A Global Error Bound for Quadratic Perturbation of Linear Programs

M. Ç. PINAR*
Service de Mathématiques de la Gestion
Université Libre de Bruxelles
Bd. du Triomphe CP 210/01
Campus de la Plaine
1050 Bruxelles, Belgium
mpinar@smg.ulb.ac.be

(Received March 1999; revised and accepted December 2000)

Abstract—We prove a global error bound result on the quadratic perturbation of linear programs. The error bound is stated in terms of function values. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords—Linear programming, Quadratic perturbation, Quadratic penalty functions, Global error bound.

The quadratic perturbation of linear programs is intimately related to the quadratic penalty functions applied to the linear program. More precisely, the quadratic perturbation is obtained in the primal problem if a quadratic penalty function is applied to the dual. The above observation was made and pursued in a series of papers by Mangasarian [1,2] and Li [3]. The application of quadratic penalty functions to linear programs was also studied by Pinar [4]. In a recent paper [5], Tseng derived a local error bound result for perturbation of linear programs. In the present note, we give a simple, global error bound result in Theorem 1 for the quadratic perturbation of linear programs. The result is inspired by early work of Güler [6] on the global convergence estimates of augmented Lagrangian algorithms on linear programs. It is given in terms of function values.

Consider the linear program

\[
\min_x \{ c^T x \mid Ax = b, x \geq 0 \} \tag{1}
\]

with its dual

\[
\max_y \{ -b^T y \mid A^T y + c \geq 0 \}. \tag{2}
\]
We use the quadratic penalty function
\[
\min_y \Phi(y, t) \equiv tb^Ty + \frac{1}{2}r^T(y)W(y)r(y),
\]  
where \( t \) is a positive scalar, \( W(y) \) is a diagonal matrix with diagonal entries defined as
\[
W_{ii}(y) = \begin{cases} 
0, & \text{if } r_i(y) \geq 0, \\
1, & \text{if } r_i(y) < 0,
\end{cases}
\]
and \( r(y) = A^Ty + c. \) The dual of this problem is precisely the perturbation problem
\[
\min_x \left\{ c^Tx + \frac{t}{2}x^Tx \mid Ax = b, x \geq 0 \right\}.
\]  
A solution \( y_t \) of the quadratic penalty problem (3) satisfies the following identity:
\[
(AW(y_t)A^T)y_t = -AW(y_t)c - tb.
\]  
Now, it was shown in [4] that
\[
W(y_t)r(y_t)W(y_t)(A^Ty_t + c)
\]
solves the perturbation problem (4) if \( y_t \) solves the quadratic penalty problem (3), and that \( W(y_t) \) is constant for any \( y_t \) which is a minimizer of \( \Phi(., t) \). Pinar [4] also shows that \( W(y_t) \) behaves as a piecewise linear function of \( t \) and there exists \( t^* > 0 \) such that \( W(y_t) \) remains constant for \( 0 < t \leq t^* \).

The main result of this note is the following theorem.

**Theorem 1.** Let \( t^* > 0 \) be such that \( W(y_t) \) remains constant for \( 0 < t \leq t^* \). For any \( t > 0 \) such that \( t > t^* \), the following bound holds:
\[
c^Tx_t - \omega^* = O\left(\frac{1}{t^*} - \frac{1}{t}\right),
\]  
where \( x_t \) solves the perturbation problem (4), and \( \omega^* \) is the optimal value of problem (1).

We will give the proof of the theorem after establishing some useful facts. We use \( N(B) \) and \( R(B) \) to denote the null space and range of a matrix \( B \), respectively.

**Lemma 1.** If the system \( Bx = b \) is consistent, then \( b \in R(BB^T) \).

**Proof.** Consider the problem \( \min \{\|x\|^2 : Bx = b\} \). The optimal solution satisfies \( x^* = B^Ty \) for some \( y \), and thus, \( b = Bx^* = BB^Ty \in R(BB^T) \).

Incidentally, Lemma 1 proves that \( R(B) = R(BB^T) \) for any matrix \( B \).

**Lemma 2.** If \( u = v - w \) such that \( v \) and \( w \) are orthogonal, then \( u^Tw = -\|w\|^2 \).

The proof of this lemma is trivial, and is therefore omitted. The following lemma is a standard result in penalty methods, which we include for completeness.

**Lemma 3.** \( c^Tx_t \) is decreasing, that is if \( 0 < t_1 < t_2 \), then \( c^Tx_{t_1} \leq c^Tx_{t_2} \).

**Proof.** Suppose \( t_2 > t_1 > 0 \). One has
\[
c^T x_{t_2} + \frac{t_2}{2}\|x_{t_2}\|^2 \leq c^T x_{t_1} + \frac{t_2}{2}\|x_{t_1}\|^2,
\]
\[
c^T x_{t_1} + \frac{t_1}{2}\|x_{t_1}\|^2 \leq c^T x_{t_2} + \frac{t_1}{2}\|x_{t_2}\|^2.
\]
Divide the inequalities by $1/t_2$ and $1/t_1$, respectively. The proof is completed by adding the resulting inequalities and simplifying the results.

Now, equipped with these facts, we can give the proof of Theorem 1. We can write $W_t = W(y_t)$, without ambiguity. Since $x_t = -W_t(A^y_t + c)/t$ and $A x_t = b$,
\[ AW_t (A^y_t + c) = -tb. \]
Thus, $b \in R(AW_t)$, and Lemma 1 implies that $AW_t(AW_t)^T d = b$ for some $d$. Since $W_t^2 = W_t$, we have $AW_tA^T d = b$. Substituting this in equation (7) and setting
\[ \tilde{y}_t = y_t + t d \]
gives
\[ AW_t (W_t A^T \tilde{y}_t + W_tC) = 0, \]
that is, $W_tA^T \tilde{y}_t + W_tC \in N(AW_t^T)$. Since $W_tA^T \tilde{y}_t \in R(W_tA^T) = N(AW_t)^\perp$, we see that
\[ W_tC = (W_tA^T \tilde{y}_t + W_tC) - W_tA^T \tilde{y}_t \]
is an orthogonal decomposition of $W_tC$ onto $N(AW_t)$ and its orthogonal complement $R(W_tA^T)$.

Suppose now that $t_2 > t_1 > 0$ such that $W_{t_2} = W_{t_1} := W$. Then, using the notation $u^T v = (u, v)$, we have
\[
0 \leq \langle c, x_{t_2} - x_{t_1} \rangle = -\frac{1}{t_2} \langle c, W (A^T y_{t_2} + c) \rangle + \frac{1}{t_1} \langle c, W (A^T y_{t_1} + c) \rangle \\
= \left(\frac{1}{t_1} - \frac{1}{t_2}\right) \|WC\|^2 - \frac{1}{t_2} \langle c, W A^T y_{t_2} \rangle + \frac{1}{t_1} \langle c, W A^T y_{t_1} \rangle .
\]
Now, we have
\[
-\frac{1}{t_2} \langle Wc, W A^T y_{t_2} \rangle + \frac{1}{t_1} \langle Wc, W A^T y_{t_1} \rangle = -\frac{1}{t_2} \langle Wc, W A^T \tilde{y}_{t_2} \rangle + \frac{1}{t_1} \langle Wc, W A^T \tilde{y}_{t_1} \rangle \\
= \frac{1}{t_2} \|WA^T \tilde{y}_{t_2}\|^2 - \frac{1}{t_1} \|WA^T \tilde{y}_{t_1}\|^2 ,
\]
where the first equality comes from (8), and the second one from (10) and Lemma 2.

Note that (9) implies that $u := \tilde{y}_{t_2} - \tilde{y}_{t_1}$ satisfies $WA A^T u = 0$. But, then $0 = u^T AWA^T u = \|WA A^T u\|^2$. Thus, $WA A^T u = 0$. Then, $WA A^T \tilde{y}_{t_2} = WA A^T \tilde{y}_{t_1}$. This shows that the quantity in (12) is nonpositive. Thus, (11) reduces to
\[ 0 \leq c^T x_{t_2} - c^T x_{t_1} \leq \left(\frac{1}{t_1} - \frac{1}{t_2}\right) \|c\|_2 .
\]
If $t_2^* > t_2 > t_1 > t_1^*$ where $t_2^*$ and $t_1^*$ are consecutive breakpoints, then $W_{t_2} = W_{t_1}$, and inequality (13) applies. Now, $x_t$ is a minimizer of the optimization problem (4). It is easy to verify that that $x_t$ is the projection of the vector $-c/t$ on the feasible set $F := \{x : Ax = b, x \geq 0\}$, that is, $x_t$ is the solution to the problem $\min \|x + (c/t)\| : x \in F$. The projection operator onto a convex set is nonexpansive, so that $\|x_{t_2} - x_{t_1}\| \leq (1/t_1 - 1/t_2) \|c\|$. This proves that $x_t$ is a continuous function of $t$ when $t > 0$. Consequently, we see that (13) also holds when $t_2$ and $t_1$ are replaced by $t_2^*$ and $t_1^*$, respectively. Since $c^T x_{t_2^*} = \omega^*$ (in fact $c^T x_t = \omega^*$ for all $t \in (0, t^*], [1,2,4]$), where $\omega^*$ is the optimal value of the original linear program, the proof is completed.

Interestingly Güler [6] first gives a global convergence rate estimate of $O(1/\sum_{i=0}^{k-1} \lambda_i)$ for the augmented Lagrangian algorithm, where $\lambda_i$ is the penalty parameter. Then he modifies the multiplier iteration and sharpens the bound to $O((1/\sum_{i=0}^{k-1} \sqrt{\lambda_i})^2)$. It is interesting that the bound we obtain in the theorem is also linear in the inverse of $t$. 
REFERENCES