0, \ldots, v_n \) which
are not on a hyperplane. So (5) has volume at least

$$\frac{1}{n!} \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ v_0 & v_1 & \cdots & v_n \end{pmatrix}.$$ 

Here, by Lemma 1, we get that

$$v_1 = \frac{1}{D_1} u_1,$$

where $u_1$ is an integer vector and $D_1$ is an integer $< 2^{r/n}$. So

$$\det \begin{pmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_n \end{pmatrix} = \frac{1}{|D_1| \cdots |D_n|} \det \begin{pmatrix} D_1 & \cdots & D_n \\ u_1 & \cdots & u_n \end{pmatrix} \geq \frac{1}{|D_1| \cdots |D_n|} \geq 2^{-nL} n^n$$

since the determinant in the second expression is a non-zero integer.

So the volume of the polytope (5) is at least $\frac{1}{n!} 2^{-nL} n^n > 2^{-nL}$.

Proof of Lemma 2. For $x \in \mathbb{R}^n$, set

$$\varrho_1(x) = a_1 x - b_1.$$ 

Let $x_0 \in \mathbb{R}^n$ be arbitrary.

Claim 1. There exists an $x_1 \in \mathbb{R}^n$ such that

(1) $\varrho_i(x_1) \leq \max(0, \varrho_i(x_0))$ (i = 1, ..., m)

and

(2) The vectors $\{a_i : \varrho_i(x_1) \geq 0\}$ span every other vector $a_i$.

To prove the claim, it suffices to show that if $x_0$ does not satisfy (2) then we can find a vector $x_1$ such that $x_1$ satisfies (1) and $\varrho_i(x_1) \geq 0$ holds, for more indices $i$ than $\varrho_i(x_0) \geq 0$. Repeating
this at most \( m \) times we must obtain an \( x_1 \) satisfying both (1) and (2).

Let, say \( \varrho_1(x_0), \ldots, \varrho_k(x_0) \geq 0 \), \( \varrho_{k+1}(x_0), \ldots, \varrho_m(x_0) < 0 \).

Suppose that \( a_v (v > k) \) is not a linear combination of \( a_1, \ldots, a_k \).

Then the system of linear equations

\[
\begin{align*}
a_1 y &= 0 \quad (i = 1, \ldots, k) \\
a_v y &= 1
\end{align*}
\]

is solvable. Let \( y_0 \) be a solution and consider

\[
x_1 = x_0 + ty_0,
\]

where

\[
t = \max\{s \in \mathbb{R} : sa_jy_0 + \varrho_j \leq 0 \quad (j = k+1, \ldots, m)\}
\]

\( t \) is finite, in fact \( t \leq -w_v \).

Then by the choice of \( t \),

\[
\varrho_i(x_1) = ta_iy_0 + \varrho_i(x_0) \begin{cases} = \varrho_i(x_0) & \text{if } 1 \leq i \leq k, \\ \leq 0 & \text{if } k+1 \leq i \leq m, \end{cases}
\]

and equality holds for at least one \( 1 \leq i \leq m \). This proves the Claim.

Assume now that \( x_0 \) is such that

\[
a_i x_0 < b_i + 2^{-L} \quad (i = 1, \ldots, m).
\]

Let, say \( a_i x_0 \geq b_i \) for \( i = 1, \ldots, k \). Choose the labelling so that \( a_1, \ldots, a_r \) are linearly independent but \( a_{r+1}, \ldots, a_k \) are spanned by them. By the Claim, we may assume that \( a_{k+1}, \ldots, a_n \) are also spanned by \( a_1, \ldots, a_r \).
Now let $z$ be a solution of the system of linear equations

$$a_i z = b_i \quad (i = 1, \ldots, r) .$$

We show that $z$ satisfies

$$a_i z \leq b_i$$

for every $1 \leq i \leq m$. We know that

$$a_i = \sum_{j=1}^r \lambda_j a_j$$

with some real numbers $\lambda_j$. In fact by Cramer's rule we also know that

$$\lambda_j = D_j / D,$$

where $D_j$ and $D$ are determinants formed by some entries of the vectors $a_i$ and hence they are integers with absolute value less than $2^L / n$. Now

$$D(a_i z - b_i) = \sum_{j=1}^r D_j a_j z - D b_i$$

$$= \sum_{j=1}^r D_j b_j - D b_i .$$

To estimate the right hand side, use that

$$\sum_{j=1}^r D_j b_j - D b_i = \sum_{j=1}^r D_j (a_j x_0 - \hat{a}_j(x_0)) - D(a_1 x_0 - \hat{a}(x_0))$$

$$= D \cdot \hat{a}_1(x_0) - \sum_{j=1}^r D_j \hat{a}_j(x_0)$$

$$\leq D \cdot 2^{-L} + \sum_{j=1}^r |D_j| 2^{-L} < 1 ,$$

and since the left hand side is an integer,

$$\sum_{j=1}^r D_j b_j - D b_i \leq 0 ,$$

which proves the assertion.
Proof of Lemma 4. We may assume that $x_0 = 0$, $A = I$ (i.e., the ellipsoid is the unit sphere about 0) and that $a = (-1, 0, ..., 0)^T$, since the contents of the lemma is invariant under affine transformations of the space.

Then

$$x'_0 = \left( \frac{1}{n+1}, 0, ..., 0 \right)^T$$

and

$$A' = \text{diag} \left( \frac{n^2}{(n+1)^2}, \frac{n^2}{n^2-1}, ..., \frac{n^2}{n^2-1} \right).$$

Suppose $x \in \frac{1}{2} F_a$. Then $|x|_2 \leq 1$, $1 \geq x_1^T = -a^T x \geq 0$. We have to show that

$$(x-x_0')^T A'^{-1} (x-x_0') \leq 1.$$ 

But

$$(x-x_0')^T A'^{-1} (x-x_0') = x^T A'^{-1} x - 2 x^T A'^{-1} x_0 + x_0'^T A'^{-1} x_0'$$

$$= \frac{n^2-1}{n^2} x^2 + \frac{2n+2}{n^2} x_1^2 - 2 \frac{n+1}{n^2} x_1 + \frac{1}{n^2}$$

$$= \frac{n^2-1}{n^2} (x^2-1) + \frac{2n+2}{n^2} x_1 (x_1-1) + 1 \leq 1.$$

Proof of Lemma 5. We may assume again that $E$ is the unit sphere about 0 and $a = (1, 0, ..., 0)^T$, since affine transformations do not change the proportion of volumes. By a well-known formula,

$$\lambda(E^n) = \frac{\sqrt{\det A'}}{\sqrt{\det A}} \cdot \lambda(E) = \sqrt{\det A'} \cdot \lambda(E)$$

$$= \frac{n}{n+1} \left( \frac{n^2}{n^2-1} \right)^{(n-1)/2} \cdot \lambda(E) = c(n) \cdot \lambda(E).$$
To estimate this factor use that

\[
\frac{n^2}{n^2 - 1} = 1 + \frac{1}{n^2 - 1} < e^{1/(n^2 - 1)}
\]

and

\[
\frac{n}{n+1} = 1 - \frac{1}{n+1} < e^{-1/(n+1)}.
\]

Substituting these bounds we get

\[
c(n) < e^{-1/(2(n+1))}.
\]