Matrix inverses

Recall...

**Definition** A square matrix $A$ is invertible (or nonsingular) if $\exists$ matrix $B$ such that $AB = I$ and $BA = I$. (We say $B$ is an inverse of $A$.)

**Remark** Not all square matrices are invertible.

**Theorem.** If $A$ is invertible, then its inverse is unique.

**Remark** When $A$ is invertible, we denote its inverse as $A^{-1}$.

**Theorem.** If $A$ is an $n \times n$ invertible matrix, then the system of linear equations given by $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$.

**Proof.** Assume $A$ is an invertible matrix. Then we have

$$A(A^{-1}\vec{b}) = (AA^{-1})\vec{b} = I\vec{b} = \vec{b}.$$  

Thus, $\vec{x} = A^{-1}\vec{b}$ is a solution to $A\vec{x} = \vec{b}$.

Suppose $\vec{y}$ is another solution to the linear system. It follows that $A\vec{y} = \vec{b}$, but multiplying both sides by $A^{-1}$ gives $\vec{y} = A^{-1}\vec{b} = \vec{x}$. □

**Theorem** (Properties of matrix inverse).

(a) If $A$ is invertible, then $A^{-1}$ is itself invertible and $(A^{-1})^{-1} = A$.

(b) If $A$ is invertible and $c \neq 0$ is a scalar, then $cA$ is invertible and $(cA)^{-1} = \frac{1}{c}A^{-1}$.

(c) If $A$ and $B$ are both $n \times n$ invertible matrices, then $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. 
“socks and shoes rule” – similar to transpose of $AB$

generalization to product of $n$ matrices

(d) If $A$ is invertible, then $A^T$ is invertible and $(A^T)^{-1} = (A^{-1})^T$.

To prove (d), we need to show that the matrix $B$ that satisfies $BA^T = I$ and $A^T B = I$ is $B = (A^{-1})^T$.

**Proof of (d).** Assume $A$ is invertible. Then $A^{-1}$ exists and we have

$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$$

and

$$A^T (A^{-1})^T = (A^{-1}A)^T = I^T = I.$$  
So $A^T$ is invertible and $(A^T)^{-1} = (A^{-1})^T$. ■

Recall...

How do we compute the inverse of a matrix, if it exists?

**Inverse of a $2 \times 2$ matrix:** Consider the special case where $A$ is a $2 \times 2$ matrix with $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then $A$ is invertible and its inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

How do we find inverses of matrices that are larger than $2 \times 2$ matrices?

**Theorem.** If some EROs reduce a square matrix $A$ to the identity matrix $I$, then the same EROs transform $I$ to $A^{-1}$.

$$\begin{bmatrix} A & I \end{bmatrix} \overset{\text{EROs}}{\rightarrow} \begin{bmatrix} I & A^{-1} \end{bmatrix}$$

If we can transform $A$ into $I$, then we will obtain $A^{-1}$. If we cannot do so, then $A$ is not invertible.
Can we capture the effect of an ERO through matrix multiplication?

**Definition** An *elementary matrix* is any matrix obtained by doing an ERO on the identity matrix.

**Examples**

\[
\begin{bmatrix}
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\[R_1 \leftrightarrow R_2\] on \(4 \times 4\) identity

\[
\begin{bmatrix}
1 & 0 & -4 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

\[R_1 - 4R_3\] on \(3 \times 3\) identity

Notice that

\[
\begin{bmatrix}
1 & 0 & -4 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix} =
\begin{bmatrix}
a_{11} - 4a_{31} & a_{12} - 4a_{32} & a_{13} - 4a_{33} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix}
\]

Left mult. of \(A\) by row vector is a linear comb. of rows of \(A\).

**Remark** An elementary matrix \(E\) is invertible and \(E^{-1}\) is elementary matrix corresponding to the “reverse” ERO of one associated with \(E\).

**Example** If \(E\) is 2nd elementary matrix above, then “reverse” ERO is \(R_1 + 4R_3\) and \(E^{-1} =
\begin{bmatrix}
1 & 0 & 4 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}\).

**Remark** When finding \(A^{-1}\) using Gauss-Jordan elimination of \([A \mid I]\), if we keep track of EROs, and if \(E_1, E_2, \ldots, E_k\) are corresponding elem. matrices, then we have

\[E_k E_{k-1} \cdots E_1 A = I \implies A = E_1^{-1} \cdots E_{k-1}^{-1} E_k^{-1}.\]
**Theorem** (Fundamental Thm of Invertible Matrices). For an \( n \times n \) matrix, the following are equivalent:

1. \( A \) is invertible.
2. \( A\vec{x} = \vec{b} \) has a unique solution for any \( \vec{b} \in \mathbb{R}^n \).
3. \( A\vec{x} = \vec{0} \) has only the trivial solution \( \vec{x} = 0 \).
4. The RREF of \( A \) is \( I \).
5. \( A \) is product of elementary matrices.

**Proof.**

(1) \( \Rightarrow \) (2):
Proven in first theorem of today’s lecture

(2) \( \Rightarrow \) (3):
If \( A\vec{x} = \vec{b} \) has unique sol’n for any \( \vec{b} \in \mathbb{R}^n \), then in particular, \( A\vec{x} = \vec{0} \) has a unique sol’n. Since \( \vec{x} = 0 \) is a solution to \( A\vec{x} = \vec{0} \), it must be the unique one.

(3) \( \Rightarrow \) (4):
If \( A\vec{x} = \vec{0} \) has unique sol’n \( \vec{x} = 0 \), then augmented matrix has no free variables and a leading one in every column:

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
\vdots & \vdots & \vdots \\
1 & 0 & 0
\end{bmatrix}
\]

so RREF of \( A \) is \( I \).

(4) \( \Rightarrow \) (5):
\[E_k \cdots E_1 A = \text{RREF of } A = I\]
and elem. matrices are invertible
\[\Rightarrow \quad A = E_1^{-1} \cdots E_k^{-1} E_k^{-1}.
\]

(5) \( \Rightarrow \) (1):
Since \( A = E_k \cdots E_1 \) and \( E_i \) invertible \( \forall i \), \( A \) is product of invertible matrices so it is itself invertible.

\[\blacksquare\]
**Theorem.** Let $A$ be a square matrix. If $B$ is a square matrix such that either $AB = I$ or $BA = I$, then $A$ is invertible and $B = A^{-1}$.

**Proof.** Suppose $A$, $B$ are $n \times n$ matrices and that $BA = I$. Then consider the homogeneous system $A\vec{x} = \vec{0}$. We have

$$B(A\vec{x}) = B\vec{0} \quad \Rightarrow \quad (BA)\vec{x} = \vec{0} \quad \Rightarrow \quad \vec{x} = \vec{0}.$$ 

Since $A\vec{x} = \vec{0}$ has only the trivial solution $\vec{x} = \vec{0}$, by the Fundamental Thm of Inverses, we have that $A$ is invertible, i.e., $A^{-1}$ exists. Thus,

$$(BA)A^{-1} = IA^{-1} \quad \Rightarrow \quad B(AA^{-1}) = A^{-1} \quad \Rightarrow \quad B = A^{-1}.$$ 

We leave the case of $AB = I$ as an exercise.  

---

**Definition**  The vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \in \mathbb{R}^n$, where $\vec{e}_i$ has a one in its $i$th component and zeros elsewhere, are called *standard unit vectors*.

**Example**  The $4 \times 4$ identity matrix can be expressed as

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mid & \mid & \mid & \mid \\ \vec{e}_1 & \vec{e}_2 & \vec{e}_3 & \vec{e}_4 \end{bmatrix}$$

**Theorem.** If some EROs reduce a square matrix $A$ to the identity matrix $I$, then the same EROs transform $I$ to $A^{-1}$.

**Why does this work?**

Want to solve $AX = I$, with $X$ unknown $n \times n$ matrix.
If $\vec{x}_1, \ldots, \vec{x}_n$ are columns of $A$, then want to solve $n$ linear systems $A\vec{x}_1 = \vec{e}_1, \ldots, A\vec{x}_n = \vec{e}_n$. Can do so simultaneously using one “super-augmented matrix.”