43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Load balanced parallel block QR decomposition

S. Shahmorad
University of Tabriz

M. Famil Barraghie
University of Tabriz

Abstract

This paper introduces a new parallel QR decomposition using Householder transformation and Givens rotation. We try to parallelize the first stage of Boleng and Raghavan's algorithm [1] using parallel block Householder transformation. Load balancing technique be used here considers total computational works. The new hybrid algorithm has two stages, IT (The Internal Transformations Stage) and BR (The Balanced Rotations Stage). In the first stage, each processor uses block Householder transformation. The second stage annihilates remaining elements using Givens rotations. Since we are desinging with shared memory, parallelism does not defray extra communication costs.

Keywords: QR decomposition, Load balanced, Householder transformation, Givens rotation

Mathematics Subject Classification: 53A15

1 Introduction

Definition 1.1. *Any real $m \times n$ matrix A can be decomposed as*

$$A = QR$$

where Q is an $m \times m$ orthogonal matrix and R is an $m \times n$ upper triangular matrix. A is assumed to be full rank: rank(A) = rank (R)=n. Householder transformation and Givens rotation can be used to compute the decomposition.

Definition 1.2. *A Householder transformation is a transformation that takes a vector and reflects it about some plane or hyperplane. We can use this operation to calculate the QR decomposition of an $m \times n$ matrix A. A Householder matrix is presented by*

$$H = I - \beta vv^T,$$

where H is an $m \times m$ matrix and $\beta = 2/v^T v$ (v is a column vector). H is orthogonal and symmetric. The Houlseholder transformation introduces large number of zeroes in one matrix operation but can not be carried out in parallel in a straight forward way.

Definition 1.3. *The QR decomposition can also be computed with a series of Givens rotations. Each rotation zeroes an element in the subdiagonal of the matrix, forming the R matrix. An $m \times m$*

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

matrix J(i,j,c,s) of the form

$$J(i,j,c,s) = \begin{bmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & c & \dots & s & & \\ & & \vdots & \ddots & \vdots & & \\ & & -s & \dots & c & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}$$

is called Givens matrix where $c^2 + s^2 = 1$ *and*

$$c = a_{jk}/\sqrt{a_{jk}^2 + a_{ik}^2}, \qquad s = a_{ik}/\sqrt{a_{jk}^2 + a_{ik}^2}$$

The number of needed Givens matrices is

$$r = n(n-1)/2$$

and the number of householder matrices is n when $m > n$ *, or n-1 when* $m = n$*.*

1.1 The Internal Transformations Stage (IT)

During this stage, the rows of the matrix are divided amoung the processors with each processor getting a block of size $(m/p \times n)$ and performing block Householder transformations that we will introduce in below.

1.1.1 Extention Of the Householder Transformation

Let us consider a full rank matrix V and introduce the matrix extension of Householder transformation

$$H(V) = I - 2V(V^TV)^{-1}V^T.$$

Theorem 1.4. *For any full column rank* $(m/p \times r)$ *matrix A,*

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

where A_1 *is an* $r/p \times r$ *nonsingular marix, if we choose*

$$V_A = \begin{bmatrix} A_1 + X \\ A_2 \end{bmatrix}$$

where X is given by

$$X = P^T \sqrt{D} P A_1$$

with $\sqrt{D} = diag_{i=1}{}^r \sqrt{d_i}$ *where the nonnegative scalar* d_i *and the orthogonal matrix P are defined by*

$$I + (A_2A_1^{-1})^T(A_2A_1^{-1}) = P^T diag_{i=1}{}^r d_i P,$$

then

$$H(V_A)A = \begin{bmatrix} -X \\ 0_{(m-r)\times r} \end{bmatrix}$$

where I is an $r \times r$ *identity matrix and* $0_{(m\times r)\times r}$ *is an* $(m \times r) \times r$ *zero matrix.*

Proof. See [2]. □

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

$$\begin{bmatrix} X & X & X & X & X & X & X & X & X & X & X & X \\ 0 & X & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & 0 & X & X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X & X & X & X & X \\ 0 & X & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & 0 & X & X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X & X & X & X & X \\ 0 & X & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & 0 & X & X & X & X & X & X & X & X \end{bmatrix}$$

The matrix A after completion of the IT stage. Example using m=15, n=12 and p=3.

1.1.2 Application To The Block QR Decomposition

Let us consider a block matrix

$$M = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix}$$

where A_1 is an $r \times r$ nonsingular matrix. Then by choosing $H_1 = H(V_A)$, we obtain

$$H_1 M = \begin{bmatrix} -X & B_1^* \\ 0_{(m-r)\times r} & B_2^* \end{bmatrix}$$

which is an upper block triangular matrix. The Householder transformation matrix can be

$$H_2 = \begin{bmatrix} H_2^1 & 0_{m\times(n-r)} \\ 0_{(m-r)\times r} & H_2^1 \end{bmatrix},$$

where H_2^1 and H_2^2 are the Householder transformation matrices which act, respectly, on X and B_2^*. The application of this result leads to parallelize the QR decomposition of the matrix M. It authorizes to calculate H_2^1 and H_2^2 in two completely independent processors.

1.2 The Balanced Rotations Stage (BR)

In this stage, remaining elements are annihilated by using Givens rotations. This consist of $m/p-1$ steps. During each step, we should consider the following sequences:

- The rows of block 1 are assigned as pivot rows to the processors.
- Processors vanish all the elements in the remaining blocks that are of the same length as their assigned pivot rows. Each step annihilates the first diagonal in block 2 to p.

The next step is to perform both the IT and BR stages on the submatrix as shown in figure 2 by Y entires. The above two stages are repeated to annihilate the elements in blocks 2 to p until the matrix is fully decomposed.

$$
\begin{bmatrix}
X & X & X & X & X & X & X & X & X & X & X & X \\
0 & X & X & X & X & X & X & X & X & X & X & X \\
0 & 0 & X & X & X & X & X & X & X & X & X & X \\
0 & 0 & 0 & X & X & X & X & X & X & X & X & X \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y \\
0 & 0 & 0 & 0 & Y & Y & Y & Y & Y & Y & Y & Y
\end{bmatrix}
$$

The matrix A after completion of the BR stage. Example using m=15, n=12 and p=3.

1.3 Details of Load Balancing Technique

Load balancing refers to the practice of distributing works among tasks so that all tasks are kept busy all of the time. It can be considered a minimization of task idle time. One of the ways for load balancing the BR stage is using the Karl Friedrich Gauss idea for obtaining sum of the integers 1 to 100. Applying this idea, the processors assignment is done in a cyclic way. Cyclic length is 2p.

2 Main Result

This paper introduced a new parallel QR decomposition using modest number of processors. Most of the parallel algorithms need at least $m/2$ processors. For the problems with large sizes, this number of processors are not available, but this algorithm could approach to optimal results by using a few procossers. We know that Householder transformation can not be parallelized, however we would parallelize each block of matrix utilizing block Householder transformation. Load balancing technique helped to balance the work of each processor. Therefore, these methods for decomposing can reduce computational time and numerical calculation.

References

[1] J. Boleng, M. Misra, Load balanced parallel QR decomposition on shared memory multiprocessors, Parallel computing 27 (2001) 1321-1345 .

[2] F. Rotella, I. Zambettakis, Block householder transformation for parallel QR factorzation, Applied mathematics letters 12 (2011) 29-34.

Email:shahmorad@tabrizu.ac.ir

Email:m-baraghi88@tabrizu.ac.ir