Computational Optimization and Applications, 35, 131–133, 2006

DOI: 10.1007/s10589-006-0311-z

COAP 2005 Best Paper Award

In each year, the Computational Optimization and Applications (COAP) editorial board selects a paper from the preceding year's COAP publications for the "Best Paper Award." The recipients of the award for papers published in 2005 are Julian Hall and Ken McKinnon of the University of Edinburgh for their paper "Hyper-sparsity in the revised simplex method and how to exploit it," published in Volume 32, pages 259–283.

Julian Hall and Ken McKinnon have been working together on the simplex method for linear programming since 1990. With the long-term goal of developing an efficient parallel implementation of the simplex method, they started by writing a serial revised simplex solver, EMSOL. By incorporating various algorithmic and computational techniques, such as the EXPAND procedure of Gill et al. [4], and others described by the authors in [8], the speed and reliability of EMSOL had improved significantly by the late 1990's. The development of EMSOL led to one notable academic result [7] and its computational techniques were applied successfully in a number of commercial systems. The computational components of EMSOL also formed the basis of the parallel algorithms ASYNPLEX [6] and PARSMI [5].

One day in 1998, the authors were profiling their implementation of the basis matrix inversion technique of Tomlin [11] and were surprised that it was not more efficient for a large sparse LP problem with significant network structure. Upon closer examination they identified that, when using Gaussian elimination to factorise the residual "bump" following the triangularisation phase of the Tomlin procedure, the computational cost was dominated by the test for zero in the right-hand-side when the pivotal column is formed by applying the current set of lower triangular multipliers. This persistence in the sparsity of the right-hand-side was initially referred to by Hall as super-sparsity. However, to avoid confusion with another, little-known, LP technique (in which the occurrence of few different values in the constraint matrix is exploited to save storage), they settled on the term hyper-sparsity.

Hall and McKinnon immediately realised that the same behaviour was likely to be observed when applying the basis matrix factors to the two linear systems that are solved in each iteration of the revised simplex method, and that the effect of this computational inefficiency on overall solution time was likely to be very much greater than its impact on basis matrix inversion. Indeed, for most hyper-sparse problems the size of the Tomlin "bump" and, thus, the cost of its factorisation, are negligible.

For LP problems that exhibit hyper-sparsity, if it is not exploited in all computational components then, drawing an analogy with Amdahl's law, the value of exploiting hyper-sparsity will depend on the performance of those components in which hyper-sparsity is not exploited. Thus, after developing and implementing their own computational technique to exploit hyper-sparsity during basis matrix inversion and solution of linear systems, the authors addressed the other major computational components.

Exploiting hyper-sparsity within the matrix-vector product and row selection operation performed in each simplex iteration was relatively straightforward. In accordance with the observation above, the computational cost of solving many LP problems was then dominated by that of column selection, a simplex operation whose cost is traditionally considered to be of no real importance. Exploiting hyper-sparsity within column selection was relatively difficult, and was achieved by observing that the vector of *changes* in the values used for column selection is typically sparse. With hyper-sparsity now exploited in all computational components, the final limit on performance was the proportion of simplex iterations in which the particular LP did not exhibit hyper-sparsity.

Hall first presented the authors' work at the Dundee Numerical Analysis Conference in July 1999. There, Nick Gould suggested that Bob Bixby was likely to be presenting similar work at a conference in Cambridge a fortnight later. Thus it became clear that hyper-sparsity in the simplex method was identified and exploited independently by both CPLEX and the Edinburgh group.

Only in November 1999, after a seminar by Hall at the Rutherford Appleton Labs, did John Reid bring to the attention of the authors that what they referred to as hyper-sparsity when solving linear systems had been identified in 1988 by Gilbert and Peierls [3], again in the context of matrix factorisation. That this phenomenon and techniques for its exploitation had been known within the sparse numerical linear algebra community for more than a decade was the main reason that it took until 2005 for the authors' work to appear as a journal publication. Bixby, meanwhile, had made a brief reference to the use of Gilbert and Peierls' work within CPLEX in two works of broader scope in 2000 [2] and 2002 [1], the former appearing in the proceedings of the 1999 Cambridge conference.

Thanks to the Edinburgh group and CPLEX identifying hyper-sparsity in the simplex method and developing techniques for exploiting it, there have been huge performance improvements in the best implementations. These have been so great that, when coupled with new algorithmic techniques in the dual simplex method, it can no longer be assumed that a barrier method is generally the fastest way to solve single LP problems. For the significant class of LP problems that exhibit hyper-sparsity, the dual simplex method is generally preferable and can be more than an order of magnitude faster. For other LP problems, the performance of barrier methods can be similarly dominant although, when a family of LP problems is to be solved, (dual) simplex is still the method of choice.

Unfortunately, for a long time this increased value of the simplex method did not lead to a revival of research interest in the academic world. Many of the computational techniques, particularly those associated with hyper-sparsity, were being developed for use within commercial solvers and, understandably, not published openly. However, the COIN-OR initiative [10] led to commercial quality source code (CLP) appearing in the public domain. Recently, computational techniques used in the very efficient MOPS system have been described in detail by Koberstein [9]. Hopefully these developments, together with the publicity for the simplex method resulting from the authors' COAP 2005 prize paper [8], will lead to a return to academic simplex research.

References

1. R.E. Bixby, "Solving real-world linear programs: A decade and more of progress," Operations Research, vol. 50, no. 1, pp. 3–15, 2002.

2. R.E. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling, "MIP: Theory and practice closing the gap," in M.J.D. Powell and S. Scholtes (Eds.), System Modelling and Optimization: Methods, Theory and Applications, Kluwer, The Netherlands, 2000, pp. 19–49.
3. J.R. Gilbert and T. Peierls, "Sparse partial pivoting in time proportional to arithmetic operations," SIAM J. Sci. Stat. Comput., vol. 9, no. 5, pp. 862–874, 1988.
4. P.E. Gill, W. Murray, M.A. Saunders, and M.H. Wright, "A practical anti-cycling procedure for linearly constrained optimization," Mathematical Programming, vol. 45, pp. 437–474, 1989.
5. J.A.J. Hall and K.I.M. McKinnon, "PARSMI, a parallel revised simplex algorithm incorporating minor iterations and Devex pricing," in J. Waśniewski, J. Dongarra, K. Madsen, and D. Olesen (Eds.), Applied Parallel Computing, volume 1184 of Lecture Notes in Computer Science, Springer, pp. 67–76, 1996.
6. J.A.J. Hall and K.I.M. McKinnon, "ASYNPLEX, an asynchronous parallel revised simplex method algorithm," Annals of Operations Research, vol. 81, pp. 27–49, 1998.
7. J.A.J. Hall and K.I.M. McKinnon, "The simplest examples where the simplex method cycles and conditions where EXPAND fails to prevent cycling," Mathematical Programming, vol. 100, no. 1, pp. 133–150, 2004.
8. J.A.J. Hall and K.I.M. McKinnon, "Hyper-sparsity in the revised simplex method and how to exploit it," Computational Optimization and Applications, vol. 32, no. 3, pp. 259–283, 2005.
9. A. Koberstein, "The Dual Simplex Method - Techniques for a fast and stable implementation," PhD thesis, University of Paderborn, 2005.
10. R. Lougee-Heimer, et al., "The COIN-OR initiative: Open source accelerates operations research progress," ORMS Today, vol. 28, no. 4, pp. 20–22, 2001.
11. J.A. Tomlin, "Pivoting for size and sparsity in linear programming inversion routines," J. Inst. Maths. Applics, vol. 10, pp. 289–295, 1972.

Julian Hall received his BA in Mathematics from the University of Oxford in 1987 and his PhD from the University of Dundee in 1992. Since 1990 he has been employed as a lecturer in the School of Mathematics at the University of Edinburgh.

Ken McKinnon received his BSc in Mathematics and Physics from Glasgow University in 1969 and his PhD from the University of Cambridge in 1974. Since 1972 he has been employed as a lecturer, reader and professor in the School of Mathematics at the University of Edinburgh.