A Quadratic Minimization Problem on Subsets of Symmetric Positive Semidefinite Matrices*

Pablo Tarazaga†

August, 1989

TR89-7

*Work supported by CONICET, Argentina.
†IMASL, Universidad Nacional de San Luis, San Luis, Argentina and Mathematical Sciences Department, Rice University, Houston, Texas 77251-1892.
A Quadratic Minimization Problem on Subsets of Symmetric Positive Semidefinite Matrices*

by

Pablo Tarazaga†

* Work supported by CONICET, Argentina
† IMASL, Universidad Nacional de San Luis, San Luis, Argentina and Mathematical Science Department, Rice University, Houston, Texas 77251-1892
1.- INTRODUCTION.

In this paper we deal with a family of constrained optimization problems in which the objective function, which is the same for each problem in the family, is quadratic and strictly convex and is defined for any matrix of order $n$. The difference among problems in this family is in the feasible sets. These feasible sets are the sets of symmetric positive semidefinite matrices with rank less than or equal to $k$, so $k$ is the parameter of the family. The first observation is that these feasible sets are closed but not convex, except for $k = n$ which satisfies both properties.

The central idea contained in our approach is to transform each problem of the family into an unconstrained optimization problem using a function that takes $\mathbb{R}^{n \times k}$ onto the feasible sets. This function allows us to consider an equivalent objective function on $\mathbb{R}^{n \times k}$.

This family of optimization problems is related with the problem of finding the nearest distance matrix to a given one and its applications. The set of symmetric positive semidefinite matrices allows a parametrization of these problems [3],[4],[6].

The organization of this paper is as follows. First we study the relation between feasible sets and the faces of the cone of symmetric positive semidefinite matrices. In this context we give a more complete and useful characterization of these faces, than the one given in [5]. This result allows us to describe the feasible sets in a convenient way.

In the second section we will see that if we minimize the quadratic and strictly convex function on $\mathbb{R}^{n \times n}$ and the solution is symmetric, the location of this solution gives us information about the solutions of our problems for special values of $k$. We prove explicitly that if this solution is positive semidefinite with rank $\tilde{k}$, then the problems stated with $k$ greater than $\tilde{k}$ have the same solution. For the case in which the minimizer of the quadratic function in $\mathbb{R}^{n \times n}$ is indefinite or negative semidefinite, the solution of the problems for every $k$ has rank less than $n$.

In the final section we introduce the function that transforms each problem into an unconstrained problem which is easier to solve, and then we study a variant of this new problem in order to get the matrix solution in its spectral representation. This can be done solving a new constrained problem or constructing an exact penalty function.

2.- THE FEASIBLE SETS.

We denote by $\Omega_n$ the set of symmetric positive semidefinite matrices of order $n$. Now
for \( k = 1, \ldots, n \), we can define the following subsets

\[
\Omega_k = \{ A \in \Omega_n \mid \text{rank}(A) \leq k \}.
\]

Our first interest is to get a useful description of \( \Omega_k, \ k = 1, \ldots, n \), since they are the feasible sets of our problems. We will point out the relation of these subsets with the faces of \( \Omega_n \). Characterizations for these faces have been given in [1],[2] and [5], but in order to get our goal, we establish a new characterization of the faces of \( \Omega_n \).

First, we need some definitions.

A subset \( C \) of \( \mathbb{R}^n \) is a cone if \( x, y \in C \) and \( \alpha, \beta \geq 0 \) implies that \( \alpha x + \beta y \in C \).

A cone \( C \) induces a partial order on its elements defined for \( x, y \in C \) by \( x \leq y \) if and only if \( y - x \in C \).

A cone is polyhedral if it is the intersection of a finite number of half spaces.

A subset \( S \) of a cone \( C \) is a subcone of \( C \) if \( S \) is a cone.

A subcone \( F \) of \( C \) is a face of \( C \) if and only if \( x, y \in C \) and \( 0 \leq x \leq y \) and \( y \in F \) imply \( x \in F \).

An extremal of a cone is a face of dimension one, where the dimension of any face is computed as the dimension of a convex set.

We enunciate now some known properties of the set of symmetric positive definite and positive semidefinite matrices. A survey of general properties of this cone can be found in [5].

a) \( \Omega_n \) is a cone but not a polyhedral cone.

b) The relative interior of \( \Omega_n \) is the set of positive definite matrices.

c) The extremals of \( \Omega_n \) are the positive semidefinite matrices of rank equal to one.

Now, given a matrix \( A \in \Omega_n \), we would like to characterize the smallest face of \( \Omega_n \) that contains \( A \). We denote by \( R(A) \) the column space of the matrix \( A \) and in the same way \( R(s_1, \ldots, s_k) \), the column space of the matrix whose columns are \( s_i, \ i = 1, \ldots, k \). We now have the following result.

Lemma 1: If \( A \in \Omega_k \), the set defined by

\[
F(A) = \{ \sum_{i=1}^{k} s_is_i^T / R(s_1, \ldots, s_k) \subseteq R(A) \}
\]

is the smallest face of \( \Omega_n \) that contains \( A \).

Proof: First of all, \( A \in F(A) \) as a consequence of its spectral representation. In the second place we will see that \( F(A) \) is a cone. If \( B, C \in F(A) \), they can be written

\[
B = \sum_{i=1}^{k} u_iu_i^T \quad C = \sum_{i=1}^{k} s_is_i^T
\]
where $R(u_1, \ldots, u_k)$ and $R(s_1, \ldots, s_k)$ are contained in $R(A)$. We need to prove that for
$\mu$ and $\nu$ greater or equal to zero, $\mu B + \nu C$ is in $F(A)$. But $\mu B + \nu C$ is equal to

$$
\sum_{i=1}^{k} \mu^{1/2} u_i \mu^{1/2} u_i^T + \sum_{i=1}^{k} \nu^{1/2} s_i \nu^{1/2} s_i^T
$$

and this expression says that $R(\mu B + \nu C) \subseteq R(u_1, \ldots, u_k, s_1, \ldots, s_k) \subseteq R(A)$, which implies that $\text{rank}(\mu B + \nu C) \leq k$. Moreover $\mu B + \nu C$ is symmetric and positive semidefinite. Hence the spectral representation of $\mu B + \nu C$ is

$$
\sum_{i=1}^{k} \lambda_i v_i v_i^T,
$$

and we can write it as

$$
\sum_{i=1}^{k} \lambda_i^{1/2} v_i \lambda_i^{1/2} v_i^T.
$$

Since $R(\mu B + \nu C) = R(v_1, \ldots, v_k)$, it is clear that $\mu B + \nu C \in F(A)$.

Finally, we prove that $F(A)$ is a face. Let $X, Y \in \Omega_n$ and consider $0 \leq X \leq Y$ and $Y \in F(A)$, then

$$
X = \sum_{i=1}^{n} s_i s_i^T \quad \text{and} \quad Y = \sum_{i=1}^{k} u_i u_i^T.
$$

But $Y - X \in \Omega_n$, and we can write

$$
Y - X = \sum_{i=1}^{n} \lambda_i v_i v_i^T,
$$

and this implies that

$$
Y = \sum_{i=1}^{k} u_i u_i^T = \sum_{i=1}^{n} s_i s_i^T + \sum_{i=1}^{n} \lambda_i v_i v_i^T. \quad (\star)
$$

Now using the fact that, if $A$ and $B$ are positive semidefinite

$$
\text{rank}(A) \leq \text{rank}(A + B)
$$

in the equality $(\star)$, we get

$$
k \geq \text{rank}(Y) \geq \text{rank}(\sum_{i=1}^{n} s_i s_i^T) = \text{rank}(X).
$$
From the expression for $X$, we know that $R(X) \subseteq R(s_1, \ldots, s_n)$, and using the equality (*), we obtain $R(s_1, \ldots, s_n) \subseteq R(u_1, \ldots, u_k) \subseteq R(A)$. The spectral representation of $X$ with these two facts just mentioned tell us that $X \in F(A)$. The next result was proved in lemma 1.

Corollary 2: Given $k$ linearly independent vectors $a_1, \ldots, a_k$, let $A$ be the matrix

$$A = \sum_{i=1}^{k} a_i a_i^T,$$

then $F(A)$ is a face.

Consider now the following function

$$\omega : R^n \rightarrow R^{n \times n}$$

$$\omega(s) = ss^T$$

We will use the Frobenius inner product in $R^{n \times n}$ defined by $<A, B>_F = tr(A^T B)$, where $tr(.)$ denotes the trace of any matrix of $R^{n \times n}$.

Now, for $s, u \in R$ we have

$$<\omega(s), \omega(u)> = tr[(ss^T)(uu^T)] = tr[s(s^T u)u^T] = s^T u [tr(su^T)] = (s^T u)^2,$$

which shows us that if $s^T u = 0$, then $\omega(s)$ and $\omega(u)$ are orthogonal.

The following two results are consequences of this fact.

Corollary 3: If $R(a_1, \ldots, a_k)$ is orthogonal to $R(a_{k+1}, \ldots, a_n)$ then

$$F(\sum_{i=1}^{k} a_i a_i^T) \quad \text{and} \quad F(\sum_{i=k+1}^{n} a_i a_i^T)$$

are orthogonal. The dimensions of these two faces are the squares of the dimensions of the subspaces $R(a_1, \ldots, a_k)$ and $R(a_{k+1}, \ldots, a_n)$ respectively (for details see [5])

Corollary 4: If $a_1, \ldots, a_n$ are orthogonal vectors, then

$$F(a_1 a_1^T), \ldots, F(a_n a_n^T)$$

are orthogonal faces of dimension one.

To finish this section we give an interesting relation between the set $\Omega_k$ and the faces of $\Omega_n$. 

5
Corollary 5: If \( k = 1, \ldots, n \), then \( \Omega_k \) is equal to the union of all faces \( F(A) \) of \( \Omega_n \) such that \( A \in \Omega_k \).

3.- THE OPTIMIZATION PROBLEM: PROPERTIES.

Let \( q \) be a strictly convex quadratic function

\[
q : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}
\]

defined for the square matrices of order \( n \). We define for \( k = 1, \ldots, n \) the following subsets of \( \Omega_n \)

\[
\Omega^=_{k} = \{ A \in \Omega_k / \text{rank}(A) = k \}.
\]

This section concerns the following optimization problems:

\[
P^=_{k} = \min_{A \in \Omega^=_{k}} q(A)
\]

and

\[
P_{k} = \min_{A \in \Omega_{k}} q(A)
\]

defined in both cases for \( k = 1, \ldots, n \).

The first natural question is, are there solutions for \( P^=_{k}, k = 1, \ldots, n \)? In other words, is it possible to minimize any strictly convex quadratic function over the symmetric positive semidefinite matrices of rank \( k \)? In the next result we will show that it is not always possible.

Theorem 6: If \( A_* \) solves \( P_{n} \) and \( A_* \in \Omega^=_{k_*} \), then for \( k \) such that \( k_* < k \leq n \), the problem \( P^=_{k} \) has no solution.

Proof: If \( A_k \) solves \( P^=_{k} \), then clearly \( q(A_k) \geq q(A_*) \). Now consider \( k \) satisfying the hypothesis of the theorem. Then \( q(A_*) < q(A_k) \), since \( P_{n} \) has a unique solution and \( A_k \neq A_* \) because \( A_* \notin \Omega^=_{k_*} \).

Since \( A_* \in \Omega^=_{k_*} \),

\[
A_* = \sum_{i=1}^{k_*} \lambda_i v_i v_i^T,
\]

where \( \lambda_1, \ldots, \lambda_{k_*} \) are positive and \( v_i, i = 1, \ldots, k_* \), are orthonormal vectors. We can obtain an orthonormal basis \( v_1, \ldots, v_{k_*}, v_{k_*+1}, \ldots, v_n \) and define the following sequence

\[
\tilde{A}_j^k = \sum_{i=1}^{k_*} \lambda_i v_i v_i^T + \frac{1}{j} \left( \sum_{i=k_*+1}^{k} v_i v_i^T \right) \quad j = 1, 2, \ldots.
\]
This sequence has the properties that $\hat{A}_j^k \in \Omega_k^=$ for all $j$, and that $\{\hat{A}_j^k\}$ converges to $A_*$. Now using the continuity of $q$, it is possible to choose $j_0$ such that

$$q(A_*) < q(\hat{A}^k_{j_0}) < q(A_k),$$

but this is a contradiction to the fact that $A_k$ solves $P_k^=$, so $P_k^=$ has no solution.

This result points out that problem $P_k^=$, might not have a solution for some values of $k$. This fact suggests that we consider instead the problem $P_k$, because $\Omega_k$ is a closed set and then $P_k$ will always have a solution. Note that $\Omega_k$ is closed because each convergent sequence of symmetric positive semidefinite matrices of rank $\leq k$ must converge to a symmetric positive semidefinite matrix of rank $\leq k$.

More information about the rank of the solution of the problem $P_n$ can be given if we know that the minimum of $q$ over $R^{n \times n}$ is a symmetric matrix but not positive semidefinite.

Theorem 7: The problem

$$\{ \min q(A) \quad \min q(A) \\
A \in R^{n \times n} \}$$

has a solution $A_*$. If $A_*$ is symmetric but $A_*$ does not belong to $\Omega_n$, then $P_n^=$ has no solution. Thus, the solution of $P_n$ has rank strictly less than $n$. This result is not true if $A_*$ is not symmetric.

Proof: We give a direct proof of this result. Suppose that $\tilde{A}$ solves $P_n^=$. Now, consider the matrix $A_\lambda = \lambda(A_* - \tilde{A}) + \tilde{A}$ and observe that $A_0 = \tilde{A}$ and $A_1 = A_*$. Since $A_*$ is not positive semidefinite, it has some negative eigenvalue, and there exists a minimum $\lambda \in (0,1)$ such that $\det(A_\lambda) = 0$ and the eigenvalues of $A_\lambda$ are nonnegative. This implies that rank $(A_\lambda) < n$.

Now from the strict convexity of $q$ and the fact that $A_*$ is the unique minimizer, $q$ is a decreasing function of $\lambda$ for $\lambda \in [0,1]$ and

$$q(A_*) < q(A_\lambda) < q(\tilde{A}).$$

But for any $\lambda \in (0,\hat{\lambda})$, $A_\lambda$ is positive definite, therefore has rank $n$ and

$$q(A_\lambda) < q(\tilde{A}),$$

which contradicts the definition of $\tilde{A}$, so $P_n^=$ has no solution.
In order to see that the result fails when $A_*$ is not symmetric, consider the following example for $n = 2$. Define $A_*$ by
\[
\begin{pmatrix}
1 & 1 \\
0 & 1 \\
\end{pmatrix}
\]
and consider the function
\[
q(A) = \| A - A_* \|^2_F
\]
where $\| \cdot \|_F$ denotes the Frobenius norm.

The unique minimizer is $A_*$ a non symmetric positive definite matrix. Since the level surfaces are spheres, the solution of
\[
\begin{cases}
\min q(A) \\
A \in \Omega_2
\end{cases}
\]
is the projection on the subspace of symmetric matrices given by
\[
\frac{A_* + A_*^T}{2} = \begin{pmatrix}
1 & 1/2 \\
1/2 & 1
\end{pmatrix},
\]
and clearly its rank is two.

The following results are consequences of the last two theorems.

Corollary 8: If $A_*$ is the solution of $P_n$ and $A_* \in \Omega_{k_*}^-$, then $A_*$ solves $P_{k_*}^-$ and $P_{k_*}$.

Corollary 9: If $A_*$ is the solution of $P_n$ and $A_* \in \Omega_{k_*}^-$, then $A_*$ solves all the problems $P_k$ with $k \geq k_*$.

This result has the following converse.

Corollary 10: Suppose that all $P_k$ with $k \geq k_*$ have the same solution $A$, then $A \in \Omega_{k_o}$ with $k_o \leq k_*$. 

4.- TRANSFORMING THE PROBLEM.

We want now to solve the problem $P_k$, for $k = 1, \ldots, n$. In order to do this, we transform it into an unconstrained problem. As a natural consequence of lemma 1 and corollary 5, we need to generalize the function defined after corollary 2 in the following way
\[
\tilde{\omega} : \mathbb{R}^{n \times k} \rightarrow \mathbb{R}^{n \times n}
\]
\[
\tilde{\omega}(a_1, \ldots, a_k) = \sum_{i=1}^{k} a_i a_i^T
\]
This function has interesting properties.

Lemma 11: The function \( \tilde{\omega} \) satisfies

a) \( \text{rank}[\tilde{\omega}(a_1, \ldots, a_k)] \leq k \)

b) \( \tilde{\omega}(R^{n \times k}) = \Omega_k \)

Proof: a) is a consequence of the fact that \( \tilde{\omega}(a_1, \ldots, a_k) = AA^T \) where \( A \) is a matrix of dimension \( n \times k \) whose columns are the vectors \( a_i, \quad i = 1, \ldots, k \).

b) Given a matrix \( A \in \Omega_k \), if its spectral decomposition is

\[
A = \sum_{i=1}^{k} \lambda_i v_i v_i^T,
\]

then clearly

\[
\tilde{\omega}(\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) = A.
\]

Using b) of lemma 11 we can define the function

\[
\bar{q} : R^{n \times k} \rightarrow R
\]

\[
\bar{q}(a_1, \ldots, a_k) = q(\tilde{\omega}(a_1, \ldots, a_k))
\]

and this new function allows us to consider the problem

\[
T_k = \left\{ \min_{(a_1, \ldots, a_k) \in R^{n \times k}} \bar{q}(a_1, \ldots, a_k) \right\}
\]

The following result gives us an interesting relation between problems \( T_k \) and \( P_k \).

Theorem 12: If \( (a_1, \ldots, a_k) \) solves \( T_k \), then \( \tilde{\omega}(a_1, \ldots, a_k) \) solves \( P_k \). If \( A \) solves \( P_k \) and its spectral representation is

\[
A = \sum_{i=1}^{k} \lambda_i v_i v_i^T,
\]

then all elements in \( \tilde{\omega}^{-1}(A) \) solve \( T_k \) and \( (\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) \in \tilde{\omega}^{-1}(A) \).

Proof: Consider the following equalities

\[
\bar{q}(a_1, \ldots, a_k) = q(\tilde{\omega}(a_1, \ldots, a_k)) = q(A) = q\left( \sum_{i=1}^{k} \lambda_i v_i v_i^T \right) = q\left( \sum_{i=1}^{k} \lambda_i^{1/2} v_i \lambda_i^{1/2} v_i^T \right) =
\]

\[= q(\tilde{\omega}(\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k)).
\]
It is easy to see that while \( k \leq \frac{n+1}{2} \), the dimension \( nk \) of the problem \( T_k \) is less or equal to the dimension \( \frac{n(n+1)}{2} \) of the problem \( P_k \).

This last theorem tells us that it is possible that by solving \( T_k \) we may get the spectral representation of the solution \( A \). This causes the following transformation.

We define the function
\[
h: \mathbb{R}^{n \times k} \rightarrow \mathbb{R}
\]
by
\[
h(a_1, \ldots, a_k) = \sum_{i<j} (a_i^T a_j)^2.
\]
This function satisfies:
\[h(a_1, \ldots, a_k) \geq 0\]
\[h(a_1, \ldots, a_k) = 0 \text{ if and only if all the } a_i, i = 1, \ldots, k \text{ different from zero are orthogonal.}\]

We consider now the problem
\[
T_k^\equiv = \begin{cases} 
\min \tilde{q}(a_1, \ldots, a_k) \\
h(a_1, \ldots, a_k) = 0.
\end{cases}
\]

The following result compares this problem with \( T_k \).

Lemma 13: Assume that \( A \) is equal to \( \bar{\omega}(a_1, \ldots, a_k) \), and its spectral representation is
\[
\sum_{i=1}^k \lambda_i v_i v_i^T.
\]
If \( (a_1, \ldots, a_k) \) solves \( T_k \), then \( (\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) \) solves \( T_k^\equiv \). Moreover, if \( (a_1, \ldots, a_k) \) is a solution of \( T_k^\equiv \), then it solves \( T_k \).

Proof: The proof is a consequence of the definition of \( \tilde{q} \), the properties of \( h \), and the fact that each matrix in \( \Omega_k \) has spectral representation of the given form.

Now we study another way to solve the problem \( T_k^\equiv \) by constructing an exact penalty function.

Theorem 14: The function \( Q \) defined by
\[
Q(a_1, \ldots, a_k) = \tilde{q}(a_1, \ldots, a_k) + h(a_1, \ldots, a_k)
\]
is an exact penalty function for the problem \( T_k^\equiv \).

Proof: First of all, it is clear that
\[
\tilde{q}(a_1, \ldots, a_k) \leq Q(a_1, \ldots, a_k).
\]
Moreover we prove that if \((a_1, \ldots, a_k)\) is a minimizer for \(Q\) on \(\mathbb{R}^{n \times k}\), then \(h(a_1, \ldots, a_k) = 0\). Suppose for a moment that \(h(a_1, \ldots, a_k) > 0\). Now the spectral representation of \(\tilde{\omega}(a_1, \ldots, a_k)\) is

\[
\tilde{\omega}(a_1, \ldots, a_k) = \sum_{i=1}^{k} \lambda_i^{1/2} v_i \lambda_i^{1/2} v_i^T
\]

and clearly

\[
h(\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) = 0,
\]

which implies

\[
Q(\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) = \tilde{q}(\lambda_1^{1/2} v_1, \ldots, \lambda_k^{1/2} v_k) = \tilde{q}(a_1, \ldots, a_k) < \tilde{q}(a_1, \ldots, a_k) + h(a_1, \ldots, a_k) = Q(a_1, \ldots, a_k)
\]

and this is a contradiction.

Now we can prove that if \((a_1, \ldots, a_k)\) is a minimizer for \(Q\), then it is a solution for \(T_k^\infty\). Suppose that \((a_1, \ldots, a_k)\) is a minimizer for \(Q\) and is not a solution for \(T_k^\infty\), then there exist \((b_1, \ldots, b_k)\) such that

\[
\tilde{q}(b_1, \ldots, b_k) < \tilde{q}(a_1, \ldots, a_k)
\]

and

\[
h(b_1, \ldots, b_k) = 0.
\]

But this implies

\[
Q(b_1, \ldots, b_k) = \tilde{q}(b_1, \ldots, b_k) < \tilde{q}(a_1, \ldots, a_k) = Q(a_1, \ldots, a_k)
\]

which is impossible.

The last two results can be applied to the important problem of find the eigenvalues and eigenvectors of a matrix \(A\). We show in the next result how to apply theorem 14 to this problem, and the same can be done with lemma 13.

Corollary 15: Let \(A\) be a matrix in \(\Omega_n\). If \((x_1, \ldots, x_n)\) solves

\[
\left\{ \begin{array}{l}
\min Q_A(x_1, \ldots, x_n) \\
(x_1, \ldots, x_n) \in \mathbb{R}^{n \times n}
\end{array} \right.
\]

where

\[
Q_A(x_1, \ldots, x_n) = \| \sum_{i=1}^{n} x_i x_i^T - A \|_F^2 + h(x_1, \ldots, x_n)
\]

11
and
\[ h(x_1, \ldots, x_n) = \sum_{i<j} (x_i^T x_j)^2, \]
then any nonzero vector in \( x_i, i = 1, \ldots, n \) is an eigenvector of a nonzero eigenvalue of \( A \).

Proof: It is easy to see from the definition that \( Q_A(x_1, \ldots, x_n) \geq 0 \), moreover we have \( Q_A(x_1, \ldots, x_n) = 0 \) if and only if \( A = \sum_{i=1}^n x_i x_i^T \) and \( x_i, i = 1, \ldots, n \) are orthogonal.

It is clear that if the rank of \( A \) is known, then the problem can be formulated in a lower dimension.

ACKNOWLEDGMENTS.

I want to express my grateful to Professor John E. Dennis for his comments and suggestions besides the correction of the draft, also to Dr. Hector Martinez for his comments.

References: