

# Existence of Left/Right/Two-sided Inverses

P. Sam Johnson

September 19, 2014

## Definition

Let  $A$  be a  $m \times n$  matrix. Then the **column space** of  $A$  is  $\mathcal{C}(A)$  is

$$\mathcal{C}(A) := \{Ax : x \in \mathbb{R}^n\}$$

and the **row space** of  $A$  is

$$\mathcal{R}(A) := \{y^T A : y \in \mathbb{R}^m\}.$$

- We call  $\dim(\mathcal{R}(A))$  the **row rank** of  $A$  and  $\dim(\mathcal{C}(A))$  the **column rank** of  $A$ .
- We refer to a basis of  $\mathcal{C}(A)$  consisting of columns of  $A$  as a **column basis**. A **row basis** is defined similarly.

# Matrix Multiplication

**Notation.**  $A_{i*}$  denotes the  $i$ -th row of  $A$  and  $A_{*j}$  denotes the  $j$ -th column of  $A$ .

Let  $A, B, C$  be matrices of orders  $m \times n, n \times p$ , and  $p \times q$  respectively.  
Then

- 1  $(AB)_{ij} = A_{i*}B_{*j}$ ,
- 2  $(AB)_{i*} = A_{i*}B$ ,
- 3  $(AB)_{*j} = AB_{*j}$ ,
- 4  $(ABC)_{ij} = A_{i*}BC_{*j}$ .
- 5 For any  $m \times n$  matrix  $A$ , we have  $A_{i*} = e_i^T A$  and  $A_{*j} = Ae_j$ .

If  $A$  has column rank  $r$ , then

- any  $r$  linearly independent columns of  $A$  form a basis for  $\mathcal{C}(A)$ ,
- every maximal linearly independent set of columns of  $A$  contains exactly  $r$  vectors,
- any  $r$  columns of  $A$  which generate  $\mathcal{C}(A)$  form a basis of  $\mathcal{C}(A)$ .

### Theorem

*For any matrix  $A$ , the row rank of  $A$  equals the column rank of  $A$ .*

**Proof.** Let  $A$  be a  $m \times n$  matrix with row rank  $r$  and column rank  $s$ . If  $A = 0$ , then  $\mathcal{R}(A) = \{0\}$  and  $\mathcal{C}(A) = \{0\}$ , so  $r = s = 0$  and we are done.

Let  $B = [x_1 : x_2 : \cdots : x_s]$  be an  $m \times s$  matrix whose columns form a basis of  $\mathcal{C}(A)$ . Then for each  $j = 1, 2, \dots, n$ , each column of  $A$ ,  $A_{*j}$  is a linear combination of the columns of  $B$ , so there exists an  $s \times 1$  vector  $y_j$  such that  $A_{*j} = By_j$ .

Now

$$\begin{aligned} A &= [A_{*1} : \cdots : A_{*n}] = [By_1 : \cdots : By_n] \\ &= B[y_1 : \cdots : y_n] = BC \end{aligned}$$

where  $C = [y_1 : \cdots : y_n]$ . Note that  $C$  is of size  $s \times n$ .

Since  $A = BC$ ,  $A_{i*} = B_{i*}C$ , and each row of  $A$  is a linear combination of the rows of  $C$ . Thus  $\mathcal{R}(A) \subseteq \mathcal{R}(C)$ .

Taking dimensions, we get  $r \leq \text{row rank}(C)$ .

As  $C$  has only  $s$  rows,  $\text{row rank}(C) \leq s$ . Hence  $r \leq s$ .

**Interchanging the roles of row rank and the column rank.** Let  $C = [y_1 : \cdots : y_r]^T$  be an  $r \times n$  matrix whose rows form a basis of  $\mathcal{R}(A)$ . Then for each  $i = 1, 2, \dots, n$ , each row of  $A$ ,  $A_{i*}$  is a linear combination of the rows of  $C$ , so there exists an  $r \times 1$  vector  $x_i$  such that  $A_{i*} = x_i C$ .

$$\begin{aligned} A &= [A_{1*} : \cdots : A_{n*}]^T = [x_1 C : \cdots : x_n C]^T \\ &= [x_1 : \cdots : x_n]^T C = BC \end{aligned}$$

where  $B = [x_1 : \cdots : x_n]^T$ . Note that  $C$  is of size  $s \times n$ .

Since  $A = BC$ ,  $A_{*j} = B \cdot C_{*j}$ , and each column of  $A$  is a linear combination of the columns of  $C$ . Thus  $\mathcal{C}(A) \subseteq \mathcal{C}(B)$ .

Taking dimensions, we get  $s \leq \text{column rank}(B)$ . As  $B$  has only  $r$  columns,  $\text{column rank}(B) \leq r$ . Hence  $s \leq r$ . Thus  $r = s$ .

## Definition

The **rank** of a matrix  $A$  is the common value of the row rank of  $A$  and the column rank of  $A$  and is denoted by  $\rho(A)$ .

- The rank of an  $m \times n$  matrix obviously lies between 0 and  $\min(m, n)$ .
- Conversely, given any non-negative integer  $r \leq \min(m, n)$ , there exists an  $m \times n$  matrix  $A$  with rank  $r$ .
- Let  $A$  be a  $m \times n$  matrix of rank  $r$  and  $B$  a submatrix of  $A$ . By considering row rank (column rank) if  $B$  is obtained from  $A$  by omitting only some rows (columns). Any submatrix can be obtained by omitting some rows and then some columns. Then  $\rho(B) \leq \rho(A)$ .

## Definition

An  $m \times n$  matrix  $A$  is said to be of **full row rank** if its rows are linearly independent, that is, if its rank is  $m$ . Similarly  $A$  is said to be of **full column rank** if its columns are linearly independent.

A **left inverse** of a matrix  $A$  is any matrix  $B$  such that  $BA = I$ . A **right inverse** of  $A$  is any matrix  $C$  such that  $AC = I$ .

A matrix  $B$  is said to be an **inverse** of  $A$  if it is both a left inverse and a right inverse of  $A$ .



## Theorem

Let  $A$  be a  $m \times n$  matrix over  $\mathbb{R}$ . Then the following statements are equivalent.

- 1  $A$  has a right inverse.
- 2 **Right cancellation law:**  $XA = YA \Rightarrow X = Y$ .
- 3  $XA = 0 \Rightarrow X = 0$ .
- 4  $A$  is of full row rank.
- 5 **The linear transformation  $f : x \mapsto Ax$  is onto:**  $\mathcal{C}(A) = \mathbb{R}^m$ .

**Question:** If  $A$  has a right inverse, how many right inverses does  $A$  have ?

## Theorem

Let  $A$  be a  $m \times n$  matrix over  $\mathbb{R}$ . Then the following statements are equivalent.

- 1  $A$  has a left inverse.
- 2 **Left cancellation law:**  $AX = AY \Rightarrow X = Y$ .
- 3  $AX = 0 \Rightarrow X = 0$ .
- 4  $A$  is of full column rank.
- 5 **The linear transformation  $f : x \mapsto Ax$  is one-to-one:**  $\mathcal{R}(A) = \mathbb{R}^n$ .

- A matrix  $B$  is a left inverse of a matrix  $A$  iff  $B^T$  is a right inverse of  $A^T$ .
- If  $B$  and  $C$  are left inverses of  $A$ , then  $\alpha B + (1 - \alpha)C$  is also a left inverse of  $A$ .

If a matrix  $A$  has a left inverse  $B$  and a right inverse  $C$ , then

- $A$  is square,
- $B = C$ ,
- $A$  has a unique left inverse, a unique right inverse and a unique inverse.

If a matrix  $A$  has an inverse, then  $A^{-1}$  is unique,  $A$  is square and  $AA^{-1} = A^{-1}A = I$ .

### Theorem

*Let  $A$  be a square matrix of order  $n$ . Then the following statements are equivalent:*

- 1  $A$  has a right inverse
- 2 rank of  $A$  is  $n$
- 3  $A$  has a left inverse
- 4  $A$  has an inverse.

## Definition

A square matrix  $A$  is said to be **non-singular** if it has an inverse. A square matrix which does not possess an inverse is said to be **singular**.

- 1  $AB$  is non-singular iff both  $A$  and  $B$  are non-singular.
- 2 If  $A$  is non-singular and  $k$  is a positive integer, then  $A^k$  is non-singular and its inverse is  $(A^{-1})^k$ .
- 3 The sum of two non-singular matrices need not be non-singular.
- 4 Let  $P$  be a permutation matrix. Then  $P$  is non-singular and  $P^{-1} = P^T$ .
- 5 If  $P$  is a permutation matrix obtained from  $I$  by interchanging two rows,  $P^{-1} = P$ .

## Left and Right Inverses

We know that if  $A$  has a left-inverse ( $BA = I$ ) and a right-inverse ( $AC = I$ ), then the two inverses are equal:  $B = B(AC) = (BA)C = C$ . Now, from the rank of a matrix, it is easy to decide which matrices actually have these inverses.

Roughly speaking, an inverse exists only when the rank is as large as possible. In other words, the matrix has an inverse if  $A$  has to have to full rank,  $\text{rank}(A) = \min\{m, n\}$ .

The rank always satisfies  $r \leq m$  and also  $r \leq n$ . As  $m$  by  $n$  matrix cannot have more than  $m$  independent rows or  $n$  independent columns. When  $r = m$  there is a right-inverse, and when  $r = n$  there is a left-inverse.

In the first case  $AX = b$  always has a solution. In the second case the solution (if it exists) is unique. Only a square matrix can have both  $r = m$  and  $r = n$ , and therefore only a square matrix can achieve both existence and uniqueness. Only a square matrix has a two-sided inverse.

## Existence and Uniqueness

**Existence.** The system  $Ax = b$  has **at least** one solution  $x$  for every  $b$  iff the columns span  $\mathbb{R}^m$ ; then  $r = m$ . In this case there exists an  $n$  by  $m$  right inverse  $C$  such that  $AC = I_m$ , the identity matrix of order  $m$ . This is possible only if  $m \leq n$ .

**Uniqueness.** The system  $Ax = b$  has **at most** one solution  $x$  for every  $b$  iff the columns are linearly independent; then  $r = n$ . In this case there exists an  $n$  by  $m$  left-inverse  $B$  such that  $BA = I_n$ , the identity matrix of order  $n$ . This is possible only if  $m \geq n$ .

In the first case, one possible solution is  $x = Cb$ , since then  $Ax = ACb = b$ . But there will be other solutions if there are other right-inverses.

In the second case, if there is a solution to  $Ax = b$ , it has to be  $x = BAx = Bb$ . But there may be no solution.

## Simple Formulas for Left and Right Inverses

There are simple formulas for left and right inverses, if they exist:

$B = (A^T A)^{-1} A^T$  and  $C = A^T (A A^T)^{-1}$ . Certainly  $BA = I$  and  $AB = I$ .

What is not so certain is that  $A^T A$  and  $A A^T$  are actually invertible. We show that  $A^T A$  does have an inverse if the rank is  $n$ , and  $A A^T$  has an inverse when the rank is  $m$ . Thus the formulas make sense exactly when the rank is as large as possible, and the one-sided inverses are found.

There is also a more basic approach. We can look, a column at a time, for a matrix  $C$  such that  $AC = I$  or  $A[x_1 \ x_2 \ \cdots \ x_m] = [e_1 \ e_2 \ \cdots \ e_m]$ . Each column of  $C$ , when multiplied by  $A$ , gives a column of the identity matrix. To solve  $Ax_i = e_i$  we need the coordinate vectors  $e_i$  to be in the column space. If it contains all those vectors, the column space must be all of  $\mathbb{R}^m$ ! Its dimension (the rank) must be  $r = m$ . This is the “existence case,” when the columns span  $\mathbb{R}^m$ .

## Example

Consider a simple 2 by 3 matrix of rank 2:  $A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}$ . Since  $r = m = 2$ , the theorem guarantees a right-inverse  $C$ :

$$AC = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \\ a & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In fact, there are many right-inverses; the last row of  $C$  is completely arbitrary. This is a case of existence but no uniqueness. The matrix  $A$  has no left-inverse because the last column of  $BA$  is certain to be zero, and cannot agree with the 3 by 3 identity matrix.



The transpose of  $A$  yields an example in the opposite direction, with infinitely many left-inverses:

$$\begin{pmatrix} 1/4 & 0 & c \\ 0 & 1/5 & d \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Now it is the last column of  $B$  that is completely arbitrary. This is typical of the “uniqueness case,” when the  $n$  columns of  $A$  are linearly independent. The rank is  $r = n$ . There are no free variables, since  $n - r = 0$ , so if there is a solution it will be the only one. You can see when this example has a solution:

$$\begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \text{ is solvable if } b_3 = 0.$$

When  $b_3$  is zero, the solution (unique!) is  $x_1 = \frac{1}{4}b_1, x_2 = \frac{1}{5}b_2$ .

For a rectangular matrix, it is not possible to have both existence and uniqueness. If  $m$  is different from  $n$ , we cannot have  $r = m$  and  $r = n$ . A square matrix is the opposite. If  $m = n$ , we cannot have one property without the other.

A square matrix has a left-inverse iff it has a right-inverse. There is only one inverse, namely  $B = C = A^{-1}$ . Existence implies uniqueness and uniqueness implies existence, when the matrix is square. The condition for this invertibility is that the rank must be as large as possible:  $r = m = n$ .

We can say that in another way: For a square matrix  $A$  of order  $n$  to be nonsingular, each of the following conditions is a necessary and sufficient test:

- 1 The columns span  $\mathbb{R}^n$ , so  $Ax = b$  has at least one solution for every  $b$ .
- 2 The columns are independent, so  $Ax = 0$  has only the solution  $x = 0$ .

## A Typical Application: Vandermonde's matrix

Consider all polynomials

$$P(x) = a_1 + a_2x + a_3x^2 + \cdots + a_nx^{n-1}$$

of degree  $n - 1$ .

The only such polynomial that vanishes at  $n$  given points  $x_1, \dots, x_n$  is the zero polynomial  $P(x) \equiv 0$ .

No other polynomial of degree  $n - 1$  can have  $n$  roots. This is a statement of uniqueness, and it implies a statement of existence: Given any values  $b_1, \dots, b_n$ , there exists a polynomial of degree  $n - 1$  interpolating these values:  $P(x_i) = b_i$ ,  $i = 1, \dots, n$ .

The point is that we are dealing with a square matrix; the number of coefficients in  $P(x)$  (which is  $n$ ) matches the number of equations. In fact the equations  $P(x_i) = b_i$  are the same as

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

The coefficient matrix  $A$  is  $n$  by  $n$ , and is known as *Vandermonde's matrix*.

Since  $Ax = 0$  has only the solution  $x = 0$  (in other words  $P(x_i) = 0$  is only possible if  $P \equiv 0$ ), it follows that  $A$  is nonsingular.

Thus  $Ax = b$  always has a solution - a polynomial can be passed through any  $n$  values  $b_i$  at distinct points  $x_i$ .

The determinant of the Vandermonde's matrix is  $\prod_{1 \leq i < j \leq n} (x_j - x_i)$ . The determinant is nonzero because all  $x_i$ 's are distinct.

## Rank-one Matrices

Finally comes the easiest case, when the rank is as small as possible (except for the zero matrix with rank zero). One of the basis themes of mathematics is, given something complicated, to show how it can be broken into simple pieces.

For linear algebra the simple pieces are matrices of **rank one**,  $r = 1$ .

The following example is typical:  $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{pmatrix}$ .

Every row is a multiple of the first row, so the row space is one-dimensional.

In fact, we can write the whole matrix in the following special way, as **the product of a column vector and a row vector**:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \\ -1 \end{pmatrix} = (2 \ 1 \ 1).$$

The product of a 4 by 1 matrix and a 1 by 3 matrix is a 4 by 3 matrix, and this product has rank one. Note that, at the same time, the columns are all multiples of the same column vector; the column space shares the dimension  $r = 1$  and reduces to a line.

The same thing will happen for any other matrix of rank one: **Every matrix of rank one has the simple form  $A = uv^T$** . The rows are all multiples of the same vector  $v^T$ , and the columns are all multiples of the same vector  $u$ . The row space and column space are lines.



## Storing Rank-One Matrix as a Matrix

If  $x, y \neq 0$ , then any matrix of the form  $B = xy^T$  has rank one; that is, its columns span a one-dimensional space. Conversely, any rank-one matrix  $B$  can be represented in the form  $xy^T$ .

Rank-one matrices arise frequently in numerical applications, and it is important to know how to deal with them.

The first thing to note is that one does not store a rank-one matrix as a matrix. For example, if  $x$  and  $y$  are  $n$ -vectors, then the matrix  $xy^T$  requires  $n^2$  locations to store, as opposed to  $2n$  locations to store  $x$  and  $y$ .

To get some idea of the difference, suppose that  $n = 1000$ . Then  $xy^T$  requires one million ( $1000^2$ ) words to store as a matrix, as opposed to 2000 to store  $x$  and  $y$  individually - the storage differs by a factor of 500.

## Product of Rank-One Matrix and Column Vector

If we always represent a rank-one matrix  $B = xy^T$  by storing  $x$  and  $y$ , the question arises of how we perform matrix operations with  $B$  - how, say, we can compute the matrix-vector product  $c = Bb$ ?

An elegant answer to this question may be obtained from the equation  $c = Bb = (xy^T)b = x(y^T b) = (y^T b)x$ , in which the last equality follows from the fact that  $y^T b$  is a scalar.

This equation leads to the following algorithm.

- 1 Compute  $z = y^T b$
- 2 Compute  $c = zx$ .

This algorithm requires  $2n$  multiplications and  $n - 1$  additions. This should be contrasted with the roughly  $n^2$  multiplications and additions required to form an ordinary matrix vector product.