Use Elementary Column Operations to Calculate the Basis of the Null Space of a Matrix

CHEN BIHONG
Mathematics College, ShenZhen University
Email: cbhmath@hotmail.com

---

abstracts: This paper gives and proofs a theorem, for any matrix A, do elementary column operations, change it to a matrix which is partitioned to two blocks which left one is column full rank and right one is zero matrix. That is, use a invertible matrix P to let AP=(B,O), O is zero matrix with n-r columns, r and n is rank and column number of A, so the P's right n-r columns is just the basis of the null space of the matrix A. On the basis of the theorem, lots of problems of linear algebra can be resolved and lots of theorems can be proofed by elementary column operations. Perhaps the textbooks used in universities will have a lot of change with the result of the paper. This result is first found by author in 2010.12.8 in http://www.paper.edu.cn/index.php/default/releasepaper/content/201012-232, but is not formal published.

keywords: linear algebra, null space, systems of linear equations

1 Intruduction

In linear algebra, the kernel or null space (also nullspace) of a matrix **A** is the set of all vectors **x** for which $\mathbf{Ax} = \mathbf{0}$. The kernel of a matrix with n columns is a linear subspace of n-dimensional Euclidean space. The dimension of the null space of **A** is called the nullity of **A**. With traditional method, for calculating the basis of the null space of **A**, we must use row reduction to find a basis for the null space. That is, first use elementary row operations to put **A** in reduced row echelon form, then interpreting the reduced row echelon form as a homogeneous linear system, determine which of the variables $x_1, x_2, \cdots, x_n$ are free, then write equations for the dependent variables in terms of the free variables, then for each free variable x_i, choose the vector in the null space for which $x_i = 1$ and the remaining free variables are zero then resulting collection of vectors is a basis for the null space of **A** [1]. In this paper we call it elementary row operations method.

Now I give another method which is called elementary column operations method. It is given by theorem as follows.

theorem 1 (Chen Bihong theorem) Given an $m \times n$ matrix $\mathbf{A}$ with $\text{rank}(\mathbf{A}) = r, r < n$, $\mathbf{E}_n$ is a unit matrix of size n, construct a partition matrix $\begin{pmatrix} \mathbf{A}_{m\times n} \\ \mathbf{E}_n \end{pmatrix}$, do elementary column operations to it to let $\mathbf{A}$ become $(\mathbf{B}_{m\times r}, \mathbf{O}_{m\times(n-r)})$ where $\mathbf{O}_{m\times(n-r)}$ is a zero matrix with $n-r$ columns, so $\begin{pmatrix} \mathbf{A}_{m\times n} \\ \mathbf{E}_n \end{pmatrix}$ is changed to $\begin{pmatrix} \mathbf{B}_{m\times r} & \mathbf{O}_{m\times(n-r)} \\ \mathbf{P}_{n\times r} & \mathbf{Q}_{n\times(n-r)} \end{pmatrix}$ and the all $n-r$ columns of $\mathbf{Q}$ are the basis of null space of matrix $\mathbf{A}$.

Proof this theorem is easy, but I will proof it use some special method in next sections to suggest some teaching methods.

2 Old and New System of Linear Equations

We call a usual system of linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1}$$

as old system, where $\mathbf{A}$ is a matrix with m rows and n columns and $\mathbf{b}$ is a column vector of n items. Give any invertible square matrix $\mathbf{P}$ of order m and invertible square matrix $\mathbf{Q}$ of order n, change the old system of (1) to

$$\mathbf{PAQy} = \mathbf{Pb} \tag{2}$$

witch are called new system. Given an solution ξ of the old system (1), so

$$\mathbf{A}\xi = \mathbf{b} \tag{3}$$

and let $\eta = \mathbf{Q}^{-1}\xi$, then η is solution of new system, because

$$\mathbf{PAQ}\eta = \mathbf{PAQQ}^{-1}\xi = \mathbf{PA}\xi = \mathbf{Pb}.$$

Otherwise, given a solution η of the new system (2), we have

$$\mathbf{PAQ}\eta = \mathbf{Pb} \tag{4}$$

Left multiply P^{-1} on two sides of the equation can get

$$\mathbf{AQ}\eta = \mathbf{b}.$$

So $\xi = \mathbf{Q}\eta$ is solution of the old system (1).

So the two sets of solutions of old and new systems have relation of one to one and invertible linear transformation. If we select suitable $\mathbf{P}$ and $\mathbf{Q}$ to let new system (2) have simple form, we can find the new method to calculate solutions of system of linear equations.

3 Standard System

We know concept of standard matrix is that zero matrix is standard matrix, and unit matrix is standard matrix, and a matrix which is got by adding some zero rows under a standard matrix or adding some zero columns on right of a standard matrix is also standard matrix. We use $\mathbf{D}$ to express a standard matrix in this paper.

We call a system of linear equations as standard system if its matrix of coefficients is standard matrix. We express it as follows.

$$\mathbf{Dx} = \mathbf{d} \tag{5}$$

where $\mathbf{D}$ is $m \times n$ standard matrix with $\mathrm{rank}(\mathbf{D}) = r$, $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$ is a column vector with n entries, and $\mathbf{d} = (d_1, d_2, \cdots, d_m)^T$ is a column vector with m entries.

So the standard system (5) can be written as

$$\begin{cases} x_1 = d_1, \\ x_2 = d_2, \\ \cdots \\ x_r = d_r, \\ 0 = d_{r+1}, \\ 0 = d_{r+2}, \\ \cdots \\ 0 = d_m. \end{cases} \tag{6}$$

The standard system (5) or (6) has solutions if and only if $d_{r+1}, d_{r+2}, \cdots, d_m$ not exist or are all zeros. If $r = n$, the system has unique solution directly given by (6). But we are more interested on the situation of $n > r$.

When system (6) has solution, it directly give the values of first r variables $x_1 = d_1, x_2 = d_2, \cdots, x_r = d_r$, but last $n - r$ variables $x_{r+1}, x_{r+2}, \cdots, x_n$ are not appeared, so that they can take any number $x_{r+1} = c_1, x_{r+2} = c_2, \cdots, x_n = c_{n-r}$, where $c_1, c_2, \cdots, c_{n-r}$ are $n - r$ numbers which can take any numbers. So the general solutions of standard system (6) can be written as

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ c_1 \\ c_2 \\ \vdots \\ c_{n-r} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} 0 \\ \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ \\ \vdots \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + c_{n-r} \begin{pmatrix} 0 \\ \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{7}$$

So let $\gamma = (d_1, d_2, \cdots, d_r, 0, \cdots, 0)^T$ which is also a solution of (6) when $c_1 = c_2 = \cdots = c_{n-r} = 0$ and γ is like $\mathbf{d}$ of (1) but is not same as $\mathbf{d}$ because it is n-dimension vector and $\mathbf{d}$ is m-dimension vector. Let $\mathbf{e}_i$ is column vector of i-th column of unit matrix $\mathbf{E}$ of order n, $i = 1, 2, \cdots, n$. So formula (7) can be written as

$$\mathbf{x} = \gamma + c_1\mathbf{e}_{r+1} + c_2\mathbf{e}_{r+2} + \cdots + c_{n-r}\mathbf{e}_n \tag{8}$$

If $\mathbf{d} = \mathbf{0}$ in system (5), so system $\mathbf{Dx} = \mathbf{0}$ is called homogeneous standard system. So we know the vectors $\mathbf{e}_{r+1}, \mathbf{e}_{r+2}, \cdots, \mathbf{e}_n$ consist one basis of null space of matrix $\mathbf{D}$.

4 General Systems

Now we discuss the general system or old system $\mathbf{Ax} = \mathbf{b}$. We assume that $\mathbf{A}$ is not zero matrix. So do some elementary operations to $\mathbf{A}$ to transform $\mathbf{A}$ to a standard matrix $\mathbf{D}$. So exist invertible matrix $\mathbf{P}$ and $\mathbf{Q}$ so that

$$\mathbf{D} = \mathbf{PAQ} \tag{9}$$

where $\mathbf{D}$ is a standard matrix. So let

$$\mathbf{d} = \mathbf{Pb} \tag{10}$$

then we get new system which is standard system $\mathbf{Dx} = \mathbf{d}$.

We change vector $\mathbf{x}$ in (5) to $\mathbf{y}$ for convenient. So the standard system is $\mathbf{Dy} = \mathbf{d}$ and its general solution is written as

$$\mathbf{y} = \gamma + c_1\mathbf{e}_{r+1} + c_2\mathbf{e}_{r+2} + \cdots + c_{n-r}\mathbf{e}_n \tag{11}$$

So the general solution of the old system is

$$\mathbf{x} = \mathbf{Q}\mathbf{y} = \mathbf{Q}\gamma + c_1\mathbf{Q}\mathbf{e}_{r+1} + c_2\mathbf{Q}\mathbf{e}_{r+2} + \cdots + c_{n-r}\mathbf{Q}\mathbf{e}_n \tag{12}$$

where to make

$$\eta = \mathbf{Q}\gamma, \xi_1 = \mathbf{Q}\mathbf{e}_{r+1}, \xi_2 = \mathbf{Q}\mathbf{e}_{r+2}, \cdots, \xi_{n-r} = \mathbf{Q}\mathbf{e}_n, \tag{13}$$

According to the nature of matrix multiplication, $\xi_1, \xi_2, \cdots, \xi_{n-r}$ are just column $r+1, r+2, \cdots, n$ of matrix $\mathbf{Q}$, so

$$\mathbf{x} = \eta + c_1\xi_1 + c_2\xi_2 + \cdots + c_{n-r}\xi_{n-r} \tag{14}$$

So it is easy to proof that η is a specific solution of old system $\mathbf{Ax} = \mathbf{b}$, and $\xi_1, \xi_2, \cdots, \xi_{n-r}$ compose basis of null space of matrix $\mathbf{A}$.

The key technology is to get invertible matrix $\mathbf{P}$ and $\mathbf{Q}$. First get the constant vector $\mathbf{d} = \mathbf{Pb}$ of standard system by elementary row operations to matrix $(\mathbf{A}, \mathbf{b})$. The matrix $\mathbf{Q}$ includes all records of elementary row operations to matrix $\mathbf{A}$. So construct a partitioned matrix $\begin{pmatrix} \mathbf{A} \\ \mathbf{E} \end{pmatrix}$, do elementary row operations to $\mathbf{A}$ part and do elementary column operations to entire matrix to change $\mathbf{A}$ part to standard matrix $\mathbf{D}$, then the part below is changed to matrix $\mathbf{Q}$, that is, $\begin{pmatrix} \mathbf{A} \\ \mathbf{E} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{D} \\ \mathbf{Q} \end{pmatrix}$, so the last $n-r$ columns of $\mathbf{Q}$ construct the basis of null space of matrix $\mathbf{A}$.

When the method is realized by computer, elementary row operations are not need to do. So we have proofed Chen Bihong theorem.

5 Examples

Now use new method to calculate a basis of null space of matrix $\mathbf{A}$ where

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 2 & -2 & -2 & 2 \\ -1 & 1 & 2 & -2 \end{pmatrix}$$

Then add a identity matrix below $\mathbf{A}$ then do elementary column operations as follows:

$$\left(\frac{\mathbf{A}}{\mathbf{E}}\right) = \left(\begin{array}{cccc} 1 & -1 & 1 & -1 \\ 2 & -2 & -2 & 2 \\ -1 & 1 & 2 & -2 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[c_4+c_1]{\substack{c_2+c_1 \\ c_3-c_1}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 0 & -4 & 4 \\ -1 & 0 & 3 & -3 \\ \hline 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 0 & -4 & 4 \\ -1 & 0 & 3 & -3 \\ \hline 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{c_4+c_3} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 0 & -4 & 0 \\ -1 & 0 & 3 & 0 \\ \hline 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

In the Horizontal line above we got two zero columns, under the two zero columns we get two vectors $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$, and they constitute a basis of the null space of matrix $\mathbf{A}$.

6 Conclusion

So after Chen Bihong theorem is discovered, a lot of concepts of linear algebra, like "free variable", "reduced row echelon", "Gaussian elimination", become stupid and will be discarded or less use. Entire textbook of linear algebra should rewrite, and a lot of computer programs to calculate the solution of system of linear equations should be reprogrammed.

References

[1] Lay, David C. (August 22, 2005), Linear Algebra and Its Applications (3rd ed.), Addison Wesley, ISBN 978-0321287137.

[2] Chen Bihong. Use Elementary Column Operations to Resolve Systems of Linear Equations. 2010-12-08. (in chinese) http://www.paper.edu.cn/index.php/default/releasepaper/content/201012-232.

[3] Chen Bihong. Adding Rows and Columns to Augmented Matrix to Solving Systems of Linear Equations. 2010-08-30. (in chinese) http://www.paper.edu.cn/index.php/default/releasepaper/content/201008-452.