Journal of
Linear and Topological Algebra
Vol. 08*, No.* 02*,* 2019, 133- 143

On the duality of quadratic minimization problems using pseudo inverses

D. Pappas[a,*], G. N. Domazakis[b]

[a] *Department of Statistics, Athens University of Economics and Business, 76 Patission Str 10434, Athens, Greece.*
[b] *Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Iroon Polytexneiou 9, 15780 Zografou, Athens, Greece.*

Received 2 September 2018; Revised 15 October 2018; Accepted 30 November 2018.

Communicated by Mohammad Sadegh Asgari

Abstract. In this paper we consider the minimization of a positive semidefinite quadratic form, having a singular corresponding matrix H. We state the dual formulation of the original problem and treat both problems only using the vectors $x \in \mathcal{N}(H)^{\perp}$ instead of the classical approach of convex optimization techniques such as the null space method. Given this approach and based on the strong duality principle, we provide a closed formula for the calculation of the Lagrange multipliers λ in the cases when (i) the constraint equation is consistent and (ii) the constraint equation is inconsistent, using the general normal equation. In both cases the Moore-Penrose inverse will be used to determine a unique solution of the problems. In addition, in the case of a consistent constraint equation, we also give sufficient conditions for our solution to exist using the well known KKT conditions.

Keywords: Moore-Penrose inverse, general normal equation, constrained optimization, Lagrange multipliers, duality principle, KKT conditions.

2010 AMS Subject Classification: 15A09, 47A05, 47N10.

1. Introduction

In many situations in physical sciences, the minimization of quadratic forms arises as a suitable formulation of certain problems, in both the finite or the infinite dimensional setting. Considering the minimization of a positive semidefinite quadratic form under

---

*Corresponding author.
E-mail address: dpappas@aueb.gr (D. Pappas); domazakisgeorge@hotmail.com (G. N. Domazakis).

Print ISSN: 2252-0201
Online ISSN: 2345-5934

http://jlta.iauctb.ac.ir

linear constraints, an interesting question arises when the corresponding matrix is singular. In this case, the problem has been treated in the last decades using pseudo-inverses (for example, see [10, 11, 14]. An example of such a case comes from electrical network analysis, considering the problem of minimizing the energy dissipation $f(x) = \langle x, Hx \rangle$, where H is the conductance matrix, most of the times positive definite. The voltage vector x satisfies Kirchoff's second law if and only if it minimizes the above quadratic form under linear constraints (for more details, see [1]). Since a positive semidefinite quadratic form has a minimum value of zero when H is positive semidefinite, another approach is to use only the vectors belonging to $\mathcal{N}(H)^{\perp}$. Moreover, if the corresponding matrix H is singular, some type of a generalized inverse must be used. The right choice is the Moore-Penrose inverse because of the orthogonal decomposition of the space resulted by its use. The general problem studied is the minimization of a positive semidefinite quadratic form under linear constraints in the form of

$$\begin{aligned} &\underset{x}{\text{minimize}} && \Phi(x) = \frac{1}{2}\langle x, Hx \rangle + \langle g, x \rangle, \\ &\text{subject to} && Ax = b, \end{aligned}$$

where H is a $n \times n$ real matrix. We will call the problem above as the primal problem. The primal problem is closely related with the well-known dual problem in the form of

$$\begin{aligned} &\max_{x} \quad \min_{\lambda} && \mathcal{L}(x, \lambda) = -\frac{1}{2}\langle x, Hx \rangle + \langle b, \lambda \rangle, \\ &\text{subject to} && Hx + g = A^{\top}\lambda. \end{aligned}$$

From the optimization theory point of view, as we are working in a finite dimensional case, due to the strong duality principle, both solutions of the primal and dual problem coincide.

In this paper based on that argument we present a method to calculate the Lagrange multipliers involved in the dual problem and finally compute the solution, inspired by the primal-dual methods in convex optimization techniques and the fact that we are using only the vectors $x \in \mathcal{N}(H)^{\perp}$. Using the strong duality principle, our result (Theorem 3.3) shows that we can compute the Langrange multipliers using the Moore-Penrose inverse, without the use of the classical methods involving derivatives of certain functions, which in many times may be difficult to compute. Using the well known KKT conditions we give sufficient conditions for our solution to be unique, despite the fact that H is positive semidefinite. In the last section of this paper, the same work is extended in the case of an inconsistent constraint equation $Ax = b$, considering the general normal equation $A^{k+1}x = A^k b$.

2. Preliminaries and notation

Let as denote the set of all $n \times r$ real matrices as $\mathbb{M}_{n\times r}(\mathbb{R})$. In the case of $n \times n$ square matrices, we denote $\mathbb{M}_n(\mathbb{R})$. Consider $T \in \mathbb{M}_n(\mathbb{R})$. We denote its kernel as $\mathcal{N}(T)$, its range as $\mathcal{R}(T)$, its rank as $rank(T)$, its conjugate transpose as T^* and its transpose as $T^{\top}$. Moreover, the smallest positive integer k for which $rank(T^{k+1}) = rank(T^k)$ is called the index of T and is denoted by $ind(T)$.

Let us recall the notion of Moore-Penrose inverse. If $T \in \mathbb{M}_n(\mathbb{R})$ with $rank(T) = r$

it can be shown that there exists a unique matrix, the Moore-Penrose inverse, which satisfies the well-known following four Penrose equations:

$$TT^{\dagger}T = T, \quad T^{\dagger}TT^{\dagger} = T^{\dagger}, \quad TT^{\dagger} = (TT^{\dagger})^*, \quad T^{\dagger}T = (T^{\dagger}T)^*,$$

which is denoted as $T^{\dagger}$. For a great description, see [2]. One can easily verify that $TT^{\dagger}$ is the orthogonal projection of $\mathbb{R}^n$ onto $\mathcal{R}(T)$ denoted by P_T and that $T^{\dagger}T$ is the orthogonal projection of $\mathbb{R}^m$ onto $\mathcal{R}(T^*)$ denoted by P_{T^*}. It is also well known that $\mathcal{R}(T^{\dagger}) = \mathcal{R}(T^*)$.

In case when T commutes with $T^{\dagger}$ or equivalently $\mathcal{R}(T) = \mathcal{R}(T^*)$, then T is called an EP matrix. EP matrices were introduced by Schwerdtfeger [13] and consist of a general class of matrices such as hermitian, skew-hermitian or normal. Since the corresponding matrices of the quadratic forms that we study are positive semidefinite, they are all EP matrices.

For a matrix A of index k, the equation $A^{k+1}x = A^k b$ is consistent irrespective of consistency of the system $Ax = b$. As we can see, this new type of constraint can be used when the linear equation $Ax = b$ is inconsistent and the new set is still a convex set: It will be the general normal equation of the system $Ax = b$ and therefore, the new constraint set will be defined as:

$$\mathfrak{S}_D = \{x|\ x \in \mathbb{R}^n, A^{k+1}x = A^k b,\ k \geqslant indA\}.$$

Standard reference books on generalized inverses are [2, 4, 6]. One of the basic problems in optimization theory, arising in several real life problems, is the quadratic programming problem with equality constraints. In his classical book on optimization theory, Luenberger [7] presents various similar optimization problems for both finite and infinite dimensions. The problem studied in general is the following:

$$\min f(x) = \langle x, Tx \rangle + \langle p, x \rangle + a, \quad x \in S,$$

where $S = \{x : Ax = b\}, a \in \mathbb{R}$, p is a given vector and T is a positive definite matrix. An interesting case to examine is when T is singular and positive semidefinite with a nonempty kernel, $\mathcal{N}(T) \neq \{0\}$. In this case we have that $\langle x, Tx \rangle = 0$ for all $x \in \mathcal{N}(T)$ and so, a first approach in both the finite and infinite dimensional case would be to look among the vectors $x \in \mathcal{N}(T)^{\perp} = \mathcal{R}(T^*) = \mathcal{R}(T)$ for a minimizing vector for $\Phi(x)$. In other words, we will study the problem $\min \Phi(x) = \langle x, Tx \rangle + \langle p, x \rangle + a$ for all $x \in S \cap \mathcal{N}(T)^{\perp}$. Because of the singularity of T, the Moore-Penrose inverse will be used. Moreover, since $\mathcal{N}(T) = \mathcal{N}(T^{\dagger})$ the vectors examined in this case satisfy also an additional condition: That $x \in \mathcal{N}(T)^{\perp}$. We present the above statement in the following theorem found in [11].

Theorem 2.1 Let T a be $n \times n$ singular positive semidefinite matrix with a non empty kernel, $\mathcal{N}(T) \neq \{0\}$, with $X^2 = T$. Let also A be a singular $m \times n$, and consider the equation $Ax = b$. If the set $S = \{x : Ax = b\}$ is not empty, then the problem :

$$\text{minimize} \quad \Phi(x) = \langle x, Tx \rangle + \langle x, p \rangle + a, \quad x \in S \cap \mathcal{N}(T)^{\perp}$$

with $p \in \mathbb{R}^n$ and $a \in \mathbb{R}$ has the unique solution $\hat{x} = X^{\dagger}(AX^{\dagger})^{\dagger}(\frac{1}{2}AT^{\dagger}p + b) - \frac{1}{2}T^{\dagger}p$.

The main motivation of this paper is to calculate the Lagrange multipliers of the classic quadratic minimization problem, only using linear algebra techniques and generalized inverses, without the use of differentiation. Using our technique, even though a positive semidefinite quadratic form is not strictly convex and may have infinite minimizing vectors, using the vectors $x \in \mathcal{N}(T)^{\perp}$ we can find a global minimum and the corresponding Lagrange multipliers. Conditions for such a solution to exist are presented in section 3. One of the applications of the Moore-Penrose inverse in the finite dimensional case is the minimization of a positive definite quadratic functional under linear constraints, presented by Manherz and Hakimi [8], having a similar point of view to the one presented in our work. Moreover, interest on quadratic minimization under linear constraints can also be seen in electrical circuits, signal processing and linear estimation applications (see e.g. [12, 15]) is an additional rationale and these practical applications can be solved and have a global minimum value following our point of view.

The paper is organized as follows: In section 3, we examine the classical problem in both the primal and the dual form and compute the solutions for both of them. We prove that these solutions coincide and calculate the Lagrange multipliers using a closed formula without the use of derivatives. We also present the difference between our method and the null space method, using the KKT conditions for such a solution to exist and the conditions for that are explicitly given. In section 4, the same work is extended in the case of an inconsistent constraint equation $Ax = b$, considering the General Normal equation $A^{k+1}x = A^k b$. Corresponding examples are given in both cases.

3. The primal and dual solutions when the constraint equation is consistent

The minimizing vectors: In many situations it is useful to treat the primal problem with the Langrange method. For this method, we introduce some additional variables $\lambda = (\lambda_1, \cdots, \lambda_n)$, one for each constraint, the well-known Langrange multipliers. The Langrange method leads us to the following dual formulation of the primal problem, which is given in the following definition.

Definition 3.1 [5] Consider the primal problem:

$$\begin{aligned} &\underset{x}{\text{minimize}} \quad \Phi(x) = \frac{1}{2}\langle x, Hx\rangle + \langle g, x\rangle, \\ &\text{subject to} \quad Ax = b. \end{aligned} \tag{1}$$

Then, the dual problem is

$$\begin{aligned} &\max_{x} \quad \min_{\lambda} \quad \mathcal{L}(x,\lambda) = -\frac{1}{2}\langle x, Hx\rangle + \langle b, \lambda\rangle, \\ &\text{subject to} \quad Hx + g = A^{\top}\lambda. \end{aligned} \tag{2}$$

In many cases the usefulness of the Langrange formulation showed up by its natural physical meaning. For example, the shadow prices in economics, or the total potential energy of a certain system. Although, treating the minimization quadratic problems by the Langrange dual method, leads us in a geometric interpretation in terms of the solution of the dual problem, being actually a saddle point of $\mathcal{L}(x, \lambda)$, the Langrange function.

Let us now discuss the solutions of the problems in Definition 3.1. Since the matrix H is singular the classical minimization techniques are collapsing, i.e. infinite minimum values are equal to zero in a trivial way. For that reason, both the primal and dual problems can be treated from another point of view, using Theorem 2.1. Since the matrix H is positive semidefinite, one possible approach to the problem would be to solve it for all vectors $x \in \mathcal{N}(H)^{\perp}$. The same approach can be used to treat the dual problem. In the following theorems we present the solutions of the primal and dual problems with this additional constraint, making use of the Moore-Penrose inverse.

Theorem 3.2 Let H be a $n \times n$ singular positive semidefinite matrix, with $X^2 = H$. Let also A be $m \times n$ matrix and consider the equation $Ax = b$ and g a given vector in $\mathbb{R}^n$. If the set $S = \{x : Ax = b, x \in \mathcal{N}(H)^{\perp}\}$ is not empty, then the primal problem (1) has the unique solution $\hat{x} = X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g + b) - H^{\dagger}g$.

Proof. The proof is straightforward by applying Theorem 2.1 ■

Theorem 3.3 Let H a $n \times n$ singular positive semidefinite matrix with $X^2 = H$. Let also A be $m \times n$ and g a given vector in $\mathbb{R}^n$. If the set $S = \{x : Hx + g = A^{\top}\lambda, x \in \mathcal{N}(H)^{\perp}\}$ is not empty, then the dual problem (2) has the unique solution for x in terms of λ: $\hat{u} = X^{\dagger}(HX^{\dagger})^{\dagger}(A^{\top}\lambda - g)$. The dimension of the Lagrange multipliers vector λ depends on the dimensions of the constraint matrix A, in order for $A^{\top}\lambda$ to be a column vector.

Proof. Obviously the problem of maximizing $-\frac{1}{2}\langle x, Hx\rangle + \langle b, \lambda\rangle$ under $Hx + g = A^{\top}\lambda, x \in S$ is equivalent to the minimization of $\frac{1}{2}\langle x, Hx\rangle - \langle b, \lambda\rangle$ under $Hx = A^{\top}\lambda - g, x \in S$. Using Theorem 2.1, by replacing T, A and b by $\frac{1}{2}H, H$ and $A^{\top}\lambda - g$ respectively and setting $p = 0$, we get the desired solution for $\hat{u}$ in terms of λ. ■

The KKT matrix and the null space method: A well known method for the solution of the primal problem is the null space method, presented explicitely in many convex optimization books, such as [9]. The difference between this method and our proposed method will be presented in details in this section. For the solution of the primal problem (1), the null space method consists of using a matrix Z, whose columns are a basis for the null space of A. Moreover, the well known Karush-Kuhn-Tucker (KKT) matrix is the following: $K = \begin{bmatrix} H & A^{\top} \\ A & 0 \end{bmatrix}$.

Lemma 3.4 [9] Let A have full row rank, and assume that the matrix $Z^{\top}HZ$ is positive definite. Then the KKT matrix K is nonsingular, and hence there is a unique vector pair (u^*, λ^*) satisfying the conditions of the Primal and Dual problems.

The null space method does not require nonsingularity of the matrix H and therefore has wide applicability. It assumes only that the conditions of Lemma 3.4 hold, namely, that A has full row rank and that $Z^{\top}HZ$ is positive definite. However, it requires knowledge of the null-space basis matrix Z. Using a random matrix Y such that $[Y|Z]$ is nonsingular, the KKT system is transformed to calculate the desired step of the solution. This system can be solved by performing a Cholesky factorization of the matrix $Z^{\top}HZ$ to obtain the Lagrange multiplier λ^*. It is now obvious that the null space method has a completely different approach for the solution of this problem. Our method consists of using the vectors perpendicular to the kernel $\mathcal{N}(H)$ of the matrix H corresponding to the quadratic form, while the null space method restricts the problem to the vectors perpendicular to the kernel $\mathcal{N}(A)$ of the restriction matrix A. We can also show using the following theorem, that x^* is a global solution of the primal problem.

Theorem 3.5 [9] Let A have full row rank and assume that the matrix $Z^{\top}HZ$ is positive definite. Then the vector u^* found from Lemma 3.4 is the unique global solution of the primal problem.

Using the above theorem, we will present sufficient conditions for our method to have a global solution:

Theorem 3.6 Let A have full row rank and $\mathcal{N}(A) \cap \mathcal{N}(H) = 0$. then, the solution $\hat{x}$ found from Theorem 3.2 is a global minimizer of the primal problem (1).

Proof. Let X be the square root of the positive semidefinite matrix H, that is, $X^2 = H$. Let us suppose that $\mathcal{N}(A) \cap \mathcal{N}(H) = 0$. Then, $\|Z^\top HZx\| = \langle Z^\top HZx, x\rangle = \langle XZw, XZw\rangle = \|XZx\| > 0$ for all $x \neq 0$ and $Z^\top HZ$ is positive definite. On the other hand, let us suppose that $\mathcal{N}(A) \cap \mathcal{N}(H) \neq 0$. Then, there exists some vector $w \in \mathcal{N}(A) \cap \mathcal{N}(H)$. The relation $\mathcal{N}(A) \cap \mathcal{N}(H) \neq 0$ is equivalent to $\mathcal{R}(Z) \cap \mathcal{N}(H) \neq 0$, so for the vector w we have that $XZw = 0$, since $Zw \in \mathcal{N}(H)$ and it holds that $\mathcal{N}(H) = \mathcal{N}(X)$. Therefore, we get that $\|XZw\| = 0$ and so, $0 = \langle XZw, XZw\rangle = \langle Z^\top HZw, w\rangle$. So, we proved that for some $w \neq 0, \|Z^\top HZw\| = 0$ and therefore the matrix $Z^\top GZ$ is positive semidefinite. Using Theorem 3.5, we get the desired result. ■

Optimal choice of λ: (the minimal norm solution) As we have proved in above theorems, the two optimal solutions for the primal and dual problems are $\hat{x} = X^\dagger(AX^\dagger)^\dagger(AH^\dagger g + b) - H^\dagger g$ and $\hat{u} = X^\dagger(HX^\dagger)^\dagger(A^\top\lambda - g)$. One can observe that the dual solution $\hat{u}$ is formed in terms of λ, $\hat{u} = \hat{u}(\lambda)$.

Before we treat the problem of choice of λ, let us discuss the relation of primal and dual solutions. In general, it is known that the primal solution dominates the dual solution, i.e. $\hat{u} \leqslant \hat{x}$ and that is the weak duality principle and the difference of primal and dual solutions $\hat{x} - \hat{u}$ is called duality gap (see e.g. [3]). This relation holds, even for nonconvex minimization problems. On the other hand, there are many classes of problems that the equality of primal and dual solutions holds, that is the strong duality principle. One class of these problems, is the class of minimization of quadratic forms. The following proposition presents the strong duality principle in our case.

Proposition 3.7 (strong duality principle) The solution of primal problem and dual problem in Theorem 3.2 and Theorem 3.3 respectively coincide, i.e. $\hat{x} = \hat{u}(\lambda)$ or in the other words, we duality gap is zero.

For a generic proof in the setting of quadratic programming, we refer the interest reader to Boyd and Vandenberghe [3]. Since the (strong) duality principle provides the equality of the optimal solutions of both the primal and dual problems, we can use this property in order to compute the optimal value of λ and finally derive the optimal solution of the dual problem via a minimal norm. Let us recall the standard minimization property of the Moore-Penrose inverse:

Proposition 3.8 Let $A \in \mathbb{M}_{r\times m}(\mathbb{R})$ and $b \in \mathbb{R}^r, b \notin \mathcal{R}(A)$, and the equation $Ax = b$. Then, if $A^\dagger$ is the generalized inverse of A, we have that $A^\dagger b = u$, where u is the minimal norm solution.

As it mentioned above and based on the provided specific duality property of Definition 3.1, we can compute the optimal λ values in the following theorem.

Theorem 3.9 The Langrange multipliers $\lambda \in \mathbb{R}^n$ in the dual problem (2) can be computed by solving the problem $\min\limits_{\lambda}||\hat{x} - \hat{u}(\lambda)||_2$, which has unique solution

$$\lambda^* = (X^\dagger(HX^\dagger)^\dagger A^\top)^\dagger \left(X^\dagger(AX^\dagger)^\dagger(AH^\dagger g + b) + X^\dagger(HX^\dagger)^\dagger g - H^\dagger g\right).$$

Proof. The problem $\min\limits_{\lambda}||\hat{x} - \hat{u}(\lambda)||_2$ is trivially equivalent with the problem

$$\min_{\lambda}\left|\left|(X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g+b)-H^{\dagger}g-X^{\dagger}(HX^{\dagger})^{\dagger}(A^{\top}\lambda-g)\right|\right|_2$$

which is equivalent to the minimal norm solution of the equation

$$(X^{\dagger}(HX^{\dagger})^{\dagger}A^{\top}\lambda=(X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g+b)+X^{\dagger}(HX^{\dagger})^{\dagger}g-H^{\dagger}g.$$

So, using Proposition 3.8, we get

$$\lambda^{*}=(X^{\dagger}(HX^{\dagger})^{\dagger}A^{\top})^{\dagger}\left(X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g+b)+X^{\dagger}(HX^{\dagger})^{\dagger}g-H^{\dagger}g\right).$$

■

Summarizing the above discussion we present the following example.

***Example* 3.10** Let the positive semidefinite matrix

$$H=\begin{bmatrix}26&10&-2\\10&8&2\\-2&2&2\end{bmatrix}$$

We are looking for the minimum of the quadratic function $f(x)=\frac{1}{2}\langle x,Hx\rangle+\langle g,x\rangle$ with $g=(1,2,3)^{\top}$ and the set of constraints S defined as $S=\{(x,y,z):3x+y+z=-1\}$ or $Ax=b$. If $X^2=H$, the matrix X is: $X=\begin{bmatrix}4.8335&1.5061&-0.6071\\1.5061&2.1932&0.9601\\-0.6071&0.9601&0.8424\end{bmatrix}$. The matrix $X^{\dagger}$ is $X^{\dagger}=\begin{bmatrix}0.1908&-0.0295&-0.0833\\-0.0295&0.2644&0.1861\\-0.0833&0.1861&0.1518\end{bmatrix}$. As presented in Theorem 3.2, when we use the Moore-Penrose inverse among all solutions of the system of linear equations we get the one with a minimal norm, belonging also to the set $\mathcal{N}(H)^{\perp}$. The set $\mathcal{N}(H)^{\perp}$ has the form $u=(2x-3y,x,y)^{\top}$, $x,y\in\mathbb{R}$. With calculations we can see that all vectors $u\in\mathcal{N}(H)^{\perp}$ satisfying the constraint $Au=b$, where $A=\begin{bmatrix}3&1&1\end{bmatrix}$ and $b=-1$, have the form $u=\left(x,\frac{-8x-3}{5},\frac{-7x-2}{5}\right)^{\top}$. Using Theorem 3.2 we can see that the minimizing vector of $f(x)$ under $\{Ax=b,x\in\mathcal{N}(H)^{\perp}\}$ is $\hat{x}=X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g+b)-H^{\dagger}g=(-0.0049,-0.5922,-0.3931)^{\top}$. It is easy to verify that the solution found satisfies the equation $A\hat{x}=b$. We will now deal with the Lagrange Dual problem using Theorem 3.3. As we have seen, this problem has the unique solution with respect to λ, $\hat{u}=X^{\dagger}(HX^{\dagger})^{\dagger}(A^{\top}\lambda-g)$. By calculations we get that $\hat{u}=\begin{bmatrix}0.06971\lambda+0.11564\\0.09864\lambda-0.42176\\0.04252\lambda-0.31971\end{bmatrix}$. If we set this solution equal to the solution found from the primal problem, we have that

$$\hat{x}=\hat{u}\Rightarrow\begin{bmatrix}0.06971\lambda+0.11564\\0.09864\lambda-0.42176\\0.04252\lambda-0.31971\end{bmatrix}=\begin{bmatrix}-0.0049\\-0.5922\\-0.3931\end{bmatrix}.$$

In all 3 equations, the value of λ is the same $\lambda=-1.72$. On the other hand, using the equation

$$\lambda^{*}=(X^{\dagger}(HX^{\dagger})^{\dagger}A^{\top})^{\dagger}\left(X^{\dagger}(AX^{\dagger})^{\dagger}(AH^{\dagger}g+b)+X^{\dagger}(HX^{\dagger})^{\dagger}g-H^{\dagger}g\right)$$

and the matrices used above, we get that $\lambda^{*}=-1.72$. Moreover, we will examine the

KKT matrix and the conditions presented in Theorem 3.6: The KKT matrix is

$$K = \begin{bmatrix} 26 & 10 & -2 & 3 \\ 10 & 8 & 2 & 1 \\ -2 & 2 & 2 & 1 \\ 3 & 1 & 1 & 0 \end{bmatrix}.$$

We have $det(K) = -192$ and also $\mathcal{N}(A) \cap \mathcal{N}(H) = 0$.

4. The primal and dual solutions when the constraint equation is inconsistent

The minimizing vectors: In this section we will deal with the same problem, considering the case when the constraint equation $Ax = b$ is inconsistent, where A is a square matrix. So, we will use the General Normal equation, $A^{k+1}x = A^k b$ where k is the index of A, which is always consistent. The new constraint set is now defined as:

$$\mathfrak{S}_D = \{x | \ x \in \mathbb{R}^n, A^{k+1}x = A^k b, \ k \geqslant indA\}.$$

In this section, we suppose that $T \in \mathbb{M}_n(\mathbb{R})$ is again a positive semidefinite matrix. Let $A \in \mathbb{M}_n(\mathbb{R})$ be such that $ind(A) = k$ and $x, b \in \mathbb{R}^n$. We consider the minimization of the functional $\Phi(x) = \langle x, Tx\rangle + \langle x, p\rangle + a$, where p is a real vector and a is a real number. Since $\mathcal{N}(T) \neq \emptyset$, we have that $\langle x, Tx\rangle = 0$ for all $x \in \mathcal{N}(T)$ and so, we will investigate the minimization problem

$$\text{minimize} \ \ \Phi(x), \ x \in \mathfrak{S}_D \cap \mathcal{N}(T)^{\perp}, \tag{3}$$

under the assumption $\mathfrak{S}_D \cap \mathcal{N}(T)^{\perp} \neq \emptyset$.

Proposition 4.1 [14] If $T \in \mathbb{M}_n(\mathbb{R})$ is a positive semidefinite matrix, then there exists an orthogonal matrix U and invertible diagonal matrix T_1 such that

$$T = U^{\top}(T_1 \oplus O)U = U^{\top} \begin{bmatrix} T_1 & 0 \\ 0 & 0 \end{bmatrix} U.$$

Also, there exists a unique matrix X such that $X^2 = T$ which is also an EP matrix, and which satisfies the $X = U^{\top}(R \oplus O)U$ and $X^{\dagger} = U^{\top}(R^{-1} \oplus O)U$, where $R^2 = T_1$.

Theorem 4.2 [14] For a given square matrix A, let $k = ind(A)$. The following vector, denoted by $\hat{x}_1$ is an approximate solution of the problem (3):

$$\hat{x}_1 = X^{\dagger}(A^{k+1}X^{\dagger})^{\dagger}A^k \left(b + \tfrac{1}{2}AT^{\dagger}p_1\right) - \tfrac{1}{2}T^{\dagger}p_1,$$

where $X^2 = T$ and $p_1 = P_{\mathcal{R}(T)}(p)$.

We will apply Theorem 4.2 to the dual optimization problems defined in 3.1 and present both solutions.

Theorem 4.3 Let H be a $n \times n$ singular positive semidefinite matrix, with $X^2 = H$. Let A be $n \times n$ and $k = ind(A)$, where A is singular and consider the inconsistent equation

$Ax = b$. Therefore, the set $\mathfrak{S}_D = \{x|\ x \in \mathbb{R}^n, A^{k+1}x = A^k b,\ k \geqslant indA\}$ and a vector $g \in \mathbb{R}^n$. Then, the primal problem

$$\begin{aligned} &\underset{x}{minimize} \quad \Phi(x) = \frac{1}{2}\langle x, Hx\rangle + \langle g, x\rangle, \\ &\text{subject to} \quad\ x \in \mathfrak{S}_D \cap \mathcal{N}(H)^{\perp}, \end{aligned}$$

has the following two vectors, denoted by $\hat{x}_1$ and $\hat{x}_2$, as approximate solutions:

$$\begin{aligned} \hat{x}_1 &= X^{\dagger}(A^{k+1}X^{\dagger})^{\dagger}A^k\left(b + AH^{\dagger}g_1\right) - H^{\dagger}g_1, \\ \hat{x}_2 &= X^{\dagger}\left(A^k X^{\dagger}\right)^D A^D A^k\left(b + AH^{\dagger}g_1\right) - AH^{\dagger}g_1, \end{aligned}$$

where $X^2 = H$ and $g_1 = P_{\mathcal{R}(H)}(g) = HH^{\dagger}g$.

Proof. It comes straightforward from Theorem 4.2, replacing p by $2g$. ■

Theorem 4.4 Let H be a $n \times n$ singular positive semidefinite matrix, with $X^2 = H$. Let A be $n \times n$ and $k = ind(A)$, where A is singular, a vector $g \in \mathbb{R}^n$ and the equation $A^{k+1}x = A^k b$. If the set $S = \{x : Hx + g = (A^{k+1})^{\top}\lambda\}$ is not empty, then the dual problem

$$\begin{aligned} &\underset{x}{maximize} \quad K(x,\lambda) = -\frac{1}{2}\langle x, Hx\rangle + \langle A^k b, \lambda\rangle, \\ &\text{subject to} \quad\ Hx + g = (A^{k+1})^{\top}\lambda, \quad x \in \mathfrak{S}_D \cap \mathcal{N}(H)^{\perp} \end{aligned}$$

has the unique solution for x in terms of λ: $\hat{u} = X^{\dagger}(HX^{\dagger})^{\dagger}((A^{k+1})^{\top}\lambda - g)$.

Proof. Obviously the solution of dual problem is equivalent to the solution of the following minimization problem:

$$\begin{aligned} &\underset{x}{minimize} \quad \frac{1}{2}\langle x, Hx\rangle - \langle b, \lambda\rangle \\ &\text{subject to} \quad\ Hx = (A^{k+1})^{\top}\lambda - g, \quad x \in \mathfrak{S}_D \cap \mathcal{N}(H)^{\perp}. \end{aligned}$$

Minimizing $\langle x, Hx\rangle = \langle Xx, Xx\rangle$ is the same as minimizing $||Xx||_2$. Let $y = Xx$, then $X^{\dagger}y = X^{\dagger}Xx = x$, since $x \in \mathcal{N}(H)^{\perp} = \mathcal{R}(H) = \mathcal{R}(X) = \mathcal{R}(H^{\top})$, so we get $x = X^{\dagger}y$. Replacing this value of x in the constraint equation we have $HX^{\dagger}y = (A^{k+1})^{\top}\lambda - g$. So, the minimal norm solution for y is given by $\hat{y} = (HX^{\dagger})^{\dagger}((A^{k+1})^{\top}\lambda - g)$. Replacing back to find the minimal norm solution for x, we get that the minimum norm solution for x in terms of λ is $\hat{u} = X^{\dagger}(HX^{\dagger})^{\dagger}((A^{k+1})^{\top}\lambda - g)$. It is obvious again that the dimension of the Lagrange multipliers vector λ depends on the dimensions of the constraint matrix A, in order for $A^T\lambda$ to be a column vector. ■

Remark 1 *Once more, we need to discuss on the KKT conditions. The difference between this case and the previous one is that the constraint $Ax = b$ has been replaced by the equation $A^{k+1}x = A^k b$, so the new KKT matrix will be*

$$K = \begin{bmatrix} H & (A^{k+1})^{\top} \\ A^{k+1} & 0 \end{bmatrix}.$$

As we have seen in Lemma 3.4 one of the conditions is that the matrix A^{k+1} must have full row rank. In this case, this condition is never satisfied, as A is a singular square matrix with index k. Therefore the KKT conditions never hold when the constraint equation is inconsistent.

Optimal choice of λ using the general normal equation: Once more, we have solved the primal and Dual problems, and the solutions found are:

$$\hat{x} = X^\dagger (A^{k+1}X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1\right) - H^\dagger g_1, \quad \hat{u} = X^\dagger (HX^\dagger)^\dagger ((A^{k+1})^\top \lambda - g).$$

So we need to minimize the 2-norm of their difference in terms of λ; that is, $\min_\lambda ||\hat{x} - \hat{u}(\lambda)||_2$ or

$$\min_\lambda \left|\left|(X^\dagger (A^{k+1}X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1\right) - H^\dagger g_1 - X^\dagger (HX^\dagger)^\dagger ((A^{k+1})^\top \lambda - g)\right|\right|_2$$

which is equivalent to the minimal norm solution of the equation

$$(X^\dagger (A^{k+1}X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1\right) - H^\dagger g_1 + X^\dagger (HX^\dagger)^\dagger g = X^\dagger (HX^\dagger)^\dagger (A^{k+1})^\top \lambda.$$

Therefore, using Proposition 3.8, again we have that

$$\lambda^* = (X^\dagger (HX^\dagger)^\dagger (A^{k+1})^\top)^\dagger \left((X^\dagger (A^{k+1}X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1\right) - H^\dagger g_1 + X^\dagger (HX^\dagger)^\dagger g\right).$$

Once more, to clarify the above discussion we present the following example.

***Example* 4.5** Let $\mathcal{H} = \mathbb{R}^3$, and the positive semidefinite matrix

$$H = \begin{bmatrix} 14 & 28 & 1 \\ 28 & 56 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

We are looking for the minimum of the functional $f(x) = \frac{1}{2}\langle x, Hx\rangle + \langle g, x\rangle$ with $g = (2,3,1)^\top$ and the set of constraints $\mathfrak{S}_D$ defined as $A^2x = Ab$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 5 & 6 & 7 \end{bmatrix}, \quad ind(A) = 1, \quad b = (-1, 1, -1)^\top, \quad g_1 = (1.6, 3.2, 1)^\top.$$

Using Theorem 4.3, we can see that the minimizing vector of $f(x)$ under $\{A^2x = Ab, x \in \mathcal{N}(H)^\perp\}$ is

$$\hat{x} = X^\dagger (A^{k+1}X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1\right) - H^\dagger g_1 = (0.6951, -0.3902, -0.3049)^\top.$$

It is easy to verify that the solution found satisfies the equation $A^2\hat{x} = Ab$. We will now deal with the dual problem using Theorem 4.4. As we have seen, this problem has the unique solution with respect to λ, $\hat{u} = X^\dagger (HX^\dagger)^\dagger ((A^{k+1})^T \lambda - g)$. By calculations we get that

$$\hat{u} = \begin{bmatrix} 687.2 & 478.5 & 1852.9 \\ -343.6 & -239.3 & -926.4 \\ 13 & 9 & 34.9 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} - \begin{bmatrix} 20.38 \\ -10.15 \\ 0.459 \end{bmatrix}.$$

If we set this solution equal to the solution found from the primal problem, we have that

$$\hat{u} = \hat{x} \Rightarrow \begin{bmatrix} 687.2 & 478.5 & 1852.9 \\ -343.6 & -239.3 & -926.4 \\ 13 & 9 & 34.9 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} - \begin{bmatrix} 20.38 \\ -10.15 \\ 0.459 \end{bmatrix} = \begin{bmatrix} 0.6951 \\ -0.3902 \\ -0.3049 \end{bmatrix}.$$

Solving these 3 equations, the value of λ is: $\lambda^* = \begin{bmatrix} -3.238 \\ 6.111 \\ -0.366 \end{bmatrix}$. On the other hand, using the equation

$$\lambda^* = (X^\dagger (HX^\dagger)^\dagger (A^{k+1})^\top)^\dagger \left((X^\dagger (A^{k+1} X^\dagger)^\dagger A^k \left(b + AH^\dagger g_1 \right) - H^\dagger g_1 + X^\dagger (HX^\dagger)^\dagger g \right),$$

we get exactly the same result: $\lambda^* = \begin{bmatrix} -3.238 \\ 6.111 \\ -0.366 \end{bmatrix}$.

5. Conclusion

In this work we presented a new way to treat the classical quadratic optimization problem, when the corresponding quadratic for is positive semidefinite. Using only the vectors perpendicular to its kernel we tackled the problem and found a unique solution. We solved both the primal and dual problems and proved that these two solutions coincide. Our main tool for this goal was the Moore-Penrose inverse. Using the proposed method, the Lagrange multipliers can be found using a closed formula without the need of differentiating. Sufficient conditions are given so that the solutions exists and is unique. Many possible applications of the presented method can be found in economics, mathematical finance, electrical engineering and other fields of applied mathematics, where we deal with quadratic optimization problems.

References

[1] A. B. Israel, Generalized inverses and the Bott-Duffin network analysis, J. Math. Anal. Appl. 7 (1963), 428-435.

[2] A. Ben-Israel, T. N. E. Grenville, Generalized Inverses: Theory and Applications, Springer-Verlag, Berlin, 2002.

[3] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[4] S. L. Campbell, C. D. Meyer, Generalized Inverses of Linear Transformations, Dover Publications Inc, New York, 1991.

[5] W. S. Dorn, Duality in quadratic programming, Quarterly of Applied Mathematics. 18 (2) (1960), 155-162.

[6] C. W. Groetsch, Generalized inverses of linear operators, Marcel Dekker Inc, New York, 1977.

[7] D. Luenberger, Optimization by Vector Space Methods, Wiley Publ, New York, 1969.

[8] R. K. Manherz, S. L. Hakimi, The generalized inverse network analysis and quadratic-error minimization problems, IEEE Trans. Circuit Theory. (1969), 559-562.

[9] J. Nocedal, S. Wright, Numerical Optimization, Springer, 2006.

[10] D. Pappas, Minimization of constrained quadratic forms in Hilbert spaces, Ann. Funct. Anal. (2) (1) (2011), 1-12.

[11] D. Pappas, Restricted linear constrained minimization of quadratic functionals, Linear and Multilinear Algebra. 61 (10) (2013), 1394-1407.

[12] D. Pappas, A. Perperoglou, Constrained matrix optimization with applications, J. Appl. Math. Comput. 40 (2012), 357-369.

[13] H. Schwerdtfeger, Introduction to Linear Algebra and the Theory of Matrices, Noordhoff, Groningen, 1950.

[14] P. Stanimirovic, D. Pappas, S. Miljkovic, Minimization of quadratic forms using the Drazin-inverse solution, Linear and Multilinear Algebra. 62 (2) (2014), 252-266.

[15] P. Stoica, A. Jakobsson, J. Li, Matrix optimization result with DSP applications, Dig. Signal Proc. 6 (1996), 195-201.