INFORMATICA, 1991, Vol.2, No.3, 352–366

PRACTICAL ERROR BOUNDS FOR A CLASS OF QUADRATIC PROGRAMMING PROBLEMS

Abraham CHARNES

Center for Cybernetic Studies, University of Texas at Austin
Austin, TX 78712–1177, U.S.A.

John SEMPLE

College of Business, University of Texas at Arlington
Arlington, TX 76019, U.S.A.

Abstract. Error bounds are developed for a class of quadratic programming problems. The absolute error between an approximate feasible solution, generated via a dual formulation, and the true optimal solution is measured. Furthermore, these error bounds involve considerably less work computationally than existing estimates.

Key words: quadratic programming, nonlinear programming, error estimates.

1. Introduction. Quadratic programming has long been the cornerstone for many numerical techniques in nonlinear programming. Typically, these methods use a quadratic function to approximate the objective function and linear equalities and (or) inequalities to approximate the constraint functions, all in some neighborhood of a specified feasible point. The resulting quadratic programming problem is solved and the optimal solution (or approximation thereof) becomes the point from which a new round of approximations is made and a new quadratic programming problem is formulated. In addition, quadratic programming merits attention in its own right as it arises naturally in such fields as engineering, economics, and game theory.

Increasingly, though, the topic has been treated as a special case of the linearly constrained variational inequality problem:
Find u^* such that

$$(u - u^*)^T f(u^*) \leqq 0 \quad \text{for all } u \text{ in } K \tag{1}$$

where K is polyhedral set

$$K = \{u : \ Cu \geqq d, \ Au = b\} \tag{2}$$

and C, A, d, and b are specified matrices and vectors of appropriate dimensions. Of special interest has been the development of error bounds for approximate solutions, x, to the true optimal solution, x^*, of (1). Particular attention has been paid to the case where $f(x)$ is a linear function

$$f(x) = Mx + c$$

with M a positive definite (p.d.) matrix. Recently, Pang (1987) provided a survey of known error estimates along with a number of extensions. These follow the general format

$$||x - x^*|| \leqq \rho \cdot r(x)$$

where ρ is a constant independent of x (but dependent on the data) and $r(x)$ is the generalized residual, a quantity depending on the value of x and the type of measure employed. The norm, unless otherwise stated, is the Euclidean norm. A common drawback to these bounds is that at least one of the two quantities is difficult to calculate (and any simplification in one seems to make the other more difficult).

The purpose of this paper is to derive a set of error estimates for a special case of the quadratic programming problem:

$$\begin{aligned} \min\ 1/2\ x^T H x + v^T x& \\ \text{s.t.} \qquad Ax &= b \\ x &\geqq 0 \end{aligned} \tag{3}$$

where H is positive definite, and $A \in R^{mxn}$ is of full row rank. We will sometimes refer to this as the primal (quadratic programming) problem.

In relationship to Pang's work we remark that the results presented here are a hybrid of those termed linear programming measures (of error) and dual measures (of error) in Pang (1987). A Corollary to our Theorem 3.3 results in one similar to Pang's Theorem 4.1. However, our bounds are relatively easy to compute and thus can be employed not only for a posteriori error analysis but also as a test for termination of the particular quadratic programming algorithm being used.

We have divided the remainder of this paper into two sections. In section 2 we develop a dual formulation to (3) using conjugate functions. This approach allows us to estimate the current duality gap (denoted by Δ) and consequently bounds the improvement we can expect in our dual objective functional. As the dual formulation is solved it generates approximate optimal solutions to the primal problem (given by the gradient of the dual objective functional). In section 3 we bound the error between these generated solutions (denoted by x) and the true optimal solution to the primal problem (denoted by x^*) as a function of the estimated duality gap (Δ).

2. The Dual Formulation. In this section we derive the dual formulation to system (3). The theory of duality in quadratic programming is not particularly new. Dorn (1960) and Lemke (1960) are among the first to address this topic.

Both formulations in these works are different from the dual in this section, yet both can be used to derive it. Lost in the derivation, however, is the notion of the duality gap which is essential to our analysis. Therefore, we will use the general theory of duality as set forth in Avriel (1976). To employ this theory we introduce perturbation variables w_1, w_2 and ν, and define

$$\phi(x, w_1, w_2, \nu) = \begin{cases} 1/2\, x^T H x + v^T x & \text{if } \begin{array}{r} Ax - b \geqq w_1 \\ -Ax + b \geqq w_2 \\ x \geqq \nu \end{array} \\ +\infty & \text{otherwise.} \end{cases}$$

The conjugate function, ϕ^*, becomes

$$\phi^*(\zeta, \lambda_1, \lambda_2, u) = \begin{cases} \sup_x [\zeta^T x + \lambda_1^T (Ax - b) \\ \quad +\lambda_2^T(-Ax + b) - u^T x \\ \quad -(1/2x^T H x + v^T x)] & \text{for } \lambda_1, \lambda_2, \\ & u \geqq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Letting $y = \lambda_1 - \lambda_2$ and substituting $\zeta = 0$ (in anticipation of the Weak Duality Theorem), we get

$$\phi^*(0, y, u) = \begin{cases} \sup_x [y^T(Ax - b) + u^T x \\ \quad -(1/2x^T H x + v^T x)] & \text{for } u \geqq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

The supremum over x can now be removed by performing the necessary maximization and inserting the optimal solution, yielding

$$\phi^*(o, y, u) = \begin{cases} (1/2y^T A H^{-1} A y \\ \quad -y^T A H^{-1}(u - v) \\ \quad +1/2y^T A H^{-1} A^T y & \text{for } u \geqq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

By the Weak Duality Theorem we have

$$\inf_{x} \phi(x,0,0,0) \geqq \sup_{(y,u)} -\phi^*(o,y,u).$$

Since y is free, it too can be eliminated by performing the required maximization of $-\phi^*$. One can easily show that the correct (optimal) choice of y (in terms of $u \geqq 0$) is

$$y^* = (AH^{-1}A^T)^{-1}(b - AH^{-1}(u-v)).$$

Substituting this result and acknowledging the dependence of ϕ^* on u yields

$$\phi^*(u) = \begin{cases} (1/2u^TPu + c^Tu + k & \text{for } u \geqq 0 \\ +\infty & \text{otherwise,} \end{cases} \tag{4A}$$

where

$$P = H^{-1} - H^{-1}A^T(AH^{-1}A^T)^{-1}AH^{-1} \tag{4B}$$

$$c = -Pv + H^{-1}A^T(AH^{-1}A^T)^{-1}b \tag{4C}$$

$$k = 1/2v^TPv - b^T(AH^{-1}A^T)^{-1}AH^{-1}v - 1/2b^T(AH^{-1}A^T)^{-1}b. \tag{4D}$$

The weak duality inequality becomes

$$\inf_{\substack{Ax=b \\ x\geqq 0}} [1/2x^THx + v^Tx] \geqq \sup_{u\geqq 0} -[1/2u^TPu + c^Tu + k]$$

and by virtue of this inequality, any x satisfying the primal constraints $[Ax = b,\ x \geqq 0]$ and any $u \geqq 0$ provides a "working duality gap", i.e., an upper bound on the possible improvement in either minimizing the left hand side (starting from x) or maximizing the right hand side (starting from u). As such, we define

$$\Delta(x,u) = 1/2x^THx + v^Tx + 1/2u^TPu + c^Tu + k \tag{5}$$

for each (x, u) satisfying $Ax = b$, $x \geqq 0$, and $u \geqq 0$. Moreover, we define the dual quadratic program to (3) as

$$\begin{array}{ll} \inf & 1/2u^T Pu + c^T u + k \\ \text{s.t.} & u \geqq 0. \end{array} \tag{6}$$

Clearly, $\Delta(x, u) = 0$ means x and u are optimal to the primal and dual problems, respectively.

The next proposition is an elementary exercise in linear algebra whose proof is left to the reader.

PROPOSITION 2.1. The following relationships hold

1. $AP = 0$
2. $||P|| \leqq ||H^{-1}||$
3. $PHP = P$
4. $A(Pu + c) = b$
5. $c^T HP = -v^T P$
6. $1/2c^T Hc + v^T c = -k.$

If we assume that the primal problem (3) has a feasible solution, then the dual problem (6) is bounded below and it is not hard to show that there exists an optimal solution u^*. This u^* also solves the (equivalent) linear complementarity problem: Find u such that

$$Pu + c \geqq 0 \qquad u \geqq 0 \qquad u^T(Pu + c) = 0.$$

By virtue of proposition 2.1 (part 4), $x^* = Pu^* + c$ is a feasible point to the primal problem and a direct calculation shows $\Delta(x^*, u^*) = 0$ demonstrating the optimality of x^* as well. As a consequence of the same proposition (part 1), we observe that the function $Pu + c$ is not strongly monotone over the nonnegative orthant, as evidenced by the nontrivial nullspace of P.

For the remainder of this paper we will make the following assumption:

(A) There exists an interior point $\hat{x}$ to (3): $A\hat{x} = b$, $\hat{x} > 0$.

We remark that assumption (A) is equivalent to either of the following assumptions:

(A') The set of optimal solutions to the dual problem (6) is bounded or

(A'') There exists a $u \geqq 0$ such that $Pu + c > 0$.

The details can be found in Semple (1990). Observe that (A'') implies (A') as previously shown by Mangasarian and McLinden (1985).

3. Error Estimates. A number of reasons warrant computing solutions to (3) via the dual formulation (6). The constraint set consists only of nonnegativity constraints allowing for a considerable reduction in the computational work load. Moreover, the dual problem preserves the number of variables, unlike many existing dualities, and has accessible feasible points.

Recall from section 2 that an optimal dual solution u^* induces the corresponding (unique) optimal solution to the primal via the formula $x^* = Pu^* + c$. As such, we want to know how close the estimate $x = Pu + c$ is to the primal optimal solution. Observe that the estimate satisfies $Ax = b$ by Proposition 2.1 (part 4) but not necessarily the nonnegativity condition, $x \geqq 0$. A number of quadratic programming algorithms applied to (6) require nonnegativity of $Pu + c$, particularly those based on adding a logarithmic barrier function (see Kojima, Mizuno and Yoshise (1989) or Monteiro and Adler (1989)). Others based on projecting (or truncating components of) the gradient do not. Of the former type, the two cited references actually enforce *positivity* which, as remarked in section 2, requires that assumption (A) hold. Error estimates in this case are handled easily by Corollary 3.4.

To start, we will derive a bound on $||x - x^*||$ (again, $x = Pu + c$) based on the true duality gap, i.e., the difference

$$1/2u^TPu + c^Tu - \left[1/2(u^*)^TPu^* + c^Tu^*\right].$$

Lemma 3.1. *Let u^* denote an optimal solution to (6), namely*

$$\min \quad 1/2u^TPu + c^Tu \quad \text{subject to} \quad u \geqq 0,$$

with P and c defined as in $(4B, 4C)$. At u suppose an upper bound δ has been acquired such that

$$1/2u^TPu + c^Tu - \left[1/2(u^*)^TPu^* + c^Tu^*\right] \leqq \delta \tag{7}$$

(as, for example, when one has an upper bound on the duality gap). Then

$$||P(u^* - u)|| \leqq \sqrt{2\lambda_m\delta}$$

where λ_m is the largest eigenvalue of the (p.s.d.) matrix P.

Proof. It is easy to show that

$$\begin{aligned} \delta &\geqq 1/2u^TPu + c^Tu - \left[1/2(u^*)^TPu^* + c^Tu^*\right] \\ &= -1/2(u - u^*)^TP(u - u^*) - (u^* - u)^T(Pu + c). \end{aligned} \tag{8}$$

Define $F(t)$ on [0, 1] by

$$F(t) = 1/2\big(u + t[u^* - u]\big)^TP\big(u + t[u^* - u]\big) + c^T\big(u + t[u^* - u]\big).$$

Observe that $u+t[u^*-u]$ is nonnegative for $t\varepsilon[0,1]$ and $F(t)$ is monotone decreasing on $[0,1]$. Thus $F'(t) \leqq 0$ on $[0,1]$ which, for $t=1$, yields the inequality

$$(u^* - u)^T P(u^* - u) \leqq -(u^* - u)^T (Pu + c). \tag{9}$$

The two inequalities (8) and (9) imply

$$(u^* - u)^T P(u^* - u) \leqq 2\delta. \tag{10}$$

Finally, since P is p.s.d. (and symmetric) we have

$$(u^* - u)^T P^T P(u^* - u) \leqq \lambda_m (u^* - u)^T P(u^* - u)$$

from which we conclude

$$||P(u^* - u)||^2 \leqq 2\delta\lambda_m.$$

Taking the square root of both sides of the inequality completes the proof.

COROLLARY 3.2. If $x = Pu+c$ is an approximate optimal solution to (3) (not necessarily nonnegative), then

$$||x - x^*|| \leqq \sqrt{2\lambda_m \delta}$$

where x^* is the (unique) optimal solution to (3) and δ satisfies (7).

Our focus now turns to obtaining good estimates of the δ in (7). The working duality gap $\Delta(x,u)$ is one such bound, but it requires that x be feasible to $Ax=b$, $x \geqq 0$. If x is not nonnegative (but satisfies $Ax=b$), then we can use a perturbation technique analogous to that in Fiacco and McCormick (1968, Thm. 29) or Robinson (1975) to create a nearby point which is. More precisely, suppose that an interior point $\widehat{x}$ has been generated (or as is the case in some problems, known in advance). Then for $x = Pu + c$, the perturbed point z given by

$$z = tx + (1-t)\widehat{x}$$

is both feasible and nonnegative provided $0 \leqq t \leqq \theta$ where

$$\theta = \min_i \left\{ \frac{\widehat{x}_i}{\widehat{x}_i - \min\{x_i, 0\}} \right\}. \tag{11}$$

Taking $t = \theta$ to define a particular z, and then using it to evaluate $\Delta(z, u)$ is the basis for the following theorem.

Theorem 3.3. *Suppose assumption (A) holds and that an interior point $\widehat{x}$ has been specified. Let x be defined by $x = Pu + c$ and let x^* be the (unique) optimal solution to the primal quadratic programming problem (3). Then there exists a constant vector ζ and constants ρ and τ, all depending on $\widehat{x}$, such that*

$$||x - x^*|| \leqq$$

$$\leqq \sqrt{2\lambda_m\{x^T u + (1-\theta)\big[\theta x^T\zeta + (1+\theta)k + (1-\theta)\rho + \tau\big]\}}$$

where θ is given by equation (11), k is the constant defined in $(4D)$, and λ_m is the largest eigenvalue of P.

Proof. In light of Lemma 3.1 it suffices to show that the quantity in brackets $\{\cdot\}$ is an upper bound similar to δ in (7). Define $z = \theta x + (1-\theta)\widehat{x}$ where θ is given (depending on $\widehat{x}$ and x) by (11). Then $\Delta(z, u)$, as given in (5), becomes

$$\begin{aligned}\Delta(z,u) = {} & 1/2\theta^2 u^T PHPu + \theta^2 c^T HPu + 1/2\theta^2 c^T Hc \\ & + \theta(1-\theta)\widehat{x}^T H(Pu + c) + 1/2(1-\theta)^2 \widehat{x}^T H\widehat{x} \\ & + \theta v^T(Pu + c) + (1-\theta)v^T\widehat{x} + 1/2u^T Pu + c^T u + k.\end{aligned}$$

Using the identities $PHP = P$, $c^T HP = -v^T P$, and $1/2c^T c = -k - v^T c$ from proposition 2.1 and regrouping:

$$\begin{aligned}\Delta(z,u) =&(1/2\theta^2 + 1/2)u^T Pu + c^T u \\ &+ (1-\theta)\big[\theta(Pu + c)^T(H\widehat{x} + v) + (1+\theta)k \\ &+ 1/2(1-\theta)\widehat{x}^T H\widehat{x} + \widehat{x}^T v\big].\end{aligned}$$

Finally, using the positive semi-definiteness of P and the definition of x, we obtain

$$\begin{aligned}\Delta(z,u) \leqq& u^T x + (1-\theta)\big[\theta x^T(H\widehat{x} + v) \\ &+ (1+\theta)k + (1-\theta)(1/2\widehat{x}^T Hx) + \widehat{x}^T v\big],\end{aligned}$$

the constants in brackets $\{\cdot\}$ now being evident if we take

$$\zeta = H\widehat{x} + v \qquad \rho = 1/2\widehat{x}^T H\widehat{x} \qquad \tau = \widehat{x}^T v. \tag{12}$$

COROLLARY 3.4. Suppose for a given $u \geqq 0$ $Pu + c \geqq 0$. Then for $x = Pu + c$

$$||x - x^*|| \leqq \sqrt{2\lambda_m x^T u}.$$

Proof: $\theta = 1$ if $Pu + c \geqq 0$, and the term under the root in Theorem 3.3 reduces accordingly.

Observe that no attempt has been made to fit the interior point $\widehat{x}$ to a particular value of $Pu + c$, nor have we tried to optimize our choice of θ for given $\widehat{x}$. For example, given $\widehat{x}$ and $Pu + c$ one might minimize the function

$$g(s) = (1-s)\big[s\, x^T(H\widehat{x}+v)+(1+s)k+(1-s)(1/2\widehat{x}^T H\widehat{x})+\widehat{x}^T v\big]$$

subject to the constraint $0 \leqq s \leqq \theta$, possibly improving the error bound. Additionally, in response to a candidate $Pu+c$, one might attempt to generate an interior point with larger positive compōnents corresponding to (and in scale with) the negative components of $Pu + c$, thus reducing the factor $1 - \theta$. However, without additional knowledge, a priori, of the particular approximate solutions $(Pu + c)$ to be encountered, it seems wise to find a "good" interior point (i.e., one with uniformly large positive entries) in an attempt to keep the factor $1 - \theta$ small. Such an interior point can be found by solving (either exactly or approximately) the linear program

$$\begin{aligned} & \text{Max } t \\ \text{s.t.} \quad & Ax = b \\ & x - te \geqq 0 \\ & x, t \geqq 0, \end{aligned} \tag{13}$$

where e is the vector of $1's$, $e^T = (1, 1, \ldots, 1)$. Observe that any feasible point $(\widehat{x}, t)$ to (13) with $t > 0$ yields a set of constants via (12) which may be used to estimate the absolute error between our current candidate $(Pu + c)$ and the true optimal solution (x^*).

Note that Theorem 3.3 (and Corollary 3.4) can be extended to incorporate the notion of the residual, i.e., the portion of $Pu + c$ which violates the nonnegativity condition. To measure the violation consider the l_p norm

$$||(Pu + c)^{-}||_p$$

where $(Pu + c)^{-}$ denotes the vector obtained by inserting $0'$s in the positive components of $Pu + c$. In particular, if the ∞-norm is chosen, and $(\widehat{x}, t)$ is a feasible solution to (13) with $t > 0$ then

$$\theta \geqq \min_i \left\{ \frac{\widehat{x}_i}{\widehat{x}_i + ||(Pu + c)^{-}||_\infty} \right\}$$

thus the factor $(1-\theta)$ in Theorem 3.3 satisfies

$$1-\theta \leqq \frac{1}{t}\,||(Pu+c)^{-}||_{\infty}.$$

Various extensions and alternate formulations of Theorem 3.3 and Corollary 3.4 are now obvious. We will not pursue these here.

We conclude with a remark about the computational efficiency of these error estimates. As previously mentioned, once an interior point to $Ax = b \quad x \geqq 0$ has been found, it may be used to bound the error between *any* candidate $(Pu + c)$ and x^*. The error bounds are, of course, more relevant to those approximate solutions which are nearly nonnegative (thus $1-\theta$ will be small) and nearly complementary (hence $x^T u$ will be small). Each of the constants in Theorem 3.3 can be computed (or bounded) with a minimal amount of work.

In contrast, error estimates which rely on the gap function (see section 4 of Pang (1987)) require the complete solution to a linear programming problem *for each approximate optimal solution considered.* The dual measures (see section 5 of Pang (1987)) also require a substantial amount of work. Here, one needs a constant defined by the maximum norm of a vector constrained to the surface of an ellipsoid (characterized by the data).

The potential of the error bounds in Theorem 3.3 can be realized with minimal computational work. They should be considered, as our title suggests, practical.

ACKNOWLEDGEMENT. The authors wish to thank Dr. Faizul Huq, Stephen Haag, and Pat Jaska (University of Texas at Arlington) for their helpful comments and suggestions.

REFERENCES

Avriel, M. (1976). *Nonlinear Programing, Analysis and Methods.* Prentice Hall, Inc., Englewood Cliffs, N.J.

Dorn, W.S. (1960). Duality in quadratic programming. *QAM,* **18**(2), 155–161.

Fiacco, A.V., and G.P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, New York.

Kojima, M., S.Mizuno and A.Yoshise (1989). A polynomial time algorithm for a class of linear complementarity problems. *Mathematical Programming,* **44**, 1–26.

Lemke, C.E. (1960). A method of solution for quadratic programs. *Management Science,* **8**, 442–453.

Mangasarian, O.L., and L.McLinden (1985). Simple bounds for solutions of monotone complementarity problems and convex programs . *Mathematical Programming,* **32**, 32–40.

Monteiro, R.D.C., and I.Adler (1989). Interior path following primal-dual algorithms. Part II: Convex Quadratic Programming. *Mathematical Programming,* **44**, 43–66.

Pang, J.S. (1987). A posteriori error bounds for the linearly-constrained variational inequality problem. *Mathematics of Operations Research,* **12**(3), 474–484.

Robinson, S.M. (1975). An application of error bounds for convex programming in a linear space. *SIAM Journal for Control and Optimization,* **13**(2), 271–273.

Semple, J. (1990). *On a Class of Dual Methods for Quadratic Programming.* Ph.D. thesis, Mathematics Department, The University of Texas at Austin, Austin, TX, December.

Received March 1991

A. Charnes is currently Professor of Mathematics, Computer Science, Management Science/Information Systems, Harbin Chair Emeritus, and Director of the Center for Cybernetic Studies. He is author or coauthor of over 450 research papers in mathematics, informatics, and managment science and six books translated into Russian, Chinese, Japanese, and Czechoslovakian. He is a Fellow of ORSA, The Econometric Society, and AAAS; cowinner of the von Neumann Theory Prize; Finalist for the Nobel Prize in Economics (1975); and lecturer for Hungary, USSR, China, Mexico Academies of Science and many European, Oriental, Latin American, Canadian and U.S. universities.

J. Semple is currently Visiting Professor of Information Systems and Management Science at The University of Texas at Arlington. He received his Ph.D. in mathematics from The University of Texas at Austin in December, 1990. He has authored or coauthored 5 research papers in mathematics and management science.