The Multi-vector Compressed Storage of Lower Half Banded Matrix

Zhiguo Ren *[1, a], Qiuju Yue and Han Zhang

[1] School of Electronics and Information Engineering,Lanzhou City University. No.11 Jiefang Road, Anning District,Lanzhou City,Gansu Province,730070, P.R.China.

[a] ren_zhiguo@qq.com,

Keywords: Multi-vector; Compressed Storage; Lower Half Banded Matrix

ABSTRACT

Matrix is a mathematical object, commonly used in scientific computing and engineering calculation. In the data structure we are not interested in data itself, but how to store the elements in the matrix, and make the various operations can run effectively. The main purpose of the compressed storage is to make more of the same nonzero elements share the same storage unit according to the distribution of matrix element, while the zero elements don't allocate storage space. In this paper, we studied the multi-vector compressed storage problems of lower half banded matrix, and obtained the row and column priority compressed storage address mapping function for the first time.

The compressed storage structures have high compression ratio. The conclusions hope to provide the basic theory of data compression storage for the computer research workers.

1.DEFINITION OF HALF BANDED MATRIX

Matrix is a mathematical object, commonly used in scientific computing and engineering calculation. We are not interested in data type or value in the data structure, but how to store the elements in the matrix. When programming in a high-level language, often use a two-dimensional array to store the elements in the matrix. If adopt this method of storage, we can random access each data element, thus can easily realize operations of the matrix. But, when there are a large number of zero elements in the matrix and have regular distribution, if still use a two-dimensional array to store the matrix, a particular element will consume large amounts of storage unit. For high order matrix, the storage method is not only waste storage unit, but also takes a lot of time for invalid computation, it is obviously not desirable. In order to save the storage space, we need to compress storage for such matrix.

The main purpose of the compressed storage is to make more of the same nonzero elements share the same storage unit according to the distribution of matrix element, while the zero elements don't allocate storage space. In this paper, we studied the compressed storage problems of lower banded matrix, and obtained the corresponding storage address mapping functions.

The lower half banded matrix is evolved from the banded matrix. The elements come from the banded matrix diagonal line and lower part only. The elements along the diagonal direction of the matrix regular distribution in the one side of the diagonal.

In a matrix A of order n, if there exists a maximum positive number m, $a_{i,j}$ is a matrix element, when it meets $0 \le i-j \le m-1$, $a_{i,j}$ is an ordinary integer, the rest elements are 0 or a constant integer. The A is called lower half banded matrix. The integer m is called bandwidth. If the order is n, the bandwidth is m, the data elements in the half band matrix A compressed storage into a multi-vector B. By the simple calculation, the size of storage space for B vector is $(m \times n)$.

2. ROW PRIORITY COMPRESSED STORAGE

When n=5,m=2,the corresponding lower half banded matrix A can be compressed into a multi-vector storage space B. As shown in figure 1.

$$A=\begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} & 0 \\ 0 & 0 & 0 & a_{54} & a_{55} \end{bmatrix} \Rightarrow B=\begin{bmatrix} 0 & a_{11} \\ a_{21} & a_{22} \\ a_{32} & a_{33} \\ a_{43} & a_{44} \\ a_{54} & a_{55} \end{bmatrix}$$

Fig.1 The results of compression storage at n=5 and m=2

When n=5,m=3, the corresponding lower half banded matrix A can be compressed into a two-dimensional array B by compression method 1. As shown in figure2.

$$A=\begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ 0 & a_{42} & a_{43} & a_{44} & 0 \\ 0 & 0 & a_{53} & a_{54} & a_{55} \end{bmatrix} \Rightarrow B=\begin{bmatrix} 0 & 0 & a_{11} \\ 0 & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} \\ a_{42} & a_{43} & a_{44} \\ a_{53} & a_{54} & a_{55} \end{bmatrix}$$

Fig.2 the results of compression storage at n=5 and m=3

The image can be seen as a conversion: The sub arrays of the lower half banded matrix A are carried out to the left, so as to achieve the compression for storage the lower half banded matrix A.

Obviously, the elements at the diagonal line are carried to the right of the two-dimensional array B. The elements at the lowest diagonal line are carried to the left of array B. The rest of the elements are stored in the array B by this method. The row index of element $a_{i,j}$ does not change, the column index is changed to (j-i+m).

Then you can get the mapping relationship of $a_{i,j}$ to $b_{i'j'}$ (formula 1):

$$\begin{cases} i' = i \\ j' = j - i + m \end{cases} \quad (1)$$

3. COLUMN PRIORITY COMPRESSED STORAGE

When n=5,m=2, the corresponding lower half banded matrix A can be compressed into a two-dimensional array B by compression method 2. As shown in figure3.

$$A=\begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} & 0 \\ 0 & 0 & 0 & a_{54} & a_{55} \end{bmatrix} \Rightarrow B=\begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{44} & a_{55} \\ a_{21} & a_{32} & a_{43} & a_{54} & 0 \end{bmatrix}$$

Fig.3 the results of compression storage at n=5 and m=2

When n=5,m=3, the corresponding lower half banded matrix A can be compressed into a two-dimensional array B by compression method 2. As shown in figure4.

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ 0 & a_{42} & a_{43} & a_{44} & 0 \\ 0 & 0 & a_{53} & a_{54} & a_{55} \end{bmatrix} \Rightarrow B = \begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{44} & a_{55} \\ a_{21} & a_{32} & a_{43} & a_{54} & 0 \\ a_{31} & a_{42} & a_{53} & 0 & 0 \end{bmatrix}$$

Fig.4 The results of compression storage at n=5 and m=3

The above image can be seen as a conversion: The sub array of the lower half banded matrix A are carried out to the upper, so as to achieve the compression for storage the lower half banded matrix A.

Obviously, the elements at the diagonal line are carried to the upper of the two-dimensional array B. The elements at the lowest diagonal line are carried to the lowest of array B. The rest of the elements are stored in the array B by this method. The column index of element $a_{i,j}$ does not change, the raw index is changed to(i-j+1).

Then you can get the mapping relationship of $a_{i,j}$ to $d_{i'j'}$ (formula 2):

$$\begin{cases} i' = i \\ j' = i - j + 1 \end{cases} \quad (2)$$

4. CONCLUSION

In this paper, we studied the multi-vector compressed storage problems of lower banded matrix, and obtained the row priority compressed storage address mapping function of the multi-vector compressed storage(formula 1), also obtained the column priority compressed storage address mapping function of the multi-vector compressed storage(formula 2). In an n order matrix with m band width, the data elements can be compressed into a multi-vector array B, its compression ratio can be obtained. $CR=(1-(m*n)/n^2)*100\%$. In a 100-order square matrix, the band width is 10. Its compression ratio is: $CR=(1-(10*100)/100^2)*100\%=90\%$.

It can be seen that these two kinds of compressed storage have a high compression ratio. These conclusions hope to provide the theory basis of data compression storage for the data processing and scientific computing algorithm design.

Zhiguo Ren. School of Electronics and Information Engineering,Lanzhou City University. No.11 Jiefang Road, Anning District,Lanzhou City,Gansu Province,730070, P.R.China.

Supported by the Education and Teaching Research Foundation(2013-JY-25) and Doc's Foundation of Lanzhou City University(LZCU-BS2013-09;LZCU-BS2013-12);

Supported by Association of Fundamental Computing Education in Chinese Universities (AFCECU:NO.2016084).

5 REFERENCES

1. Thomas H.Cormen,Charles E.Leiserson,Ronald L.Rivest,Clifford Stein. Introduction to Algorithm, the third edition. The MIT Press,2009.
2. Alfred V.Aho,John E.Hopcroft,and Jeffrey D.Ullman.Data structures and Algorithms. Addison-Wesley, 1983.
3. Donald E.Knuth.Fundamental Algorithms,volume 1 of The Art of Computer Programming. Addison-Wesley, 1968.Third edition,1997.
4. Donald E.Knuth.Seminumerical Algorithms,volume 2 of The Art of Computer Programming. Addison- Wesley,1969.Third edition,1997.

5. Don Coppersmith and Shmuel Winograd.Matrix Multiplication via arithmetic progression. Journal of Symbolic Computation,9(3):251-280,1990.
6. T.C.Hu and M.T.Shing.Computation of Matrix chian products.Part 1,SIAM Journal on Computing, 11(2):362-373,1982.
7. T.C.Hu and M.T.Shing.Computation of Matrix chian products.Part 2,SIAM Journal on Computing, 13(2):228-251,1984.
8. Mark Allen Weiss.Data Structures and Algorithm analysis in Java. Addison-Wesley,third edition,2007.
9. Zhiguo Ren, Data Structure(C Language Description), Science Press, Peking China, 2016.06.
10. Ludeña-Choez, J., & Gallardo-Antolín, A. (2016). Acoustic event classification using spectral band selection and non-negative matrix factorization-based features. Expert Systems with Applications,46(C), 77-86.
11. Osipov, & Andrey. (2016). A study of resolvent set for a class of bandoperators with matrix elements. Concrete Operators, 3(1), 85-93.