The Least Norm Solution of the Linear System via a Quadratic Penalty Function

SULEYMAN SAFAK
Division of Mathematics,
Faculty of Engineering,
Dokuz Eylül University,
35160 Tınaztepe, Buca, İzmir,
TURKEY.
suleyman.safak@deu.edu.tr

Abstract: - In this paper, the linear system of m equations in n unknowns is formulated as a quadratic programming problem, and the least norm solution for the consistent or inconsistent system of the linear equations is investigated using the optimality conditions of the quadratic penalty function (QPF). In addition, several algebraic characterizations of the equivalent cases of the QPFs are given using the orthogonal decompositions and the generalized inverses of the coefficient matrices obtained from optimality conditions. It is seen that the least norm solution of the consistent or inconsistent system of the linear equations can be found with the penalty method. In addition, it is shown that the method can be applied to all systems of the linear equations to find the least norm solution. Numerical examples are presented by using the analytic results that were obtained.

Key-Words: - Least norm solution, Linear systems, Quadratic programming, Penalty function

1 Introduction

In this paper, the system of the linear equations $A\boldsymbol{x}=\boldsymbol{b}$ is considered, where A is an $m\times n$ matrix and $\boldsymbol{b}$ is an $m\times 1$ vector. It is assumed that the elements of A and $\boldsymbol{b}$ are real. We consider the general case of a rectangular matrix with rank $r\leq n$, where the system $A\boldsymbol{x}=\boldsymbol{b}$ is underdetermined ($m<n$), overdetermined ($m>n$) or square ($m=n$). The system of the linear equations is investigated depending on the rank of the coefficient matrix of the system. Since the linear system $A\boldsymbol{x}=\boldsymbol{b}$ is consistent and the rank of the matrix A is $r<n$, the system has infinite solutions. In this case, the minimum Euclidian-norm solution of the system happens to be of great interest [2, 3, 4, 14, 15, 19, 27]. In addition, the system has a unique solution when $rank(A)=n$. If the system $A\boldsymbol{x}=\boldsymbol{b}$ is inconsistent, then there is no solution for the system of the linear equations. We want to find the best approximate solution with minimum norm of the inconsistent system [3, 15, 27].

The least norm problem of the consistent linear system is formulated as an optimization problem and various methods and algorithms for obtaining minimum norm solution to the consistent linear system are developed by using the singular value decomposition, orthogonal decomposition, LU factorization and especially the generalized inverse of a matrix. [2, 3, 4, 8, 9, 19]. The optimal solution via QR factorization and Lagrange multipliers are investigated and the least norm solution of a consistent linear equation $A\boldsymbol{x}=\boldsymbol{b}$ is given in the form a determinant, which reduces to Cramer's rule if A is nonsingular [2, 9, 14, 30].

It has been known for many years that the best approximate solution with minimum norm of the inconsistent system of the linear equations $A\boldsymbol{x}\approx\boldsymbol{b}$ is obtained by several methods using singular value decomposition of a matrix and the generalized inverses, especially the least squares and the regularization methods [3, 5, 7, 12, 13, 15, 16, 17, 21, 27]. Tikhonov regularization, which is the most popular regularization method, in its simplest form, replaces the linear system of $A\boldsymbol{x}=\boldsymbol{b}$ by the minimization problem

$$\underset{\boldsymbol{x}}{Min}\left\{\left\|A\boldsymbol{x}-\boldsymbol{b}\right\|^{2}+\frac{1}{\mu}\left\|\boldsymbol{x}\right\|^{2}\right\},$$

where $\mu > 0$ is a regularization parameter [13, 18, 21]. The least squares solution of the inconsistent system of the linear equations is computed by the use of the method of the normal equations and also the least solution of the system is found via QR decomposition, bidiagonal decomposition and Householder algorithm [8, 14, 17]. The best approximate solution problem of the singular system is one of the most interesting topics of active researchers in the computational mathematics and mathematical programming and has been widely applied in various areas such as engineering problems and other related areas [3, 15, 27]. The paper by Penrose [23] describes the generalized inverse of a matrix, as the unique solution of a certain set of equations. The best approximate solution of the system of linear equations is found by the method of least squares and a further relevant application is depicted in [24]. The best approximate solution with minimum norm of the inconsistent linear system can be computed and found by the penalty method [28].

The least squares method is commonly used in the linear, quadratic and mathematical programming problems [7, 12, 13, 29] and this method is applied to the best approximate solution for the inconsistent system of the linear equations [13, 16, 21]. The analytical and approximate methods for consistent and inconsistent systems of linear equations are developed by using the methods of the singular value decomposition and generalized inverse of a matrix [3, 12, 15, 20, 21, 23, 24]. Moreover, the optimal solutions of the linear, nonlinear and quadratic programming problems are found and investigated by applying the penalty method [1, 11, 25, 26].

In this paper, the linear system of m equations in n unknowns is considered. We first express a consistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ for any $m \times n$ matrix A of rank r as a quadratic programming problem with the minimum norm and we formulate the QPF as an unconstrained optimization problem. Then we investigate the least norm solution of the problem using the optimality conditions of the QPF. In similar manner, the best approximate solution with minimum norm of the inconsistent linear system is formulated as the QPF using the least squares method and then the solution is found from the necessary and sufficient conditions of the QPF. In addition, several algebraic characterizations of the equivalent cases of the QPF are given using the orthogonal decompositions of the coefficient matrices obtained from optimality conditions.

2The linear system and its formulation as a quadratic programming problem

We now consider the linear system $A\boldsymbol{x} = \boldsymbol{b}$ for any $m \times n$ matrix A of rank r. A necessary and sufficient condition for the equation $A\boldsymbol{x} = \boldsymbol{b}$ to have a solution is

$$AA^{+}\boldsymbol{b} = \boldsymbol{b}, \tag{2.1}$$

in this case, the general solution is

$$\boldsymbol{x} = A^{+}\boldsymbol{b} + \left(I - A^{+}A\right)\boldsymbol{y}, \tag{2.2}$$

where $\boldsymbol{y}$ is arbitrary vector and the $n \times m$ matrix A^{+} is the generalized inverse of A [3, 15, 27].

If $A\boldsymbol{x} = \boldsymbol{b}$ is a consistent linear system and $r = n$, then the unique solution of the system is

$$\boldsymbol{x} = A^{+}\boldsymbol{b} = \left(A^{T}A\right)^{-1}A^{T}\boldsymbol{b}. \tag{2.3}$$

where A^T is the transpose of A. In this case, $rank(A) = rank[A,\boldsymbol{b}]$ and the linear system $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, where $[A,\boldsymbol{b}]$ augmented matrix by $\boldsymbol{b}$. If the rank of A is less than n, the matrix $A^T A$ is not invertible and $\boldsymbol{x}$ is not uniquely determined by $A\boldsymbol{x} = \boldsymbol{b}$. Then, we have to choose one with the minimum norm of those many vectors that satisfy $A\boldsymbol{x} = \boldsymbol{b}$. Furthermore, this problem known as the least norm solution of the consistent system $A\boldsymbol{x} = \boldsymbol{b}$ is formulated as a quadratic programming problem:

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^{T}\boldsymbol{x} \mid A\boldsymbol{x} = \boldsymbol{b} \right\}, \tag{2.4}$$

where $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^{T}\boldsymbol{x}}$.

In addition, if A is right- invertible matrix, then

$$\boldsymbol{x} = A^{T}\left(AA^{T}\right)^{-1}\boldsymbol{b}$$

is the optimal solution of the least norm problem (2.4), where $rank(A) = m = Min(m,n)$ and $A^{+} = A^{T}\left(AA^{T}\right)^{-1}$ is the right g-inverse of the matrix A. Otherwise if $rank(A) < m = Min(m,n)$, $\boldsymbol{x}$ is not uniquely determined by $A\boldsymbol{x} = \boldsymbol{b}$ [3, 5, 15, 27].

Now, let an inconsistent system of m equations in n unknowns be $A\boldsymbol{x} = \boldsymbol{b}$, where $rank(A) \neq rank[A,\boldsymbol{b}]$. The least squares solution

(LSS) to the inconsistent linear system $A\boldsymbol{x}=\boldsymbol{b}$ satisfies

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b}\,, \tag{2.5}$$

which is known as the normal equation of the linear system [3, 5, 15, 27]. If the rank of A is n, then the solution defined in (2.3) is the unique LSS. Note that the solution defined in (2.3) is known as the approximate solution of the inconsistent system.

If the rank of A is less than n, the matrix $A^T A$ is not invertible and $\boldsymbol{x}$ is not uniquely determined by $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$. Then, this system of the normal equations has infinitely many solutions, but we have to seek the solution such that $\|x\|$ is minimized.
The optimal solution, among all solutions of $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$, is the one that has the minimum length of errors $\|A\boldsymbol{x}-\boldsymbol{b}\|$. This solution is also called the best approximate LSS of any inconsistent linear system $A\boldsymbol{x}=\boldsymbol{b}$ [3, 5 - 7, 12, 13, 16 - 21]. Furthermore, the minimum norm solution problem of the inconsistent linear system $A\boldsymbol{x}=\boldsymbol{b}$ can be expressed as the following quadratic problem:

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^T\boldsymbol{x} \;\middle|\; A^T A\boldsymbol{x} = A^T\boldsymbol{b} \right\}. \tag{2.6}$$

Note that the quadratic programming problems (2.4) and (2.6) have the consistent linear constraints and common algebraic characterizations. When the system of $A\boldsymbol{x}=\boldsymbol{b}$ is consistent, the problems (2.4) and (2.6) are equivalent and have the same solution. It is well known that if the system of $A\boldsymbol{x}=\boldsymbol{b}$ is inconsistent, the solution of the problem (2.6) is the best approximate solution with minimum norm of the system. In section 3, we present the QPFs of the problems defined in (2.4) and (2.6) and give the main results using their optimality conditions. Numerical examples are given in the forthcoming sections of the study and calculated with the use of the results obtained.

3 The least norm solution via a quadratic penalty function

Many efficient methods have been developed for solving the quadratic programming problems [1, 11, 18, 22, 29], one of which is the penalty method. In this class of methods we replace the original constrained problem with unconstrained problem that minimizes the penalty function [1, 10, 11, 22, 28].

We assume that the quadratic programming problems (2.4), (2.6) have feasible solutions. To find the least norm solution for the consistent system of the linear equations, the QPF of the problem (2.4) can be defined as

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} + \frac{1}{2}q\|A\boldsymbol{x}-\boldsymbol{b}\|^2\,, \tag{3.1}$$

where the scalar quantity q is the penalty parameter.

From the first order necessary conditions for the unconstrained minimum of the QPF defined in (3.1), we obtain

$$\nabla f(\boldsymbol{x}) = \boldsymbol{x} + qA^T A\boldsymbol{x} - qA^T\boldsymbol{b} = 0\,. \tag{3.2}$$

We also obtain the Hessian matrix of (3.1), which represents the sufficient condition as

$$H(\boldsymbol{x}) = I + qA^T A\,, \tag{3.3}$$

where $\nabla f(\boldsymbol{x}) = \left[\dfrac{\partial f}{\partial x_1},\ \dfrac{\partial f}{\partial x_2},\ \ldots,\ \dfrac{\partial f}{\partial x_n}\right]^T$ and $H(\boldsymbol{x}) = \left[\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}\right]$ for $i, j = 1, 2, \ldots, n$.

Corollary 3.1. The Hessian matrix $H(\boldsymbol{x}) = I + qA^T A$ of the QPF defined in (3.1) is positive definite.

Proof. Let eigenvalues be $\lambda_1, \lambda_2, ..., \lambda_n$ of the $n \times n$ matrix $A^T A$. It is clear that eigenvalues of the matrix $A^T A$ are $\lambda_i \geq 0$ for $i = 1, 2, \ldots, n$. The eigenvalues of the Hessian matrix $H(\boldsymbol{x})$ are $1 + q\lambda_i > 0$, where the penalty parameter $q > 0$. So the Hessian matrix is positive definite.

Now we can establish the following theorem for the least norm solution of the consistent linear system $A\boldsymbol{x}=\boldsymbol{b}$.

Theorem 3.1. Let the system of the linear equations $A\boldsymbol{x}=\boldsymbol{b}$ be consistent. Then the solution with minimum norm of $A\boldsymbol{x}=\boldsymbol{b}$ is

$$\boldsymbol{x} = \underset{q\to\infty}{Lim}\left(\frac{1}{q}I + A^T A\right)^{-1} A^T\boldsymbol{b}\,, \tag{3.4}$$

where $\mathbf{det}\left(\frac{1}{q}I + A^T A\right) \neq 0$ for large number $q > 0$ and $rank(A) \leq n$.

Proof. From (3.2), we obtain

$$\left(\frac{1}{q}I + A^T A\right)\boldsymbol{x} = A^T\boldsymbol{b}. \tag{3.5}$$

Using $\mathbf{det}\left(\frac{1}{q}I + A^T A\right) \neq 0$ for $q \neq 0$ and applying (2.2), we get

$$\boldsymbol{x} = \left(\frac{1}{q}I + A^T A\right)^{-1} A^T\boldsymbol{b}.$$

This vector is the solution with minimum norm of the QPF (3.1) and its Hessian matrix H is positive definite. Using (2.3), we see that

$$\lim_{q\to\infty}\left(\frac{1}{q}I + A^T A\right)^{-1} = \left(A^T A\right)^+$$

and

$$\boldsymbol{x} = \left(A^T A\right)^+ A^T\boldsymbol{b} = A^+\boldsymbol{b}. \tag{3.6}$$

In addition, let the orthogonal decomposition of the matrix $A^T A$ be

$$A^T A = VDV^{-1},$$

where $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$, $D = diag\ \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\boldsymbol{v}_i$ are eigenvectors corresponding to eigenvalues λ_i of the matrix $A^T A$. Using the decomposition of $A^T A$, the solution (3.4) can be expressed as

$$\boldsymbol{x} = \lim_{q\to\infty} V\left(\frac{1}{q}I + D\right)^{-1} V^{-1}A^T\boldsymbol{b},$$

where

$\frac{1}{q}I + D = diag\left\{\lambda_1 + \frac{1}{q}, \lambda_2 + \frac{1}{q}, \ldots, \lambda_n + \frac{1}{q}\right\}$. From (3.6), we see that

$$\left(A^T A\right)^+ = \lim_{q\to\infty} V\left(\frac{1}{q}I + D\right)^{-1} V^{-1} = VD^+V^{-1},$$

where D^+ is the generalized inverse of D and $D^+ = D^{-1}$ when $rank(A) = n$. Then the proof is completed.

The formula of the least norm solution of the consistent system given in Theorem 3.1 is the same formula with the best approximate solution of the inconsistent system obtained by Safak [28]. This result shows that the same formula can be applied to all systems which are either consistent or inconsistent.

We can calculate the solution with the minimum norm correct to desired decimal places of the consistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ using

$$\boldsymbol{x}_k \approx \left(\frac{1}{q_k}I + A^T A\right)^{-1} A^T\boldsymbol{b}, \tag{3.7}$$

where $q_k = 10^k$ for $k = 0, 1, 2, \ldots$.

In addition, the generalized inverse of the matrix A is computed approximately by any q_k in (3.7) as follows:

$$A^+ \approx \left(\frac{1}{q_k}I + A^T A\right)^{-1} A^T \tag{3.8}$$

The solution $\boldsymbol{x}_k$ converges to the least norm solution $\boldsymbol{x}$ as $q_k \to \infty$. In matrix term, as $q_k \to \infty$

$$\left(\frac{1}{q_k}I + A^T A\right)^{-1} A^T \to A^+.$$

The following example is presented by using (3.7). First of all, the solution with minimum norm of the consistent system is calculated using (3.4) and secondly, the approximate solution correct to six decimal places of the given system is computed directly for $q_k = 10^6$. In addition, the generalized inverse of the matrix A is computed approximately by taking $q_k = 10^6$ in (3.8) as the following example.

Example 3.1. Let $A\boldsymbol{x} = \boldsymbol{b}$ be a consistent system, where the matrix A of rank 2 and $\boldsymbol{b}$ are

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

We computed that the least norm solution $\boldsymbol{x}$ for full row rank of the given system of the linear equations using the right inverse of the coefficient matrix A as follows:

$$\boldsymbol{x} = A^T(AA^T)^{-1}\boldsymbol{b} = \begin{bmatrix} 1/4 & -1/4 \\ 1/4 & -1/4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.5 \end{bmatrix}$$

By Theorem 3.1 and from (3.4), we computed as

$$\boldsymbol{x} = \lim_{q\to\infty} \frac{q^3}{8q^2+6q+1}\begin{bmatrix} \frac{4q^2+4q+1}{q^2} & -\frac{4q+2}{q} & 0 \\ -\frac{4q+2}{q} & \frac{4q^2+4q+1}{q^2} & 0 \\ 0 & 0 & \frac{4q+1}{q^2} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \lim_{q\to\infty}\begin{bmatrix} \frac{2q^2+q}{8q^2+6q+1} & -\frac{2q^2+q}{8q^2+6q+1} \\ \frac{2q^2+q}{8q^2+6q+1} & -\frac{2q^2+q}{8q^2+6q+1} \\ \frac{4q^2+q}{8q^2+6q+1} & \frac{4q^2+q}{8q^2+6q+1} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \lim_{q\to\infty}\begin{bmatrix} \frac{2q^2+q}{8q^2+6q+1} \\ \frac{2q^2+q}{8q^2+6q+1} \\ \frac{4q^2+q}{8q^2+6q+1} \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.5 \end{bmatrix}.$$

It is seen that

$$A^+ = \lim_{q\to\infty}\begin{bmatrix} \frac{2q^2+q}{8q^2+6q+1} & -\frac{2q^2+q}{8q^2+6q+1} \\ \frac{2q^2+q}{8q^2+6q+1} & -\frac{2q^2+q}{8q^2+6q+1} \\ \frac{4q^2+q}{8q^2+6q+1} & \frac{4q^2+q}{8q^2+6q+1} \end{bmatrix} = \begin{bmatrix} 1/4 & -1/4 \\ 1/4 & -1/4 \\ 1/2 & 1/2 \end{bmatrix}$$

$$= A^T(AA^T)^{-1}.$$

In Table 1, some values for the approximate solution of the given consistent system using (3.7) are listed with seven decimals for $k = 0,1,2,\ldots,6$.

k	x_1	x_2	x_3
0	0.2000000	0.2000000	0.3333333
1	0.2439024	0.24390243	0.4761904
2	0.2493765	0.2493765	0.4975124
3	0.2499375	0.2499375	0.4997501
4	0.2499937	0.2499937	0.4999750
5	0.2499993	0.2499993	0.4999975
6	0.2499999	0.2499999	0.4999997

Table 1 The least norm solution correct to desired decimal places of the consistent system

In addition, the generalized inverse of the matrix A is computed approximately by taking $q_k = 10^6$ in (3.8) as follows:

$$A^+ \approx \begin{bmatrix} 0.2499999 & -0.2499999 \\ 0.2499999 & -0.2499999 \\ 0.4999997 & 0.4999997 \end{bmatrix}.$$

Using this matrix, we can compute the least norm solution correct to six decimal places of the given consistent system as

$$\boldsymbol{x}^T = [0.2499999 \quad 0.2499999 \quad 0.4999997].$$

It is obvious that the exact least norm solution justifies being extremely close to this approximate solution.

We now assume that the linear system $A\boldsymbol{x} = \boldsymbol{b}$ for any $m\times n$ matrix A of rank r is inconsistent. To find the best approximate solution with the minimum norm for this inconsistent system of the linear equations, the penalty function of the problem (2.6) can be defined as

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} + \frac{1}{2}q\left\|A^TAx - A^T\boldsymbol{b}\right\|^2. \tag{3.9}$$

From the first order necessary conditions for the unconstrained minimum of the QPF (3.9), we obtain

$$\nabla g(\boldsymbol{x}) = \boldsymbol{x} + qA^TA\left(A^TA\boldsymbol{x} - A^T\boldsymbol{b}\right) = 0. \tag{3.10}$$

We also obtain the Hessian matrix of (3.9), which represents the sufficient condition as

$$H_g(\boldsymbol{x}) = I + q\left(A^TA\right)^2. \tag{3.11}$$

Corollary 3.2. The Hessian matrix $H_g(\boldsymbol{x}) = I + q\left(A^TA\right)^2$ of the QPF (3.9) is positive definite.

Proof. Let eigenvalues be $\lambda_1,\lambda_2,\ldots,\lambda_n$ of the $n\times n$ matrix A^TA. It is clear that eigenvalues of the matrix A^TA are $\lambda_i \geq 0$. The eigenvalues of the Hessian matrix $H_g(\boldsymbol{x})$ are $1+q\lambda_i^2 > 0$, where the penalty parameter $q > 0$. So the Hessian matrix is positive definite.

Now we can establish the following theorem for the best approximate solution of the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$.

Theorem 3.2. Let the system of the linear equations $A\boldsymbol{x} = \boldsymbol{b}$ be inconsistent. Then the best approximate solution of $A\boldsymbol{x} = \boldsymbol{b}$ is

$$\boldsymbol{x} = \lim_{q\to\infty}\left(\frac{1}{q}I + (A^TA)^2\right)^{-1}(A^TA)A^T\boldsymbol{b}, \tag{3.12}$$

where $\det\left(\frac{1}{q}I + (A^TA)^2\right) \neq 0$ for large number $q > 0$ and $rank(A) \leq n$.

Proof. From (3.10), we obtain

$$\left(\frac{1}{q}I + (A^TA)^2\right)\boldsymbol{x} = (A^TA)A^T\boldsymbol{b}. \tag{3.13}$$

Using $\det\left(\frac{1}{q}I + (A^TA)^2\right) \neq 0$ for $q \neq 0$ and applying (2.2), we get

$$\boldsymbol{x} = \left(\frac{1}{q}I + (A^TA)^2\right)^{-1}(A^TA)A^T\boldsymbol{b}.$$

This solution is the best approximate solution with minimum norm of the QPF (3.9) and its Hessian matrix H_g is positive definite. Using (2.3), we see that

$$\boldsymbol{x} = \left[\left(A^TA\right)^2\right]^+\left(A^TA\right)A^T\boldsymbol{b} = \left(A^TA\right)^+A^T\boldsymbol{b} \tag{3.14}$$

and

$$\lim_{q\to\infty}\left(\frac{1}{q}I + \left(A^TA\right)^2\right)^{-1} = \left[\left(A^TA\right)^2\right]^+.$$

Then the proof is completed.

Using the orthogonal decomposition of the matrix A^TA, we also express the matrix $\frac{1}{q}I + \left(A^TA\right)^2$ as

$$\frac{1}{q}I + \left(A^TA\right)^2 = V\left(\frac{1}{q}I + D^2\right)V^{-1}, \tag{3.15}$$

where

$$\frac{1}{q}I + D^2 = diag\left\{\lambda_1{}^2 + \frac{1}{q}, \lambda_2{}^2 + \frac{1}{q}, \ldots, \lambda_n{}^2 + \frac{1}{q}\right\}.$$

We know that $\mu_i = \lambda_i{}^2 + \frac{1}{q}$ are eigenvalues of the matrix $\frac{1}{q}I + \left(A^TA\right)^2$ when λ_i are eigenvalues of the matrix A^TA for $i = 1, 2, \ldots, n$.

This leads to the following theorem.

Theorem 3.3. Let be the eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the matrix A^TA, respectively. The best approximate solution of the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ is

$$\begin{aligned}\boldsymbol{x} &= V\lim_{q\to\infty}\left(\frac{1}{q}I + D^2\right)^{-1}DV^{-1}A^T\boldsymbol{b} \\ &= \left(A^TA\right)^+A^T\boldsymbol{b} \\ &= A^+\boldsymbol{b},\end{aligned} \tag{3.16}$$

where $D = diag\ \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$.

Proof. Applying Theorem 3.2 and using (3.15), we get

$$\boldsymbol{x} = \lim_{q\to\infty}V\left(\frac{1}{q}I + D^2\right)^{-1}DV^{-1}A^T\boldsymbol{b}.$$

If $rank\ (A) = n$,

$$\boldsymbol{x} = \lim_{q\to\infty}V\begin{bmatrix}\frac{q}{q\lambda_1{}^2+1} & 0 & \cdots & 0 \\ 0 & \frac{q}{q\lambda_2{}^2+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{q}{q\lambda_n{}^2+1}\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}V^{-1}A^T\boldsymbol{b}$$

$$= \lim_{q\to\infty}V\begin{bmatrix}\frac{q\lambda_1}{q\lambda_1{}^2+1} & 0 & \cdots & 0 \\ 0 & \frac{q\lambda_2}{q\lambda_2{}^2+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{q\lambda_n}{q\lambda_n{}^2+1}\end{bmatrix}V^{-1}A^T\boldsymbol{b}$$

$$= V\begin{bmatrix}\frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n}\end{bmatrix}V^{-1}A^T\boldsymbol{b}$$

$$= VD^{-1}V^{-1}A^T\boldsymbol{b} = (A^TA)^{-1}A^T\boldsymbol{b} = A^+\boldsymbol{b}.$$

If $rank\ (A) < n$,

$$\boldsymbol{x} = \underset{q\to\infty}{Lim}\, V\begin{bmatrix} \frac{q\lambda_1}{q\lambda_1^{\,2}+1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \frac{q\lambda_2}{q\lambda_2^{\,2}+1} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{q\lambda_r}{q\lambda_r^{\,2}+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} V^{-1}A^T\boldsymbol{b}$$

$$= V\begin{bmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_r} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} V^{-1}A^T\boldsymbol{b}$$

$$= VD^{+}V^{-1}A^T\boldsymbol{b} = (A^TA)^{+}A^T\boldsymbol{b},$$

where $D = diag\ \{\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0\}$ and

$$\left(\frac{1}{q}I + D^2\right)^{-1} = \begin{bmatrix} \frac{q}{q\lambda_1^{\,2}+1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \frac{q}{q\lambda_2^{\,2}+1} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{q}{q\lambda_r^{\,2}+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & q & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & q \end{bmatrix}$$

Thus the proof is completed.

Here we establish the following corollary which is proved easily using the optimality conditions of the QPF defined in (3.9) and results on the generalized inverses of matrices.

Corollary 3.3. Let λ_i be eigenvalues of the matrix A^TA and let $\mu_i = \lambda_i^{\,2} + \frac{1}{q}$ be eigenvalues of the matrix $\frac{1}{q}I + \left(A^TA\right)^2$. A necessary and sufficient condition for the QPF defined in (3.9) to have a best approximate solution with minimum norm is $\mu_i = \lambda_i^{\,2} + \frac{1}{q} > 0$, in which case the optimal solution is $\boldsymbol{x} = A^{+}\boldsymbol{b}$.

Proof. Since the Hessian matrix given in Corollary 3.2 is a positive definite matrix and eigenvalues of the matrix $H_g(\boldsymbol{x})$ are $1 + q\lambda_i^{\,2} > 0$, then $\mu_i = \lambda_i^{\,2} + \frac{1}{q} > 0$ and the QPF defined in (3.9) has a best approximate solution with minimum norm. From Theorem 3.3, we find the solution $\boldsymbol{x} = A^{+}\boldsymbol{b}$. This completes the proof.

Corollary 3.4. Let any linear system of m equations in n unknowns be $A\boldsymbol{x} = \boldsymbol{b}$ and $rank(A)$ be $r \le n$.
(i) If $A\boldsymbol{x} = \boldsymbol{b}$ is a consistent linear system, then the solution defined in (3.4)

$$\boldsymbol{x} = \underset{q\to\infty}{Lim}\left(\frac{1}{q}I + A^TA\right)^{-1}A^T\boldsymbol{b}$$

is the least norm solution of the consistent system.
(ii) If $A\boldsymbol{x} = \boldsymbol{b}$ is an inconsistent linear system, then the solution $\boldsymbol{x}$ is the best approximate solution with minimum norm of the inconsistent system.

Proof. From Theorem 3.1, Theorem 3.3, the corollary is easily proved.

The results show that the optimum solutions of the consistent and inconsistent system of linear equations can be computed with the same formula.

We can calculate the approximate solution with the minimum norm of the consistent or inconsistent linear systems $A\boldsymbol{x} = \boldsymbol{b}$ using (3.7) and

$$\boldsymbol{x}_k = \left(\frac{1}{q_k}I + \left(A^TA\right)^2\right)^{-1}\left(A^TA\right)A^T\boldsymbol{b}, \qquad (3.17)$$

where $q_k = 10^k$ for $k = 0, 1, 2, \ldots$.

Example 3.2. Suppose that the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ is

$$x_1 + x_2 = 2$$
$$2x_1 + 2x_2 = 2$$
$$3x_1 + 3x_2 = 3$$

where $rank(A) = 1$ and the best approximate solution of the system computed by Graybill [15] is $x_1 = x_2 = 15/28$.

The following solution, which is obtained using (3.6) and (3.14), is the best approximate solution with minimum norm of the given system $A\boldsymbol{x} = \boldsymbol{b}$.

$$\boldsymbol{x} = A^{+}\boldsymbol{b} = \left(A^T A\right)^{+} A^T \boldsymbol{b},$$

$$= \frac{1}{28}\begin{bmatrix}15\\15\end{bmatrix} = \begin{bmatrix}0.535714285714286\\0.535714285714286\end{bmatrix}$$

where A^{+} and $\left(A^T A\right)^{+}$ is computed using (3.6) and (3.14), respectively as

$$A^{+} = \lim_{q\to\infty}\begin{bmatrix}\frac{q}{28q+1} & \frac{2q}{28q+1} & \frac{3q}{28q+1}\\ \frac{q}{28q+1} & \frac{2q}{28q+1} & \frac{3q}{28q+1}\end{bmatrix}$$

$$= \frac{1}{28}\begin{bmatrix}1 & 2 & 3\\1 & 2 & 3\end{bmatrix}$$

and

$$\left(A^T A\right)^{+} A^T = \lim_{q\to\infty}\begin{bmatrix}\frac{28q}{784q+1} & \frac{56q}{784q+1} & \frac{84q}{784q+1}\\ \frac{28q}{784q+1} & \frac{56q}{784q+1} & \frac{84q}{784q+1}\end{bmatrix}$$

$$= \frac{1}{28}\begin{bmatrix}1 & 2 & 3\\1 & 2 & 3\end{bmatrix}.$$

Table 2 and 3 represent the best approximate solution for the optimal solution of the inconsistent linear equation $A\boldsymbol{x} = \boldsymbol{b}$ by using (3.7) and (3.17).

In Table 2, some values for the approximate solution using (3.6) are listed with seven decimals for $k = 0, 1, 2, \ldots, 6$.

k	x_1	x_2
0	0.5172413	0.5172413
1	0.5338078	0.5338078
2	0.5355230	0.5355230
3	0.5356951	0.5356951
4	0.5357123	0.5357123
5	0.5357140	0.5357140
6	0.5357142	0.5357142
7	0.5357142	0.5357142

Table 2 The best approximate solution correct to desired decimal places of the problem

In Table 3, some values for the approximate solution using (3.17) are listed with seven decimals for $k = 0, 1, 2, \ldots, 4$.

k	x_1	x_2
0	0.5350318	0.5350318
1	0.5356459	0.5356459
2	0.5357074	0.5357074
3	0.5357136	0.5357136
4	0.5357142	0.5357142
5	0.5357142	0.5357142

Table 3 The best approximate solution of the example

It is clear that the optimal solution proves to be extremely close to approximate solution. From Table 2 and Table 3, we easily see that the approximate solution correct to seven decimal places of the given system is computed directly for q_k. If the penalty parameter q is taken big enough, the approximate solution correct to desired decimal places of the system can be calculated by using (3.6) and (3.17).

4 Conclusion

In this study, the consistent and inconsistent linear systems of m equations in n unknowns are formulated as a quadratic programming problem, and the least norm solution for the consistent or inconsistent system of the linear equations is investigated using the optimality conditions of the QPFs. Additionally, several algebraic characterizations of the equivalent cases of the QPF are given using the orthogonal decomposition of the coefficient matrices obtained from optimality conditions and the analytic results are compared with numerical examples. It is seen that the least norm solution of the consistent or inconsistent system of the linear equations can be found by the penalty method. The results show that the optimum solutions of the consistent and inconsistent system of linear equations can be computed with the same formula.

References:

[1] Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms. 2nd edn. Wiley, New York, 1993.

[2] Ben-Isreal, A.: A Cramer rule for least-norm solutions of consistent linear equations. Linear Algebra Appl.,Vol 43,1982, pp. 223-226.

[3] Ben-Israel, A., Greville, T.N.E.: Generalized Inverses: Theory and Applications. Wiley, New York, 1974.

[4]Cadzow, J.A.: A finite algorithm for the minimum l_∞ solution to a system of consistent linear equations. SIAM J. Numer. Anal., Vol 10, No 4,1973, pp. 607-614.

[5]Cadzow, J.A.: Minimum l_1, l_2 and l_∞ norm approximate solutions to an overdetermined system of linear equations. Digit. Signal Process. Vol 12, 2002, pp. 524-560.

[6]Cao, Z.H.: On the converge of iterative methods for solving singular linear systems. J. Comput. Appl. Math. Vol 145, 2002, pp 1-9.

[7] Cline, A.K.: An elimination method for the solution of linear least squares problems, SIAM J. Numer. Anal. Vol 10, 1973, pp. 283-289.

[8]Cline, R.E., Plemmons, J.: l_2 Solutions to Undetermined Linear Systems. SIAM Review, Vol 18, No 1, 1976, pp.92-106.

[9] Dax, A., Elden, L.: Approximating minimum norm solutions of rank-deficient least squares problems. Numer. Linear Algebra Appl. Vol **5**,1998, pp. 79–99.

[10]Di Pillo, G., Liuzzi, G., Lucidi, S.: An exact penalty-Lagrangian approach for large-scale nonlinear programming. Optimization, Vol 60, No 1-2, 2011, pp. 223-252.

[11]Dostal, Z.: On penalty approximation of quadratic programming problem. Kybernetika (Prague), Vol 27, No 2, 1991, pp. 151-154.

[12] Elfving, T.: Block-iterative methods for consistent and inconsistent linear equations, Numer. Math., Vol **35**, 1980, pp. 1-12.

[13]Golub, G.H., Hansen P.C., O'leary D.C.: Tikhonov regularization and total least squares. SIAM J. Matrix Anal. Appl.Vol 21, 1999, pp. 1-10.

[14]Golub, G.H., Van Loan, C.F.: Matrix computations. Johns Hopkins Studies in the Mathematical Sciences, 3rd edn. Johns Hopkins Univ. Press, Baltimore, MD 1996.

[15]Graybill, A.: Introduction to Matrices with Applications in Statistics. Wadsworth, Belmont,1969.

[16]Kim, S.J., Koh K., Lustig M., Boyd S., Gorinevsky D.: An interior- point method for large-scale l_1- regularized least squares. IEEE J. Sel. Top. Signal Process. Vol 1, No 4, 2007, pp. 606-617.

[17]Lawson, C.L. , Hanson R.J.: Solving Least Squares Problems, Prentice Hall, Englewood Cliffs, 1974.

[18]Lewis, B., Reichel, L.: Arnoldi-Tikhonov regularization methods. J. Comput. Appl. Math.Vol **226, No** 1,2009, pp. 92–102.

[19]Marlow, W. H.: Mathematics for Operations Research. John Wiley and Sons, Canada, 1978.

[20]Miao, J., Ben-Israel, A.: On l_p – approximate solutions of linear equations. Linear Algebra Appl., Vol **199**, 1994, pp. 305-327.

[21]Neumair, A.: Solving ill-conditioned and singular linear systems: A tutorial on regularization, SIAM Rev., Vol **40**, 1998, pp.636-666.

[22]Özdemir, N., Evirgen F.: A dynamic system approach to quadratic programming problems with penalty method. Bull. Malays. Math. Sci. Soc., Vol **33, No** 1, 2010, pp. 79-91.

[23]Penrose, R.: A generalized inverse for matrices, Proc. Cambridge Philos. Soc., Vol 51, 1955, pp.406-413.

[24]Penrose, R.: On the best approximate solutions of linear matrix equations. Proc. Cambridge Philos. Soc. Vol 52, 1956, pp. 17-19.

[25]Pınar, M.Ç.: Linear programming via a quadratic penalty function, Math. Methods Ope. Res., Vol 44, 1996, pp. 345-370.

[26]Pınar, M.Ç., Elhedhli, S.: A penalty continuation method for the l_∞ solution of the overdetermined linear systems. BIT, Vol 38, No 1, 1998, pp. 127-150.

[27]Rao, C.R., Mitra, S.K.: Generalized Inverse of Matrices and Its Applications. Wiley, New York, 1971.

[28] Safak, S.: The best approximate solution of the inconsistent system via a quadratic penalty function. Bull. Malays. Math. Sci. Soc., Vol 38, No 2, 2015, pp. 683-694.

[29]Übi, E.:A numerically stable least squares solution to the quadratic programming problem. Cent. Eur. J. Math. Vol 6**,** 2008, pp. 171-178.

[30]Wang, G.R.: A Cramer rule for finding the solution of a class of singular equations. Linear Algebra and its Appl. Vol 116, 1989, pp. 27-34.