Math 313 Lecture #11
§2.3: Characterizations of Invertible Matrices

We have already used (and will prove shortly) one characterization of the invertibility of a square matrix A, namely that it is row equivalent to I.

There are many, many more characterizations of the invertibility of A.

Theorem 8 (Inverse Matrix Theorem). Let A be an $n \times n$ matrix. The following are equivalent.

a. A is invertible.
b. A is row equivalent to I.
c. A has n pivot positions.
d. The equation $A\vec{x} = \vec{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one.
g. The equation $A\vec{x} = \vec{b}$ as at least one solution for each $\vec{b} \in \mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\vec{x} \mapsto A\vec{x}$ 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 invertible.

Proof. (a) $\Rightarrow$ (b). Suppose that A is invertible.

By Theorem 5, the equation $A\vec{x} = \vec{b}$ has a (unique) solution for each $\vec{b}$ in $\mathbb{R}^n$.

By Theorem 4 in Section 1.4, each row of A has a pivot position in it.

Since A is square, each column has a pivot position in it.

The only place for these n pivot positions in an $n \times n$ matrix are on the diagonal entries of A.

This implies that the reduced echelon form of A is I, and hence A is row equivalent to I.

(b) $\Rightarrow$ (a). Suppose that A is row equivalent to I.

Then there are finitely many elementary matrices $E_1, E_2, \ldots, E_p$ such that

$$(E_p \cdots E_2 E_1)A = I.$$

To get the invertibility of A we only have to show that $A(E_p \cdots E_2 E_1) = I$.

Since each elementary matrix is invertible, the product $E_p \cdots E_2 E_1$ is invertible.

Hence multiplying both sides of $E_p \cdots E_2 E_1 A = I$ on the *left* by $(E_p \cdots E_2 E_1)^{-1}$ gives $A = (E_p \cdots E_2 E_1)^{-1} I = I(E_p \cdots E_2 E_1)^{-1}$.

Since $(E_p \cdots E_2 E_1)^{-1}$ is invertible with inverse $E_p \cdots E_2 E_1$, multiplying both sides of $A = I(E_p \cdots E_2 E_1)$ by $E_p \cdots E_2 E_1$ on the *right* gives $A(E_p \cdots E_2 E_1) = I$.

Thus A is invertible with inverse $A^{-1} = E_p \cdots E_2 E_1$.

Equivalence of (a) and/or (b) with remaining statements: all of these are restatements of previous Theorems we have seen.

For instance, (b) and (c) are equivalent because the only $n \times n$ reduced row echelon form with n pivot positions is the identity matrix.

And (a) and (l) are equivalent because when A is invertible then A^T is invertible, and when A^T is invertible, then $(A^T)^T = A$ is invertible. $\square$

What is also impressive about all of the logically equivalent statements in the Inverse Matrix Theorem is that of the negations of the statement are logically equivalent as wel (the contrapositive of each implication).

For instance, we know that A is not invertible when $A\vec{x} = \vec{0}$ has more than the trivial solution, or when the columns of A do not span $\mathbb{R}^n$.

Probably the easiest way to determine if a matrix A is invertible or not is to simply row reduce it to see if it has enough pivot positions or not.

Example. Determine the invertibility of

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

With enough row reduction we identify the pivot positions:

$$\begin{aligned}
\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} R_1 \leftrightarrow R_4 &\sim \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} R_3 + R_1 \to R_3 \\
&\sim \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix} R_2 \leftrightarrow R_3 \\
&\sim \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} R_4 + R_2 \to R_4 \\
&\sim \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} R_4 + R_3 \to R_4 \\
&= \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix}
\end{aligned}$$

The matrix A has four pivot positions and so it is invertible.

Invertible Linear Transformation. We say that a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\vec{x})) = \vec{x} \text{ for all } x \in \mathbb{R}^n, \text{ and } T(S(\vec{x})) = \vec{x} \text{ for all } x \in \mathbb{R}^n.$$

This is equivalent to saying the T is a bijection, i.e., T is both injective (one-to-one) and surjective (onto).

When T is invertible, we call S the inverse of T and write $T^{-1} = S$.

We will use the Inverse Matrix Theorem to characterize an invertible linear transformation.

Theorem 9. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and let A be its standard matrix. Then T is invertible if and only A is an invertible matrix. When T is invertible, the inverse of T is the unique linear transformation, given by $T^{-1}(\vec{x}) = A^{-1}\vec{x}$, that satisfies $T(S(\vec{x})) = \vec{x}$ and $S(T(\vec{x})) = \vec{x}$ for all $\vec{x}$ in $\mathbb{R}^n$.

Proof. Suppose T is invertible.

Then there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T(S(\vec{y})) = \vec{y}$ for all $\vec{y}$ in $\mathbb{R}^n$.

For an arbitrary $\vec{b}$ in $\mathbb{R}^n$ set $\vec{x} = S(\vec{b})$.

Then $T(\vec{x}) = T(S(\vec{b})) = \vec{b}$.

Since $T(\vec{x}) = A\vec{x}$, we have shown that $A\vec{x} = \vec{b}$ has a solution for each $\vec{b}$ in $\mathbb{R}^n$.

By the Inverse Matrix Theorem, the matrix A is invertible.

Now suppose that A is invertible.

Define the matrix transformation $S : \mathbb{R}^n \to \mathbb{R}^n$ by $S(\vec{x}) = A^{-1}\vec{x}$.

Then $T(S(\vec{x})) = A(A^{-1}\vec{x}) = (AA^{-1})\vec{x} = I\vec{x} = \vec{x}$ and $S(T(\vec{x})) = A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = I\vec{x} = \vec{x}$ for all $\vec{x}$ in $\mathbb{R}^n$.

Thus T is invertible.

The uniqueness part of the result is a homework problem (#38). $\square$

In the proof of Theorem 9, when showing that T being an invertible linear transformation implies A being an invertible matrix, we used the existence of S such that $T(S(\vec{x})) = \vec{x}$ but not $S(T(\vec{x})) = \vec{x}$.

This should make you suspicious about this part of the proof as we did not make use of all the available information connected with the assumed invertibility of T.

Either we have made a mistake, or there is something else undiscovered that is true.

If we instead worked with $S(T(\vec{x})) = \vec{x}$, we could prove that T is one-to-one.

How? Well for $\vec{u}, \vec{v}$ in $\mathbb{R}^n$ with $T(\vec{u}) = T(\vec{v})$ we have

$$\vec{u} = S(T(\vec{u})) = S(T(\vec{v})) = \vec{v},$$

and so T is one-to-one, meaning that the matrix transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one.

By the Inverse Matrix Theorem, the matrix A is invertible, and we reach the same conclusion.

The proof of Theorem 9 is correct, so what is undiscovered truth here? That a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one if and only if T is onto.

This follows by the Inverse Matrix Theorem because $\vec{x} \mapsto A\vec{x}$ is one-to-one if and only if $\vec{x} \mapsto A\vec{x}$ is onto.