Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 Characterizations of invertible matrices

Lay–Lay–McDonald §2.3 pp. 113 – 116

Today we look at various ways of deciding if a square matrix is invertible.

Example 1. *Determine if the following matrices are invertible:*

1. $\begin{bmatrix} -4 & 6 \\ 6 & -9 \end{bmatrix}$

2. $\begin{bmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{bmatrix}$

3. $\begin{bmatrix} 1 & -5 & -4 \\ 0 & 3 & 4 \\ -3 & 6 & 0 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 & 7 & 4 \\ 0 & 5 & 9 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 10 \end{bmatrix}$

Here is the main result:

The Invertible Matrix Theorem
Let A be a square $n \times n$ matrix. Then the following statements are equivalent:

1. A is an invertible matrix.
2. A is row equivalent to the $n \times n$ identity matrix.
3. A has n pivot positions.
4. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
5. The columns of A form a linearly independent set.
6. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
7. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b} \in \mathbb{R}^n$.
8. The columns of A span $\mathbb{R}^n$.
9. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
10. There is an $n \times n$ matrix C such that $CA = I$.
11. There is an $n \times n$ matrix D such that $AD = I$.
12. A^T is an invertible matrix.

Example 2. *Use the Invertible Matrix Theorem to decide if* $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$ *is invertible.*

Recall that matrix multiplication corresponds to composition of linear transformations. When a matrix A is invertible, the equation $A^{-1}A\mathbf{x} = \mathbf{x}$ can be viewed as a statement about linear transformations.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *invertible* if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \tag{1}$$

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{2}$$

The next result shows that if such an S exists, it is unique and must be a linear transformation. We call S the *inverse* of T and write it as T^{-1}.

Theorem 3. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying (1) and (2).*

Example 4. *The following transformation T is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$:*

$$T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2).$$

Show that T is invertible and find a formula for T^{-1}.

Example 5. *What can you say about a one-to-one linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$?*

2 Matrix factorization

Lay–Lay–McDonald §2.5 pp. 125 – 129

A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices. In the language of computer science, the expression of A as a product amounts to a *preprocessing* of the data in A, organizing that data into two or more parts whose structures are more useful in some way, perhaps more accessible for computation.

The *LU* factorization (described below) is motivated by the common real-life problem of solving a sequence of equations, all with the same coefficient matrix:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \dots, \quad A\mathbf{x} = \mathbf{b}_p. \tag{3}$$

When A is invertible, one could compute A^{-1} and then compute $A^{-1}\mathbf{b}_1, A^{-1}\mathbf{b}_2$, and so on. However, it is more efficient to solve the first equation in sequence (3) by row reduction and obtain an LU factorization at the same time. Then the remaining equations in sequence (3) are solved with the LU factorization.

First, assume that A is an $m \times n$ matrix that can be row reduced to echelon form *without row swaps*. Then A can be written in the form $A = LU$ where L is an $m \times m$ lower triangular matrix with 1s on the diagonal and U is an $m \times n$ echelon form of A:

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

This is called an *LU factorization* of A. The matrix L is invertible and is called a unit lower triangular matrix.

Before giving an algorithm for finding L and U, we look at why this decomposition is useful. When $A = LU$, the equation $A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$. Then writing $\mathbf{y} := U\mathbf{x}$, we can find $\mathbf{x}$ by solving the following pair of equations:

$$\begin{aligned} L\mathbf{y} &= \mathbf{b} \\ U\mathbf{x} &= \mathbf{y}. \end{aligned}$$

First solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$ and then solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. Each equation is easy to solve because L and U are triangular.

Example 6. *It can be checked that*

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU.$$

Use this LU factorization of A to solve $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.