Using pivot positions to prove the Invertible Matrix Theorem in Lay's Linear Algebra

11/4/2011

This handout provides an alternate proof to the Invertible Matrix Theorem from Chapter 2 of Lay's Linear Algbra with the intention of helping linear algebra students organize the theorem into easy to understand parts. As this is a companinion to Lay's book, the perspective is centered around the theory of pivot positions that is well developed by Lay. The various theorems regarding pivot positions from Chapter 1 of Lay's book are digested in another handout titled "Interpretting Pivot Positions". That handout will be cited in the proof presented here.

Following the proof of the theorem will be some commentary on the consequences of the proof on the construction of left and right inverses of not-necessarily-invertible matrices.

While this handout is written with the 4^{th} edition ([2]) in mind, references for the 3^{rd} edition ([1]) are also included if they differ from those for the 4^{th}.

Statement of the Theorem

Here is the theorem in question.

Theorem (The Invertible Matrix Theorem)**.** *[2, Theorem 8 from Chapter 2, page 112] Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A the statements are either all true or all false.*

a. A is an invertible matrix.

b. A is row equivalent to the $n \times n$ identity matrix.

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T is an invertible matrix.

Proof of the Theorem

Throughout this proof the fact that only one pivot position can be found in a particular row or columns is used. In light of this a matrix with n columns (or rows) with n pivot postitions must have a pivot postition in each column (or row). This fact follows from the definition of reduced row echelon form in section 1.2 of [2].
In this proof I use the following lemmas. The first two include part of the content of "Interpretting Pivot Positions."

Lemma 1. *Let A be a matrix with n rows. The following are equivalent.*

c. A has n pivot positions.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.

k. There is an $n \times n$ matrix C such that $AD = I$.

Proof. Let m denote the number of columns of A.
As in Lay's proof, $(k) \Rightarrow (h)$ is Exercise 26 of Section 2.1 of [2] (Exercise 24 in [1]).
To prove the converse, let $\mathbf{e}_i$ denote the i^{th} column of the identity matrix I_n. If (h) holds, the equation $A\mathbf{x} = \mathbf{e}_i$ always has a solution. Let D be a $m \times n$ matrix, the i^{th} column of which is a solution to $A\mathbf{x} = \mathbf{e}_i$. In this case $AD = I_n$. This proves $(k) \Leftrightarrow (h)$.
The conditions other than (k) are shown to be equivalent in Chapter 1 of [2]. These equivalences are organized in the "Interpretting Pivot Positions" handout. The equivalence of (c), (g) and (h) is Theorem 4 of Chapter 1 of [2]. That $(i) \Leftrightarrow (h)$ is part (a) of Theorem 12 of Chapter 1 of [2]. $\square$

Lemma 2. *Let A be a matrix with n columns. The following are equivalent.*

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

j. There is an $n \times n$ matrix C such that $CA = I$.

Proof. The conditions other than (c) and (j) are shown to be equivalent in Chapter 1 of [2]. These equivalences are organized in the "Interpretting Pivot Positions" handout. The equivalence $(d) \Leftrightarrow (e)$ is Theorem 11 of Chapter 1 of [2]. That $(f) \Leftrightarrow (e)$ is part (b) of Theorem 12 of Chapter 1 of [2].
The equivalence of (c) and (d) is implicitly noted in a boxed statement at the beginning of Section 1.5 of [2], "The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a

nontrivaial solution if and only if the equation has at least one free variable." Free variables are represented by columns without pivot positions. As a result, if A has a pivot position in each column there are no free variables and no non-trivial solution to the homogeneous system. Conversely, no non-trivial solution implies no free variables and thereby a pivot position in each column.
As in Lay's proof, $(j) \Rightarrow (d)$ is Exercise 23 of Section 2.1 of [2].
Let m denote the number of rows of A.
Denote by R_A the matrix which is the product of the elementary matrices that perform the reduced row reduction on a matrix A by multiplication on the left. Let C' be the $n \times m$ matrix with entries c_{ij}, where $c_{ij} = 1$ if the j^{th} row of A contains the pivot position of the i^{th} pivot column of A and $c_{ij} = 0$ otherwise.[1]
If A has n pivot postitions, $C'R_AA = I$ so $C = C'R_A$ satisfies (j). This proves $(c) \Rightarrow (j)$. □

Lemma 3. *Let A be a matrix. Then the following statements are equivalent.*

a. A is an invertible matrix.

b. A is row equivalent to the $n \times n$ identity matrix.

c.' A has dimensions $n \times n$ and has n pivot positions.

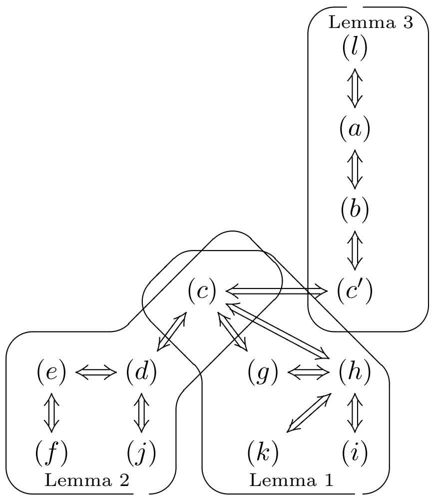
l. A^T is an invertible matrix.

Proof. As is pointed out in Lay's proof, $(a) \Rightarrow (k)$ is a consequence of part (c) of Theorem 6 from Chapter 2 of [2]. To prove the other statements are equivalent I follow Lay's lead, nearly quoting him.
A is $n \times n$ and has n pivot positions if and only if the pivots lie on the main diagonal, if and only if the reduced echelon form of A is I_n. Thus $(c') \Leftrightarrow (b)$. Also $(b) \Leftrightarrow (a)$ by Theorem 7 of Section 2.2. □

These three lemmas in fact complete the proof of the theorem as A being an $n \times n$ matrix means it has n rows and n columns. `q.e.d.`

[1] For examples of this construction see the Comments section of this handout.

Here is a diagram of the structure of this proof.



(1)

Comments

In addition to providing the granularity of the three Lemma's, this proof provides the construction for left and right inverses if they exist.

Constructing Right Inverses

To construct a right inverse of a matrix, first find the reduced row-echelon form of the partitioned matrix $[A|I]$. For this discussion label this matrix $[A'|B]$. This matrix will reveal whether A has a right inverse as it will establish the number of pivot positions in A. Let $\mathbf{b}_i$ denote the columns of B. Note that a solution of $A'\mathbf{x} = \mathbf{b}_i$ is a solution to $A\mathbf{x} = \mathbf{e}_i$. Any such solution is a good candidate for the i^{th} column of a right inverse of A. In particular on can construct a solution for which all of the parameters are zero. The columns may then have a non-trivial solution to the corresponding homogeneous system added to them. Here is an example. Let

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

Now I row reduce the matrix

$$[A|I_4] = \left[\begin{array}{ccccc|cccc} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right], \tag{3}$$

which yields

$$[A'|B] = \left[\begin{array}{ccccc|cccc} 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & 1 \end{array}\right]. \tag{4}$$

This reveals that A has four pivot positions and therefor has a right inverse. Note that the fourth column represents the parameter.
Setting the parameter to 0, the following vectors are solutions

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{5}$$

which provide the following right inverse of A,

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 \end{bmatrix}. \tag{6}$$

Note that this is B with a zero row inserted in the row of the parameter. That is, the fourth row.
To employ a non-zero parameter r in any column would be to add to it the vector

$$\begin{bmatrix} r \\ -r \\ -r \\ r \\ 0 \end{bmatrix}. \tag{7}$$

Setting r_i for the i^{th} column, the general form for a right inverse of A is

$$\begin{bmatrix} 1+r_1 & r_2 & r_3 & r_4-1 \\ -r_1 & 1-r_2 & -r_3 & -r_4 \\ 1-r_1 & 1-r_2 & -1-r_3 & -r_4 \\ r_1 & r_2 & r_3 & r_4 \\ -1 & -1 & 1 & 1 \end{bmatrix}. \tag{8}$$

By specifying the values of each parameter,

$$r_1 = 3,\ r_2 = -2,\ r_3 = 0 \text{ and } r_4 = 1 \tag{9}$$

one specifies the following right inverse of A

$$\begin{bmatrix} 4 & -2 & 0 & 0 \\ -3 & 3 & 0 & -1 \\ -2 & 3 & -1 & -1 \\ 3 & -2 & 0 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}. \tag{10}$$

Constructing Left Inverses

In order to construct left inverses on again begins with reduced row-echelon form of the partitioned matrix $[A|I]$. The result will again be denoted $[A'|B]$. Here $B = R_A$ from the proof of Lemma 2. Multiplying B by C' as specified in the proof[2] produces a left inverse. In fact any non-pivot column of C' may be filled with whatever entries one like and the result will still be a left inverse. That is because that column will be multiplied by a row of zeroes in the reduced matrix BA. These non-pivot columns will always be on the right-most side of C', since the non-pivot rows of BA are at the bottom of the matrix.
Here is an example. Let

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}. \tag{11}$$

The reduced row echelon form of $[A|I_5]$ is

$$[A'|B] = \left[\begin{array}{cccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 2 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & 0 \end{array}\right] \tag{12}$$

Here

$$C' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{13}$$

As a result the general form of a left inverse of A is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & r_1 \\ 0 & 1 & 0 & 0 & r_2 \\ 0 & 0 & 1 & 0 & r_3 \\ 0 & 0 & 0 & 1 & r_4 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & -1 & -1 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 0 \end{bmatrix}. \tag{14}$$

[2]Let C' be the $n \times m$ matrix with entries c_{ij}, where $c_{ij} = 1$ if the j^{th} row of A contains the pivot position of the i^{th} pivot column of A and $c_{ij} = 0$ otherwise.

Setting

$$r_1 = 3,\ r_2 = -2,\ r_3 = 0 \text{ and } r_4 = 1, \tag{15}$$

produces the left inverse

$$\begin{bmatrix} 3 & -2 & -1 & 2 & -1 \\ -2 & 3 & 3 & -2 & -1 \\ 0 & 0 & -1 & 0 & 1 \\ 1 & -2 & -2 & 2 & 1 \end{bmatrix}. \tag{16}$$

Using the transpose

Due to Theorem 3 of Chapter 2 of [2], these methods are interchangable. That is, to find s left-inverse C of A one may find the right inverse D of A^T. Setting $C = D^T$ will suffice since

$$D^T A = (A^T D)^T = I^T = I. \tag{17}$$

Likewise finding the right-inverse of A can be reduced to finding the left inverse of A.

To Be Continued

In Section 2.9 of Lay's Linear Algebra, the Invertible Matrix Theorem is continued. If this is discussed in class a companion handout building on this one will be distributed.

Bibliography

[1] David E. Lay. *Linear Algebra and Its Applications.* Addison-Wesley, 3rd edition, 2006. update.

[2] David E. Lay. *Linear Algebra and Its Applications.* Addison-Wesley, 4th edition, 2012. update.