A Probabilistic Consideration of Sparse Matrix

Tsutomu Shiino*

1. INTRODUCTION

In sparse matrix calculations, the reduction of computational steps where only a minimum level of storage capacity is required is most impotant. Even when an initial matrix contains many zero elements, failure to select proper pivoting order will increase non-zero elements during the computational process and thus one can not take advantage of sparsity. For this reason, various techniques for optimal ordering have been widely reported. However the effects of sparse matrix method have rarely been discussed quantiatively and usually they have been discussed only with special problems in these papers.

This paper presents a probabilistic consideration of ordinary sparse matrices, and studies quantitatively the change of the number of non-zero elements during the process of calculations and the effects of non-zero element distribution on calclation process. It also discusses the technique of predicting the maximum number of non-zero elements. It shows that sufficiently effective results can be obtained by initial ordering of sparse matrix alone without subsequent pivot selections. It also shows that remarkable improvement can be made by initial column interchange in addition to initial row interchange, and discusses the optimum distribution of non-zero elements of the initial matrix.

2. PROBABILISTIC TREATMENT OF SPARSE MATRIX

Let us assume to calculate a simultaneous equation having n x n coefficient sparse matrix X by the Gaussian elimination method. The initial matrix is represented by $X^{(0)}$ and the element of ith row and jth column is represented by $x_{ij}^{(0)}$. If the probability that $x_{ij}^{(0)}$ is non-zero is expressed as $P_{ij}^{(0)}$, the number of non-zero elements (the expected value) in $X^{(0)}$ is

This paper first appeared in Japanese in Joho－Shori (Journal of the Information Processing Society of Japan), Vol. 14, No. 9 (1973), pp. 677～683.

* Software Systems Division, Oki Electric Industry Co., Ltd., Tokyo, Japan

$$N^{(0)} = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}^{(0)}. \qquad (1)$$

Consider the increase of non-zero elements at an elimination step with the first row as a pivot row. It is assumed here that $x_{11}^{(0)}$ is not zero. The number of non-zero elements in the ith row increases only when x_{i1} is non-zero and j satisfies the conditions of $x_{1j} \neq 0$ and $x_{ij} = 0$ exists. Therefore, the increase of non-zero elements in the ith row can be expressed by the following equation.

$$P_{i1}^{(0)} \left\{ \sum_{j=2}^{n} P_{1j}^{(0)} (1 - P_{ij}^{(0)}) \right\} \qquad (2)$$

The number of increased non-zero elements by the first elimination is

$$C^{(1)} = \sum_{j=2}^{n} P_{i1}^{(0)} \left\{ \sum_{j=2}^{n} P_{1j}^{(0)} (1 - P_{ij}^{(0)}) - 1 \right\} \qquad (3)$$

The sub-matrix obtained by eliminating the first row and the first column of the post-elimination matrix $X^{(1)}$ represented by $X^{*(1)}$. The probability that the element $x_{ij}^{(1)}$ (i, j = 2, ..., n) of $X^{*(1)}$ is non-zero can be given by

$$P_{ij}^{(1)} = P_{ij}^{(0)} + P_{i1}^{(0)} \times P_{1j}^{(0)} \times (1 - P_{ij}^{(0)}). \qquad (4)$$

Generally, the number of non-zero elements which is increased by the kth elimination is given by

$$C^{(k)} = \sum_{i=k+1}^{n} P_{ik}^{(k-1)} \times \left\{ \sum_{j=k+1}^{n} P_{kj}^{(k-1)} (1 - P_{ij}^{(k-1)}) - 1 \right\} \qquad (5)$$

The probability that the element $x_{ij}^{(k)}$ (i, j = k + 1, k + 2, ..., n) of the submatrix $X^{*(k)}$ is non-zero is obtained by the following equation.

$$P_{ij}^{(k)} = P_{ij}^{(k-1)} + P_{ik}^{(k-1)} \times P_{kj}^{(k-1)} \times (1 - P_{ij}^{(k-1)}) \qquad (6)$$

The number of non-zero elements in the row $X^{(k)}$ can be expressed as

$$N^{(k)} = N^{(k-1)} + C^{(k)} \qquad (7)$$

where k = 1, 2, ..., n.

3. MATRIX HAVING RANDOM DISTRIBUTION OF NON-ZERO ELEMENTS AT FIXED PROBABILITY

Let us assume that the randomly distributed non-zero elements in a matrix has fixed probability. Since diagonal elements must be non-zero, it is assumed that $P_{ii}^{(0)} = 1$ (i = 1, 2, ..., n).

The number of non-zero elements that are increased by the kth elimination is

$$C^{(k)} = (n - k) P^{(k-1)} \times \left\{ (n-k-1) P^{(k-1)} \times (1-P^{(k-1)}) - 1 \right\} \qquad (8)$$

Note that $P^{(k-1)}$ denotes the non-zero probability of the non-diagonal elements of the sub-matrix $X^{*(k)}$ obtained by the (k-1)th elimination. The non-zero probability $P^{(k)}$ of the elements other than the diagonal elements in $X^{*(k)}$ is given by

$$P^{(k)} = P^{(k-1)} + \left\{P^{(k-1)}\right\}^2 \times \left\{1 - P^{(k-1)}\right\}, \tag{9}$$

and the total number of non-zero elements is

$$N^{(k)} = N^{(k-1)} + C^{(k)} \tag{10}$$

Fig. 1 and Fig. 2 show the values of $N^{(k)}$ and $P^{(k)}$ obtained by equations (8), (9) and (10) for n = 101 and n = 1001. In these figures, m is the number of non-zero elements in each row, excluding the diagonal element. The figures show that the number of non-zero elements reaches a peak near $P^{(k)} = 1$, and the peak value $Nmax^{(k)}$ can be observed:

$$Nmax^{(k)} \underset{\sim}{} (n - k)^2 \tag{11}$$

All the subsequent sub-matrices $X^{*(k)}$ will only have non-zero elements. If $P^{(0)}$ is given, the value of k corresponding to $P^{(k)} = 1$ can be obtained from equation (9). The results are shown in Fig. 3.

4. MATRIX HAVING UNIFORM ROWWISE AND NON-UNIFORM COLUMNWISE DISTRIBUTION OF NON-ZERO ELEMENTS

In such a case, the number of non-zero elements which is increased at elimination steps can be decreased sharply by selecting pivotal rows. The number of non-zero elements in an entire matrix after the kth step of Gaussian elimination is

$$N^{(k)} = N^{(k-1)} + \sum_{i=k+1}^{n} P_i^{(k-1)} \times \left\{(n-k-1)P_k^{(k-1)}\left(1 - P_i^{(k-1)}\right) - 1\right\}. \tag{12}$$

The probability that the elements, with the exception of the diagonal element, in the ith row of the sub-matrix $X^{*(k)}$ is given by

$$P_i^{(k)} = P_i^{(k-1)} + P_k^{(k-1)} \times P_i^{(k-1)} \times \left\{1 - P_i^{(k-1)}\right\} \tag{13}$$

where i = k + 1, k + 2, ..., n, and $P_i^{(k-1)}$ is the non-zero probability of the elements in the ith row of the sub-matrix $X^{*(k-1)}$. Fig. 4 shows $N^{(k)}$ of the matrices which have n = 1001 and $N^{(0)} = 5005$ and three types of columnwise distribution.

Line (a) in Fig. 4 represents a matrix in which non-zero elements are distributed at a fixed probability $P^{(0)} = 4/1000$ in the entire matrix (except the diagonal elements.)

Line (b) represents a matrix with the following distribution.

$Pi1^{(0)} = 1/1000 : i_1 = 1 - 143$ $P_{i5}^{(0)} = 5/1000 : i_5 = 573 - 715$

$P_{i2}^{(0)} = 2/1000 : i_2 = 144 - 286$ $P_{i6}^{(0)} = 6/1000 : i_6 = 716 - 858$

$P_{i3}^{(0)} = 3/1000 : i_3 = 287 - 429$ $P_{i7}^{(0)} = 7/1000 : i_7 = 859 - 1001$

$P_{i4}^{(0)} = 4/1000 : i_4 = 430 - 572$

Line (c) represents the completely reverse distribution ($i_1 - i_7$ of (b)).

In the case of ordinary matrices whose non-zero element distribution is not uniform, remarkable improvements can be made by interchanging rows in the order of increasing number of non-zero elements (from small to large as in (b)).

5. MATRIX HAVING UNIFORM COLUMNWISE AND NON-UNIFORM ROWWISE DISTRIBUTIONS OF NON-ZERO ELEMENTS

The number of non-zero elements in the entire matrix after the kth step of elimination is given by

$$N^{(k)} = N^{(k-1)} + (n-k-1)P_k^{(k-1)} \times \left\{ \sum_{j=k+1}^{n} P_j^{(k-1)}(1 - P_j^{(k-1)}) - 1 \right\} \tag{14}$$

where $P_j^{(k)}$ is the non-zero probability of the jth column of $X^{*(k)}$. This equation can be transformed as below.

$$N^{(k)} = N^{(k-1)} + (n-k-1) \times \sum_{j=k+1}^{n} \left\{ P_k^{(k-1)} P_j^{(k-1)}(1 - P_j^{(k-1)}) - 1 \right\} \tag{15}$$

It becomes clear that it is completely identical with the equation (12). The non-zero probability, $P_j^{(k)}$, of the elements of the column j is

$$P_j^{(k)} = P_j^{(k-1)} + P_k^{(k-1)} P_j^{(k-1)} \left\{ 1 - P_j^{(k-1)} \right\} \tag{16}$$

It is seen that this is identical to eq. (13). In other words, the distributions of rowwise non-zero elements and of columnwise non-zero elements have an identical effect in the overall calculation of sparse matrix.

6. OPTIMAL NON-ZERO ELEMENT DISTRIBUTION OF INITIAL MATRIX AND ESTIMATION OF MAXIMUM STORAGE CAPACITY

The distribution of non-zero elements in the initial matrices has a large effect on the memory capacity and the number of calculation steps of a sparse matrix. Considering a general purpose program which does not make use of the correlation between the elements, a sufficient condition for an optimal matrix calculation is that the non-zero elements of the initial matrix are ordered in an increasing order of non-zero probability for rows and columns. No pivot selection is needed for sufficiently large matrices if there is no correlation between elements. The worst possible distribution of non-zero elements in a matrix is a uniform distribution. In this case, however, the maximum possible number of the increase non-zero elements can be obtained by equation (11) as mentioned in chapter 2.

With the above considerations the maximum number of non-zero elements of a matrix can be estimated by ordering the rows and the columns of the initial matrix and comparing the density of non-zero elements in each column and row with $P^{(0)}$ of Fig. 1 or Fig. 2.

7. EFFECTIVENESS OF SPARSE MATRIX METHOD

The choise between the sparse matrix method or conventional matrix reduction method always poses a problem in calculating some matrix. It seems quite useful to clarify the scope and range in effectiveness of the sparse matrix method from the point of view of both storage capacity and computational time.

If the locations of non-zero elements of a sparse matrix are stored by the row number and the column number, the storage capacity for non-zero elements locations is given by the number of rows and the number of non-zero elements. The total storage capacity for the sparse matrix is n + 2N, where N is the number of non-zero elements. From the point of view of the storage capacity, the sparse matrix method is useful if

$$n^2 > n + 2N_{max}$$

where N_{max} is the maximum number of non-zero elements.

Let us now consider the processing time required. Let α be the ratio of relocation processing time and elimination processing time. The ratio of the processing time of sparse matrix method vs. that of non-sparse matrix method can be given by $A(1 + \alpha)/A_o$. It is assumed here that the processing time is proportional to the total number of non-zero elements. Here A is the area enclosed by the curve of sparse matrix method of figure 2 and the horizontal axis. A_0 is the area enclosed by the curve of non-sparse matrix method (100% non-zero probability) of figure 2 and the horizontal axis. If $\alpha = 1$, the sparse matrix method is useful in the range of $2A/A_o < 1$.

8. CONCLUSION

The sparse matrix method has been used extensively for computation of matrices containing many non-zero elements. This paper has quantitatively studied possible changes of numbers of non-zero elements during calculations through probabilistic consideration. It is clearly established that an ordering of initial matrix by row and column interchanges is quite effective for the general-purpose calculation program for large-scale sparse matrices. Methods of prediction of the maximum fill-in of non-zero elements, the total number of non-zero elements generated, and the volume of processing and effectiveness of the sparse matrix method are also discussed.

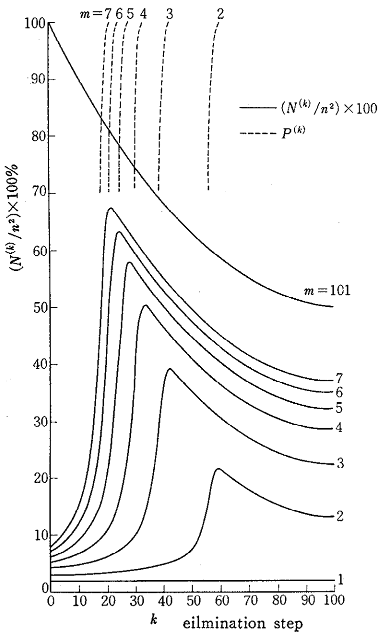

Fig. 1 Number of non-zero elements vs. elimination step. (n = 101)

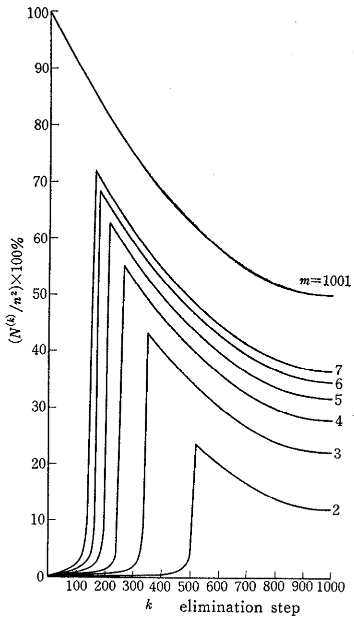

Fig. 2 Number of non-zero elements vs. elimination step. (n = 1001)

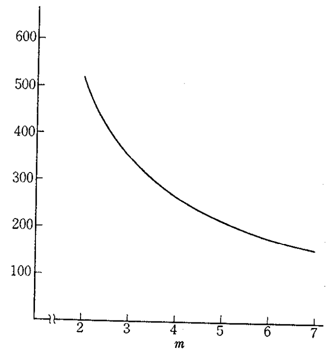

Fig. 3 m (number of non-zero elements of each row in the initial matrix) vs. the elimination step at which $P^{(k)}$ becomes 1. (n = 1001)

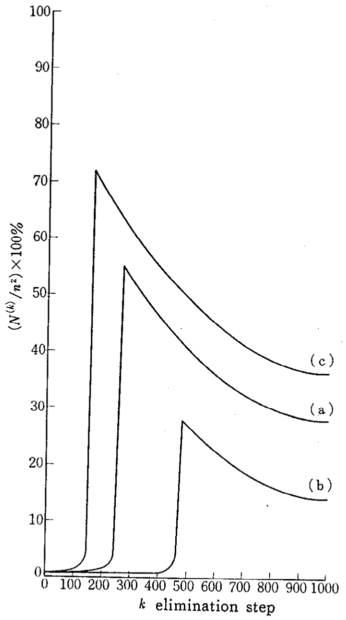

Fig. 4 Effect of the distribution of non-zero elements in the initial matrix on elimination process.