CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION, CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

Accession # 604 711

☒ 1. We regret that legibility of this document is in part unsatisfactory. Reproduction has been made from best available copy.

☐ 2. A portion of the original document contains fine detail which may make reading of photocopy difficult.

☐ 3. The original document contains color, but distribution copies are available in black-and-white reproduction only.

☐ 4. The initial distribution copies contain color which will be shown in black-and-white when it is necessary to reprint.

☐ 5. Limited supply on hand; when exhausted, document will be available in Microfiche only.

☐ 6. Limited supply on hand; when exhausted document will not be available.

☐ 7. Document is available in Microfiche only.

☐ 8. Document available on loan from CFSTI (TT documents only).

☐ 9.

Processor:

NBS____ 9/64

THE ELIMINATION FORM OF THE INVERSE AND ITS APPLICATION TO LINEAR PROGRAMMING

Harry Markowitz

P-680

8 April 1955

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

20)

ACKNOWLEDGEMENT

The writer is indebted to George Dantzig and Alan Manne for valuable suggestions.

THE ELIMINATION FORM OF THE INVERSE AND ITS APPLICATION TO LINEAR PROGRAMMING

The inverse (A^{-1}) of a matrix (A) is valuable when a number of sets of equations AX = b are to be solved using different b's and the same A. A^{-1}, like any matrix, may be expressed as the product of other matrices

$$A^{-1} = M_m M_{m-1} \cdots M_1$$

in an infinite number of ways. E.g. (2) = (1/2) (4) = (1/8) (16) etc. If we have such $M_1, \ldots, M_m$ we can solve AX = b, $X = A^{-1}b$ in a series of steps:

$$X^{(1)} = M_1 b$$
$$X^{(2)} = M_2 X^{(1)}$$
$$\vdots$$
$$X = M_m X^{(m-1)}$$

The expression $M_m \cdots M_1$ is referred to as a "product form" of inverse. In some problems there may be M_i which are easier to obtain and apply than A^{-1} itself.

This paper will discuss a particular product form of inverse which is closely related to the Gaussian elimination method of solving a set of simultaneous equations. This "elimination form of the inverse," as we shall call it, is especially valuable when A has a large number of zero coefficients. If A has no zero coefficients, on the other hand, the elimination form of inverse is still generally as convenient as the conventional A^{-1}.

The elimination form of inverse can be illustrated in terms of the

solution of three equations in three unknowns:

1) $a_{11} X_1 + a_{12} X_2 + a_{13} X_3 = r_1$

2) $a_{21} X_1 + a_{22} X_2 + a_{23} X_3 = r_2$

3) $a_{31} X_1 + a_{32} X_2 + a_{33} X_3 = r_3$

For the moment we will let the k^{th} diagonal element be the k^{th} pivotal element. From equation 1) we get the first equation of our back solution

B1) $X_1 = \frac{r_1}{a_{11}} - \frac{a_{12}}{a_{11}} X_2 - \frac{a_{13}}{a_{11}} X_3$

We eliminate X_1 from equations 2) and 3) by adding $\left(- \frac{a_{i1}}{a_{11}} \right)$ times the first equation to the i^{th} equation, thus obtaining

2') $b_{22} X_2 + b_{23} X_3 = r_2^*$

3') $b_{32} X_2 + b_{33} X_3 = r_3^*$

where

$$r_2^* = r_2 - \left(\frac{a_{21}}{a_{11}}\right) r_1$$

$$r_3^* = r_3 - \left(\frac{a_{31}}{a_{11}}\right) r_1$$

Similarly we get

B2) $X_2 = \frac{1}{b_{22}} r_2^* - \frac{b_{23}}{b_{22}} X_3$

and $c_{33} X_3 = r_3^{**}$

where $r_3^{**} = r_3^* - \frac{b_{32}}{b_{22}} r_2^*$.

Finally

B3) $X_3 = \frac{1}{c_{33}} r_3^{**}$

B3) gives us X_3; X_3 and B2) give X_2; X_2, X_3 and B1) give X_1.

Consider the transformations which occurred to our original right hand side. First we formed

$$\begin{pmatrix} r_1 \\ r_2^* \\ r_3^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$$

then

$$\begin{pmatrix} r_1 \\ r_2^* \\ r_3^{**} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{b_{32}}{b_{22}} & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2^* \\ r_3^* \end{pmatrix}$$

then

$$\begin{pmatrix} r_1 \\ r_2^* \\ X_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{b_{22}} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2^* \\ r_3^{**} \end{pmatrix}$$

then

$$\begin{pmatrix} r_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{b_{22}} - \frac{b_{23}}{b_{22}} & \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2^* \\ X_3 \end{pmatrix}$$

and finally

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{a_{11}} - \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ X_2 \\ X_3 \end{pmatrix}$$

Thus

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{a_{11}} & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{b_{22}} & -\frac{b_{32}}{b_{22}} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{c_{33}} \end{pmatrix}$$

$$\cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{b_{32}}{b_{22}} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$$

Since the a_{ij} b_{ij} c_{ij} do not depend on the r_i we have

$$A^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & 1 \end{pmatrix}$$

Similarly if A is any n X n non-singular matrix we have

$$A^{-1} = B_1 B_2 \cdots B_n R_{n-1} \cdots R_1$$

where R_k is a matrix of the form

$$\begin{pmatrix} 1 & & & & & & 0 \\ & 1 & & & & & \vdots \\ & & \ddots & & & & \vdots \\ & & & \ddots & & & \vdots \\ & & & & 1 & & 0 \\ & & & & & 1 & \\ & & & & -\frac{v_{k+1,k}}{v_{kk}} & & \\ & & & & \vdots & & \\ & & & & -\frac{v_{nk}}{v_{kk}} & & \end{pmatrix}$$

$$\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

and B_k is of the form

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{1}{v_{kk}} & -\frac{v_{kk+1}}{v_{kk}} & \cdots & -\frac{v_{km}}{v_{kk}} \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Although this elimination form of inverse consists of 2m-1 matrices it contains only n^2 numbers which are not known a-priori to be 0 or 1. It requires the same number of multiplications and additions to apply the elimination form as it does to apply the conventional A^{-1}. The arithmetic operations required to obtain the $B_1 \ldots B_m \; R_{m-1} \ldots R_1$ and to apply them once are exactly the same as those required to solve a set of linear equations by Gaussian elimination.

Suppose that the k^{th} pivot is not v^k_{kk} but $v^k_{i_o j_o}$ (where v^k_{ij} is the value of the parameter of i^{th} equation j^{th} variable at the k^{th} step). This step transforms the right-hand side by

$$(r)^{new} = \begin{pmatrix} 1 & & & & \\ & \ddots & \gamma_1 & & \\ & & \vdots & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ & & \gamma_m & & 1 \end{pmatrix} (r)^{old}$$

where at least k-1 of the $\gamma_i = 0$ and where the γ vector is in the i_o^{th} column of the matrix. The k^{th} step also gives rise to a back solution

$$X_{j_o} = \frac{1}{v^k_{i_o j_o}} r^k_{i_o} + \Sigma \eta_j X_j$$

where at least m-k of the η_j equal zero.

The elimination form of the inverse in this case is

$$A^{-1} = B_1 \ldots B_m \, P \, R_{m-1} \ldots R_1$$

where

$$B_k = \begin{matrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ \eta_1 & \cdots & \eta_{i_o} & \cdots & \eta_m & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \end{matrix}$$

$$R_k = \begin{matrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \gamma_1 & & \\ & & \vdots & & \\ & & \gamma_m & & \\ & & & \ddots & \\ & & & & 1 \end{matrix}$$

and P is permutation matrix such that if v^k_{ij} is a pivotal element P makes the old i^{th} component of (r) become the new j^{th} component of r.

That $A^{-1} = B_1 \ldots P \ldots R_1$, i.e., that $AX = r$ implies $X = B_1 \ldots P \ldots R_1 \; r$ may be seen as follows. The transformations $R_{m-1} \ldots R_1 \; r$ give the right hand sides of the back solutions in the original equation order. We can imagine reordering the equations so that new first back solution is the one of the form

$$X_1 = \frac{1}{v^k_{il}} \; r^k_i + \Sigma \eta X$$

and similarly the new k^{th} back solution is the one in which X_k was eliminated. This rearrangement changes the order of the r_i exactly as does P. The last back solution is of the form

$$X_{j_m} = \frac{1}{v} \; r_{o_m}$$

Since r_{i_m} is now the j_mth component of (r), B_m (as described above) will transform (r) into a vector with X_{j_m} in the j_mth spot. The next back solution is of the form

$$X_{j_{m-1}} = -\frac{1}{\nu} r_{i_{m-1}} + \eta X_{j_m}$$

Since, thanks to B_m and P, $r_{i_{m-1}}$ is the j_{m-1}th component and X_{j_m} is the j_mth component of (r), B_{m-1} transforms (r) into a vector with $X_{j_{m-1}}$ in the j_{m-1}th spot as well as X_{j_m} in the j_mth spot. And so on.

In recording B_k or R_k in, say, a hand computation it is not necessary to write out the entire matrix. In the case of an R_k it is only necessary to record the i eliminated and the non-zero γ_i. In the case of a B_k it only is necessary to record the j eliminated, the non-zero n_j and 1/pivotal element. If, for example, the pivotal elements v^1_{13}, v^2_{31}, v^3_{22} are used in inverting a 3 x 3 matrix, the elimination form of the inverse might be recorded as in Table 1.

Table 1.

Right Hand Side Transformations

k	i eliminated	i	γ	i	γ
1	1	2	6.42	3	8.15
2	3	2	9.10		
3	2				

m Back Solution

k	j eliminated	1/(pivotal element)	j	η	j	η
1	3	.84	1	1.08	2	6.45
2	1	5.92	2	3.18		
3	2	1.08				

In machine computation as well only the non-trivial parts of B_k and R_k need be stored.

If the matrix A has a large number of zero α_{ij}, the elimination form of inverse may have appreciably less than n^2 non zero γ s and η s. This may be so even though A^{-1} has no zeros. Thus the matrix

$$A = \begin{matrix} 1 & 0 & 0 & -1/2 \\ -1/2 & 1 & 0 & 0 \\ 0 & -1/2 & 1 & 0 \\ 0 & 0 & -1/2 & 1 \end{matrix}$$

has a conventional inverse

$$A^{-1} = \begin{matrix} 16/15 & 2/15 & 4/15 & 8/15 \\ 8/15 & 16/15 & 2/15 & 4/15 \\ 4/15 & 8/15 & 16/15 & 2/15 \\ 2/15 & 4/15 & 8/15 & 16/15 \end{matrix}$$

and has an elimination form, using the diagonal elements as pivots, of

Right Hand Side Transformations

k = i elim.	i	γ
1	2	1/2
1	3	1/2
1	4	1/2

Back Solution

k = j elim.	1/pivotal elem.	j	η
1	1	4	1/2
2	1	4	1/4
3	1	4	1/8
4	16/15		

Thus while A has 8 non-zeros, A^{-1} has 16 non-zeros and the elimination form has 10 non-zeros.

The number of non-zero η and γ in an elimination form of inverse may depend on which pivotal elements are used. Suppose the *'s below represent the non-zero elements of a 5 x 5 matrix.

```
* * * * *
* * * *
* * *
* *
*
```

If α_{11} is the first pivotal element the non-zeros at step two are (barring accidental zeros) as follows

```
* * * *
* * * *
* * * *
* * * *
```

But if α_{15} or α_{51} is the pivotal element the pattern of non-zeros is

```
* * * *
* * *
* *
*
```

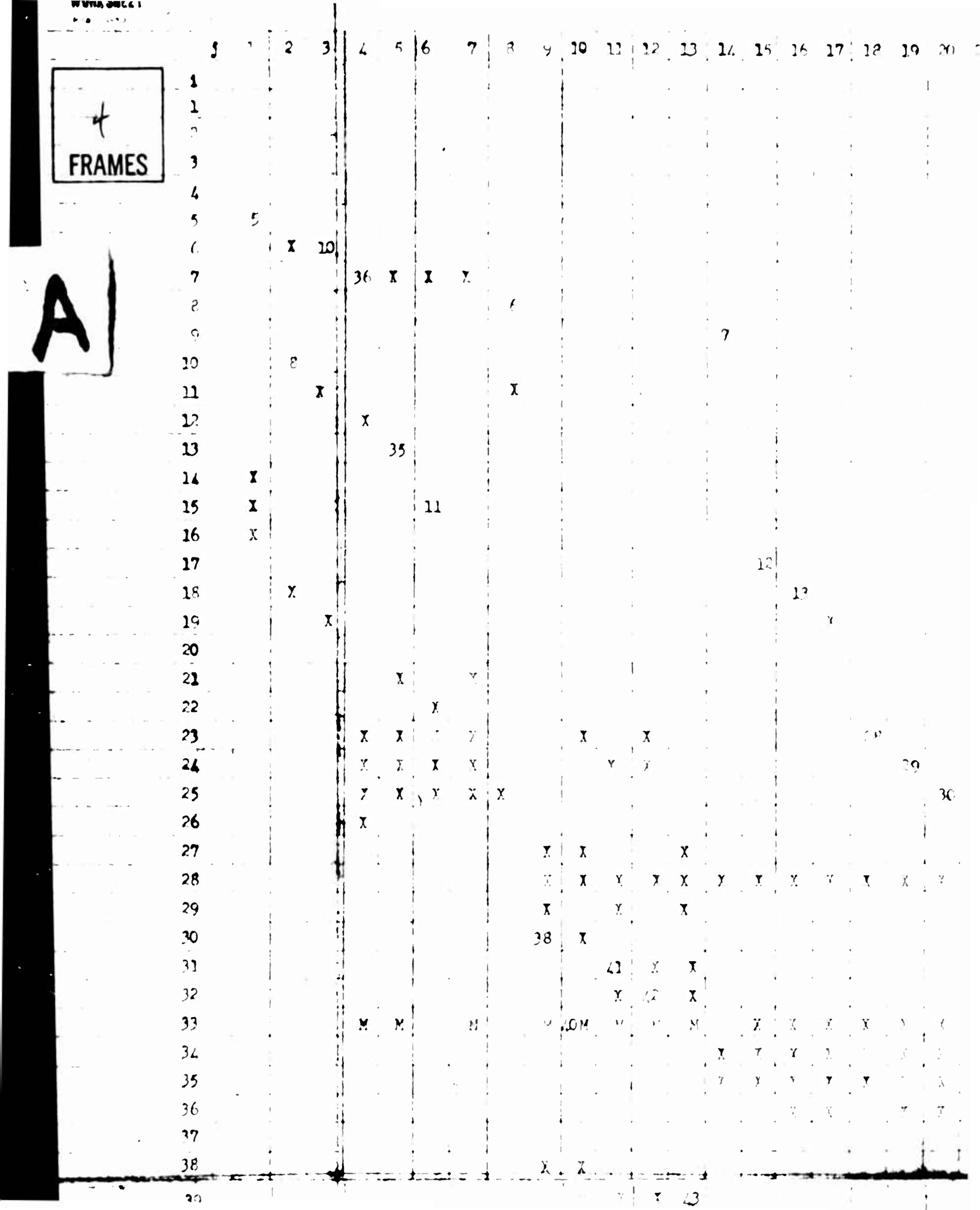
A table indicating zero and non-zero coefficients is a valuable aid in the choice of pivotal elements. From such a table an agenda (i.e., a complete set of pivotal elements) can be chosen before computation begins. This separation of the choice of the agenda and actual elimination is convenient both in hand and machine computation. There are, however, two dangers attached to deciding on an agenda beforehand. Some pivotal element may accidentally become zero, in which case the agenda cannot be followed to the end. Or some pivotal element may turn out to be so small that its use would adversely affect the accuracy of the results. The solution to these difficulties seems to be to have some test of the acceptability of a pivotal element; form the agenda beforehand and follow it as long as each pivotal element meets the test. If a pivotal element fails the test, a new agenda can be worked out for the

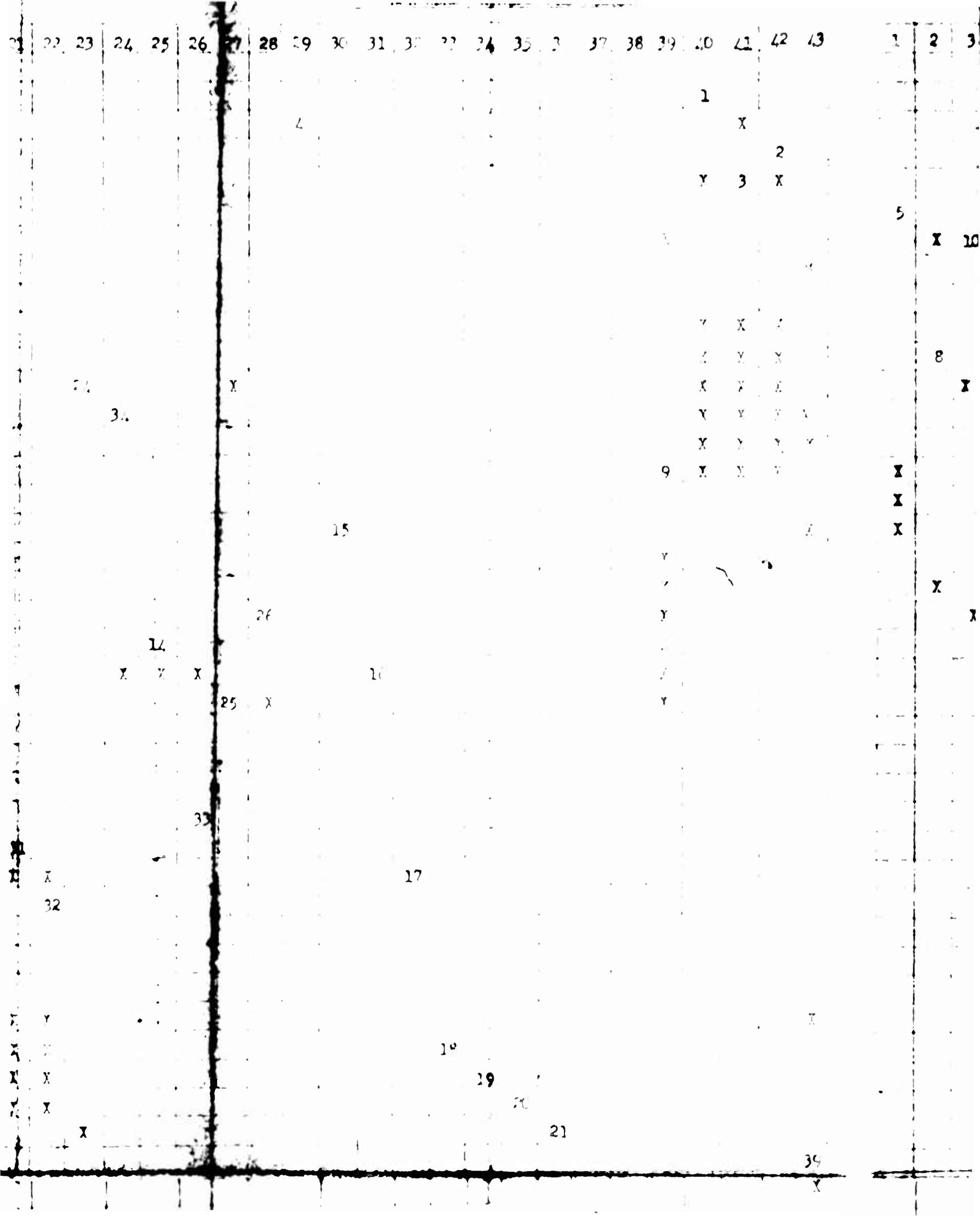
remaining equations and variables. When A has many zero coefficients the accuracy of the and is improved by the fact that less operations are required to obtain them.

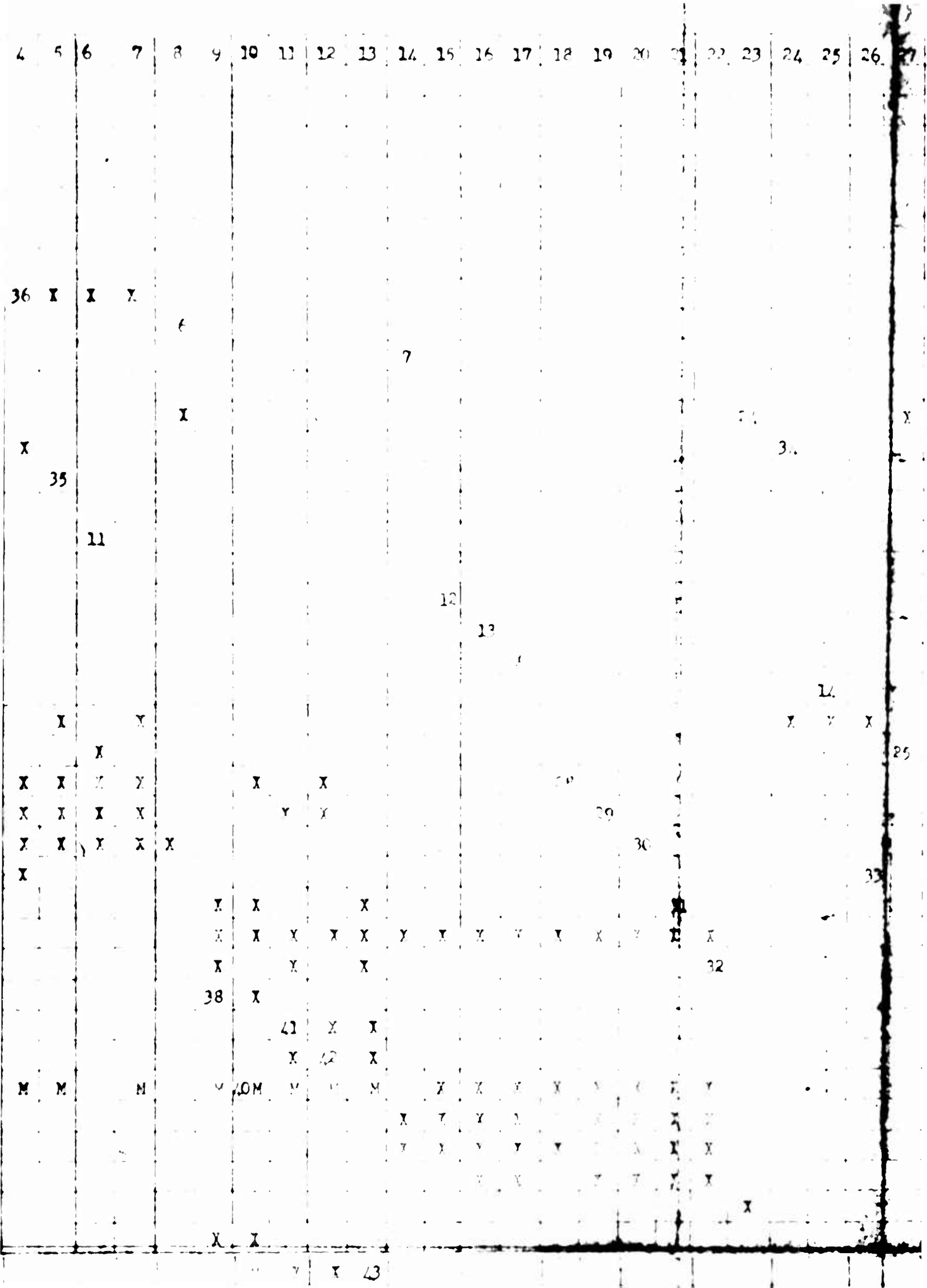
An example agenda is presented in Table 2. The X's represent the original non-zero elements of the matrix. The M's represent coefficients which began as zeros but ended as non-zeros. The numbers k = 1, . . .,43 in the matrix indicate the k^{th} pivotal element. The number (o_i) at the right of each row indicates the number of elements of that row which were not already eliminated when the row was eliminated. The number o_j at the bottom of each column indicates the number of elements of that column which were not already eliminated when the column was eliminated. One of the by-products of making an agenda beforehand is foreknowledge of all the variables which will appear in any equation and all the equations in which a variable will ever appear.

The matrix which Table 2 represents was the optimum basis of a linear programming problem involving a 43-equation model of petroleum industry. This matrix has 197 non-zero elements. As compared with a possible $(43)^2$ = 1849, the number of non-zero elements in the elimination form of inverse is $\Sigma \rho_i + \Sigma (\sigma_j - 1) = \Sigma \rho_i + \Sigma \sigma_j - 43 = 201$. To derive this inverse requires $\sum_{(i,j \text{ pivotal})} \rho_i \sigma_j = 247$ multiplications or divisions and somewhat less additions or subtractions.

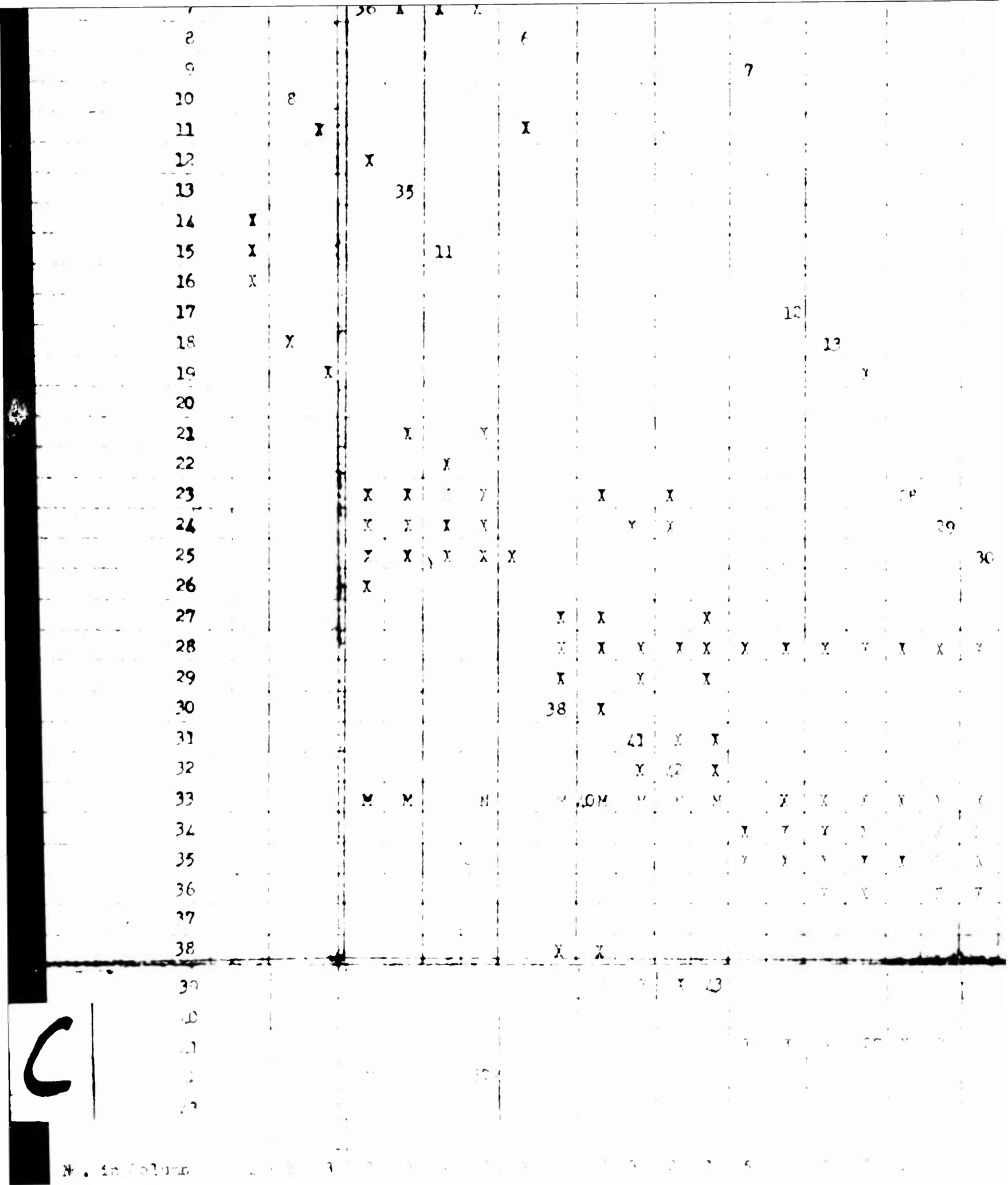
It would be desirable to choose an agenda so as to minimize the number of zeros which become non-zero. In some cases it is harder to find such an "optimum" agenda than to invert a matrix. An alternative is to choose at each step the pivot which minimizes the number of zeros which become non-zero at that step. A still simpler alternative, which seems adequate generally, is to choose the pivot which minimizes the number of coefficients modified at

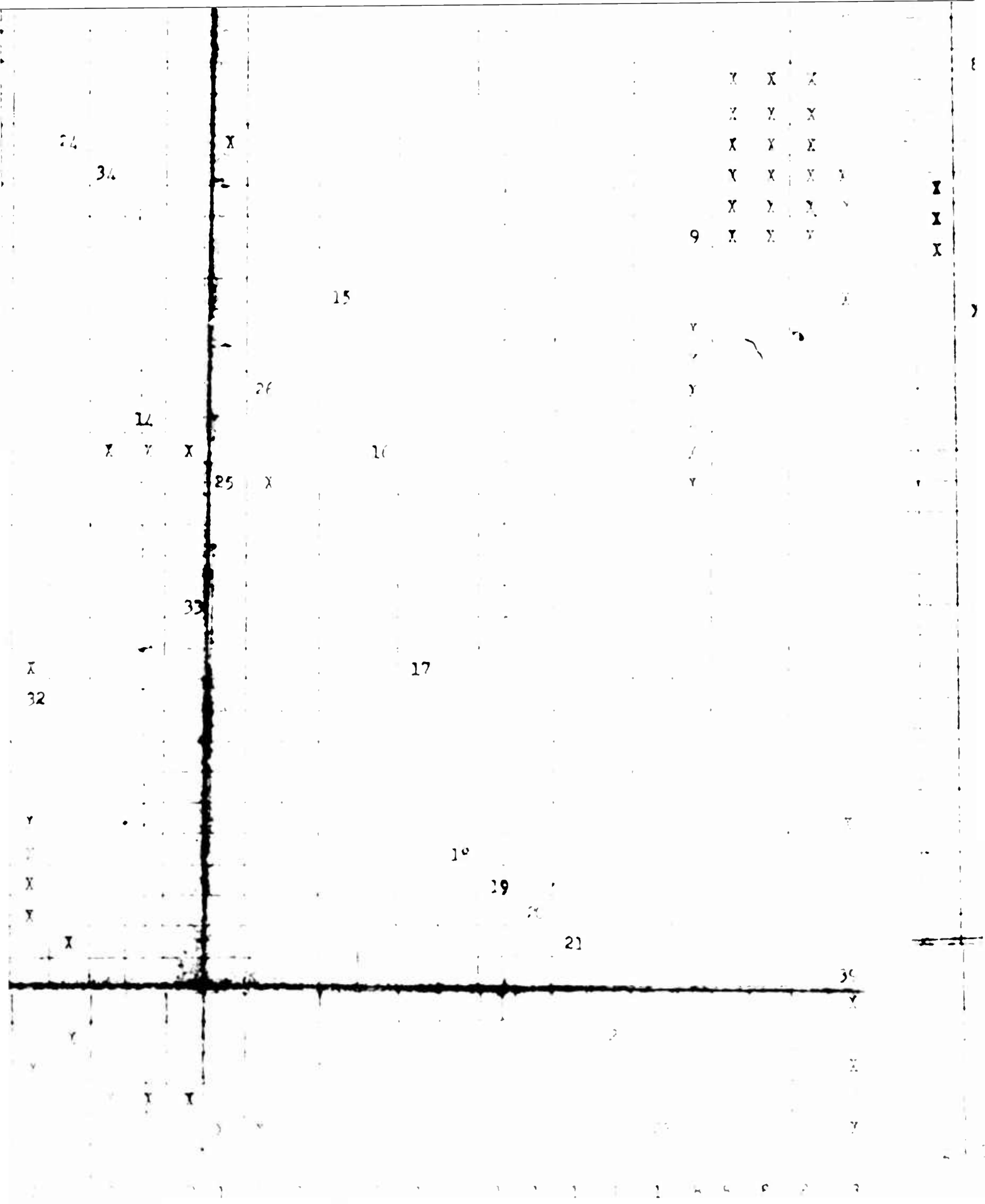

FRAMES

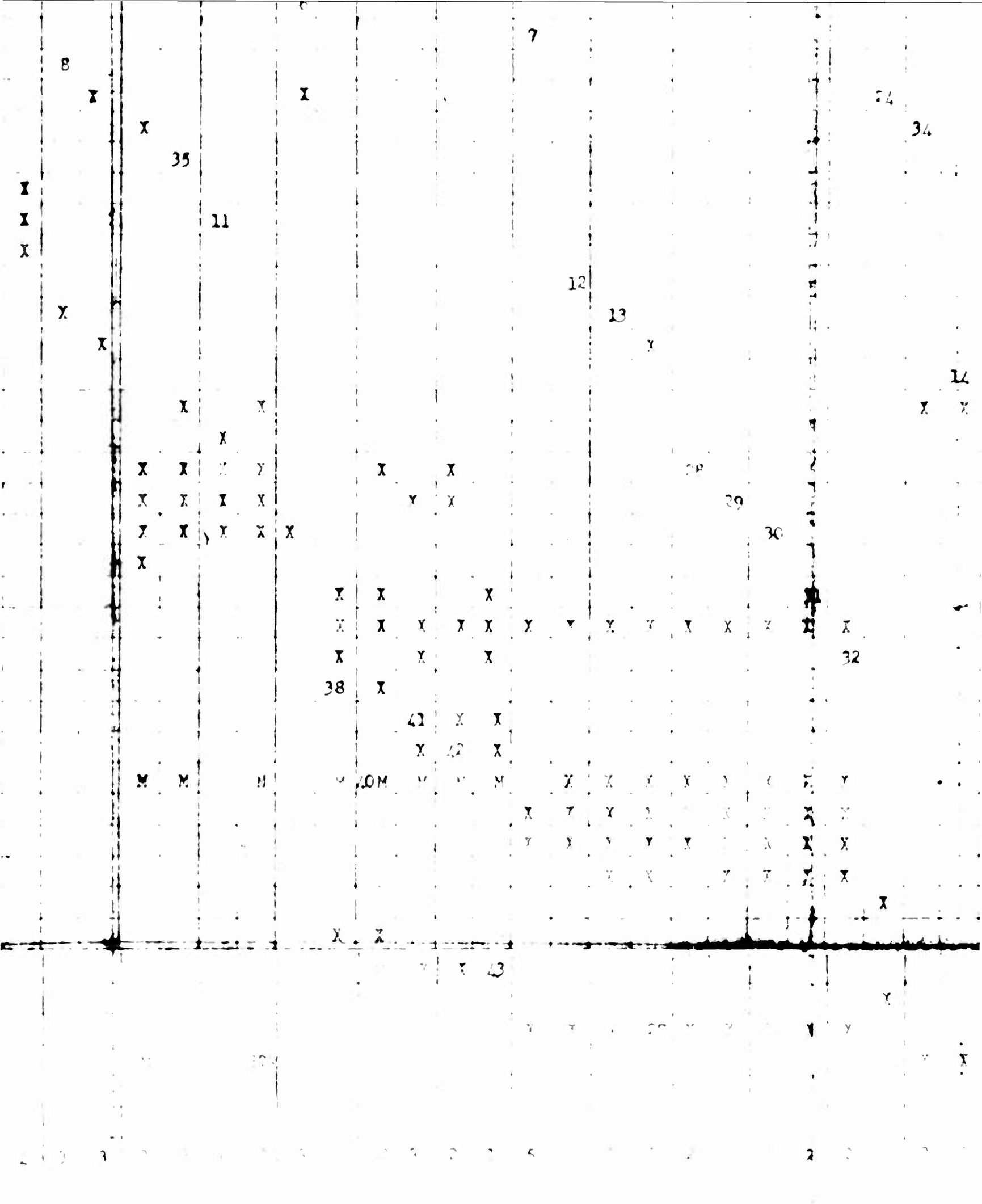


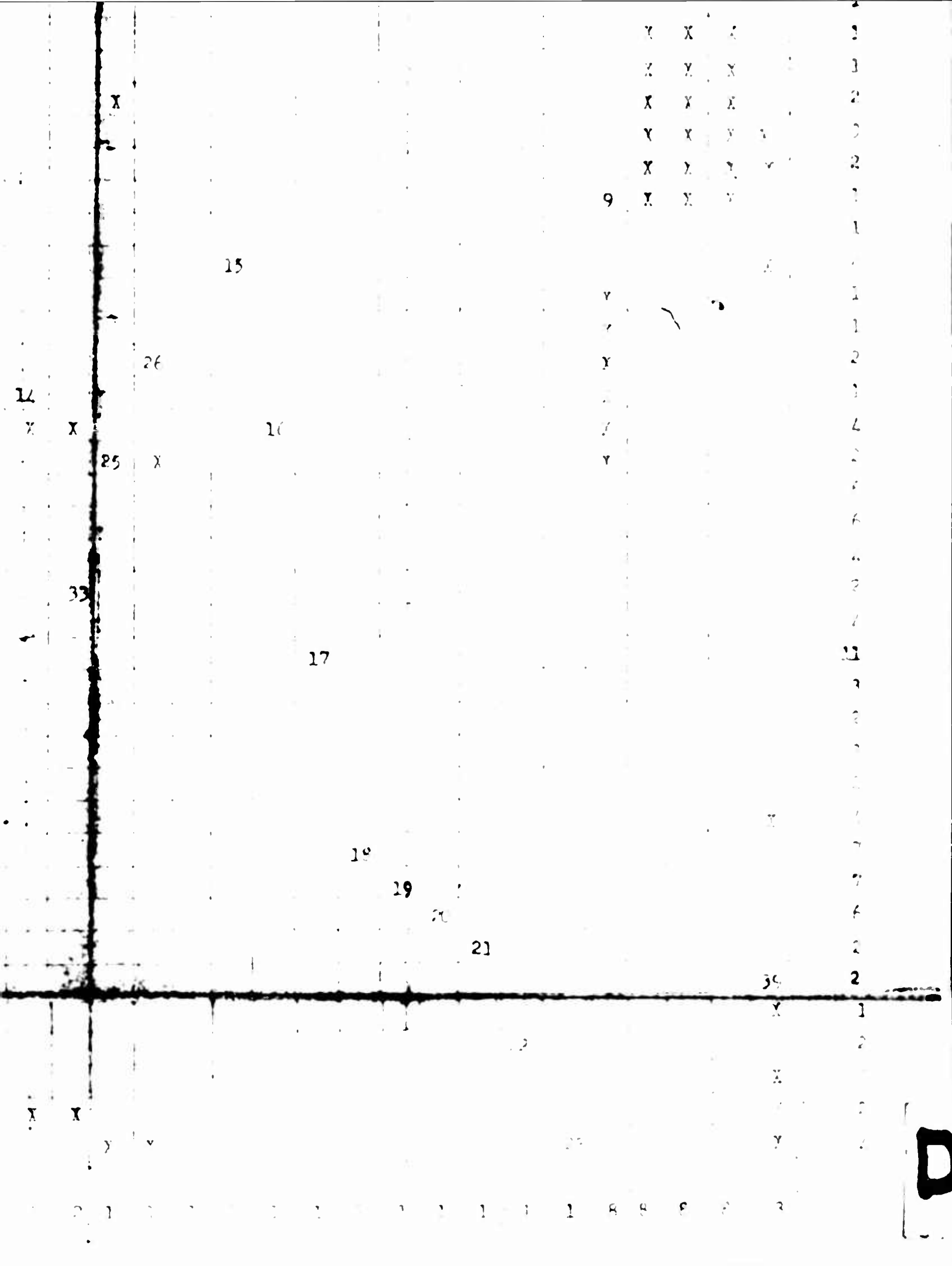


27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	No. in Row when Eliminated
													1				1
		4												X			1
															2		
													Y	3	X		
																	1
																	1
																	1
X													X				
													Y	Y			
													Y				
												9	X				
			15														
	26																
				16													
5	X											Y					
					17												11
						18											
							19										
									21								
																37	2









each step (excluding those which are eliminated at the particular step). This is equivalent to choosing the non-zero element with minimum $(\rho_i - 1)(\sigma_j - 1)$.

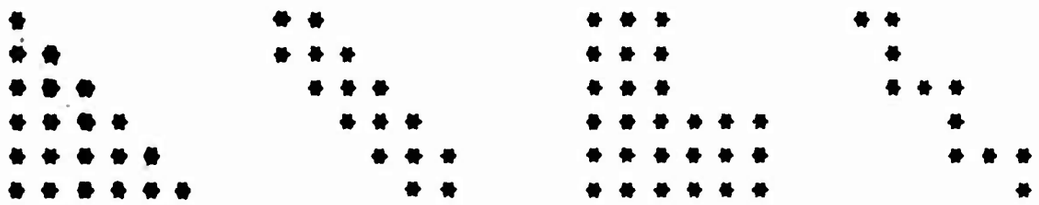
Figure 1 illustrates matrices with the following properties:

a) all diagonal elements α_{kk} are non-zero

b) if an element α_{ik} or α_{kj} in the k^{th} column or row is non-zero, then all elements between α_{ik} (or α_{kj}) and α_{kk} are non-zero.

It is trivial to find an optimum agenda for such matrices. If the k^{th} diagonal element is used as the k^{th} pivot, no zero coefficient will become non-zero. If a matrix is almost of the above form except that a few zeros are mixed in with the non-zeros, then using the k^{th} diagonal element as the k^{th} pivot may cause the "misplaced" zeros (and only these) to become non-zero.

Figure 1.



APPLICATION TO LINEAR PROGRAMMING

The simplex method for solving linear programming problems has a number of variants.* A recent version requires the solution of two sets of equations. The first set of equations

$$p'A = \gamma' \quad \text{or} \quad A'p = \gamma$$

yields prices p which are used to select a variable X_s, currently equal to zero, to be increased until some "basis" variable becomes zero. The second set of equations is

$$Ae = P_s$$

where P_s is a column vector associated with the variable X_s. e is used to

*G. B. Dantzig, Alex Orden, & Philip Wolfe, "The Generalized Simplex Method," RAND P-392-1, 4 August 1953.

determine which basis variable I_r first becomes zero as X_s is increased. The matrix A at one iteration differs from that at the next only in that the column vector P_r, which we will assume to be the r^{th} column of A, is replaced by the column vector P_s. I.e.,

$$A^{(k)} = (P_{j_1} \ldots P_r \ldots P_{j_m})$$

becomes

$$A^{(k+1)} = (P_{j_1} \ldots P_s \ldots P_{j_m})$$

Or

$$A^{(k+1)} = (P_{j_1} \ldots P_r \ldots P_{j_m}) \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & e_1 & & \\ & & \vdots & & \\ & & e_r & & \\ & & \vdots & \ddots & \\ & & e_m & & 1 \end{pmatrix}$$

Letting $E = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & e_1 & & \\ & & \vdots & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & e_m \end{pmatrix}$ we have

$$A^{(k+1)} = A^k E$$

The inverse of $A^{(k+1)}$ is

$$A^{(k+1)^{-1}} = E^{-1} A^{(1)^{-1}}$$

where

$$E^{-1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & -e_1/e_r & & \\ & & \vdots & & \\ & & 1/e_r & & \\ & & \vdots & \ddots & \\ & & -e_m/e_r & & 1 \end{pmatrix}$$

In an early variant of the simplex method the new inverse $A^{(k+1)^{-1}}$ was obtained at each step by multiplying out $E^{(k)^{-1}} A^{(k)^{-1}}$.* In a more recent version the $E^{(k)^{-1}}$ are carried along and used as product form of inverse.** If the first iteration starts with the identity matrix then

$$A^{(k)^{-1}} = E^{(k-1)^{-1}} \cdot E^{(k-2)^{-1}} \cdots E^{(1)^{-1}}.$$

If the product form of inverse is used, as k increased computing time required per iteration also increases. A point is eventually reached when it is desirable to reinvert the current basis $A^{(k_0)}$ and let

$$A^{(k_0+k)^{-1}} = E^{(k_0+k-1)^{-1}} \cdots E^{(k_0)^{-1}} A^{(k_0)^{-1}}$$

At this point the elimination form of inverse can be of value, especially if A has a large number of zeros, since this form requires less time to obtain and apply.

Reinverting A is only part of the operations involved in solving a linear programming problem. We therefore cannot expect to obtain, by the use of the elimination form, the same percentage reduction in computing time for the whole linear programming problem as we obtain for the reinversion of A. When a more convenient form of inverse is available it may be desirable to reinvert more frequently. To see the effect of a more convenient form of inverse on the frequency of reinversion and the time required to solve a linear programming problem, we must explore the question of optimum reinversion policy.

We will first derive some neat results under several restrictive assumptions. Afterwards we will show computing procedures for obtaining an optimum reinversion policy under more general assumptions.

---

*See George B. Dantzig, "Maximization of a Linear Function of Variables Subject to Linear Inequalities," pp. 339-347, and Robert Dorfman, "Application of the Simplex Method to a Game Theory Problem," p. 358 in Activity Analysis of Production and Allocation, T. C. Koopmans, Ed., New York, 1951

**See George B. Dantzig, "The Product Form for the Inverse in the Simplex Method," in Mathematical Tables and Other Aids to Computation, VIII, No. 46, April, 1954.

It has been observed, with stop watch as well as theory, that, with the RAND linear programming code for the IBM 701, the computing time required per iteration increases linearly with the number of iterations. Let us assume, for the moment, that (a) the problem starts with a first basis to be inverted; (b) the time required for this first inversion is the same as that for any subsequent reinversion; (c) the computing time required since the beginning of a (re)inversion is a quadratic function of the number of iterations since then. $t = \alpha + \beta I + \gamma I^2$ where $\alpha, \beta, \gamma > 0$. Let us further assume, for the moment, that the number of iterations $\bar{I}$ required to solve the problem is known beforehand.

Suppose it were decided that there would be n inversions (n - 1 reinversions). Let ΔI_i = the number of iterations between the i^{th} and $i+1^{th}$ inversion (for i = 1,..,n - 1). Let ΔI_n = the number of iterations from the n^{th} inversion to the end of the problem. Total time (T) required is

$$T = \sum_{i=1}^{n} \alpha + \beta \Delta I_i + \gamma (\Delta I_i)^2$$

where

$$\sum_{i=1}^{n} \Delta I_i = \bar{I}$$

The optimum solution must satisfy the Lagrangian equations

$$\frac{\partial \sum_i (\alpha + \beta \Delta I_i + \gamma (\Delta I_i)^2) - \lambda(\Delta I_i)}{\partial \Delta I_{i_0}} = 0$$

$\therefore \beta + 2\gamma \Delta I_i - \lambda = 0$ for all i

$\therefore \Delta I_i$ is the same for all i

We can therefore rewrite the expression for T as

$$T = n(\alpha + \beta I + \gamma I^2)$$

where $I = \Delta I_i = \frac{\bar{I}}{n}$

Or

$$T = \alpha n + \beta \bar{I} + \frac{\gamma \bar{I}^2}{n}$$

$$\frac{dT}{dn} = \alpha - \frac{\gamma \bar{I}^2}{n^2}$$

$$\frac{d^2T}{dn^2} = 2\frac{\gamma \bar{I}^2}{n^3}$$

Since $\frac{d^2T}{dn^2} > 0$ for all $n > 0$, and since $T \to \infty$ as $n \to 0$, any $n > 0$ with $\frac{dT}{dn} = 0$ gives a minimum value of T for all $n > 0$. If such an n is non-integral the best integral value is either that immediately above $\hat{n}$, or that immediately below $\hat{n}$, or both

When $\frac{dT}{dn} = 0$

$$\hat{n} = \sqrt{\frac{\gamma}{\alpha}}\ \bar{I}$$

Let us assume that n is integral. Then the optimum

$$\hat{I} = \hat{\Delta} I = \sqrt{\frac{\alpha}{\gamma}}$$

$$\hat{T} = \alpha\sqrt{\frac{\gamma}{\alpha}}\ \bar{I} + \beta\bar{I} + \gamma\sqrt{\frac{\alpha}{\gamma}}\ \bar{I} = (\beta + 2\sqrt{\alpha\gamma})I$$

The last expression can be used for estimating the time to be saved by using a more convenient form of inverse. Thus--given our various assumptions--if a new method of inversion could produce an inverse in one-fourth the time (α) and because of its compactness it permitted the first subsequent iteration to be done in one-half the time (β, roughly), the whole linear programming problem could be done in one-half the time.

Let us now suppose that $\bar{I}$ is not known but has an a-priori probability distribution (derived presumably from past linear programming problems). We may as well also drop the quadratic assumption on t. We define

a_{ij} = the expected value of r_{ij}

where

$$r_{ij} = \begin{cases} 0 \text{ when } \bar{I} < i \\ \text{The (expected) time required to (re)invert the matrix at } i \text{ and iterate through } j\text{-}1 \text{ without reinverting--if } \bar{I} \geq j. \\ \text{The expected time from the beginning of the reinversion at it to the end of the problem if } i \leq I < j. \end{cases}$$

Suppose the points of reinversion are before iterations $I_1, \ldots, I_K$ (since reinversion points can be chosen with I_K so large that there is a zero probability that this iteration will occur, there is no loss of generality in assuming a fixed K). Expected time, to be minimized, is

$$E = a_{1 I_1} + a_{I_1 I_2} + \ldots + a_{I_{K-1} I_K}$$

The optimal value of $\hat{I}_1$ of I_1 could be calculated if $\hat{I}_2$ were known. It is given by the function $\hat{I}_1(I_2)$ which minimizes $a_{1 I_1} + a_{I_1 I_2}$ for various values of I_2 and which can be readily computed.

Define

$$a_{I_{2/I_1}} = a_{1 \hat{I}_1(I_2)} + a_{\hat{I}_1(I_2), I_2}$$

We now only need to minimize

$$E = a_{I_{2/I_1}} + a_{I_2 I_3} + \ldots + a_{I_{K-1} I_K}$$

We repeat the process until we have

$$E = a_{I_{K/I_1, \ldots I_{K-1}}}$$

from which we get I_K and work back through

$\hat{I}_{K-1}(\hat{I}_K)$, $\hat{I}_{K-2}(\hat{I}_{K-1})$, etc.

POSTSCRIPT

It is common for the matrices in industrial applications of linear programming to have a large proportion of zero coefficients. While every item (raw material, intermediate material, end item, equipment item) in, say, a steel plant may be in some manner related to every other, any particular process uses very few of these. Thus the matrix describing steel technology has a small percentage of non-zeros. If spatial or temporal distinctions are introduced into the model the percentage of non-zeros generally falls further. Thus if a one-time period, one-place model uses m equations and has q non-zero coefficients, an S place, T time period model usually has about STm equations and the proportion of non-zero coefficients is roughly

$$\frac{KTSq}{S^2 T^2 m^2} = \frac{Kq}{ST}$$

where K is usually about 2.

If S or T is doubled it may generally be expected that the time required per iteration by a linear programming procedure which does not take advantage of zeros will increase by a factor of 4. If an elimination form of inverse is used we may expect α, β and γ in $t = \alpha + \beta I + \gamma I^2$ to roughly double, and therefore time per iteration (approximately $\beta + \sqrt{\alpha\gamma}$) to also double.