Solving Linear Diophantine Matrix Equations Using the Smith Normal Form (More or Less)

Raymond N. Greenwell¹ and Stanley Kertzner²

¹Department of Mathematics, Hofstra University, Hempstead, NY 11549

matrng@hofstra.edu

²Department of Mathematics, Hofstra University, Hempstead, NY 11540

²Department of Mathematics, Hofstra University, Hempstead, NY 11549 matszk@hofstra.edu

May 16, 2009

Abstract

Using a modification of the Smith normal form of a matrix, we give necessary and sufficient conditions for the integer matrix equation AX = B to have an integer solution, and we provide an explicit formula for that solution. We provide an algorithm for generating the solution, and we also show how the solution can be generated with Maple. **AMS Subject Classification:** 15A06, 15A24, 15A36 **Key Words and Phrases:** matrix equations, linear systems, Smith normal form, diophantine equations, integer solutions

1 Introduction

Solving a system of linear equations with integer coefficients is of both theoretical and practical importance. This is why the topic is taught, not only to mathematics majors, but also to majors in business, engineering, and other disciplines. Finding solutions is equivalent to solving the matrix equation AX = B. The matrix B is often a single column vector, but it need not be so. When the unknowns represent numbers of objects, then the solution must be an integer matrix with nonnegative entries. Thus the primary problem of finding all integer solutions is of intrinsic theoretical interest. The substance of this paper is to resolve the primary problem. Thus we give necessary and sufficient

conditions for the integer matrix equation AX = B to have an integer solution, where A is $m \times n$, B is $m \times p$, and X is $n \times p$, and we provide an explicit formula for that solution.

Example 1. The entries in each column of the matrix A below give the units of fiber, protein, and fat in a container of brands J, K, L, and M of feed for Vietnamese pot-bellied pigs. A breeder would like to know how many containers of each brand of feed he should use to end up with a mixture that contain the units of fiber, protein, and fat given by each column of matrix B below. Column 1 of B gives the nutrient requirements for a male, and column 2 for a female.

This problem boils down to finding all nonnegative integer solutions to the system AX = B. Such a system can be solved by the Gauss-Jordan method. But in an underdetermined system such as Example 1, with two free variables, it is not clear whether or not integer solutions exist, or what they are if they do exist. The reader may wish to try finding all integer solutions to Example 1 using Gauss-Jordan. Our procedure, which gives those solutions explicitly, is similar to those of Lazebnik [5], Newman [6], and Ward [10], but has the advantage of giving a simplified formula for the solutions.

2 An Old Theorem and a New One

The procedure is based on a modification of the following theorem found in Jacobson (see [4], p. 79), which originated with a paper of Smith [9].

Theorem 1. If A is any integer matrix, there exist invertible integer matrices P and Q, whose inverses are also integer matrices, such that

$$PAQ = D$$
,

where D is a diagonal matrix of integers with the property that

$$D = diag\{d_{11}, d_{22}, \cdots, d_{rr}, 0, \cdots, 0\}$$

with d_{ii} a factor of d_{jj} for i < j and for which $d_{ii} \neq 0$ for $i \leq r$.

The matrix D is known as the Smith normal form of the matrix A. Returning to our problem, for given matrices A and B of respective sizes $m \times n$ and $m \times p$, let P, Q, and D be as in Theorem 1. Assume for the moment that r < m and r < n; we'll later consider the cases in which one or both of these inequalities is not true. Define

$$\overline{D} = \operatorname{diag}\{d_{11}, d_{22}, \cdots, d_{rr}\}\$$

and the $(m-r) \times p$ integer matrix U such that

$$PB = \left[\begin{array}{c} \overline{PB} \\ U \end{array} \right],$$

where \overline{PB} is the matrix consisting of the first r rows of PB. Then we have the following theorem.

Theorem 2. AX = B for the integer matrix X if and only if

- a) U = O, a matrix of zeros, and
- b) \overline{D}^{-1} \overline{PB} is an integer matrix.

Then the solutions are all matrices of the form $X = Q \begin{bmatrix} \overline{D}^{-1} \overline{PB} \\ Z \end{bmatrix}$ for some $(n-r) \times p$ integer matrix Z.

Proof. Suppose AX=B for the integer matrix X. Then by Theorem 1, $P^{-1}DQ^{-1}X=B, \text{ and } DQ^{-1}X=PB.$ Let

$$Q^{-1}X = Y = \left[\begin{array}{c} \overline{Y} \\ Z \end{array} \right],$$

where \overline{Y} consists of the first r rows of the $n \times p$ matrix Y. Note that since Q^{-1} and X are integer matrices, so are \overline{Y} and the $(n-r) \times p$ matrix Z. Then DY = PB translates into

$$\left[\begin{array}{cc} \overline{D} & O \\ O & O \end{array}\right] \left[\begin{array}{c} \overline{Y} \\ Z \end{array}\right] = \left[\begin{array}{c} \overline{PB} \\ U \end{array}\right].$$

Thus U = O and $\overline{D} \overline{Y} = \overline{PB}$, so \overline{Y} is equal to $\overline{D}^{-1} \overline{PB}$, which must then be an integer matrix. We also have that

$$X = QY = Q \left[\begin{array}{c} \overline{D}^{-1} \overline{PB} \\ Z \end{array} \right].$$

Suppose now that

$$X = Q \left[\begin{array}{c} \overline{D}^{-1} \overline{PB} \\ Z \end{array} \right]$$

for the integer matrices \overline{D}^{-1} \overline{PB} and Z. Also suppose that U=O. Thus

$$PB = \left[\begin{array}{c} \overline{PB} \\ O \end{array} \right].$$

X is an integer matrix, and by Theorem 1,

$$AX = P^{-1}DQ^{-1}Q \begin{bmatrix} \overline{D}^{-1} \overline{PB} \\ Z \end{bmatrix}$$

$$= P^{-1} \begin{bmatrix} \overline{D} & O \\ O & O \end{bmatrix} \begin{bmatrix} \overline{D}^{-1} \overline{PB} \\ Z \end{bmatrix}$$

$$= P^{-1} \begin{bmatrix} \overline{PB} \\ O \end{bmatrix}$$

$$= P^{-1}PB = B.$$

Remark 1. Note that Theorem 2 is valid even if the non-zero entries d_{11}, \dots, d_{rr} do not divide each other as in the Smith normal form.

3 Jacobson's Algorithm

The algorithm described in Jacobson proceeds by multiplying A on the left and right by integer elementary matrices with integer inverses, which is equivalent to 1) interchanging two rows or two columns, 2) multiplying a row or column by -1, and 3) adding an integer multiple of a row or column to another row or column. We refer the reader to Jacobson for details of the algorithm. When we perform the algorithm by hand, we modify it by allowing division of a row by a constant. One consequence is that we may eliminate a common factor in a row. More significantly, this allows us to make $d_{ii} = 1$ for $i \leq r$, so that the matrix \overline{D} in the theorem, and hence \overline{D}^{-1} , is the identity matrix I_r . As a result, P may not necessarily be an integer matrix, but with P and Q determined in this manner, we have the following corollary.

Corollary 1. AX = B for the integer matrix X if and only if

- a) U = O, a matrix of zeros, and
- b) \overline{PB} is an integer matrix.

Then the solutions are all matrices of the form $X=Q\left[\begin{array}{c}\overline{PB}\\Z\end{array}\right]$ for some integer matrix Z.

When conditions a) and b) of Corollary 1 are true, then all integer solutions of AX = B are found from Corollary 1 by letting the entries of Z range over the integers. In this case, if we create an $n \times r$ matrix Q_1 out of the first r columns of Q, and an $n \times (n-r)$ matrix Q_2 out of the last n-r columns of Q, then it is straightforward to show that the solution X can be written as $Q_1\overline{PB} + Q_2Z$, where the first term is a particular integer solution to the original equation AX = B. Since $PAQ_2 = O$ and P is invertible, Q_2 is an integer solution to the homogeneous equation. That is, $AQ_2 = O$, where O here is $m \times (n-r)$.

Remark 2. Note that Corollary 1 is true whenever invertible integer matrices P and Q can be found for which $PAQ = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ and Q^{-1} is an integer matrix.

4 The Other Cases

We now return to the cases we ignored before. The reader may verify the details for these cases.

Case 1: r=m=n. In this case, there is no U and no Z. Corollary 1 is still valid, but part a) is no longer relevant, \overline{PB} in part b) is just PB, and part c) reduces to X=QPB. Neither the row or column of zero matrices in PAQ will appear.

Case 2: r = m < n. In this case, there is no U. Corollary 1 is still valid, but without part a), and with \overline{PB} replaced by PB in parts b) and c). The row of zero matrices in PAQ will not appear.

Case 3: r = n < m. In this case, there is no Z. Corollary 1 is still valid, but part c) reduces to $X = Q\overline{PB}$. The column of zero matrices in PAQ will not appear.

5 The Modified Algorithm

We now describe our modified version of the algorithm found in Jacobson, and we will then illustrate the procedure by completing Example 1.

Step 1: Form the augmented matrix

$$\left[\begin{array}{cc} A & B \\ I_n & * \end{array}\right],$$

where * is left blank.

Step 2: Let k = 1.

Step 3: Carry through the columns the repeated application of the division algorithm to the entries in row k of A. (If row k is all zeros but a later row is not, interchange rows to correct this.) With each application, retain the remainder when each entry is divided by the entry corresponding to the entry in row k of A with smallest nonzero absolute value.

Step 4: Continue Step 3 until it produces a matrix in which the entries of row k of A are all 0, except for one entry, which is the greatest common divisor of this row.

- Step 5: Divide row k of the augmented matrix by the entry described in Step 4.
- Step 6: Interchange columns of A (as well as the matrix below it) to produce a matrix with a 1 in the pivot position and 0 elsewhere in the row.
- Step 7: Use row operations to change the other elements in column k of A to 0.
- Step 8: Increase k by 1 and repeat Steps 3 through 7.
- Step 9: Continue Steps 3 through 8 until the matrix A has been changed to the form

$$\left[\begin{array}{cc} I_r & O \\ O & O \end{array}\right],$$

where the row of zero matrices will only appear if r < m, and the column of zero matrices will only appear if r < n.

This sequence of row and column operations is equivalent to producing invertible matrices P and Q for which Q and Q^{-1} are integer matrices and

$$PAQ = \left[\begin{array}{cc} I_r & O \\ O & O \end{array} \right].$$

When dividing a row of the augmented matrix by the greatest common divisor of that row, that element must divide into the elements of that row where B originally appeared. If it doesn't, then there are no integer solutions. In that case, the columns of B for which the conditions are not met can be eliminated if one wants to find integer solutions for the remaining columns of B.

Because P is the product of the matrices corresponding to the row operations, and Q is the product of the matrices corresponding to the column operations, the augmented matrix at the end of the algorithm will be of the form

$$\left[\begin{array}{cc} I_r \ O & \overline{PB} \\ O \ O & U \\ Q & * \end{array}\right],$$

where the row of zero matrices will only appear if r < m, and the column of zero matrices will only appear if r < n. It is now immediately revealed whether or not U is a matrix of zeros and whether or not \overline{PB} is an integer matrix. If either of these conditions is not met, then there are no integer solutions. If the conditions are met, the integer solutions are explicitly given by the equation for X in the corollary.

We should add that there are various ways in which one can carry out the procedure. It doesn't matter whether one generates the standard D described by Jacobson and then divides through to convert \overline{D} to I_r , or one converts the diagonal elements to 1 as they are generated. Similarly it doesn't matter whether one simplifies the rows by dividing through by a common factor. Also, rather than strictly following the algorithm, someone might do the allowable operations in a different order, leading to a different P and Q, as long as the ultimate effect is to convert the matrix A to the form described in Step 9 of the algorithm.

Example 1 (revisited). Find all integer solutions to the system
$$AX = B$$
, where $A = \begin{bmatrix} 36 & 10 & 16 & 9 \\ 102 & 80 & 152 & 113 \\ 63 & 95 & 188 & 147 \end{bmatrix}$ and $B = \begin{bmatrix} 624 & 525 \\ 4818 & 4065 \\ 5667 & 4785 \end{bmatrix}$.

We create the augmented matrix

$$\begin{bmatrix} 36 & 10 & 16 & 9 & 624 & 525 \\ 102 & 80 & 152 & 113 & 4818 & 4065 \\ 63 & 95 & 188 & 147 & 5667 & 4785 \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \end{bmatrix}$$

and start with row 1. Following Step 3 of the algorithm, we reduce the elements in the first column with the column operations $C_1 - 4C_4 \rightarrow C_1$, $C_2 - C_4 \rightarrow C_2$, and $C_3 - C_4 \rightarrow C_3$. In Step 4, we reduce the other elements in row 1 to 0 with $C_3 - 7C_2 \rightarrow C_3$ and $C_4 - 9C_2 \rightarrow C_4$. Step 5 is unnecessary, since the remaining element in row 1 is 1. In Step 6, we swap columns 1 and 2. In Step 7, we perform the row operations $R_2 + 33R_1 \rightarrow R_2$ and $R_3 + 52R_1 \rightarrow R_3$ to get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -350 & 270 & 410 \\ 0 & -525 & 405 & 615 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -7 & -9 \\ 0 & 0 & 1 & 0 \\ -1 & -4 & 6 & 10 \end{bmatrix} \begin{bmatrix} 624 & 525 \\ 25,410 & 21,390 \\ 38,115 & 32,085 \end{bmatrix} .$$

Rather than continue to strictly follow the algorithm, we will reduce the numbers in rows 2 and 3 by dividing these rows by 10 and 5, respectively. We then let k=2 and go through Steps 3 through 7 for row 2. We first perform $C_2+C_3\to C_2$ and $C_4-C_3\to C_4$. We next perform $C_3+3C_2\to C_3$ and $C_4+2C_2\to C_4$. We finally get a 1 in row 2 with $C_2-4C_4\to C_2$ and $C_3+C_4\to C_3$. To finish Step 4, we only need $C_4+2C_3\to C_4$. Step 5 is again unnecessary, although it would have been necessary had we not divided the rows earlier. We perform Step 6 by interchanging columns 1 and 2. Step 7

only requires $R_3 - 3R_2 \rightarrow R_3$. The result is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 624 & 525 \\ 0 & 1 & 0 & 0 & 2541 & 2139 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & -7 & 12 \\ 1 & -44 & 57 & -104 \\ 0 & 5 & -3 & 11 \\ -1 & 20 & -30 & 48 \end{bmatrix}.$$

Now
$$\overline{PB} = \begin{bmatrix} 624 & 525 \\ 2541 & 2139 \end{bmatrix}$$
, and

$$X = QY = \begin{bmatrix} 0 & 5 & -7 & 12 \\ 1 & -44 & 57 & -104 \\ 0 & 5 & -3 & 11 \\ -1 & 20 & -30 & 48 \end{bmatrix} \begin{bmatrix} 624 & 525 \\ 2541 & 2139 \\ z_1 & z_3 \\ z_2 & z_4 \end{bmatrix}$$
$$= \begin{bmatrix} 12,705 - 7z_1 + 12z_2 & 10,695 - 7z_3 + 12z_4 \\ -111,180 + 57z_1 - 104z_2 & -93,591 + 57z_3 - 104z_4 \\ 12,705 - 3z_1 + 11z_2 & 10,695 - 3z_3 + 11z_4 \\ 50,196 - 30z_1 + 48z_2 & 42,255 - 30z_3 + 48z_4 \end{bmatrix}$$

where z_1 , z_2 , z_3 , and z_4 are arbitrary integers.

This procedure sometimes generates numbers in the final answer that are unnecessarily large. This may be remedied by a change of variable. Consider the first element $12,705+7z_1-12z_2$. Since $12,705/12 \approx 1059$, it helps to replace z_2 with $-z_2-1059$. (The negative sign in front of z_2 is to make the new z_2 nonnegative when we look for nonnegative solutions. For the same reason, we will replace z_1 with $-z_1$.) Similarly, the numbers in the second column can be reduced by replacing z_4 with $-z_4-891$, and we will replace z_3 with $-z_3$ so that z_2 will have a nonnegative value. The result is

$$X = \begin{bmatrix} -3 + 7z_1 - 12z_2 & 3 + 7z_3 - 12z_4 \\ -1044 - 57z_1 + 104z_2 & -927 - 57z_3 + 104z_4 \\ 1056 + 3z_1 - 11z_2 & 894 + 3z_3 - 11z_4 \\ -636 + 30z_1 - 48z_2 & -513 + 30z_3 - 48z_4 \end{bmatrix}, z_1, z_2, z_3, z_4 \in \mathbb{Z},$$

the complete set of integer solutions. As we mentioned after the corollary, this could also be written as

$$X = \begin{bmatrix} -3 & 3 \\ -1044 & -927 \\ 1056 & 894 \\ -636 & -513 \end{bmatrix} + \begin{bmatrix} 7 & -12 \\ -57 & 104 \\ 3 & -11 \\ 30 & -48 \end{bmatrix} \begin{bmatrix} z_1 & z_3 \\ z_2 & z_4 \end{bmatrix},$$

where the first term is a particular integer solution to the original equation AX = B, and the second term is of the form Q_2Z , where Q_2 is an integer solution to the homogeneous equation.

The primary problem of finding all integer solutions to the matrix equation has now been solved. In the application to finding how many containers of each brand of feed should be used, we need to determine all nonnegative integer solutions, which is equivalent to finding the feasible region for an integer programming problem. Specifically, we must find all integers z_1 , z_2 , z_3 , and z_4 such that all entries in the solution matrix for X are nonnegative. In this case, the solutions are not hard to find by graphing four inequalities. The first column has three solutions. The first is $z_1 = 306$ and $z_2 = 178$, from which we calculate that the numbers of containers of brand J, K, L, and M of feed are 3, 26, 16, and 0, respectively. The other two solutions are $z_1 = 308$, $z_2 = 179$ and $z_1 = 310$, $z_2 = 180$. These lead to the following numbers of containers of each brand of feed: 5, 16, 11, and 12, and 7, 6, 8, and 24, respectively. The second column has two solutions: $z_3 = 259$, $z_4 = 151$ and $z_3 = 261$, $z_4 = 152$. These lead to the following numbers of containers of each brand of feed: 4, 14, 10, and 9, and 6, 4, 5, and 21, respectively.

6 Using Maple

Rather than carry out the algorithm by hand to get the solution in the form given by Corollary 1, we could take advantage of Maple's command for finding the Smith normal form of a matrix, as well as the matrices P and Q, and then apply Theorem 2. Using the matrix A of Example 1, the command

$$(D, P, Q) := SmithForm(A, method = 'integer', output = ['S', 'U', 'V'])$$

gives the result

$$D, P, Q := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 1 \\ 5 & -3 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \\ -31 & -41 & -27 & -44 \\ 29 & 38 & 18 & 41 \\ -17 & -22 & -6 & -24 \end{bmatrix}.$$

(Actually, since D is reserved in Maple, another variable name must be used. Notice also that the matrix D given by Maple is not square.) To see whether or not integer solutions exist, calculate

$$PB = \left[\begin{array}{cc} 624 & 525 \\ 225 & 195 \\ 0 & 0 \end{array} \right].$$

Notice that both D and PB have exactly two nonzero rows, so according to Theorem 2, there are indeed integer solutions, and the complete set is given by

$$X = Q \begin{bmatrix} \overline{D}^{-1} \overline{PB} \\ Z \end{bmatrix} = Q \begin{bmatrix} 624 & 525 \\ 45 & 39 \\ z_1 & z_3 \\ z_2 & z_4 \end{bmatrix}$$
$$= \begin{bmatrix} z_1 & z_3 \\ -21,189 - 27z_1 - 44z_2 & -17,874 - 27z_3 - 44z_4 \\ 19,806 + 18z_1 + 41z_2 & 16,707 + 18z_3 + 41z_4 \\ -11,598 - 6z_1 - 24z_2 & -9,783 - 6z_3 - 24z_4 \end{bmatrix}.$$

As before, we can replace z_2 with $-z_2 - 482$ and z_4 with $-z_4 - 406$. The result

 is

$$X = \begin{bmatrix} z_1 & z_3 \\ 19 - 27z_1 + 44z_2 & -10 - 27z_3 + 44z_4 \\ 44 + 18z_1 - 41z_2 & 61 + 18z_3 - 41z_4 \\ -30 - 6z_1 + 24z_2 & -39 - 6z_3 + 24z_4 \end{bmatrix}, z_1, z_2, z_3, z_4 \in \mathbb{Z}.$$

The values $(z_1, z_2) = (3, 2)$, (5, 3), and (7, 4) lead to the three solutions for column 1 mentioned before. The values $(z_3, z_4) = (4, 3)$ and (6, 4) lead to the two solutions for column 2 mentioned before.

7 Other issues

The problem of finding solutions to linear Diophantine matrix equations has also been addressed by Contejean and Devie [1], Domenjoud [2], and Pottier [8], as well as by others who have refined their methods. Our method is quite different from these in that it gives an explicit formula for the set of solutions, and also because it allows B to have more than one column. Papp and Vizvári [7] investigate methods for applications in chemistry. Dumas et al. [3] have investigated ways to calculate the Smith normal form of a sparse matrix more efficiently.

Our method, of course, also applies to matrix equations of the form XA = B, since this is equivalent to the equation $A^tX^t = B^t$. Our method can also find all integer solutions to the equation AX = B where A and B are matrices with rational entries.

One remaining issue we wish we could resolve is to give an intuitive explanation to what this procedure is doing to the system. It is well known that the row operations of Gauss-Jordan transform the solution hyperplanes so they are parallel to the coordinate axes. We would like to provide a similar insight into the column operations of the procedure we have described, but we are not able to do so. We invite others to explore this question.

References

- [1] E. Contejean and H. Devie, "An efficient incremental algorithm for solving systems of linear Diophantine equations," *Information and Computation* 113 (1994) 143-172.
- [2] E. Domenjoud, "Solving systems of linear Diophantine equations: an algebraic approach," Mathematical Foundations of Computer Science 1991,
 A. Tarlecki, ed., Vol. 520 of Lecture Notes in Computer Science, Springer,
 1991.
- [3] J.-G. Dumas, B. D. Saunders, and G. Villard, "On efficient sparse integer matrix Smith normal form computations," *Journal of Symbolic Computa*tion, 32 (2001) 71-99.
- [4] N. Jacobson, Lectures in Abstract Algebra, Vol. 2: Linear Algebra, Van Nostrand, 1953.
- [5] F. Lazebnik, "On systems of linear Diophantine equations," The Mathematics Magazine 69 (1996) 261-266.
- [6] M. Newman, *Integral Matrices*, Academic Press, 1972.
- [7] D. Papp and B. Vizvári, "Effective solution of linear Diophantine equation systems with an application in chemistry," *Journal of Mathematical Chemistry* 39 (2006) 15-31.
- [8] L. Pottier, "Minimal solutions of linear diophantine systems: bounds and algorithms," Rewriting Techniques and Applications, R. V. Book, ed., Vol. 488 of Lecture Notes in Computer Science, Springer, 1991.

- [9] H. J. S. Smith, "On systems of linear indeterminate equations and congruences," *Philosophical Transactions of the Royal Society of London* 151 (1861) 293-326.
- [10] A. J. B. Ward, "A matrix method for a system of linear Diophantine equations," The Mathematical Gazette 84 (2000) 81-84.