Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 Matrix operations

Lay–Lay–McDonald §2.1 pp. 101 – 102

We finish this section by recapping some properties of matrix transposes:

Theorem 1. *Let A and B denote matrices whose sizes are appropriate for the following sums and products.*

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. *For any scalar r,* $(rA)^T = rA^T$
4. $(AB)^T = B^T A^T$

2 The inverse of a matrix

Lay–Lay–McDonald §2.2 pp. 104 – 111

Today we discuss what it means to invert a matrix A; that is, to compute a matrix A^{-1} such that

$$A^{-1}A = AA^{-1} = I.$$

An $n \times n$ matrix A is said to be *invertible* if there is an $n \times n$ matrix C such that $CA = I$ and $AC = I$, where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an inverse of A. The inverse of a matrix A is unique, and we denote it as A^{-1}. A matrix that is not invertible is sometimes called a singular matrix, and an invertible matrix is called a nonsingular matrix.

Example 2. *Let* $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}, C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$. *Compute AC and CA.*

Below is a formula for the inverse of a 2×2 matrix, along with a test for when a 2×2 matrix is invertible:

Theorem 3. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If* $ad - bc \neq 0$, *then* A *is invertible, and*

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

If $ad - bc = 0$, *then* A *is not invertible.*

The quantity above of $ad - bc$ is called the *determinant* of A (in the case of a 2×2 matrix), and we write $\det A = ad - bc$.

Invertible matrices are very useful for solving matrix equations. In fact, we have the following theorem:

Theorem 4. *If* A *is an invertible* $n \times n$ *matrix, then for each* $\mathbf{b} \in \mathbb{R}^n$, *the equation* $A\mathbf{x} = \mathbf{b}$ *has the unique solution* $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof. Let $\mathbf{b} \in \mathbb{R}^n$. Since A is invertible, we may compute $\mathbf{x} = A^{-1}\mathbf{b}$, and we see that

$$A\mathbf{x} = AA^{-1}\mathbf{b} = I\mathbf{b} = \mathbf{b},$$

so $\mathbf{x}$ is certainly a solution to the equation. To check uniqueness, suppose that we have another solution $\mathbf{u}$ of the equation; that is $A\mathbf{u} = \mathbf{b}$. Then multiplying both sides of the equation by A^{-1} yields

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b} \quad \Rightarrow \quad I\mathbf{u} = A^{-1}\mathbf{b} \quad \Rightarrow \quad \mathbf{u} = A^{-1}\mathbf{b},$$

and we see that $\mathbf{u} = \mathbf{x}$. $\square$

Example 5. *Use an inverse matrix to solve the system*

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7. \end{aligned}$$

Here are some useful results about invertible matrices:

1. If A is an invertible matrix, then A^{-1} is invertible, and

$$(A^{-1})^{-1} = A.$$

2. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,
$$(AB)^{-1} = B^{-1}A^{-1}.$$
More generally, the product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

3. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,
$$(A^T)^{-1} = (A^{-1})^T.$$

We will soon see that an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by tracking the row reduction of A to I. Before that, we describe how elementary row operations can be expressed in terms of matrices.

An *elementary matrix* is one that is obtained by performing a single elementary row operation (scale, replace, swap) on an identity matrix. The next example illustrates the three kinds of elementary matrices.

Example 6. *Let* $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}, A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$
Compute E_1A, E_2A, *and* E_3A, *and describe how these products can be obtained by elementary row operations on* A.

It turns out that each elementary matrix E is invertible. The inverse of E is the elementary matrix that transforms E back to I.

Example 7. *Find the inverse of* $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}.$

The following theorem tells us how to see if a matrix is invertible, and it leads to a method for finding the inverse of a matrix.

Theorem 8. *An* $n \times n$ *matrix* A *is invertible if and only if* A *is row equivalent to* I_n, *and in this case, any sequence of elementary row operations that reduces* A *to* I_n *also transforms* I_n *into* A^{-1}.

If we place A and I side by side to form an augmented matrix $[A \quad I]$, then row operations on this matrix produce identical operations on A and on I. By Theorem 8, either there are row operations that transform A to I_n, and I_n to A^{-1} or else A is not invertible.

Algorithm for finding A^{-1}
Row reduce the augmented matrix $[A \quad I]$. If A is row equivalent to I, then $[A \quad I]$ is row equivalent to $[I \quad A^{-1}]$. Otherwise A does not have an inverse.

Example 9. *Find the inverse of the matrix* $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, *if it exists.*

In real life, one might need some, but not all of the entries of A^{-1}. In general, it's an expensive computation to produce all of the entries of A^{-1}. Here's how to get a few columns' worth of A^{-1}. Denote the columns of I_n by $\mathbf{e}_1, \ldots \mathbf{e}_n$. Then row reduction of $[A \quad I]$ to $[I \quad A^{-1}]$ can be viewed as the simultaneous solution of the n systems

$$A\mathbf{x} = \mathbf{e}_1, \quad A\mathbf{x} = \mathbf{e}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{e}_n, \tag{1}$$

where the "augmented columns" of these systems have all been placed next to A to form

$$[A \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n] = [A \quad I].$$

The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in 1. Thus if we are just after a few columns of A^{-1}, it is enough to solve the corresponding systems in (1).