A SIMPLE HEURISTIC APPROACH TO SIMPLEX EFFICIENCY

by

SHELDON M. ROSS
A SIMPLE HEURISTIC APPROACH TO SIMPLEX EFFICIENCY

by

Sheldon M. Ross
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

DISTRIBUTION STATEMENT A
Approved for public release; Distribution Unlimited
# A Simple Heuristic Approach to Simplex Efficiency

Sheldon M. Ross

Operations Research Center
University of California
Berkeley, California 94720

United States Air Force
Air Force Office of Scientific Research
Bolling Air Force Base, D.C. 20332

August 1981

11

Approved for public release; distribution unlimited.

Linear Programming
Simplex Algorithm
Probabilistic Analysis
Poisson

(SEE ABSTRACT)
Consider the standard linear program:

\[
\text{Minimize } c^T x \\
\text{subject to: } A x = b \\
x \geq 0
\]

where \( A \) is an \( m \times n \) matrix. The simplex algorithm solves this linear program by moving from extreme point of the feasibility region to a better (in terms of the objective function \( c^T x \)) extreme point (via the pivot operation) until the optimal is reached. In order to obtain a feel for the number of necessary iterations, we consider a simple probabilistic (Markov chain) model as to how the algorithm moves along the extreme points. At first we suppose that if at any time the algorithm is at the \( j \)th best extreme point then after the next pivot the resulting extreme point is equally likely to be any of the \( j-1 \) best. Under this assumption, we show that the time to get from the \( N \)th best to the best extreme point is approximately, for large \( N \), a Poisson distribution with mean equal to the logarithm (base \( e \)) of \( N \). We also consider a more general probabilistic model in which we drop the uniformity assumption and suppose that when at the \( j \)th best the next one is chosen probabilistically according to weights \( w_i, i = 1, ..., j-1 \).
A SIMPLE HEURISTIC APPROACH TO SIMPLEX EFFICIENCY

by

Sheldon M. Ross

1. INTRODUCTION

Consider the standard linear program:

Minimize \( c^T x \)

subject to: \( A x = b \)

\( x \geq 0 \)

where \( A \) is an \( m \times n \) matrix. The simplex algorithm solves this linear program by moving from extreme point of the feasibility region to a better (in terms of the objective function \( c^T x \)) extreme point (via the pivot operation) until the optimal is reached. As there are roughly \( N \leq \binom{n}{m} \) such extreme points it would seem that this method might take many iterations but, surprisingly to some, this does not appear to be the case in practice.

In order to obtain a feel for whether or not the above is surprising, we consider a simple probabilistic (Markov chain) model as to how the algorithm moves along the extreme points. At first we suppose that if at any time the algorithm is at the \( j \)-th best extreme point then after the next pivot the resulting extreme point is equally likely to be any of the \( j - 1 \) best. Under this assumption, we show that the time to get from the \( N \)-th best to the best extreme point has approximately, for large \( N \), a Poisson distribution with mean equal to the logarithm (base \( e \)) of \( N \). We also consider a more general probabilistic model.
in which we drop the uniformity assumption and suppose that when at
the jth best the next one is chosen probabilistically according to
weights $w_i$, $i = 1, \ldots, j-1$. 
2. THE UNIFORM MARKOV CHAIN

Consider a Markov chain for which \( P_{1i} = 1 \) and

\[
P_{ij} = \frac{1}{i-1}, \quad j = 1, \ldots, i-1, \quad i > 1
\]

and let \( T_N \) denote the number of transitions to get from state \( N \) to state 1. Then \( T_N \) can be expressed as

\[
T_N = \sum_{j=1}^{N-1} I_j
\]

where

\[
I_j = \begin{cases} 
1 & \text{if the process ever enters } j \\
0 & \text{otherwise}
\end{cases}
\]

**Proposition 1:**

\( I_1, \ldots, I_{N-1} \) are independent and

\[
P(I_j = 1) = \frac{1}{j}, \quad 1 \leq j \leq N - 1.
\]

**Proof:**

Given \( I_{j+1}, \ldots, I_N \) let \( n = \min \{i : i > j, I_i = 1\} \). Then

\[
P(I_j = 1 \mid I_{j+1}, \ldots, I_N) = \frac{1/(n-1)}{j/(n-1)} = \frac{1}{j}.
\]
Corollary 2:

(i) \( E[T_N] = \sum_{j=1}^{N-1} \frac{1}{j} \)

(ii) \( \text{Var}(T_N) = \sum_{j=1}^{N-1} \frac{1}{j} \left( 1 - \frac{1}{j} \right) \)

(iii) For \( N \) large, \( T_N \) has approximately a Poisson distribution with mean \( \log N \).

Proof:

Parts (i) and (ii) follow from Proposition 1 and the representation \( T_N = \sum_{j=1}^{N-1} I_j \). Part (iii) follows from the Poisson limit theorem since

\[
\int_{1}^{N} \frac{dx}{x} < \sum_{j=1}^{N-1} \frac{1}{j} < 1 + \int_{1}^{N} \frac{dx}{x}
\]

or

\[
\log N < \sum_{j=1}^{N-1} \frac{1}{j} < 1 + \log (N-1)
\]

and so

\[
\log N \approx \sum_{j=1}^{N-1} \frac{1}{j}.
\]
3. APPLICATION TO SIMPLEX

Assuming that \( n, m \) and \( n - m \) are all large, we have by Stirling's approximation that

\[
N = \binom{n}{m} \approx \frac{n^{n+1/2}}{(n - m)^{n-m+1/2} \sqrt{2\pi m}} \]

and so letting \( c = \frac{n}{m} \)

\[
\log N \sim (mc + 1/2) \log (mc) - (m(c - 1) + 1/2) \log (m(c - 1))
- (m + 1/2) \log m - 1/2 \log (2\pi)
\]

or

\[
\log N \sim m \left[ c \log \frac{c}{c-1} + \log (c - 1) \right].
\]

Now, as \( \lim_{x \to \infty} x \log (x/x - 1) = 1 \), it follows that when \( c \) is large

\[
\log N \sim m[1 + \log (c - 1)].
\]

Thus for instance if \( n = 8000 \), \( m = 1000 \), then the number of necessary transitions is approximately Poisson distributed with mean \( 1000(1 + \log 7) \approx 3000 \). As the variance is equal to the mean, we see by the normal approximation to the Poisson that the number of necessary transitions would be roughly between

\[
3000 \pm 2\sqrt{3000} \text{ or, roughly, } 3000 \pm 110
\]

95 percent of the time.
4. A WEIGHTED MARKOV CHAIN MODEL

Suppose now that \( P_{11} = 1 \) and

\[
P_{ij} = \frac{w_j}{w_1 + \ldots + w_{i-1}} \quad j \leq i - 1.
\]

With this model we are thus able to give more weight to those states closest to the one presently at by letting \( w_j \) increase in \( j \).

Analogously with Proposition 1, we have

**Proposition 2:**

If

\[
I_j = \begin{cases} 
1 & \text{if } j \text{ is ever visited} \\
0 & \text{otherwise}
\end{cases}
\]

Then \( I_1, \ldots, I_{N-1} \) are independent and

\[
P(I_j = 1) = \frac{w_j}{\sum_{i=1}^{j} w_i}, \quad 1 \leq j \leq N - 1.
\]

In addition, if \( T_N = \sum_{j=1}^{N-1} I_j \). Then

\[
E[T_N] = \sum_{j=1}^{N-1} \left( \frac{w_j}{\sum_{i=1}^{j} w_i} \right)
\]

\[
\text{Var}(T_N) = \sum_{j=1}^{N-1} \frac{w_j}{\sum_{i=1}^{j} w_i} \left( 1 - \frac{w_j}{\sum_{i=1}^{j} w_i} \right).
\]
If for instance we use polynomial weights—\( w_j = j^\alpha \), \( 0 < \alpha < \infty \), then

\[
\frac{1}{i} \sum_{i=1}^{j} w_i = \frac{1}{i} \int_{1}^{i} x^\alpha dx
\]

\[
\approx \int_{1}^{j} x^\alpha dx = \frac{j^{\alpha+1} - 1}{\alpha + 1}
\]

and so

\[
\frac{1}{i} \sum_{i=1}^{j} w_i \approx \frac{(\alpha + 1)j^\alpha}{j^{\alpha+1} - 1} \approx \frac{\alpha + 1}{j}.
\]

Hence

\[
E[T_N] \approx \int_{1}^{N-1} \frac{\alpha + 1}{x} dx = (\alpha + 1) \log (N - 1)
\]

and thus in this case \( T_N \) has, for large \( N \), approximately a Poisson distribution with mean \( (\alpha + 1) \log N \). Thus when \( N = \binom{n}{m} \), the number of transitions (i.e., simplex pivot iterations) is approximately Poisson with mean

\[
(\alpha + 1) m \left[ c \log \left( \frac{c}{c - 1} \right) + \log (c - 1) \right], \quad c = \frac{n}{m}
\]

which when \( c \) is large is approximately

\[
(\alpha + 1) m[1 + \log (c - 1)]
\]
REFERENCES

