How Linear Programming Became Practical

Robert Fourer
President, AMPL Optimization Inc.
Professor Emeritus, Northwestern University
4er@ampl.com — +1 773-336-2675

IFORS 2021: The 22nd Conference of the
International Federation of Operational Research Societies
Seoul, Korea (online) — 23-27 August 2021
The Evolution of Computationally Practical Linear Programming

Although a recognizable simplex approach to linear programming was being studied by Dantzig and others by 1947, the initially proposed algorithms were (and have remained) computationally impractical. Drawing on a series of obscure RAND technical reports, this talk tells the story of how the “revised” simplex method subsequently emerged to make today’s powerful solvers possible. The presentation concludes by considering how the earlier, impractical simplex algorithms have come to be adopted by almost all textbooks, while computationally practical versions remain known mainly to experts.
1948

Programming of Interdependent Activities II: Mathematical Model

- George B. Dantzig
- *Econometrica* 17 (1949)

“Linear Programming”

I called my first paper: *Programming in a Linear Structure*. In the summer of 1948, Koopmans and I visited the RAND Corporation. One day we took a walk near the Santa Monica beach. Koopmans said: “Why not shorten Programming in a Linear Structure to Linear Programming?” I replied: “That’s it! From now on that will be its name.” Later that same day I saved a talk at RAND entitled Linear Programming.
1948

Programming of Interdependent Activities II: Mathematical Model

- George B. Dantzig
- *Econometrica* 17 (1949)

“Linear Programming”
- Formulations & applications
- No algorithm

“It is proposed to solve linear programming problems . . . by means of large scale digital computers . . . . Several computational procedures have been evolved so far and research is continuing actively in this field.”
1949

Maximization of a Linear Function of Variables Subject to Linear Inequalities

- George B. Dantzig
- Activity Analysis of Production and Allocation (1951)

“Simplex Method”

The term Simplex Method arose out of a discussion with T. Motzkin who felt that the approach that I was using in the geometry of the columns was best described as a movement from one simplex to a neighboring one.
Maximization of a Linear Function of Variables Subject to Linear Inequalities

- George B. Dantzig
- Activity Analysis of Production and Allocation (1951)

"Simplex Method"

- Proof of convergence
- No computers

As a practical computing matter the iterative procedure of shifting from one basis to the next is not as laborious as would first appear..."
1953

An Introduction to Linear Programming

- W.W. Cooper, A. Henderson
- A. Charnes

“Simplex Tableau”

- Symbolic description
- Numerical example

“As far as computations are concerned it is most convenient to arrange the data at each stage in a ‘simplex tableau’ as shown in Table I."\(^{12}\)

\(^{12}\)A. Orden suggested this efficient arrangement developed by himself, Dantzig, and Hoffman.”
Terminology

*Linear program*

Minimize \( c \cdot x \)

Subject to \( A \cdot x = b \)

\( x \geq 0 \)

\( m \) constraints on \( n \) variables: \( m < n \)

*Data*

\( b = (b_1, \ldots, b_m) \)

\( c = (c_1, \ldots, c_n) \)

\( A = [a_{ij}] \), with \( m \) rows \( a^i \) and \( n \) columns \( a_j \)

*Variables*

\( x = (x_1, \ldots, x_n) \)
Terminology (cont’d)

Basis

- $\mathcal{B}, \mathcal{N}$, sets of basic and nonbasic column indices
  - $|\mathcal{B}| = m$, $|\mathcal{N}| = n - m$
- $c_\mathcal{B}, x_\mathcal{B}$, corresponding subvectors of $c, x$

Basis matrix

- $B$, nonsingular $|\mathcal{B}| \times |\mathcal{B}|$ submatrix of $A$
- $B^{-1} = [z_{ij}]$, with $|\mathcal{B}|$ rows $z^i$ and $|\mathcal{B}|$ columns $z_j$
Tableau Simplex Method

*Set up a* $(|B| + 1) \times (|N| + 1)$ *table of values*

$y_{ij}$, $i \in B$, $j \in N$: the transformed columns $y_j = B^{-1}a_j$

$y_{i0} \equiv x_i$, $i \in B$: the basic solution $x_B = B^{-1}b$

$y_{0j} \equiv d_j$, $j \in N$: the reduced costs $c_j - c_By_j$

*Choose an entering variable*

$p \in N$: $y_{0p} < 0$

*Choose a leaving variable*

$q \in B$: $y_{q0}/y_{qp} = \min_{y_{ip}>0} y_{i0}/y_{ip}$

*“Pivot” on the tableau*

$y_{ij} \leftarrow y_{ij} - y_{qj}y_{ip}/y_{qp}$: subtracts multiples of row $q$ from other rows

$y_{ip} \leftarrow -y_{ip}/y_{qp}$, $y_{qp} \leftarrow 1/y_{qp}$
Impracticalities

**Computational inefficiency**
- \(|B| \times |N| = m(n - m)|\) additions & multiplications
- \(|B| \times |N| |\) numbers to write and store

**Numerical instability**
- Fixed rules for choosing \(p\) and \(q\)
- *risking* small denominators in \(y_{ij} - y_{qj}y_{ip}/y_{qp}\)
- *causing* loss of precision in pivot steps
1953

The Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints

- George B. Dantzig, Alex Orden, Philip Wolfe
- Project RAND Research Memorandum RM-1264

“Lexicographic Simplex Method”

- Prevent cycling due to degeneracy
- Adapt computations accordingly

“The $k+1$st iterate is closely related to the $k$th by simple transformations that constitute the computational algorithm [6], . . .”
1953

Computational Algorithm of the Revised Simplex Method

- George B. Dantzig
- Project RAND Research Memorandum RM-1266

“Revised Simplex Method”

- Tableau replaced by basis inverse
- Computations streamlined

“The transformation of just the inverse (rather than the entire matrix of coefficients with each cycle) has been developed because it has several important advantages over the old method: . . .”
Revised Simplex Method

Given a matrix of inverse values
\[ z_{ij}, \ i \in \mathcal{B}, \ j \in \mathcal{B} \]: the basis inverse \( B^{-1} \)
\[ z_{i0} \equiv x_i, \ i \in \mathcal{B} \] (the basic solution)
\[ z_{0i} \equiv \pi_i, \ i \in \mathcal{B} \] (the dual prices)

Choose an entering variable
\[ p \in \mathcal{N}: \ c_p - z^0 \cdot a_p < 0 \]

Choose a leaving variable
\[ y_{ip} = z^i \cdot a_p \]
\[ q \in \mathcal{B}: \ z_{q0}/y_{qp} = \min_{y_{ip} > 0} z_{i0}/y_{ip} \]

“Pivot” on the inverse
\[ z_{ij} \leftarrow z_{ij} - z_{qj}z_{ip}/z_{qp} : \text{subtracts multiples of row } q \text{ from other rows} \]
\[ z_{ip} \leftarrow -z_{ip}/z_{qp}, \ z_{qp} \leftarrow 1/z_{qp} \]
Advantages

Smaller update

“... In the original method (roughly) $m \times n$ new elements have to be recorded each time. In contrast, the revised method (by making extensive use of cumulative sums of products) requires the recording of about $m^2$ elements ...”

Sparse operations

“In most practical problems the original matrix of coefficients is largely composed of zero elements. ... The revised method works with the matrix in its original form and takes direct advantage of these zeros.”

\[
Z_{ij} \leftarrow Z_{ij} - z_{qj} z_{ip} / Z_{qp}
\]

\[
d_p = c_p - z^0 \cdot a_p
\]

\[
y_{ip} = z^i \cdot a_p
\]
Impracticalities

Inefficiency

- $|\mathcal{B}| \times |\mathcal{B}| = m^2$ additions & multiplications
- $|\mathcal{B}| \times |\mathcal{B}|$ numbers to write and store

Numerical instability

- Fixed rules for choosing $p$ and $q$
- *risking* small denominators in $z_{ij} - z_{jq}z_{ip}/z_{qp}$
- *causing* loss of precision in pivot steps

However . . .

“In contrast, the revised method (by making extensive use of cumulative sums of products) requires the recording of about $m^2$ elements (and an alternative method [5] can reduce this to $m$ . . .).”
1953

Alternate Algorithm for the Revised Simplex Method

- George B. Dantzig, Wm. Orchard-Hays
- Project RAND Research Memorandum RM-1268

“Product Form for the Inverse”

- Fully exploit sparsity of coefficients
- Solve practical problems

“Using the I.B.M. Card Programmed Calculator, . . . where the inverse matrix is needed at one stage and its transpose at another, this is achieved simply by turning over the deck of cards representing the inverse.”
Product-Form Simplex Method

*Given*

\[ \mathbf{x}_B \text{ (the basic solution)} \]
\[ B^{-1} = E_k^{-1} E_{k-1}^{-1} \cdots E_2^{-1} E_1^{-1} \] (factorization of the basis inverse)

*Choose an entering variable*

\[ \mathbf{\pi} = \mathbf{c}_B E_k^{-1} E_{k-1}^{-1} \cdots E_2^{-1} E_1^{-1} \]
\[ p \in N: c_p - \mathbf{\pi} \cdot \mathbf{a}_p < 0 \]

*Choose a leaving variable*

\[ \mathbf{y}_p = E_k^{-1} E_{k-1}^{-1} \cdots E_2^{-1} E_1^{-1} \mathbf{a}_p \]
\[ q \in B: \frac{x_q}{y_{qp}} = \min_{y_{ip} > 0} \frac{x_i}{y_{ip}} \]

*Update*

- add a factor \( E_{k+1}^{-1} \) derived from \( \mathbf{y}_p \)
- update the basic solution: \( \mathbf{x}_B \rightarrow \mathbf{x}_B - (x_q/y_{qp}) \mathbf{y}_p \)
Factorization of the Inverse

Form of the factors
- $E_i^{-1}$ is an identity matrix except for one column

Computation of the factors
- Gauss-Jordan elimination
- Elimination ordering can be chosen to promote sparsity and stability

Storage of the factors
- nonzeros only, in (row,value) pairs
- diagonal element first

Update of the factors
- $E_{k+1}$ is an identity matrix except for $y_p$ in one column
Practical Simplex Method

*Given*
- $x_B$ (the basic solution)
- a factorization of $B$ suitable for solving equations fast

*Choose an entering variable*
- solve $B^T \pi = c_B$
- $p \in \mathcal{N}$: $c_p - \pi \cdot a_p < 0$

*Choose a leaving variable*
- solve $By_p = a_p$
- $q \in \mathcal{B}$: $x_q/y_{qp} = \min_{y_{ip} > 0} x_i/y_{ip}$

*Update*
- update factorization to reflect change of basis
- update basic solution to $x_B \leftarrow (x_q/y_{qp}) y_p$
“. . . the simplex algorithm . . . starts with a canonical form, consists of a sequence of pivot operations, and forms the main subroutine of the simplex method.”

“Because some readers might find that the matrix notation of §8.5 [The Simplex Algorithm in Matrix Form] obscures the computational aspects, we have tended to avoid its use here.”
Tableau Simplex Method Revisited

Simple
- No linear algebra
- No matrices & inverses
- All computations in one “pivot” step
- Easy to set up for hand calculation

Familiar
- Textbooks presented it
- Students learned it
- Some students wrote new textbooks . . .

But still impractical
1968

**Advanced Linear-Programming Computing Techniques**

- William Orchard-Hays

“Except for [a few sections], the contents of the book reflect actual and extensive experience.”

“I hope that the many users of mathematical programming systems implemented on today’s large computers find the book valuable as background for the largely undocumented algorithms embedded in these systems. If it should also be found useful as a course text, all objectives will have been achieved.”
Essential Simplex Method

**Given**
- \(x_B\) (the basic solution)
- \(B\) (the basis)

**Choose an entering variable**
- solve \(B^T\pi = c_B\)
- \(p \in \mathcal{N}: c_p - \pi \cdot a_p < 0\)

**Choose a leaving variable**
- solve \(By_p = a_p\)
- \(q \in \mathcal{B}: x_q/y_{qp} = \min_{y_{ip}>0} x_i/y_{ip}\)

**Update**
- update basic solution to \(x_B - (x_q/y_{qp}) y_p\)
Essential Simplex Method

Course notes for use in teaching

- Optimization Methods I: Solving Linear Programs by the Simplex Method
- https://www.4er.org/CourseNotes

Slides Available at ampl.com

Recent and upcoming events

- https://ampl.com/resources/calendar/

News & events archive

1978

History of Mathematical Programming Systems

- William Orchard-Hays

“Overview of an Era”

- Better implementations
- More powerful computers

“One cannot clearly comprehend the development of mathematical programming software without reference to the development of the computing field itself.”
First, mathematical programming and computing have been contemporary in an almost uniquely exact sense. Their histories parallel each other year by year in a remarkable way.

Furthermore, mathematical programming simply could not have developed without computers. Although the converse is obviously not true, still linear programming was one of the important and demanding applications for computers from the outset.

The quarter century from the late 1940s to the early 1970s constituted an era, one of the most dynamic in the history of mankind. Among the many technological developments of that period — and indeed of any period — the computing field has been the most virulent and astounding.

. . . the nature of the computing industry, profession, and technology has by now been determined — all their essential features have existed for perhaps five years. One hopes that some of the more recent developments will be applied more widely and effectively but the technology that now exists is pretty much what will exist, leaving aside a few finishing touches to areas already well developed, such as minicomputers and networks.
1981

Reminiscences About the Origins of Linear Programming

- George B. Dantzig
- Operations Research Letters
  1 (1982)

“Linear programming is viewed as a revolutionary development”

- System of linear inequalities
- Objective function
- Practical computational method

“Before closing let me tell some stories about how various linear programming terms arose.”