8th International Conference on DEVELOPMENT AND APPLICATION SYSTEMS
S u c e a v a, R o m a n i a, M a y 25 – 27, 2 0 0 6

# A NEWTON-TYPE METHOD FOR CONVEX QUADRATIC PROGRAMMING

**Vasile MORARU[1], Radu MELNIC[2]**
*Technical University of Moldova*
*168, Bd. Stefan cel Mare si Sfint*
*2012 Chisinau, Moldova*
*[1]moraru@mail.utm.md*
*[2]rmelnic@mail.utm.md*

***Abstract.*** *In this paper it is talked over Newton-type method for the problem of minimizing a convex quadratic function subject to linear constraints. The determination of the Karush-Kuhn-Tucker system solution is reduced to the solving of a nonlinear system of equations with continuously differentiable functions.*
***Keywords:*** *quadratic convex programming, Karush-Kuhn-Tucker conditions, Lagrange multipliers, max operator, Newton's method.*

## Introduction

In this paper we consider the convex quadratic programming problem in a standard form:

$$\left.\begin{array}{l} f(x)=\dfrac{1}{2}x^TQx+q^Tx\to\min \\ \text{subject to}: \ Ax=b, x\geq 0. \end{array}\right\} \quad (1)$$

We'll suppose that $Q\in R^{n\times n}$ is a symmetric matrix and positive semidefinite, $x^TQx\geq 0$, $\forall x\in R^n$ but $A\in R^{m\times n}$ and $rank(A)=m\leq n$, $b\in R^n$. The "$^T$" symbol indicates the operation of transposition.

The interest for the considered problem is very big. From the first point of view, the quadratic models are frequently met in different real applications, like VLSI design, the structural analysis, the support of the vector machines, graphs theory, etc. On the second part, almost in all Newton methods and quasi-Newton for solving the constrained optimization problems, it is requested at each step the solving of the type (1) problems, where $q$ and $Q$ are respectively the gradient and the Hesse matrix of the objective function [3-6]. A rich bibliography concerning to the problems of (1) type can be found in (7), a work-paper that contains more than 1000 references.

The Karush-Kuhn-Tucker conditions for the problem (1) will be written [5, 8]:

$$\left.\begin{array}{l} Qx+q+A^T\mu-\lambda=0, \\ Ax-b=0, \\ x^T\lambda=0, \\ x\geq 0,\ \lambda\geq 0. \end{array}\right\} \quad (2)$$

where: $\mu\in R^m$ and $\lambda\in R^n$ are Lagrange multipliers.

Because *f(x)* is a convex function, the Karush-Kuhn-Tucker (2) conditions are necessary and sufficient in order that $x^*$ will be considered aggregate minimum point in the problem (1).

One of the most well-known and widespread solving quadratic programming problems scheme is based on the direct solving of the system with equations (2). Nowadays exist different methods of solving of Karush-Kuhn-Tucker system (see, for example, [3-5, 8-11]).

In this paper is presented one transformational technique of the Karush-Kuhn-Tucker system (2) in one equivalent system with differentiable continuous functions. More than that, it is not necessary to follow the non-negative conditions $x\geq 0, \lambda\geq 0$, they have been accomplished automatically following the application of the considered procedure. This thing is done by the plus function (that is not continuous differentiable), and then it is replaced by another convenient function.

**Equivalence of the quadratic program to a system of smooth equations**

We note that the max and min operators:
$(a)^+ = \max\{o,a\}$ and $(a)^- = \min\{o,a\}$ ,
for $a \in R$. Let's suppose that
$y_i \in R, i = 1,2,...m,\ z_i \in R, i = 1,2,...n$
and consider the function
$F : R^n \times R^m \times R^n \to R^n \times R^m \times R^n$
defined by

$$F(x,y,z) = \begin{pmatrix} \nabla f(x) + A^T y - z^+ \\ Ax - b \\ x_1 + z_1^- \\ x_2 + z_2^- \\ . \\ . \\ x_n + z_n^- \end{pmatrix}$$

The next theorem is following [8]:

**Theorem 1.** *The vector* $x^* \in R^n$ *verifies the Karush-Kuhn-Tucker (2) conditions if only if exists the vector* $y^* \in R^m$ $z^* \in R^n$, *thus* $F(x^*, y^*, z^*) = 0$.

As a result, we can reduce the solving of the problem (1) to the solving of the non-linear system of equations:

$$F(x,y,z) = 0 \tag{3}$$

The functions $(\cdot)^+$ and $(\cdot)^-$ are non-differentiable everywhere, and then the function *F(x, y, z)* is also not differentiable. This asks big problems at the solving of the system (3), because we can not use effective methods, like the Newton's.

Let's consider the Chen-Harker-Kanzow-Smale smoothing function (CHKS):

$$\varphi(u,v) = \frac{1}{2}\left(v + \sqrt{v^2 + \psi(u)}\right),$$

where: $\psi : R \to R$, a function with the properties:

1. $\psi(u)$ is continuous differentiable;
2. $\psi(u) \geq 0$ for $\forall u \in R$;
3. $\psi(0) = 0$.

As a function $\psi$ can be considered, for example [12], the function $\psi(u) = cu^2$, $c = const \neq 0$.

We notice that the function $\varphi(u,v)$ is continuously differentiable for $\forall(u,v) \in R \times R$, $(u,v) \neq (0,0)$ and verifies the relations:

1. $\varphi(0,v) = \frac{1}{2}(v + |v|) = (v)^+ = \max\{0,v\}$;
2. $\varphi(u,0) = \frac{1}{2}\sqrt{\psi(u)}$;
3. $v - \varphi(0,v) = (v)^- = \min\{0,v\}$
4. $\nabla\varphi(u,v) = \begin{bmatrix} \dfrac{\psi'(u)}{4\sqrt{v^2 + \psi(u)}} \\ \dfrac{1}{2}\left[1 + \dfrac{2v}{\sqrt{v^2 + \psi(u)}}\right] \end{bmatrix}$.

Let's define the function:

$$\widetilde{F}(x,y,z,u) = \begin{bmatrix} Qx + q + A^T y - \Phi(u,z) \\ Ax - b \\ x + z - \Phi(u,z) \\ u \end{bmatrix},$$

where

$$\Phi(u,z) = \begin{bmatrix} \varphi(u,z_1) \\ \varphi(u,z_2) \\ . \\ . \\ . \\ \varphi(u,z_n) \end{bmatrix}$$

We notice immediately that $(x^*, y^*, z^*)$ solve (3), if only $(x^*, y^*, z^*, 0)$ solve the system

$$\widetilde{F}(x,y,z,u) = 0 \tag{4}$$

Unlike the equation (3), the functions that intervene in the system of equations (4) are differential. Thus the problem (1) was reduced to the solving of the system (4) with $(2n+m+1)$ equations and as many unknowns.

**Newton's method**

For any $d \in R^n$, let $Diag(d_1, d_2, \ldots, d_n)$ denote an $n \times n$ diagonal matrix whose *i*-th diagonal element is $d_i$, $I_{n\times n}$ denotes the $n \times n$ identity matrix and $O_{m\times n}$ denotes the $m \times n$ zero matrix. Then the Jacobean of $\widetilde{F}$ at $(x, y, t, u)$ is given by:

$$\nabla\widetilde{F}(x,y,z,u) = \begin{pmatrix} Q & A^T & D & d \\ A & O_{m\times n} & O_{n\times n} & O_{m\times 1} \\ I_{n\times n} & O_{m\times n} & I_{n\times n}+D & d \\ O_{1\times m} & O_{1\times n} & O_{1\times m} & 1 \end{pmatrix},$$

where

$$D = Diag(w_1, w_2, \ldots, w_n),$$

$$w_i = -\frac{1}{2}\left(1 + \frac{2z_i}{\sqrt{z_i^2 + \psi(u)}}\right),$$

$$d = (d_1, d_2, \ldots, d_n)^T,$$

$$d_i = -\frac{\psi'(u)}{4\sqrt{z_i^2 + \psi(u)}},\ i = 1,2,\ldots,n.$$

**Theorem 2.** *If A is of full rank matrix then for any* $(x, y, z, u)$ *with* $u \neq 0$ *the matrix* $\nabla\widetilde{F}$ *is nonsingular.*

Proof. Let $\Delta p = (\Delta x, \Delta y, \Delta z, \Delta u)$ satisfy $\nabla\widetilde{F}(x,y,z,u)\Delta p = 0$. Then we immediately have: $\Delta u = 0$ and taking in consideration that

$$(A^T\Delta y, \Delta x) = 0,$$

$$Q\Delta x + D\Delta z = 0,$$

$$\Delta x + (I + D)\Delta z = 0.$$

From here we obtain:

$$(Q\Delta x, \Delta x) - \left(D(I+D)^{-1}\Delta x, \Delta x\right) = 0.$$

As the matrix $D(I+D)^{-1}$ is negatively definite for $u \neq 0$, the last relation gives us $\Delta x = 0$, this implies $\Delta z = 0$, $\Delta y = 0$. Thus matrix $\nabla\widetilde{F}$ is non-singular, to solve the system of equations (4). Thus, we can use Newton method. To generate the global convergence the technique proposed in [13] can be used.

**References**


[1] Pshenichnyi B.N., Danilin Yu.M. (1977) *Numerical Methods in Extremal Problems* - Mir Publishers, Moscow.

[2] Pshenichnyi B.N. (1983) *Metod linearizatii* - Nauka, Moscow, (in russian).

[3] *** (1974) *Numerical Methods for Constrained Optimization.* Edited by P.E. Gill and W. Murrey. - Academic Press.

[4] Gill P.E., Murray W., Wright M. (1981) *Practical Optimization*. - Academic Press,.

[5] Bertsekas D. (1982) *Constrained Optimization and Lagrange Multiplier Methods*. - Academic Press.

[6] Bogs P.T., Tolle J.W. (1996) *Sequential quadratic programming.* Acta Numerica, , pp. 1-100.

[7] Gould N.I.M., Toint Ph.L. (2001) *A Quadratic Programming Bibliography*. Numerical Analysis Group Internal Report 200-1 Rutherford Appleton Laboratory, Chilton, England,. http://www.optimization-online.org/DB_4TML/2001/02/285.html

[8] Hiriart-Urruty J.-B. (1998) Optimisation et analyse convexe. Presses Universitaires de France.

[9] Coleman Th.F., Liu J. (1999) *An interior Newton method for quadratic programming.* Mathematical Programming, Ser. A85, pp.491-523.

[10] Moraru V. (1997) *An Algorithm for Solving Quadratic Programming Problems.* Computer Science Journal of Moldova, vol.5, No. 2, pp.223-235.

[11] Bazaraa M.C., Shetty C.M. (1993) *Nonlinear Programming. Theory and Algorithms*. John Wiley & Sons, Inc.

[12] Qi L., Sun D., Zhou G.(2000) *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities* Mathematical Programming, vol. 87, pp. 1–35.

[13] Curtis F., Nocedal J. (2005) *Steplength Selection in Interior-Point Methods for Quadratic Programming.* Technical Report, Optimization Technology Center.