Journal of Convex Analysis
Volume 2 (1995), No.1/2, 263–276

Lagrangian Stability and Global Optimality in Nonconvex Quadratic Minimization Over Euclidean Balls and Spheres

Pham Dinh Tao, Le Thi Hoai An
Mathematical Modelling and Applied Optimization Group,
LMI, INSA Rouen, CNRS, URA 1378.
BP 8, 76131 Mont Saint Aignan Cedex, France.

Received 19 July 1994
Revised manuscript received 28 November 1994

Dedicated to R. T. Rockafellar on his 60th Birthday

We prove in this paper the stability of the Lagrangian duality in nonconvex quadratic minimization over Euclidean balls and spheres. As direct consequences we state both global optimality conditions in these problems and detailed descriptions of the structure of their solution sets. These results are essential for devising solution algorithms.

Keywords : nonconvex quadratic minimization, stability of Lagrangian duality, global optimality conditions.

1. Introduction

In this paper we shall be concerned with the following nonconvex quadratic optimization problems

$$(Q_1) \quad \min\{\frac{1}{2}x^T Ax + b^T x : \quad \|x\| \leq r\}$$

$$(Q_2) \quad \min\{\frac{1}{2}x^T Ax + b^T x : \quad \|x\| = r\}$$

where A is an $n \times n$ real symmetric matrix, $b \in I\!R^n$, r is a positive number. If A is positive semi-definite then (Q_1) is a convex quadratic problem. In general (Q_1) is nonconvex. Problem (Q_2), whose feasible domain is a sphere, is always nonconvex even if A is positive semi-definite. These are among the few nonconvex optimization problems which possess a complete characterization of their optimal solutions. These problems play an important role in optimization and numerical analysis ([1], [4], [8], [9], [11], [12], [16]). Golub et al. studied (Q_2) from both theoretical and computational viewpoints. In [3] the sensitivity of the solutions to primal problem was discussed.

ISSN 0944-6532 / \$ 2.50

Very recently, a characterization of local-nonglobal minimizers of (Q_1) and (Q_2) are investigated in [7].

In this paper we study the stability of the Lagrangian duality and global optimality relative to (Q_1) and (Q_2) in the case where their respective constraints are analytically expressed as $\{x \in \mathbb{R}^n : \frac{1}{2}\|x\|^2 \leq \frac{1}{2}r^2\}$ and $\{x \in \mathbb{R}^n : \frac{1}{2}\|x\|^2 = \frac{1}{2}r^2\}$ (although these are one and the same problem). Their original analytical expressions will lead to a duality gap equal to infinity. By using Lagrangian duality we obtain some basic properties on the solutions of (Q_1), (Q_2) and their dual problems. We show that these nonconvex problems have no duality gap. The optimality conditions obtained by different ways in Gay [4], Moré-Sorensen [9], Fletcher ([2]) and Pham Dinh Tao et al. ([1], [10], [12]) immediately follow from these results.

This paper constitutes the first part (Part I) of our work relative to nonconvex quadratic minimization over Euclidean balls and spheres. It is organized as follows. In the next section we collect some important problems which can be formulated in the forms of (Q_1) or (Q_2). The last section is dealing with the stability of Lagrangian duality and global optimality in (Q_1) and (Q_2). As direct consequences, complete characterizations of solutions for (Q_1) and (Q_2) are pointed out together with detailed descriptions of the structure of solution sets for (Q_1) and (Q_2). These results are essential for solution algorithms.

2. Examples

In this section we formulate the following problems in the forms of (Q_1) or (Q_2).

$$\min\{\frac{1}{2}x^T Ax + b^T x : Bx = c, \|x\| = 1\} \quad \text{(Constrained Eigenvalue Problem, [3])} \tag{2.1}$$

where B is $(m \times n)$ matrix $(m < n)$ and $c \in I\!R^m$,

$$\min\{\|Ax - b\| : \|Cx - d\| = r\} \quad \text{(Quadratically Constrained Least Squares Problem, [6])} \tag{2.2}$$

$$\min\{\|Ax - b\| : \|Cx - d\| \leq r\} \tag{2.3}$$

where A is a $(m \times n)$ matrix, C is a $(p \times n)$ matrix, $b \in I\!R^m$ and $d \in I\!R^p$,

$$\min\{\frac{1}{2}x^T Ax + b^T x : \frac{1}{2}x^T Cx + d^T x \leq r^2\} \tag{2.4}$$

where A, C are $(n \times n)$ symmetric positive semi-definite matrices.

For problem (2.1) we assume that it has at least one solution and that $rank(B) = m$. Let Z be an orthogonal basis of Ker(B). Let $y \in Im(B^T)$ such that $By = c$ (since $rank(B) = m$, such a point y always exists). Thus for each $x \in I\!R^n$ there exists $z \in I\!R^m$ such that $x = y + Zz$. By a simple calculation we can show that (2.1) is equivalent to the following problem of the form (Q_2):

$$\min\{\frac{1}{2}z^T A'z + b'^T z : \quad \|z\| = s\}$$

where $A' = Z^TAZ$, $b' = Z^T(Ay+c)$ and $s^2 = 1-\|y\|^2$. Note that, since (2.1) has at least one solution, $\|y\| \leq 1$. Now we show that (2.4) is of the form of (Q_1). We distinguish two cases:

Case 1. C is positive definite. In this case the equation $Cx = -d$ has a unique solution $w = -C^{-1}d$ which minimizes the function

$$\frac{1}{2}x^TCx + d^Tx.$$

By Cholesky's factorization we have $C = R^TR$ ([5]). Using the variable $y = R(x-w)$ we can write (2.4) in the form

$$\min\{\frac{1}{2}y^TA'y + b'^Ty : \quad \|y\|^2 \leq (1/2)s^2\}$$

where $A' = (R^{-1})^TAR, b' = R^{-1}(Aw+b)$ and $s^2 = r^2 + (1/2)w^TCw$.

Case 2. C is positive semi-definite and $d \in Im(C) = \mathrm{Ker}(C^T)$. In this case we solve the equation $Cx = -d$ by procedure QR ([5]). Suppose $rank(C) = m$, then we can find a submatrix C_J of rank m such that the equation $Cx = -d$ is equivalent to $C_Jx = -d_J$. This submatrix is given by

$$C_J^T = [\bar{Q}_1, \bar{Q}_2]\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = \bar{Q}_1R_1$$

where R_1 is an $m \times m$ upper-triangular matrix. Noting that $\bar{Q}_1$ is a basis of $Im(C_J^T)$ and $\bar{Q}_2$ is basis of $\mathrm{Ker}(C_J) = \mathrm{Ker}(C)$ we can write

$$x = \bar{Q}_1y + \bar{Q}_2z, \quad y \in \mathbb{R}^m, \; z \in \mathbb{R}^{n-m}.$$

Let $y^* = R_1^{-T}d_J$. Then it is clear that $w = \bar{Q}_1y^*$ is a solution to the equation $Cx = -d$ which minimizes the form

$$\frac{1}{2}x^TCx + d^Tx.$$

Moreover $\bar{Q}_1y^* \in Im(C_J^T)$ has the minimal norm among the elements in $Im(C_J^T)$.
Problem (2.4) can thus be written as:

$$\min\{q(y,z) = [\frac{1}{2}y^T\bar{Q}_1^TA\bar{Q}_1y + b^T\bar{Q}_1y] \; + \; [\frac{1}{2}z^T\bar{Q}_2^TA\bar{Q}_2z + b^T\bar{Q}_2z] \; + \; z^T\bar{Q}_2^TA\bar{Q}_1y\}$$

subject to

$$y \in U = \{y \in \mathbb{R}^m : (y-y^*)^T\bar{Q}_1^TC\bar{Q}_1(y-y^*) \leq s^2\}, \; z \in \mathbb{R}^{n-m},$$

where $s^2 = r^2 + 1/2y^{*T}\bar{Q}_1^TC\bar{Q}_1y^*$. Note that $\bar{Q}_1^TC\bar{Q}_1$ is positive definite and $q(y,z)$ is a convex function of (y,z). Hence we have

$$\min_{y \in U, z \in \mathbb{R}^{n-m}} q(y,z) = \min_{y \in U} h(y)$$

where

$$h(y) = \min\{q(y,z) : z \in \mathbb{R}^{n-m}\} \tag{2.5}$$

is a convex function.
For each fixed y, z solves (2.5) if and only if

$$\bar{Q}_2^T A \bar{Q}_2 z \;+\; \bar{Q}_2^T b = -\bar{Q}_2^T A \bar{Q}_1 y. \tag{2.6}$$

For simplicity we assume that A is positive definite (if A is positive semi- definite we can use the pseudo-inverse of A to obtain an analogous formula). In this case the solution of (2.5) takes the form

$$z = -(\bar{Q}_2^T A \bar{Q}_2)^{-1}(\bar{Q}_2^T b + \bar{Q}_2^T A \bar{Q}_1 y).$$

Hence

$$h(y) = \frac{1}{2} y^T G y + g^T y$$

where

$$\begin{aligned} G &= \bar{Q}_1^T [A - A\bar{Q}_2(\bar{Q}_2^T A \bar{Q}_2)^{-1} \bar{Q}_2^T A] \bar{Q}_1 \\ g &= [\bar{Q}_1^T - \frac{1}{2} \bar{Q}_1^T A \bar{Q}_2 (\bar{Q}_2^T A \bar{Q}_2)^{-1} \bar{Q}_2^T] b. \end{aligned}$$

Finally, setting $u = y - y^*$ it is easy to verify that (2.4) is equivalent to the problem given by

$$\min\{\frac{1}{2} u^T A' u + b'^T u : u^T C' u \le s^2\}$$

where A' is positive semi-definite and C' is positive definite. The result then follows from case 1.
Note that (2.3) is a special case of (2.4).
In a similar way we can formulate problem (2.2) in the form of (Q_2).

3. Stability of the Lagrangian duality and global optimality conditions for Problems $(\boldsymbol{Q}_1)$ and $(\boldsymbol{Q}_2)$

3.1. Stability of the Lagrangian duality for problem $(\boldsymbol{Q}_1)$

In the first two parts of this section we study Lagrangian duality for Problems (Q_1) and (Q_2) by writing their constraints in the equivalent forms $\{x \in \mathbb{R}^n : \frac{1}{2}\|x\|^2 \le \frac{1}{2} r^2\}$ and $\{x \in \mathbb{R}^n : \frac{1}{2}\|x\|^2 = \frac{1}{2} r^2\}$. We shall establish a characterization for the solutions of these problems and their duals. Stability of Lagrangian duality and optimality conditions can be easily derived from these results.
Denote by $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ the eigenvalues of A and by $u_1, u_2 \ldots, u_n$ their corresponding eigenvectors which constitute an orthogonal basis of $\mathbb{R}^n$. It is easy to see that for $\lambda > -\lambda_1$ the solution to the linear equation

$$(A + \lambda I)x = -b. \tag{3.1}$$

is given by

$$u_i^T x(\lambda) = \frac{-u_i^T b}{\lambda_i + \lambda} \quad i = 1, \dots, n. \tag{3.2}$$

Let us describe now (Q_1) in the equivalent form which, for simplicity of notation, we also denote by (Q_1)

$$\alpha_1 = \min\{f(x) = \frac{1}{2}x^T Ax + b^T x : \quad \frac{1}{2}\|x\|^2 \leq \frac{1}{2}r^2\}. \tag{Q_1}$$

The Lagrangian function for this problem is defined by

$$L_1(x, \lambda) = \begin{cases} \frac{1}{2}x^T Ax + b^T x + \frac{\lambda}{2}(\|x\|^2 - r^2) & \text{if } \lambda \geq 0 \\ -\infty & \text{otherwise} . \end{cases}$$

For each fixed $\lambda \geq 0$ we define the problem

$$g_1(\lambda) = \inf_{x \in I\!R^n}\{\frac{1}{2}x^T(A + \lambda I)x + b^T x - \frac{\lambda}{2}r^2\}. \tag{Q^1_λ}$$

Clearly, g_1 is concave.
The dual problem of (Q_1) is given as

$$\beta_1 = \sup\{g_1(\lambda) : \quad \lambda \geq 0\}. \tag{D_1}$$

We denote by $\mathcal{Q}_1$, $\mathcal{Q}^1_\lambda$,$\mathcal{D}_1$ the solution sets of (Q_1), (Q^1_λ),(D_1) respectively. First observe the following:

- If $\lambda > -\lambda_1$ then (3.1) admits a unique solution given by

$$x(\lambda)^T u_i = -\frac{b^T u_i}{\lambda + \lambda_i}, \ i = 1, \dots, n \ \text{ and } \|x(\lambda)\|^2 = \sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda + \lambda_i)^2} \tag{3.3}$$

- If $\lambda = -\lambda_1$ and $b \in \text{Ker}(A - \lambda_1 I)^\perp$ then $\mathcal{P}^1_\lambda = x^+ + \text{Ker}(A - \lambda_1 I)$ where $x^+ = -(A - \lambda_1 I)^+ b$ given by $((A - \lambda_1 I)^+$ stands for the pseudo-inverse of $(A - \lambda_1 I))$

(i) If $b = 0$ then $x^+ = -(A - \lambda_1 I)^+ b = 0$.
(ii) If $b \neq 0$ then the complement in $\{1, \dots, n\}$ of $J_1 = \{i = 1, \dots, n : \lambda_i = \lambda_1\}$ is nonempty and

$$x^{+T} u_i = -\frac{b^T u_i}{\lambda + \lambda_i}, \ i \notin J_1 \text{ and } \|x^+\|^2 = \sum_{i \notin J_1} \frac{(b^T u_i)^2}{(\lambda + \lambda_i)^2}. \tag{3.4}$$

Proposition 3.1.
(i) $\text{dom}\, g_1 = \{\lambda \in \mathrm{I\!R} : \lambda \geq 0, \lambda + \lambda_1 \geq 0$ *and* $b \in Ker(A + \lambda I)^\perp\}$.
(ii)

$$g_1(\lambda) \ = \ \frac{1}{2}b^T x - \frac{r^2}{2}\lambda \ = \ -\frac{1}{2}b^T(A + \lambda I)^+ b - \frac{r^2}{2}\lambda, \ \ \forall x \in \mathcal{Q}^1_\lambda, \ \forall \lambda \in \text{dom}\, g_1.$$

Proof. $\operatorname{dom} g_1 := \{\lambda \geq 0 : g_1(\lambda) > -\infty\}$.
(i) If $\lambda + \lambda_1 < 0$ then $\lambda \notin \operatorname{dom} g_1$. Indeed $\bar{x} \in \operatorname{Ker}(A - \lambda_1 I)$ implies

$$\lim_{\|\bar{x}\| \to \infty} \langle (A + \lambda I)\bar{x}, \bar{x}\rangle + \langle b, \bar{x}\rangle = \lim_{\|\bar{x}\| \to \infty} (\lambda + \lambda_1)\|\bar{x}\|^2 + \langle b, \bar{x}\rangle = -\infty.$$

Hence $g_1(\lambda) = -\infty$.
If $\lambda \geq 0$ and $\lambda + \lambda_1 \geq 0$, then (Q^1_λ) is a convex problem and $\bar{x} \in \mathcal{Q}^1_\lambda \Leftrightarrow (A + \lambda I)\bar{x} = -b$. Thus $\lambda \in \operatorname{dom} g_1$ if and only if (3.1) has a solution, i.e., $b \in \operatorname{Ker}(A + \lambda I)^\perp$.
(ii) This is immediate from the definition of g_1 and the fact

$$\forall \lambda \in \operatorname{dom} g_1 : \mathcal{Q}^1_\lambda = (A + \lambda I)^+ b + \operatorname{Ker}(A + \lambda I).$$

□

Set $\tilde{g}_1(\lambda) = -g_1(\lambda)$. Clearly $\tilde{g}_1$ is convex and $\operatorname{dom} \tilde{g}_1 = \operatorname{dom} g_1$.

The following proposition is a direct consequence of Proposition 3.1 and (3.3), (3.4):

Proposition 3.2.
(i) *If $\lambda \geq 0$ and $\lambda + \lambda_1 > 0$ then $\lambda \in \operatorname{dom} \tilde{g}_1$, and*

$$\tilde{g}_1(\lambda) = \frac{1}{2} \sum_{i=1}^{n} \frac{(b^T u_i)^2}{\lambda_i + \lambda} + \frac{r^2}{2}\lambda. \tag{3.5}$$

In this case $\tilde{g}_1$ is differentiable for every $\lambda > 0$, and

$$\tilde{g}_1'(\lambda) = -\frac{1}{2}\|x(\lambda)\|^2 + \frac{r^2}{2}. \tag{3.6}$$

In particular, if $\lambda_1 > 0$ then $\operatorname{dom} \tilde{g}_1 = [0, \infty[$ and (3.5) is satisfied for every $\lambda \in \operatorname{dom} \tilde{g}_1$. In this case $\tilde{g}_1$ is subdifferentiable at 0, and

$$\partial \tilde{g}_1(0) =]-\infty, -\frac{1}{2}\|x(0)\|^2 + \frac{r^2}{2}].$$

(ii) *If $\lambda_1 \leq 0$ and $b \in Ker(A - \lambda_1 I)^\perp$ then $\operatorname{dom} \tilde{g}_1 = [-\lambda_1, +\infty[$, and for every $\lambda \in \operatorname{dom} \tilde{g}_1$ we have:*
- *If $b = 0$ then $\tilde{g}_1(\lambda) = \frac{r^2}{2}\lambda$.*
- *If $b \neq 0$ then $\{1, \ldots, n\} \backslash J_1$ is nonempty and*

$$\tilde{g}_1(\lambda) = \frac{1}{2} \sum_{i \notin J_1} \frac{(b^T u_i)^2}{\lambda_i + \lambda} + \frac{r^2}{2}\lambda.$$

In this case $\tilde{g}_1$ is subdifferentiable at $-\lambda_1$, and

$$\partial \tilde{g}_1(-\lambda_1) =]-\infty, -\frac{1}{2}\|x^+\|^2 + \frac{r^2}{2}].$$

(iii) *If* $\lambda_1 \leq 0$ *and* $b \notin Ker(A - \lambda_1 I)^\perp$, *then* $\operatorname{dom}\tilde{g}_1 =]-\lambda_1, +\infty[$ *and for each* $\lambda \in \operatorname{dom}\tilde{g}_1$ *the function* $\tilde{g}_1(\lambda)$ *is defined by* (3.5).
In this case $\tilde{g}_1$ *is differentiable at every* $\lambda \in \operatorname{dom}\tilde{g}_1$, *and* $\tilde{g}_1'$ *is defined by* (3.6).

Now we give some results concerning the characterizations of the solution to the dual problem.

Theorem 3.3.
(i) $\mathcal{D}_1$ *is a singleton.*
(ii) *Let* $\lambda^* \geq 0$, $\lambda^* > -\lambda_1$. *Then* $\lambda^* \in \mathcal{D}_1$ *if and only if* $\|x(\lambda^*)\| = r$. *In particuliar, if* $\lambda_1 > 0$ *then* $0 \in \mathcal{D}_1$ *if and only if* $\|x(0)\| \leq r$.
(iii) *Let* $\lambda^* = -\lambda_1 \geq 0$. *Then* $\lambda^* \in \mathcal{D}_1$ *if and only if* $b \in Ker(A - \lambda_1 I)^\perp$ *and* $\|x^+\| \leq r$.

Proof. First we remark that (D_1) is a convex program having always one solution, since $\tilde{g}_1$ is coercive. In fact

$$(D_1) \Leftrightarrow (\tilde{D}_1) \qquad \inf_{\lambda \geq 0} \tilde{g}_1(\lambda) = \inf_{\lambda \in \operatorname{dom}\tilde{g}_1} \tilde{g}_1(\lambda)$$

and by (3.5), $\lim_{\lambda \to +\infty} \tilde{g}_1(\lambda) = +\infty$. Again by (3.5): $\tilde{g}_1''(\lambda) > 0$, $\forall \lambda > -\lambda_1$, i.e., $\tilde{g}_1$ is strictly convex in $]-\lambda_1, +\infty[$, consequently (D_1) contains one element. Hence (i).
(ii) and (iii) are immediate from Proposition 3.2 and from the fact $\lambda^* \in \mathcal{D}_1 \Leftrightarrow 0 \in \partial\tilde{g}_1(\lambda^*)$. □

The following corollary shows the stability of the duality for (Q_1):

Corollary 3.4. $\alpha_1 = \beta_1$ *and*

$$\mathcal{Q}_1 = \{x^* \in \mathcal{Q}^1_{\lambda^*} : \ \lambda^*(\|x^*\| - r) = 0,\ \|x^*\| \leq r\},\ \mathcal{D}_1 = \{\lambda^*\}.$$

Proof. Let $E = \{x \in X : \|x\| \leq r\}$ and χ_E be its indicator function, i.e. $\chi_E(x) = 0$ if $x \in E, \chi_E(x) = +\infty$ otherwise. From the definition of $L_1(x, \lambda)$ we have $\sup\{L_1(x, \lambda) : \lambda \geq 0\} = (f + \chi_E)(x)$. Hence

$$\alpha_1 = \inf_{x \in X}(f + \chi_E)(x) = \inf_{x \in X} \sup_{\lambda \geq 0} L_1(x, \lambda) \geq \sup_{\lambda \geq 0} \inf_{x \in X} L(x, \lambda) = \sup_{\lambda \geq 0} g_1(\lambda) = \beta_1.$$

Therefore to prove the Corollary it is sufficient to point out a point $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}$ satisfying $L_1(x^*, \lambda) \leq L_1(x^*, \lambda^*) \leq L_1(x, \lambda^*)$, $\forall (x, \lambda) \in X \times \mathbb{R}$, i.e. (x^*, λ^*) is a saddle point of L_1.
Let $\lambda^* \in \mathcal{D}_1$. If $\lambda^* > -\lambda_1$ we have from Theorem 3.3 $\|x(\lambda^*)\| = r$ (or $\|x(\lambda^*)\| \leq r$ if $\lambda^* = 0$). Thus in this case $\mathcal{Q}^1_{\lambda^*}$ contains a unique point x^* which satisfies the complementarity condition (CP) $\lambda^*(\|x^*\| - r) = 0$.
If $\lambda^* = -\lambda_1$, again by Theorem 3.3, $\|x^+\| \leq r$ and $b \in \operatorname{Ker}(A - \lambda_1 I)^\perp$ which implies

$$\mathcal{Q}^1_{\lambda^*} = x^+ + \operatorname{Ker}(A - \lambda_1 I).$$

Hence we can choose $x^* \in \mathcal{Q}^1_{\lambda^*}$, $\|x^*\| = r$ which implies $\lambda^*(\|x^*\| - r) = 0$.

Consequently, in both cases we have

$$\alpha_1 \ \leq \ \frac{1}{2}x^{*T}Ax^* + b^T x^* \ = \ \frac{1}{2}x^{*T}Ax^* + b^T x^* + \frac{\lambda^*}{2}(\|x^*\|^2 - r^2) \ = \ \beta_1.$$

□

Remark 3.5. We remark that in the above presentation the form $\{x \in \mathbb{R}^n : \frac{1}{2}\|x\|^2 \leq \frac{1}{2}r^2\}$ is essential. If we take the equivalent constraint $\{x \in \mathbb{R}^n : \|x\| \leq r\}$ the dual optimal value will be $-\infty$ if $\lambda_1 < 0$. Indeed, in this case the function $g_1(\lambda)$ is defined by

$$g_1(\lambda) = \inf_{x \in \mathbb{R}^n} \{\frac{1}{2}\langle x, Ax\rangle + \langle x, b\rangle + \frac{1}{2}\lambda(\|x\| - r)\}, \text{ for } \lambda \geq 0.$$

Then if $\bar{x} \in \mathrm{Ker}(A - \lambda_1 I)$ we have

$$\lim_{\|\bar{x}\| \to +\infty} \frac{1}{2}\langle \bar{x}, A\bar{x}\rangle + \langle \bar{x}, b\rangle + \frac{1}{2}\lambda(\|\bar{x}\| - r) = \lim_{\|\bar{x}\| \to +\infty} \frac{1}{2}\lambda_1\|\bar{x}\|^2 + \langle \bar{x}, b\rangle + \frac{1}{2}\lambda(\|\bar{x}\| - r) = -\infty$$

and therefore $g_1(\lambda) = -\infty$ for every $\lambda \geq 0$. Hence $\beta_1 = -\infty$.

3.2. Stability of the Lagrangian duality for Problem $(\boldsymbol{Q}_2)$

The Lagrangian function of the problem

$$\alpha_2 = \min\{\frac{1}{2}x^T Ax + b^T x : \quad \frac{1}{2}\|x\|^2 = \frac{1}{2}r^2\} \tag{Q_2}$$

is

$$L_2(x, \lambda) = \frac{1}{2}x^T Ax + b^T x + \frac{\lambda}{2}(\|x\|^2 - r^2), \ \lambda \in \mathbb{R}.$$

Let $g_2(\lambda)$ be the function on $\mathbb{R}$ given by

$$g_2(\lambda) = \inf\{L_2(x, \lambda) : \quad x \in \mathbb{R}^n\}. \tag{Q_λ^2}$$

Then the dual problem of (Q_2) takes the form

$$\beta_2 = \sup\{g_2(\lambda) : \quad \lambda \in \mathbb{R}\}. \tag{D_2}$$

As before we denote by $\mathcal{Q}_2$, $\mathcal{Q}_\lambda^2$ and $\mathcal{D}_2$ the solution sets of problems (Q_2), (Q_λ^2) and (D_2) respectively. Let $\tilde{g}_2(\lambda) = -g_2(\lambda), \forall \lambda \in \mathbb{R}$.

In a similar way one can show the following results concerning the stability of the Lagrangian duality for problem (Q_2). The first ones correspond to Proposition 3.1 and 3.2.

Proposition 3.6.

(i) $\mathrm{dom}\, g_2 = \{\lambda \in \mathbb{R} : \lambda \geq -\lambda_1, \ \textit{and}\, b \in \textit{Ker}(A + \lambda I)^\perp\}$.

(ii)

$$g_2(\lambda) = \frac{1}{2}b^T x - \frac{r^2}{2}\lambda = -\frac{1}{2}b^T(A + \lambda I)^+ b - \frac{r^2}{2}\lambda, \ \forall x \in \mathcal{Q}_\lambda^2.$$

Proposition 3.7.
(i) *If* $b \notin Ker(A - \lambda_1 I)^{\perp}$ *then* $\operatorname{dom}\tilde{g}_2 =]-\lambda_1, +\infty[$, *and* $\forall\lambda \in \operatorname{dom}\tilde{g}_2$

$$\tilde{g}_2(\lambda) = \frac{1}{2}\sum_{i=1}^{n}\frac{(b^T u_i)^2}{\lambda_i + \lambda} + \frac{r^2}{2}\lambda.$$

(ii) *If* $b \in Ker(A - \lambda_1 I)^{\perp}$ *then* $\operatorname{dom}\tilde{g}_2 = [-\lambda_1, +\infty[$ *and for every* $\lambda \in \operatorname{dom}\tilde{g}_2$ *we have:*

- *If* $b = 0$ *then* $\tilde{g}_2(\lambda) = \frac{r^2}{2}$.
- *If* $b \neq 0$ *then* $\{1, \ldots, n\}\backslash J_1$ *is nonempty and*

$$\tilde{g}_2(\lambda) = \frac{1}{2}\sum_{i\notin J_1}\frac{(b^T u_i)^2}{\lambda_i + \lambda} + \frac{r^2}{2}\lambda.$$

In this case $\tilde{g}_2$ *is subdifferentiable at* $-\lambda_1$, *and*

$$\partial\tilde{g}_2(\lambda_1) =]-\infty, -\frac{1}{2}\|(A - \lambda_1 I)^{+} b\|^2 + \frac{r^2}{2}].$$

(iii) $\forall\lambda > -\lambda_1$ *we have* $\lambda \in \operatorname{dom}\tilde{g}_2$, $\tilde{g}_2$ *is differentiable at* λ, *and*

$$\tilde{g}_2'(\lambda) = -\frac{1}{2}\|x(\lambda)\|^2 + \frac{r^2}{2}.$$

Finally we state below the analogues of Theorem 3.3 and its Corollary.

Theorem 3.8.
(i) $\mathcal{D}_2$ *is a singleton.*
(ii) *Let* $\lambda^* > -\lambda_1$. *Then* $\lambda^* \in \mathcal{D}_2$ *if and only if* $\|x(\lambda^*)\| = r$.
(iii) *Let* $\lambda^* = -\lambda_1$. *Then* $\lambda^* \in \mathcal{D}_2$ *if and only if* $b \in Ker(A - \lambda_1 I)^{\perp}$ *and* $\|x^*\| \leq r$

Corollary 3.9. $\alpha_2 = \beta_2$; $\mathcal{Q}_2 = \{x^* \in \mathcal{Q}^2_{\lambda^*} : \|x^*\| = r\}$, $\mathcal{D}_2 = \{\lambda^*\}$.

3.3. Global optimality conditions for $(\boldsymbol{Q}_1)$ and $(\boldsymbol{Q}_2)$

The preceding results concerning the stability of Lagrangian duality relative to (Q_1) and (Q_2) imply, as direct consequences, the following two theorems on global optimality conditions for these problems.

Theorem 3.10. ([4]) x^* *is an optimal solution to Problem* (Q_1) *if and only if there exists* $\lambda^* \geq 0$ *such that*
(i) $(A + \lambda^* I)$ *is positive semi-definite,*
(ii) $(A + \lambda^* I)x^* = -b$,
(iii) $\lambda^*(\|x^*\| - r) = 0, \|x^*\| \leq r$. *This* λ^* *is unique.*

We note that (Q_2) is equivalent to the following problem of the form (Q_1):

$$\min\{\frac{1}{2}x^T(A + \gamma I)x + b^T x \quad : \quad \|x\| \leq r\}, \tag{3.8}$$

where γ is a real number so that $A + \gamma I$ is nonpositive semi- definite.
Indeed, (Q_2) is equivalent to

$$\min\{\frac{1}{2}x^T(A+\gamma I)x + b^T x \quad : \quad \|x\| = r\} \tag{3.9}$$

because $\langle x, \gamma x\rangle = \gamma r^2$ if $\|x\| = r$.
On the other hand, if $A + \gamma I$ is nonpositive semi-definite, then the solution set of (3.1) is necessarily contained in the boundary $\{x : \quad \|x\| = r\}$, according to Theorem 3.10. It implies the equivalence of Problems (3.1) and (3.2).
In [16] Sorensen gave a sufficient condition for optimality to problem (Q_2). In [10] Pham Dinh Tao proved that this condition is necessary and sufficient. This result was also given partly in Fletcher ([2]) by using the optimality condition for (Q_1).

Theorem 3.11. ([10], [2], [12]) *x^* is an optimal solution to (Q_2) if and only if there exists λ^* such that*
(i) *$(A + \lambda^* I)$ is positive semi-definite*
(ii) *$(A + \lambda^* I)x^* = -b$,*
(iii) *$\|x^*\| = r$. This number λ^* is unique.*

3.4. More on the structure of solution sets for $(\boldsymbol{Q}_1)$ and $(\boldsymbol{Q}_2)$

Let us consider now the following two functions ϕ and ψ definied on $\mathbb{R}\backslash\{-\lambda_i : i = 1, \dots, n\}$

$$\phi(\lambda) = \|x(\lambda)\|$$

where $x(\lambda)$ is the solution of $(A + \lambda I)x = -b$,

$$\psi(\lambda) = \frac{1}{r} - \frac{1}{\phi(\lambda)} \text{ if } b \neq 0.$$

Numerical approaches by Moré & Sorensen [9], Pham Dinh Tao [10] – [12] and Golub et al [6] for solving (Q_1) and (Q_2) involve the following equation (referred to as secular equation by Golub et al. [6] in case there is $\lambda^o \geq 0, \lambda^o > -\lambda_1$ such that $\phi(\lambda^o) \geq r$ (such a case implies $b \neq 0$):

$$\phi(\lambda) = r. \tag{3.10}$$

The Safeguarding by Moré & Sorensen and the Dichotomy algorithm use, for solving (3.10), an adapted Newton-like algorithm (called Hebden algorithm) which is nothing else than the Newton method applied to solving

$$\psi(\lambda) = 0.$$

The hard case corresponds to the situation where $\lambda^* = -\lambda_1$. One must then adapt other techniques to compute primal and dual solutions.
The stability of Lagrangian duality relative to (Q_1) and the main properties of ϕ, ψ given below show that solving (3.10) amounts to computing the dual solution λ^* (of (D_1)) and the corresponding primal solution $x(\lambda^*)$ (of (Q_1)). Moreover in the hard case, ϕ is multivalued at $\lambda^* = -\lambda_1$ since $\tilde{g}_1$ is no longer differentiable at $-\lambda_1$.

Proposition 3.12. *Assume that $b \neq 0$, then*
(i) $\phi \in C^\infty(\mathbb{R}\backslash\{-\lambda_i : i = 1, \ldots, n\})$ *is strictly convex on each open interval contained in $\mathbb{R}\backslash\{-\lambda_i : i = 1, \ldots, n\}$ and strictly decreasing on $(-\lambda_1, +\infty)$. More precisely one has, for $\lambda \notin \{-\lambda_i : i = 1, \ldots, n\}$,*

$$\phi'(\lambda) = -\frac{1}{\phi(\lambda)} \sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} \tag{3.11}$$

$$\phi''(\lambda) = \frac{2}{\phi(\lambda)} \sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} + \frac{1}{\phi(\lambda)} \left\{ \sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^4} - \frac{1}{\phi^2(\lambda)} \left[\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} \right]^2 \right\} \tag{3.12}$$

If $b \notin \mathbb{N}(A - \lambda_1 I)^\perp$ then $\lim_{\lambda \downarrow -\lambda_1} \phi(\lambda) = +\infty$.
If $b \in \mathbb{N}(A - \lambda_1 I)^\perp$ then $\phi(\lambda) = \|(A - \lambda_1 I)^+ b\|$ for $\lambda > -\lambda_1$ and $\lim_{\lambda \downarrow -\lambda_1} \phi(\lambda) = \|(A - \lambda_1 I)^+ b\|$.
(ii) $\psi \in C^\infty(\mathbb{R}\backslash\{-\lambda_i : i = 1, \ldots, n\})$ *is strictly convex on each open interval contained in $\mathbb{R}\backslash\{-\lambda_i : i = 1, \ldots, n\}$ and strictly decreasing on $(-\lambda_1, +\infty)$. More precisely, if $\theta(\lambda) = -1/\phi(\lambda) = \psi(\lambda) - (1/r)$ then one has for $\lambda \notin \{-\lambda_i : i = 1, \ldots, n\}$*

$$\theta'(\lambda) = \phi(\lambda)^{-2} \phi'(\lambda), \tag{3.13}$$

$$\theta''(\lambda) = \frac{3}{\phi(\lambda)^3} \left\{ \sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^4} - \frac{1}{\phi(\lambda)^2} \left[\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} \right]^2 \right\}. \tag{3.14}$$

(iii) *For $\lambda \in \operatorname{dom} \tilde{g}_1$ such that $\lambda \geq -\lambda_1$, one has*

$$\phi^2(\lambda) = \frac{r^2}{2} - \tilde{g}_1(\lambda), \ if \ \lambda > -\lambda_1.$$

So $\phi(\lambda)$ can be extended at $\lambda = -\lambda_1 \in \operatorname{dom} \tilde{g}_1$ by

$$\phi^2(-\lambda_1) = [\|(A - \lambda_1 I)^+ b\|^2, +\infty).$$

Proof. It suffices to prove that $\phi''(\lambda) \geq 0$ and $\theta''(\lambda) \geq 0$ for $\lambda \notin \{-\lambda_i : i = 1, \ldots, n\}$. The other results are either direct computational derivatives or direct consequences of Proposition 2. For this we shall prove that

$$\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^4} - \frac{1}{\phi^2(\lambda)} [\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3}]^2 \geq 0$$

for $\lambda \notin \{-\lambda_i : i = 1, \ldots, n\}$.
We can write for such a λ

$$\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} = \sum_{i=1}^{n} \frac{b^T u_i}{(\lambda_i + \lambda)^2} \frac{b^T u_i}{\lambda_i + \lambda}$$

So by Schwarz's inequality we obtain

$$\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^3} \leq [\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^4}]^{1/2} [\sum_{i=1}^{n} \frac{(b^T u_i)^2}{(\lambda_i + \lambda)^2}]^{1/2}$$

that is equivalent to the above inequality.

Remark 3.13. If $b = 0$ then $\phi(\lambda) = 0$ for every $\lambda \notin \{-\lambda_i : i = 1, \dots, n\}$ and $\psi(\lambda)$ is no longer defined for such λ. In this case the solution set of (Q_1) will be given in Proposition 3.14

Now let us deduce from the preceding results some direct consequences about the structure of primal and dual solutions of (Q_1) (resp. (Q_2)) and (D_1) (resp. (D_2)). These properties are of great usefulness for their solution methods.

Proposition 3.14.
1. *Assume that* $b \neq 0$.
(i) *If* $\lambda_1 > 0$ *(i.e. A is positive definite) one has*

- *If* $\|A^{-1}b\| \leq r$ *then* $\mathcal{D}_1 = \{\lambda^* = 0\}$ *and* $\mathcal{Q}_1 = \{x(0) = -A^{-1}b\}$.
- *If* $\|A^{-1}b\| = r$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{\lambda^* = 0\}$ *and* $\mathcal{Q}_1 = \mathcal{Q}_2 = \{x(0) = -A^{-1}b\}$.
- *If* $\|A^{-1}b\| < r$ *then*

(a) *If* $b \notin Ker(A - \lambda_1 I)^\perp$ *or* $b \in Ker(A - \lambda_1 I)^\perp$ *and* $\|(A - \lambda_1 I)^+ b\| > r$ *then* $\mathcal{D}_2 = \{\lambda^*\}$ *with* $-\lambda_1 < \lambda^* < 0$ *and* $\mathcal{Q}_2 = \{x(\lambda^*) = -(A + \lambda^* I)^{-1} b\}$, *where* λ^* *is the unique solution of the equation*

$$\phi(\lambda) = r. \tag{3.15}$$

(b) *If* $b \in Ker(A - \lambda_1 I)^\perp$ *and* $\|(A - \lambda_1 I)^+ b\| \leq r$ *then* $\mathcal{D}_2 = \{\lambda^* = -\lambda_1\}$ *and*

$$\mathcal{Q}_2 = \{x = -x^+ + u : \|x\|^2 = \|x^+\|^2 + \|u\|^2 = r^2\}.$$

- *If* $\|A^{-1}b\| > r$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{\lambda^* > 0\}$ *where* λ^* *is the unique solution of (3.15) and* $\mathcal{Q}_1 = \mathcal{Q}_2 = \{x(\lambda^*) = -(A + \lambda^* I)^{-1} b\}$.

(ii) *If* $\lambda_1 = 0$ *(i.e. A is positive semi-definite) one has*

- *If* $\|A^+ b\| \leq r$ *and* $b \in Ker(A)^\perp$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{0\}$

$$\mathcal{Q}_1 = \{x = -x^+ + u, \quad u \in Ker(A) \text{ such that } \|x\|^2 = \|x^+\|^2 + \|u\|^2 \leq r^2\},$$

$$\mathcal{Q}_2 = \{x = -x^+ + u, \quad u \in Ker(A) \text{ such that } \|x\|^2 = \|x^+\|^2 + \|u\|^2 = r^2\}.$$

- *If* $b \notin Ker(A)^\perp$ *or* $b \in Ker(A)^\perp$ *and* $\|A^+ b\| > r$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{\lambda^* > 0\}$ *where* λ^* *is the unique solution of (3.15) and* $\mathcal{Q}_1 = \mathcal{Q}_2 = \{x(\lambda^*) = -(A + \lambda^* I)^{-1} b\}$.

(iii) *If* $\lambda_1 < 0$ *(i.e. A is nonpositive semi-definite) then*

- *If* $\|(A - \lambda_1 I)^+ b\| \leq r$ *and* $b \in Ker(A - \lambda_1 I)^\perp$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{-\lambda_1\}$ *and*

$$\mathcal{Q}_1 = \mathcal{Q}_2 = \{x = -x^+ + u, \quad u \in Ker(A - \lambda_1 I), \text{ such that } \|x\|^2 = \|x^+\|^2 + \|u\|^2 = r^2\}$$

- *If* $b \notin Ker(A - \lambda_1 I)^\perp$ *or* $b \in Ker(A - \lambda_1 I)^\perp$ *and* $\|(A - \lambda_1 I)^+ b\| > r$ *then* $\mathcal{D}_1 = \mathcal{D}_2 = \{\lambda^* > 0\}$ *where* λ^* *is the unique solution of (3.15) and* $\mathcal{Q}_1 = \mathcal{Q}_2 = \{x(\lambda^*) = -(A + \lambda^* I)^{-1} b\}$.

2. *Assume that* $b = 0$.
(i) *If* $\lambda_1 > 0$ *then* $\mathcal{D}_1 = \{0\}, \mathcal{Q}_1 = \{0\}$, $\mathcal{D}_2 = \{-\lambda_1\}$ *and* $\mathcal{Q}_2 = \{x \in Ker(A - \lambda_1 I) : \|x\| = r\}$.
(ii) *If* $\lambda_1 = 0$ *then* $\mathcal{D}_1 = \{0\}$, $\mathcal{Q}_1 = \{x \in Ker(A) : \|x\| \leq r\}$, $\mathcal{D}_2 = \{0\}$, $\mathcal{Q}_2 = \{x \in Ker(A) : \|x\| = r\}$.
(iii) *If* $\lambda_1 < 0$ *then* $\mathcal{Q}_1 = \mathcal{Q}_2 = \{x \in \mathcal{N}(A - \lambda_1 I) : \|x\| = r\}$.

Remark 3.15. It is worth noting that $(A-\lambda_1 I)(A-\lambda_1 I)^+ b = b \Leftrightarrow b \in \mathrm{Ker}(A-\lambda_1 I)^\perp$. The following result concerning finiteness of the solution set to (Q_1) is very useful for proving the convergence of the whole sequence $\{x^k\}$ (generated by DCA) to a solution of (Q_1).

Corollary 3.16. *$\mathcal{Q}_1$ is finite if and only if $\mathcal{Q}_1$ is a singleton. More precisely $\mathcal{Q}_1$ is finite if and only if $\|(A + \lambda^* I)^+ b\| = r$ where $\mathcal{D}_1 = \{\lambda^*\}$. In this case we have*

$$\mathcal{Q}_1 = \{-(A + \lambda^* I)^+ b\} \tag{3.16}$$

$\mathcal{Q}_2$ is finite if and only if either of the following properties holds:
(i) $\|A + \lambda^* I)^+ b\| = r$ *where* $\mathcal{D}_2 = \{\lambda^*\}$.
This condition is in fact necessary and sufficient for $\mathcal{D}_2$ to be a singleton:

$$\mathcal{Q}_2 = \{-(A + \lambda^* I)^+ b\} \tag{3.17}$$

(ii) $\|A + \lambda^* I)^+ b\| < r$ *where* $\lambda^* \in \mathcal{D}_2$ *and* $Ker(A + \lambda^* I)$ *is a one dimensional subspace. This condition is in fact necessary and sufficient for* $\mid \mathcal{D}_2 \mid$ *to be equal to 2.* ($\mid \mathcal{D}_2 \mid$ *denotes the number of elements in* $\mathcal{D}_2$).

Finally the following nice result which has been stated very recently by Martinez [7] strengthens the ability for DCA to reach a solution (global minimum) of (Q_1):

Theorem 3.17.
(i) *If x^* is local-nonglobal minimum for (Q_1) or (Q_2) then $(A + \lambda^* I)x^* = -b$ with $\lambda^* \in]-\lambda_2, -\lambda_1[$ and $\phi'(\lambda^*) \geq 0$.*
If x^* is a local-nonglobal minimum for (Q_1) then $\lambda^* \geq 0$.
(ii) *There exists at most one local-nonglobal minimum for (Q_1) or (Q_2).*
(iii) *If $\|x^*\| = r$, $(A + \lambda^* I)x^* = -b$ for some $\lambda^* \in]-\lambda_2, -\lambda_1[$ and $\phi'(\lambda^*) > 0$ then x^* is a strict local minimum for (Q_2).*
If, in addition, $\lambda^* > 0$, x^* is also a strict local minimum for (Q_1).
(iv) *If b is orthogonal to some eigenvector associated with λ_1, then there are no local-nonglobal minimum for (Q_1) and (Q_2).*

Conclusion.
We have completed a thorough study on the stability of Lagrangian duality and global optimality conditions for nonconvex quadratic minimization over Euclidean balls and spheres. These results, especially the main properties, in Proposition 3.12, of ϕ and ψ, the detailed description of solution sets for (Q_1) and (Q_2) in Proposition 3.14 and the finiteness of their solution sets in Corollary 3.16, should be essential for devising solution algorithms.
This paper constitutes the theoretical part (Part I) of our work relative to nonconvex quadratic minimization over Euclidean balls and spheres.

Acknowledgment. We would like to thank Professors Hoang Tuy and Le Dung Muu for their careful reading and constructive criticism. Their useful comments substantially clarified the presentation and corrected some inaccuracies.

References

[1] J.R. Clermont, M.E de la Lande and Pham Dinh Tao: Analysis of plane and axisymmetric flows of incompressible fluids with the stream tube method: Numerical simulation by Trust region algorithm, Inter. J. for Numer. Method in Fluids, vol. 13, 371–399, 1991.

[2] R. Fletcher: Practical methods of optimization, A Wiley-Interscience publication, Second Edition, 1991.

[3] W. Gander, G.H. Golub and U. von Matt: A constrained eigenvalue problem, Linear Algebra and its Application, 114/115, 815–839, 1989.

[4] D.M. Gay: Computing optimal locally constrained steps, SIAM J. Sci. Stat. Comput., 2, 186–197, 1981.

[5] G.H Golub, C. van Loan: Matrix computation, North Oxford Academic, Oxford, 1989.

[6] G.H. Golub, U. von Matt: Quadratically constrained least squares and quadratic problems, Numer. Math. 59, 561–580, 1991.

[7] J. Martinez: Local minimizers of quadratic functions on euclidean balls and spheres, SIAM J. Optimization, Vol.4, No 1, 159–176, 1994.

[8] J.J. More: Recent developments in algorithm and software for Trust Region Methods. Mathematical Programming, The State of the Art, Springer, Berlin, 258–287, 1983.

[9] J.J. More, D.C. Sorensen: Computing a trust region step. SIAM J. Sci. Stat. Comput., 4, 553–572, 1983.

[10] Pham Dinh Tao: Méthodes numériques pour la minimisation globale d'une forme quadratique (convexe ou non convexe) sur une boule et une sphère euclidiennes. Rapport de recherche. Université Joseph-Fourier, Grenoble, 1989.

[11] Pham Dinh Tao, S.Wang: Training multi-layered neural network with a Trust region based algorithm, Math. Modell. Numer. Anal. (M^2AN), 24(4), 523–553, 1990.

[12] Pham Dinh Tao, Le Thi Hoai An: Minimisation globale d'une forme quadratique sur une boule et une sphère euclidiennes. Stabilité de la dualité lagrangienne. Optimalité globale. Méthodes numériques. Rapport de Recherche, L.M.I, CNRS URA 1378, INSA-Rouen, 1992.

[13] Pham Dinh Tao, Le Thi Hoai An and Thai Quynh Phong: Numerical methods for globally solving a class of nonconvex quadratic programming, Submitted

[14] B. Polyak: Introduction to Optimization. Optimization Software, Inc., Publication Division, New York, 1987.

[15] R.T. Rockafellar: Monotone operators and the proximal point algorithm in convex programming, SIAM J. on Control and Optim., 14, 877–898, 1976.

[16] D.C. Sorensen: Newton's method with a model trust region modification, SIAM J. Numer. Anal., 19(2), 409–426, 1982.

[17] J.F. Toland: Duality in Nonconvex Optimization. J. Math. Anal. Appl., 66, 399–415, 1978.