Parallel Implementation of the Large Sparse and Symmetric Eigenproblem

E. M. Garzón and I. García

The aim of this work is to solve the eigenproblem for large symmetric and sparse matrices through a divide–and–conquer strategy put forward by Cuppen [1, 2]. This method has been implemented on a Cray T3E multiprocessor system, using PVM parallel interface and a detailed analysis of the obtained performance has been carried out.

The solution of the eigenproblem has been decomposed in the next phases:

- *Obtaining a structured matrix from the input matrix* (A). In this phase, both a tridiagonal matrix (T) and an orthonormal matrix Q are obtained from A, where T is similar to A. These matrices hold the relation $A = QTQ^T$. The Lanczos Method with Complete Reorhtogonalization has been implemented to be carried out at this stage [3, 4, 5].
- *Solving the eigenproblem of* T. The goal of this stage is to compute the diagonal D and orthonormal M matrices. These matrices hold the relation $T = MDM^T$. The divide-and-conquer method (dc) has been implemented to solve this problem. This method is decomposed in the following stages:
  1. Decomposition in Sp subproblems (with $Sp = 2p$), through the Rank-One Modification of T.
  2. Solving eigenproblem of each subproblem through the QR algorithm.
  3. Reconstructing. The solution of the eigenproblem for each link of subproblems is computed through the solution of each. Once it has been determined the solution of all the links of the same dimension, other stage of the reconstruction begins in which pairs of previous stage subproblems are linked. As the process develops, the number of subproblems is halved, until just one link of is solved. In this way, the graph of tasks connected with the development of this stage can be represented through a binary tree.
- *Computing the eigenvectors of* A. If G denotes the orthonormal matrix whose columns are the eigenvectors of A, then $A = GDG^T$. On the other hand, the output matrices generated in the previous stages hold the relation $A = QTQ^T = QMDM^TQ^T$, so $G = QM$. Thus, this stage is carried out with a matrices product.

The eigenproblem of a sparse matrix is decomposed in several procedures whose performance is sequential, therefore it has been established a parallel implementation for each procedure in an independent way. Nevertheless, in order to link all the procedures in parallel, the output data distribution of a procedure establishes the parallel implementation of the procedure that receives these data as input.

The parallel implementation in phase 1 (Lanczos) is based on a decomposition in domains of input matrix. The input data are irregular, since A is sparse. Thus, a stage of preprocessing has been designed, namely Pivoting-Block, that is quick and guarantees that the input data partitions are balanced and also the operations associated to them [6, 7].

Then the parallel implementation of the Cuppen's method is based on decomposition in domains in both, the input and output data. The binary tree of tasks is distributed so that the same number of branches of the tree is allocated to each PE. If P refers to the number of PEs of the multiprocessor system and $P < Sp$ then initially each PE starts the reconstruction process independently, until $Sp = P$. Then groups of PEs that collaborate in the solution of a pair of subproblems are defined. As the reconstruction stages develop the subproblems are of a greater dimension and the number of PEs that collaborate is larger too.

The product $QM = G$ is carried out starting with a partition of Q and M by rows among PEs. This partition causes a remarkable penalization in the communications that should be established. It will be proved that the phases with the largest computational cost are the first and the last. This implementation will be evaluated for matrices of dimension between 782 and 7168, obtaining superlineal speed-up in most cases. It will be checked how the superlineal speed-up is due to low efficiency management of memory hierarchy with few PEs. It wil be proved that the management improves substantially as the number of PEs is increased.

References

[1] Cuppen, J.J.M, "A Divide and Conquer Method for the Symmetric Tridiagonal Eigenproblem", Numer. Math, vol. 36, pag. 177-195, 1981.

[2] Dongarra, J.J. and Sorensen, D.C " Fully Parallel Algorithm for the Symmetric Eigenvalue Problem", SIAM J. Sci. Stat. Comput., vol 8, num. 2, pag. 139-154, 1987.

[3] Paige, C.C. "Practical Use of the Symmetric Lanczos Process with Re-Orthogonalization.", BIT, 10, pag. 183-195. 1970.

[4] Paige, C.C. "Accuracy and Effectiveness of the Lanczos Algorithm for the Symmetric Eigenproblem.", Linear Algebra and its Applications, 34, pag. 235-258, 1980.

[5] Parlett, B.N. and Nour-Omid, B. "Towards a Black Box Lanczos Program.", Computers Physics Communications. Vol 53, pag 169-179, 1989.

[6] Garzón, E.M. and García, I. "Parallel Implementation of the Lanczos Method for Sparse Matrices: Analysis of the Data Distributions." Proceedings ICS'96. 1996.

[7] Garzón, E.M. and García, I. "Evaluation of the Work Load Balance in Irregular Problems Using Value Based Data Distributions." Proceedings of Euro-PDS'97, pag 137-143, 1997.