Integral solution of linear equations using integer arithmetic

S. K. SEN AND A. A. SHAMIM
Computer Centre, Indian Institute of Science, Bangalore 560 012, India

Received on October 13, 1977 ; Revised on February 2, 1978

Abstract

The paper describes the use of integer arithmetic on a method of transforming a matrix to a Smith Normal Form and hence computing a generalized inverse that gives all the integral solutions to linear equations.

Key words: Elementary row/column operations, Generalized matrix inverse, Integer arithmet , Integral matrix, Residue arithmetic, Smith Normal Form, Unimodular matrix.

1. Introduction

Hurt and Waid[4] propose a generalized inverse which gives all the integral solutions to linear equations. An exact computational approach for computing such a generalized inverse based on modular arithmetic is suggested by Adegbeyeni and Krishnamurthy[1] for integral solutions of linear equations. This is costly from computing power and programming points of view in a general purpose computing system. Also, choice of n primes and combining the resulting n outputs (in the last stage) using the Chinese Remainder Theorem are added problems.

We present here the method due to Marcus and Minc,[3] and Hurt and Waid,[4] mention the main results and then describe the use of integer arithmetic to obtain any integral solution (exact) economically. Illustrative numerical examples are given.

2. Definitions

(i) *Integral vector and integral matrix*

Let

(a) $K =$ the ring of integers 0, $\pm$ 1, $\pm$ 2,...,

(b) $K^m =$ the m dimensional vector space over K,

(c) $K^{m \times n} =$ the $m \times n$ matrices over K, and

(d) $K_r^{m \times n} =$ the $m \times n$ matrices with rank r over K.

Any element of K^m is an integral vector. Any element of $K^{m \times n}$ is an integral matrix and any element of $K_r^{m \times n}$ is an integral matrix of rank r.

(ii) *Unimodular matrix*

Any nonsingular matrix $P \in K^{m \times m}$ whose inverse P^{-1} is also in $K^{m \times m}$ is a unimodular matrix.

(iii) *Elementary row and column operations*

A sequence of elementary row and column operations used here consists of

(*a*) *Type* 1 : interchanging two rows (columns), and

(*b*) Type 2 : subtracting an integral multiple of one row (column) from another row (column).

(iv) *Equivalent matrices*

Two matrices A, $S \in K^{m \times n}$ are equivalent over K if there exist two unimodular matrices $P \in K^{m \times m}$ and $Q \in K^{n \times n}$ such that

$$PAQ = S.$$

(v) *Smith Normal Form*

A matrix $S = (s_{ij}) \in K_r^{m \times n}$ is the Smith Normal Form (SNF) if

(*a*) $s_{ii} \neq 0,\ i = 1\,(1)\,r$

(*b*) $s_{ij} = 0$ otherwise, and

(*c*) s_{ii} divides $s_{i+1,\ i+1}$, $\quad i = 1\,(1)\,r - 1.$

(vi) *Generalized Inverse of Smith Normal Form*

The generalized inverse (g-inverse) of SNF S is the matrix $S^{+} = (s_{ij}^{+}) \in K_r^{n \times m}$ if

(*a*) $s_{ii}^{+} = s_{ii}^{-1}, \quad i = 1\,(1)\,r$ and

(*b*) $s_{ij}^{+} = 0$ otherwise.

(vii) *Integer arithmetic*

Let $+$, $-$, $.$, $/$ be integer add, subtract, multiply and divide operations respectively. Let a, $b \in K$.

Then

Add :	$a + b = c \in K$
Subtract :	$a - b = d \in K$
Multiply :	$a\,.\,b = e \in K$
Divide :	$a/b = f \in K\,(b \neq 0)$

provided

- $(c+d)/2=a$ and $(c-d)/2=b$ (for add and subtract operations)
- $e/a=b\,(a\neq 0)$ and $e/b=a\,(b\neq 0)$ and $e=0$ whenever $a=0$ or $b=0$ or both are zero ; e is negative when a or b negative (for multiply operation).
- $|a|\geqslant|b|\,.\,|f|$;

  For $|a|>|b|$, f is negative when a or b is negative, otherwise f is positive ;

  For $0\leqslant|a|<|b|$, $f=0$;

  $r=|a|-|b|\,.\,|f|\in K$;

  $0\leqslant r<|b|$ (for divide operation).

Note : If $r=0$ then b divides a.

Examples : $a=-5\in K,\ b=2\in K$

$$a+b=-5+2=-3=c\in K$$
$$a-b=-5-2=-7=d\in K$$
$$a\,.\,b=-5\,.\,2=-10=e\in K$$
$$a/b=-5/2=-2=f\in K$$

since

$(-3+(-7))/2=-5=a,\ (-3-(-7))/2=2=b.$

$-10/(-5)=2=b,\ -10/2=-5=a,\ |a|=5>2.2=4,$

$f=2$ is negative as a is negative.

$r=5-2.2=1\in K,\ 0\leqslant r<2.$

3. The method

Let $A\in K_r^{m\times n}$.

Step 1 : *Computation of SNF* $S=(s_{ij})\in K_r^{m\times n}$:

(i) Find the greatest common divisor (GCD) of the elements of A.

(ii) Bring it to the position (1, 1) by using Type 1 and/or Type 2 operations.

(iii) Make zeros of all other elements in the first column and first row using Type 2 operations.

Note : The matrix $C=(c_{ij})$ so obtained is that

(*a*) $C\in K_r^{m\times n}$ is equivalent over K to A, (*b*) c_{11} divides c_{ij} $(i > 1, j > 1)$, and

(*c*) $c_{i1} = c_{1j} = 0$ $(i > 1, j > 1)$.

(iv) Setting $s_{11} = c_{11}$ repeat the algorithm for the $(m-1) \times (n-1)$ matrix (c_{ij}) $(i > 1, j > 1)$.

(v) Repeat the algorithm thus r times and stop when the bottom right $(m-r) \times (n-r)$ submatrix is zero giving the SNF.

Step 2 : *Computation of* A^- :

(i) Compute the unimodular matrix $P\,(Q)$ defined in Sec. 2 (iv), which is the product of all the elementary row (column) matrices, in the right order. Thus $PAQ = S$.

(ii) Compute $A^- = QS^+P$.

Note : A^- satisfies $AA^-A = A$, $A^-AA^- = A^-$, $AA^- \in K^{m \times m}$, $A^-A \in K^{n \times n}$.

Step 3 : *Computing a solution vector*

(i) If $AA^-b \neq b$ then $Ax = b$ is inconsistent, *i.e.*, it has no solution.

(ii) If $AA^-b = b$ but $x = A^-b$ is not integral then $Ax = b$ has no integral solution (it has nonintegral rational solutions though).

(iii) If $AA^-b = b$ and $x = A^-b$ is integral then compute any integral solution.

$$x = A^-b + y - A^-Ay$$

by assigning a value to $y \in K^{n \times n}$ which is arbitrary.

4. Main results

The method follows from the theorem and corollaries below. The theorem states that any integral matrix is equivalent over K to a diagonal integral matrix.

Theorem

Let $A \in K_r^{m \times n}$. Then A is equivalent over K to SNF $S \in K_r^{m \times n}$. For a constructive proof see Marcus and Minc.[3]

Corollary 1 : Let P and Q be unimodular matrices and $PAQ = S$ be the SNF of $A \in K^{m \times n}$. Also, let $A^- = QS^+P$. Then $AA^-A = A$, $A^-AA^- = A^-$, $A^-A \in K^{n \times n}$ $AA^- \in K^{m \times m}$.

Proof : $PAQ = S = SS^+S = PAQS^+PAQ = PAA^-AQ$. Hence $A = AA^-A$. A^-AA^- is proved similarly. Since $A^-AQ = QS^+PAQ\ QS^+S$ and $PAA^- = PAQS^+P = SS^+P$, the integrality of A^-A and AA^- follows.

Corollary 2: (Hurt and Waid[4]) Let $A \in K^{m \times n}$, $b \in K^m$.

Let $Ax = b$ be consistent. Then $Ax = b$ has an integral solution if and only if $A^- b$, is integral, in which case the general integral solution of $Ax = b$ is

$$x = A^- b + y - A^- Ay$$

where $y \in K^n$ is arbitrary.

5. Use of integer arithmetic

Let $Ax = b$ where $A \in K_r^{m \times n}$, $b \in K^m$. The foregoing method (Sec. 3) involves (i) finding the GCD, (ii) transforming A to SNF S, and computing P and Q where $PAQ = S$, (iii) obtaining $A^- = QS^- P$, (iv) checking if $AA^- b = b$, and (v) computing $x = A^- b + y - A^- Ay, y \in K^n$.

(i) To find the GCD of the elements of A

(*a*) obtain the smallest element in modulus,

(*b*) if it does not divide any one element of A then compute the remainder r_1; divide the smallest element by r_1. If it divides then r_1 is the GCD, otherwise divide r_1 by the remainder r_2 and repeat the process.

(*c*) If the smallest magnitude element divides all the elements of A then the modulus of it is the GCD.

(ii) To transform A to SNF and compute P and Q we evidently need only integer arithmetic.

iii) To obtain $A^- = QS^{+} P$,

(*a*) Compute $\alpha = \overset{r}{\underset{t=1}{\pi}} s_{tt}$,

(*b*) compute (i, j)th element of $\alpha A^- =$

$$\sum_{k=1}^{r} q_{ik} p_{kj} \overset{r}{\underset{\substack{t=1\\t \neq k}}{\pi}} s_{tt}, \; i = 1\,(1)\, r, \; j = 1\,(1)\, r$$

$A^- \in K_r^{n \times m}$ has zero for the other elements.

(iv) Compute $A\alpha A^- b$. If it is equal to αb then solution exists. Otherwise the system has no solution (either non-integral rational or integral).

(v) If $A\alpha A^- b = \alpha b$ then compute $\alpha A^- b$. If α divides all the elements of $\alpha A^- b$, then integral solution exists. Then obtain

$$x = (\alpha A^- b/\alpha) + y - A^- Ay \text{ for any } y \in K^n.$$

6. Present algorithm *versus* finite-field algorithms

Finite-field computational techniques using residue arithmetic are almost always advocated for exact computation for the following reasons:

(i) the parallelism in computation, and

(ii) the disadvantage of integer arithmetic which demands very long precision operands and hence makes computation slow.

As to (i), if we use n primes then the problem is solved n times independently. A multiprocessing system or many independent (monoprocessing) systems are needed to use the parallelism. The processing power, however, is actually not saved.

As to (ii), the present algorithm finds the GCD of the elements of A to compute the SNF S. In many problems, it avoids very long precision operands and thus is not that slow. Precisely, the integer arithmetic is inherent to this algorithm. The modular arithmetic need to be used only when the integer arithmetic fails.

The efficiency of the present algorithm compared to the algorithms based on residue arithmetic varies to some extent with the entry (element) size and the problem size. Assume that the computer used is a general purpose one, and the number (n) of prime bases used is reasonable. Then the present algorithm is nearly (n) times faster.

7. Numerical examples

(i) $$A = \begin{bmatrix} 5 & 3 & 7 \\ 2 & 4 & 3 \\ 7 & 7 & 10 \end{bmatrix}$$

The GCD of the elements of A is 1. To bring it to the position (1, 1) subtract 2 times the second row from the first :

$$P_1A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & -5 & 1 \\ 2 & 4 & 3 \\ 7 & 7 & 10 \end{bmatrix}$$

To reduce the first column elements below diagonal zero premultiply P_1A by P_2:

$$P_2P_1A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix} P_1A = \begin{bmatrix} 1 & -5 & 1 \\ 0 & 14 & 1 \\ 0 & 42 & 3 \end{bmatrix}$$

To reduce the first row elements above diagonal zero postmultiply P_2P_1A by Q_1:

$$P_2P_1AQ_1 = P_2P_1A \begin{bmatrix} 1 & 5 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 14 & 1 \\ 0 & 42 & 3 \end{bmatrix}$$

The GCD of the elements of the trailing 2×2 submatrix is 1 which is itself an element. To bring it to position (2, 2) interchange second and third columns, *i.e.*, postmultiply $P_2P_1AQ_1$ by Q_2 :

$$P_2P_1AQ_1Q_2 = P_2P_1AQ_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 14 \\ 0 & 3 & 42 \end{bmatrix}$$

To reduce the second column element below diagonal zero premultiply $P_2P_1AQ_1Q_2$ by P_3:

$$P_3P_2P_1AQ_1Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} P_2P_1AQ_1Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 14 \\ 0 & 0 & 0 \end{bmatrix}$$

To reduce the second row element above diagonal zero postmultiply $P_3P_2P_1AQ_1Q_2$ by Q_3:

$$P_3P_2P_1AQ_1Q_2Q_3 = P_3P_2P_1AQ_1Q_2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -14 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = S\,(\text{SNF})$$

$$P = P_3P_2P_1 = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ -1 & -1 & 1 \end{bmatrix},\ Q = Q_1Q_2Q_3 = \begin{bmatrix} 1 & -1 & 19 \\ 0 & 0 & 1 \\ 0 & 1 & -14 \end{bmatrix}, PAQ = S.$$

$$S^+ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A^- = QS^+ P = \begin{bmatrix} 3 & -7 & 0 \\ 0 & 0 & 0 \\ -2 & 5 & 0 \end{bmatrix}$$

If $b_1 = (15\ 9\ 24)^t$ then $AA^-b_1 = b_1$. Hence a solution exists. If $y = 0$ (null column vector), then $x = A^-b_1 = (-18\ \ 0\ \ 15)^t$. If $y = (19\ 1\ -14)^t$, then $x = (1\ 1\ 1)^t$. If $b_2 = (20\ 2\ 25)^t$, then $AA^-b_2 \neq b_2$. Hence no solution exists, *i.e.*, the system is inconsistent.

$$\text{(ii)}\quad A = \begin{bmatrix} -1 & 2 & 3 & 3 \\ 2 & 5 & 6 & 3 \\ -5 & -8 & -9 & -3 \end{bmatrix}$$

The GCD of the elements of A is 1. To bring it to position (1.1), subtract -1 time the second row from the first :

$$P_1A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 7 & 9 & 6 \\ 2 & 5 & 6 & 3 \\ -5 & -8 & -9 & -3 \end{bmatrix}$$

To reduce the first column elements below diagonal zero premultiply P_1A by P_2 :

$$P_2P_1A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} P_1A = \begin{bmatrix} 1 & 7 & 9 & 6 \\ 0 & -9 & -12 & -9 \\ 0 & 27 & 36 & 27 \end{bmatrix}$$

To reduce the first row elements above diagonal zero postmultiply P_2P_1A by Q_1 :

$$P_2P_1AQ_1 = P_2P_1A \begin{bmatrix} 1 & -7 & -9 & -6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -9 & -12 & -9 \\ 0 & 27 & 36 & 27 \end{bmatrix}$$

The GCD of the elements of the trailing 2 × 3 submatrix is 3. To bring it to the position (2, 2), subtract 1 time the second column from the first :

$$P_2P_1AQ_1Q_2 = P_2P_1AQ_1 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & -12 & -9 \\ 0 & -9 & 36 & 27 \end{bmatrix}$$

To reduce the second column element below diagonal zero premultiply $P_2P_1AQ_1Q_2$ by P_3:

$$P_3P_2P_1AQ_1Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} P_2P_1AQ_1Q_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & -12 & -9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

To reduce the second row elements above the diagonal zero postmultiply $P_3P_2P_1AQ_1Q_2$ by Q_3 :

$$P_3P_2P_1AQ_1Q_2Q_3 = P_3P_2P_1AQ_1Q_2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = S\,(\mathrm{SNF})$$

$$P = P_3P_2P_1 = \begin{bmatrix} 1 & 1 & 0 \\ -2 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}, \; Q = Q_1Q_2Q_3 = \begin{bmatrix} 1 & 2 & -55 & 0 \\ 0 & 1 & 4 & 3 \\ 0 & -1 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; PAQ = S.$$

$$3.A^- = 3.QS^-P = \begin{bmatrix} -1 & 1 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ where } S^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

If $b_1 = (7\ 16\ -25)^t$, then $AA^- b_1 = b_1$. Hence a solution exists. $x = A^- b_1 = (3\ -10\ 10\ 0)^t$ for $y = 0$ (null column vector). If $y = (-2\ 11\ -9\ 1)^t$, then $x = (1\ 1\ 1\ 1)^t$.

References

1. ADEGBEYENI, E. O. AND KRISHNAMURTHY, E. V. — Finite field computational technique for the exact solution of systems of linear equations and interval linear programming problems. *Int. J. Systems Sci.*, 1977, **8** (10), 1181-92.

2. BEN-ISRAEL, A. AND GREVILLE, T. N. E. — *Generalized Inverses : Theory and Applications*, 1974, Wiley Interscience, New York.

3. MARCUS, M. AND MINC, H. — *A Survey of Matrix Theory and Matrix Inequalities*, 1964, Allyn and Bacon, Boston, Mass., USA.

4. HURT, M. F. AND WAID, C. — A generalized inverse which gives all the integral solutions to a system of linear equations. *SIAM J. Appl. Math.*, 1970, **19**, 547-50.