A Global Error Bound for Quadratic Perturbation of Linear Programs

Mustafa Ç. Pınar*
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

December 1998
Revised November 1999, June 2000

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 Pınar [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 O. 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

$$\min_x \{c^T x | Ax = b, x \geq 0\} \tag{1}$$

with its dual

$$\max_y \{-b^T y | A^T y + c \geq 0\}. \tag{2}$$

We use the quadratic penalty function:

$$\min_y \ \Phi(y,t) \equiv t b^T y + \frac{1}{2} r^T(y) W(y) r(y) \tag{3}$$

*On leave from Department of Industrial Engineering, Bilkent University, 06533 Ankara, Turkey 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 $r(y) = A^T y + c$. The dual of this problem is precisely the perturbation problem

$$\min_x \{c^T x + \frac{t}{2} x^T x | Ax = b, x \geq 0\}. \tag{4}$$

A solution y_t of the quadratic penalty problem (3) satisfies the following identity:

$$(AW(y_t)A^T)y_t = -AW(y_t)c - tb. \tag{5}$$

Now, it was shown in [4] that

$$x_t \equiv -W(y_t)r(y_t)/t = -W(y_t)(A^T y_t + c)/t$$

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 $\Phi(.,t)$. Pınar [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*

$$c^T x_t - \omega^* = \mathcal{O}(\frac{1}{t^*} - \frac{1}{t}), \tag{6}$$

where x_t solves the perturbation problem (4), and ω^* 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 $x^* = B^T y$ for some y, and thus $b = Bx^* = BB^T y \in R(BB^T)$. ∎

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 $u^T w = -\|w\|^2$.*

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 *$c^T x_t$ is decreasing, that is if $0 < t_1 < t_2$, then $c^T x_{t_1} \leq c^T x_{t_2}$.*

Proof: Suppose $t_2 > t_1 > 0$. One has

$$c^T x_{t_2} + \frac{t_2}{2}\|x_{t_2}\|^2 \le c^T x_{t_1} + \frac{t_2}{2}\|x_{t_1}\|^2,$$

$$c^T x_{t_1} + \frac{t_1}{2}\|x_{t_1}\|^2 \le c^T x_{t_2} + \frac{t_1}{2}\|x_{t_2}\|^2.$$

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 $x_t = -W_t(A^T y_t + c)/t$, and $Ax_t = b$,

$$AW_t(A^T y_t + c) = -tb. \tag{7}$$

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 $AW_tA^Td = b$. Substituting this in equation (7) and setting

$$\bar{y}_t = y_t + td \tag{8}$$

gives

$$AW_t(W_tA^T\bar{y}_t + W_tc) = 0, \tag{9}$$

that is, $W_tA^T\bar{y}_t + W_tc \in N(AW^T)$. Since $W_tA^T\bar{y}_t \in R(W_tA^T) = N(AW_t)^\perp$, we see that

$$W_tc = (W_tA^T\bar{y}_t + W_tc) - W_tA^T\bar{y}_t \tag{10}$$

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 $u^Tv = \langle u, v\rangle$, we have

$$\begin{aligned} 0 \ \le\ \langle c, x_{t_2} - x_{t_1}\rangle &= -\frac{1}{t_2}\langle c, W(A^T y_{t_2} + c)\rangle + \frac{1}{t_1}\langle c, W(A^T y_{t_1} + c)\rangle \\ &= (\frac{1}{t_1} - \frac{1}{t_2})\|Wc\|^2 - \frac{1}{t_2}\langle c, WA^T y_{t_2}\rangle + \frac{1}{t_1}\langle c, WA^T y_{t_1}\rangle. \end{aligned} \tag{11}$$

Now,

$$\begin{aligned} -\frac{1}{t_2}\langle Wc, WA^T y_{t_2}\rangle + \frac{1}{t_1}\langle Wc, WA^T y_{t_1})\rangle &= -\frac{1}{t_2}\langle Wc, WA^T\bar{y}_{t_2}\rangle + \frac{1}{t_1}\langle Wc, WA^T\bar{y}_{t_1}\rangle \\ &= \frac{1}{t_2}\|WA^T\bar{y}_{t_2}\|^2 - \frac{1}{t_1}\|WA^T\bar{y}_{t_1}\|^2, \end{aligned} \tag{12}$$

where the first equality from (8), and the second one from (10) and Lemma 2.

Note that (9) implies that $u := \bar{y}_{t_2} - \bar{y}_{t_1}$ satisfies $AWA^T u = 0$. But, then $0 = u^T AWA^T u = \|WA^T u\|^2$. Thus, $WA^T u = 0$. Thus, $WA^T \bar{y}_{t_2} = WA^T \bar{y}_{t_1}$. This shows that the quantity in (12) is non-positive. Thus (11) reduces to

$$0 \leq c^T x_{t_2} - c^T x_{t_1} \leq (\frac{1}{t_1} - \frac{1}{t_2})\|c\|_2. \tag{13}$$

If $t_2^* > t_2 > t_1 > t_1^*$ where t_2^* and t_1^* are consecutive breakpoints, the $W_{t_2} = W_{t_1}$, 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 non-expansive, 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 $c^T x_{t^*} = \omega^*$ (in fact $c^T x_t = \omega^*$ for all $t \in (0, t^*]$, [1, 2, 4]), where ω^* is the optimal value of the original linear program, the proof is completed.

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

Acknowledgments. The paper benefited from careful reviews of two anonymous referees. I am particularly grateful to the referee who suggested a simplification of the proofs which I incorporated into the final version. This led to a more concise presentation.

References

[1] O.L. Mangasarian, 1981, Iterative Solution of Linear Programs, *SIAM J. Numer. Anal.* 18, 606-614.

[2] O.L. Mangasarian, 1984, Normal Solution of Linear Programs, *Math. Prog. Study* 22, 206-216.

[3] W.Li, 1993, The Sharp Lipschitz Constants for Feasible and Optimal Solutions of a Perturbed Linear Program, *Linear Algebra and Appl.*, 187, 15–40.

[4] M.Ç. Pınar, 1997, Piecewise Linear Pathways to the Optimal Solution Set of Linear Programs, *J. Optim. Theory and Appl.*, 93, 619-634.

[5] P. Tseng, 1999, Convergence and Error Bound for Perturbation of Linear Programs, *Computational Optim. and Appl.* 13, 221–230.

[6] O. Güler, 1992, Augmented Lagrangian Algorithms for Linear Programs, *J. Optim. Theory and Appl.* 75, 445-470