The best approximate solution of the inconsistent linear system via a quadratic penalty function

S. Safak*

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

In this paper, the inconsistent linear system of m equations in n unknowns is formulated as a quadratic programming problem, and the best approximate solution with the minimum norm for the 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 QPF are given using the orthogonal decompositions of the coefficient matrices obtained from optimality conditions, and analytic results we obtained are satisfied with numerical examples.

AMS Subject Classifications: 90C90, 15A06, 15A09

Keywords: Inconsistent linear equations; least squares problem; penalty function; generalized inverses; quadratic programming.

1. Introduction

In this paper, the solutions of the system of the linear equations $A\boldsymbol{x}=\boldsymbol{b}$ are 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 $r\le n$, where the system $A\boldsymbol{x}=\boldsymbol{b}$ is inconsistent and therefore there is no solution of the system of the linear equations [2, 11, 22].

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 various methods using singular value decomposition of a matrix and the generalized inverses, especially the least squares and the regularization methods [2, 3, 5, 8, 9, 11, 12, 13, 15, 16, 22, 23]. 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

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

where $\mu>0$ is a regularization parameter [9, 14, 16]. The least squares solution of the inconsistent system of the linear equations is computed by use of the method of the normal equations and also the least solution of the system is found via QR decomposition and Householder algorithm [10, 13]. The best approximate solution problem of the singular system is one of the most interest 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 [2, 11, 22]. The paper by Penrose [18] 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 [19]. Rosen [23] gives minimum and basic solutions to singular linear systems and has developed an algorithm for computation.

The least squares method is commonly used in the linear, quadratic and mathematical programming problems [5, 8, 9, 24]. The least squares method is applied to the best approximate solution for the inconsistent system of the linear equations [9, 12, 16]. 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 [2, 8, 11, 15, 16, 18, 19, 23]. Moreover, the optimal solutions of the linear, nonlinear and quadratic programming problems are found and investigated by applying the penalty method [1, 7, 20, 21].

In this paper, the inconsistent linear system of linear equations in n unknowns is considered. Let an inconsistent system of m equations in n unknowns be $A\boldsymbol{x} = \boldsymbol{b}$. The least squares solution (LSS) to the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ satisfies

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b} \tag{1}$$

which is known as the normal equation [2, 3, 11, 22]. If the rank of A is n, then the unique LSS is

$$\boldsymbol{x} = \left(A^T A\right)^{-1} A^T\boldsymbol{b}. \tag{2}$$

The projection of $\boldsymbol{b}$ onto the column space of A is therefore

$$\boldsymbol{p} = A\boldsymbol{x} = A\left(A^T A\right)^{-1} A^T\boldsymbol{b} = AA^{+}\boldsymbol{b}, \tag{3}$$

where

$$A^{+} = \left(A^T A\right)^{-1} A^T \tag{4}$$

is the generalized inverse of A and A^T is the transpose of A [2, 11, 18, 22].

If the columns of A are not linearly independent, the null space of A does not contain the zero vector, the rank of A is less then n, the matrix $A^T A$ is not invertible and $\boldsymbol{x}$ is not uniquely determined by $A\boldsymbol{x} = \boldsymbol{p}$. Then, we have to choose one of those many vectors that satisfy $A\boldsymbol{x} = \boldsymbol{p}$.

The optimal solution, among all solutions of $A\boldsymbol{x} = \boldsymbol{p}$, is the one that has the minimum length of errors $\left\|A\boldsymbol{x} - \boldsymbol{b}\right\|$. This solution is also called the best approximate solution with minimum norm [2-5, 8, 9, 11, 12, 14-16].

Note that the best approximate solution is the optimal LSS of any inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$. If $A\boldsymbol{x} = \boldsymbol{p}$ and $\boldsymbol{x} \in \Re\left(A^T\right)$, then $\boldsymbol{x} = A^{+}\boldsymbol{b}$ is the optimal least squares solution, where $\Re\left(A^T\right)$ is the row space of A [2, 8, 11, 22].

In this study, we first express an inconsistent linear system of m equations in n unknowns as a quadratic programming problem using the least squares method and we formulate the QPF as an unconstrained optimization problem

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

where q is a positive large number which is a penalty parameter. Note that the relationship this optimization problem and the minimization problem given by Tikhonov regularization method is of great interest. Then we investigate the equivalent solutions of the problem using the optimality 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.

2. The inconsistent 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

$$AA^{+}\boldsymbol{b} = \boldsymbol{b}, \tag{5}$$

in this case, the general solution is

$$\boldsymbol{x} = A^{+}\boldsymbol{b} + \left(I - A^{+}A\right)\boldsymbol{y}, \tag{6}$$

where $\boldsymbol{y}$ is arbitrary vector [2, 11, 22].

If $A\boldsymbol{x} = \boldsymbol{b}$ is an inconsistent linear system and $r < n$, then the solution of the normal equation $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ defined in (1) is

$$\boldsymbol{x} = \left(A^T A\right)^{+} A^T\boldsymbol{b} = A^{+}\boldsymbol{b}. \tag{7}$$

Note that the following equations are consistent and equivalent.

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b}, \ A^{+}A\boldsymbol{x} = A^{+}\boldsymbol{b} \text{ and } A\boldsymbol{x} = AA^{+}\boldsymbol{b}, \tag{8}$$

where the $n \times m$ matrix A^{+} satisfies four conditions below [2, 11, 22]

$$(AA^{+})^T = AA^{+}, \ (A^{+}A)^T = A^{+}A, \ AA^{+}A = A \text{ and } A^{+}AA^{+} = A^{+}. \tag{9}$$

$P_A = A(A^T A)^{+} A^T = AA^{+}$ is the matrix that projects a vector $\boldsymbol{b}$ onto the space spanned by the columns of A. If $\boldsymbol{b}$ is not in the column space of A, then the linear system $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent. Using the projection matrix P_A and the properties of the generalized inverse of A, the following results are obtained as

$$P_A A\boldsymbol{x} = P_A\boldsymbol{b}, \ AA^{+}A\boldsymbol{x} = AA^{+}\boldsymbol{b} \text{ and } A\boldsymbol{x} = AA^{+}\boldsymbol{b} = \boldsymbol{p},$$

where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of A. As mentioned in Introduction section, the optimal solution, among all solutions of $A\boldsymbol{x} = \boldsymbol{p}$, is the one that has the minimum length one of errors $\left\| A\boldsymbol{x} - \boldsymbol{b} \right\|$. This solution with minimum norm is also the best approximate LSS of any inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ and the linear systems $A\boldsymbol{x} = p$ and $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ are consistent and equivalent [2, 8, 11, 13, 18, 19, 22]. Furthermore, the minimum norm solution problem of the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ can be expressed as the following quadratic problem:

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^T\boldsymbol{x} \,\middle|\, A\boldsymbol{x} = \boldsymbol{p} \right\}. \tag{10}$$

Using the system of the linear equations defined in (8), the minimum norm problem can be investigated with the following quadratic programming problems

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^T\boldsymbol{x} \,\middle|\, A^T A\boldsymbol{x} = A^T\boldsymbol{b} \right\}, \tag{11}$$

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^T\boldsymbol{x} \,\middle|\, A\boldsymbol{x} = AA^{+}\boldsymbol{b} \right\} \tag{12}$$

and

$$\underset{\boldsymbol{x}}{Min}\left\{ \boldsymbol{x}^T\boldsymbol{x} \,\middle|\, A^{+}A\boldsymbol{x} = A^{+}\boldsymbol{b} \right\}, \tag{13}$$

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

DEFINITION 2.1 (Penalty Function Method): Consider the equality constrained optimization problem as

$$\underset{\boldsymbol{x}}{Min}\left\{ z(\boldsymbol{x}) \,\middle|\, \phi_i(\boldsymbol{x}) = 0, \ i = 1, 2, \ldots, n \right\}.$$

Methods using penalty functions transform a constrained problem into a single unconstrained problem. The most simple and straightforward approach to the constrained problems of the above form is to apply a suitable unconstrained optimization algorithm. The formulation of the penalty function of the constrained problem is

$$P(\boldsymbol{x}, q) = z(\boldsymbol{x}) + \sum_{i=1}^{n} q_i \phi_i^{\,2}(\boldsymbol{x})$$

where q_i are the penalty parameters and $q_i >> 0$. The solution to this unconstrained minimization problem is denoted by $\boldsymbol{x}(q)$, where q denotes the respective parameters. Often $q_i = q$ for all i where q is constant. Thus the penalty function is denoted by $P(\boldsymbol{x}, q)$ and the corresponding minimum by $\boldsymbol{x}_o(q)$. It can be shown that under normal continuity conditions $\underset{q \to \infty}{Lim}\, \boldsymbol{x}_o(q) = \boldsymbol{x}_o$.

Typically, the overall penalty parameter q is the set at $q = 10^4$ if the constraints functions are normalized in some sense [1, 6, 7, 17].

In section 3, we present the QPF of the problem defined in (10) and give the main results using its optimality conditions. A numerical example is given in the forthcoming sections of the study and calculated with the use of the results obtained.

3. The best approximate solution via a quadratic penalty function

Many efficient methods have been developed for solving the quadratic programming problems [1, 4, 7, 14, 17, 24], 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, 6, 7, 17].

We assume that the quadratic programming problems (10), (11) and (12) have feasible solutions and their constraint regions are bounded.

The penalty function of the problem (10) can be defined as

$$P(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} + \frac{1}{2}q\|A\boldsymbol{x} - \boldsymbol{p}\|^2,$$

where the scalar quantity q is the penalty parameter. It is clear that quadratic programming problems defined in (10)-(13) have the same optimal solution because $A\boldsymbol{x} = \boldsymbol{p}$ and the equations defined in (8) are consistent and equivalent. Since $\|A\boldsymbol{x} - \boldsymbol{p}\| = \|\boldsymbol{b} - \boldsymbol{p}\| = \|\boldsymbol{b} - AA^+\boldsymbol{b}\| = \|\boldsymbol{b} - A\boldsymbol{x}\|$ for the optimal solution of the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$, we can formulate the QPF as

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} + \frac{1}{2}q\|A\boldsymbol{x} - \boldsymbol{b}\|^2 \tag{14}$$

to find the best approximate solution with the minimum norm for the inconsistent system of the linear equations.

From the first order necessary conditions for the unconstrained minimum of the QPF (14), we obtain

$$\nabla f(\boldsymbol{x}) = \boldsymbol{x} + qA^TA\boldsymbol{x} - qA^T\boldsymbol{b} = 0\,. \tag{15}$$

We also obtain the Hessian matrix of (14), which represents the sufficient condition as

$$H(\boldsymbol{x}) = I + qA^TA\,, \tag{16}$$

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

COROLLARY 3.1 *The Hessian matrix* $H(\boldsymbol{x}) = I + qA^TA$ *of the QPF defined in* (14) *is positive definite.*

Proof Let eigenvalues be $\lambda_1, \lambda_2, ..., \lambda_n$ 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 $H(\boldsymbol{x})$ 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 best approximate solution of the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$.

THEOREM 3.1 *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*

$$\boldsymbol{x} = \lim_{q\to\infty}\left(\frac{1}{q}I + A^T A\right)^{-1} A^T\boldsymbol{b}\,, \tag{17}$$

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

Proof From (15), we obtain

$$\left(\frac{1}{q}I + A^T A\right)\boldsymbol{x} = A^T\boldsymbol{b}\,. \tag{18}$$

Using $\det\left(\frac{1}{q}I + A^T A\right) \neq 0$ for $q \neq 0$ and applying (6), we get

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

This solution is the best approximate solution with minimum norm of the QPF (14) and its Hessian matrix H is positive definite. Using (2) and (7) we see that

$$\boldsymbol{x} = \left(A^T A\right)^{+} A^T\boldsymbol{b} = A^{+}\boldsymbol{b} \tag{19}$$

and

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

Then the proof is completed.

We now consider the characteristic equations defined in (20) and (21) of the matrices $A^T A$ and $\left(\frac{1}{q}I + A^T A\right)$, respectively

$$\det\left(A^T A - \lambda I\right) = (\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n) = 0 \tag{20}$$

and

$$\det\left(\frac{1}{q}I + A^T A - \mu I\right) = \left(\mu-\lambda_1-\frac{1}{q}\right)\left(\mu-\lambda_2-\frac{1}{q}\right)\ldots\left(\mu-\lambda_n-\frac{1}{q}\right) = 0\,, \tag{21}$$

where $\mu_i = \lambda_i + \frac{1}{q} > 0$ for $i = 1, 2, \ldots, n$.

Let $\boldsymbol{v}_i$ be eigenvectors corresponding to eigenvalues λ_i of the matrix $A^T A$ and let $\boldsymbol{u}_i$ be eigenvectors corresponding to eigenvalues $\lambda_i + \frac{1}{q}$ of the matrix $\frac{1}{q}I + A^T A$, where $\|\boldsymbol{v}_i\| = 1$ and $\boldsymbol{v}_i^T\boldsymbol{v}_j = 0$ for $i \neq j$. From

$$\left(\frac{1}{q}I + A^T A - \mu_i I\right)\boldsymbol{u}_i = \left[\frac{1}{q}I + A^T A - \left(\lambda_i + \frac{1}{q}\right)I\right]\boldsymbol{u}_i = \left(A^T A - \lambda_i I\right)\boldsymbol{u}_i = 0,$$

we see that

$$\boldsymbol{u}_i = \boldsymbol{v}_i \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Let the orthogonal decomposition of the matrix $A^T A$ be

$$A^T A = VDV^{-1}\,, \tag{22}$$

where $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$, $D = diag\,\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\frac{1}{q}I + D = diag\left\{\lambda_1 + \frac{1}{q}, \lambda_2 + \frac{1}{q}, \ldots, \lambda_n + \frac{1}{q}\right\}$.

Using (22), we also express the matrix $\frac{1}{q}I + A^T A$ as

$$\frac{1}{q}I + A^T A = V\left(\frac{1}{q}I + D\right)V^{-1}. \tag{23}$$

This leads to the following theorem.

THEOREM 3.2 *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 $A^T A$, respectively. The best approximate solution of the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ is*

(i) *If $rank\,(A) = n$, $\boldsymbol{x} = VD^{-1}V^{-1}A^T\boldsymbol{b} = (A^T A)^{-1} A^T \boldsymbol{b}$,*

(ii) *If $rank\,(A) < n$, $\boldsymbol{x} = VD^{+}V^{-1}A^T\boldsymbol{b} = (A^T A)^{+} A^T \boldsymbol{b}$,*

where D^{-1} is the inverse of the diagonal matrix $D = diag\,\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and D^{+} is the generalized inverse of $D = diag\,\{\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0\}$.

Proof Applying Theorem 3.1 and using (23), we get

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

If $rank\,(A) = n$,

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

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

If $r < n$,

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

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

Thus the proof is completed.

Here we establish the following corollary which is proved easily using the optimality conditions of the QPF defined in (14) and results on the generalized inverses of matrices.

COROLLARY 3.2 *Let* λ_i *be eigenvalues of the matrix* $A^T A$ *and let* $\mu_i = \lambda_i + \frac{1}{q}$ *be eigenvalues of the matrix* $\frac{1}{q} I + A^T A$. *A necessary and sufficient condition for the QPF defined in* (14) *to have a best approximate solution with minimum norm is* $\mu_i = \lambda_i + \frac{1}{q} > 0$ *, in which case the optimal solution is* $\boldsymbol{x} = A^{+}\boldsymbol{b}$.

Proof Since the Hessian matrix given in Corollary 3.1 is a positive definite matrix and eigenvalues of the matrix $H(\boldsymbol{x})$ are $1 + q\lambda_i > 0$, $\mu_i = \lambda_i + \frac{1}{q} > 0$ and the QPF defined in (14) has a best approximate solution with minimum norm. Using (22) and (23), we find the solution $\boldsymbol{x} = A^{+}\boldsymbol{b}$. This completes the proof.

4. Numerical Example

We can also calculate the approximate solution with the minimum norm of the inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$ using

$$\boldsymbol{x}_{k+1} = \left(\frac{1}{q_k} I + A^T A \right)^{-1} A^T \boldsymbol{b} \tag{24}$$

or

$$\boldsymbol{x}_{k+1} = V \left(\frac{1}{q_k} I + D \right)^{-1} V^{-1} A^T \boldsymbol{b}, \tag{25}$$

where

$$q_k = 10^k \text{ for } k = 0, 1, 2, \ldots .$$

Let $A\boldsymbol{x} = \boldsymbol{b}$ be an inconsistent system computed by Rosen in [23], where A of rank 3 and $\boldsymbol{b}$ are

$$A = \begin{bmatrix} 1 & -1 & 3 & 1 \\ 2 & 4 & 5 & 1 \\ -1 & 2 & -1 & 1 \\ 4 & 1 & 9 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1.0 \\ 3.0 \\ 2.5 \\ 2.5 \end{bmatrix}.$$

Table 1 and 2 represent the best approximate solution for optimal solution of the inconsistent linear equation $A\boldsymbol{x} = \boldsymbol{b}$. The following solution, which is obtained using (19), is the exact solution of $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$. This solution is the optimal solution with minimum norm of $A\boldsymbol{x} = \boldsymbol{b}$:

$$\boldsymbol{x} = (A^T A)^{+} A^T \boldsymbol{b} = \begin{bmatrix} -0.2115384 & 0.0448717 & -0.2243589 & 0.0576923 \\ -0.1923076 & 0.1923076 & 0.0384615 & -0.0384615 \\ 0.0865384 & -0.0032051 & 0.0160256 & 0.0673076 \\ 0.5096153 & -0.0929487 & 0.4647435 & -0.0480769 \end{bmatrix} \begin{bmatrix} 1.0 \\ 3.0 \\ 2.5 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.4935897 \\ 0.3846153 \\ 0.2852564 \\ 1.2724358 \end{bmatrix}.$$

In addition, the generalized inverse of the matrix A is computed approximately by taking $q_k = 10^6$ in (24) as follows:

$$\left(\frac{1}{10^6} I + A^T A \right)^{-1} A^T = \begin{bmatrix} -0.2115383 & 0.0448717 & -0.2243588 & 0.0576922 \\ -0.1923076 & 0.1923076 & 0.0384615 & -0.0384615 \\ 0.0865384 & -0.0032051 & 0.0160256 & 0.0673076 \\ 0.5096150 & -0.0929486 & 0.4647433 & -0.0480768 \end{bmatrix}.$$

Using this matrix, we can compute the approximate solution correct to six decimal places of the given inconsistent system as

$$\boldsymbol{x}^T = \left[-0.4935894 \quad 0.3846155 \quad 0.2852563 \quad 1.2724351 \right].$$

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

In Table 1, some values for the approximate solution using (24) are listed with seven decimals for $k = 0,1,2, \ldots, 7$.

k	x_1	x_2	x_3	x_4
0	-0.2942274	0.4464388	0.2449741	0.8334291
1	-0.4635347	0.3954889	0.2789696	1.2060391
2	-0.4904223	0.3857787	0.2845914	1.2654362
3	-0.4932712	0.3847325	0.2851895	1.2717321
4	-0.4935578	0.3846271	0.2852497	1.2723654
5	-0.4935865	0.3846165	0.2852557	1.2724288
6	-0.4935894	0.3846155	0.2852563	1.2724351
7	-0.4935897	0.3846153	0.2852564	1.2724358
8	-0.4935898	0.3846153	0.2852562	1.2724359

Table 1: The best approximate solution correct to desired decimal places of the problem

In Table 2, some values for the approximate solution using (25) are listed with seven decimals for $k = 0, 1, 2, \ldots, 6$.

k	x_1	x_2	x_3	x_4
0	-0.2942274	0.4464388	0.2449741	0.8334291
1	-0.4635347	0.3954889	0.2789696	1.2060390
2	-0.4904223	0.3857787	0.2845914	1.2654362
3	-0.4932712	0.3847325	0.2851895	1.2717321
4	-0.4935578	0.3846271	0.2852497	1.2723654
5	-0.4935865	0.3846165	0.2852557	1.2724288
6	-0.4935894	0.3846155	0.2852563	1.2724351
7	-0.4935901	0.3846153	0.2852566	1.2724356

Table 2: The best approximate solution of the example

From Table 1 and Table 2, we see that the approximate solution correct to six decimal places of the given system is computed directly for $q_k = 10^7$. Should the penalty parameter q is taken big enough, the approximate solution correct to desired decimal places of the system can be calculated using (24) and (25).

5. Conclusions

In this study, the inconsistent linear system of m equations in n unknowns is formulated as a quadratic programming problem, and the best approximate solution with the minimum norm for the inconsistent system of the linear equations is investigated using the optimality conditions of the QPF. 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 analytic results are compared with numerical examples. Numerical results in this paper show that the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ can be solved using (24) or (25) when $rank(A) < n$. It is shown that, with minimum norm, the approximate solution correct to desired decimal places of the inconsistent system can be computed by the penalty method.

References

[1] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, Nonlinear Programming: Theory and Algorithms, Second Edition, John Wiley and Sons, New York, 1993.

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

[3] J. A. Cadzow, Minimum l_1, l_2 and l_∞ norm approximate solutions to an overdetermined system of linear equations, Digital Signal Proc., 12(2002) 524-560.

[4] Z.-H. Cao, On the converge of iterative methods for solving singular linear systems, J. Comput. Appl. Math., 145(2002) 1-9.

[5] A. K. Cline, An elimination method for the solution of linear least squares problems, SIAM J. Numer. Anal., 10(1973) 283-289.

[6] G. Di Pillo, G. Liuzzi and S. Lucidi, An exact penalty-Lagrangian approach for large-scale nonlinear programming, Optimization, Vol.60, No1-2(2011) 223-252.

[7] Z. Dostal, On penalty approximation of quadratic programming problem, Kybernetika, 27(1991) 151-154.

[8] T. Elfving, Block-iterative methods for consistent and inconsistent linear equations, Numer. Math., 35(1980) 1-12.

[9] G. H. Golub, C. Per Hansen and D. O'leary, Tikhonov regularization and total least squares, SIAM J. on Matr. Anal. and Appl., 21(1999) 1-10.

[10] G. H. Golub and C. F. V. Loan, Matrix computations, Third Edition, J. Hopkins Univ. Press, London, 1996.

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

[12] S.-J. Kim, K. Koh, M. Lustig, S. Boyd and D. Gorinevsky, An interior- point method for large-scale l_1- regularized least squares, IEEE Journal on Selected Topics in Signal Processing, 1(4)(2007) 606-617.

[13] C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, Prentice Hall, Inc. Englewood Cliffs, N.J., 1974.

[14] B. Lewis and R. Lothar, Arnoldi-Tikhonov regularization methods, J. Comput. Appl. Math., 226(2009) 92-102.

[15] J. Miao and A. Ben-Israel, On l_p – approximate solutions of linear equations, Lin. Algeb. and its Appl., 199(1994) 305-327.

[16] A. Neumair, Solving ill-conditioned and singular linear systems: A tutorial on regularization, SIAM Rev., 40(1998) 636-666.

[17] N. Özdemir and F. Evirgen, A dynamic system approach to quadratic programming problems with penalty method, Bull. Malays. Math. Sci. Soc., (2) 33(1) (2010) 79-91.

[18] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc., 51(1955) 406-413.

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

[20] M.Ç. Pınar, Linear programming via a quadratic penalty function, Math. Method of Ope. Res., 44(1996) 345-370.

[21] M.Ç. Pınar and S. Elhedhli, A penalty continuation method for the l_∞ solution of the overdetermined linear systems, BIT, 38(1)(1998) 127-150.

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

[23] J. B. Rosen, Minimum and basic solutions to singular linear systems, J. Soc. Indust. Appl. Math., 12(1964) 156-162.

[24] E. Übi, A numerically stable least squares solution to the quadratic programming problem, Cent. Eur. J. Math., 6(2008) 171-178.