Nonlinear Dynamics and Systems Theory, **19** (1-SI) (2019) 193–199

A Recursive Solution Approach to Linear Systems with Non-Zero Minors

A. Moradi [1*], A. Kameli [2], H. Jafari [2,3] and A. Valinejad [1]

[1] *Department of Computer Science, University of Mazandaran, Babolsar, Iran*
[2] *Department of Mathematics, University of Mazandaran, Babolsar, Iran*
[3] *Department of Mathematical Sciences, University of South Afirca, UNISA0003, South Africa*

Received: June 3, 2018;   Revised: February 4, 2019

Abstract: In this paper, we introduce a recursive solution approach to linear systems of the form $Ax = b$, where A is non-singular and its corner minors are all non-zero. For the first time in the literature, we show how one can exploit (possible) useful information provided by corner sub-matrices of A towards an efficient solution approach to the linear system. This is going to initiate a thorough study of solution methods whose goals are to fully exploit available information within the given linear system.

Keywords: *linear system of equations; corner minors; matrix inversion; recursive methods.*

Mathematics Subject Classification (2010): 15A06, 15A09.

1 Introduction

The problem of solving a linear system $Ax = b$ is central to scientific computation [1], a subject which is used in most parts of modern mathematics. Computational solution methods of such system are often an important part of numerical linear algebra (see [2,3]), and play an important role in engineering, physics, chemistry, computer science, and economics [4]. Even more, systems of non-linear equations are often approximated by linear ones with the aim of linearization, a helpful technique while making a mathematical model or computer simulation of a relatively complex system. A reader interested in the applications of linear systems is referred to [4–7].

Iterative vs. direct solution methods for solving general linear systems have been gaining popularity in many areas of scientific computing [8, 9]. Until recently, direct

* Corresponding author: mailto:a.moradi@umz.ac.ir

© 2019 InforMath Publishing Group/1562-8353 (print)/1813-7385 (online)/http://e-ndst.kiev.ua

solution methods were often preferred to iterative methods in real applications because of their robustness and predictable behavior [9]. However, to the best of our knowledge, none of the existing methods is capable of exploiting special information provided by the underlying linear system. This information could appear in an application setting within which a linear system with known solution is going to be expanded to a larger linear system. Other than that, simple matrix operations often reveal sub-matrices of A whose inverse are quickly computable. This paper initiates the study of linear systems when such information is available. We limit our attention to a special class of non-singular matrices and build necessary algebraic tools to study linear systems with such coefficient matrices.

The rest of the paper is organized as follows. In Section 2, we define and elaborate on the necessary notations and definitions needed in the paper. In Section 3, we build algebraic tools to derive matrix inverse while fully exploiting available information of inverse of a sub-matrix. We elaborate on the method by algorithmic restatement and also by giving an example. In Section 4, we explain how the result obtained in Section 3 can naturally result in a solution method to linear systems. Finally, in Section 5 we draw some conclusions and outline some possible avenues for further improvement.

2 Terminology

We consider a matrix $A = (a_{i,j})_{n\times m}$ of n rows and m columns. For any $1 \le i \le n$ and any $1 \le j \le m$, the i-th row and the j-th column of A are denoted by A^i and A_j, respectively. The index sequence of rows and columns of A are the sequence $\langle 1, 2, \cdots, n\rangle$ and $\langle 1, 2, \cdots, m\rangle$, respectively. Let us refer to A's index sequence of *rows* as A's *r-sequence* and A's index sequence of *columns* as A's *c-sequence*. Having a sub-sequence $\langle r_1, r_2, \cdots, r_p\rangle$ of the A's r-sequence and a sub-sequence $\langle c_1, c_2, \cdots, c_q\rangle$ of A's c-sequence, one can define a sub-matrix $S = (s_{i,j})_{p\times q}$ of A as $s_{i,j} = a_{r_i,c_j}$. Conversely, for any sub-matrix S of A, S's r-sequence and c-sequence are proper sub-sequences of A's r-sequence and A's c-sequence, respectively. In this setting, crossing off the i-th index in A's r-sequence defines a sub-matrix of A denoted by $\mathsf{del}^i(A)$. Similarly, crossing off the j-th index in A's c-sequence defines a sub-matrix denoted by $\mathsf{del}_j(A)$. If the deletion operations happen simultaneously, we get the sub-matrix $\mathsf{del}^i_j(A)$. We also need to define a matrix obtained by adding a new row and simultaneously a new column to A. Given indexes $1 \le i \le n+1$ and $1 \le j \le m+1$, and vectors $F_{1\times(n+1)}, G_{(m+1)\times 1}$ with $f_{1,j} = g_{i,1}$, the unique matrix B, defined by

$$B^i = F, \quad B_j = G, \quad \mathsf{del}^i_j(B) = A,$$

is denoted by $\mathsf{add}^i_j(A, F, G)$. The operators del and add will be extensively used in the following.

3 Computing A^{-1}

Given a non-singular $n \times n$ matrix A, suppose that there exists a square sub-matrix of A, say S, whose inverse is known (or quickly computable). The core question in this work asks: how can A^{-1} be computed using the available information on (the inverse of) the sub-matrix S? In this paper, we build our results on a special class of non-singular matrices for which every corner minor is non-zero.

Let us limit our attention and assume every corner minor of A is non-zero. Let $S = \mathsf{del}_n^n(A)$ and suppose that its inverse, S^{-1}, is known. Note that by the assumption on A, S^{-1} does exist. Define

$$B = \mathsf{add}_n^n(S, I^n, I_n) = \left(\begin{smallmatrix} S & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

whose inverse is simply

$$B^{-1} = \mathsf{add}_n^n(S^{-1}, I^n, I_n) = \left(\begin{smallmatrix} S^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

and consider the $n \times n$ square matrix C given by the equation $A = B.C$. Then C is simply given by

$$C = \left(\begin{smallmatrix} I_{(n-1)\times(n-1)} & V \\ a_{n,1}\cdots a_{n,n-1} & a_{n,n} \end{smallmatrix}\right) \text{ where } V = S^{-1}.(a_{1,n}, \cdots, a_{n-1,n})^T \tag{1}$$

and I is the identity matrix. Matrix C has the property that its inverse can be easily computed by means of the following lemma.

Lemma 3.1 Let $p = A^n.\left(\begin{smallmatrix} V \\ -1 \end{smallmatrix}\right)$, then p is non-zero and the $i-$th row of C^{-1} is given by

$$(C^{-1})^i = \begin{cases} \frac{1}{p}(A^n - (1 + a_{nn})I^n), & i = n \\ -v_i(C^{-1})^n + I^i, & i \neq n. \end{cases} \tag{2}$$

Proof. Knowing $C^{-1}C = I$, let us expand the equations obtained by $(C^{-1})^n C = I^n$:

$$\begin{aligned}
&c^{-1}_{n,1} + c^{-1}_{nn}a_{n,1} = 0,\\
&c^{-1}_{n,2} + c^{-1}_{nn}a_{n,2} = 0,\\
&\vdots\\
&c^{-1}_{n,n-1} + c^{-1}_{n,n}a_{n,n-1} = 0,\\
&(c^{-1}_{n,1}, c^{-1}_{n,2}, \cdots, c^{-1}_{n,n-1}) \cdot V + c^{-1}_{n,n}a_{n,n} = 1.
\end{aligned} \tag{3}$$

Now, the j-th equation gives $c^{-1}_{n,j} = -c^{-1}_{nn}a_{n,j}$ for each $j = 1, \cdots, n-1$. Then we write the last equation as

$$1 - c^{-1}_{n,n}a_{n,n} = -c^{-1}_{n,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,n-1}).V$$

and we get

$$\begin{aligned}
&c^{-1}_{n,n}((a_{n,1}a_{n,2}\cdots a_{n,n-1}) \cdot V - a_{n,n}) = -1,\\
&c^{-1}_{n,n}\left((a_{n,1}, a_{n,2}, \cdots a_{n,n-1}, a_{n,n})\begin{pmatrix} V \\ -1 \end{pmatrix}\right) = -1,\\
&c^{-1}_{n,n}\left(A^n\begin{pmatrix} V \\ -1 \end{pmatrix}\right) = -1.
\end{aligned} \tag{4}$$

This, in turn, implies that $c^{-1}_{n,n} = -\frac{1}{p}$, where $p = A^n\binom{V}{-1} \neq 0$. As a result

$$\begin{aligned}
(C^{-1})^n &= (c^{-1}_{n,1}, c^{-1}_{n,2}, \cdots, c^{-1}_{n,n-1}, c^{-1}_{n,n})\\
&= (-c^{-1}_{n,n}a_{n,1}, -c^{-1}_{n,n}a_{n,2}, \cdots - c_{n,n-1}a_{n,n-1}, c^{-1}_{n,n})\\
&= -c^{-1}_{n,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,n-1}, -1)\\
&= -c^{-1}_{n,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,n-1}, -1 - a_{nn} + a_{nn})\\
&= \frac{1}{p}\left(A^n - (1 + a_{n,n}I^n)\right).
\end{aligned} \tag{5}$$

Now, in order to compute other rows of C^{-1}, let us expand the equations obtained by $(C^{-1})^iC = I^i, i \neq n$, as

$$\begin{aligned}
&c^{-1}_{i,1} + c^{-1}_{i,n}a_{n,1} = 0,\\
&c^{-1}_{i,2} + c^{-1}_{i,n}a_{n,2} = 0,\\
&\vdots\\
&c^{-1}_{i,i} + c^{-1}_{i,n}a_{n,i} = 1,\\
&\vdots\\
&c^{-1}_{i,n-1} + c^{-1}_{i,n}a_{n,n-1} = 0,\\
&(c^{-1}_{i,1}c^{-1}_{i,2}\cdots c^{-1}_{i,i}\cdots c^{-1}_{i,n-1}) \cdot V + c^{-1}_{i,n}a_{n,n} = 0.
\end{aligned} \tag{6}$$

One can write the first $n-1$ equations as

$$(c^{-1}_{i,1}, c^{-1}_{i,2}, \cdots, c^{-1}_{i,i}, \cdots, c^{-1}_{i,n-1}) = I^i - c^{-1}_{i,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,i}, \cdots, a_{n,n-1}).$$

Now, using the last equation in (6), we get

$$\begin{aligned}
&-c^{-1}_{i,n}a_{n,n} = I^i \cdot V - c^{-1}_{i,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,i}, \cdots, a_{n,n-1}) \cdot V,\\
&c^{-1}_{i,n}\left(A^n\begin{pmatrix} V\\ -1\end{pmatrix}\right) = v_i.
\end{aligned}$$

This, in turn, implies that $c^{-1}_{i,n} = \frac{1}{p}.v_i$. As a result

$$\begin{aligned}
(C^{-1})^i &= (c^{-1}_{i,1}, c^{-1}_{i,2}, \cdots, c^{-1}_{i,i}, \cdots c^{-1}_{i,n-1}, c^{-1}_{i,n})\\
&= -c^{-1}_{i,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,i}, \cdots, a_{n,n-1}, -1) + I^i\\
&= -c^{-1}_{i,n}(a_{n,1}, a_{n,2}, \cdots, a_{n,i}, \cdots, a_{n,n-1}, -1 + a_{n,n} - a_{n,n}) + I^i\\
&= -c^{-1}_{i,n}(A^n - (1 + a_{n,n})I^n) + I^i\\
&= -\frac{1}{p} \cdot v_i(A^n - (1 + a_{n,n})I^n) + I^i\\
&= -v_i.(C^{-1})^n + I^i.
\end{aligned}$$

This completes the proof. □

Having computed B^{-1} and C^{-1}, the inverse of A can be computed as $A^{-1} = C^{-1}B^{-1}$. Note how S^{-1} is used in computing A^{-1}. The equation also suggests a recursive procedure to obtain A^{-1} via its corner sub-matrices as described in Algorithm 3.1.

Algorithm 3.1 Computing A^{-1}

procedure $\textsc{Inverse}(A, n)$

$S \leftarrow \mathsf{del}_n^n(A)$

$S^{-1} \leftarrow \textsc{Inverse}(S, n-1)$

$V \leftarrow S^{-1}.(a_{1,n}, \cdots, a_{n-1,n})^T$

$B^{-1} \leftarrow \left(\begin{smallmatrix} S^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right)$

$p = A^n \left(\begin{smallmatrix} V \\ -1 \end{smallmatrix}\right)$

$(C^{-1})^i \leftarrow \begin{cases} \frac{1}{p}(A^n - (1 + a_{n,n}).I^n) & \text{if } i = n, \\ -v_i.(C^{-1})^n + I^i & \text{if } i \neq n. \end{cases}, \qquad \forall \quad i = 1, 2, \cdots, n$

return $C^{-1}B^{-1}$

Example 3.1 Let

$$A = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 0 & 3 \end{array}\right)$$

and set $S = \mathsf{del}_4^4(A)$, which is simply $I_{3\times 3}$. Then

$$\begin{aligned}
&V = I_{3\times3}.(2, -1, 1)^T = (2, -1, 1)^T, \\
&p = (1, 1, 0, 3).(2, -1, 1, -1)^T = -2, \\
&(C^{-1})^4 = -\frac{1}{2}\{(1, 1, 0, 3) - 4(0, 0, 0, 1)\} = (-0.5, 0.5, 0, 0.5), \\
&(C^{-1})^1 = -2.(\frac{-1}{2}, \frac{-1}{2}, 0, \frac{1}{2}) + (1, 0, 0, 0) = (2, 1, 0, -1), \\
&(C^{-1})^2 = +1.(\frac{-1}{2}, \frac{-1}{2}, 0, \frac{1}{2}) + (0, 1, 0, 0) = (\frac{-1}{2}, \frac{1}{2}, 0, \frac{1}{2}), \\
&(C^{-1})^3 = -1.(\frac{-1}{2}, \frac{-1}{2}, 0, \frac{1}{2}) + (0, 0, 1, 0) = (\frac{1}{2}, \frac{1}{2}, 1, \frac{-1}{2}).
\end{aligned}$$

Putting all together

$$C^{-1} = \begin{pmatrix} 2 & 1 & 0 & -1 \\ \frac{-1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{-1}{2} \\ \frac{-1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

we have $A^{-1} = C^{-1}$ as computed above. $\square$

4 Solving Linear System of Equations

Having a procedure to compute A^{-1}, as introduced above, automatically results in a solution procedure of the linear system $A.x = b$, where $x = (x_1, \cdots, x_n)^T$ and $b = (b_1, \cdots, b_n)^T$. Algorithm 3.1 will immediately translate to a recursive solution procedure of the linear system as follows.

Here again, we try to find a connection between the solution of the linear system and the solution of the subsystem $\mathsf{del}_n^n(A).y = \mathsf{del}^n(b)$. Recall that $S = \mathsf{del}_n^n(A)$.

Solution x of the linear system $Ax = b$ simply satisfies

$$S.\mathsf{del}^n(x) = \mathsf{del}^n(b) - x_n.\mathsf{del}^n(A_n), \tag{7}$$

$$\mathsf{del}_n(A^n).\mathsf{del}^n(x) = b_n - a_{n,n}.x_n. \tag{8}$$

Having S^{-1} available, one can rewrite (7) as

$$\mathsf{del}^n(x) = S^{-1}.\mathsf{del}^n(b) - x_n.S^{-1}.\mathsf{del}^n(A_n). \tag{9}$$

Note that the term $S^{-1}.\mathsf{del}^n(b)$ is the solution to the subsystem $S.y = \mathsf{del}^n(b)$. Then the solution to the system $Ax = b$ can be easily computed using equations (8) and (9). In this way, the solution process of the system $Ax = b$ can carefully make use of the information (possibly) available through the subsystem $S.y = \mathsf{del}^n(b)$.

Example 4.1 Let A be the matrix given in Example 1 and $b = (1, -2, 1, 4)^T$. In order to solve the system $Ax = b$, set $S = \mathsf{del}_4^4(A)$ which is simply $I_{3\times 3}$. Computing $\mathsf{del}^4(x)$ by equation (9) and putting it in equation (8) give

$$(1, 1, 0)\left(\left(\begin{smallmatrix} 1 \\ -2 \\ 1 \end{smallmatrix}\right) - x_4\left(\begin{smallmatrix} 2 \\ -1 \\ 1 \end{smallmatrix}\right)\right) = 4 - 3x_4,$$

then $x_4 = 2.5$ and equation (9) computes

$$\mathsf{del}^4(x) = \left(\begin{smallmatrix} 1 \\ -2 \\ 1 \end{smallmatrix}\right) - x_4\left(\begin{smallmatrix} 2 \\ -1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -4 \\ 0.5 \\ -1.5 \end{smallmatrix}\right).$$

So, we have $x = (\mathsf{del}^4(x), x_4)^T = (-4, 0.5, -1.5, 2.5)^T$. $\square$

Note that the way we solved the above linear system has an important capability with which different solution procedures of a linear system can be combined.

5 Conclusion

In this paper, we studied linear systems of the form $Ax = b$. When A admits non-zero corner minors, we showed a solution method could be devised capable of using available information provided by the corner submatrices of A. This, in turn, asks for a more detailed study of solution methods whose goals are to fully exploit available information within the given linear system having a general coefficient matrix.

Acknowledgment

We would like to thank the referees for their careful review of our manuscript and some helpful suggestions.

References

[1] G. H. Golub, and C. F. Van Loan. *Matrix Computations.* Vol. 3. JHU Press, 2012.

[2] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra.* Vol. 71. Siam, 2000.

[3] D. Poole. *Linear Algebra: A Modern Introduction.* Cengage Learning, 2014.

[4] L. Hogben. *Handbook of Linear Algebra.* Chapman and Hall/CRC, 2006.

[5] G. Chatzarakis, J. Diblk, and I. Stavroulakis. Explicit integral criteria for the existence of positive solutions of first order linear delay equations. *Electronic Journal of Qualitative Theory of Differential Equations* **45** (2016) 1–23.

[6] A. A. Martynyuk, A. G. Mazko, S. N. Rasshyvalova, and K. L. Teo. On the Past Ten Years and the Future Development of Nonlinear Dynamics and Systems Theory (ND&ST). *Nonlinear Dynamics and Systems Theory* **11**(4) (2011) 337–340.

[7] A. A. Martynuk. Stability in the models of real world phenomena. *Nonlinear Dynamics and Systems Theory* **11**(1) (2011) 7–52.

[8] T. A. Davis. *Direct Methods for Sparse Linear Systems*. Vol. 2. Siam, 2006.

[9] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Vol. 82. Siam, 2003.