Updating Lower Bounds

when Using Karmarkar's Projective Algorithm

for Linear Programming[1]

J. E. MITCHELL[2]

Communicated by R. Tapia

---

[1]This research was partially supported by ONR Grant number N00014–90–J–1714.

[2]Assistant Professor, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York.

Abstract.

We give two results related to Gonzaga's recent paper showing that lower bounds derived from the Todd-Burrell update can be obtained by solving a one-variable linear programming problem involving the centering direction and the affine direction. We show how these results may be used to update the primal solution when using the dual affine variant of Karmarkar's algorithm. This leads to a dual projective algorithm.

Key Words.

Karmarkar's algorithm, lower bounds, linear programming, interior point methods.

this stuff is here to get the labelling to work.

1 Introduction

2 Todd's two variable update has the Gonzaga form

3 The combination of the affine and centering directions used in Karmarkar's algorithm

4 A dual projective algorithm

5 Derivation of the algorithm

5.1 Rescaling the problem

5.2 Finding the affine and centering directions

5.3 Updating the primal solution x

5.4 The Primal Problem Corresponding to (D_{wz}).

1. Introduction

We give two results related to Gonzaga's recent paper (Ref. 1) showing that lower bounds derived from the Todd-Burrell update (Ref. 2) can be obtained by solving a one-variable linear programming problem involving the centering direction and the affine direction. Firstly, Gonzaga showed that the Todd-Burrell update is obtained by solving the one-variable problem in the homogenized space; in Section 2 we show that Todd's two-variable update (Ref. 3) is equivalent to Gonzaga's one-variable problem in the original space. Secondly, in Section 3 we show that the resulting direction for Karmarkar's algorithm (Ref. 4) can be expressed very simply in terms of the optimal solution to Gonzaga's one-variable linear program.

In Section 4 we give the centering direction for a dual algorithm, when the primal variables have simple upper bounds. A feasible primal solution can then be obtained by solving a 1-variable LP. This leads to a dual projective algorithm. The centering direction and affine directions are derived using a QR-factorization in Section 5.

We use the following notation throughout: e denotes a vector of ones of the appropriate dimension. Given a lower-case vector such as $x = (x_1, ..., x_n)^T$, the diagonal matrix $\mathrm{diag}(x_1, ..., x_n)$ is denoted by the upper-case letter X. The matrix operator which projects a vector u onto the null space of a full row rank matrix M is denoted by P_M. Thus

$$P_M u = (I - M^T(MM^T)^{-1}M)u \tag{1}$$

for any vector u.

2. Todd's Two Variable Update has the Gonzaga Form

In Sections 2 and 3 we work with the standard form linear programming problem

$$(P) \qquad \begin{array}{lrcl} \min & c^T x, & & \\ \text{s.t.} & Ax & = & b, \\ & x & \geq & 0. \end{array}$$

Here, A is a full row rank $m \times n$ matrix, c and x are n-vectors, and b is an m-vector. The dual problem to (P) is

$$(D) \qquad \begin{array}{lrcrcl} \max & b^T y, & & & & \\ \text{s.t.} & A^T y & + & z & = & c, \\ & & & z & \geq & 0, \end{array}$$

where y is an m-vector and z is an n-vector of dual slacks. We assume we have a current primal feasible solution $\bar{x} > 0$ and a lower bound $\bar{v}$ on the optimal value of (P). The matrix $A\bar{X}$ is denoted by $\bar{A}$ and the vector $\bar{X}c$ by $\bar{c}$. Projection onto the null space of the matrix $\bar{A}$ is denoted by the subscript p; thus, $P_{\bar{A}} u = u_p$ for any vector u.

Gonzaga updates the lower bound to (P) by solving the one-variable linear program

$$(D_G) \qquad \begin{array}{lrcl} \max & \mu, & & \\ \text{s.t.} & \mu(e - e_p) & \leq & \bar{c}_p. \end{array}$$

If μ is feasible in (D_G) then $z = \bar{X}^{-1}(\bar{c}_p - \mu(e - e_p))$ is a vector of dual slacks which is feasible in (D). If the optimal solution is $\hat{\mu}$ then $(\bar{c} - \bar{c}_p)^T e + a\hat{\mu}$ is a valid lower bound for (P), where $a = \|e - e_p\|^2$. This lower bound comes from the fact that the duality gap between $\bar{x}$ and z is $\bar{x}^T z$.

The problem (P) is equivalent to the problem

$$(\tilde{P}) \qquad \begin{array}{lrcl} \min & \tilde{c}^T x & & \\ \text{s.t.} & \tilde{A}x & = & 0 \\ & g^T x & = & 1 \\ & x & \geq & 0. \end{array}$$

where the $m \times (n+1)$ matrix $\tilde{A} := [A| - b]$, the $(n+1)$-vector $\tilde{c} := (c^T, 0)^T$, and the $(n+1)$-vector g is the $(n+1)$st unit vector $g = (0^T, 1)^T$. (Hereafter, we abuse notation and write $g = (0, 1)$, etc.) Note that x is now an $(n+1)$-vector. Projective algorithms work on problems of this type. The Todd-Burrell lower bound update for this problem is obtained by solving the one-variable problem

$$(D_{TB}) \qquad \begin{array}{lrcl} \max & \nu, & & \\ \text{s.t.} & \nu P_{\tilde{A}\tilde{X}} g & \leq & P_{\tilde{A}\tilde{X}} \tilde{X}\tilde{c}, \end{array}$$

where $\tilde{x} = (\bar{x}, 1)$ is the current primal solution. Writing H for the $(m+1) \times (n+1)$ matrix consisting of $\tilde{A}$ with the extra row g^T appended, Gonzaga shows that solving (D_{TB}) is equivalent to solving the one variable problem

$$\begin{array}{lrcl} \max & \mu, & & \\ \text{s.t.} & \mu(e - e_{\tilde{p}}) & \leq & (\tilde{X}\tilde{c})_{\tilde{p}}, \end{array}$$

where $u_{\tilde{p}} := P_{H\tilde{X}} u$ for any $(n+1)$-vector u.

In (Ref. 3) , Todd defines a refinement of the Todd-Burrell procedure. This involves solving the two-variable linear program

$$
(D_{T2}) \qquad \begin{aligned} \max \quad & \nu, \\ \text{s.t.} \quad & \nu g \; + \; \eta(-g + P_{\tilde{A}\tilde{X}} g) \; \leq \; P_{\tilde{A}\tilde{X}} \tilde{X} \tilde{c}. \end{aligned}
$$

(This bound was also derived independently by de Ghellinck and Vial (Ref. 5) and by Ye and Kojima (Ref. 6).) We now show that solving this linear program is exactly equivalent to solving the linear program (D_G).From the Sherman-Morrison-Woodbury formula,

$$
\begin{aligned}
(\tilde{A}\tilde{X}^2\tilde{A}^T)^{-1} &= (A\bar{X}^2A^T + bb^T)^{-1} \\
&= (A\bar{X}^2A^T)^{-1} - \frac{1}{1 + b^T(A\bar{X}^2A^T)^{-1}b}(A\bar{X}^2A^T)^{-1}bb^T(A\bar{X}^2A^T)^{-1}. \quad (2)
\end{aligned}
$$

Writing r for $1/(1 + b^T(A\bar{X}^2A^T)^{-1}b)$, it follows that

$$
\begin{aligned}
g - P_{\tilde{A}\tilde{X}} g &= \tilde{X}\tilde{A}^T(\tilde{A}\tilde{X}^2\tilde{A}^T)^{-1}\tilde{A}\tilde{X}g = -\tilde{X}\tilde{A}^T(\tilde{A}\tilde{X}^2\tilde{A}^T)^{-1}b = -r\tilde{X}\tilde{A}^T(A\bar{X}^2A^T)^{-1}b \\
&= -r \begin{bmatrix} \bar{X}A^Tw_2 \\ -b^Tw_2 \end{bmatrix} \qquad (3)
\end{aligned}
$$

and

$$
\begin{aligned}
P_{\tilde{A}\tilde{X}} \tilde{X}\tilde{c} &= \tilde{X}\tilde{c} - \tilde{X}\tilde{A}^T(\tilde{A}\tilde{X}^2\tilde{A}^T)^{-1}\tilde{A}\tilde{X}^2\tilde{c} = \tilde{X}\tilde{c} - \tilde{X}\tilde{A}^T(\tilde{A}\tilde{X}^2\tilde{A}^T)^{-1}A\bar{X}^2c \\
&= \begin{bmatrix} \bar{X}c \\ 0 \end{bmatrix} - \begin{bmatrix} \bar{X}A^Tw_1 \\ -b^Tw_1 \end{bmatrix} + rb^Tw_1 \begin{bmatrix} \bar{X}A^Tw_2 \\ -b^Tw_2 \end{bmatrix} \\
&= \begin{bmatrix} \bar{c}_p \\ 0 \end{bmatrix} + rb^Tw_1 \begin{bmatrix} \bar{X}A^Tw_2 \\ 1 \end{bmatrix} \qquad (4)
\end{aligned}
$$

where

$$
w_1 = (A\bar{X}^2A^T)^{-1}A\bar{X}^2c \qquad (5)
$$

$$
w_2 = (A\bar{X}^2A^T)^{-1}b. \qquad (6)
$$

Note that $\bar{c}_p = \bar{c} - \bar{X}A^T w_1$ and $e - e_p = \bar{X}A^T w_2$. Therefore, (D_{T2}) is equivalent to

$$\begin{array}{lrcl} \max & \nu, & & \\ \text{s.t.} & \eta r(e - e_p) & \leq & \bar{c}_p + rb^T w_1(e - e_p), \\ & \nu \; - \; \eta r b^T w_2 & \leq & rb^T w_1. \end{array}$$

The last constraint will be active at the optimal solution $(\hat{\nu}, \hat{\eta})$, so $\hat{\nu} = rb^T w_1 + \hat{\eta} r b^T w_2$. Since $r > 0$ and $b^T w_2 > 0$, the objective function can be restated as $\max \eta$ and the last constraint can be dropped. Making the substitution $\gamma = r(\eta - b^T w_1)$ gives the equivalent problem

$$\begin{array}{lrcl} \max & \gamma, & & \\ \text{s.t.} & \gamma(e - e_p) & \leq & \bar{c}_p, \end{array}$$

which is exactly the problem (D_G). If γ is feasible in this problem then $y = w_1 + \gamma w_2$ is feasible in (D) with corresponding dual slacks $\bar{X}^{-1}(\bar{c}_p - \gamma(e - e_p))$. Ye (Ref. 7) and Freund (Ref. 8) (among others) also update the dual solution by using a combination of w_1 and w_2.

3. Combinations of the Affine and Centering Directions Used in Interior Point Algorithms

The directions used in primal interior point algorithms are combinations $-\bar{X}(\bar{c}_p - \beta e_p)$ of the centering direction e_p and the affine direction $\bar{c}_p$. For example, Gonzaga (Ref. 9) uses

$$\beta = \frac{c^T \bar{x} - \bar{v} - \bar{c}^T e_p}{b^T (A\bar{X}^2 A^T)^{-1} b} \tag{7}$$

and de Ghellinck and Vial (Ref. 5), Ye and Kojima (Ref. 6), Gonzaga (Refs. 9 and 10), and Mitchell and Todd (Ref. 11) all use

$$\beta = \frac{c^T \bar{x} - \bar{v} - \bar{c}^T e_p}{1 + b^T (A\bar{X}^2 A^T)^{-1} b}. \tag{8}$$

(See den Hertog and Roos (Ref. 12) for a survey.) In this section, we show how these expressions may be simplified using the optimal solution $\hat{\mu}$ to the dual problem (D_G).

If solving (D_G) led to an improvement in the lower bound then

$$\bar{v} = (\bar{c} - \bar{c}_p)^T e + a\hat{\mu}, \tag{9}$$

where

$$a = \|e - e_p\|^2 = b^T (A\bar{X}^2 A^T)^{-1} b. \tag{10}$$

Thus

$$c^T \bar{x} - \bar{v} - \bar{c}^T e_p = \bar{c}^T e - (\bar{c} - \bar{c}_p)^T e - a\hat{\mu} - \bar{c}^T e_p = -a\hat{\mu}. \tag{11}$$

Therefore, we have the following theorem:

Theorem 3.1 If the lower bound $\bar{v}$ is updated by solving (D_G) and this leads to an improved lower bound then Gonzaga (Ref. 9) uses the primal direction $-\bar{X}(c_p + \hat{\mu} e_p)$ and de Ghellinck and Vial (Ref. 5), Ye and Kojima (Ref. 6), Gonzaga (Refs. 9 and 10), and Mitchell and Todd (Ref. 11) all use the primal direction $-\bar{X}(c_p + \frac{a\hat{\mu}}{1+a} e_p)$, where $a = \|e - e_p\|^2$.

4. A Dual Projective Algorithm

We now consider linear programs where the primal variables have both upper and lower bounds. We develop a dual projective algorithm for problems of this type and use the results of the earlier sections to show how primal feasible solutions may be constructed. The results given in this section are proved in Section 5. An affine algorithm for problems of this type was described in Marsten et al. (Ref. 13). The centering direction has been given in the documentation to the package OB1 (Ref. 14). Gay (Ref. 15) described a dual projective algorithm for problems in standard form. Instead of using Gay's algorithm to derive our algorithm, we use Mitchell and Todd (Ref. 16) , which describes a projective algorithm for problems where some of the variables are unrestricted in sign. This algorithm eliminates the variables which are unrestricted in sign. This derivation enables us to formulate a constrained version of the primal problem, which can be solved easily to give primal estimates.

By adding a constant to the primal variables, we can assume the lower bounds are zero. Therefore, we work with the problem

$$(P_U) \qquad \begin{array}{lrcl} \min & c^T x & & \\ \text{s.t.} & Ax & = & b \\ & x & \leq & u \\ & x & \geq & 0. \end{array}$$

Here, A is an $m \times n$ matrix, c, u, and x are n-vectors, and b is an m-vector. The dual problem to (P_U) is

$$(D_U) \qquad \begin{array}{lrcrcrcl} \max & b^T y & - & u^T w & & & & \\ \text{s.t.} & A^T y & - & w & + & z & = & c \\ & & & w, z & & \geq & & 0, \end{array}$$

where w and z are n-vectors.

We make the following assumptions concerning (P_U) and (D_U).

Assumption A1: The rows of A have full rank.

Assumption A2: The sets of optimal solutions to (P_U) and (D_U) are bounded.

Assumption A3: We have a vector $(\bar{y}, \bar{w}, \bar{z})$ which is strictly feasible in (D_U).

Assumption A4: We have an upper bound $\bar{B}$ on the optimal value of (D_U).

When the unrestricted variables y are eliminated from (D_U), the resulting problem is an equality constrained problem where all the variables are constrained to be nonnegative, ie, a standard form problem. It follows that this problem can be solved using a standard method.

As in Marsten et al. (Ref. 13), we define

$$\bar{D}^2 := (\bar{W}^2 + \bar{Z}^2)^{-1}. \tag{12}$$

As we show in Section 5, the upper bound can be updated by solving the one variable linear program

$$(P(\lambda)) \qquad \begin{array}{lrcccl} \min & & & \lambda & & \\ \text{s.t.} & v^1 & \leq & \lambda v^2 & \leq & u + v^1, \end{array}$$

where

$$\begin{aligned} v^1 &= -\bar{D} P_{A\bar{D}} \bar{D} \bar{W}^2 u - \bar{D}^2 A^T (A \bar{D}^2 A^T)^{-1} b \qquad (13) \\ &= -\bar{D}^2 \bar{W}^2 u - \bar{D}^2 A^T (A \bar{D}^2 A^T)^{-1} (b - A \bar{D}^2 \bar{W}^2 u) \\ v^2 &= \bar{D} P_{A\bar{D}} \bar{D} c. \qquad (14) \end{aligned}$$

The scalar λ corresponds to the primal solution

$$x(\lambda) := -v^1 + \lambda v^2 \tag{15}$$

so $\bar{B}$ can be updated if the optimal solution $\bar{\lambda}$ to $(P(\lambda))$ satisfies $c^T x(\bar{\lambda}) < \bar{B}$. Notice that $x(\lambda)$ satisfies $Ax = b$ for any λ and that λ is feasible in $(P(\lambda))$ if and only if $x(\lambda)$ is feasible in (P_U).

The affine direction for the problem (D_B) is given by

$$\begin{aligned} dy^{aff} &:= (A\bar{D}^2A^T)^{-1}(b - A\bar{D}^2\bar{W}^2u) & (16)\\ dw^{aff} &:= -\bar{Z}^2\bar{D}^2\bar{W}^2u + \bar{W}^2\bar{D}^2A^Td_y^{aff} = -\bar{W}^2(u+v^1) & (17)\\ dz^{aff} &:= -\bar{Z}^2\bar{D}^2\bar{W}^2u - \bar{Z}^2\bar{D}^2A^Td_y^{aff} = \bar{Z}^2v^1. & (18) \end{aligned}$$

The centering direction is given by

$$\begin{aligned} dy^{cent} &:= (A\bar{D}^2A^T)^{-1}A\bar{D}^2(\bar{w} - \bar{z}) & (19)\\ dw^{cent} &:= \bar{w} + \bar{W}^2\bar{D}P_{A\bar{D}}\bar{D}c = \bar{w} + \bar{W}^2v^2 & (20)\\ dz^{cent} &:= \bar{z} - \bar{Z}^2\bar{D}P_{A\bar{D}}\bar{D}c = \bar{z} - \bar{Z}^2v^2. & (21) \end{aligned}$$

As mentioned in Section 3, projective directions are obtained by combining the centering direction and the affine direction appropriately. We define each iteration of the extended dual projective algorithm as follows:

<u>Step 1:</u> Solve $(P(\lambda))$ to obtain $\bar{\lambda}$.

<u>Step 2:</u> If $c^T x(\bar{\lambda}) \geq \bar{B}$ then $\bar{\lambda} = (\bar{B} + c^T v^1)c^T v^2$. Otherwise, update x to $x(\bar{\lambda})$ and $\bar{B}$ to $c^T x$.

Step 3: Use equations (16), (17), (18), (19), (20), and (21) to find dy^{aff}, dw^{aff}, dz^{aff}, dy^{cent}, dw^{cent} and dz^{cent} respectively.

Step 4: Let $\gamma = \bar{\lambda} c^T v^2/(1 + c^T v^2)$, with v^2 given in (14).

Step 5: Set $dy = dy^{aff} + \gamma dy^{cent}$, $dw = dw^{aff} + \gamma dw^{cent}$, and $dz = dz^{aff} + \gamma dz^{cent}$.

Step 6: Choose a step length α which ensures $w + \alpha dw > 0$ and $z + \alpha dz > 0$.

Step 7: Update $y \leftarrow y + \alpha dy$, $w \leftarrow w + \alpha dw$, and $z \leftarrow z + \alpha dz$.

Step 2 tests whether the new $\bar{\lambda}$ found in Step 1 leads to an improvement in the upper bound $\bar{B}$. If it does not, we choose $\bar{\lambda}$ to correspond to the incumbent x. In Step 4, the quantity $c^T v^2$ is strictly positive, provided the current primal-dual point is not optimal. Thus, Step 4 is well defined. The choice of γ in Step 4 follows from arguments similar to those in Section 3.

As presented, three systems of equations of the form $(A\bar{D}^2A^T)w = p$ are solved at each iteration for different right hand sides p, one when finding v^1, one when finding v^2 and one when finding dy^{cent}. It is possible to solve the problem without calculating y, obtaining it at the end by solving the equation

$$(A\bar{D}^2A^T)y = A\bar{D}^2(c + w - z),$$

provided sufficient accuracy is maintained in w and z. This makes it unnecessary to calculate dy^{cent} at each iteration.

The complete extended dual projective algorithm is as follows:

Step 1: Input data, including initial strictly dual feasible solution $(\bar{y}, \bar{w}, \bar{z})$, upper bound $\bar{B}$ and duality gap tolerance ϵ.

Step 2: If $\bar{B} - b^T\bar{y} + u^T\bar{w} < \epsilon$, STOP.

Step 3: Perform an iteration of the algorithm and return to Step 2.

The standard comments regarding projective algorithms apply for this algorithm. In particular, the algorithm is polynomially convergent; it generates feasible primal solutions; it requires the solution of two systems of equations (with the same matrix) per iteration whereas the corresponding affine algorithm only requires one; and the algorithm is less likely to get trapped in corners than the affine algorithm. These comments suggest that it may be useful to use the projective algorithm initially to get a reasonably well centered point, then switch to the affine algorithm to exploit the cheaper cost of iterations, and then generate new primal solutions at the end of the run using the projective step.

5. Derivation of the Algorithm

In this section, we derive the centering direction and the update of the upper bound. We also give a different derivation of the affine direction from the one in Marsten et al. (Ref. 13). We will use the results of Mitchell and Todd (Ref. 16).

A projective algorithm for problems with some variables unrestricted in sign was presented in (Ref. 16). The problem (D_U) is exactly of this form, so the results of (Ref. 16) can be applied to the current situation. The algorithm of (Ref. 16) was obtained in two different ways: by eliminating the unrestricted variables, and by solving a constrained least squares problem. We will obtain the directions and the problem $(P(\lambda))$ given in Section 4 by eliminating the unconstrained variables y, but first we use the constrained least squares problem to justify a rescaling of the problem by the diagonal matrix $\bar{D}$.

5.1. Rescaling the Problem

The Least Squares Approach of Mitchell and Todd (Ref. 16)

The notation in this paragraph is unrelated to that used in the rest of this paper. The results in (Ref. 16) are stated in terms of problems of the form $\min\{c_A^T x_A + c_F^T x_F : Ax_A + Fx_F = 0, g_A^T x_A + g_F^T x_F = 1, x_A \geq 0\}$. Standard form problems with some unrestricted variables can be transformed into this form by adding an extra variable corresponding to the right hand side and a homogenizing constraint requiring the extra variable to take the value one. The direction is obtained by considering a family of constrained least squares problems $\min\{\|A^T y - (c_A - zg_A)\| : F^T y = c_F - zg_F\}$, where z is a lower bound on the linear program. The direction followed by the variables x_A is the negative of the optimal residual of the constrained least squares problem for an appropriate choice of z. Notice that the optimal residual is not altered if the constraints $Ax_A + Fx_F = 0$ are multiplied by a diagonal matrix D, giving $DAx_A + DFx_F = 0$; this merely rescales the optimal y in the constrained least squares problem.

Using Mitchell and Todd (Ref. 16) to Rescale the Current Problem

Let $(\bar{y}, \bar{w}, \bar{z})$ be a current feasible solution to (D_U), with $\bar{w} > 0$ and $\bar{z} > 0$. Since the direction obtained is not affected by scaling the constraints, we can replace (D_U)

by the problem

$$
(D_U(\bar{D})) \qquad \begin{array}{rrcrcrcl}
\max & b^T y & - & u^T w & & & & \\
\text{s.t.} & \bar{D}A^T y & - & \bar{D}w & + & \bar{D}z & = & \bar{D}c \\
& & & w, z & \geq & 0, & &
\end{array}
$$

where $\bar{D}$ is as defined in equation (12). As will be seen, this rescaling considerably simplifies calculation of the affine and centering directions.

5.2. Finding the Affine and Centering Directions

Following (Ref. 16), we eliminate the unrestricted variables y to obtain a standard form problem; we then derive the affine direction and centering direction in this equivalent problem. In order to eliminate y, we find a QR-factorization of the matrix $\bar{D}A^T$:

$$
\bar{D}A^T = [Q_1|Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R, \tag{22}
$$

where $Q := [Q_1|Q_2]$ is an orthogonal matrix and R is a square upper-triangular matrix. By assumption, A has full rank and every diagonal element of $\bar{D}$ is positive; it follows that R has full rank. Multiplying the equality constraints of $(D_U(\bar{D}))$ by Q^T gives the linear program

$$
\begin{array}{rrcrcrcl}
\max & b^T y & - & u^T w & & & & \\
\text{s.t.} & Ry & - & Q_1^T \bar{D}w & + & Q_1^T \bar{D}z & = & Q_1^T \bar{D}c \\
& & - & Q_2^T \bar{D}w & + & Q_2^T \bar{D}z & = & Q_2^T \bar{D}c \\
& & & w, z & \geq & 0, & &
\end{array}
$$

or equivalently,

$$
(D_{wz}) \qquad \begin{array}{lrcrcl}
\min & c_W^T w & + & c_Z^T z & & \\
\text{s.t.} & -\;Q_2^T \bar{D} w & + & Q_2^T \bar{D} z & = & Q_2^T \bar{D} c \\
 & w, z & \geq & 0, & &
\end{array}
$$

with

$$
\begin{aligned}
c_W &= u - \bar{D} Q_1 R^{-T} b && (23)\\
c_Z &= \bar{D} Q_1 R^{-T} b && (24)\\
y &= R^{-1} Q_1^T \bar{D} c + R^{-1} Q_1^T \bar{D} w - R^{-1} Q_1^T \bar{D} z. && (25)
\end{aligned}
$$

The problem (D_{wz}) is a standard form linear programming problem, so the centering direction and affine direction are defined as in den Hertog and Roos (Ref. 12). Before deriving these directions, we state two equalities which will prove useful in the following:

$$
\begin{aligned}
Q_1 R^{-T} &= \bar{D} A^T (A \bar{D}^2 A^T)^{-1} && (26)\\
Q_2 Q_2^T &= P_{A\bar{D}}. && (27)
\end{aligned}
$$

These equalities follow directly from (22). Using (26), we can rewrite (25) as

$$
y = (A\bar{D}^2 A^T)^{-1} A \bar{D}^2 (c + w - z). \tag{28}
$$

The Affine Direction

The affine direction for the standard form problem (D_{wz}) is defined to be

$$
\begin{bmatrix} dw^{aff} \\ dz^{aff} \end{bmatrix} = -\begin{bmatrix} \bar{W} & 0 \\ 0 & \bar{Z} \end{bmatrix} P_{[-Q_2^T \bar{D} \bar{W} | Q_2^T \bar{D} \bar{Z}]} \begin{bmatrix} \bar{W} c_W \\ \bar{Z} c_Z \end{bmatrix}. \tag{29}
$$

Now, from equation (12),

$$[-Q_2^T\bar{D}\bar{W}|Q_2^T\bar{D}\bar{Z}][-Q_2^T\bar{D}\bar{W}|Q_2^T\bar{D}\bar{Z}]^T = Q_2^T\bar{D}^2(\bar{W}^2+\bar{Z}^2)Q_2 = Q_2^TQ_2 = I. \tag{30}$$

It follows that

$$\begin{aligned} P_{[-Q_2^T\bar{D}\bar{W}|Q_2^T\bar{D}\bar{Z}]} &= I - [-Q_2^T\bar{D}\bar{W}|Q_2^T\bar{D}\bar{Z}]^T[-Q_2^T\bar{D}\bar{W}|Q_2^T\bar{D}\bar{Z}] \\ &= I - [-\bar{W}|\bar{Z}]^T\bar{D}Q_2Q_2^T\bar{D}[-\bar{W}|\bar{Z}], \end{aligned} \tag{31}$$

so

$$\begin{aligned} \begin{bmatrix} dw^{aff} \\ dz^{aff} \end{bmatrix} &= -\begin{bmatrix} \bar{W}^2c_W \\ \bar{Z}^2c_Z \end{bmatrix} + \begin{bmatrix} -\bar{W}^2\bar{D} \\ \bar{Z}^2\bar{D} \end{bmatrix} Q_2Q_2^T\bar{D}(-\bar{W}^2c_W + \bar{Z}^2c_Z) \\ &= -\begin{bmatrix} \bar{W}^2c_W \\ \bar{Z}^2c_Z \end{bmatrix} + \begin{bmatrix} \bar{W}^2\bar{D}Q_2Q_2^T\bar{D}\bar{W}^2u \\ -\bar{Z}^2\bar{D}Q_2Q_2^T\bar{D}\bar{W}^2u \end{bmatrix} \\ &\quad + \begin{bmatrix} -\bar{W}^2\bar{D} \\ \bar{Z}^2\bar{D} \end{bmatrix} Q_2Q_2^T\bar{D}^2(\bar{W}^2+\bar{Z}^2)Q_1R^{-T}b \\ &= -\begin{bmatrix} \bar{W}^2c_W - \bar{W}^2\bar{D}Q_2Q_2^T\bar{D}\bar{W}^2u \\ \bar{Z}^2c_Z + \bar{Z}^2\bar{D}Q_2Q_2^T\bar{D}\bar{W}^2u \end{bmatrix} \\ &= \begin{bmatrix} -\bar{Z}^2\bar{W}^2\bar{D}^2u + \bar{W}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}(b - A\bar{D}^2\bar{W}^2u) \\ -\bar{Z}^2\bar{W}^2\bar{D}^2u - \bar{Z}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}(b - A\bar{D}^2\bar{W}^2u) \end{bmatrix}, \end{aligned} \tag{32}$$

using equations (23), (24), (12), (26), (27), and the facts that $Q_2^TQ_1 = 0$ and $\bar{W}^4\bar{D}^2u - \bar{W}^2u = -\bar{Z}^2\bar{W}^2\bar{D}^2u$. (This last fact is a consequence of (12).) It then follows from equation (25) that

$$dy^{aff} = R^{-1}Q_1^T\bar{D}dw^{aff} - R^{-1}Q_1^T\bar{D}dz^{aff}$$

$$= (A\bar{D}^2A^T)^{-1}A\bar{D}^2(dw^{aff} - dz^{aff})$$

$$= (A\bar{D}^2A^T)^{-1}(b - A\bar{D}^2\bar{W}^2u), \tag{33}$$

using equations (12) and (26). This direction $(dy^{aff}, dw^{aff}, dz^{aff})$ is exactly the direction used in Marsten et al. (Ref. 13).

From equation (32), we obtain

$$\begin{aligned} dw^{aff} &= -\bar{Z}^2\bar{W}^2\bar{D}^2u + \bar{W}^2\bar{D}^2A^T dy^{aff} \\ &= -\bar{Z}^2\bar{W}^2\bar{D}^2u + \bar{W}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}b + \bar{W}^2\bar{D}P_{A\bar{D}}\bar{D}\bar{W}^2u - \bar{W}^4\bar{D}^2u \\ &= \bar{W}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}b + \bar{W}^2\bar{D}P_{A\bar{D}}\bar{D}\bar{W}^2u - \bar{W}^2u \\ &= -\bar{W}^2(v^1 + u), \end{aligned}$$

using equations (12) and (13). Similarly,

$$\begin{aligned} dz^{aff} &= -\bar{Z}^2\bar{W}^2\bar{D}^2u - \bar{Z}^2\bar{D}^2A^T dy^{aff} \\ &= -\bar{Z}^2\bar{W}^2\bar{D}^2u - \bar{Z}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}b - \bar{Z}^2\bar{D}P_{A\bar{D}}\bar{D}\bar{W}^2u + \bar{Z}^2\bar{W}^2\bar{D}^2u \\ &= -\bar{Z}^2\bar{D}^2A^T(A\bar{D}^2A^T)^{-1}b - \bar{Z}^2\bar{D}P_{A\bar{D}}\bar{D}\bar{W}^2u \\ &= \bar{Z}^2v^1, \end{aligned}$$

as stated in Section 4, equations (17) and (18).

The Centering Direction

The centering direction for the standard form problem (D_{wz}) is

$$\begin{bmatrix} dw^{cent} \\ dz^{cent} \end{bmatrix} = \begin{bmatrix} \bar{W} & 0 \\ 0 & \bar{Z} \end{bmatrix} P_{[-Q_2^T\bar{D}\bar{W} | Q_2^T\bar{D}\bar{Z}]} \begin{bmatrix} e \\ e \end{bmatrix}. \tag{34}$$

From equation (31), we obtain

$$\begin{aligned}
\begin{bmatrix} dw^{cent} \\ dz^{cent} \end{bmatrix} &= \begin{bmatrix} \bar{w} \\ \bar{z} \end{bmatrix} - \begin{bmatrix} -\bar{W}^2\bar{D}Q_2 \\ \bar{Z}^2\bar{D}Q_2 \end{bmatrix} (-Q_2^T\bar{D}\bar{w} + Q_2^T\bar{D}\bar{z}) \\
&= \begin{bmatrix} \bar{w} \\ \bar{z} \end{bmatrix} - \begin{bmatrix} -\bar{W}^2\bar{D}Q_2 \\ \bar{Z}^2\bar{D}Q_2 \end{bmatrix} Q_2^T\bar{D}c \\
&= \begin{bmatrix} \bar{w} + \bar{W}^2\bar{D}P_{A\bar{D}}\bar{D}c \\ \bar{z} - \bar{Z}^2\bar{D}P_{A\bar{D}}\bar{D}c \end{bmatrix} \\
&= \begin{bmatrix} \bar{w} + \bar{W}^2v^2 \\ \bar{z} - \bar{Z}^2v^2 \end{bmatrix},
\end{aligned}$$

using equations (27), (14), and the fact that $(\bar{w}, \bar{z})$ is feasible in the problem (D_{wz}).

It follows from equation (25) that

$$\begin{aligned}
dy^{cent} &= R^{-1}Q_1^T\bar{D}dw^{cent} - R^{-1}Q_1^T\bar{D}dz^{cent} \\
&= (A\bar{D}^2A^T)^{-1}A\bar{D}^2(\bar{w} - \bar{z}) + R^{-1}Q_1^T\bar{D}^2(\bar{W}^2 + \bar{Z}^2)Q_2Q_2^T\bar{D}c \\
&= (A\bar{D}^2A^T)^{-1}A\bar{D}^2(\bar{w} - \bar{z}), \qquad (35)
\end{aligned}$$

using equation (12) and the fact that $Q_1^TQ_2 = 0$. This direction is as stated in (Ref. 14).

5.3. Updating the Primal Solution x

A lower bound on the optimal value of (D_{wz}) can be obtained by solving the problem

$$(P(\mu)) \qquad \begin{aligned}
\max \quad & \mu \\
\text{s.t.} \quad & \mu(w - dw^{cent}) \le -dw^{aff} \\
& \mu(z - dz^{cent}) \le -dz^{aff}
\end{aligned}$$

Using equations (17), (18), (20), and (21) and dividing the constraints by $\bar{W}^2$ and $\bar{Z}^2$, we obtain the equivalent problem

$$
\begin{array}{lll}
(P(\lambda)) & \min & \lambda \\
 & \text{s.t.} & v^1 \;\leq\; \lambda v^2 \;\leq\; u + v^1,
\end{array}
$$

with v^1 and v^2 as in (13) and (14) respectively. As noted in Section 4, we can define

$$x(\lambda) = -v^1 + \lambda v^2. \tag{36}$$

This point is feasible in (P_U) if and only if λ is feasible in $(P(\lambda))$, since $Av^1 = -b$ and $Av^2 = 0$. Hence, the optimal value of $\bar{\lambda}$ leads to the upper bound $-c^T v^1 + \bar{\lambda} c^T v^2$ on the optimal value of (D_U). Since $c^T v^2 > 0$, minimizing $c^T x(\lambda)$ is equivalent to minimizing λ.

It may happen that the optimal value $\bar{\lambda}$ for $(P(\lambda))$ leads to a point $x(\lambda)$ which has worse objective value than the current upper bound $\bar{B}$. This happens if $c^T x(\bar{\lambda}) > \bar{B}$, that is, if

$$\bar{\lambda} > \frac{\bar{B} + c^T v^1}{c^T v^2}. \tag{37}$$

Hence, the "incumbent" value of λ is the ratio $\frac{\bar{B}+c^T v^1}{c^T v^2}$, so it only makes sense to update λ if the optimal value is smaller than this ratio. This explains Step 2 of the iteration given in Section 4.

5.4. The Primal Problem Corresponding to (D_{wz}).

The primal problem obtained by taking the dual of (D_{wz}) is

$$\begin{array}{lrcl} \min & -c^T \bar{D} Q_2 \pi & & \\ \text{s.t.} & -\bar{D} Q_2 \pi & \leq & c_W \\ & \bar{D} Q_2 \pi & \leq & c_Z \end{array}$$

or equivalently

$$(P_{wz}) \qquad \begin{array}{lrcl} \min & -c^T \bar{D} Q_2 \pi & & \\ \text{s.t.} & \bar{D} Q_1 R^{-T} b - \bar{D} Q_2 \pi & \leq & u \\ & \bar{D} Q_1 R^{-T} b - \bar{D} Q_2 \pi & \geq & 0 \end{array}$$

Of course, this is exactly equivalent to (P_U), with $x = \bar{D}Q_1R^{-T}b - \bar{D}Q_2\pi$, since $A\bar{D}Q_1R^{-T}b = b$ and the columns of $\bar{D}Q_2$ form a basis for the null space of A. It follows that eliminating y from $(D_U(\bar{D}))$ corresponds to taking $x = \bar{D}^2A^T(A\bar{D}^2A^T)^{-1}b$ and then adding to it some quantity in the null space of A.

References

1. GONZAGA, C. C., *On Lower Bound Updates in Primal Potential Reduction Methods for Linear Programming*, Mathematical Programming, Vol. 52, pp. 415–428, 1991.

2. TODD, M. J., and BURRELL, B. P., *An Extension of Karmarkar's Algorithm for Linear Programming Using Dual Variables*, Algorithmica, Vol. 1, pp. 409–424, 1986.

3. TODD, M. J., *Improved Bounds and Containing Ellipsoids in Karmarkar's Linear Programming Algorithm*, Mathematics of Operations Research, Vol. 13, pp. 650–659, 1988.

4. KARMARKAR, N. K., *A New Polynomial-Time Algorithm for Linear Programming*, Combinatorica, Vol. 4, pp. 373–395, 1984.

5. DE GHELLINCK, G., and VIAL, J. P., *A Polynomial Newton Method for Linear Programming*, Algorithmica, Vol. 1, pp. 425–453, 1986.

6. YE, Y., and KOJIMA, M., *Recovering Optimal Dual Solutions in Karmarkar's Polynomial Algorithm for Linear Programming*, Mathematical Programming, Vol. 39, pp. 305–317, 1987.

7. YE, Y., *An $O(n^3L)$ Potential Reduction Algorithm for Linear Programming*, Mathematical Programming, Vol. 50, pp. 239–258, 1991.

8. FREUND, R. M., *Polynomial Algorithms for Linear Programming Based only on Primal Scaling and Projected Gradients of a Potential Function*, Mathematical Programming, Vol. 51, pp. 203–222, 1991.

9. GONZAGA, C. C., *Search Directions for Interior Linear-Programming Methods*, Algorithmica, Vol. 6, pp. 153–181, 1991.

10. GONZAGA, C. C., *Conical Projection Algorithms for Linear Programming*, Mathematical Programming, Vol. 43, pp. 151–174, 1989.

11. MITCHELL, J. E., and TODD, M. J., *On the Relationship Between the Search Directions in the Affine and Projective Variants of Karmarkar's Linear Programming Algorithm*, Contributions to Operations Research and Economics:

The Twentieth Anniversary of CORE, Edited by Bernard Cornet and Henry Tulkens, MIT Press, Cambridge, Massachusetts, pp. 237–250, 1989.

12. DEN HERTOG, D., and ROOS, C., *A Survey of Search Directions in Interior Point Methods for Linear Programming*, Mathematical Programming, Vol. 52, pp. 481–510, 1991.

13. MARSTEN, R. E., SALTZMANN, M. J., SHANNO, D. F., PIERCE, G. S., and BALLINTIJN, J. F., *Implementation of a Dual Affine Interior Point Algorithm for Linear Programming*, ORSA Journal on Computing, Vol. 1, pp. 287–297, 1989.

14. MARSTEN, R. E., *OB1 User Documentation*, Technical report, Department of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 1990.

15. GAY, D. M., *A Variant of Karmarkar's Linear Programming Algorithm for Problems in Standard Form*, Mathematical Programming, Vol. 37, pp. 81–90, 1987.

16. MITCHELL, J. E., and TODD, M. J., *A Variant of Karmarkar's Linear Programming Algorithm for Problems with some Unrestricted Variables*, SIAM Journal on Matrix Analysis and Applications, Vol. 10, pp. 30–38, 1989.

References

[1] C. C. Gonzaga. On lower bound updates in primal potential reduction methods for linear programming. *Mathematical Programming*, 52(2):415–428, 1991.

[2] M. J. Todd and B. P. Burrell. An extension of Karmarkar's algorithm for linear programming using dual variables. *Algorithmica*, 1:409–424, 1986.

[3] M. J. Todd. Improved bounds and containing ellipsoids in Karmarkar's linear programming algorithm. *Mathematics of Operations Research*, 13(4):650–659, 1988.

[4] N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[5] G. de Ghellinck and J.-P. Vial. A polynomial Newton method for linear programming. *Algorithmica*, 1:425–453, 1986.

[6] Y. Ye and M. Kojima. Recovering optimal dual solutions in Karmarkar's polynomial algorithm for linear programming. *Mathematical Programming*, 39(3):305–317, 1987.

[7] Y. Ye. An $O(n^3L)$ potential reduction algorithm for linear programming. *Mathematical Programming*, 50(2):239–258, 1991.

[8] R. M. Freund. Polynomial algorithms for linear programming based only on primal scaling and projected gradients of a potential function. *Mathematical Programming*, 51:203–222, 1991.

[9] C. C. Gonzaga. Search directions for interior linear-programming methods. *Algorithmica*, 6:153–181, 1991.

[10] C. C. Gonzaga. Conical projection algorithms for linear programming. *Mathematical Programming*, 43:151–174, 1989.

[11] J. E. Mitchell and M. J. Todd. On the relationship between the search directions in the affine and projective variants of Karmarkar's linear programming algorithm. In Bernard Cornet and Henry Tulkens, editors, *Contributions to Operations Research and Economics: The Twentieth Anniversary of CORE*, pages 237–250. MIT Press, Cambridge, 1989.

[12] D. den Hertog and C. Roos. A survey of search directions in interior point methods for linear programming. *Mathematical Programming*, 52(2):481–510, 1991.

[13] R. E. Marsten, M. J. Saltzmann, D. F. Shanno, G. S. Pierce, and J. F. Ballintijn. Implementation of a dual affine interior point algorithm for linear programming. *ORSA Journal on Computing*, 1(4):287–297, 1989.

[14] R. E. Marsten. Ob1 user documentation. Technical report, Department of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 1990.

[15] D. M. Gay. A variant of Karmarkar's linear programming algorithm for problems in standard form. *Mathematical Programming*, 37:81–90, 1987.

[16] J. E. Mitchell and M. J. Todd. A variant of Karmarkar's linear programming algorithm for problems with some unrestricted variables. *SIAM Journal on Matrix Analysis and Applications*, 10(1):30–38, 1989.