A PARALLEL OPTIMAL ASSIGNMENT ALGORITHM BASED ON DIAGONAL SCALING

STÉPHANE GAUBERT, LAURA GRIGORI, AND MEISAM SHARIFY

An important problem in the solution of very large sparse linear systems of equations is to permute large elements in magnitude on the diagonal of a matrix. In the context of sparse LU factorization, the purpose of the large-diagonal permutation is to decrease the probability of encountering small pivots during factorization, and hence not pivot during the factorization. This problem can be formulated as finding a permutation σ of n elements maximizing the product

$$\prod_{1\leq i\leq n} a_{i\sigma(i)}$$

where a_{ij} denotes the absolute value of the (i,j) entry in the matrix of the $n\times n$ linear system.

This is equivalent to solving the classical optimal assignment problem with weights $\log a_{ij}$. The latter can be solved by efficient specific algorithms (Hungarian method, auction algorithm, ...) or by various network flow algorithms. However, the classical methods, which are inherently sequential, are not adapted to situations in which the system is very large.

We propose a new optimal assignment algorithm, based on completely different ideas, which can be slower than the classical methods, but which scales well and can be implemented in parallel.

We start with the following optimization problem, which consists in finding an $n\times n$ bistochastic matrix $X=(x_{ij})$ maximizing the relative entropy

$$S_p(X) := -\sum_{1\leq i,j\leq n} x_{ij}(\log(x_{ij}/a_{ij}^p)-1)\ ,$$

where $p>1$ is a parameter which we shall adapt during the course of the algorithm. If $a_{ij}=0$, it is understood that the latter entropy is $-\infty$ unless $x_{ij}=0$.

When $p=1$, the same entropy maximization problem arises in the celebrated "DAD" scaling problem [MS69], which consists in finding diagonal matrices D and D' with positive diagonal entries such that the scaled matrix DAD' is bistochastic. When the matrix $A:=(a_{ij})$ is fully indecomposable, such diagonal matrices are known to exist, and their diagonal entries are given by the exponentials of the optimal Lagrange multipliers of the dual of the previous entropy maximization problem.

An immediate strict concavity argument shows that as soon as the matrix A has full term rank (meaning that the bipartite graph which is associated to it admits a perfect matching), the previous entropy maximization problem admits a unique solution $x^{(p)}$.

We first show that, when the permutation σ which is the solution of the above optimal assignment problem, is unique, the solution $x^{(p)}$ converges to the matrix representing the permutation σ. This is partly inspired by max-plus or tropical algebra, in which a similar deformation plays a central role.

For a fixed p, the solution $x^{(p)}$ can be computed classically by iterative algorithms, the simplest of which is Sinkhorn balancing (see [SK67]). The latter consists in starting from the matrix $A_p := (a_{ij}^p)$, dividing every row of A_p by its sum, then dividing every column of the new matrix by its sum, and so on, until the matrix obtained in this way converges to a bistochastic matrix.

The idea of our algorithm is to perform Sinkhorn balancing type iterations, while at the same time increasing the value of p. If the matrix A is fully indecomposable, if it has a unique optimal assignment, and if the growth of p is moderate enough, then, the sequence of matrices produced by the algorithm converges to the matrix representing the optimal permutation. The interest of this method is that it can be performed in parallel, since it only requires computing successive rows and column sums.

Date: May 22, 2009, submitted to the SIAM Workshop on Combinatorial Scientific Computing (CSC09).

When there are several optimal permutations, the sequence of matrices produced by the algorithm converges to a point x^* in the relative interior of the face of the polytope of bistochastic matrices containing the optimal permutations. Then, we are reduced to the question of finding a perfect matching in the bipartite graph associated to x^*. We note that the idea of using matrix balancing to detect the existence of a perfect matching was used by Linial, Samorodnitsky and Wigderson [LSW00]. A key new idea here is the introduction of the entropy maximization problem with the deformation parameter p, which allows us to get the optimal matching.

Finding the optimal growth of p as a function of the number of steps k and of the matrix characteristics appears to be a difficult problem. We performed our experiments by chosing a “polynomial” type growth $p := p_k$, with $p_k = p_0 k^{\beta}$, and typically $\beta = 1.5$. Higher values of the exponent β lead to a faster convergence in some cases, but they lead to a loss of convergence in other cases. We warn the reader that the algorithm is not adapted to the case of full random matrices, which have (with an overwhelming probability) several “nearly optimal” permutations. In such circumstances, the convergence to the optimal permutation is very slow. However, in many practical cases, the ratio between the values of the best and second best permutation is large enough, and then the algorithm turns out to be efficient.

We implemented the algorithm as an auxiliary routine in SuperLU, and tested it on matrices from the Harwell Boeing Sparse Matrix Collection. Our numerical results confirm the interest of the algorithm as a preprocessing tool before running a sparse LU solver without pivoting.

References

[LSW00] Nathan Linial, Alex Samorodnitsky, and Avi Wigderson. A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents. *Combinatorica*, 20(4):545–568, 2000.

[MS69] M. V. Menon and Hans Schneider. The spectrum of a nonlinear operator associated with a matrix. *Linear Algebra and Appl.*, 2:321–334, 1969.

[SK67] Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific J. Math.*, 21:343–348, 1967.

INRIA Saclay – Île-de-France & Centre de Mathématiques appliquées, Ecole Polytechnique, 91128 Palaiseau, France

E-mail address: Stephane.Gaubert@inria.fr

INRIA Saclay – Île-de-France & Laboratoire de Recherche en Informatique, Université Paris-Sud 11, France

E-mail address: Laura.Grigori@inria.fr

INRIA Saclay – Île-de-France & Centre de Mathématiques appliquées, Ecole Polytechnique, 91128 Palaiseau, France

E-mail address: Meisam.Sharify@inria.fr