Matrices, transposes, and inverses

Math 40, Introduction to Linear Algebra
Wednesday, February 1, 2012

Matrix-vector multiplication: two views

- 1st perspective: \( A\vec{x} \) is linear combination of columns of \( A \)

\[
\begin{bmatrix}
1 & -2 & 3 \\
2 & 1 & 5 \\
\end{bmatrix}
\begin{bmatrix}
4 \\
3 \\
2 \\
\end{bmatrix}
= 4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} -2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 5 \end{bmatrix}
= \begin{bmatrix} 4 \\ 21 \end{bmatrix}
\]

- 2nd perspective: \( A\vec{x} \) is computed as dot product of rows of \( A \) with vector \( \vec{x} \)

\[
\begin{bmatrix}
1 & -2 & 3 \\
2 & 1 & 5 \\
\end{bmatrix}
\begin{bmatrix}
4 \\
3 \\
2 \\
\end{bmatrix}
= \text{dot product of } \begin{bmatrix} 4 \\ 2 \end{bmatrix} \text{ and } \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}
= \begin{bmatrix} 4 \\ 21 \end{bmatrix}
\]

Notice that \# of columns of \( A = \# \) of rows of \( \vec{x} \).
This is a requirement in order for matrix multiplication to be defined.
Matrix multiplication

What sizes of matrices can be multiplied together?

For $m \times n$ matrix $A$ and $n \times p$ matrix $B$, the matrix product $AB$ is an $m \times p$ matrix.

If $A$ is a square matrix and $k$ is a positive integer, we define

$$A^k = A \cdot A \cdots A$$

$k$ factors

Properties of matrix multiplication

Most of the properties that we expect to hold for matrix multiplication do....

$$A(B + C) = AB + AC$$

$$(AB)C = A(BC)$$

$k(AB) = (kA)B = A(kB)$ for scalar $k$

.... except commutativity!!

In general, $AB \neq BA$. 
Matrix multiplication not commutative

Problems with hoping $AB$ and $BA$ are equal:

- $BA$ may not be well-defined.
  
  (e.g., $A$ is $2 \times 3$ matrix, $B$ is $3 \times 5$ matrix)

- Even if $AB$ and $BA$ are both defined, $AB$ and $BA$ may not be the same size.
  
  (e.g., $A$ is $2 \times 3$ matrix, $B$ is $3 \times 2$ matrix)

- Even if $AB$ and $BA$ are both defined and of the same size, they still may not be equal.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix} \neq \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Truth or fiction?

**Question 1** For $n \times n$ matrices $A$ and $B$, is

$$A^2 - B^2 = (A - B)(A + B)$$

**No!!**

$$ (A - B)(A + B) = A^2 + \overbrace{AB - BA} - B^2 \neq 0 $$

**Question 2** For $n \times n$ matrices $A$ and $B$, is $(AB)^2 = A^2 B^2$?

**No!!**

$$(AB)^2 = ABAB \neq AABB = A^2 B^2$$
Matrix transpose

**Definition**  The *transpose* of an $m \times n$ matrix $A$ is the $n \times m$ matrix $A^T$ obtained by interchanging rows and columns of $A$,

\[ (A^T)_{ij} = A_{ji} \quad \forall \ i, j. \]

**Example**

\[
A = \begin{bmatrix}
1 & 3 & 5 & -2 \\
5 & 3 & 2 & 1
\end{bmatrix}
\]
\[
A^T = \begin{bmatrix}
1 & 5 \\
3 & 3 \\
5 & 2 \\
-2 & 1
\end{bmatrix}
\]

Transpose operation can be viewed as flipping entries about the diagonal.

**Definition**  A square matrix $A$ is *symmetric* if $A^T = A$.

Properties of transpose

(1)  $(A^T)^T = A$

(2)  $(A + B)^T = A^T + B^T$

(3)  For a scalar $c$, $(cA)^T = cA^T$

(4)  $(AB)^T = B^TA^T$

To prove this, we show that

\[
[(AB)^T]_{ij} = \\
\vdots = [(B^TA^T)]_{ij}
\]

**Exercise**

Prove that for any matrix $A$, $A^T A$ is symmetric.
Special matrices

**Definition** A matrix with all zero entries is called a *zero matrix* and is denoted 0.

\[
A = \begin{bmatrix}
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

**Definition** A square matrix is *upper-triangular* if all entries below main diagonal are zero.

\[
A = \begin{bmatrix}
    2 & \frac{1}{4} & 5 \\
    0 & 6 & 0 \\
    0 & 0 & -3 \\
\end{bmatrix}
\]

**Definition** A square matrix is *lower-triangular* if all entries above main diagonal are zero.

**Definition** A square matrix whose off-diagonal entries are all zero is called a *diagonal matrix*.

\[
A = \begin{bmatrix}
    -\frac{3}{8} & 0 & 0 & 0 \\
    0 & -2 & 0 & 0 \\
    0 & 0 & -4 & 0 \\
    0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

**Definition** The *identity matrix*, denoted \( I_n \), is the \( n \times n \) diagonal matrix with all ones on the diagonal.

\[
I_3 = \begin{bmatrix}
    1 & 0 & 0 \\
    0 & 1 & 0 \\
    0 & 0 & 1 \\
\end{bmatrix}
\]

Identity matrix

**Definition** The *identity matrix*, denoted \( I_n \), is the \( n \times n \) diagonal matrix with all ones on the diagonal.

\[
I_3 = \begin{bmatrix}
    1 & 0 & 0 \\
    0 & 1 & 0 \\
    0 & 0 & 1 \\
\end{bmatrix}
\]

Important property of identity matrix

If \( A \) is an \( m \times n \) matrix, then \( I_mA = A \) and \( AI_n = A \).

If \( A \) is a square matrix, then \( IA = A = AI \).
The notion of inverse

**Exploration**  Consider the set of real numbers, and say that we have the equation

\[ 3x = 2 \]

and we want to solve for \( x \).

What do we do?

We multiply both sides of the equation by \( \frac{1}{3} \) to obtain

\[ \frac{1}{3}(3x) = \frac{1}{3}(2) \quad \implies \quad x = \frac{2}{3}. \]

\( \frac{1}{3} \) is the multiplicative inverse of 3 since \( \frac{1}{3}(3) = 1 \).

Now, consider the linear system

\[
\begin{align*}
3x_1 - 5x_2 &= 6 \\
-2x_1 + 3x_2 &= -1
\end{align*}
\]

Notice that we can rewrite equations as

\[
\begin{pmatrix} 3 & -5 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \end{pmatrix}
\]

How do we isolate the vector \( \vec{x} \) by itself on LHS?

The notion of inverse

Now, consider the linear system

\[
\begin{align*}
3x_1 - 5x_2 &= 6 \\
-2x_1 + 3x_2 &= -1
\end{align*}
\]

Notice that we can rewrite equations as

\[
\begin{pmatrix} 3 & -5 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \end{pmatrix}
\]

How do we isolate the vector \( \vec{x} \) by itself on LHS?

want this equal to identity matrix, \( I \)

\[
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -3 & -5 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} 6 \\ -1 \end{pmatrix} = \begin{pmatrix} -13 \\ -9 \end{pmatrix}
\]
Matrix inverses

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 $A$.)

Example

$$A = \begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix}$$ is invertible because for $B = \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix}$, we have $AB = \begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

and likewise $BA = \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$.

The notion of an inverse matrix only applies to square matrices.

- For rectangular matrices of full rank, there are one-sided inverses.
- For matrices in general, there are pseudoinverses, which are a generalization to matrix inverses.

Example Find the inverse of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. We have

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\implies a + c = 1 \text{ and } a + c = 0 \quad \text{IMPOSSIBLE!}$$

Therefore, $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is not invertible (or singular).

Take-home message: Not all square matrices are invertible.
Important questions:

- When does a square matrix have an inverse?
- If it does have an inverse, how do we compute it?
- Can a matrix have more than one inverse?

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

*Proof.* Assume $A$ is invertible. Suppose, by way of contradiction, that the inverse of $A$ is not unique, i.e., let $B$ and $C$ be two distinct inverses of $A$. Then, by def’n of inverse, we have

$$BA = I = AB \quad (1)$$
$$\text{and } CA = I = AC. \quad (2)$$

It follows that

$$B = BI \quad \text{by def’n of identity matrix}$$
$$= B(AC) \quad \text{by (2) above}$$
$$= (BA)C \quad \text{by associativity of matrix mult.}$$
$$= IC \quad \text{by (1) above}$$
$$= C. \quad \text{by def’n of identity matrix}$$

Thus, $B = C$, which contradicts the previous assumption that $B \neq C$.  
$\Rightarrow\Leftarrow$ So it must be that case that the inverse of $A$ is unique.  

**Take-home message:** The inverse of a matrix $A$ is unique, and we denote it $A^{-1}$.

**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 there is some matrix ___ such that

___ $A^T = I$ and $A^T$ ___ = $I$.

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 $(A^T)^{-1} = (A^{-1})^T$. □

Question: If $A$ and $B$ are invertible $n \times n$ matrices, what can we say about $A + B$?

There is no guarantee $A + B$ is invertible even if $A$ and $B$ themselves are invertible! In other words, we CANNOT say that $(A + B)^{-1} = A^{-1} + B^{-1}$.

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

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

★ Exercise: Check that $AA^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A^{-1}A$.

Example For $A = \begin{bmatrix} -2 & 1 \\ 3 & -3 \end{bmatrix}$, we have

$$A^{-1} = \frac{1}{3} \begin{bmatrix} -3 & -1 \\ -3 & -2 \end{bmatrix} = \begin{bmatrix} -1 & -\frac{1}{3} \\ -1 & -\frac{2}{3} \end{bmatrix}.$$ We can easily check that

$$AA^{-1} = \begin{bmatrix} -2 & 1 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} -1 & -\frac{1}{3} \\ -1 & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$A^{-1}A = \begin{bmatrix} -1 & -\frac{1}{3} \\ -1 & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$ 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 $A^{-1}$.

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

If we can transform $A$ into $I$, then we will obtain $A^{-1}$. If we cannot do so, then $A$ is not invertible.
Example: Find the inverse of the matrix $A = \begin{bmatrix} -1 & -3 & 1 \\ 3 & 6 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

$$
\begin{bmatrix}
-1 & -3 & 1 & \mid & 1 & 0 & 0 \\
3 & 6 & 0 & \mid & 0 & 1 & 0 \\
1 & 0 & 1 & \mid & 0 & 0 & 1
\end{bmatrix}
\xrightarrow{R_2+3R_1}
\begin{bmatrix}
-1 & -3 & 1 & \mid & 1 & 0 & 0 \\
0 & -3 & 3 & \mid & 3 & 1 & 0 \\
0 & -3 & 2 & \mid & 1 & 0 & 1
\end{bmatrix}
\xrightarrow{R_3-R_1}
\begin{bmatrix}
1 & 3 & -1 & \mid & -1 & 0 & 0 \\
0 & -3 & 3 & \mid & 3 & 1 & 0 \\
0 & 0 & -1 & \mid & -2 & -1 & 1
\end{bmatrix}
\xrightarrow{R_3-R_2}
\begin{bmatrix}
1 & 0 & 2 & \mid & 2 & 1 & 0 \\
0 & -3 & 3 & \mid & 3 & 1 & 0 \\
0 & 0 & 1 & \mid & 2 & 1 & -1
\end{bmatrix}
\xrightarrow{R_1+R_2}
\begin{bmatrix}
1 & 0 & 2 & \mid & 2 & 1 & 0 \\
0 & 1 & -1 & \mid & -1 & -\frac{1}{3} & 0 \\
0 & 0 & 1 & \mid & 2 & 1 & -1
\end{bmatrix}
\xrightarrow{-\frac{1}{3}R_2}
\begin{bmatrix}
1 & 0 & 0 & \mid & -2 & -1 & 2 \\
0 & 1 & 0 & \mid & 1 & \frac{2}{3} & -1 \\
0 & 0 & 1 & \mid & 2 & 1 & -1
\end{bmatrix}
\xrightarrow{R_1-2R_3}
\xrightarrow{R_2+R_3}
\begin{bmatrix}
-2 & -1 & 2 \\
1 & 2 & -1 \\
2 & 1 & -1
\end{bmatrix}
\]

Thus, $A$ is invertible and its inverse is

$$
A^{-1} = \begin{bmatrix} -2 & -1 & 2 \\ 1 & \frac{2}{3} & -1 \\ 2 & 1 & -1 \end{bmatrix}.
$$

Why does this work? $\implies$ discussion next class