WAVELET DECOMPOSITION VIA THE STANDARD TABLEAU SIMPLEX METHOD OF LINEAR PROGRAMMING

Gabriel Yarmish, Brooklyn College (yarmish@sci.brooklyn.cuny.edu)

Abstract

Wavelet decomposition problems have been modeled as linear programs – but only as extremely dense problems. Both revised simplex and interior point methods have difficulty with dense linear programs. The question then is how to get around that issue. In our experiments the standard method outperforms a revised implementation for these problems. Moreover, the standard method can be easily and scalably distributed. Hence the standard simplex method should be useful in solving wavelet decomposition problems.

Keywords: Wavelet decomposition, linear programming; Standard simplex method; Dense matrices; Distributed computing

1. INTRODUCTION

To date there have been numerous models for wavelet decomposition each with advantages and disadvantages [Daubechies, 1988; Mallat and Zhong, 1993; Coifman and Wickerhauser, 1992]. One method introduced by Chen et al [1998] is to model wavelet decomposition as a linear program. One major issue is that the resulting linear programs are very dense thus making both the revised and interior point methods unsuitable for solving these linear programs in an efficient manner. Chen et al finessed this issue by restricting the wavelet decomposition problems to those with dictionaries having a special structure and by tailoring an implementation of an interior point method to take advantage of that special structure.

Our research approaches this issue by refocusing attention on the simplex method. Dantzig's simplex algorithm for linear programming has two major variants: the original tableau, or *standard method*, and the *revised method*. Today, virtually all serious implementations are based on the revised method because it takes advantage of the sparsity that is inherent in most linear programming applications. The revised method is also advantageous for problems with a high aspect ratio; that is, for problems with many more columns than rows.

However, the standard method has advantages as well. First, the standard method is effective for dense problems [Nash and Sofer, 1996 p. 115]. While dense problems are uncommon in general, they do occur frequently in some important applications within linear programming [Eckstein et al, 1995]. Included in these are wavelet decomposition [Chen et al, 1998], Image processing [Selesnick et al, 2004], and digital filter design [Hu & Rabiner, 1972; Steiglitz et al, 1992; Gislason et al, 1993]. All these problem groups are well suited to the standard method. Second, the standard method can be easily and effectively extended to parallel and coarse grained distributed algorithms. (There are no scalable distributed versions of the revised simplex method.) When the standard method is distributed, aspect ratio becomes less of an issue. We simply divide the extra columns among more processors. If done properly, parallelization of the standard method pays off even on small problems [Yarmish, 2001].

Although we focus on the simplex method, it should be noted that it is difficult to deal with dense problems even with interior point methods [Bertsimas & Tsitsiklis, 1997, p. 440, pp. 536-537; Chen et al, 1998, p. 57].

We have written a standard implementation of the simplex method (retroLP) and compared it to the commonly used revised method as implemented by the well-known *MINOS* optimization package [Murtagh, & Saunders, 1998]. The theoretical fact that the standard method is faster than the revised method for dense problems is well known [Hadley, 1962 p. 216; Nash and Sofer, 1996, p.115]. We compare the efficiencies of the standard and revised simplex methods in solving wavelet decomposition problems. We empirically show that although for sparse problems the revised method is superior, the standard method is actually better suited and should be used. for the wavelet decomposition problems.

Another motivation for applying the standard method to wavelet decomposition problems is parallelization. It is relatively straightforward to distribute a linear program amongst multiple processors when employing the original standard method whereas it is not straightforward using modern implementations of the revised method [Yarmish, 2001].

Section 2 provides a brief overview of the standard and revised simplex methods. In section 3 we provide a brief review of previous research. Section 4 describes our experimental configuration. Section 5 describes empirical results for dense Wavelet Decomposition applications. In particular we describe empirical tests comparing the two methods for varying problem densities. Section 6 gives a brief summary and conclusion.

2. THE REVISED AND STANDARD SIMPLEX METHODS

We consider linear programs in the general form:

$$\begin{aligned}&Max\ z = cx\\&b^l \le Ax \le b^u\\&l_j \le x_j \le u_j \quad for\ \ j = 1,\ldots,n\end{aligned}$$

Or with $y = Ax$ we have:

$$\underset{x}{Maximize}\ z = \sum_{j=1}^{n} c_j x_j$$
$$Subject\ to\ y_i = \sum_{j=1}^{n} a_{ij} x_j \quad (i = 1, 2, \ldots, m)$$
$$l_j \le x_j \le u_j \ \text{for}\ j = 1, \ldots, n; \quad b_i^l \le y_i \le b_i^u \ \text{for}\ i = 1, \ldots, m$$

$A = \{a_{ij}\}$ is a given m x n matrix, x is an n-vector of decision variables x_j, each with given lower bound l_j and upper bound u_j. The m-vectors b^l and b^u are given data that define constraints. The lower bound, l_j, may take on the value $-\infty$ and the upper bound, u_j, may take on the value $+\infty$. Similarly, some or all of the components of b^l may be $-\infty$, and some or all of b^u may be $+\infty$.

Table 1 summarizes the main qualitative differences between the standard and revised simplex method that affect wavelet decomposition linear programming problems.

Revised Simplex Method	Standard Simplex Method
Takes better advantage of sparsity in problems	Is more effective for dense problems
Is more efficient for problems with large aspect ratio (n/m)	Is more efficient for problems with low aspect ratio.
Is difficult to perform efficiently in parallel, especially, in loosely coupled systems.	Very easy to convert to a distributed version with a loosely coupled system.

Table 1: Comparison of Revised and Standard Forms of the Simplex Method

3. PREVIOUS RESEARCH

3.1 Wavelet Decomposition

Chen, Donoho and Saunders [1998] have modeled Wavelet Decomposition as Linear Programs. Their goal was to improve on previous methods such as the method of Frames, the method of Matching Pursuits and the method of Best Orthogonal Basis. These methods, in the opinion of the authors, produce dictionaries with sufficient computational speed but “lack qualities of sparsity preservation and of stable super-resolution.”

In order to compensate for this, the authors introduced a new decomposition method called “Atomic Decomposition by Basis Pursuit.” This method translates into large linear programs. For example, a typical wave signal of length 8192 results in an equivalent Linear Program of size 8192 by 212,992. Unfortunately the Linear Programs produced are not only extremely large but are dense. In order to solve these problems they could use neither the revised algorithm nor a straightforward implementation of an interior-point method. To quote the authors [Chen et al, 1998, p. 57] “However, the optimization problems we are interested in have a key difference from [other] successful large-scale applications…. The matrix A we deal with is not at all sparse; it is generally completely dense…”

In order to deal with this the authors implemented a specialized interior point method to derive a unique wavelet dictionary from an over complete dictionary. (In particular they chose an algorithm based on the primal-dual log barrier interior point algorithm. In most implementations, the most computationally intensive step of interior point methods is solving large linear systems of equations using Cholesky factorization. Heuristics are used to take advantage of data sparsity to achieve sparse Cholesky factorizations. See, for example, Bertsimas & Tsitsiklis, 1997, p. 440, pp. 536-537. Among other things they took advantage of fast implicit algorithms for representations in the dictionaries they considered. They used this to develop a substitute approach for efficiently solving the systems of equations and restricted the class of wavelet dictionaries used.

3.2 Standard Simplex Method for Dense Linear Programs

It is well known that there are two factors that determine the efficiency of the standard method vs the revised method. These are density and aspect ratio (column/row ratio) [Hadley, 1962; Nash and Sofer, 1996; Chvátal, 1983]. Low density and high aspect ratio favor the revised method whereas high density and low aspect ratio favor the standard method. Most problems are extremely sparse and have a moderate aspect ratio. They are therefore much more efficiently solved via the revised method. On the other hand, linear programs resulting from wavelet decomposition, although possessing a high aspect ratio, are also extremely dense, often approaching 100% density.

3.3 Scalable Parallel Algorithms for the Standard Simplex Method

Recently there has been much research on methods to parallelize the simplex method. The standard method has proven to be more amenable to distributed and parallel algorithms than the revised method. A number of parallel algorithms have been produced both for massively parallel machines and for distributed networks of workstations.

Thomadakis & Liu [1996] worked on the standard method utilizing the MP-1 and MP-2 MasPar. Eckstein et al [1995] showed in the context of the parallel connection machine CM-2 that the iteration time for the parallel revised method tended to be significantly higher than for the parallel tableau algorithm even when the revised method is implemented very carefully. Stunkel [1988] worked on a way to parallelize both the revised and standard methods so that both would obtain a similar advantage in the context of the parallel Intel iPSC hypercube. Yarmish [2001] describes a coarse grained distributed standard simplex method, dpLP, especially optimized for loosely coupled workstations.

Dense applications, such as wavelet decomposition, for which the standard method yields lower iteration times, have particular potential for increased efficiency through the use of these parallel algorithms. We therefore propose that it is both possible and advantageous to use the general purpose standard method to solve these wavelet decomposition problems without having to be limited to wavelet dictionaries with fast representations.

4. EXPERIMENTAL CONFIGURATION

4.1 retroLP

retroLP is an implementation of the standard simplex method that directly implements linear programs of the form shown in Equations (1) and (2) in Section 2. *retroLP* accepts any file in the standard MPS format. It is quite stable; for example, it has successfully solved all linear programming problems in the Netlib repository.

4.2 MINOS

We use MINOS 5.5 [Murtagh & Saunders, 1998] as a representative implementation of the revised method. We installed it to run in the same environment as retroLP. This allowed us to make reasonable comparisons between the standard and revised methods. The purpose of these comparisons is not so much to compare running times but to examine the relative behavior of these approaches as the parameters of interest, primarily density, are varied. We chose the parameters for both solvers to try to eliminate factors extraneous to the direct comparison of the revised and standard methods.

5. EXPERIMENTAL RESULTS

Below we report on three separate groups of test runs. Data for these experiments were procured from the Wavelet and Atomizer packages provided by Chen at al [http://www-stat.stanford.edu/~atomizer/]. It is therefore important to point out a few issues to bear in mind while reviewing the results of our experiments. First, the ability to test problems with specific chosen densities on specific chosen problem sizes is severely hampered. In particular the Atomizer package will only provide wavelet dictionaries that are a power or two in height and width. Furthermore the dictionaries do not directly allow for control of densities.

A second thing to keep in mind is the structure inherent in these wavelet dictionaries. As mentioned above the authors themselves noted that they are only working with wavelet dictionaries that have a structure that lend themselves to fast implicit algorithms to solve their systems of equations. We were automatically limited to dictionaries provided by the authors' atomizer package which only gave these specially structured dictionaries. Nevertheless, by using the standard method as we have, other dictionaries without fast implicit algorithms could be used as well, since neither retoLP nor MINOS take advantage of any special structure in the wavelet dictionaries.

The first group of problems tested is listed in Table 2. In it are listed seven Wavelet Decomposition problems with varying problem sizes. The denser problems are solved more efficiently by the standard simplex method. Note that the three 87.5% problems are comprised of the "Discrete Sine Transform" wavelet dictionary. The other problems are comprised of Wavelet Packet, Discrete Cosine and combinations of dictionaries made by using the MakeList command provided by the Atomizer package. The 512 x 2048 problem size implies an aspect ratio of 4.

Figures 1 and 2 compare the revised and tableau methods as density varies for 20 separate Cosine Packet(CP) wavelet problems. We were able to get varying densities from a CP wavelet by combining high density basis parts of the CP dictionary with low density basis parts of the CP dictionary to get a medium density basis within the CP dictionary.

Results for the second and third groups of problems are shown in Figures 1 and 2 respectively. In the second group, all problems were size 512 x 1024, which implies an aspect ratio of 2. The wavelet dictionaries used are "cosine packet" dictionaries. The breakeven point for this problem group is at approximately 60% density.

The third group of problems used a problem size of 512 x 2048. This has an aspect ratio of 4, double the aspect ratio of the last set of problems. As can be seen from Figure 2, the breakeven point of 75% is higher than the breakeven point of Figure 1 as expected.

Problem	M	N	Density	Minos time/iter	retroLP time/iter
1	1024	4096	0.20%	0.0002637	0.0950439
2	512	2048	0.39%	0.0003320	0.0464786
3	256	1536	33.85%	0.0055506	0.0163869
4	128	3246	34.37%	0.0015652	0.0034923
5	512	2048	87.52%	0.0532897	0.0452119
6	512	2048	87.52%	0.0553762	0.0450486
7	512	2048	87.52%	0.0527936	0.0451510

Table 2: Comparison of *retroLP* and *MINOS* for Wavelet Decomposition

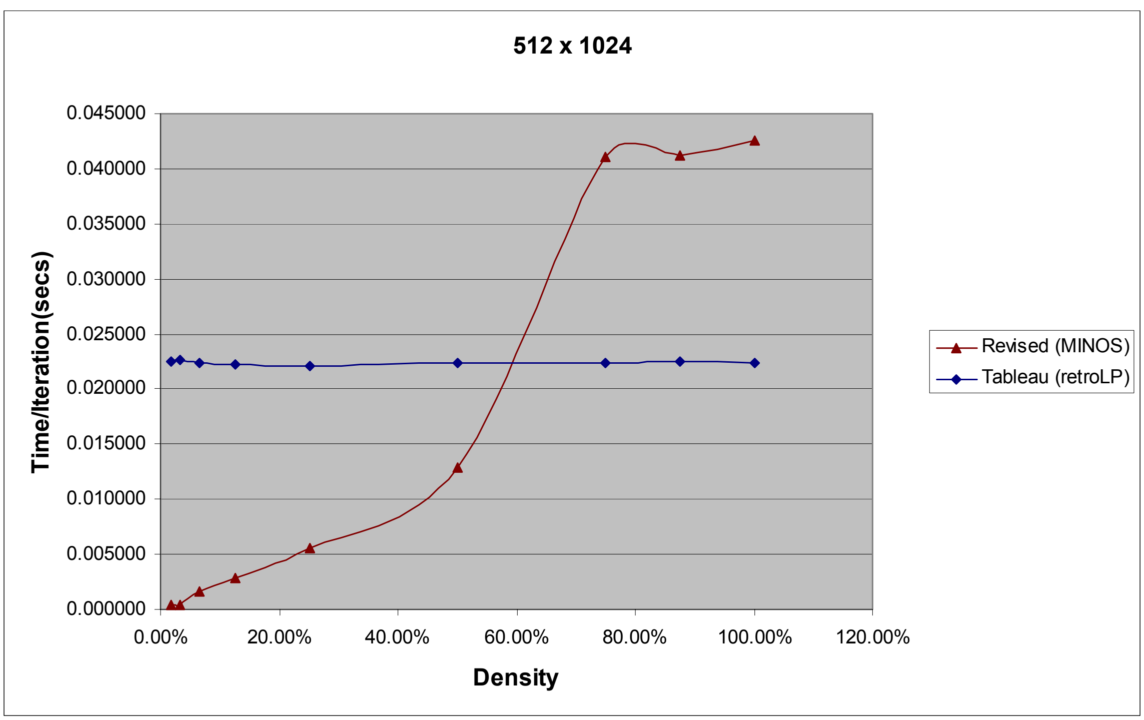

Figure 1: Density Breakeven point for 512 x 1024 wavelet decomposition

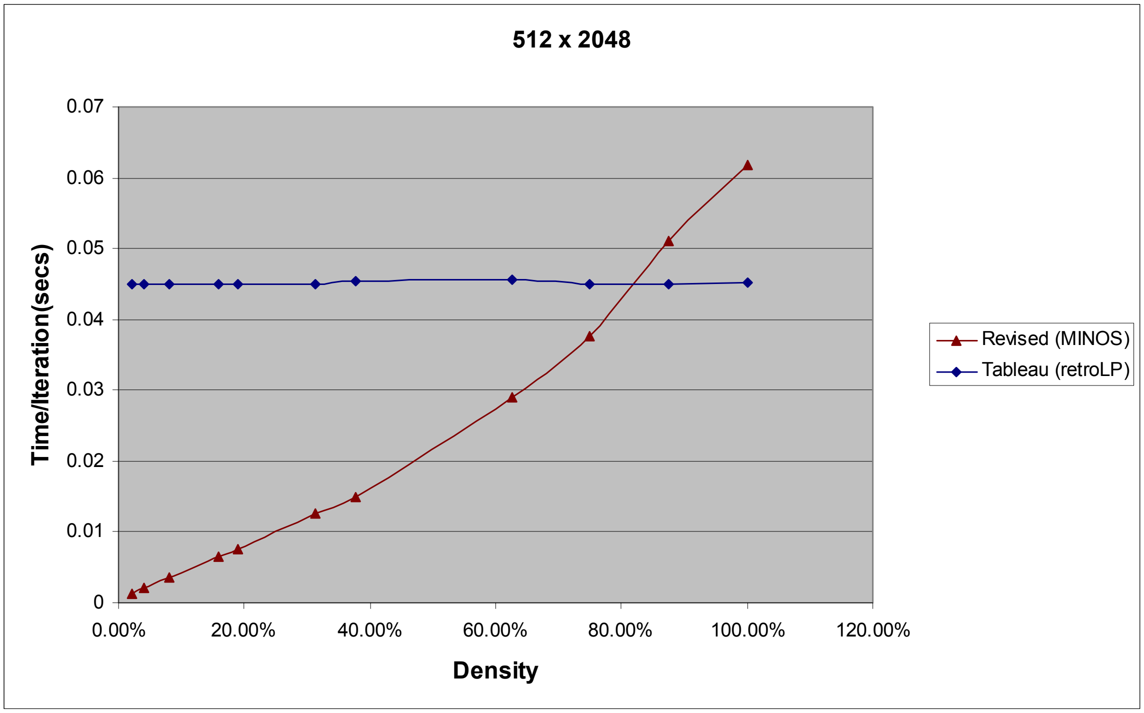

Figure 2: Density Breakeven point for 512 x 2048 wavelet decomposition

6. SUMMARY AND CONCLUSIONS

In this paper we discussed Wavelet Decomposition problems modeled as linear programs. We focused in particular on the model offered by Chen et al [1998]. We provided empirical data and experiments that compare the standard algorithm with the revised algorithm and that show density breakeven points. Our experiments comparing *MINOS* and *retroLP* indicate that for moderate values of density the standard method is competitive, and that Wavelet Decomposition can take advantage of the standard method as indicated by our results when using retroLP.

An implementation of the standard method makes possible a natural Single Program Multiple Data (SPMD) approach for a distributed simplex method. Partition the columns among a number of workstations. Each iteration, each workstation prices out its columns, and makes a "bid" to all the workstations. The winning bid defines a pivot column, then all the workstations pivot on their columns in parallel, and so on. This is important for such problems as Wavelet Decomposition that are suited to the standard method.

REFERENCES

Bertsimas, Dimitris, and John N. Tsitsiklis, Introduction to Linear Optimization, Athena, 1997.

Bixby, Robert E., "Implementing the Simplex Method: The Initial Basis," *ORSA J. on Computing*, Vol. 4, No. 3, pp. 267-284, Summer, 1992.

Chen S.S., D.L. Donoho, and M.A. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM J. on Scientific Computing*, 20, 1, pp. 33-61, 1998. http://www-stat.stanford.edu/~atomizer/

Chvátal, Vasek, Linear Programming, Freeman, 1983.

Coifman, R. and M.V. Wickerhauser, "Entropy-based Algorithms for Best-basis Selection," *IEEE Transactions on Information Theory*, v. 38, pp. 713-718, 1992

Daubechies, I, "Time-frequency Localization Operators: A Geometric Phase Space Approach," *IEEE Transactions on Information Theory*, v. 34, pp. 605-612, 1988

Eckstein, J., I. Boduroglu, L. Polymenakos, and D. Goldfarb, "Data-Parallel Implementations of Dense Simplex Methods on the Connection Machine CM-2," *ORSA Journal on Computing*, v. 7, n. 4, pp. 402-416, Fall 1995.

Gislason, Eyjolfur, et al, "Three Different Criteria for the Design of Two-Dimensional Zero Phase FIR Digital Filters," *IEEE Trans. on Signal Processing*, v. 41, n. 10, October, 1993.

Hadley, G., Linear Programming, Addison-Wesley, 1962

Hu, J.V., and L.R. Rabiner, "Design techniques for two-dimensional digital filters," *IEEE Trans. Audio Electroacoust.*, v. AU-20, pp. 249-257, Oct. 1972.

Mallet, S. and Z. Zhong, "Matching Pursuit in a Time-frequency Dictionary," *IEEE Transactions on Signal Processing,* v. 41, pp. 3397-3415, 1993

Murtagh, Bruce A., and Michael Saunders, "MINOS 5.5 Users Guide," *Technical Report* SOL 83-20R, Revised July, 1998.

Nash, Stephen and Ariella Sofer, Linear and Nonlinear Programming, McGraw-Hill, 1996

Selesnick Ivan W. Richard Van Slyke, and Onur G. Guleryuz, "Pixel Recovery via l_1 Minimization in the Wavelet Domain,", *IEEE International Conference on Image Processing (ICIP) 2004*, Singapore.

Steiglitz, Kenneth, Thomas W. Parks, and James F. Kaiser, "METEOR: A Constraint-Based FIR Filter Design Program," *IEEE Trans. On Signal Proc*., v. 40, n. 8, pp. 1901-1901, August 1992.

Yarmish, Gavriel, "A Distributed Implementation of the Simplex Method," *Ph.D. Dissertation*, Polytechnic University, Brooklyn, NY, March, 2001.