Università di Pisa
Dipartimento di Statistica e Matematica
Applicata all'Economia

Report n. 299

Global optimization of a generalized quadratic program

Riccardo Cambini and Claudio Sodini

Pisa, settembre 2007
- Stampato in proprio –

Via Cosimo Ridolfi, 10 - 56124 PISA - Tel. Segr. Amm. 050 2216231 Segr. Stud. 050 2216317 Fax 050 2216375
Cod. Fisc. 80003670504 - P. IVA 00286820501 - Web http://statmat.ec.unipi.it- E-mail: dipstat@ec.unipi.it

Global optimization of a generalized quadratic program

Riccardo Cambini and Claudio Sodini

Department of Statistics and Applied Mathematics
Faculty of Economics, University of Pisa
Via Cosimo Ridolfi 10, 56124 Pisa, ITALY
e-mail: cambric@ec.unipi.it, csodini@ec.unipi.it

Abstract

The aim of this paper is to propose a solution algorithm for a class of generalized quadratic programs having a polyhedral feasible region. The algorithm is based on the so called "optimal level solutions" method. Various global optimality conditions are discussed and implemented in order to improve the efficiency of the algorithm.

Key words: generalized quadratic programming, optimal level solutions, global optimization.
AMS - 2000 Math. Subj. Class. 90C20, 90C26, 90C31.
JEL - 1999 Class. Syst. C61, C63.

1 Introduction

The aim of this paper is to study, from both a theoretical, an algorithmic and a computational point of view, the following class of generalized quadratic problems:

$$P: \begin{cases} \inf f(x) = \phi(\frac{1}{2}x^TQx + q^Tx, d^Tx) \\ x \in X = \{x \in \Re^n: \ Ax \leq b\} \end{cases}$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, $q, d \in \Re^n$, $Q \in \Re^{n \times n}$ is positive definite and $X \neq \emptyset$. The scalar function $\phi(y_1, y_2)$ is assumed to be continuous and strictly increasing with respect to variable y_1, and is defined for all the values in $(\Omega_1 \times \Omega_2)$ where:

$$\begin{aligned} \Omega_1 &= \{y \in \Re: \ y = \frac{1}{2}x^TQx + q^Tx, \ x \in X\} \\ \Omega_2 &= \{y \in \Re: \ y = d^Tx, \ x \in X\} \end{aligned}$$

The considered class of objective functions $f(x)$ is extremely wide and it covers both multiplicative, fractional and d.c. quadratic functions. Just as an example, given any strictly increasing function $g_1 : \Omega_1 \rightarrow \Re$, any positive function $g_2 : \Omega_2 \rightarrow \Re_+$ and any real function $g_3 : \Omega_2 \rightarrow \Re$, then the following function $f(x)$ verifies the assumptions of problem P by using $\phi(y_1, y_2) = g_1(y_1)g_2(y_2) + g_3(y_2)$ (see also [9]):

$$f(x) = g_1 \left(\frac{1}{2} x^T Q x + q^T x \right) g_2 \left(d^T x \right) + g_3 \left(d^T x \right) \tag{1}$$

Various particular problems belonging to this class have been studied in the literature of mathematical programming and global optimization, from both a theoretic and an applicative point of view ([2, 13, 14, 15, 16, 18, 23]). In particular, it is worth noticing that this class covers several multiplicative, fractional, d.c. and generalized quadratic problems (see for all [4, 6, 7, 8, 9, 12, 19, 21]) which are very used in applications, such as location models, tax programming models, portfolio theory, risk theory, Data Envelopment Analysis (see for all [1, 10, 12, 17, 19, 24]).

The solution method proposed to solve this class of problems is based on the so called "optimal level solutions" method (see [3, 4, 5, 6, 7, 8, 9, 11, 20, 21, 22]). It is known that this is a parametric method, which finds the optimum of the problem by determining the minima of particular subproblems. In particular, the optimal solutions of these subproblems are obtained by means of a sensitivity analysis aimed to maintain the Karush-Kuhn-Tucker optimality conditions. Applying the optimal level solutions method to problem P we obtain some strictly convex quadratic subproblems which are independent of function $\phi(y_1, y_2)$. In other words, different problems share the same set of optimal level solutions, and this allow us to propose an unifying method to solve all of them.

In Section 2 we describe how the optimal level solutions method can be applied to problem P; in Section 3 a solution algorithm is proposed and fully described; in Section 4 some results are proposed in order to improve the performance of the method; finally, in Section 5 the results of a deep computational test are provided and discussed.

2 A parametric approach

In this section we show how problem P can be solved by means of the so called *optimal level solutions approach* (see for all [6, 7, 9, 20]). With this aim, let $\xi \in \Re$ be a real parameter and let us define the corresponding parametrical subset of X:

$$X_\xi = \{x \in \Re^n : Ax \leq b, \ d^T x = \xi\}$$

In the same way, the following further subset of X can be defined:

$$X_{[\xi_1,\xi_2]} = \{x \in \Re^n : \ Ax \leqq b, \ \xi_1 \leq d^T x \leq \xi_2\}$$

The following parametric subproblem can then be obtained just by adding to problem P the constraint $d^T x = \xi$:

$$P_\xi : \begin{cases} \min \phi(\frac{1}{2}x^T Q x + q^T x, \xi) \\ x \in X_\xi = \{x \in \Re^n : \ Ax \leqq b, \ d^T x = \xi\} \end{cases}$$

The parameter ξ is said to be a *feasible level* if the set X_ξ is nonempty. An optimal solution of problem P_ξ is called an *optimal level solution.* Since $\phi(y_1, y_2)$ is strictly increasing with respect to variable y_1, then for any feasible level ξ the optimal solution of problem P_ξ coincides with the optimal solution of the following strictly convex quadratic problem $\overline{P}_\xi$:

$$\overline{P}_\xi : \begin{cases} \min \frac{1}{2}x^T Q x + q^T x \\ x \in X_\xi = \{x \in \Re^n : \ Ax \leqq b, \ d^T x = \xi\} \end{cases}$$

Obviously, an optimal solution of problem P is also an optimal level solution and, in particular, it is the optimal level solution with the smallest value; the idea of this approach is then to scan all the feasible levels, studying the corresponding optimal level solutions, until the minimizer of the problem is reached. Starting from an incumbent optimal level solution, this can be done by means of a sensitivity analysis on the parameter ξ, which allows us to move in the various steps through several optimal level solutions until the optimal solution is found (see [9]).

Remark 2.1 Notice that problem $\overline{P}_\xi$ admits one and only one minimum point since its objective function is quadratic and positive definite and the feasible region X_ξ is closed. Since function $\phi(y_1, y_2)$ is strictly increasing with respect to variable y_1 and is defined for all the values in $(\Omega_1 \times \Omega_2)$, then the problem P_ξ admits one and only one minimum point too, the same of $\overline{P}_\xi$. As a consequence, the following logical implication holds:

$$\xi \in \Re \text{ is a feasible level} \quad \Rightarrow \quad \arg \min_{x \in X_\xi} f(x) \neq \emptyset$$

2.1 Sensitivity analysis

Let x' be the optimal solution of problem $\overline{P}_{\xi'}$, where $d^T x' = \xi'$, and let us consider the following Karush-Kuhn-Tucker conditions for $\overline{P}_{\xi'}$:

$$\begin{cases} Qx' + q = A^T\mu + d\lambda & \\ d^T x' = \xi' & \\ Ax' \leqq b & \textit{feasibility} \\ \mu \leqq 0 & \textit{optimality} \\ \mu^T(Ax' - b) = 0 & \textit{complementarity} \\ \mu \in \Re^m, \ \lambda \in \Re & \end{cases} \tag{2}$$

Since $\overline{P}_{\xi'}$ is a quadratic strictly convex problem, the previous system has at least one solution (μ', λ'). By means of a sort of sensitivity analysis, we now aim to study the optimal level solutions of problems $\overline{P}_{\xi'+\theta}$, $\theta \in (0, \epsilon)$ with $\epsilon > 0$ small enough. This can be done by maintaining the consistence of the Karush-Kuhn-Tucker systems corresponding to these problems. Since the Karush-Kuhn-Tucker systems are linear whenever the complementarity conditions are implicitly handled, then the solution of the optimality conditions regarding to $\overline{P}_{\xi'+\theta}$, $\theta \in (0, \epsilon)$ with $\epsilon > 0$ small enough, is of the kind:

$$x'(\theta) = x' + \theta\Delta_x \ , \ \mu'(\theta) = \mu' + \theta\Delta_\mu \ , \ \lambda'(\theta) = \lambda' + \theta\Delta_\lambda \tag{3}$$

It is worth pointing out that the strict convexity of problem $\overline{P}_{\xi'+\theta}$ guarantees for any $\theta \in [0, \epsilon)$ the uniqueness of the optimal level solution $x'(\theta) = x' + \theta\Delta_x$; this implies also the following important property:

vector Δ_x is unique and different from 0.

Clearly, the Karush-Kuhn-Tucker conditions are verified for values of $\theta \geq 0$ such that:

$$\begin{aligned} \textit{feasibility conditions} &: \ Ax' + \theta A\Delta_x \leqq b \ , \\ \textit{optimality conditions} &: \ \mu' + \theta\Delta_\mu \leqq 0 \ . \end{aligned}$$

Our aim is to determine the values of x', Δ_x, μ', Δ_μ, λ' and Δ_λ. By means of these parameters it can be computed also the value $\theta_m = \min\{F, O\}$ where:

$$\begin{aligned} F &= \sup\{\theta \geq 0 : Ax' + \theta A\Delta_x \leqq b\} \\ O &= \sup\{\theta \geq 0 : \mu' + \theta\Delta_\mu \leqq 0\} \end{aligned}$$

Observe that in [9] it has been proved that F and O are positive values whenever $\xi' < \xi_{max}$, so that θ_m results to be positive too.

Notice that for $\theta \in [0, \theta_m]$ both the optimality and the feasibility of $x'(\theta)$ are guaranteed, so that $x'(\theta)$ represents a segment of optimal level solutions. Starting from $x'(\theta_m)$ we can iterate the process determining a new segment of optimal level solutions. As a consequence, it yields that the set of the optimal level solutions is nothing but a connected set given by the union of segments.

In [9] various results are given for determining the values of the feasibility and optimality parameters. In this paper we aim to propose a simplified approach for determining them from a computational point of view, taking into account that the starting optimal level solution $x' = x'(0)$ is known.

Let x' be the optimal level solution corresponding to the level ξ' and let $x'(\delta) = x' + \delta\Delta_x$ be the optimal level solution corresponding to the level $\xi' + \delta$, with $\delta > 0$ small enough to guarantee that x' and $x'(\delta)$ belong to the same segment of optimal level solutions. Hence, it is:

$$\Delta_x = \frac{x'(\delta) - x'}{\delta}$$

Once Δ_x is computed, the set of binding constraints for $\theta \in [0, \delta]$ can be easily determined, so that the complementarity conditions in the Karush-Kuhn-Tucker system can be implicitly handled.

With this aim, let A_B the largest submatrix of A (made by rows of A) such that $A_B(x' + \theta\Delta_x) = b_B$ for all $\theta \in [0, \delta]$, where b_B is the subvector of b corresponding to A_B. Notice that the positivity of δ implies that such a condition is equivalent to the following one:

$$A_B x' = b_B \quad \text{and} \quad A_B\Delta_x = 0$$

By implicitly handling the complementarity conditions, the Karush-Kuhn-Tucker system becomes for $\theta \in [0, \delta]$ the following one:

$$\begin{cases} -A_B^T\mu_B & -d\lambda & +Qx & = -q \\ & & A_B x & = b_B \\ & & d^T x & = \xi' + \theta \end{cases}$$

which can be expressed in matrix form as:

$$S \begin{bmatrix} \mu_B \\ \lambda \\ x \end{bmatrix} = \begin{bmatrix} -q \\ b_B \\ \xi' + \theta \end{bmatrix}, \quad \text{where } S = \begin{bmatrix} -M^T & Q \\ 0 & M \end{bmatrix}, \quad M = \begin{bmatrix} A_B \\ d^T \end{bmatrix} \qquad (4)$$

Assuming the rows of matrix M to be linearly independent (which can be obtained by eventually deleting some redundant rows of A_B), we have that

Procedure **Parameters**(*inputs*: x'; *outputs*: Δ_x, μ', Δ_μ, λ', Δ_λ, F, O, θ_m)
let $\delta > 0$ be the step parameter; set $\xi' := d^T x'$;
let $x'_\delta := \arg\min\{\overline{P}_{\xi'+\delta}\}$ and set $\Delta_x := \frac{x'_\delta - x'}{\delta}$;
let A_B be the submatrix of A such that $A_B x' = b_B$ and $A_B \Delta_x = 0$;
if rank $\begin{bmatrix} A_B \\ d^T \end{bmatrix} <$ *rows* $\begin{bmatrix} A_B \\ d^T \end{bmatrix}$ *then* delete the redundant rows of A_B;
set $M := \begin{bmatrix} A_B \\ d^T \end{bmatrix}$, $S := \begin{bmatrix} -M^T & Q \\ 0 & M \end{bmatrix}$ and compute S^{-1};
set $\begin{bmatrix} \mu'_B \\ \lambda' \\ x' \end{bmatrix} := S^{-1} \begin{bmatrix} -q \\ b_B \\ \xi' \end{bmatrix}$ and $\begin{bmatrix} \Delta_{\mu_B} \\ \Delta_\lambda \\ \Delta_x \end{bmatrix} := S^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$;
set $F := \sup\{\theta \geq 0 : Ax' + \theta A\Delta_x \leq b\}$;
set $O := \sup\{\theta \geq 0 : \mu'_B + \theta\Delta_{\mu_B} \leq 0\}$ and $\theta_m := \min\{F, O\}$;
end proc.

matrix S is nonsingular (for the positive definiteness of Q). As a consequence the solution of (4) is unique and is given by:

$$\begin{bmatrix} \mu'_B(\theta) \\ \lambda'(\theta) \\ x'(\theta) \end{bmatrix} = \begin{bmatrix} \mu'_B \\ \lambda' \\ x' \end{bmatrix} + \theta \begin{bmatrix} \Delta_{\mu_B} \\ \Delta_\lambda \\ \Delta_x \end{bmatrix} = S^{-1} \begin{bmatrix} -q \\ b_B \\ \xi' \end{bmatrix} + \theta S^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Clearly, the parameters μ_i and Δ_{μ_i} corresponding to the nonbasic rows of A are equal to zero. Notice also that the value O can be computed by using the parameters μ'_B and Δ_{μ_B} only. The described approach is summarized in the following procedure "*Parameters*()".

In the solution algorithm there will be the need to evaluate the objective function $f(x) = \phi(\frac{1}{2}x^T Q x + q^T x, d^T x)$ along the obtained segment of optimal level solutions $x'(\theta)$, $\theta \in [0, \theta_m]$. With this aim, it is worth defining the following restriction function:

$$\begin{aligned} z(\theta) &= f(x' + \theta\Delta_x) = \\ &= \phi\left(\frac{1}{2}\theta^2\Delta_\lambda + \theta\lambda' + \frac{1}{2}x'^T Q x' + q^T x', \xi' + \theta\right) \end{aligned}$$

where we took into account (see for example [9]) that from the Karush-Kuhn-Tucker conditions it yields $d^T\Delta_x = 1$, $\Delta_x^T Q \Delta_x = \Delta_\lambda$, $\Delta_x^T A^T \Delta_\mu = 0$, $\Delta_x^T(Qx' + q) = \lambda'$.

2.2 Underestimation function

A key role in the study of problem P will be played by the use of a proper underestimation function, that is a function $\psi(\xi)$ which verifies the following property for all the feasible levels ξ:

$$\min_{x \in X_\xi} f(x) \geq \psi(\xi)$$

In order to determine such an underestimation function the following notations can be introduced:

$$\gamma = 1/d^T Q^{-1} d \quad , \quad \xi_u = -d^T Q^{-1} q$$

where γ is positive due to the positive definiteness of Q.

Lemma 2.1 *The following strictly convex quadratic parametric problem*

$$\begin{cases} \min \frac{1}{2} x^T Q x + q^T x \\ d^T x = \xi \end{cases}$$

attains the minimum at $\hat{x}(\xi) = \gamma(\xi - \xi_u) Q^{-1} d - Q^{-1} q$ *with minimum value* $\hat{g}(\xi) = \frac{1}{2}\gamma(\xi - \xi_u)^2 - \frac{1}{2} q^T Q^{-1} q$.

Proof The minimum point of the problem verifies the following necessary and sufficient optimality condition:

$$\begin{cases} Qx + q = \lambda d \\ d^T x = \xi \end{cases}$$

Since Q is positive definite it is also non singular, hence $\hat{x}(\xi) = \lambda Q^{-1} d - Q^{-1} q$. By means of simple calculations, from $\xi = d^T \hat{x}(\xi)$ we then have:

$$\begin{aligned} \lambda &= \gamma(\xi - \xi_u) \\ \hat{x}(\xi) &= \gamma(\xi - \xi_u) Q^{-1} d - Q^{-1} q \\ \hat{g}(\xi) &= \frac{1}{2} \hat{x}(\xi)^T Q \hat{x}(\xi) + q^T \hat{x}(\xi) = \\ &= \frac{1}{2} \lambda^2 d^T Q^{-1} d - \frac{1}{2} q^T Q^{-1} q = \frac{1}{2} \gamma(\xi - \xi_u)^2 - \frac{1}{2} q^T Q^{-1} q. \end{aligned}$$

□

The previous lemma shows that it is possible to explicitly determine the line of unconstrained minima corresponding to problem P, which from now on will be denoted as follows:

$$U_P = \{x \in \Re^n : x = \hat{x}(\xi), \xi \in \Re\}$$

The positiveness of γ implies that function $\hat{g}(\xi)$ is a convex parabola with minimum value $\hat{g}(\xi_u) = -\frac{1}{2} q^T Q^{-1} q$. The following result suggests a first possible underestimation function for problem P.

Theorem 2.1 *Consider problem P. Then, for any feasible level ξ it is:*

$$\min_{x \in X_\xi} f(x) \geq \phi(\hat{g}(\xi), \xi).$$

Proof Since the scalar function $\phi(y_1, y_2)$ is strictly increasing with respect to variable y_1 and by means of Lemma 2.1 it results:

$$\begin{aligned}
\min_{x \in X,\ d^T x = \xi} f(x) &\geq \min_{x \in \Re^n,\ d^T x = \xi} f(x) = \\
&= \min_{x \in \Re^n,\ d^T x = \xi} \phi(\frac{1}{2} x^T Q x + q^T x, \xi) = \\
&= \phi\left(\min_{x \in \Re^n,\ d^T x = \xi} \left\{ \frac{1}{2} x^T Q x + q^T x \right\}, \xi \right) = \\
&= \phi(\hat{g}(\xi), \xi)
\end{aligned}$$

□

In the case the line of unconstrained minima U_P does not intersect the feasible region X the underestimation function can be furthermore improved.

With this aim, assume $U_P \cap X = \emptyset$ and let $x_s \in X$ and $v_s \in (Q^{-1}d)^\perp$ be such that $\{x \in \Re^n : v_s^T x = v_s^T x_s\}$ is a support hyperplane for X separating region X itself and the unconstrained minima line U_P, with $v_s^T x \leq v_s^T x_s$ for all $x \in X$. Notice that:

$$(Q^{-1}d)^\perp = \{v \in \Re^n : v = Mw, w \in \Re^n\}$$

where

$$M = I - \frac{Q^{-1} d d^T Q^{-1}}{d^T Q^{-1} Q^{-1} d}$$

is a symmetric singular positive semidefinite matrix such that $M^2 = M$, with one eigenvalue equal to 0 (and corresponding eigenvector $Q^{-1}d$) and $n-1$ eigenvalues equal to 1 (and corresponding eigenvectors in $(Q^{-1}d)^\perp$).

From $MQ^{-1}d = 0$ it yields $M\hat{x}(\xi) = -MQ^{-1}q$ for all $\xi \in \Re$, so that given a point $x \in X$ it results:

$$v = M(\hat{x}(\xi) - x) = M(-Q^{-1}q - x)$$

Such a vector v is nothing but the vector starting from point $x \in X$ and reaching the unconstrained minima line in orthogonal way. To determine the separating hyperplane we are then left to determine the point $x_s \in X$ which is as close as possible to the unconstrained minima line, that is the one providing the smallest vector $M(-Q^{-1}q - x)$. In other words we have to minimize the quadratic form $(M(-Q^{-1}q-x))^T(M(-Q^{-1}q-x))$ and this can be done by solving the following equivalent convex quadratic problem (recall that $M^2 = M$):

$$x_s = \arg\min_{x\in X}\left\{\frac{1}{2}x^T Mx + q^T Q^{-1}Mx\right\}$$

From now on we can then assume:

$$v_s = M(-Q^{-1}q - x_s)$$

Notice that from $M^2 = M$ and $M\hat{x}(\xi) = -MQ^{-1}q$ it yields:

$$v_s^T\hat{x}(\xi) = -v_s^T Q^{-1}q \quad \text{and} \quad v_s^T(\hat{x}(\xi) - x_s) = v_s^T v_s$$

Clearly, it is $v_s \neq 0$ if and only if the unconstrained minima line U_P does not intersect the feasible region X. To determine a tighter underestimation function let us define, in the case $X \cap U_P = \emptyset$, the following notation:

$$\nu = \frac{v_s^T v_s}{v_s^T Q^{-1} v_s} > 0$$

Lemma 2.2 *The following strictly convex quadratic parametric problem*

$$\begin{cases} \min \ \frac{1}{2}x^T Qx + q^T x \\ d^T x = \xi \\ v_s^T x \le v_s^T x_s \end{cases}$$

attains the minimum at $\ddot{x}(\xi) = \hat{x}(\xi) - \nu(Q^{-1}v_s)$ *with minimum value* $\ddot{g}(\xi) = \hat{g}(\xi) + \frac{1}{2}\nu(v_s^T v_s)$.

Proof The minimum point of the problem verifies the following necessary and sufficient optimality condition:

$$\begin{cases} Qx + q = \lambda d + \alpha v_s \\ d^T x = \xi \ , \ v_s^T x \le v_s^T x_s \\ \alpha(v_s^T x - v_s^T x_s) = 0 \\ \alpha \le 0 \ , \ \lambda \in \Re \end{cases}$$

Since Q is positive definite it is also non singular, hence $\ddot{x}(\xi) = \lambda Q^{-1}d - Q^{-1}q + \alpha Q^{-1}v_s$. By means of simple calculations, from $\xi = d^T\ddot{x}(\xi)$ and $v_s \in (Q^{-1}d)^{\perp}$ we then have $\lambda = \gamma(\xi - \xi_u)$ so that $\ddot{x}(\xi) = \hat{x}(\xi) + \alpha Q^{-1}v_s$. From the feasibility conditions we have also:

$$v_s^T x_s \ge v_s^T\ddot{x}(\xi) = v_s^T\hat{x}(\xi) + \alpha\frac{v_s^T v_s}{\nu}$$

and hence, from $v_s^T(\hat{x}(\xi) - x_s) = v_s^T v_s$, it yields

$$\alpha \leq \frac{v_s^T(x_s - \hat{x}(\xi))}{v_s^T v_s}\nu = -\nu < 0$$

Being $\alpha < 0$ from the complementarity conditions we get $v_s^T \ddot{x}(\xi) = v_s^T x_s$, which yields $\alpha = -\nu$ and $\ddot{x}(\xi) = \hat{x}(\xi) - \nu(Q^{-1}v_s)$. Finally, it results:

$$\begin{aligned}
\ddot{g}(\xi) &= \frac{1}{2}\ddot{x}(\xi)^T Q\ddot{x}(\xi) + q^T\ddot{x}(\xi) = \\
&= \hat{g}(\xi) - \nu(v_s^T\hat{x}(\xi)) + \frac{1}{2}\nu(v_s^T v_s) - \nu(q^T Q^{-1}v_s) = \\
&= \hat{g}(\xi) + \nu(v_s^T Q^{-1}q) + \frac{1}{2}\nu(v_s^T v_s) - \nu(q^T Q^{-1}v_s) = \hat{g}(\xi) + \frac{1}{2}\nu(v_s^T v_s)
\end{aligned}$$

□

The proof of the following result is analogous to the one of Theorem 2.1.

Theorem 2.2 *Consider problem P and assume that $X \cap U_P = \emptyset$. Then, for any feasible level ξ it is:*

$$\min_{x \in X_\xi} f(x) \geq \phi\left(\hat{g}(\xi) + \frac{1}{2}\nu(v_s^T v_s), \xi\right).$$

As a conclusion, the following underestimation function can be defined:

$$\psi(\xi) = \phi\left(\hat{g}(\xi) + \hat{g}_0, \xi\right)$$

where:

$$\hat{g}_0 = \begin{cases} 0 & \text{if } X \cap U_P \neq \emptyset \\ \frac{1}{2}\nu(v_s^T v_s) & \text{if } X \cap U_P = \emptyset \end{cases}$$

Notice that in the case $X \cap U_P = \emptyset$ it is

$$\phi\left(\hat{g}(\xi) + \hat{g}_0, \xi\right) > \phi\left(\hat{g}(\xi), \xi\right)$$

since function $\phi(y_1, y_2)$ is strictly increasing with respect to variable y_1, $v_s \neq 0$ and $\nu > 0$. Notice also that the continuity of $\phi(y_1, y_2)$ implies the continuity of $\psi(\xi)$. From a theoretical point of view, the previous underestimation function $\psi(\xi)$ allows to prove the following result which generalizes the one provided in Remark 2.1.

Corollary 2.1 *Consider problem P. Then, for any compact interval of feasible levels $[\xi_1, \xi_2]$ it results:*

$$\arg\min_{x \in X_{[\xi_1,\xi_2]}} f(x) \neq \emptyset \quad \text{and} \quad \min_{x \in X_{[\xi_1,\xi_2]}} f(x) \geq \min_{\xi \in [\xi_1,\xi_2]} \psi(\xi)$$

Proof Since $X_{[\xi_1,\xi_2]}$ is a closed set and $f(x)$ is a continuous function, then the image set $f(X_{[\xi_1,\xi_2]})$ is closed too. Since $\psi(\xi)$ is continuous and the interval $[\xi_1,\xi_2]$ is compact, then $\min_{\xi\in[\xi_1,\xi_2]}\psi(\xi)$ exists. For Theorem 2.1 it then yields that:

$$f(x) \geq \min_{\xi\in[\xi_1,\xi_2]} \psi(\xi) \quad \forall x \in X_{[\xi_1,\xi_2]}$$

As a consequence, the set $f(X_{[\xi_1,\xi_2]})$ is closed and lower bounded, so that the result is proved. □

From the previous corollary it yields that problem P can be unbounded only along extreme rays with feasible levels ξ going towards $+\infty$ or $-\infty$, while it admits minimum in any compact set of feasible levels.

3 Solution algorithm

In order to find a global minimum (assuming that one exists) it would be necessary to solve problems $\overline{P}_\xi$ for all the feasible levels. In this section we will show that this can be done by means of a finite number of iterations, using the results of the previous section.

The method scans all the feasible levels looking for the optimal solution starting from a certain feasible level ξ_F. With this aim, there will be the need of visiting the feasible levels lower than ξ_F in decreasing order. This can be done by reversing the problem itself, observing that problem P can be equivalently rewritten in the following form:

$$P \equiv \tilde{P} : \begin{cases} \inf f(x) = \tilde{\phi}(\frac{1}{2}x^T Qx + q^T x, \tilde{d}^T x) \\ \qquad x \in X \end{cases}$$

where $\tilde{\phi}(y_1, y_2) = \phi(y_1, -y_2)$ and $\tilde{d} = -d$. In this light, the decreaseness of the feasible levels of P corresponds to the increaseness of the feasible levels of $\tilde{P}$.

The following procedures "*Main*()" and "*Visit*()" can then be proposed. Procedure "*Main*()" initialize the algorithm by determining the set of feasible levels and a "good" starting incumbent solution, then it uses procedure "*Visit*()" to obtain the global optimal solution (if it exists). As it will be deepened on in the next section, a "good" incumbent solution is useful in order to reduce the set of feasible levels to be explicitly scanned, thus improving the performance of the proposed method.

In particular, the optimal level solutions x_1' and x_2' are determined in order to have a good starting incumbent solution. The obtained starting incumbent solution results to be extremely effective in the case the objective

Procedure **Main**(*inputs*: P; *outputs*: Opt, $OptVal$)

Compute the values $\xi_{min} := \inf_{x \in X} d^T x$ and $\xi_{max} := \sup_{x \in X} d^T x$;
Let $\xi_{big} >> 0$ and set $\xi_1 := \max\{-\xi_{big}, \xi_{min}\}$; $\xi_2 := \min\{\xi_{big}, \xi_{max}\}$;
Compute $x'_1 := \arg\min\{\overline{P}_{\xi_1}\}$ and $x'_2 := \arg\min\{\overline{P}_{\xi_2}\}$;
if $f(x'_1) < f(x'_2)$ *then* $\overline{x} := x'_1$ *else* $\overline{x} := x'_2$ *end if*;
Set $UB := f(\overline{x})$ and let $I_P = \{\xi \in \Re : A\hat{x}(\xi) \leq b\}$;
if $I_P = \emptyset$ *then* $\xi_F := d^T x_s$; $x_F := \arg\min\{\overline{P}_{\xi_F}\}$;
  else if $I_P \cap [\xi_1, \xi_2] = \emptyset$ *then* $\xi_F := \frac{\xi_1 + \xi_2}{2}$; $x_F := \arg\min\{\overline{P}_{\xi_F}\}$;
    else $\xi_F := \arg \min_{\xi \in I_P \cap [\xi_1, \xi_2]} \{\psi(\xi)\}$; $x_F := \hat{x}(\xi_F)$;
  end if;
end if;
if $f(x_F) < UB$ *then* $\overline{x} := x_F$ and $UB := f(x_F)$ *end if*;
if $d^T \overline{x} \geq \xi_F$ *then*
  $[\overline{x}, UB] := Visit(P, \xi_F, \xi_{max}, \overline{x}, UB)$;
  $[\overline{x}, UB] := Visit(\tilde{P}, -\xi_F, -\xi_{min}, \overline{x}, UB)$;
else
  $[\overline{x}, UB] := Visit(\tilde{P}, -\xi_F, -\xi_{min}, \overline{x}, UB)$;
  $[\overline{x}, UB] := Visit(P, \xi_F, \xi_{max}, \overline{x}, UB)$;
end if;
$Opt := \overline{x}$ and $OptVal := UB$;

end proc.

function of problem P is unbounded along a feasible extremum ray. The starting feasible level ξ_F and its corresponding optimal level solution x_F are determined taking into account of the possibility to have $U_P \cap X = \emptyset$ or not.

Procedure "*Visit*()" scans iteratively the given set of feasible levels obtaining the best solution. Notice that "*Visit*()" uses two subprocedures, the first one is procedure "*Parameters*()" which has been already described in Section 3, the latter one is procedure "*MinRestriction*()" which determines the minimum of the continuous single valued function $z(\theta)$ in the closed interval $[0, \theta_m]$. Observe that procedure "*MinRestriction*()" can be implemented numerically, and eventually improved for specific functions $f(x)$ (see [6, 7, 9, 20]). Notice finally that in procedure "*Visit*()" there is also one more optional subprocedure, namely "*ImplicitVisit*()", which is aimed to improve the performance of the solution algorithm by implicitly visiting some of the feasible levels to be scanned. This optional procedure will be discussed in the next section.

The correctness of the proposed algorithm follows since all the feasible levels are scanned and the optimal solution, if it exists, is also an optimal level solution.

Procedure **Visit**(*inputs*: P, ξ_F, ξ_{max}, $\overline{x}$, UB; *outputs*: Opt, $OptVal$)
  $\xi' := \xi_F$; $x' := x_F$;
  # $[\xi', x'] := ImplicitVisit(\xi', x', \xi_{max}, false)$;
  while $\xi' < \xi_{max}$
    set $[\Delta_x, \mu', \Delta_\mu, \lambda', \Delta_\lambda, F, O, \theta_m] := Parameters(x')$;
    let $z(\theta) = \phi\left(\frac{1}{2}\theta^2\Delta_\lambda + \theta\lambda' + \frac{1}{2}x'^TQx' + q^Tx', \xi' + \theta\right)$;
    set $[\overline{\theta}, z_{inf}] := MinRestriction(z(\theta), [0, \theta_m])$;
    if $z_{inf} = -\infty$ then $\overline{x} := []$; $UB := -\infty$; $\xi' := +\infty$ *else*
      if $z_{inf} < UB$ *then*
        $UB := z_{inf}$;
        if $\overline{\theta} = +\infty$ *then* $\overline{x} := []$ *else* $\overline{x} := x' + \overline{\theta}\Delta_x$ *end if*;
      end if;
      set $\xi' := \xi' + \theta_m$ and $x' := x' + \theta_m\Delta_x$;
    end if;
    # $[\xi', x'] := ImplicitVisit(\xi', x', \xi_{max}, true)$;
  end while;
  $Opt := \overline{x}$; $OptVal := UB$;
end proc.

It remains to verify the convergence (finiteness), that is to say that the procedure stops after a finite number of steps. First note that, at every iterative step of the proposed algorithm, the set of binding constraints changes; note also that the level is increased from ξ' to $\xi' + \theta_m > \xi'$, so that it is not possible to obtain again an already used set of binding constraints; the convergence then follows since we have a finite number of possible sets of binding constraints.

Remark 3.1 Let us point out that problems $\overline{P}_\xi$ are independent of the function ϕ. This means that problems having the same feasible region, the same Q, q and d, but different function ϕ (either multiplicative or fractional or d.c.), they share the same set of optimal level solutions. As a consequence, when procedure "*Main*()" explicitly visits all the feasible levels, these different problems are solved by means of the same iterations of the while cycle in procedure "*Visit*()".

4 Algorithm improvements

In this section we aim to discuss how the proposed algorithm can be improved in the visit of the feasible levels.

First of all, let us notice that in the various iterations of procedure "*Visit*()" some feasible levels could be implicitly visited in the case $O > F$. With this aim, first note that for all $\theta \in [0, O]$, the value $z(\theta)$ is a lower bound for the parametric problem $\bar{P}_{\xi'+\theta}$; in fact if $\theta \in [0, \theta_m]$ then $x'(\theta)$ is an optimal level solution, while if $\theta \in (F, O]$ then $x'(\theta)$ is unfeasible for $\bar{P}_{\xi'+\theta}$ but is an optimal solution of a problem with the same objective function as $\bar{P}_{\xi'+\theta}$ and a feasible region containing $X_{\xi'+\theta}$. As a consequence, if the minimum value of $z(\theta)$ in the interval $(F, O]$ is greater than or equal to UB then the feasible levels $(F, O]$ can be skipped.

Analogously, some more feasible levels can be implicitly visited by using the underestimation function $\psi(\xi)$. In fact, given $\xi_a \in [\xi', \xi_{max}]$ it can be easily proved that:

$$\psi(\xi) \geq UB \ \forall \xi \in [\xi', \xi_a] \quad \Rightarrow \quad \min_{x \in X_{[\xi', \xi_{max}]}} f(x) = \min_{x \in X_{[\xi_a, \xi_{max}]}} f(x)$$

This property suggests another way to improve the algorithm by reducing in the various iterations of procedure "*Visit*()" the set of feasible levels to be scanned, that is to say by implicitly visiting some of the feasible levels.

As a conclusion, the following procedure "*ImplicitVisit*()" can be proposed in order to improve the visit of the feasible levels. Notice that in the procedure the *lower-level sets* of function $\psi(\xi)$ have been denoted with $L(\psi, UB) = \{\xi \in \Re : \psi(\xi) \leq UB\}$.

Finally notice that procedure "*ImplicitVisit*()" is as more effective as smaller is the value UB of the incumbent solution. For this very reason, in order to improve the algorithm performance it is important to initialize the method with a "good" starting incumbent solution, as it has been described in the previous section.

5 Computational results

The previously described procedures have been fully implemented with the software MatLab 7.4 R2007a on a computer having 2 Gb RAM and two Xeon dual core processors at 2.66 GHz.

The following four different objective functions have been used in the computational test:

Procedure **ImplicitVisit**(*inputs*: ξ', x', ξ_{max}, *inside*; *outputs*: ξ', x')
$\xi'_{old} := \xi'$;
if $\xi' < \xi_{max}$ *and* $\psi(\xi') > UB$ *then*
let $\mathcal{L} := [\xi', \xi_{max}] \cap L(\psi, UB)$;
if $\mathcal{L} = \emptyset$ *then* $\xi' := \xi_{max}$ *else* $\xi' := \min\{\mathcal{L}\}$ *end if*;
end if;
if $\xi' < \xi_{max}$ *and inside = true and* $O - F > \xi' - \xi'_{old}$ *then*
$[\tilde{\theta}, \tilde{z}_{inf}] := MinRestriction(z(\theta), [F + \xi' - \xi'_{old}, \min\{O, \xi_{max} - \xi'_{old} + F\}])$;
if $\tilde{z}_{inf} >= UB$ *then*
$\xi' := \xi'_{old} + O - F$;
if $\xi' < \xi_{max}$ *and* $\psi(\xi') > UB$ *then*
let $\mathcal{L} := [\xi', \xi_{max}] \cap L(\psi, UB)$;
if $\mathcal{L} = \emptyset$ *then* $\xi' := \xi_{max}$ *else* $\xi' := \min\{\mathcal{L}\}$ *end if*;
end if;
end if;
end if;
if $\xi' < \xi_{max}$ *and* $\xi' > \xi'_{old}$ *then* $x' := \arg\min\{\overline{P}_{\xi'}\}$ *end if*;
end proc.

	$\phi(y_1, y_2)$	$f(x)$
P_1	$y_1 - y_2^2$	$\left(\frac{1}{2}x^TQx + q^Tx\right) - \left(d^Tx\right)^2$
P_2	$y_1 y_2^3$	$\left(\frac{1}{2}x^TQx + q^Tx\right)\left(d^Tx\right)^3$
P_3	y_1/y_2^2	$\left(\frac{1}{2}x^TQx + q^Tx\right) / \left(d^Tx\right)^2$
P_4	$y_2^2 log(y_1)$	$\left(d^Tx\right)^2 log\left(\frac{1}{2}x^TQx + q^Tx\right)$

where in P_2 and P_3 function d^Tx is positive over the feasible region, while in P_4 function $\frac{1}{2}x^TQx + q^Tx$ is positive over the feasible region.

The problems have been randomly created; in particular, matrices and vectors $Q \in \Re^{n\times n}$, $q, d \in \Re^n$, $A \in \Re^{m\times n}$, $b \in \Re^m$, $m = 3n$, have been generated with components in the interval [-10,10] by using the "rand()" MatLab function (numbers generated with uniform distribution). Within the procedures, the linear problems and the convex quadratic problems have been solved with the "linprog()" and "quadprog()" MatLab functions, respectively.

For each amount "n" of variables a number "num" of problems have been randomly generated and each of these problems have been solved for both the objective functions in P_1, P_2, P_3 and P_4. The average number of iterations and the CPU times spent by the algorithm to solve the problems are given as the result of the test (see Tables 1 and 2).

In order not to waste time, the complete visit of the feasible levels have been tested for dimensions up to $n = 50$, while the use of procedure "ImplicitVisit()" have been tested reaching dimension $n = 100$. Clearly, in the case procedure "ImplicitVisit()" is not used (that is all the feasible levels are explicitly scanned) we provide only the results related to problem P_1 since all the problems are solved in the same number of iterations (see Remark 3.1).

n	num	Complete	With Implicit Visit			
			P_1	P_2	P_3	P_4
10	1000	24.987	3.874	3.929	12.124	9.839
15	1000	39.72	5.545	4.283	19.103	16.148
20	1000	53.725	7.325	4.5	25.477	22.389
25	1000	68.135	9.198	4.783	32.465	28.784
30	1000	83.094	10.586	4.841	38.017	35.666
35	1000	99.184	11.406	5.3674	41.631	40.53
40	600	114.46	11.678	5.6167	43.21	46.02
45	600	134.63	11.28	6.195	44.662	51.573
50	600	158.65	11.517	7.0617	47.462	57.552
60	600	–	10.808	9.13	48.792	74.19
70	600	–	11.29	8.655	50.252	93.698
80	600	–	11.727	10.468	51.43	116.93
90	400	–	13.492	12.255	56.083	140.03
100	400	–	15.01	15.098	55.3	165.14

Table 1: Amount of iterations

The obtained results point out the effectiveness of the improvements proposed in Section 5; in particular, the performance is strongly improved for problems P_1 and P_2, for both the number of iterations and the spent CPU time.

6 Conclusions

The proposed algorithm allows to solve a wide range of nonconvex problems. The computational test shows that it is possible to efficiently handle problems with up to 100 variables. In particular, the improvement criteria suggested in Section 5 resulted to be extremely effective in making the algorithm efficient.

Further improvements could be based on the study of the quasiconvexity of functions $f(x)$ and $\psi(\xi)$. The quasiconvexity of $f(x)$ suggests to stop the algorithm when a local minimum is found, while the quasiconvexity of $\psi(\xi)$ makes the condition $\psi(\xi') > UB$ a global optimality condition and a concrete stopping criterion. In this light, notice that if $f(x)$ is quasiconvex then $\psi(\xi)$ is quasiconvex too. Improvements could be obtained also by iteratively

n	num	Complete	With Implicit Visit			
			P_1	P_2	P_3	P_4
10	1000	0.90392	0.64764	0.68338	1.0331	0.92926
15	1000	1.8236	0.84843	0.79404	1.5577	1.4047
20	1000	12.747	2.6592	1.9635	7.2688	6.4555
25	1000	21.04	4.1972	2.7459	11.63	10.322
30	1000	33.285	6.1779	3.7045	17.205	15.966
35	1000	60.252	10.051	5.9661	28.141	26.716
40	600	89.185	13.284	7.9423	37.125	37.893
45	600	135.08	16.888	11.088	48.872	53.576
50	600	195.65	21.374	14.843	63.532	71.894
60	600	–	29.563	25.292	93.901	129.16
70	600	–	42.358	34.813	133.51	217.86
80	600	–	57.579	52.35	177.72	346.28
90	400	–	82.368	74.708	244.82	518.6
100	400	–	114.04	110.88	307.25	758.13

Table 2: CPU time

updating the underestimation function $\psi(\xi)$ over the feasible subset $X_{[\xi',\xi_{max}]}$ which remains to be visited.

References

[1] Barros A.I. (1998) Discrete and Fractional Programming Techniques for Location Models. Combinatorial Optimization, vol.3, Kluwer Academic Publishers, Dordrecht.

[2] Bomze I.M., Csendes T. and R. Horst (eds) (1997) Developments in Global Optimization. Nonconvex Optimization and its Applications, vol.18, Kluwer Academic Publishers, Dordrecht.

[3] Cambini A. and L. Martein (1986) A modified version of Martos' algorithm. Methods of Operation Research, 53:33-44.

[4] Cambini A., Martein L. and C. Sodini (1983) An algorithm for two particular nonlinear fractional programs. Methods of Operations Research, 45:61-70.

[5] Cambini R. (1994) A class of non-linear programs: theoretical and algorithmic results. In: Generalized Convexity, S. Komlósi, T. Rapcsák and S. Schaible (eds), Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, 405:294-310.

[6] Cambini R. and C. Sodini (2002) A finite algorithm for a particular d.c. quadratic programming problem. Annals of Operations Research, 117:33-49.

[7] Cambini R. and C. Sodini (2003) A finite algorithm for a class of non-linear multiplicative programs. Journal of Global Optimization, 26:279-296.

[8] Cambini R. and C. Sodini (2005) A sequential method for a class of box constrained quadratic programming problems. Report n.272, Department of Statistics and Applied Mathematics.

[9] Cambini R. and C. Sodini (2007) An unifying approach to solve a class of parametrically-convexifiable problems. In: Generalized Convexity and Related Topics, I.V. Konnov, D.T. Luc and A.M. Rubinov (eds), Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, 583:149-166.

[10] Cooper W.W., Seiford L.M. and J. Zhu (eds) (2004) Handbook on Data Envelopment Analysis. International Series in Operations Research & Management Science, vol.71, Kluwer Academic Publishers, Boston.

[11] Ellero A. (1996) The optimal level solutions method. Journal of Information & Optimization Sciences, 17:355-372.

[12] Frenk J.B.G. and S. Schaible (2005) Fractional programming. In: Handbook of Generalized Convexity and Generalized Monotonicity, Nonconvex Optimization and Its Applications, 76:335-386, Springer, New York.

[13] Horst R. and P.M. Pardalos (eds) (1995) Handbook of Global Optimization. Nonconvex Optimization and its Applications, vol.2, Kluwer Academic Publishers, Dordrecht.

[14] Horst R. and H. Tuy (1996) Global Optimization: Deterministic Approaches. 3rd rev., Springer, Berlin.

[15] Horst R., Pardalos P.M. and N.V. Thoai (2001) Introduction to Global Optimization. 2nd ed., Nonconvex Optimization and Its Applications, vol.48, Kluwer Academic Publishers, Dordrecht.

[16] Konno H. and T. Kuno (1995) Multiplicative programming problems. In: Handbook of Global Optimization, R. Horst and P.M. Pardalos (eds), Nonconvex Optimization and Its Applications, Kluwer Academic Publishers, Dordrecht, 2:369-405.

[17] Mjelde K.M. (1983) Methods of the Allocation of Limited Resources. Wiley, New York.

[18] Ryoo H.-S. and N.V. Sahinidis (2003) Global optimization of multiplicative programs. Journal of Global Optimization, 26:387-418.

[19] Schaible S. (1995) Fractional programming. In: Handbook of Global Optimization, Horst R. and P.M. Pardalos (eds), Nonconvex Optimization and Its Applications, 2:495-608, Kluwer Academic Publishers, Dordrecht.

[20] Schaible S. and C. Sodini (1995) A finite algorithm for generalized linear multiplicative programming. Journal of Optimization Theory and Applications, 87:441-455.

[21] Sodini C. (1986) Minimizing the sum of a linear function and the square root of a convex quadratic form. Methods of Operations Research, 53:171-182.

[22] Sodini C. (1990) Equivalence and parametric analysis in linear fractional programming. In: Generalized Convexity and Fractional Programming with Economic Applications, Cambini A., Castagnoli E., Martein L., Mazzoleni P. and S. Schaible (eds), Lecture Notes in Economics and Mathematical Systems, 345:143-154, Springer-Verlag, Berlin.

[23] Tuy H. (1998) Convex Analysis and Global Optimization. Nonconvex Optimization and Its Applications, vol.22, Kluwer Academic Publishers, Dordrecht.

[24] Yitzakhi S. (1982) A tax programming model. Journal of Public Economics, 19:107-120.