A Non-simplex Active-set Method for Linear Programs in Standard Form

Ángel Santos-Palomo

Department of Applied Mathematics,

University of Málaga,

29071 Málaga,

santos@ctima.uma.es

SPAIN

Pablo Guerrero-García

Department of Applied Mathematics, University of Málaga,

29071 Málaga,

pablito@ctima.uma.es

http://www.satd.uma.es/matap/personal/pablito/

SPAIN

Abstract: In 1973 Gill and Murray proposed a non-simplex active-set method (also known as basis-deficiency-allowing dual simplex variation since P.-Q. Pan's 1998 work) for the primal linear problem in inequality form $\min\{c^T x : A^T x \geq b\}$. In this note a non-simplex active-set method for its dual linear program in standard form $\max\{b^T y : Ay = c, y \geq O\}$ is given, so we obtain a different but equivalent description of the basis-deficiency-allowing primal simplex variation of Pan, which is an exterior method that allows us to work with a subset of independent columns of A which is not necessarily a square basis. The new description reveals more clearly the relations with Gill and Murray's work; moreover, we focus our attention in providing several alternatives to address several features not properly addressed by Pan, namely their sparse implementation, suitable Phase I's and Farkas' connection.

Keywords: linear programming, basis-deficiency-allowing simplex variations, non-simplex active-set methods, Phase I, Farkas' lemma, sparsity.

Ángel Santos-Palomo received his MSc in Mathematics from the University of Sevilla in 1975, and his PhD in Econometrics from the University of Málaga in 1995. He started as a teaching assistant in Numerical Analysis at the University of Málaga in 1975, and became Professor (current position) in 1999. His main interest is in all aspects of numerical optimization.

Pablo Guerrero-García received his MSc in Computing from the University of Málaga in 1994, his MSc in Mathematics from the Complutense University of Madrid in 1998, and his PhD in Mathematics from the University of Málaga in 2002. He has been a teaching assistant in Numerical Analysis at the University of Málaga since 1994. His main interest is in numerical linear algebra and its applications to optimization software.

1. The Proposed Algorithm

Let us consider the usual unsymmetric primal-dual pair of linear programs using a non-standard notation (we have deliberately exchanged the usual roles of b and c, x and y, n and m, and (P) and (D), as in e.g. [1, §2]):

$$(P) \quad \min \ \ell(x) \doteq c^T x \quad , x \in R^n \qquad\qquad (D) \quad \max \ L(y) \doteq b^T y \quad , y \in R^m$$
$$\text{s.t} \quad A^T x \geq b \qquad\qquad\qquad\qquad \text{s.t} \quad Ay = c \quad , y \geq O$$

where $A \in R^{n\times m}$ with $m \geq n$ and $\operatorname{rank}(A) = n$. We denote with F and G the feasible regions of (P) and (D), respectively.

In [2] Gill and Murray proposed a non-simplex active-set algorithm (see also [1, §2.5–2.6], [3] and [4, §8.5.4]) for solving problem (P) starting with a feasible point $x^{(0)}$. In a similar way, we can develop a non-simplex active-set method for its dual linear program in standard form (D). *We allow that the basic* *(dual) variable set has less than n indices* and so we start with a dual feasible point $y^{(0)}$ that can have more than $m-n$ zero elements. We shall indicate with $m_k \le n$ the number of elements of the ordered basic set $B^{(k)}$, with $N^{(k)} \doteq [1:m] \setminus B^{(k)}$ and with $A_k \in R^{n\times m_k}$ and $N_k \in R^{n\times(m-m_k)}$ the submatrices of A formed with the columns included in $B^{(k)}$ and $N^{(k)}$, respectively, with $\mathrm{rank}(A_k) = m_k$; as usual, the i th column of A is denoted by a_i, and consistently the matrix whose i th row is a_i^T is denoted by A^T. We shall indicate with $Z_k \in R^{n\times(n-m_k)}$ a matrix whose columns form a basis (not necessarily orthonormal) of $N(A_k^T)$. Superindex $^{(k)}$ will be withheld when it is clear from context.

The algorithm, using "pseudo-MATLAB" notation, is as follows:

S0. *Initialize*. Let $k \leftarrow 0$ and let $y^{(0)}$ be a given dual feasible point with basic set $B^{(0)}$ and basic matrix A_0.

S1. *Check for optimality*. Compute $x^{(k)}$ as a solution of $A_k^T x = b_B$ and compute the residues $r_N^{(k)} = N_k^T x^{(k)} - b_N$ of the non-basic constraints (residues $r_B^{(k)}$ of the basic ones are zero). If $r_N^{(k)} \ge O$ then stop, because $x^{(k)}$ is optimal.

S2. *Pick a non-basic variable to add to the basic set and define (if possible) the search direction for the basic variables*. Using some appropriate rule, choose an index $j \in 1:(m-m_k)$ such that $r_{N_j}^{(k)} < 0$. With $p \doteq N_j^{(k)}$, determine if it exists a search direction $d_B^{(k)} \in R^{m_k}$ for the basic variables (the search direction for the non-basic ones is $d_N^{(k)} = -e_j \in R^{m-m_k}$) by checking the compatibility of the system $A_k d_B = a_p$. If $Z_k^T a_p = O$ then we can obtain $d_B^{(k)}$ and go on to **S3**. If $Z_k^T a_p \ne O$ such system is not compatible, so $a_p \notin R(A_k)$ and p can be added to $B^{(k)}$ without deleting constraints from $B^{(k)}$; we indicate this fact by letting $i \leftarrow 0$, $q \leftarrow \varnothing$ and going on to **S5** with $\tau = 0$.

S3. *Calculate the maximum feasible step along the search direction*. For $i \in 1:m_k$ calculate the step τ_i such that the i th basic variable would be zero:

$$\tau_i = \begin{cases} y_{B_i}^{(k)} / d_{B_i}^{(k)} & ,\text{if } d_{B_i}^{(k)} > 0 \\ +\infty & ,\text{if } d_{B_i}^{(k)} \le 0 \end{cases} , \quad \text{and let } \tau = \min_{1\le i\le m_k} \tau_i .$$

S4. *Test for unbounded solution and pick a basic variable to delete from the basic set*. If $\tau = +\infty$ then stop, because L is unbounded above in G (i.e., $F = \varnothing$). Else, using an appropriate rule, choose index $i \in 1:m_k$ of a basic variable such that $\tau = \tau_i$.

S5. *Prepare the next iteration*. If $i \ne 0$ then

$$q \doteq B_i^{(k)} \quad ; \quad y_{B_i}^{(k)} \leftarrow \varnothing \quad ; \quad d_{B_i}^{(k)} \leftarrow \varnothing \quad ; \quad B_i^{(k)} \leftarrow \varnothing$$

and delete the i th column from A_k. Then let $N_j^{(k)} \leftarrow \varnothing$ and

$$y_B^{(k)} \leftarrow \begin{bmatrix} y_B^{(k)} - \tau d_B^{(k)} \\ \tau \end{bmatrix} \quad ; \quad B^{(k)} \leftarrow [B^{(k)}, p] \quad ; \quad N^{(k)} \leftarrow [N^{(k)}, q].$$

Append a_p as the last column of A_k and let $k \leftarrow k+1$. Go back to **S1**.

The proposed algorithm is a different but equivalent description of the basis-deficiency-allowing primal simplex variation of [5]. It is an exterior method (so it does not need any primal feasible point to start with) with respect to which Dantzig has exhibited "enthusiastic encouragement and advice" [6]. It ends with optimal solutions for (P) and (D) but these solutions do not necessarily correspond to a square basis, because we maintain here dual feasibility and complementary slackness (not strictly because we can have $y_i = 0$ and $r_i = 0$) and stop when we reach primal feasibility. We could obtain an optimal square basis starting with these optimal solutions by applying the strongly polynomial algorithm of Megiddo [7].

Note that *we are generalizing the well-known primal simplex method (in which* $m_k = n$ *for all* k *) to allow the case* $m_k < n$, so we can directly deal with basic feasible solutions of $Ay = c, y \geq O$ (which correspond to degenerate vertices of G if $m_k < n$) *without including* $n - m_k$ *additional indices in* $B^{(k)}$ *to conform a square basis*. Dense examples of distinctive behaviour with respect to the primal simplex method can be found in [8,5]. The new description reveals more clearly the relations with Gill and Murray's work; in fact, the generalization of the well-known dual simplex method was already done in [2] by allowing to perform non-simplex steps!

The rest of the note is organized as follows. A proof of the finite termination of the §1 algorithm under dual nondegeneracy assumption on iterations in which an exchange is involved is included in §2, as well as several comments on suitable Phase I's (§3) and their sparse implementation (§4), solving either a sequence of sparse compatible systems or else a sequence of sparse least squares problems; the computational behaviour of two of the Phase I's proposed are highly encouraging. These Phase I's are new elementary proofs of the Farkas' lemma (§3), which constitute an alternative to that given by Dax in [9].

2. Proof of Convergence

Under dual nondegeneracy assumption, the convergence of the proposed algorithm can be proved being aware of the following facts. When $y^{(k+1)} = y^{(k)} - \tau d^{(k)}$ with $\tau > 0$, in order to be $b^T y^{(k+1)} \doteq L^{(k+1)} > L^{(k)} \doteq b^T y^{(k)}$ it must hold that $b^T d^{(k)} < 0$; this can be derived easily from $Ad^{(k)} = O$, since

$$Ad^{(k)} = A_k d_B^{(k)} + N_k d_N^{(k)} = a_p - N_k e_j = O,$$

and from $r_B^{(k)} = O$, because

$$b^T d^{(k)} = (A^T x^{(k)} - r^{(k)})^T d^{(k)} = x^{(k)T} A d^{(k)} - r^{(k)T} d^{(k)} =$$
$$= -(r_B^{(k)T} d_B^{(k)} + r_N^{(k)T} d_N^{(k)}) = r_N^{(k)T} e_j = r_{N_j}^{(k)} < 0.$$

The vector $y_B^{(k)}$ of multipliers such that $A_k y_B^{(k)} = c$ is unique because A_k has full column rank. So when $a_p \notin R(A_k)$ and $A_{k+1} = [A_k, a_p]$, the only choice for $y_B^{(k+1)}$ keeps on verifying $A_k y_B^{(k+1)} = c$ is that $y_B^{(k+1)} = [y_B^{(k)}; 0]$, because A_{k+1} will have full column rank too and then $y_B^{(k+1)}$ has to be unique. In this case we have no improvement in the dual space because $L^{(k+1)} = L^{(k)}$. But observe that when $m_k = n$ it is not possible that $a_p \notin R(A_k)$. Finally, when $d_B^{(k)} \leq O$ we have that $(-d^{(k)})$ is a direction (of ascent) of G, for

$$-d^{(k)} \geq O \qquad \text{and} \qquad b^T(-d^{(k)}) > 0 \qquad \text{and} \qquad A(-d^{(k)}) = O,$$

so we can derive the infeasibility of (P) from the dual form of the Farkas' lemma.

Note that we have to resort to some anti-cycling rule (e.g., Bland's rule) to deal with degenerate steps. Nevertheless, to prove finite termination (i.e., that no cycling occurs) we only need to assume dual nondegeneracy on iterations in which an exchange is involved, *because basic variable sets cannot be repeated if we only add an index*.

3. Suitable Phase I's and Farkas' Connection

In order to obtain a dual feasible solution to start with we can easily adapt the Gass' single artificial variable Phase I (e.g., [4, §7.9.4] and [6]). Given a point $y^{(0)} \geq O$ such that $Ay^{(0)} \neq c$, we can define a Phase I by using a single artificial variable z and the initial residue $s^{(0)} = Ay^{(0)} - c$:

$$\begin{array}{lll} \text{minimize} & z & , y \in R^m, z \in R \\ \text{subject to} & Ay - s^{(0)} z = c, y \geq O, z \geq 0 & , A \in R^{n \times m} \end{array}$$

so we have to maximize $\overline{b}^T \overline{y}$ subject to $\overline{A}\overline{y} = \overline{c}$ and $\overline{y} \geq O$, where:

$\overline{b} = [O_m; -1] = -e_{m+1}$; $\overline{A} = [A, -s^{(0)}]$; $\overline{y} = [y; z]$; $\overline{c} = c$.

The point $\overline{y}^{(0)} = [y^{(0)}; 1]$ is feasible with respect to the constraints of the Phase I, but with $m_0 < n$. This is the reason why we solve this Phase I using the proposed non-simplex active-set method, e.g. with $y^{(0)} = O$ and $B^{(0)} = [m+1]$. In this case the primal problem corresponding to the proposed Phase I is

$$\begin{array}{lll} \text{minimize} & c^T d^P & , d^P \in R^n \\ \text{subject to} & \begin{bmatrix} A^T \\ c^T \end{bmatrix} d^P \geq \begin{bmatrix} O_m \\ -1 \end{bmatrix} & , A \in R^{n \times m}, \end{array}$$

and then a primal null-space descent direction d^P come into play [10, 8].

As an alternative Phase I we can proceed as in [11]. We can apply the NNLS algorithm [12, §23.3] or the Dax algorithm [9] to the positive semidefinite quadratic problem

$$\text{minimize} \quad 1/2 \cdot \| Ay - c \|_2^2 \quad \text{subject to} \quad y \geq O, \tag{1}$$

which implies solving a sequence of unconstrained least squares problems of the form $\min \| A_k z - c \|_2$ (NNLS) or of the form $\min \| A_k w - s^{(k)} \|_2$ (Dax). In this case we do not need any artificial variable and what we obtain is a direction of F or a dual basic feasible solution with $m_k \leq n$. It turns out that the computation can also be paraphrased in terms of a primal null-space descent direction [13, 10]. Instead of (1) we can deal with its Wolfe dual (least distance) strictly convex quadratic problem

$$\text{minimize} \quad 1/2 \cdot \| u \|_2^2 \quad \text{subject to} \quad A^T u \geq A^T c \doteq v, \quad u \in R^n, \tag{2}$$

which also implies solving a sequence of least squares problems. To solve (2) we can particularize any of the primal, dual and primal-dual active-set methods introduced in [14].

We could also have adopted the dual point of view, namely using the non-simplex active-set algorithm of [2] or [4, §8.5] and the Gass' Phase I with a single artificial variable (e.g., [4, p. 314]),

$$\begin{array}{lll} \text{minimize} & z & , x \in R^n, z \in R \\ \text{subject to} & A^T x - r^{(0)} z \geq b, z \geq 0 & , A \in R^{n \times m}, \end{array}$$

where $r^{(0)} = A^T x^{(0)} - b$ and $\overline{x}^{(0)} = [x^{(0)}; 1]$, e.g. with $x^{(0)} = O$ (a primal degenerate vertex) and $B^{(0)} = \varnothing$ (the case $B^{(0)} = 1:m$ and $N^{(0)} = [m+1]$ needs a further generalization of the simplex method [15] to allow $m_k \geq n$ and so it is intended to deal mainly with primal degeneracy). Note that we also need an anti-cycling rule to prove finiteness, and that in this case the dual problem corresponding to the proposed Phase I is

$$\begin{array}{lll} \text{maximize} & [b; 0]^T \delta & , \delta \in R^{m+1} \\ \text{subject to} & \begin{bmatrix} A & O_n \\ b^T & 1 \end{bmatrix} \delta = \begin{bmatrix} O_n \\ 1 \end{bmatrix}, \delta \geq O & , A \in R^{n \times m}. \end{array}$$

From this point of view, we can also use a least-squares-based Phase I, solving

$$\text{minimize} \quad 1/2 \cdot \| x \|_2^2 \quad \text{subject to} \quad A^T x \geq b, \quad x \in R^n$$

with a quadratic algorithm or using the Cline's method described in [12, §23.4].

All the Phase I's given above are new elementary proofs of the Farkas' lemma, which constitute an alternative to the proof given by Dax in [9]. Note that in the NNLS-based case, we do not have to resort to any anti-cycling rule in exact arithmetic. An application of this idea can be found in [16].

4. Sparse Implementation Details and Computational Results

The implementation of the algorithmic schemes given above for dense problems can be accomplished with the QR factorization of A_k (adding and deleting columns); but nowadays we are mainly interested in sparse problems, and hence we have tried two sparse orthogonal approaches.

Our first sparse orthogonal approach use the sparse QR factorization of A_k^T (adding and deleting rows). Since A_k is formed by a subset of the columns of a fixed matrix A, we have adapted [17] to matrices with more rows than columns Saunders' techniques for square matrices [18] using the static data structure of George and Heath (e.g., see [19] and the references therein) but allowing row downdating on it; this way we can take advantage of the intermediate results obtained when dealing with the problem $\min \| A^T x - b \|_2$ by processing A^T row by row in relationship with (D).

The disadvantage of the least-squares-based Phase I's with respect to the Phase I's with a single artificial variable is that the necessity to obtain least squares solutions does not allow us to work with the QR factorization of A_k^T. In [3, p. 321] iterative relaxation techniques are recommended, but in this way we do not take advantage of the existing relationship between A_k and A_{k+1}. A suitable (second) sparse orthogonal approach is to adapt the methodology of Björck [20] and Oreborn [21] to be able to apply the sparse NNLS algorithm with a "short-and-fat" matrix A. They proposed an active set algorithm for the sparse least squares problem

$$\text{minimize} \qquad 1/2 \cdot y^T C y + d^T y, \quad y \in R^m \qquad \text{subject to} \qquad l \le y \le u$$

with $C > 0$. In our problem

$$C = A^T A \quad \wedge \quad d = -A^T c \quad \wedge \quad \forall i \in 1:m, l_i = 0 \wedge u_i = +\infty$$

but $C \ge 0$, hence to maintain a sparse QR factorization of A_k we have had to adapt [13] the proposed technique as in [22], but without forming C.

We have conducted [23, §5] some computational experiments comparing slight modifications of two of the Phase I's given above against the TOMLAB LPSOLVE V3.0 [24] sparse implementation of the usual primal simplex method. The sparse technique chosen depends on the primal null-space descent direction used. The computational results, both for sparse adaptations of classical test problems and for the first 31 smallest NETLIB problems [25], are the subject of a forthcoming paper, in which we will show the clear advantage obtained in number of iterations, quality of solutions and execution time when suitable pivot strategies are used and with no special anti-cycling tools.

The sparse orthogonal approaches that we have tried have the added bonus that they are parallelisable and they fit well within a mixed interior-point simplex methodology. This is due to the fact that, in spite of using active-set methods, we work on top of the static structure of the Cholesky factor of AA^T (our first approach, as in interior-point algorithms using the normal equations approach) or else on that of $A^T A$ (our second approach, as in the interior-point method described in [26]).

Acknowledgements

The authors thank M. Saunders who, with respect to the technical report prior to [11], commented us: "Their method is equivalent to the known (and very simple) NNLS algorithm for non-negative least squares. Their real contribution was to apply it to linear programs to find a feasible solution". We also thank A. Dax and Å. Björck for providing us with references [3] and [20].

REFERENCES

1. OSBORNE, M.R., **Finite Algorithms in Optimization and Data Analysis**, John Wiley and Sons, Chichester, 1985.

2. GILL, P.E. and MURRAY, W., **A numerically stable form of the simplex algorithm**, Linear Algebra Appl., 7:99–138, 1973.

3. DAX, A., **Linear programming via least squares**, Linear Algebra Appl., 111:313–324, 1988.

4. GILL, P.E., MURRAY, W. and WRIGHT, M.H., **Numerical Linear Algebra and Optimization,** Vol. 1, Addison-Wesley, Redwood City, 1991.

5. PAN, P.-Q., **A basis-deficiency-allowing variation of the simplex method for linear programming**, Computers Math. Applic., 36(3):33–53, August 1998.

6. PAN, P.-Q. and PAN, Y.-P., **A phase-1 approach for the generalized simplex algorithm**, Computers Math. Applic., 42(10/11):1455–1464, November 2001.

7. MEGIDDO, N., **On finding primal- and dual-optimal bases**, ORSA J. Computing, 3(1):63–65, Winter 1991.

8. SANTOS-PALOMO, Á., **The sagitta method for solving linear programs**, Europ. J. Operl. Res., 157(3):527–539, September 2004.

9. DAX, A., **An elementary proof of Farkas' lemma**, SIAM Rev., 39(3):503–507, September 1997.

10. SANTOS-PALOMO, Á. and GUERRERO-GARCÍA, P., **Sagitta method with guaranteed convergence**, Technical report, Department of Applied Mathematics, University of Málaga, 2005. Submitted for publication.

11. DANTZIG, G.B. and LEICHNER, S.A. and DAVIS, J.W., **A strictly improving linear programming Phase I algorithm**, Ann. Oper. Res., 46/47:409–430, 1993.

12. LAWSON, C.L. and HANSON, R.J., **Solving Least Squares Problems**, Original edition, Prentice-Hall, Englewood Cliffs, NJ, USA, 1974.

13. SANTOS-PALOMO, Á. and GUERRERO-GARCÍA, P., **Solving a sequence of sparse least squares problems**, Technical report MA-03-03, Department of Applied Mathematics, University of Málaga, September 2003. Submitted for publication.

14. SANTOS-PALOMO, Á., **New range-space active-set methods for strictly convex quadratic programming**, In OLIVARES-RIEUMONT, P. (ed.), Proceedings of the III Conference on Operations Research, volume 2, page 27, Universidad de La Habana, Cuba, March 1997.

15. GUERRERO-GARCÍA, P. and SANTOS-PALOMO, Á., **Working with non-simplex active-set full row rank sparse matrices**, Technical report, Department of Applied Mathematics, University of Málaga, 2005. In preparation.

16. GUERRERO-GARCÍA, P. and SANTOS-PALOMO, Á., **Gyula Farkas would also feel proud**, Technical report MA-02-01, Department of Applied Mathematics, University of Málaga, November 2002. Submitted for publication.

17. GUERRERO-GARCÍA, P. and SANTOS-PALOMO, Á., **Solving a sequence of sparse compatible systems**, In GRIFFITHS, D. and WATSON, G.A. (eds.), Proceedings of the 19th Biennial Conference on Numerical Analysis, pages 23–24, Dundee, Scotland, June 2001. Submitted for publication.

18. SAUNDERS, M.A., **Large-scale linear programming using the Cholesky factorization**, Tech. report CS-TR-72-252, Computer Science Department, Stanford University, CA, January 1972.

19. HEATH, M.T., **Numerical methods for large sparse linear least squares problems**, SIAM J. Sci. Statist. Comput., 5:497–513, 1984.

20. BJÖRCK, Å., **A direct method for sparse least squares problems with lower and upper bounds**, Numer. Math., 54:19–32, 1988.

21. OREBORN, U., **A Direct Method for Sparse Nonnegative Least Squares Problems**, Licentiat thesis, Department of Mathematics, Linköping University, Sweden, 1986.

22. COLEMAN, T.F. and HULBERT, L.A., **A direct active set algorithm for large sparse quadratic programs with simple lower bounds**, Math. Progr., 45:373–406, 1989.

23. GUERRERO-GARCÍA, P., **Range-Space Methods for Sparse Linear Programs** (Spanish), Ph. D. thesis, Departamento de Matemática Aplicada, Universidad de Málaga, Spain, July 2002.

24. HOLMSTRÖM, K., **The TOMLAB optimization environment v3.0 user's guide**, Technical report, Mälardalen University, Sweden, April 2001.

25. GAY, D.M., **Electronic mail distribution of linear programming test problems**, Committee on Algorithms (COAL) Newsletter, 13:10–12, 1985.

26. BIRGE, J.R. and FREUND, R.M. and VANDERBEI, R.J., **Prior reduced fill-in solving equations in interior point algorithms**, Oper. Res. Lett., 11:195–198, May 1992.