Using Consistency Conditions to Solve Linear Systems

Geza Schay (gs@cs.umb.edu), University of Massachusetts at Boston, Boston, MA 02425

We present a new strategy of row reduction to obtain a basis for the left nullspace of a matrix (and, by transposition, one for the nullspace too). The method can also be extended to solve a linear system, and to invert a nonsingular square matrix.

Most introductory linear algebra books contain examples and exercises in which they ask for conditions on the right-hand side of a linear system to ensure consistency. The usual recommendation is to reduce the system $A\mathbf{x} = \mathbf{b}$ by elementary row operations to a form $U\mathbf{x} = \mathbf{c}$, where U is an echelon matrix, and set the components of $\mathbf{c}$ that correspond to the zero rows of U equal to zero. Since $\mathbf{c}$ is obtained from $\mathbf{b}$ by elementary row operations, we thus obtain a set of homogeneous linear equations for the components of $\mathbf{b}$. It is then natural to ask questions about this system such as: Are its equations independent and what characterizes its coefficient matrix?

To avoid the components of $\mathbf{b}$ being hidden in $\mathbf{c}$ and to obtain their coefficient matrix explicitly we write the system as $A\mathbf{x} = I\mathbf{b}$ and reduce the latter or, equivalently, the augmented matrix $[A|I]$. Let us look at an example:

Example 1. Find conditions on $\mathbf{b}$ that ensure the consistency of $A\mathbf{x} = \mathbf{b}$ with

$$A = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & -1 \\ 3 & -3 & 0 \\ 2 & 0 & -2 \end{bmatrix}.$$

According to the above discussion we can do this by the following reduction:

$$\left[\begin{array}{rrr|rrrr} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ -2 & 3 & -1 & 0 & 1 & 0 & 0 \\ 3 & -3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & -2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{rrr|rrrr} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 3 & -3 & 2 & 1 & 0 & 0 \\ 0 & -3 & 3 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{rrr|rrrr} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 3 & -3 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 1 \end{array}\right].$$

Thus if we put

$$M = \begin{bmatrix} -1 & 1 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix},$$

the consistency conditions can be written as $M\mathbf{b} = \mathbf{0}$. Clearly the rows of this M are independent. On the other hand we know that the equation $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A, and so the solutions of $M\mathbf{b} = \mathbf{0}$ must make up the column space of A. Hence the rows of M must be orthogonal to the column space of A. It is straightforward to check that in this case indeed $MA = O$. ■

We can generalize the results of the above example as follows:

Theorem 1. *Let A be any real $m \times n$ matrix of rank r. Consider the block matrix $[A \quad I]$, where I is the unit matrix of order m. This matrix can be reduced by elementary row operations to a form $\begin{bmatrix} U & L \\ O & M \end{bmatrix}$, in which U is an $r \times n$ echelon matrix and O the $(m-r) \times n$ zero matrix. Then the transposed rows[1] of the $(m-r) \times m$ matrix M form a basis for the left nullspace of A.* ■

Proof. For any A as stated, the equation $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b} \in \mathrm{Col}(A)$. For any $\mathbf{b} \in \mathrm{Col}(A)$ write the above equation as $A\mathbf{x} = I\mathbf{b}$ and reduce the latter, by elementary row operations, until A is in an echelon form $\begin{bmatrix} U \\ O \end{bmatrix}$, with U having no zero rows. On the right-hand side denote the result of this reduction of the matrix I by $\begin{bmatrix} L \\ M \end{bmatrix}$. Thus we get the equations $U\mathbf{x} = L\mathbf{b}$ and $\mathbf{0} = M\mathbf{b}$. The last equation shows that the rows of M must be orthogonal to any vector in the column space of A, and so their transposes are in the left nullspace of A. Furthermore, the matrix $\begin{bmatrix} L \\ M \end{bmatrix}$ has full rank, since it is obtained from I by elementary row operations, which are invertible. Consequently the rows of M are independent. On the other hand, since the dimension of the left nullspace of A is $m - r$ and M has $m - r$ independent rows, their transposes span the left nullspace of A. ■

The construction of Theorem 1 applied to the transpose A^T of A in place of A yields a basis for the nullspace of A, as illustrated in the next example.

Example 2. Find a basis for the nullspace of

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & -1 & 5 & 1 \end{bmatrix}.$$

[1]We follow the convention of considering the left nullspace of A to be a space of column vectors.

We row-reduce $[A^T \;\; I]$ as follows:

$$\left[\begin{array}{rr|rrrr} 1 & 2 & 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{rr|rrrr} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & -5 & -2 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 & 1 & 0 \\ 0 & -5 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{rr|rrrr} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & -5 & -2 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right].$$

From here we can read off the matrix M as

$$M = \begin{bmatrix} -2 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix},$$

whose transposed rows form a basis for Null(A).

For comparison let us find a basis by the usual method as well:

We reduce A to

$$U = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & -1 & 1 & -1 \end{bmatrix},$$

and solve $U\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$. We set the free variables equal to parameters, that is, set $x_3 = s$ and $x_4 = t$. Then $U\mathbf{x} = \mathbf{0}$ becomes $x_1 + 2x_2 + 3t = 0$ and $-x_2 + s - t = 0$. Hence $x_2 = s - t$ and $x_1 = -2s - t$, and so the general solution is

$$\mathbf{x} = s\begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix},$$

resulting in the same basis vectors for Null(A) as before.

Our method seems to be more straightforward, but on the other hand it requires working with a bigger matrix. ■

We can also extend the above procedure to one for solving $A\mathbf{x} = \mathbf{b}$, by reformulating it as $[x^T \; 1]\begin{bmatrix} A^T \\ -\mathbf{b}^T \end{bmatrix} = \mathbf{0}$ and finding the left nullspace of $\begin{bmatrix} A^T \\ -\mathbf{b}^T \end{bmatrix}$.

We can further modify the last idea so as to obtain a new way of computing the inverse of a matrix. We may solve $AX = I$ for an invertible matrix A by solving the systems $A\mathbf{x_i} = \mathbf{e_i}$ with the above method, where $\mathbf{x_i}$ and $\mathbf{e_i}$ are the columns of X and I respectively. An efficient organization of this yields the following theorem for computing A^{-1}.

Theorem 2. *Let A be any real, nonsingular $n \times n$ matrix. Consider the block matrix $\begin{bmatrix} A & I \\ -I & O \end{bmatrix}$, where I is the unit matrix of order n and O is the $n \times n$ zero matrix. We can row-reduce this, without exchanging any of the last n rows, to a form $\begin{bmatrix} U & L \\ O & M \end{bmatrix}$, in which U is an upper triangular matrix row-equivalent to A. Then $M = A^{-1}$.* ■

Example 3. Find the inverse of

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}.$$

Applying Theorem 2 we can do this by the following reduction:

$$\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 10 & -3 & 1 \\ \hline 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 10 & -3 & 1 \\ \hline 0 & 0 & 4/10 & 2/10 \\ 0 & 0 & -3/10 & 1/10 \end{array}\right].$$

Thus

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix}. \quad \blacksquare$$

There seems to be no computational advantage or disadvantage to the proposed methods, since they require just as many operations as the standard ones. What we gain in avoiding back substitution, we lose because of the enlarged matrices.

Acknowledgment. The author wishes to thank the referee for suggesting several valuable improvements.

An Application of Number Theory

Professor James C. Kirby (Tarleton State University, kirby@tarleton.edu) sends the following.

In practicing baseball with my two boys, I found that I was too predictable in throwing them a fly or grounder. Generally, if I threw one a fly, then I also threw the other one a fly. I sought a (somewhat) random way of deciding how to do this. I came up with the following procedure, and it works quite well. Beginning with the day of the week if it is odd, or the next day if it is even, I throw a ground ball if the integer is a prime and a fly ball if it is a non-prime. For the next one, I use the next odd integer. After 99, I go to one (hence the reason for non-prime and not composite) and continue until I get to the number before the starting point. This gives each one of them twenty-five opportunities with no pattern.

My little girl hasn't yet started softball, so she is not included. One might, however, consider what would be a good technique for three fielders.

How many fathers use mathematics while playing baseball with their children? Let me rephrase that. How many fathers know the difference between a prime and a composite number?